



(12) 发明专利

(10) 授权公告号 CN 102307897 B

(45) 授权公告日 2016. 01. 20

(21) 申请号 200980156194. 1

WO 02095076 A2, 2002. 11. 28,

(22) 申请日 2009. 12. 04

B D Green. N-terminal His7-modification of glucagon-like peptide-1(7-36) amide generates dipeptidyl peptidase IV-stable analogues with potent antihyperglycaemic activity. 《Journal of Endocrinology》. 2004, 第 380-384 页 .

(30) 优先权数据

61/120, 135 2008. 12. 05 US

(85) PCT国际申请进入国家阶段日

2011. 08. 05

(86) PCT国际申请的申请数据

PCT/EP2009/066395 2009. 12. 04

(87) PCT国际申请的公布数据

W02010/063818 EN 2010. 06. 10

B D Green. N-terminal His7-modification of glucagon-like peptide-1(7-36) amide generates dipeptidyl peptidase IV-stable analogues with potent antihyperglycaemic activity. 《Journal of Endocrinology》. 2004, 第 380-384 页 .

(73) 专利权人 葛兰素集团有限公司

地址 英国米德尔塞克斯

Jianbing Zhang. Pentamerization of Single-domain Antibodies from Phage Libraries: A Novel Strategy for the Rapid Generation of High-avidity Antibody Reagents. 《J. Mol. Biol. 》. 2004, 第 52-54 页 .

(72) 发明人 卡罗琳·埃尼弗 劳伦特·杰斯珀斯

马戈扎塔·普佩卡

伊恩·M. 汤姆林森

审查员 苏保卫

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 闵丹

(51) Int. Cl.

C07K 16/00(2006. 01)

C12Q 1/37(2006. 01)

(56) 对比文件

WO 02095076 A2, 2002. 11. 28,

权利要求书1页 说明书44页

序列表13页 附图12页

(54) 发明名称

选出蛋白酶抗性多肽的方法

(57) 摘要

本发明涉及从肽和多肽的文库或集合(例如展示系统)中挑选、分离和/或回收抗蛋白酶(比如血清中可见的蛋白酶)降解的肽或多肽的方法。总体上,所述方法包括提供肽或多肽文库或集合;在适合蛋白酶活性的条件下将文库或集合与蛋白酶一起温育;并挑选、分离和/或回收抗蛋白酶降解并具有所需生物活性的肽或多肽。选中的肽和多肽可以作为治疗剂用于例如治疗人的疾病。

1. 分离的蛋白酶抗性肽, 其由 SEQ ID NO:16 或 21 所示氨基酸序列组成。
2. 结合血清白蛋白的域抗体, 其由 SEQ ID NO:22 所示氨基酸序列组成。
3. 权利要求 1 的肽, 在下述条件与蛋白酶温育时, 抗选自胰蛋白酶和胰凝乳蛋白酶的一或多种蛋白酶: (i) 10 μ g/ml 到 3mg/ml 蛋白酶, (ii) 20°C 到 40°C 和 (iii) 温育至少 30 分钟。
4. 权利要求 3 的肽, 在下述条件与所述蛋白酶温育时, 抗选自胰蛋白酶和胰凝乳蛋白酶的一种或多种蛋白酶: 100 μ g/ml 蛋白酶, 37°C, 温育至少 1 小时。
5. 权利要求 1 的肽或权利要求 2 所述的结合血清白蛋白的域抗体在制备用于治疗 and / 或预防糖尿病, 糖尿病相关疾病或肥胖的药物中的用途。
6. 权利要求 5 所述的用途, 其中所述药物是用于口服, 舌下, 直肠或胃肠外给药。
7. 权利要求 5 所述的用途, 其中所述药物通过皮下、肌内或静脉注射给予。
8. 分离的或重组的核酸分子, 其编码权利要求 1 所述的分离的肽或权利要求 2 所述的结合血清白蛋白的域抗体。
9. 分离的或重组的核酸分子, 其由 SEQ ID NO:23 所示核酸序列组成。
10. 载体, 包含权利要求 8 或 9 所述的核酸分子。
11. 宿主细胞, 包含权利要求 8 或 9 所述的核酸分子或者权利要求 10 所述的载体。
12. 权利要求 11 的宿主细胞, 是毕赤酵母。

选出蛋白酶抗性多肽的方法

[0001] 发明背景

[0002] 多肽和肽已成为包括工业应用和作为医学、治疗以及诊断剂的众多用途中日益重要的物质。但是,许多治疗用肽和蛋白质在体内特别容易受到天然蛋白酶的降解。而且,一些生理状态下,比如炎症状态(例如 COPD)和癌症,组织、器官或动物(例如在肺中、肿瘤内或者肿瘤附近)中存在的蛋白酶的量可能会增加。蛋白酶的这种增加会导致内源蛋白以及为了治疗疾病而给予的治疗用肽、多肽和蛋白的降解和灭活。因此,一些有潜力用于体内的药剂由于被蛋白酶迅速地降解和灭活,而只有有限的药效。

[0003] 蛋白酶抗性多肽具有多种优势。例如,蛋白酶抗性多肽在体内比蛋白酶敏感性制剂保持更长时间的活性,因此能够在一段时间内维持功能,足以产生生物学效应。需要选择能够抗蛋白酶降解并具备所需生物活性的多肽的改进方法。

[0004] 胰高血糖素样肽 (Glucagon-like peptide, GLP)-1 是一种肠促胰岛素激素,它具有强大的葡萄糖依赖性促胰岛素和抑制高血糖素作用、对胰岛 β 细胞的滋养效果,以及对胃肠分泌和蠕动的抑制效应,这些结合起来使得血浆葡萄糖降低并减少血糖波动。GLP-1 是 GLP-1 受体的激动剂。此外,通过其增加饱腹感的能力,GLP-1 减少食物摄入,从而限制体重增加,甚至可能造成体重下降。综合起来,这些作用使得 GLP-1 很独特,被认为很有希望作为抗糖尿病药物,尤其是因为它的抗高血糖效应对葡萄糖的依赖性应当能使严重低血糖的任何风险最小化。但是,它的药物代谢动力学 / 药效动力学特性使得它不能用于治疗。因此, GLP-1 在连续给药时最有效,而单次皮下注射只有短期效果。GLP-1 在体内对酶促降解高度易感,其中最相关的是被二肽基肽酶 IV (DPP-IV) 切割,因为发生的非常迅速并且产生非促胰岛素代谢产物。因此基于对影响 GLP-1 的代谢稳定性以及药物代谢动力学 / 药效动力学特性的因素的理解,增强其治疗潜力的策略一直以来是大量研究的重点。

[0005] 已有大量工作试图抑制蛋白酶或者对 GLP-1 进行修饰,从而使得在维持生物活性的同时减慢它的降解。W005/027978 公开的 GLP-1 衍生物具有延长的作用(通过引用并入本文作为可用于本发明的 GLP-1 衍生物和类似物的例子)。WO 02/46227 公开了异源融合蛋白,其包含与 GLP-1 或类似物(关于这些类似物的公开通过引用并入本文作为可以用于本发明的 GLP-1 类似物的例子)融合的多肽(例如白蛋白)。W005/003296、W003/060071、W003/059934 公开了氨基融合蛋白,其中 GLP-1 融合了白蛋白以便提高激素的半衰期。

[0006] 然而,尽管有这些努力,还没有生产出有长期活性的 GLP-1。

[0007] 发明概述

[0008] 本发明涉及选择蛋白酶抗性肽或多肽的方法,以及选择与靶配体高亲和力结合的肽或多肽的方法。发明还涉及制备蛋白酶抗性肽或多肽集合的方法。

[0009] 一个方面,发明是用于选择蛋白酶抗性肽或多肽的方法。所述方法包括提供肽或多肽集合,将集合与蛋白酶在适合蛋白酶活性的条件下温育,以及回收具有所需生物活性的肽或多肽,从而选择到蛋白酶抗性肽或多肽。

[0010] 一个实施方案中,在展示系统中表达所述肽或多肽集合,所述蛋白酶是该展示系统或表达宿主中表达的蛋白酶。例如,在一个实施方案中,肽或多肽集合表达在细菌细胞

内,蛋白酶是该细菌的内源蛋白酶。适合集合表达的条件使表达和内源蛋白酶(比如细菌蛋白酶)的活性最大化。蛋白酶表达和活性的最大化可以通过例如提高细菌中集合的蛋白表达时间和/或温度。例如,温育时间可以从1小时到过夜(例如从12到24小时)或者更长(例如从24到48小时,或者更长)。在一个实施方案中,温度可以是30到37摄氏度或者更高。此外,可以通过使用不同细菌菌株和/或改变培养基成分来增加蛋白酶表达。还可以使用不同细菌培养物密度。另一个实施方案中,可以通过例如基因修饰改变展示系统从而增强蛋白酶表达。

[0011] 在一个实施方案中,集合作为噬菌体展示系统来提供,其中噬菌体集合在大肠杆菌细菌细胞,比如大肠杆菌 TB1 细胞、TC1 细胞或来自大肠杆菌菌株 HB2151 的细胞中表达和/或扩增。相应地,在该实施方案中,细菌蛋白酶是大肠杆菌细胞中内源表达的蛋白酶。在一个实施方案中,细菌蛋白酶可以是表达在细菌细胞周质的蛋白酶。在一个实施方案中,蛋白酶表达可以是在噬菌体产生、分泌期间或者在细菌上清中。

[0012] 因此,发明的一个实施方案中提供了这样的方法,所述方法步骤包括选择噬菌体文库、在细菌蛋白酶活性适宜的条件下在细菌中表达该文库、将表达的文库与目标配体温育,从而挑选出蛋白酶抗性的目标结合肽。任选地,所述与目标配体进行的温育包括另一种蛋白酶的存在。

[0013] 另一个实施方案中,展示系统是酵母菌展示系统,比如毕赤氏酵母;蛋白酶是酵母细胞内表达的内源蛋白酶。

[0014] 在一个实施方案中,依照发明的方法还包括将集合与另一种蛋白酶在适合所述另一种蛋白酶的活性的条件下组合,并回收具有所需生物活性的肽或多肽,从而挑选到蛋白酶抗性肽或多肽。在一个实施方案中,蛋白酶与集合在溶液中组合(即,蛋白酶被固定在支持物上)。可以在血清、痰、粘液(例如胃粘液、鼻粘液、支气管粘液)、支气管肺泡灌洗液、肺匀浆、肺提取物、胰提取物、胃液、唾液或眼泪中的一或多种中找到合适的另一种蛋白酶。

[0015] 另一方面中,提供了选择蛋白酶抗性肽或多肽的方法。所述方法包括提供肽或多肽集合、将集合与第一蛋白酶在适合蛋白酶活性的条件下温育,方法还包括将集合与第二蛋白酶在适合蛋白酶活性的条件下组合并回收具有所需生物活性的肽或多肽,从而挑选到蛋白酶抗性肽或多肽。这一方面的一个实施方案中,所述第一蛋白酶是集合展示系统的内源蛋白酶,第二蛋白酶选自可以在血清、痰、粘液(例如胃粘液、鼻粘液、支气管粘液)、支气管肺泡灌洗液、肺匀浆、肺提取物、胰提取物、胃液、唾液或眼泪中找到的蛋白酶。但是应当明白,可以按照任何顺序进行所述“第一”和“第二”蛋白酶步骤。此外,可以理解,任何这种步骤的多个重复都涵盖在本发明方法的范围内。

[0016] 在发明任何方面的一个实施方案中,所述适合另一种或第二蛋白酶活性的条件是(i) 大约 10 μ g/ml 到大约 3mg/ml 蛋白酶, (ii) 大约 20°C 到大约 40°C, 以及 (iii) 至少大约 30 分钟。在一个实施方案中,这些严紧条件使得能够挑选到具有高亲和力和/或 T_m 增加的肽或多肽。这种情况中,肽和多肽可能以单体形式展示出高亲和力。

[0017] 在一个实施方案中,依照发明任何方面的方法中,对于适合蛋白酶活性的条件使用了大约 10 到大约 100 μ g/ml 蛋白酶。对于所述适合蛋白酶活性的条件可以使用大约 30 到大约 37°C 的温度(例如大约 37°C 或者大约室温)。在一个实施方案中,可以将集合和蛋白酶组合在一起至少大约 1 小时(例如,大约 1 小时、大约 2 小时、过夜,例如 18 到 24 小时)。

本发明的方法中,集合和蛋白酶在一个实施方案中温育了至少约 30 分钟的时期。在一个实施方案中,使用的蛋白酶在大约 100 μ g/ml,组合在一起的集合和蛋白酶在大约 37°C 温育了至少约 1 小时。

[0018] 在发明任何方面的一个实施方案中,蛋白酶(例如胰蛋白酶)与多肽或可变区的比率(基于摩尔/摩尔)是 8,000 到 80,000 蛋白酶:可变区。在一个实施方案中,蛋白酶(例如胰蛋白酶)与多肽或可变区的比率(基于重量/重量,例如微克/微克)是 16,000 到 160,000 蛋白酶:可变区。在一个实施方案中,使用的蛋白酶浓度为至少 100 或 1000 微克/ml 蛋白酶。

[0019] 任何需要的蛋白酶都可以用于依据本发明任何方面的方法中,比如下述中的一种或多种:丝氨酸蛋白酶、半胱氨酸蛋白酶、天冬氨酸蛋白酶、巯基蛋白酶、基质金属蛋白酶、羧肽酶(例如羧肽酶 A、羧肽酶 B)、胰蛋白酶、胰凝乳蛋白酶、胃蛋白酶、木瓜蛋白酶、弹性蛋白酶、leucozyme、胰酶(pancreatin)、凝血酶、纤溶酶、组织蛋白酶(cathepsin)(例如组织蛋白酶 G)、蛋白水解酶(proteinase)(例如蛋白水解酶 1、蛋白水解酶 2、蛋白水解酶 3)、嗜热菌蛋白酶、凝乳酶、肠肽酶、半胱天冬氨酸酶(caspase)(例如半胱天冬氨酸酶 1、半胱天冬氨酸酶 2、半胱天冬氨酸酶 4、半胱天冬氨酸酶 5、半胱天冬氨酸酶 9、半胱天冬氨酸酶 12、半胱天冬氨酸酶 13)、需钙蛋白酶(calpain)、无花果蛋白酶(ficain)、梭菌蛋白酶(clostripain)、奇异果蛋白酶(actinidain)、菠萝蛋白酶(bromelain)、分离酶(separase)和二肽基肽酶 IV(DPP-IV)。具体实施方案中,蛋白酶是胰蛋白酶、弹性蛋白酶或者 leucozyme。还可以通过生物提取物、生物匀浆(biological homogenate)或者生物制品,例如体外完整细胞来提供蛋白酶。如果需要,方法还包括在温育结束后向集合与蛋白酶的组合中加入蛋白酶抑制剂。

[0020] 本发明任何方法的一个实施方案中,当与集合合并时,蛋白酶处于溶液中。

[0021] 发明任何方面的一个实施方案中,所需生物活性是结合活性,例如与配体(例如靶配体或者通用配体(generic ligand))的结合活性。

[0022] 一些实施方案中,具有所需生物活性的肽或多肽的回收是基于结合活性。例如可以基于与诸如蛋白 A、蛋白 G 或蛋白 L 的通用配体的结合来回收肽或多肽。结合活性还可以是与靶配体的特异结合。示范性靶配体包括 ApoE、Apo-SAA、BDNF、心肌营养素(cardiotrophin)-1、CEA、CD40、CD40 配体、CD56、CD38、CD138、EGF、EGF 受体、ENA-78、嗜酸性粒细胞趋化因子(Eotaxin)、嗜酸性粒细胞趋化因子-2、Exodus-2、AP、酸性 FGF、碱性 FGF、成纤维细胞生长因子-10、FLT3 配体、Fractalkine(CX3C)、GDNF、G-CSF、GM-CSF、GF- β 1、人血清白蛋白、胰岛素、IFN- γ 、IGF-I、IGF-II、IL-1 α 、IL-1 β 、IL-1 受体、I 型 IL-1 受体、IL-2、IL-3、IL-4、IL-5、IL-6、IL-7、IL-8(72a. a.)、IL-8(77a. a.)、IL-9、IL-10、IL-11、IL-12、IL-13、IL-15、IL-16、IL-17、IL-18(IGIF)、抑制素 α 、抑制素 β 、IP-10、角质形成细胞生长因子-2(KGF-2)、KGF、瘦素、LIF、淋巴细胞趋化因子、穆勒抑制物(Mullerian inhibitory substance)、单核细胞集落抑制因子(monocyte colony inhibitory factor)、单核细胞趋化蛋白(monocyte attractant protein)、M-CSF、MDC(67a. a.)、MDC(69a. a.)、MCP-1(MCAF)、MCP-2、MCP-3、MCP-4、MDC(67a. a.)、MDC(69a. a.)、MIG、MIP-1 α 、MIP-1 β 、MIP-3 α 、MIP-3 β 、MIP-4、骨髓造血祖细胞抑制因子(myeloid progenitor inhibitor factor-1, MPIF-1)、NAP-2、神经营养因子(Neurturin)、神经生长因子、 β -NGF、NT-3、NT-4、

制瘤素 (Oncostatin)M、PDGF-AA、PDGF-AB、PDGF-BB、PF-4、RANTES、SDF1 α 、SDF1 β 、SCF、SCGF、干细胞因子 (SCF)、TARC、TGF- α 、TGF- β 、TGF- β 2、TGF- β 3、肿瘤坏死因子 (TNF)、TNF- α 、TNF- β 、TNF 受体 I、TNF 受体 II、TNIL-1、TPO、VEGF、VEGF A、VEGF B、VEGF C、VEGF D、VEGF 受体 1、VEGF 受体 2、VEGF 受体 3、GCP-2、GRO/MGSA、GRO- β 、GRO- γ 、HCC1、1-309、HER 1、HER 2、HER 3、HER 4、血清白蛋白、vWF、淀粉样蛋白 (例如 α 淀粉样蛋白)、MMP12、PDK1、IgE、IL-13R α 1、IL-13Ra2、IL-15、IL-15R、IL-16、IL-17R、IL-17、IL-18、IL-18R、IL-23、IL-23R、IL-25、CD2、CD4、CD11a、CD23、CD25、CD27、CD28、CD30、CD40、CD40L、CD56、CD138、ALK5、EGFR、FcER1、TGFb、CCL2、CCL18、CEA、CR8、CTGF、CXCL12 (SDF-1)、胰凝乳蛋白酶、FGF、弗林蛋白酶 (Furin)、内皮素 -1、嗜酸性粒细胞趋化因子 (例如嗜酸性粒细胞趋化因子、嗜酸性粒细胞趋化因子 -2、嗜酸性粒细胞趋化因子 -3)、GM-CSF、ICAM-1、ICOS、IgE、IFNa、I-309、整合素、L- 选择素 (selectin)、MIF、MIP4、MDC、MCP-1、MMPs、嗜中性细胞弹性蛋白酶、骨桥蛋白、OX-40、PARC、PD-1、RANTES、SCF、SDF-1、siglec8、TARC、TGFb、凝血酶、Tim-1、TNF、TRANCE、类胰蛋白酶 (Trypsin)、VEGF、VLA-4、VCAM、 α 4 β 7、CCR2、CCR3、CCR4、CCR5、CCR7、CCR8、alphavbeta6、alphavbeta8、cMET、CD8、vWF、淀粉样蛋白 (例如 α 淀粉样蛋白)、MMP12、PDK1 和 IgE。另一个实施方案中,靶配体是 GLP-1 受体或其部分。例如在依照本发明任何方面的方法中,所述配体可以是 GLP-1 受体胞外结构域。

[0023] 在发明任何方面的具体实施方案中,通过淘选 (panning) 回收肽或多肽。

[0024] 在本发明任何方法的一个实施方案中,在有蛋白酶的情况下,集合与配体 (靶配体 ;通用配体) 进行接触,根据与配体的结合选择集合中的一或多个成员。

[0025] 本发明任何方法的一些实施方案中,集合包含展示系统。例如展示系统可以是噬菌体展示、核糖体展示、乳液区室化和展示 (emulsion compartmentalization and display)、酵母展示、嘌呤霉素展示 (puromycin display)、细菌展示、在质粒上进行展示或者共价展示。优选的展示系统将核酸的编码功能与该核酸编码的肽或多肽的功能性特征联系起来。具体实施方案中,展示系统包含可复制的遗传包 (genetic package)。

[0026] 本发明任何方法的一些实施方案中,展示系统包含噬菌体展示。例如,噬菌体可以是 fd、M13、 λ 、MS2 或 T7。具体实施方案中,噬菌体展示系统是多价的。一些实施方案中,肽或多肽展示为 pIII 融合蛋白。

[0027] 本发明任何方法的一个实施方案中,肽或多肽 (例如可变域) 集合展示在噬菌体上,噬菌体文库大小是例如 10^6 到 10^{13} ,例如 10^8 到 10^{12} 复制单位 (感染性病毒粒子)。在一个实施方案中,当与第二或其它蛋白酶一起温育时,集合展示在噬菌体上。

[0028] 在发明任何方面的其他实施方案中,方法还包括对编码具有所需生物活性的肽或多肽的核酸进行扩增。在具体实施方案中,核酸通过噬菌体扩增、细胞生长或者聚合酶链式反应来扩增。

[0029] 在发明任何方面的一个实施方案中,肽或多肽集合展示在噬菌体上,后者在诸如大肠杆菌的细菌细胞中扩增和表达。该实施方案中,肽或多肽集合在细菌细胞中表达时与细菌蛋白酶接触。

[0030] 一些实施方案中,集合是免疫球蛋白单个可变域的集合。具体实施方案中,免疫球蛋白单个可变域是重链可变域。更具体实施方案中,重链可变域是人重链可变域。其他实施方案中,免疫球蛋白单个可变域是轻链可变域。具体实施方案中,轻链可变域是人轻链可

变域。

[0031] 本发明另一方面是用于从肽或多肽集合中挑选能够以高亲和力结合靶配体的肽或多肽的方法。所述方法包括提供肽或多肽集合、将集合与蛋白酶在适合蛋白酶活性的条件下组合,以及回收与靶配体结合的肽或多肽。

[0032] 对于以上发明方法,当所需生物活性是结合活性时,所结合的配体(靶配体、通用配体)与蛋白酶不同。

[0033] 本发明另一方面是制备蛋白酶抗性肽或多肽集合的方法。所述方法包括提供肽或多肽集合、将肽或多肽集合与蛋白酶在适合蛋白酶活性的条件下组合,以及回收具有所需生物活性的多种肽或多肽,从而产生蛋白酶抗性肽或多肽集合。

[0034] 一些实施方案中,具有所需生物活性的多种肽或多肽是基于结合活性回收的。例如可以根据与通用配体,比如蛋白 A、蛋白 G 或蛋白的结合来回收多种肽或多肽。

[0035] 发明另一方面是用于从集合中挑选蛋白酶抗性多肽的方法,所述多肽包含能够与靶配体结合的免疫球蛋白单个可变域(dAb)。在一个实施方案中,方法包括提供含有多肽集合的噬菌体展示系统,所述多肽包含免疫球蛋白单个可变域;将噬菌体展示系统与选自弹性蛋白酶、leucozyme 和胰蛋白酶的蛋白酶在适合蛋白酶活性的条件下组合;以及回收展示多肽的噬菌体,所述多肽包含能够结合靶配体的免疫球蛋白单个可变域。适宜地,在这个方面的一个实施方案中方法还包括在适合表达内源蛋白酶的条件下进行温育。例如,所述内源蛋白酶是展示系统所表达的蛋白酶。

[0036] 一些实施方案中,使用 100 μ g/ml 的蛋白酶,合并在一起的噬菌体展示系统和蛋白酶在大约 37°C 温育过夜。

[0037] 一些实施方案中,展示包含能够结合靶配体的免疫球蛋白单个可变域之多肽的噬菌体通过与所述目标分子的结合来回收。其他实施方案中,展示包含能够结合靶配体的免疫球蛋白单个可变域之多肽的噬菌体通过淘选来回收。

[0038] 发明还涉及可以通过本文描述的方法挑选或者通过本文描述的方法挑选到的分离的蛋白酶抗性肽或多肽。具体实施方案中,发明涉及 GLP-1 受体激动剂,比如文中描述的 GLP-1 肽。实施例和图 1 中列出了合适的 GLP-1 肽和 GLP-1 肽衍生物。其他合适的肽包括 GLP-1 同源物或衍生物,比如毒蜥外泌肽(exendin)及其同源物和衍生物。其他合适的衍生物包括 GLP-1 的二肽基肽酶 IV 抗性衍生物。一个优选的肽是氨基酸序列 DMS7148(图 1 中的序列 6)所限定的肽。另一个优选的肽是氨基酸序列 DMS7161(图 1 中的序列 11)所限定的肽。适宜地,这些 GLP-1 肽与 AlbuAb™序列融合。另一个实施方案中,发明涉及可以通过本文描述的方法挑选或者通过本文描述的方法挑选到的分离的蛋白酶(例如胰蛋白酶、弹性蛋白酶、leucozyme)抗性免疫球蛋白单个可变域(例如人抗体重链可变域、人抗体轻链可变域)。

[0039] 有益地,符合本发明的肽或多肽可能表现出在低成本宿主表达方面的改良特性,因此更适合用于工业规模生产。

[0040] 发明还涉及编码这样的蛋白酶抗性肽或多肽(例如胰蛋白酶、弹性蛋白酶或 leucozyme 抗性免疫球蛋白单个可变域)的分离的或重组的核酸,所述肽或多肽是可以通过本文描述的方法挑选或者通过本文描述的方法挑选到的;还涉及包含所述核酸的载体(例如表达载体)和宿主细胞。

[0041] 发明还涉及制备可以通过本文描述的方法挑选或者通过本文描述的方法挑选到的蛋白酶抗性肽或多肽（例如胰蛋白酶、弹性蛋白酶或 leucozyme 抗性免疫球蛋白单个可变域）的方法，所述方法包括将含有编码蛋白酶抗性肽或多肽的重组核酸的宿主细胞维持在适合表达的条件下，从而产生蛋白酶抗性肽或多肽。

[0042] 因此，在本发明的任何方面中，蛋白酶可以是诸如细菌蛋白酶的展示系统的内源蛋白酶，或者是可以在血清、痰、粘液（例如胃粘液、鼻粘液、支气管粘液）、支气管肺泡灌洗液、肺匀浆、肺提取物、胰提取物、胃液、唾液或眼泪中的一或多种中找到的。在一个实施方案中，蛋白酶是在眼睛和 / 或眼泪中找到的。诸如文中讨论的，所选蛋白酶抗性肽或多肽在哺乳动物（例如人）的疾病或状况中有治疗、预防和诊断用途。具体来说，肽和多肽可以作为那些当给予患者（比如人）时很可能遭遇蛋白酶的药物的基础。

[0043] 例如，向胃肠（GI）道给药（例如口腔、舌下、直肠给药）时，肽或多肽可能接触到上 GI 道、下 GI 道、口、胃、小肠和大肠中一或多处中的蛋白酶。因此一个实施方案中提供了蛋白酶抗性肽或多肽用于经口、舌下或者直肠给予到患者的 GI 道，从而治疗和 / 或预防患者的疾病或状况。

[0044] 例如，一个实施方案中，发明涉及口服可以通过本发明的方法挑选的或者通过本发明的方法挑选到的 TNF α 拮抗剂肽或多肽，用于治疗和 / 或预防 TNF α - 介导的状况或疾病，比如关节炎（例如类风湿性关节炎）、IBD、银屑病或克罗恩氏（Crohn's）病。该实施方案中，所述拮抗剂可以是抗 TNFR1 免疫球蛋白单个可变域（dAb）。另一个例子中，肽或多肽当给予肺部组织（例如肺或呼吸道）时可能遇到蛋白酶。因此一个实施方案提供了蛋白酶抗性肽或多肽用于通过吸入或鼻内给药到患者的肺部组织，从而治疗和 / 或预防患者的疾病或状况。这类状况可能是哮喘（例如过敏性哮喘）、COPD、流感或 W02006038027（通过引用并入本文）中公开的任何其他肺部疾病或状况。

[0045] 另一个例子中，肽或多肽在经肠胃外（例如经皮下注射）给药时可能遇到血清中的蛋白酶。因此一个实施方案提供了蛋白酶抗性肽或多肽用于通过注射给药，从而治疗和 / 或预防患者的疾病或状况。这类状况可能是糖尿病。一个实施方案中，发明提供了肠胃外给予可以通过本发明的方法来挑选的或者通过本发明的方法挑选到的胰高血糖素样肽 1 受体激动剂（比如 GLP-1 或其同源物和衍生物，比如毒蜥外泌素或其衍生物）来治疗和 / 或预防糖尿病或糖尿病相关的疾病。

[0046] 符合发明的肽或多肽可能表现出提高的或者较高的解链温度（ T_m ），提供增强的稳定性。肽和多肽还可能具有高亲和力靶结合的特性。这些特性，结合蛋白酶抗性，使得所述肽或多肽适合作为哺乳动物（比如人）的药物，这种情况中通过例如 GI 道、肺部组织或肠胃外给药时很可能遇到蛋白酶。

[0047] 另一个例子中，肽或多肽（例如可变域或拮抗剂）在（通过例如眼内注射或作为滴眼液）给药到患者眼睛时可能遇到蛋白酶。因此，一个实施方案提供了蛋白酶抗性肽、多肽、免疫球蛋白单个可变域、激动剂或拮抗剂经眼给予患者，从而治疗和 / 或预防患者的疾病或状况（例如眼睛的疾病或状况）。给药可以是对眼睛的局部给药、以滴眼液的形式或者通过注射到眼睛里，例如注射到玻璃体液中。

[0048] 在一个实施方案中，发明提供了用于递送到肺的肺部制剂，其中所述制剂包含本发明的激动剂、拮抗剂、肽、多肽或可变域，颗粒大小范围在小于 5 微米，例如小于 4.5、4、

3.5 或 3 微米（例如，在 Britton-Robinson 缓冲液中，例如 pH 为 6.5 到 8.0，例如 pH 为 7 到 7.5，例如在 pH7 或 pH7.5）。

[0049] 在一个实施方案中，发明的制剂和组合物提供为 pH 在 6.5 到 8.0，例如 7 到 7.5，例如 7，例如 7.5。

[0050] 根据发明任何方面的肽或多肽（例如可变域）可具有至少 50°C，或至少 55°C，或至少 60°C，或至少 65°C，或至少 70°C 的 T_m。本发明的激动剂、拮抗剂、用途、方法、组合物、装置或制剂可包含这样的肽或多肽。

[0051] 在发明的一个方面中，本发明的肽、多肽、可变域、激动剂、拮抗剂、组合物或制剂（多肽或可变域浓度为 1mg/ml）在 Britton-Robinson 缓冲液中，于 37 到 50°C 温育 14 天后是基本稳定的。在一个实施方案中，在 37°C 这样温育后，至少 65、70、75、80、85、86、87、88、90、91、92、93、94、95、96、97、98、99% 的肽、多肽、激动剂、拮抗剂或可变域保持不聚集。在一个实施方案中，在 37°C 这样温育后，至少 65、70、75、80、85、86、87、88、90、91、92、93、94、95、96、97、98、99% 的肽、多肽或可变域保持单体状态。在一个实施方案中，在 50°C 这样温育后，至少 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、86、87、88、90、91、92、93、94、95、96、97、98、99% 的肽、多肽、激动剂、拮抗剂或可变域保持不聚集。在一个实施方案中，在 50°C 这样温育后，至少 5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、86、87、88、90、91、92、93、94、95、96、97、98、99% 的肽、多肽或可变域保持单体状态。在一个实施方案中，任何这样的温育后，都没有见到肽、多肽、可变域、激动剂、拮抗剂的聚集。在一个实施方案中，以 1mg/ml 的浓度在 Britton-Robinson 缓冲液中于 37°C 温育后，肽、多肽或可变域的 pI 没有改变。

[0052] 发明的一个方面中，本发明的肽、多肽、可变域、激动剂、拮抗剂、组合物或制剂（多肽或可变域浓度为 100mg/ml）在 pH 为 7 到 7.5（例如在 pH7 或 pH7.5）的 Britton-Robinson 缓冲液中于 4°C 温育 7 天后，基本稳定。在一个实施方案中，这样的温育后，至少 95、95.5、96、96.5、97、97.5、98、98.5、99 或 99.5% 的肽、多肽、激动剂、拮抗剂、或可变域保持不聚集。在一个实施方案中，这样的温育后，至少 95、95.5、96、96.5、97、97.5、98、98.5、99 或 99.5% 的肽、多肽或可变域保持单体状态。在一个实施方案中，任何这样的温育后，都没有见到肽、多肽、可变域、激动剂、拮抗剂的聚集。

[0053] 发明的一个方面中，本发明的肽、多肽、可变域、激动剂、拮抗剂、组合物或制剂（多肽或可变域浓度为 40mg/ml）于例如室温下（20°C）或者 37°C 在例如喷射雾化器（例如 Pari LC+cup）中雾化 1 小时后，基本稳定。在一个实施方案中，这样的雾化后，至少 65、70、75、80、85、86、87、88、90、91、92、93、94、95、96、97、98、99 或 99.5% 的肽、多肽、激动剂、拮抗剂、或可变域保持不聚集。在一个实施方案中，至少 65、70、75、80、85、86、87、88、90、91、92、93、94、95、96、97、98、99 或 99.5% 的肽、多肽或可变域保持单体状态。在一个实施方案中，任何一种这样的雾化后，都没有见到肽、多肽、可变域、激动剂、拮抗剂的聚集。

[0054] 肽或多肽可以是分离的和 / 或重组的。

[0055] 适宜地，在发明任何方面的一个实施方案中，蛋白酶抗性肽或多肽选自肽或多肽集合。

[0056] 发明还涉及可以通过本文描述的方法来挑选的或者通过本文描述的方法挑选到的蛋白酶抗性肽或多肽（例如胰蛋白酶、弹性蛋白酶、或 leucozyme 抗性免疫球蛋白单个可

变域)用于(例如治疗或诊断的)药物。发明还涉及可以通过本文描述的方法来挑选的或者通过本文描述的方法挑选到的蛋白酶抗性肽或多肽(例如胰蛋白酶、弹性蛋白酶、或 Leucozyme 抗性免疫球蛋白单个可变域)用于制备治疗疾病的药物。发明还涉及疾病的治疗方法,所述方法包括将可以通过本文描述的方法来挑选的或者通过本文描述的方法挑选到的蛋白酶抗性肽或多肽(例如胰蛋白酶、弹性蛋白酶、或 leucozyme 抗性免疫球蛋白单个可变域)以有效量给予有需要的受试对象。

[0057] 在发明任何方面的一个实施方案中,方法还包括将第二蛋白酶与蛋白酶抗性肽或多肽集合在适合第二蛋白酶活性的条件下组合;以及回收具有所需生物活性的至少一种肽或多肽,从而挑选到对第二蛋白酶有抗性的至少一种肽或多肽。第一和第二蛋白酶不同。第二蛋白酶可能是如上定义的。在一个实施方案中,第一或第二蛋白酶是集合展示系统内源的。

[0058] 发明还提供了包含所述肽或多肽的分离的 GLP-1 受体激动剂用于给予患者治疗和/或预防糖尿病,所述激动剂当与蛋白酶在适合发明方法的条件下温育时,对上文提到的一或多种蛋白酶有抗性,所述条件是例如(i)大约 10 μ g/ml 到大约 3mg/ml 蛋白酶,(ii)大约 20°C 到大约 40°C 以及(iii)温育至少大约 30 分钟(例如,在 100 μ g/ml 蛋白酶,于 37°C 温育至少 1 小时的条件)。激动剂可以用于通过注射给药。

[0059] 本发明方法的一个实施方案中,进一步评估挑选到的肽或多肽对第二蛋白酶的抗性,或者对第一蛋白酶但是在与挑选方法中所用不同的一系列条件下的抗性。第二蛋白酶不同于第一蛋白酶,但除此之外可以是上文描述的任何蛋白酶。在一个实施方案中,发明方法中挑选到了一种以上的蛋白酶抗性肽或多肽,之后的步骤用于确定这些肽或多肽中哪些显示对第二蛋白酶的抗性,或者显示对第一蛋白酶但是在与挑选方法中所用不同的一系列条件下的抗性。第二蛋白酶不同于第一蛋白酶,但除此之外可以是上文描述的任何蛋白酶。通过这种方法,得到抗一种以上蛋白酶的一或多种肽或多肽。在一个实施方案中,第一或第二蛋白酶是集合展示系统内源表达的蛋白酶。

[0060] 在本发明方法的一个实施方案中,挑选到蛋白酶抗性的单体形式的肽或多肽(例如免疫球蛋白单个可变域单体)。

[0061] 本发明的药物、激动剂和拮抗剂可以包含与所述肽或多肽融合在一起的抗体恒定区(例如 Fc)。

[0062] 在一个实施方案中,发明提供了蛋白酶抗性肽或多肽在制备药物中的用途,所述药物用于给予哺乳动物从而提供具有提高的 PK 的药物。提高的 PK 可以是提高的 AUC(曲线下面积)和/或提高的半衰期。在一个实施方案中,蛋白酶抗性肽或多肽是通过发明的方法挑选到的或者是可以通过发明的方法来挑选的。在一个实施方案中,所述肽或多肽是免疫球蛋白单个可变域。药物可以包含与所述肽或多肽融合在一起的抗体恒定区,例如抗体 Fc。

[0063] 发明提供了包含蛋白酶抗性肽或多肽的药物,用于给予哺乳动物以便提供在哺乳动物中 PK 提高的药物。在一个实施方案中,蛋白酶抗性肽或多肽是通过本发明的方法挑选到的或者可以通过本发明的方法来挑选。在一个实施方案中,所述肽或多肽是免疫球蛋白单个可变域。药物可以包含与所述肽或多肽融合在一起的抗体恒定区(例如 Fc)。

[0064] 附图简述

- [0065] 图 1 :显示了 GLP-1-AlbudAb 融合变体 1-10 的序列。
- [0066] 图 2 :显示了 GLP-1-AlbudAb 融合变体 6-10 的凝胶。
- [0067] 图 3 :显示了 GLP-1-AlbudAb 融合变体 6-10 的凝胶 (浓缩)。
- [0068] 图 4 :显示了 GLP-1-AlbudAb 融合变体 11 的凝胶。
- [0069] 图 5 :显示了 GLP-1-AlbudAb 融合变体 6-11 的 MS 结果。
- [0070] 图 5a) 显示了 DMS7148 (变体 6) ;(分析备注 :测量质量与预期含有单个二硫键的质量吻合 (15245.88)) ;b) 显示了 DMS7149 (变体 7) (分析备注 :测量质量与残基 24-142(12860.56)、26-142(12603.26) 和 28-142(12390.97) 吻合,均含有单个二硫键。每个峰有一个 +42Da 的相关峰 - 最可能是乙酰化的) ;c) 显示了 DMS7150 (变体 8) (分析备注 :测量质量与含有单个二硫键的残基 26-142 吻合 (12603.26)) ;d) 显示了 DMS7151 (变体 9) (分析备注 :未能对 12960 进行计算。12890.5、12603 和 12391.50 分别与残基 24-142、26-142 和 28-142 接近吻合,每个含有单个二硫键 (12862.53、12605.24 和 12392.94)。但是,测量质量与计算质量之间有 2Da 的差别) ;e) 显示了 DMS7152 (变体 10) (分析备注 :12790.5 和 12320.5 分别与残基 24-142 和 28-142 吻合,含有单个二硫键 (12790.42 和 12320.84).) f) 显示了 DMS7161 (变体 11)。
- [0071] 图 6 :显示了 GLP-1-AlbudAb 融合变体 6 的检测结果。
- [0072] 图 7 :显示了 GLP-1-AlbudAb 融合变体 11 的检测结果。
- [0073] 发明详述
- [0074] 参考实施方案,本说明书中以能够清楚和简介地描写的方式对发明进行了描述。其意图在于,并且应当明白这些实施方案可以各种方式组合或者分开,而不背离发明。
- [0075] 本文中“肽”是指经由肽键连在一起的大约两个到大约 50 个氨基酸。
- [0076] 本文中,“多肽”是指经由肽键连在一起的至少大约 50 个氨基酸。多肽通常包含三级结构,并折叠形成功能性结构域。
- [0077] 本文中,“抗蛋白酶降解”的肽或多肽 (例如域抗体 (dAb)) 与蛋白酶在适合蛋白酶活性的条件下温育时,基本不被该蛋白酶降解。当多肽与蛋白酶在适合蛋白酶活性的温度 (例如在 37 或 50℃) 温育大约 1 小时后,不超过大约 25%、不超过大约 20%、不超过大约 15%、不超过大约 14%、不超过大约 13%、不超过大约 12%、不超过大约 11%、不超过大约 10%、不超过大约 9%、不超过大约 8%、不超过大约 7%、不超过大约 6%、不超过大约 5%、不超过大约 4%、不超过大约 3%、不超过大约 2%、不超过大约 1% 的蛋白或者基本没有蛋白被降解,即多肽 (例如 dAb) 基本不被降解。蛋白质降解可以利用任何合适的方法进行评估,例如通过 SDS-PAGE 或者本文描述的功能检验法 (例如配体结合)。
- [0078] 本文中,“展示系统”是指这样的系统,其中一系列多肽或肽可供在所需特征 (比如物理、化学或功能特征) 的基础上进行挑选。展示系统可以是合适的多肽或肽的集合 (例如在溶液中、固定在合适的支持物上)。展示系统还可以是采用细胞表达系统 (例如在例如转化、感染、转染或转导的细胞中表达核酸文库,并在细胞表面展示所编码的多肽) 或者非细胞表达系统 (例如乳液区室化并展示) 的生物化学系统。优选的展示系统将核酸的编码功能和由该核酸编码的多肽或肽的物理、化学和 / 或功能特征联系在一起。采用这种展示系统时,可以挑选出具有所需物理、化学和 / 或功能特征的多肽或肽,编码所选多肽或肽的核酸可以容易地分离或回收。将核酸的编码功能和多肽或肽的物理、化学和 / 或功能

特征联系在一起许多展示系统是本领域已知的,例如,噬菌体展示(噬菌体展示)、核糖体展示、乳液区室化和展示、酵母展示、嘌呤霉素展示、细菌展示、在质粒上展示、共价展示等(参见例如 EP 0436597(Dyax)、美国专利 6,172,197(McCafferty et al.)、美国专利 6,489,103(Griffiths et al.))。

[0079] 本文中,“集合(repertoire)”是指以氨基酸序列多样性为特征的一系列多肽或肽。集合的个体成员可以具有共有特性,比如共有的结构特征(例如共有的核心结构)和/或共有的功能性特征(例如与共有的配体(例如通用配体或靶配体)结合的能力)。

[0080] 本文中,“功能性”描述了具有生物活性,比如特异结合活性的多肽或肽。例如,名词“功能性多肽”包括通过其抗原结合位点与靶抗原结合的抗体或其抗原结合片段,以及与底物结合的酶。

[0081] 本文中,“通用配体”是指与给定集合的相当一部分(例如基本全部)功能性成员结合的配体。通用配体(例如常见通用配体)可以结合给定集合中的许多成员,虽然这些成员可能对常见靶配体没有结合特异性。一般来说,多肽上存在功能性通用配体结合位点(其结合通用配体的能力所表明的)说明该多肽正确地折叠了,并且具有功能。合适的通用配体的例子包括超级抗原、与集合的相当一部分功能性成员上表达的表位结合的抗体,等。

[0082] “超级抗原”是本领域的一个名词,是指与免疫球蛋白超家族成员在某位点相互作用的通用配体,其中所述位点与这些蛋白的靶配体结合位点不同。葡萄球菌肠毒素是与 T 细胞受体相互作用的超级抗原的例子。结合抗体的超级抗原包括蛋白 G,该蛋白与 IgG 恒定区结合(Bjorck and Kronvall, J. Immunol., 133:969(1984));蛋白 A,该蛋白与 IgG 恒定区和 V_H 结构域结合(Forsgren and Sjoquist, J. Immunol., 97:822(1966));以及蛋白 L,该蛋白与 V_L 结构域结合(Bjorck, J. Immunol., 140:1194(1988))。

[0083] 本文中,“靶配体”是指被多肽或肽特异性或选择性结合的配体。例如,当多肽是抗体或其抗原结合片段时,靶配体可以是希望的任何抗原或表位;当多肽是酶时,靶配体可以是希望的任何底物。与靶抗原的结合取决于具有功能的多肽或肽。

[0084] 本文中,“抗体形式(antibody format)”是指任何合适的多肽结构,其中可以引入抗体可变域从而给该结构赋予对抗原的结合特异性。各种合适的抗体形式是本领域已知的,比如嵌合抗体、人源化抗体、人抗体、单链抗体、双特异性抗体、抗体重链、抗体轻链、抗体重链和/或轻链的同二聚体和异二聚体、任何以上分子的抗原结合片段(例如 Fv 片段(例如单链 Fv(scFv)、二硫键连接的 Fv)、Fab 片段、Fab' 片段、 $F(ab')_2$ 片段)、单抗体可变域(例如 adAb、 V_H 、 V_{HH} 、 V_L),以及任何以上分子的修饰形式(例如通过共价连接聚乙二醇或其他合适的聚合物进行的修饰)。

[0085] 短语“免疫球蛋白单个可变域”是指独立于其他 V 区或结构域,与抗原或表位特异结合的抗体可变域(V_H 、 V_{HH} 、 V_L)。免疫球蛋白单个可变域可以与其他可变区或可变域一起表现为一种形式(例如同-或异多聚体),其中所述其他区或结构域不是单免疫球蛋白可变域与抗原结合所必需的(即免疫球蛋白单个可变域不依赖其他可变域与抗原结合)。“域抗体”或“dAb”与用于本文的名词“免疫球蛋白单个可变域”相同。免疫球蛋白单个可变域优选是人抗体可变域,但也不可来自其他物种的单抗体可变域,比如啮齿动物(例如,WO 00/29004 中公开的,该文献内容通过引用全文并入本文)、护士鲨和 Camelid V_{HH} dAbs。Camelid V_{HH} 是来源于包括骆驼、羊驼、美洲驼、单峰驼和原驼的物种的免疫球蛋白单个可变

域多肽,这些物种产生天然不含有轻链的重链抗体。

[0086] “域 (domain)”是经过折叠的蛋白质结构,该结构具有与蛋白其他部分独立的三级结构。通常,结构域造成蛋白质的不同功能性性能,许多情况中可以添加、去除或转移给其他蛋白质,而不会失去蛋白剩余部分和 / 或该结构域的功能。“单抗体可变域”是含有抗体可变域特有序列的折叠的多肽结构域。因此其包括完整的抗体可变域和修饰的可变域,例如,其中一或多个环已被不是抗体可变域的特征序列所代替;或者被截短的或者包含 N- 或 C- 末端延伸的抗体可变域;以及可变域的折叠片段,所述片段至少保持着全长域的结合活性和特异性。

[0087] 名词“文库 (library)”是指同源多肽或核酸的混合物。文库由成员组成,每个成员含有单个多肽或核酸序列。从这个程度上说,“文库”与“集合”同义。文库成员间的序列差异造成文库的多样性。文库可以处于多肽或合适的简单混合物的形式,或者可以是生物体或细胞的形式,例如转化了核酸文库的细菌、病毒、动物或植物细胞等。优选地,每个个体生物体或细胞只含有 1 个或者有限数量的文库成员。有益地,核酸被并入表达载体,以便表达核酸所编码的多肽。因此,优选方面中,文库可以采取宿主生物体群体的形式,每个生物体含有一或多个拷贝的表达载体,后者含有核酸形式的单个文库成员,能够被表达而产生其对应的多肽成员。因此,宿主生物体群体有可能编码含有各种多肽的大集合。

[0088] “通用框架 (universal framework)”是与 Kabat (“Sequences of Proteins of Immunological Interest”, US Department of Health and Human Services, 1991) 定义的序列保守的抗体区域相对应,或者与 Chothia and Lesk ((1987) J. Mol. Biol. 196 :910-917) 定义的人种系免疫球蛋白集合或结构相对应的单抗体框架序列。发明提供了单框架或一组这种框架的用途,我们发现这样仅通过高变区的变化即可衍生出几乎任何结合特异性。

[0089] 氨基酸和核苷酸序列比对和如文中定义的同源性、相似性或同一性优选是利用算法 BLAST 2 Sequences, 采用缺省参数 (Tatusova, T. A. et al., FEMSMicrobiol Lett, 174 : 187-188 (1999)) 准备和确定的。

[0090] 发明涉及挑选具有所需生物活性的蛋白酶抗性肽和多肽的方法。方法中使用了至少两种选择压来产生有效地挑选具有高度稳定和抗蛋白酶降解的多肽的程序。本文中,蛋白酶抗性肽和多肽通常能够保持生物活性。相反,蛋白酶敏感肽和多肽会被本文描述的方法中的蛋白酶所切割或消化,从而失去它们的生物活性。相应地,蛋白酶抗性肽或多肽通常是基于它们的生物活性,比如结合活性来挑选的。

[0091] 文中描述的方法具有多个优点。例如,正如本文公开和举例的,挑选出来对一个蛋白酶 (例如胰蛋白酶) 的蛋白水解降解有抗性的肽或多肽也能抗其他蛋白酶 (例如弹性蛋白酶、leucozyme) 的降解。此外,蛋白酶抗性与肽或多肽的更高解链温度 (T_m) 相关。较高的解链温度表明肽和多肽更稳定。对蛋白酶降解的抗性还与对靶配体的高亲和力结合相关。因此,本文描述的方法提供了挑选、分离和 / 或回收具有所需生物活性并且因为是蛋白酶抗性和稳定的,因此非常适合体内治疗和 / 或诊断用途的多肽的有效方式。

[0092] 挑选方法

[0093] 一个方面中,发明是从肽和多肽文库或集合中挑选、分离和 / 或回收抗蛋白酶 (例如一或多种蛋白酶) 降解的肽或多肽的方法。优选地,方法是从肽和多肽文库或集合 (例如展示系统) 中挑选、分离和 / 或回收抗蛋白酶 (例如一或多种蛋白酶) 降解的多肽的方

法。通常,方法包括提供肽或多肽文库或集合;将文库或集合在有蛋白酶(例如细菌蛋白酶或者外源添加的蛋白酶,比如胰蛋白酶、弹性蛋白酶、leucozyme、胰酶、痰)的情况下,以适合蛋白酶活性的条件进行温育;以及挑选、分离和/或回收抗蛋白酶降解并具有所需生物活性的肽或多肽。被蛋白酶降解的肽或多肽通常由于蛋白酶的活性而生物活性降低或者丧失生物活性。相应地,可以利用所述方法根据它们的生物活性来挑选、分离和/或回收抗蛋白酶降解的肽或多肽,所述生物活性是比如结合活性(例如结合普通配体、结合特异配体、结合底物)、催化活性或其他生物活性。

[0094] 正如本文所描述和举例的,蛋白酶抗性 dAbs 通常与它们的靶配体以高亲和力结合。因此,发明另一方面是挑选、分离和/或回收与配体,优选靶配体以高亲和力结合的肽或多肽的方法。优选地,方法是挑选、分离和/或回收与配体,优选靶配体以高亲和力结合的多肽的方法。通常,方法包括提供肽或多肽文库或集合;在适合蛋白酶活性的条件下,将文库或集合与蛋白酶(例如胰蛋白酶、弹性蛋白酶、leucozyme、胰酶、痰)合并;以及挑选、分离和/或回收与配体(例如靶配体)结合的肽或多肽。正如本文描述的,方法还可以包括将肽或多肽文库或集合在适合展示系统内源蛋白酶(比如细菌蛋白酶,其中展示系统包括在细菌中表达)活性的条件下温育。因为文库或集合已与蛋白酶在蛋白酶敏感肽或多肽会被降解的条件下接触过,蛋白酶的活性可以除去具有低结合亲和力的较不稳定的多肽,从而得到一组高亲和力结合肽或多肽。例如,所选的肽或多肽可以与靶配体以 $1 \mu\text{M}$ 或更强的亲和力 (KD ; $KD = K_{\text{off}}(kd)/K_{\text{on}}(ka)$, 通过表面等离子共振确定) 结合, 优选大约 500nM 到大约 0.5pM 。例如,高亲和力的肽或多肽能够与靶配体以大约 500nM 、大约 100nM 、大约 10nM 、大约 1nM 、大约 500pM 、大约 100pM 、大约 10pM 、大约 1pM 或者大约 0.5pM 的亲和力结合。抗蛋白酶的肽和多肽据信具有更低的熵和/或更高的稳定能。因此蛋白酶抗性与高亲和力结合之间的相关性可能与通过本发明方法挑选到的肽或多肽的紧密性和表面稳定性有关。

[0095] 肽或多肽文库或集合与蛋白酶(例如一或多种蛋白酶)在适合蛋白酶蛋白水解活性的条件下合并。适合蛋白酶蛋白水解活性的条件,以及含有蛋白水解活性的生物制品或混合物是本领域公知的,或者可以由本领域普通技术人员很容易地确定。如果需要,可以通过例如评估蛋白酶在不同 pH 条件、蛋白酶浓度、温度和/或通过改变文库或集合与蛋白酶允许反应的时间下的活性,鉴定或者优化合适的条件。例如,一些实施方案中,蛋白酶(例如胰蛋白酶)对肽或多肽(例如可变域)的比率(基于摩尔/摩尔)是 800 到 80,000(例如 8,000 到 80,000) 蛋白酶:肽或多肽,例如使用 $10 \mu\text{g}/\text{ml}$ 蛋白酶时,比率是 800 到 80,000 蛋白酶:肽或多肽;或者使用 $100 \mu\text{g}/\text{ml}$ 蛋白酶时,比率是 8,000 到 80,000 蛋白酶:肽或多肽。在一个实施方案中,蛋白酶(例如胰蛋白酶)对肽或多肽(例如可变域)的比率(基于重量/重链,例如微克/微克)是 1,600 到 160,000(例如 16,000 到 160,000) 蛋白酶:肽或多肽,例如使用 $10 \mu\text{g}/\text{ml}$ 蛋白酶时,比率是 1,600 到 160,000 蛋白酶:肽或多肽;或者使用 $100 \mu\text{g}/\text{ml}$ 蛋白酶时,比率是 16,000 到 160,000 蛋白酶:肽或多肽。在一个实施方案中,使用的蛋白酶浓度是至少 100 或 1000 $\mu\text{g}/\text{ml}$, 蛋白酶:肽比率(基于摩尔/摩尔),即蛋白酶(例如胰蛋白酶)对肽或多肽(例如可变域)的比率是 8,000 到 80,000 蛋白酶:肽或多肽。在一个实施方案中,使用的蛋白酶浓度为至少 $10 \mu\text{g}/\text{ml}$, 蛋白酶(例如胰蛋白酶)与肽或多肽(例如可变域)的蛋白酶:肽比率(基于摩尔/摩尔)为 800 到 80,000 蛋白酶:肽或多肽。在一个实施方案中,蛋白酶(例如胰蛋白酶)对肽或

多肽（例如可变域）的比率（基于重量 / 重量，例如微克 / 微克）是 1600 到 160,000 蛋白酶：肽或多肽，例如当 C 是 10 微克 / ml；或者当 C 或 C' 是 100 微克 / ml 时，比率是 16,000 到 160,000 蛋白酶：肽或多肽。在一个实施方案中，浓度（c 或 c'）是至少 100 或 1000 微克 / ml 蛋白酶。为了检测个体或者分离的肽或多肽（例如，免疫球蛋白可变域），例如从集合或文库中分离到的肽或多肽，可以向溶于合适缓冲液（例如 PBS）的肽或多肽溶液中加入蛋白酶产生肽或多肽 / 蛋白酶溶液，比如至少大约 0.01% (w/w) 蛋白酶 / 肽或多肽，大约 0.01% 到大约 5% (w/w) 蛋白酶 / 肽或多肽、大约 0.05% 到大约 5% (w/w) 蛋白酶 / 肽或多肽、大约 0.1% 到大约 5% (w/w) 蛋白酶 / 肽或多肽、大约 0.5% 到大约 5% (w/w) 蛋白酶 / 肽或多肽、大约 1% 到大约 5% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.01% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.02% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.03% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.04% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.05% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.06% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.07% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.08% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.09% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.1% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.2% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.3% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.4% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.5% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.6% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.7% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.8% (w/w) 蛋白酶 / 肽或多肽、至少大约 0.9% (w/w) 蛋白酶 / 肽或多肽、至少大约 1% (w/w) 蛋白酶 / 肽或多肽、至少大约 2% (w/w) 蛋白酶 / 肽或多肽、至少大约 3% (w/w) 蛋白酶 / 肽或多肽、至少大约 4% (w/w) 蛋白酶 / 肽或多肽或者大约 5% (w/w) 蛋白酶 / 肽或多肽的溶液。混合液可以在适合蛋白酶活性的温度（例如室温，大约 37°C）下温育，以一定时间间隔（例如在 1 小时、2 小时、3 小时等）取样。样品可以利用任何合适的方法（比如 SDS-PAGE 分析或配体结合）来分析蛋白降解情况，结果可用于建立降解随时间的变化过程。

[0096] 本文描述的方法中可以使用任何希望的蛋白酶。例如单个蛋白酶、不同蛋白酶的任何所需组合、或者含有蛋白水解活性的任何生物制品、生物提取物或者生物匀浆。所用蛋白酶的身份不需要是已知的。可以单独或者以任何希望的组合使用的合适蛋白酶的例子包括丝氨酸蛋白酶、半胱氨酸蛋白酶、天冬氨酸蛋白酶、巯基蛋白酶、基质金属蛋白酶、羧肽酶（例如羧肽酶 A、羧肽酶 B）、胰蛋白酶、胰凝乳蛋白酶、胃蛋白酶、木瓜蛋白酶、弹性蛋白酶、leucozyme、胰酶、凝血酶、纤溶酶、组织蛋白酶（cathepsins）（例如组织蛋白酶 G）、蛋白水解酶（例如蛋白水解酶 1、蛋白水解酶 2、蛋白水解酶 3）、嗜热菌蛋白酶、凝乳酶、肠肽酶、半胱天冬氨酸酶（例如半胱天冬氨酸酶 1、半胱天冬氨酸酶 2、半胱天冬氨酸酶 4、半胱天冬氨酸酶 5、半胱天冬氨酸酶 9、半胱天冬氨酸酶 12、半胱天冬氨酸酶 13）、需钙蛋白酶、无花果蛋白酶、梭菌蛋白酶、奇异果蛋白酶、菠萝蛋白酶、分离酶和二肽基肽酶 IV (DPP-IV) 等。含有蛋白水解活性的合适的生物提取物、匀浆和制品包括血清、痰、粘液（例如胃粘液、鼻粘液、支气管粘液）、支气管肺泡灌洗液、肺匀浆、肺提取物、胰提取物、胃液、唾液或眼泪等。在一个实施方案中，蛋白酶是在眼睛和 / 或眼泪中发现的。蛋白酶的使用量适合进行蛋白水解性降解。例如，正如文中描述的，可以使用大约 0.01% 到大约 5% (w/w, 蛋白酶 / 肽或多肽) 的蛋白酶。当蛋白酶与包含肽或多肽的展示系统（例如噬菌体展示系统）合并时，蛋白酶的使用浓度可以是例如大约 10 μ g/ml 到大约 3mg/ml、大约 10 μ g/ml、大约 20 μ g/ml、

大约 30 $\mu\text{g/ml}$ 、大约 40 $\mu\text{g/ml}$ 、大约 50 $\mu\text{g/ml}$ 、大约 60 $\mu\text{g/ml}$ 、大约 70 $\mu\text{g/ml}$ 、大约 80 $\mu\text{g/ml}$ 、大约 90 $\mu\text{g/ml}$ 、大约 100 $\mu\text{g/ml}$ 、大约 200 $\mu\text{g/ml}$ 、大约 300 $\mu\text{g/ml}$ 、大约 400 $\mu\text{g/ml}$ 、大约 500 $\mu\text{g/ml}$ 、大约 600 $\mu\text{g/ml}$ 、大约 700 $\mu\text{g/ml}$ 、大约 800 $\mu\text{g/ml}$ 、大约 900 $\mu\text{g/ml}$ 、大约 1000 $\mu\text{g/ml}$ 、大约 1.5mg/ml、大约 2mg/ml、大约 2.5mg/ml 或者大约 3mg/ml。合适的浓度是大约 10 $\mu\text{g/ml}$ 到 1mg/ml ;10 $\mu\text{g/ml}$ 到 100、90、80、70、60、50 或 40 $\mu\text{g/ml}$;或者 10、20、30、40 或 50 $\mu\text{g/ml}$ 到 100、90、80、70、60 $\mu\text{g/ml}$ 。

[0097] 蛋白酶与肽或多肽集（文库或集合）在适合蛋白酶活性的温度下温育。例如，蛋白酶和肽或多肽集可以在大约 20°C 到大约 40°C 的温度（例如室温，大约 20°C、大约 21°C、大约 22°C、大约 23°C、大约 24°C、大约 25°C、大约 26°C、大约 27°C、大约 28°C、大约 29°C、大约 30°C、大约 31°C、大约 32°C、大约 33°C、大约 34°C、大约 35°C、大约 36°C、大约 37°C、大约 38°C、大约 39°C、大约 40°C）下温育。蛋白酶和肽或多肽集一起温育足够发生蛋白水解性降解的一段时间。例如，可以将肽或多肽集与蛋白酶一起温育大约 30 分钟到大约 24 或大约 48 小时。一些实施例中，肽或多肽集与蛋白酶一起温育过夜，或者至少大约 30 分钟、大约 1 小时、大约 1.5 小时、大约 2 小时、大约 3 小时、大约 4 小时、大约 5 小时、大约 6 小时、大约 7 小时、大约 8 小时、大约 9 小时、大约 10 小时、大约 11 小时、大约 12 小时、大约 13 小时、大约 14 小时、大约 15 小时、大约 16 小时、大约 17 小时、大约 18 小时、大约 19 小时、大约 20 小时、大约 21 小时、大约 22 小时、大约 23 小时、大约 24 小时、大约 48 小时或者更长。

[0098] 通常至少在早期几轮挑选（例如当使用展示系统时）中，希望与不包括和蛋白酶一起温育的挑选方法相比，蛋白酶导致具有被挑选的所需生物活性的克隆数量减少至少一个数量级。特定实施例中，方法中使用的蛋白酶的量和条件足够使回收到的克隆数量减少至少大约 1 个 \log （级数为 10）、至少大约 2 \log s（级数为 100）、至少大约 3 \log s（级数为 1000）或者至少大约 4 \log s（级数为 10,000）。利用常规方法和 / 或本文提供的教导可以容易地确定能够导致回收克隆发生所希望的减少的合适蛋白酶用量和温育条件。

[0099] 可以采用任何合适的方法（例如体外、体内或离体）将蛋白酶和肽或多肽集合并和温育。例如，可以将蛋白酶和肽或多肽集合并到合适的容器内，在适合蛋白酶活性的温度下保持静止、晃动、震荡、回旋等。如果需要，可以将蛋白酶和肽或多肽集合并到体内或离体系统内，比如通过将多肽集（例如噬菌体展示文库或集合）导入合适的动物（例如小鼠）内，并允许足够的时间令蛋白酶活性结束后，回收肽或多肽集。另一个实施例中，用多肽集（例如噬菌体展示文库或集合）灌注器官或组织，允许足够的时间令蛋白酶活性结束后，回收多肽集。

[0100] 温育之后，可以在所需生物活性（比如结合活性）的基础上挑选蛋白酶抗性肽或多肽。如果需要，可以在挑选前加入蛋白酶抑制剂。可以使用对挑选方法没有较大干扰的任何合适的蛋白酶抑制剂（或者两种或多种蛋白酶抑制剂的组合）。合适的蛋白酶抑制剂的实例包括 α 1- 抗胰蛋白酶、 α 2- 巨球蛋白、抑氨肽酶素、抗蛋白酶、抗凝血酶 III、抑肽酶、4-(2- 氨基乙基) 苯磺酰氟酸盐 (AEB SF)、(4- 脒基 - 苯基) - 甲烷 - 磺酰氟 (APMSF)、抑氨肽酶素 (bestatin)、苯甲酰胺、抑糜酶素、3,4- 二氯异香豆素、氟代磷酸二异丙酯 (DIFP)、E-64、乙二胺四乙酸 (EDTA)、弹性蛋白酶抑制剂、亮肽酶素、N- 乙基马来酰胺、苯甲磺酰氟 (PMSF)、胃酶抑素、1,10- 菲咯啉、磷酸二肽、丝氨酸蛋白酶抑制剂、N- 甲苯磺酰 -L- 赖氨酸 - 氯甲基酮 (TLCK)、Na- 甲苯磺酰 - 苯丙氨酸 - 氯甲基酮 (TPCK) 等。此外，有许多含有

多类蛋白酶的抑制剂的商业制品（例如 Roche Complete Protease Inhibitor Cocktail Tablets™ (Roche Diagnostics Corporation ;Indianapolis, IN, USA), 其抑制胰凝乳蛋白酶、嗜热菌蛋白酶、木瓜蛋白酶、链霉蛋白酶、胰腺提取物和胰蛋白酶）。

[0101] 挑选蛋白酶抗性肽或多肽可以利用所需生物活性挑选方法, 即能够将具有所需生物活性的肽和多肽与不具有所需生物活性的肽和多肽区别开并挑选出来。通常被蛋白酶消化或者切割过的肽或多肽会失去其生物活性, 而蛋白酶抗性肽或多肽能够保持功能。因此, 可以利用针对所述生物活性的合适检验法来挑选蛋白酶抗性肽或多肽。例如, 利用合适的结合检验法（例如 ELISA、淘选）可以评估常见结合功能（例如结合通用配体、结合特异配体、或者结合底物）。例如, 结合靶配体或者通用配体（比如蛋白 A、蛋白 L 或抗体）的多肽可以通过淘选或者利用合适的亲和基质来挑选、分离和 / 或回收。淘选的进行可以通过将配体（例如通用配体、靶配体）溶液加入合适的器皿（例如试管、平板）并允许配体沉积或者包被到器皿壁上。多余的配体可以洗去, 然后可以向器皿中加入多肽（例如噬菌体展示文库）, 并将器皿保持在适合多肽与固定化配体结合的条件下。可以洗去未结合的多肽, 利用诸如刮拭或者降低 pH 等任何适当方法来回收结合的多肽。

[0102] 使用噬菌体展示系统时, 可以在噬菌体 ELISA 中测试结合情况。噬菌体 ELISA 可以按照任何合适的程序进行。在一个实例中, 每轮挑选产生的噬菌体群可以通过 ELISA 来筛选与所选靶配体或通用配体的结合情况, 从而坚定到展示蛋白酶抗性肽或多肽的噬菌体。如果需要, 可以测试可溶性肽和多肽与靶配体或通用配体的结合, 例如通过使用试剂（例如针对 C- 或 N- 末端的标签）进行 ELISA（参见例如 Winter et al. (1994) Ann. Rev. Immunology 12, 433-55 以及本文引用的参考文献）。所选噬菌体的多样性还可以通过 PCR 产物的凝胶电泳 (Marks et al. 1991, 同前 ; Nissim et al. 1994 同前)、探针检测 (Tomlinson et al., 1992) J. Mol. Biol. 227, 776) 或者通过对载体 DNA 进行测序来评估。

[0103] 蛋白酶抗性肽和多肽还可以基于例如催化活性进行挑选, 其中催化活性可以利用催化活性检验法（例如蛋白水解活性检验、磷酸转移酶检验、磷酸水解酶检验、聚合酶活性检验）来测量。

[0104] 蛋白酶抗性肽或多肽（例如单抗体可变域）可以具有对通用配体或者任何希望的靶配体的结合特异性, 其中所述靶配体是比如人或动物蛋白, 包括细胞因子、生长因子、细胞因子受体、生长因子受体、酶（例如蛋白酶）、酶的辅酶、DNA 结合蛋白 ; 脂类和碳水化合物。合适的靶抗原包括细胞因子、生长因子、细胞因子受体、生长因子受体以及其他如本文描述的蛋白。应当明白, 这个清单绝不是穷举的。

[0105] 一些实施方案中, 蛋白酶抗性肽或多肽结合肺部组织中的目标分子, 比如选自以下的靶分子 : TNFR1、IL-1、IL-1R、IL-4、IL-4R、IL-5、IL-6、IL-6R、IL-8、IL-8R、IL-9、IL-9R、IL-10、IL-12、IL-12R、IL-13、IL-13R α 1、IL-13Ra2、IL-15、IL-15R、IL-16、IL-17R、IL-17、IL-18、IL-18R、IL-23、IL-23R、IL-25、CD2、CD4、CD11a、CD23、CD25、CD27、CD28、CD30、CD40、CD40L、CD56、CD138、ALK5、EGFR、FcER1、TGFb、CCL2、CCL18、CEA、CR8、CTGF、CXCL12 (SDF-1)、胰凝乳蛋白酶、FGF、弗林蛋白酶、内皮素 -1、嗜酸性粒细胞趋化因子（例如嗜酸性粒细胞趋化因子、嗜酸性粒细胞趋化因子 -2、嗜酸性粒细胞趋化因子 -3）、GM-CSF、ICAM-1、ICOS、IgE、IFNa、I-309、整合素、L- 选择素、MIF、MIP4、MDC、MCP-1、MMPs、嗜中性细胞弹性蛋白酶、骨桥蛋白、OX-40、PARC、PD-1、RANTES、SCF、SDF-1、siglec8、TARC、TGFb、凝血酶、Tim-1、

TNF、TRANCE、类胰蛋白酶、VEGF、VLA-4、VCAM、 $\alpha 4 \beta 7$ 、CCR2、CCR3、CCR4、CCR5、CCR7、CCR8、alphavbeta6、alphavbeta8、cMET、CD8、vWF、淀粉样蛋白（例如 α 淀粉样蛋白）、MMP12、PDK1 和 IgE。

[0106] 当本文描述的方法中使用了展示系统（例如将核酸的编码功能与该核酸所编码的肽或多肽的功能性特征联系起来的展示系统）的时候，通常将编码所选肽或多肽的核酸扩增或者提高其拷贝数是有益的。这样提供了有效的方法来获得足够量的核酸和 / 或肽或多肽用于以本文描述的方法或者其他合适的方法进行更多轮的挑选，或者用于制备更多集合（例如亲和力成熟集合）。因此，一些实施方案中，发明的方法包括使用展示系统（例如将核酸的编码功能与该核酸所编码的肽或多肽的功能性特征联系起来的展示系统）并且进一步包括对编码所选肽或多肽的核酸进行扩增或者提高其拷贝数。核酸扩增可以使用任何合适的方法，比如通过噬菌体扩增、细胞生长或者聚合酶链式反应。

[0107] 本文描述的方法可以作为分离蛋白酶抗性肽或多肽的程序的一部分，如果需要，所述程序可以包括其他合适的挑选方法。这类情况中，文中描述的方法可以使用在程序中的任何地方，比如在使用其他挑选方法之前或之后。本文描述的方法还可以用于如文中所述并举例的那样，提供两轮或更多轮挑选。

[0108] 发明的另一方面是制备蛋白酶抗性肽或多肽集合的方法。方法包括提供肽或多肽集合；将肽或多肽集合与蛋白酶在适合蛋白酶活性的条件下合并；以及回收具有所需生物活性的多种肽或多肽，从而得到蛋白酶抗性肽或多肽集合。优选地，具有所需生物活性的多种肽或多肽是基于结合活性回收的，比如基于与通用配体或靶配体的结合。文中相对于发明的其他方法，描述了适合用于方法中的蛋白酶、展示系统、蛋白酶活性的条件，以及挑选肽或多肽的方法。

[0109] 一些实施方案中，使用了包含肽或多肽集合的展示系统（例如将核酸的编码功能与该核酸所编码的肽或多肽的功能性特征联系起来的展示系统），方法还包括对编码所选多种肽或多肽的核酸进行扩增或者提高其拷贝数。核酸扩增可以使用任何合适的方法，比如通过噬菌体扩增、细胞生长或者聚合酶链式反应。在一个实施方案中，展示系统是噬菌体展示，扩增是通过在大肠杆菌中进行的表达。在该实施方案中，大肠杆菌中的蛋白酶表达可以提供用于挑选蛋白酶抗性肽或多肽的蛋白酶。

[0110] 具体实施方案中，发明是制备包含 dAbs 的蛋白酶抗性多肽集合的方法。方法包括提供包含 dAbs 的多肽集合；将肽或多肽集合与蛋白酶（例如胰蛋白酶、弹性蛋白酶、leucozyme）在适合蛋白酶活性的条件下合并；以及回收对通用配体（例如蛋白 A、蛋白 G、蛋白 L）或者靶配体有结合特异性并包含 dAbs 的多种多肽。方法可以用于制备新集合，或者偏向所需结合特异性的集合，比如对所需靶配体有结合特异性的基于亲代 dAb 的亲和力成熟集合。

[0111] 多肽展示系统

[0112] 优选地，提供用于本发明方法的肽或多肽集合或文库包含合适的展示系统。展示系统优选抗蛋白酶（例如单个蛋白酶或者蛋白酶的组、以及包含蛋白水解活性的任何生物提取物、匀浆或制品（例如血清、痰、粘液（例如胃粘液、鼻粘液、支气管粘液）、支气管肺泡灌洗液、肺匀浆、肺提取液、胰腺提取物、胃液、唾液、眼泪等）的降解。优选展示系统以及展示系统和被展示的多肽之间的连接物至少与集合中最稳定的肽或多肽一样抗蛋白酶。这

使得能够很容易地分离和 / 或扩增编码所选被展示多肽的核酸。

[0113] 在一个实施例中,可以从肽或多肽集合中挑选、分离和 / 或回收蛋白酶抗性肽或多肽,所述集合处于溶液中,或者是共价或非共价地附着于适宜的表面,比如塑料或玻璃(例如微量滴定板、多肽阵列(比如芯片))。例如可以使用表面肽阵列,阵列的方式使得每个不同文库成员(例如独特的肽序列)被放置在阵列中的独立、预定的位置上。这种阵列中的每个文库成员的身份可以通过它在阵列中的空间位置确定。发生靶配体和起反应文库成员之间结合相互作用的文库位置可以被确定,从而在空间位置的基础上鉴定出起反应成员的序列(参见例如美国专利 5,143,854、WO 90/15070 和 WO 92/10092)。

[0114] 优选地,方法采用的展示系统能够将核酸的编码功能与该核酸所编码的多肽的物理、化学和 / 或功能性特征联系起来。这种展示系统可以包含多种可复制的基因包装遗传包,比如噬菌体或细胞(细菌)。优选地,展示系统包含文库,比如噬菌体展示文库。噬菌体展示是特别优选的展示系统。

[0115] 许多合适的噬菌体展示系统(例如单价展示和多价展示系统)已有描述(参见例如 Griffiths et al., 美国专利 6,555,313B1(通过引用并入本文);Johnson et al., 美国专利 5,733,743(通过引用并入本文);McCafferty et al., 美国专利 5,969,108(通过引用并入本文);Mulligan-Kehoe, 美国专利 5,702,892(通过引用并入本文);Winter, G. et al., *Annu. Rev. Immunol.* 12:433-455(1994);Soumillion, P. et al., *Appl. Biochem. Biotechnol.* 47(2-3):175-189(1994);Castagnoli, L. et al., *Comb. Chem. High Throughput Screen*, 4(2):121-133(2001))。在噬菌体展示系统中展示的肽或多肽可以展示在任何合适的噬菌体上,比如丝状噬菌体(例如 fd、M13、F1)、裂解噬菌体(例如 T4、T7、 λ)或者 RNA 噬菌体(例如 MS2)。

[0116] 通常,可以制备或提供将肽或噬菌体多肽集合与合适的噬菌体衣壳蛋白(例如 fd pIII 蛋白)作为融合蛋白展示的噬菌体文库。融合蛋白可以将肽或多肽展示在噬菌体衣壳蛋白的尖端,或者如果希望可以展示在内部位置。例如,被展示的肽或多肽可以处于 pIII 的结构域 1 的氨基端位置(pIII 的结构域 1 又称为 N1)。被展示的多肽可以直接与 pIII(例如 pIII 中结构域 1 的 N 末端)融合或者利用连接分子与 pIII 融合。如果需要,融合还可以包含标签(例如 myc 表位、His 标签)。肽或多肽集合与噬菌体衣壳蛋白以融合蛋白展示,包含这种集合的文库可以利用任何合适的方法制备,比如通过将编码展示肽或多肽的噬菌体载体或噬粒载体文库导入合适的宿主细胞,并培养得到的细菌来产生噬菌体(例如根据需要利用合适的辅助噬菌体或者互补质粒)。合适地,在发明的一个实施方案中,选择适合蛋白酶在细菌中表达的条件。噬菌体文库可以利用任何合适的方法从培养物中回收,比如沉淀和离心。

[0117] 展示系统可以包含含有任何希望量的多样性的肽或多肽集合。例如,集合可以含有这样的肽或多肽,其氨基酸序列与生物体、一群生物体、所需组织或所需细胞类型表达的天然多肽对应;或者集合可以含有具有随机或被随机化的氨基酸序列的肽或多肽。如果需要,多肽可以有共同的核心或支架。例如,集合或文库中的所有多肽可以基于选自蛋白 A、蛋白 L、蛋白 G、纤连蛋白结构域、anticalin、CTLA4、所需酶(例如聚合酶、纤维素酶)、或者来自免疫球蛋白超家族的多肽(比如抗体或抗体片段,例如抗体可变域)的骨架。这种集合或文库中的多肽可以包含界定的随机或被随机化的氨基酸序列区域,和共有氨基酸序列区

域。一些实施方案中,结合中所有或者基本所有多肽都是希望的类型,比如希望的酶(例如聚合酶)或者希望的抗体的抗原结合片段(例如人 V_H 或人 V_L)。在优选实施方案中,多肽展示系统包含其中每个多肽都含有抗体可变域的多肽集合。例如,集合中的每个多肽可以含有 V_H 、 V_L 或者Fv(例如单链Fv)。正如本文描述的,集合可以是基于亲代分子,比如GLP-1或其衍生物(比如二肽基肽酶IV-抗性衍生物)的多肽的文库。

[0118] 可以利用任何合适的方法给肽或多肽的任何所需区域或骨架引入氨基酸序列多样性。例如,可以利用任何合适的突变方法(例如低忠实度PCR、寡核苷酸介导的突变或者定点突变、利用NNK密码子多样化)或者任何其他合适的方法,通过制备编码多样化多肽的核酸文库给靶区域,比如抗体可变域的互补决定区或者疏水结构域,引入氨基酸序列多样性。如果需要,可以将希望多样化的多肽区域随机化。

[0119] 构成集合的多肽的大小基本是个选择的问题,不需要统一的多肽大小。优选地,集合中的多肽至少具有三级结构(形成至少一个结构域)。

[0120] 挑选 / 分离 / 回收

[0121] 蛋白酶抗性肽或多肽(例如蛋白酶抗性多肽群体)可以利用任何合适的方法从集合或文库(例如在展示系统中)中挑选、分离和 / 或回收。优选地,蛋白酶抗性多肽是在可选择特性(例如物理特性、化学特性、功能性特征)的基础上挑选或分离的。合适的可选择功能特征包括集合中的肽或多肽的生物活性,例如,结合通用配体(例如超级抗原)、结合靶配体(例如抗原、表位、底物)、结合抗体(例如经由表达在肽或多肽上的表位)和催化活性(参见例如Tomlinson et al.、WO 99/20749、WO 01/57065、WO 99/58655)。

[0122] 一些实施方案中,蛋白酶抗性肽或多肽挑选和 / 或分离自这样的肽或多肽文库或集合,其中基本所有蛋白酶抗性肽或多肽拥有共同的可选择特征。例如,蛋白酶抗性肽或多肽可以选自文库或集合,其中基本所有蛋白酶抗性肽或多肽能够结合共有的通用配体、结合共有的靶配体、结合共有抗体(或者被其结合)、或者拥有相同的催化活性。这种类型的挑选特别可用于在具有所需生物活性的亲代肽或多肽的基础上制备蛋白酶抗性肽或多肽集合,例如,进行免疫球蛋白单个可变域的亲合力成熟时。

[0123] 基于与共有通用配体的结合进行选择可以产生含有是原来文库或集合的全体成员的全部或基本全部蛋白酶抗性肽或多肽的肽或多肽集或群体。例如,结合靶配体或通用配体(比如蛋白A、蛋白L或抗体)的肽或多肽可以通过淘选或者利用合适的亲和基质来挑选、分离和 / 或回收。淘选的实现可以通过将配体(例如通用配体、靶配体)溶液加入合适的器皿,并允许配体沉积或者包被到器皿壁上。多余的配体可以洗去,然后可以向器皿中加入肽或多肽(例如已经和蛋白酶一起温育过的集合),并将器皿保持在适合肽或多肽与固定化配体结合的条件下。可以洗去未结合的多肽,利用诸如刮拭或者降低pH等任何适当方法来回收结合的肽或多肽。

[0124] 合适的配体亲和基质通常含有固态支持物或者柱子(例如琼脂糖),配体可以共价或非共价地附着其上。亲和基质可以利用批量程序、过柱程序或者其他合适的程序,在适合肽或多肽结合到基质上的配体上的条件下,与肽或多肽(例如已经和蛋白酶一起温育过的集合)合并在一起。没有与亲和基质结合的肽或多肽可以洗去,结合的肽或多肽可以利用任何适当的方法洗脱并回收,所述方法是比如用较低pH缓冲液、用温和的变性剂(例如脲)、或者用竞争结合配体的肽来洗脱。在一个实例中,生物素化的靶配体与集合在适合集

合中的肽或多肽结合靶配体的条件下合并。利用固定化的抗生物素蛋白或者链霉抗生物素蛋白（例如在珠子上）回收结合的肽或多肽。

[0125] 一些实施方案中,通用或靶配体是抗体或其抗原结合片段。能够结合文库和集合中的肽或多肽中相当保守的肽或多肽结构特征的那些抗体或抗原结合片段尤其适合作为通用配体。适合作为配体用于分离、挑选和 / 或回收蛋白酶抗性肽或多肽的抗体以及抗原结合片段可以是单价或多价的,并且可以利用任何适当的方法来制备。

[0126] 文库 / 集合

[0127] 发明其他方面涉及蛋白酶抗性肽和多肽集合,涉及编码蛋白酶抗性肽和多肽的文库,还涉及制备这些文库和结合的方法。

[0128] 编码和 / 或含有蛋白酶抗性肽和多肽的文库可以利用任何合适的方法来制备或得到。利用本文描述的方法,可以在目标肽或多肽（例如选自文库的肽或多肽）的基础上设计本发明的文库编码蛋白酶抗性肽或多肽,或者文库可以选自另一个文库。例如,可以利用合适的多肽展示系统来制备富含蛋白酶抗性多肽的文库。

[0129] 在一个实例中,如本文所述,将噬菌体展示文库与蛋白酶在适合蛋白酶活性的条件下合并,所述展示文库包含含有免疫球蛋白单个可变域（例如 V_H 、 V_K 、 V_L ）的展示多肽的集合。基于所需生物活性,比如结合活性（例如结合通用配体、结合靶配体）回收蛋白酶抗性多肽,从而产生富含蛋白酶抗性多肽的噬菌体展示文库。

[0130] 另一个实例中,首先筛选包含展示多肽（含有免疫球蛋白单个可变域（例如 V_H 、 V_K 、 V_L ））集合的噬菌体展示文库来鉴定集合中对所需靶抗原具有结合特异性的成员。如文中所述,回收具有所需结合特异性的多肽集,并将多肽集与蛋白酶在适合蛋白水解活性的条件下合并。回收具有所需靶结合特异性的蛋白酶抗性多肽集,得到富含蛋白酶抗性和高亲和力多肽的文库。正如文中描述过的,本挑选方法中蛋白酶抗性与高亲和力结合相关联。

[0131] 利用任何合适的方法都可以容易地产生编码所需类型多肽的集合的文库。例如,可以获得编码所需类型多肽（例如聚合酶、免疫球蛋白可变域）的核酸序列,通过利用易错聚合酶链式反应 (PCR) 系统扩增核酸、通过化学突变 (Deng et al., J. Biol. Chem., 269 : 9533(1994)) 或者使用细菌致突变株 (Low et al., J. Mol. Biol., 260 :359(1996)) 可以制备每个含有一或多个突变的核酸的集。

[0132] 其他实施方案中,可以针对核酸的特定区域进行多样化。突变选定位点的方法也是本领域公知的,包括例如,采用错配寡核苷酸或者简并寡核苷酸,使用或者不使用 PCR。例如,已经通过对抗原结合环进行靶向突变制备到合成抗体文库。将随机或半随机抗体 H3 和 L3 区域添加到种系免疫球蛋白 V 基因区段上产生带有未突变框架区的大文库 (Hoogenboom and Winter(1992) 同前 ;Nissim et al.(1994) 同前 ;Griffiths et al.(1994) 同前 ;DeKruif et al.(1995) 同前)。这类多样化方法已扩展到包括一些或全部其他抗原结合环 (Cramer et al.(1996) Nature Med., 2 :100 ;Riechmann et al.(1995) Bio/Technology, 13 :475 ;Morphosys, WO 97/08320, 同前)。在其他实施方案中,可以通过例如两步 PCR 策略靶向核酸的特定区域进行多样化,这种策略中把第一次 PCR 的产物作为“大引物”（参见例如 Landt, O. et al., Gene 96 :125-128(1990)）。定向多样化还可以通过例如 SOE PCR 来实现（参见例如 Horton, R. M. et al., Gene 77 :61-68(1989)）。

[0133] 通过改变限定多肽序列的编码序列,从而给该位点引入多种可能的氨基酸（例如

全部 20 种或者其中一部分), 这样即可达到选定位点的序列多样化。采用 IUPAC 命名法, 最多功能的密码子是 NNK, 其编码所有的氨基酸和 TAG 终止密码子。优选使用 NNK 密码子来引入希望的多样性。也可以使用能够达到相同目的的其他密码子, 包括 NNN 密码子, 该密码子导致产生另外的终止密码子 TGA 和 TAA。这类定向手段可以保证目标区域内的整个序列空间得到研究。

[0134] 优选的文库包含是免疫球蛋白超家族成员的蛋白酶抗性多肽(例如抗体或其部分)。例如文库可以包含具有已知主链构象的蛋白酶抗性抗体多肽(参见例如 Tomlinson et al., WO 99/20749)。文库可以在合适的质粒或载体中制备。本文中, 载体是指用于将异源 DNA 导入细胞以便进行表达和 / 或复制的独特元件。可以使用任何合适的载体, 包括质粒(例如细菌质粒)、病毒或噬菌体载体、人工染色体和附加型载体。这些载体可以用于简单克隆和突变, 或者表达载体可以用于驱动文库的表达。载体和质粒通常含有一或多个克隆位点(例如多位点接头)、复制起点和至少一个选择标记基因。表达载体还可以含有驱动多肽的转录和翻译的元件, 比如增强子元件、启动子、转录终止信号、信号序列等。这些元件的排列方式使得它们可操纵地连接着克隆的编码多肽的插入片段, 这样当这类表达载体被维持在适合表达的条件(例如在合适的宿主细胞中)下时, 多肽将被表达和产生。

[0135] 克隆和表达载体通常含有使载体能够在一或多种所选宿主细胞内发生复制的核酸序列。一般在克隆载体中, 这个序列是使载体能够独立于宿主染色体 DNA 而复制的序列, 其包括复制起点或自主复制序列。已知在多种细菌、酵母和病毒中有这类序列。来自质粒 pBR322 的复制起点适用于多数革兰氏阴性细菌, 2 μ 质粒起点适用于酵母, 多种病毒起点(例如 SV40、腺病毒)可以用于哺乳动物细胞中的克隆载体。一般来说, 哺乳动物表达载体不需要复制起点, 除非这些载体要用在能够高度复制 DNA 的哺乳动物细胞(比如 COS 细胞)内。

[0136] 克隆或表达载体可以含有选择基因, 又称为选择标记。这类标记基因编码在选择性培养基中生长的转化宿主细胞存活或生长所必需的蛋白。因此未转化含有选择基因的载体的宿主细胞不能在培养基中存活。典型的选择基因编码赋予对抗生素和其他毒素的抗性的蛋白, 例如氨苄青霉素、新霉素、甲基喋呤或四环素; 补偿营养缺陷; 或者提供生长培养基中未提供的关键营养物。

[0137] 合适的表达载体可以含有多种组分, 例如复制起点、选择标记基因、一或多个表达调控元件(比如转录调控元件, 例如启动子、增强子、终止子)和 / 或一或多个翻译信号(比如信号序列或前导序列)等。表达调控元件和信号或前导序列如果存在, 可以由载体或其他来源提供。例如, 编码抗体链的克隆核酸中的转录和 / 或翻译调控序列可以用于引导表达。

[0138] 可以给在希望的宿主细胞内进行的表达提供启动子。启动子可以是组成型的或者诱导型的。例如, 启动子可以可操纵地连接着编码抗体、抗体链或其部分的核酸, 因此它引导所述核酸的转录。有大量适合原核细胞(例如对于大肠杆菌 β -内酰胺酶和乳糖启动子系统; 碱性磷酸酶; 色氨酸(trp)启动子系统; lac、tac、T3、T7 启动子)和真核细胞(例如猴病毒 40 早期或晚期启动子、Rous 肉瘤病毒长末端重复启动子、巨细胞病毒启动子、腺病毒晚期启动子、EG-1a 启动子)宿主的启动子可以使用。

[0139] 此外, 表达载体一般包含用于挑选携带载体的宿主细胞的选择标记, 并且在可复

制表达载体的情况中,还含有复制起点。编码赋予抗生素或药物抗性的产物的基因是常见的选择标记,可以用于原核(例如 β -内酰胺酶基因(氨苄青霉抗性)、赋予四环素抗性的Tet基因)和真核细胞(例如新霉素(G418或遗传霉素(geneticin))、gpt(霉酚酸)、氨苄青霉或者潮霉素抗性基因)。二氢叶酸还原酶标记基因允许在多种宿主内用氨甲喋呤进行挑选。编码宿主营养缺陷型标记的基因产物(例如LEU2、URA3、HIS3)的基因常在酵母中用作选择标记。还考虑了病毒(例如杆状病毒)或噬菌体载体,以及能够整合到宿主细胞染色体中的载体,比如逆转录载体。

[0140] 适合在原核(例如细菌细胞,比如大肠杆菌)或哺乳动物细胞内进行表达的表达载体包括例如,pET载体(例如pET-12a、pET-36、pET-37、pET-39、pET-40,Novagen和其他公司)、噬菌体载体(例如pCANTAB 5E,Pharmacia)、pRIT2T(蛋白A融合载体,Pharmacia)、pCDM8、pCDNA1.1/amp、pCDNA3.1、pRc/RSV、pEF-1(Invitrogen,Carlsbad,CA)、pCMV-SCRIPT、pFB、pSG5、pXT1(Stratagene,La Jolla,CA)、pCDEF3(Goldman,L.A.,et al.,Biotechniques,21:1013-1015(1996))、pSVSPORT(GibcoBRL,Rockville,MD)、pEF-Bos(Mizushima,S.,et al.,Nucleic Acids Res.,18:5322(1990))等。有适合在各种表达宿主,比如原核细胞(大肠杆菌)、昆虫细胞(果蝇(*Drosophila*)Schnieder S2细胞、Sf9)、酵母细胞(嗜甲醇毕赤酵母(*P. methanolica*)、巴斯德毕赤酵母(*P. pastoris*)、酿酒酵母)和哺乳动物细胞(例如COS细胞)中使用的表达载体。

[0141] 优选的载体是能够表达对应多肽文库成员的核苷酸序列的表达载体。因此,以通用和/或靶载体进行的挑选可以通过分开的增殖和表达多肽文库成员的单个克隆的表达来进行。正如上文描述的,优选的挑选展示系统是噬菌体展示。因此,可以使用噬菌体或噬粒载体。优选的载体是含有大肠杆菌复制起点(用于双链复制)和噬菌体复制起点(用于产生单链DNA)的噬粒载体。这类载体的操作和表达是本领域公知的(Hoogenboom and Winter(1992)同前;Nissim et al.(1994)同前)。简单来说,载体可以含有 β -内酰胺酶基因以赋予对噬粒的选择性,并且在表达盒上游含有lac启动子,表达盒可以含有合适的前导序列、多克隆位点、一或多个肽标签、一或多个TAG终止密码子和噬菌体蛋白pIII。这样,利用大肠杆菌的各种抑制因子和非抑制因子菌株,并添加葡萄糖、异丙基硫代- β -D-半乳糖苷(IPTG)或辅助噬菌体(比如VCSM13),载体能够在不表达的情况下像质粒一样复制,产生大量多肽文库成员或者噬菌体产物,其中的一些在其表面含有至少一个拷贝的多肽-pIII融合。

[0142] 本发明的文库和集合可以含有多种抗体形式。例如,文库和集合内含有的多肽可以是整个抗体或其片段,比如Fab、F(ab)₂、Fv或scFv片段、分开的V_H或V_L结构域,其中的任何一种带有修饰或未被修饰。scFv片段,以及其他抗体多肽可以利用任何适当的方法容易地产生。许多合适的抗体工程化方法是本领域已知的。例如,scFv可以通过将编码两个可变域的核酸和合适的寡核苷酸连接起来而形成,所述寡核苷酸编码适当的连接肽,比如(Gly-Gly-Gly-Gly-Ser)₃或其他合适的连接肽。连接肽连接着第一个V区的C末端和第二个V区的N末端。可以使用类似的技术构建其他抗体形式,比如Fv、Fab和F(ab)₂片段。为了形成Fab和F(ab)₂片段,可以将V_H和V_L多肽与恒定区区段组合,后者可以分离自重排基因、种系C基因或者由抗体序列信息合成。根据本发明的文库或集合可以是V_H或V_L文库或集合。

[0143] 包含蛋白酶抗性可变域的多肽优选包含靶配体结合位点和 / 或通用配体结合位点。一些实施方案中,通用配体结合位点是超级抗原(比如蛋白 A、蛋白 L 或蛋白 G)的结合位点。所述可变域可以基于任何所需可变域,例如人 V_H (例如 V_H1a 、 V_H1b 、 V_H2 、 V_H3 、 V_H4 、 V_H5 、 V_H6)、人 V_L (例如 $V_L I$ 、 $V_L II$ 、 $V_L III$ 、 $V_L IV$ 、 $V_L V$ 、 $V_L VI$ 或 $V_K 1$) 或人 V_K (例如 $V_K 2$ 、 $V_K 3$ 、 $V_K 4$ 、 $V_K 5$ 、 $V_K 6$ 、 $V_K 7$ 、 $V_K 8$ 、 $V_K 9$ 或 $V_K 10$)。

[0144] 核酸、宿主细胞和产生蛋白酶抗性多肽的方法

[0145] 发明还涉及分离的和 / 或重组的核酸,所述核酸编码可以通过本文描述的方法挑选的或者通过本文描述的方法挑选到的蛋白酶抗性肽或多肽。

[0146] 文中被称为“分离的”是已经与其天然环境(例如在细胞或者诸如文库的核酸混合物中)中的其他物质(例如其他核酸,比如基因组 DNA、cDNA 和 / 或 RNA)分开的核酸。分离的核酸可以作为载体(例如质粒)的一部分而分离。

[0147] 文中被称为“重组的”是通过重组 DNA 方法产生的核酸和利用聚合酶链式反应(PCR)制备的核酸,所述重组 DNA 方法包括依赖人工重组,比如利用限制性内切酶、同源重组、病毒等克隆到载体或者染色体中。

[0148] 发明还涉及包含(一或多个)重组核酸或者表达构建体的重组宿主细胞,所述表达构建体包含编码例如可以通过本文描述的方法挑选的或者通过本文描述的方法挑选到的蛋白酶抗性肽或多肽的核酸。发明还包括制备蛋白酶抗性肽或多肽的方法,包括将本发明的重组宿主细胞维持在适合蛋白酶抗性肽或多肽表达的条件下。如果需要,方法还可以包括分离或回收所述蛋白酶抗性肽或多肽的步骤。

[0149] 例如,可以利用适合所选宿主细胞的任何方法(例如转化、转染、电穿孔、感染),将编码蛋白酶抗性肽或多肽的核酸分子(即,一或多个核酸分子)或者包含所述核酸分子的表达构建体(即一或多个构建体)导入合适的宿主细胞,从而产生重组宿主细胞,使得核酸分子可操纵地连接着一或多个表达调控元件(例如在载体中,在通过细胞内的过程形成的构建体中,整合到宿主细胞染色体中)。得到的重组宿主细胞可以保持在适合表达的条件下(例如在有诱导物的情况下,在合适的动物中,在合适的补充了适当的盐、生长因子、抗生素、营养物等的培养基中),从而产生所编码的肽或多肽。如果需要,可以(例如从动物、宿主细胞、培养基、奶汁中)分离或回收所编码的肽或多肽。这个过程包括在转基因动物的宿主细胞内进行表达(参见例如 W092/03918, GenPharm International)。

[0150] 通过本文描述的方法挑选到的蛋白酶抗性肽或多肽还可以在合适的体外表达系统中产生,通过化学合成或者其他合适的方法产生。

[0151] 多肽、dAbs、激动剂 & 拮抗剂

[0152] 正如文中描述和举例的,本发明的蛋白酶抗性多肽、肽或 dAbs 通常与其靶配体以高亲和力结合。因此,另一方面,发明提供了挑选、分离和 / 或回收与靶抗原以高亲和力结合的发明的多肽或 dAb。一般来说,方法包括提供肽或多肽(例如 dAbs)文库或集合、将文库或集合与蛋白酶(例如胰蛋白酶、弹性蛋白酶、leucozyme、胰酶、痰)在适合蛋白酶活性的条件下合并,和挑选、分离和 / 或回收与配体(例如靶配体)结合的肽或多肽。因为文库或集合已经与蛋白酶在蛋白酶敏感肽或多肽会被消化的条件下接触过,蛋白酶的活性会除去具有低结合亲和力的较不稳定多肽,从而得到高亲和力结合肽或多肽集。例如,本发明的多肽或 dAb 能够以 $1 \mu M$ 或更强、或者大约 $500nM$ 到大约 $0.5pM$ 的亲和力 ($K_D; K_D = K_{off}(kd) /$

K_{on} (ka), 通过表面等离子共振确定的) 结合靶抗原。例如, 发明的多肽或 dAb 能够以大约 500nM、大约 100nM、大约 10nM、大约 1nM、大约 500pM、大约 100pM、大约 10pM、大约 1pM 或大约 0.5pM 的亲和力结合靶抗原 (例如 TNFR1)。虽然我们不拘束于任何特定理论, 抗蛋白酶的肽和多肽被认为具有较低的熵和 / 或较高的稳定能。因此, 蛋白酶抗性和高亲和力结合之间的相关性可能与通过本文描述的方法挑选到的肽或多肽以及 dAbs 的完整性和稳定性有关。

[0153] 多肽、dAb、激动剂或拮抗剂可以在大肠杆菌或毕赤酵母 (例如巴斯德毕赤酵母) 中表达。在一个实施方案中, 当在大肠杆菌或毕赤酵母 (例如巴斯德毕赤酵母) 中表达时, 配体或 dAb 单体至少分泌大约 0.5mg/L。尽管本文描述的配体和 dAb 单体在大肠杆菌或毕赤酵母 (例如巴斯德毕赤酵母) 中表达时可以被分泌出来, 它们也可以利用任何合适的方法来产生, 比如合成化学方法或者不使用大肠杆菌或毕赤酵母的生物制备方法。

[0154] 一些实施方案中, 多肽、dAb、激动剂或拮抗剂不包含 Camelid 免疫球蛋白可变域, 或者对 Camelid 种系抗体基因区段所编码的免疫球蛋白可变域来说独特的一或多个框架氨基酸, 例如在位点 108、37、44、45 和 / 或 47。

[0155] 发明的激动剂或拮抗剂可以是单价或多价的。一些实施方案中, 激动剂或拮抗剂是单价的, 含有与靶抗原相互作用的一个结合位点, 所述结合位点由发明的多肽或 dAb 提供。单价激动剂或拮抗剂结合一个靶抗原, 不能诱导细胞表面的靶抗原 (例如受体抗原) 发生会导致受体和信号转导被激活的交联或聚集。

[0156] 在其他实施方案中, 发明的激动剂或拮抗剂是多价的。多价激动剂或拮抗剂可以含有两个或多个拷贝的靶抗原特定结合位点, 或者含有两个或多个结合靶抗原的不同结合位点, 所述结合位点中的至少一个由本发明的多肽或 dAb 提供。例如, 如本文所述, 激动剂或拮抗剂可以是这样的二体、三体或者多体, 其包含两个或更多个拷贝的结合靶抗原的发明所述特定多肽或 dAb, 或者是包含两个或更多个不同的结合靶抗原的发明所述多肽或 dAbs。在一个实施方案中, 标准细胞检验中, 多价拮抗剂结合了细胞表面受体抗原, 未能较大地给抗原促效 (表现为抗原的激动剂)。

[0157] 一些实施方案中, 多价激动剂或拮抗剂含有两个或更多个结合所需表位或靶抗原结构域的位点。

[0158] 其他实施方案中, 多肽可能是促胰岛素剂 (insulinotropic agent), 比如 GLP-1 衍生肽。确定促胰岛素剂的效力、对诸如 DPP-IV 的蛋白酶的抗性、给药后的半衰期以及体内效果的合适方法参见例如 WO 2006/059106。

[0159] 其他实施方案中, 多价激动剂或拮抗剂含有由本发明的多肽或 dAbs 提供的两个或更多个结合位点, 所述结合位点结合靶抗原的不同表位或结构域。

[0160] 一些实施方案中, 当给予足够量时, 发明的多肽、dAb、激动剂或拮抗剂在慢性炎症性疾病模型中有药效。通常, 有效量是大约 1mg/kg 到大约 10mg/kg (例如大约 1mg/kg、大约 2mg/kg、大约 3mg/kg、大约 4mg/kg、大约 5mg/kg、大约 6mg/kg、大约 7mg/kg、大约 8mg/kg、大约 9mg/kg 或者大约 10mg/kg)。本领域技术人员认同慢性炎症性疾病模型 (参见 WO2006038027 中描述的那些) 可以预测对人的治疗效果。

[0161] 通常, 所述配体 (例如激动剂、拮抗剂) 会以纯化形式与药物适合的载体一起使用。一般地, 这些载体包括水或醇 / 水溶液、乳液或悬浮液, 包括盐和 / 或缓冲培养基。肠

胃外媒介包括氯化钠溶液、林格氏 (Ringer' s) 葡萄糖、葡萄糖和氯化钠和乳酸林格氏液。如果需要将多肽复合体保持在悬浮液中, 可以从诸如羧甲基纤维素、聚乙烯吡咯烷酮、明胶和海藻酸等增稠剂中选择合适的生理可接受佐剂。

[0162] 静脉内媒介包括液体和营养补充剂以及电解质补充剂, 比如基于林格氏葡萄糖的那些。还可以有防腐剂和添加剂, 比如抗微生物剂、抗氧化剂、螯合剂和惰性气体 (Mack(1982) Remington' s Pharmaceutical Sciences, 16th Edition)。有多种合适的剂型可以使用, 包括缓释剂型。

[0163] 本发明的配体 (例如拮抗剂) 可以作为单独给予的组合物或者与其他药剂联合使用。包括各种免疫治疗药物, 比如环孢素、氨甲喋呤、阿霉素或顺铂, 以及免疫毒素。药物组合物可以包括各种细胞毒剂或其他试剂连同本发明的配体, 或者甚至是本发明具有不同特异性的配体的组合所构成的“鸡尾酒混合液”, 所述配体是比如利用不同靶抗原或表位挑选到的配体, 不管它们是否在给药前汇合在一起。

[0164] 本发明的药物组合物的给药途径可以是本领域技术人员公知的任何途径。用于疗法, 包括但不限于免疫疗法, 发明所选的抗体可以按照标准技术给予任何患者。

[0165] 给药可以通过任何合适的方式, 包括肠胃外、静脉内、肌内、腹腔内、透皮、经肺部途径、或者也适合借助导管通过直接输液。给药的剂量和频率取决于患者的年龄、性别和身体状况, 是否同时给予其他药物, 以及临床医师要考虑的其他参数。给药可以是局部的 (例如通过肺部给药局部递送到肺, 例如鼻内给药) 或者如指明的是系统性给药。

[0166] 本发明的配体可以冻干以便保存, 然后使用前在合适的载体中重新配制。该技术已证实对常规免疫球蛋白是有效的, 可以采用领域公知的冻干和重组技术。本领域技术人员明白, 冻干和重组可以造成抗体活性不同程度的损失 (例如对于常规免疫球蛋白, IgM 抗体倾向比 IgG 抗体有更大的活性损失), 因此用量水平可能需要上调来补偿损失。

[0167] 含有所述配体 (例如激动剂、拮抗剂) 或其混合液的组合物可以用于预防和 / 或治疗的给药。一些治疗应用中, 足够对所选细胞群体实现至少部分抑制、压抑、调整、杀死或其他可测量参数的量被定义为“治疗有效量”。达到这个剂量所选的用量取决于疾病的严重程度和患者自身免疫系统的整体状况, 但一般在每公斤体重 0.005 到 5.0mg 范围内的配体, 例如 dAb、激动剂或拮抗剂, 其中更常用的是 0.05 到 2.0mg/kg/ 用药的剂量。对于预防应用, 含有所述配体或其混合液的组合物也可以按照类似或稍低的剂量给药来预防、抑制疾病或者延迟发病 (例如维持缓解或静止期, 或者防止急性期)。临床医师能够确定合适的用药间隔来治疗、抑制或防止疾病。如果与治疗前的症状相比, 或者相对没有用所述组合物治疗的个体或者其他合适的对照中的症状, 一或多种症状减轻了 (例如减轻至少 10% 或临床评估量表中的至少一点), 则利用本文描述的组合物进行的治疗或疗法被认为是“有效的”。症状显然根据针对的疾病或紊乱而不同, 但可以由普通的临床医师或技术人员来测量。这类症状的测量可以通过例如监测所述疾病或紊乱的一或多种生化指征 (例如与该疾病相关的酶或代谢物的水平、受影响的细胞数量等)、监测物理征象 (例如炎症、肿瘤大小等), 或者通过经认可的临床评估量表, 例如扩展残疾状态量表 (Expanded Disability Status Scale) (用于多发性硬化症)、Irvine 炎症性肠病问卷 (Inflammatory Bowel Disease Questionnaire) (32 点评估就肠功能、系统症状、社会功能和情绪状态来评价生活质量 - 评分在 32 到 224 的范围内, 较高评分表明生活质量更好)、类风湿性关节炎生活质量量表

(Quality of Life Rheumatoid Arthritis Scale)、或者本领域已知的其他经认可的临床评估量表。疾病或者紊乱的症状持续(例如一天或更多天,或者更长时间)减轻至少 10% 或者给定临床评级减小一或多个点,指示“有效的”治疗。类似地,如果与没有用组合物治疗的类似个体(人或动物模型)中的症状相比,一或多种症状的发作或严重程度被延迟、减轻或消除,则利用本文描述的组合物进行的预防是“有效的”。

[0168] 本发明含有配体(例如激动剂、拮抗剂)或其混合液的组合物可以用于预防和治疗情景来协助改变、灭活、杀死或除去哺乳动物中的选定靶细胞群。此外,本文描述的选定多肽集合可以身体外(extracorporeally)或者体外(invitro)使用,从异源细胞集合中将靶细胞群选择性地杀死、耗尽或者有效地去除。可以将来自哺乳动物的血在身体外与配体组合,从而不需要的细胞被杀死或者以其他方式从血液中除去,按照标准技术将血液返回哺乳动物。

[0169] 本发明的含有配体(例如激动剂或拮抗剂)的组合物可以用于预防和治疗中来协助改变、灭活、杀死或除去哺乳动物中的选定靶细胞群。

[0170] 配体(例如抗靶抗原拮抗剂、激动剂、dAb 单体)可以与一或多种其他治疗或活性试剂一起给药或成剂。当配体(例如 dAb)与其他治疗剂一起给药时,配体可以在给予其他试剂之前、同时或随后给予。通常,配体和其他试剂的给药方式使得它们的治疗效果有重叠。

[0171] 在发明的优选实施方案中,本发明的含有 GLP-1 药或者 GLP-1 类似物或衍生物的药物组合物可以经肠胃外给予需要这种治疗的患者。肠胃外给药可以通过借助注射器,任选类似钢笔的注射器经皮下、肌肉或静脉内注射来执行。替代地,肠胃外给药可以借助输液泵来实行。另一个选择是可能是粉末或液体的组合物用于以鼻腔或肺喷雾剂的形式给予 GLP-1 药或者 GLP-1 类似物或衍生物。作为再一个选择,本发明的 GLP-1 药或 GLP-1 类似物或衍生物也可以经皮给药(例如由贴片(patch),任选离子导入贴片),或者跨粘膜(例如经颊粘膜)。其他实施方案中,组合物作为药片、胶囊、饮料(例如作为治疗肥胖的减肥饮料上市)来口服。

[0172] 用于肠胃外给予 GLP-1 化合物的组合物可以例如如 WO 03/002136(通过引用并入本文)中所述制备。

[0173] 另一个实施方案中,本发明涉及发明的化合物在制备治疗高血糖、1 型糖尿病、2 型糖尿病或者 β -细胞缺陷的药物中的用途。在关于这些适应症的具体实施方案中,药物选自促胰岛素分泌剂和肠促胰岛素、胰升糖素样 1 肽、GLP-1 肽、GLP-1 类似物、GLP-1 衍生物、PYY、PYY 肽、PYY 类似物、PYY 衍生物、毒蜥外泌肽 -3、毒蜥外泌肽 -3 肽、毒蜥外泌肽 -3 类似物、毒蜥外泌肽 -3 衍生物、毒蜥外泌肽 -4、毒蜥外泌肽 -4 肽、毒蜥外泌肽 -4 类似物、毒蜥外泌肽 -4 衍生物,或者它们中两种或多种的组合(例如 GLP-1 肽和 PYY 肽)。

[0174] 用本发明的化合物进行的治疗还可以与第二或更多药物活性物质联用,所述活性物质可以是或者不是偶联或者融合药物的一部分。例如,活性物质选自抗糖尿病药、抗肥胖药、食欲调节药、抗高血压药、用于治疗 and / 或防止由糖尿病导致的或者与糖尿病相关的并发症,以及用于治疗 and / 或防止由肥胖导致的或者与肥胖相关的并发症和紊乱。在本文语境中,表达方式“抗糖尿病药”包括用于治疗 and / 或预防胰岛素抗性以及胰岛素抗性是其病理学机制的疾病的化合物。

[0175] 各种形式 (FORMATS)

[0176] 提高的半衰期对免疫球蛋白,特别是抗体,更尤其是体积小的抗体片段的体内应用是有益的。这类片段 (Fvs、二硫键连接的 Fvs、Fabs、scFvs、dAbs) 遭遇被快速从身体清除,因此,尽管它们能够很快地到达身体的多数部位,生产快,易于操作,但只能在体内短期驻留限制了它们的体内应用。发明的一个实施方案解决这个问题是通过提供体内半衰期提高的配体,从而在体内配体的功能活性可以坚持更长时间。

[0177] 药代动力学分析和确定配体半衰期的方法是本领域技术人员熟悉的。可以在 Kenneth, A et al: *Chemical Stability of Pharmaceuticals: A Handbook for Pharmacists* 和 Peters et al, *Pharmacokinetic analysis: A Practical Approach* (1996) 中找到详细描述。还可参考“Pharmacokinetics”, M Gibaldi & D Perron, published by Marcel Dekker, 2nd Rev. edition (1982), 该书描述了药代动力学参数,比如 t_{α} 和 t_{β} 半衰期以及曲线下面积 (AUC)。

[0178] 半衰期 ($t_{1/2\alpha}$ 和 $t_{1/2\beta}$) 和 AUC 可以由配体的血清浓度相对时间的曲线来确定。可以利用 WinNonlin 分析包 (可以从 Pharsight Corp., Mountain View, CA94040, USA 购买) 来给曲线建模。在第一期 (α 期), 配体主要进行患者体内分布,和一些清除。第二期 (β 期) 是结束期,当配体已经分散,随着配体被从患者清除,血清浓度下降。 t_{α} 半衰期是第一期的半衰期, t_{β} 半衰期是第二期的半衰期。因此,在一个实施方案中,本发明提供了发明的配体或包含配体的组合物,其 t_{α} 半衰期在 15 分钟或更长范围。在一个实施方案中,范围的下限是 30 分钟、45 分钟、1 小时、2 小时、3 小时、4 小时、5 小时、6 小时、7 小时、10 小时、11 小时或 12 小时。此外,或者替代地,发明的配体或组合物会有 t_{α} 半衰期在最多包括 12 小时的范围内。在一个实施方案中,范围的上限是 11、10、9、8、7、6 或 5 小时。合适范围的例子是 1 到 6 小时、2 到 5 小时或者 3 到 4 小时。

[0179] 在一个实施方案中,本发明提供了具有 t_{β} 半衰期在 2.5 小时或更长范围的发明所述配体 (多肽、dAb、激动剂或拮抗剂) 或包含配体的组合物。在一个实施方案中,范围的下限是 3 小时、4 小时、5 小时、6 小时、7 小时、10 小时、11 小时或 12 小时。此外,或者替代地,本发明的配体或组合物具有 t_{β} 半衰期在最多包括 21 天的范围内。在一个实施方案中,范围的上限是 12 小时、24 小时、2 天、3 天、5 天、10 天、15 天或者 20 天。在一个实施方案中,发明的配体或组合物会有在 12 到 60 小时范围内的 t_{β} 半衰期。另一个实施方案中,在 12 到 48 小时的范围内。再一个实施方案中,在 12 到 26 小时的范围内。

[0180] 此外,或者替代上述标准的,本发明提供了发明的配体或包含配体的组合物,具有 AUC 值 (曲线下面积) 在 $1\text{mg}\cdot\text{min}/\text{ml}$ 或更多的范围。在一个实施方案中,范围的下限是 5、10、15、20、30、100、200 或 $300\text{mg}\cdot\text{min}/\text{ml}$ 。此外,或者替代地,发明的配体或组合物具有最多 $600\text{mg}\cdot\text{min}/\text{ml}$ 范围的 AUC。在一个实施方案中,范围的上限是 500、400、300、200、150、100、75 或 $50\text{mg}\cdot\text{min}/\text{ml}$ 。在一个实施方案中,发明的配体会有选自以下范围的 AUC: 15 到 $150\text{mg}\cdot\text{min}/\text{ml}$ 、15 到 $100\text{mg}\cdot\text{min}/\text{ml}$ 、15 到 $75\text{mg}\cdot\text{min}/\text{ml}$ 和 15 到 $50\text{mg}\cdot\text{min}/\text{ml}$ 。

[0181] 发明的多肽和 dAbs 以及包含它们的激动剂或拮抗剂可以配制成具有较大流体动力学体积,例如通过附着 PEG 基团、血清白蛋白、转铁蛋白、转铁蛋白受体或者至少其转铁蛋白结合部分、抗体 Fc 区,或者通过偶联上抗体结构域。例如,多肽、dAbs、激动剂和拮抗剂制成为抗体的较大抗原结合片段或者为抗体 (例如制成 Fab、Fab'、 $F(ab)_2$ 、 $F(ab')_2$ 、IgG、

scFv)。

[0182] 本发明配体(例如 dAb 单体和多体)的流体动力学大小可以利用本领域已知的方法来确定。例如,可以利用凝胶过滤层析来确定配体的流体动力学大小。确定配体流体动力学大小的合适凝胶过滤基质,比如交联琼脂糖基质是公知的,很容易得到。

[0183] 配体形式的大小(例如附着在 dAb 单体上的 PEG 部分的大小)可以根据预期应用而变化。例如,想要配体离开血液循环,进入外周组织的情况中,保持较低的配体流体动力学大小以便协助从血流中外渗是可取的。替代地,希望配体在系统循环中保留较长时间的情况中,可以通过例如形成 Ig 样蛋白来提高配体的大小。

[0184] 通过靶向使体内半衰期增加的抗原或表位来延长半衰期

[0185] 如文中所述,配体的流体动力学大小及其血清半衰期也可以通过将本发明的靶抗原结合多肽、dAb、激动剂或拮抗剂偶联或关联到结合结构域(例如抗体或抗体片段)上来提高,所述结合结构域所结合的抗原或表位能够提高体内半衰期。例如,靶抗原结合剂(例如多肽)可以偶联或者连接到抗血清白蛋白或抗新生儿 Fc 受体抗体或抗体片段上,例如抗 SA 或抗新生儿 Fc 受体 dAb、Fab、Fab' 或 scFv;或者抗抗 SA 亲和体(affibody)或抗新生儿 Fc 受体亲和体或抗 SA avimer;或者抗 SA 结合结构域,所述结构域包含的支架选自,但优选不限于 CTLA-4、脂笼蛋白(lipocallin)、SpA、亲和体、avimer、GroEl 和纤连蛋白(参见 2008 年 2 月 8 日提交的 PCT/GB2008/000453(WO2008096158;US2009259026)中关于这些结合结构域的公开,这些结构域及其序列通过引用并入本文,构成本说明书公开内容的一部分)。偶联是指包含发明的多肽、dAb、激动剂或拮抗剂的组合物与能够结合血清白蛋白的结合结构域(共价或非共价地)形成连接。

[0186] 提高体内血清半衰期的合适多肽包括,例如转铁蛋白受体的特异配体-神经药物(neuropharmaceutical agent)融合蛋白(参见美国专利 5,977,307,该专利的教导通过引用并入本文)、脑毛细血管内皮细胞受体、转铁蛋白、转铁蛋白受体(例如可溶性转铁蛋白受体)、胰岛素、胰岛素样生长因子 1(IGF 1)受体、胰岛素样生长因子 2(IGF 2)受体、胰岛素受体、凝血因子 X、 α 1-抗胰蛋白酶和 HNF 1 α 。提高血清半衰期的合适多肽还包括 α -1 糖蛋白(血清类粘蛋白;AAG)、 α -1 抗胰凝乳蛋白酶(ACT)、 α -1 微球蛋白(蛋白 HC;AIM)、抗凝血酶 III(AT III)、脱脂脂蛋白 A-1(Apo A-1)、脱脂脂蛋白 B(Apo B)、铜蓝蛋白(Cp)、补体成分 C3(C3)、补体成分 C4(C4)、C1 酯酶抑制因子(C1INH)、C-反应蛋白(CRP)、铁蛋白(FER)、血红素结合蛋白(hemopexin)(HPX)、脂蛋白(a)(Lp(a))、甘露糖结合蛋白(MBP)、肌红蛋白(Myo)、前白蛋白(甲状腺素运载蛋白;PAL)、视黄醇结合蛋白(RBP)和类风湿因子(RF)。

[0187] 来自胞外基质的合适蛋白包括,例如胶原蛋白、层粘连蛋白、整合蛋白和纤连蛋白。胶原蛋白是胞外基质的主要蛋白。目前已知在身体不同部位存在大约 15 种胶原蛋白分子,例如存在于骨、皮肤、肌腱、韧带、角膜、内脏的 I 型胶原蛋白(构成身体胶原蛋白的 90%),或者存在于软骨、椎间盘、脊索和眼睛的玻璃体液的 II 型胶原蛋白。

[0188] 来自血液的合适蛋白包括,例如血浆蛋白(例如纤维蛋白、 α -2 巨球蛋白、血清白蛋白、纤维蛋白原(例如纤维蛋白原 A、纤维蛋白原 B)、血清淀粉样蛋白 A、亲血球蛋白、肌球蛋白抑制蛋白、泛素、胚泡激肽和 β -2-微球蛋白)、酶和酶抑制因子(例如血纤维蛋白溶酶原、溶菌酶、胱抑素 C、 α -1-抗胰蛋白酶和胰腺胰蛋白酶抑制因子)、免疫系统的蛋白(比

如免疫球蛋白（例如 IgA、IgD、IgE、IgG、IgM、免疫球蛋白轻链（ κ/λ ））、转运蛋白（例如视黄醇结合蛋白、 α -1 微球蛋白）、防御素（例如 β -防御素 1、嗜中性粒细胞防御素 1、嗜中性粒细胞防御素 2 和嗜中性粒细胞防御素 3）等。

[0189] 在血脑屏障或神经组织中可以发现的合适蛋白包括，例如黑素皮质素受体、髓磷脂、抗坏血酸转运蛋白等。

[0190] 增强体内半衰期的合适多肽还包括定位到肾中的蛋白（例如多囊蛋白、IV 型胶原蛋白、有机阴离子转运蛋白 K1、Heymann' s 抗原）；定位到肝脏中的蛋白（例如乙醇脱氢酶、G250）；定位到肺中的蛋白（例如结合 IgA 的分泌成分）；定位到心脏中的蛋白（例如与扩张型心肌病相关联的 HSP-27）；定位到皮肤的蛋白（例如角蛋白）；骨特异蛋白，比如形态发生蛋白（BMPs），它是展示成骨活性的转化生长因子 β 超家族蛋白的一个亚群（例如 BMP-2、BMP-4、BMP-5、BMP-6、BMP-7、BMP-8）；肿瘤特异蛋白（例如滋养层抗原、赫赛汀（herceptin）受体、雌激素受体、组织蛋白酶（例如可以在肝和肾中发现的组织蛋白酶 B））。

[0191] 合适的疾病特异蛋白包括，例如只在活化的 T 细胞上表达的抗原，包括 LAG-3（淋巴细胞活化基因）、护骨蛋白配体（OPGL；参见 Nature 402, 304-309 (1999)）、OX40（TNF 受体家族的成员，其表达在活化的 T 细胞上，在人 T 细胞白血病 I 型病毒产生细胞中被特异上调；参见 Immunol. 165 (1) :263-70 (2000)）。合适的疾病特异蛋白还包括，例如包括 CG6512 果蝇、人截瘫蛋白（paraplegin）、人 FtsH、人 AFG3L2、小鼠 ftsH 的金属蛋白酶（与关节炎 / 癌症相关联）；和血管生成生长因子，包括酸性成纤维细胞生长因子（FGF-1）、碱性成纤维细胞生长因子（FGF-2）、血管内皮生长因子 / 血管通透性因子（VEGF/VPF）、转化生长因子- α （TGF α ）、肿瘤坏死因子- α （TNF- α ）、血管生成素、白介素-3（IL-3）、白介素-8（IL-8）、血小板源性内皮生长因子（PD-ECGF）、胎盘生长因子（PlGF）、中期因子血小板源生长因子-BB（PDGF）和 fractalkine。

[0192] 提高体内血清半衰期的合适多肽还包括应激蛋白，比如热休克蛋白（HSPs）。HSPs 通常可见于细胞内。当在细胞外发现它们时，表明细胞已死亡，泄漏了内容物。当这种非程序化细胞死亡（坏死）是由于外伤、疾病或损伤而发生的时，细胞外的 HSPs 引发免疫系统的反应。与胞外 HSP 结合可以导致本发明的组合物定位到疾病部位。

[0193] 参与 Fc 转运的合适蛋白包括，例如 Brambell 受体（又称为 FcRB）。该 Fc 受体有两个功能，每一个都可能对递送有帮助。这两种功能是（1）将 IgG 跨过胎盘从母亲转运给孩子，（2）保护 IgG 免于降解，从而延长其血清半衰期。该受体被认为能够从内体循环利用 IgG（参见 Holliger et al, NatBiotechnol 15 (7) :632-6 (1997)）。

[0194] 结合血清白蛋白的 dAbs (AlbudAbs™)

[0195] 发明的一个实施方案中提供了结合靶抗原和第二 dAb 的多肽、激动剂或拮抗剂（例如包含抗靶抗原 dAb（第一 dAb）的双特异性配体），其中第二 dAb 结合血清白蛋白（SA）的 K_D 通过表面等离子共振确定是 1nM 到 1、2、3、4、5、10、20、30、40、50、60、70、100、200、300、400 或 500 μ M（即， $\times 10^9$ 到 5×10^4 ），或者 100nM 到 10 μ M，或者 1 到 5 μ M，或者 3 到 70nM，或者 10nM 到 1、2、3、4 或 5 μ M。例如通过表面等离子共振确定为 30 到 70nM。在一个实施方案中，第一 dAb（或 dAb 单体）结合 SA（例如 HSA）的 K_D 通过表面等离子共振确定为大约 1、50、70、100、150、200、300nM 或者 1、2 或 3 μ M。在一个实施方案中，对于包含第一抗 SA 的 dAb 和抗靶抗原的第二 dAb 的双特异性配体，第二 dAb 与其靶分子的亲和力（例如使用 BiaCore，

通过表面等离子共振测量到的 K_D 和 / 或 K_{off}) 是第一 dAb 与 SA 的亲力的 1 到 100000 倍 (例如, 100 到 100000, 或者 1000 到 100000, 或者 10000 到 100000 倍)。在一个实施方案中, 血清白蛋白是人血清白蛋白 (HSA)。例如, 第一 dAb 以大约 $10 \mu\text{M}$ 的亲力结合 SA, 而第二 dAb 以 100pM 的亲力结合它的靶分子。在一个实施方案中, 血清白蛋白是人血清白蛋白 (HSA)。在一个实施方案中, 第一 dAb 以大约 50, 例如 70、100、150 或 200nM 的 K_D 结合 SA (例如 HSA)。W003002609、W004003019 和 W004058821 中有关于双特异性配体的细节。

[0196] 在一个实施方案中, 发明的配体可以包含这样的 dAb, 所述 dAb 结合血清白蛋白 (SA) 的 K_D 通过表面等离子共振确定为 1nM 到 1、2、3、4、5、10、20、30、40、50、60、70、100、200、300、400 或 $500 \mu\text{M}$ (即 $\times 10^9$ 到 5×10^4), 或者 100nM 到 $10 \mu\text{M}$, 或者 1 到 $5 \mu\text{M}$, 或者 3 到 70nM , 或者 10nM 到 1、2、3、4 或 $5 \mu\text{M}$ 。例如通过表面等离子共振确定为 30 到 70nM 。在一个实施方案中, 第一 dAb (或 dAb 单体) 结合 SA (例如 HSA) 的 K_D 通过表面等离子共振确定为大约 1、50、70、100、150、200、300 nM 或者 1、2 或 $3 \mu\text{M}$ 。在一个实施方案中, 第一和第二 dAbs 通过连接分子连接, 例如连接分子含有 1 到 4 个氨基酸、或者 1 到 3 个氨基酸, 或者大于 3 个氨基酸, 或者多于 4、5、6、7、8、9、10、15 或 20 个氨基酸。在一个实施方案中, 使用较长的连接分子 (多于 3 个氨基酸) 来增强效力 (激动剂或拮抗剂中的一个或两个 dAb 的 K_D)。在一个实施方案中, 连接分子是螺旋连接分子。

[0197] 在配体、激动剂和拮抗剂的具体实施方案中, dAb 结合人血清白蛋白, 并与选自下述的 dAb 竞争结合白蛋白:

[0198] MSA-16、MSA-26 (参见 W004003019 (US2006106203) 中关于这些序列的公开, 这些序列及其核酸序列通过引用并入本文, 构成本说明书公开内容的一部分)、

[0199] DOM7m-16 (SEQ ID NO :473)、DOM7m-12 (SEQ ID NO :474)、DOM7m-26 (SEQ ID NO :475)、DOM7r-1 (SEQ ID NO :476)、DOM7r-3 (SEQ ID NO :477)、DOM7r-4 (SEQ ID NO :478)、DOM7r-5 (SEQ ID NO :479)、DOM7r-7 (SEQ ID NO :480)、DOM7r-8 (SEQ ID NO :481)、DOM7h-2 (SEQ ID NO :482)、DOM7h-3 (SEQ ID NO :483)、DOM7h-4 (SEQ ID NO :484)、DOM7h-6 (SEQ ID NO :485)、DOM7h-1 (SEQ ID NO :486)、DOM7h-7 (SEQ ID NO :487)、DOM7h-22 (SEQ ID NO :489)、DOM7h-23 (SEQ ID NO :490)、DOM7h-24 (SEQ ID NO :491)、DOM7h-25 (SEQ ID NO :492)、DOM7h-26 (SEQ ID NO :493)、DOM7h-21 (SEQ ID NO :494)、DOM7h-27 (SEQ ID NO :495)、DOM7h-8 (SEQ ID NO :496)、DOM7r-13 (SEQ ID NO :497)、DOM7r-14 (SEQ ID NO :498)、DOM7r-15 (SEQ ID NO :499)、DOM7r-16 (SEQ ID NO :500)、DOM7r-17 (SEQ ID NO :501)、DOM7r-18 (SEQ ID NO :502)、DOM7r-19 (SEQ ID NO :503)、DOM7r-20 (SEQ ID NO :504)、DOM7r-21 (SEQ ID NO :505)、DOM7r-22 (SEQ ID NO :506)、DOM7r-23 (SEQ ID NO :507)、DOM7r-24 (SEQ ID NO :508)、DOM7r-25 (SEQ ID NO :509)、DOM7r-26 (SEQ ID NO :510)、DOM7r-27 (SEQ ID NO :511)、DOM7r-28 (SEQ ID NO :512)、DOM7r-29 (SEQ ID NO :513)、DOM7r-30 (SEQ ID NO :514)、DOM7r-31 (SEQ ID NO :515)、DOM7r-32 (SEQ ID NO :516)、DOM7r-33 (SEQ ID NO :517) (参见 W02007080392 (US20070003549) 中关于这些序列的公开, 这些序列及其对应的核酸序列通过引用并入本文, 构成本说明书公开内容的一部分; 本段落中的 SEQ ID 编号是 W02007080392 中出现的那些)、

[0200] dAb8 (dAb10)、dAb10、dAb36、dAb7r20 (DOM7r20)、dAb7r21 (DOM7r21)、

dAb7r22 (DOM7r22)、dAb7r23 (DOM7r23)、dAb7r24 (DOM7r24)、dAb7r25 (DOM7r25)、dAb7r26 (DOM7r26)、dAb7r27 (DOM7r27)、dAb7r28 (DOM7r28)、dAb7r29 (DOM7r29)、dAb7r29 (DOM7r29)、dAb7r31 (DOM7r31)、dAb7r32 (DOM7r32)、dAb7r33 (DOM7r33)、dAb7r33 (DOM7r33)、dAb7h22 (DOM7h22)、dAb7h23 (DOM7h23)、dAb7h24 (DOM7h24)、dAb7h25 (DOM7h25)、dAb7h26 (DOM7h26)、dAb7h27 (DOM7h27)、dAb7h30 (DOM7h30)、dAb7h31 (DOM7h31)、dAb2 (dAbs 4、7、41)、dAb4、dAb7、dAb11、dAb12 (dAb7m12)、dAb13 (dAb15)、dAb15、dAb16 (dAb21、dAb7m16)、dAb17、dAb18、dAb19、dAb21、dAb22、dAb23、dAb24、dAb25 (dAb26、dAb7m26)、dAb27、dAb30 (dAb35)、dAb31、dAb33、dAb34、dAb35、dAb38 (dAb54)、dAb41、dAb46 (dAbs 47、52 和 56)、dAb47、dAb52、dAb53、dAb54、dAb55、dAb56、dAb7m12、dAb7m16、dAb7m26、dAb7r1 (DOM 7r1)、dAb7r3 (DOM7r3)、dAb7r4 (DOM7r4)、dAb7r5 (DOM7r5)、dAb7r7 (DOM7r7)、dAb7r8 (DOM7r8)、dAb7r13 (DOM7r13)、dAb7r14 (DOM7r14)、dAb7r15 (DOM7r15)、dAb7r16 (DOM7r16)、dAb7r17 (DOM7r17)、dAb7r18 (DOM7r18)、dAb7r19 (DOM7r19)、dAb7h1 (DOM7h1)、dAb7h2 (DOM7h2)、dAb7h6 (DOM7h6)、dAb7h7 (DOM7h7)、dAb7h8 (DOM7h8)、dAb7h9 (DOM7h9)、dAb7h10 (DOM7h10)、dAb7h11 (DOM7h11)、dAb7h12 (DOM7h12)、dAb7h13 (DOM7h13)、dAb7h14 (DOM7h14)、dAb7p1 (DOM7p1) 和 dAb7p2 (DOM7p2) (参见 W02008096158 (US2009259026) 中关于这些序列的公开, 这些序列及其对应的核酸序列通过引用并入本文, 构成本说明书公开内容的一部分)。dAb 后面的括号中显示了替代名称, 例如 dAb8 的另一个名称是 dAb10, 即 dAb8 (dAb10)。

[0201] 一些实施方案中, dAb 结合人血清白蛋白, 并且其包含的氨基酸序列与选自下述的 dAb 有至少大约 80%, 或至少大约 85%, 或至少大约 90%, 或至少大约 95%, 或至少大约 96%, 或至少大约 97%, 或至少大约 98%, 或至少大约 99% 的氨基酸序列同一性:

[0202] MSA-16、MSA-26、

[0203] DOM7m-16 (SEQ ID NO :473)、DOM7m-12 (SEQ ID NO :474)、DOM7m-26 (SEQ ID NO :475)、DOM7r-1 (SEQ ID NO :476)、DOM7r-3 (SEQ ID NO :477)、DOM7r-4 (SEQ ID NO :478)、DOM7r-5 (SEQ ID NO :479)、DOM7r-7 (SEQ ID NO :480)、DOM7r-8 (SEQ ID NO :481)、DOM7h-2 (SEQ ID NO :482)、DOM7h-3 (SEQ ID NO :483)、DOM7h-4 (SEQ ID NO :484)、DOM7h-6 (SEQ ID NO :485)、DOM7h-1 (SEQ ID NO :486)、DOM7h-7 (SEQ ID NO :487)、DOM7h-22 (SEQ ID NO :489)、DOM7h-23 (SEQ ID NO :490)、DOM7h-24 (SEQ ID NO :491)、DOM7h-25 (SEQ ID NO :492)、DOM7h-26 (SEQ ID NO :493)、DOM7h-21 (SEQ ID NO :494)、DOM7h-27 (SEQ ID NO :495)、DOM7h-8 (SEQ ID NO :496)、DOM7r-13 (SEQ ID NO :497)、DOM7r-14 (SEQ ID NO :498)、DOM7r-15 (SEQ ID NO :499)、DOM7r-16 (SEQ ID NO :500)、DOM7r-17 (SEQ ID NO :501)、DOM7r-18 (SEQ ID NO :502)、DOM7r-19 (SEQ ID NO :503)、DOM7r-20 (SEQ ID NO :504)、DOM7r-21 (SEQ ID NO :505)、DOM7r-22 (SEQ ID NO :506)、DOM7r-23 (SEQ ID NO :507)、DOM7r-24 (SEQ ID NO :508)、DOM7r-25 (SEQ ID NO :509)、DOM7r-26 (SEQ ID NO :510)、DOM7r-27 (SEQ ID NO :511)、DOM7r-28 (SEQ ID NO :512)、DOM7r-29 (SEQ ID NO :513)、DOM7r-30 (SEQ ID NO :514)、DOM7r-31 (SEQ ID NO :515)、DOM7r-32 (SEQ ID NO :516)、DOM7r-33 (SEQ ID NO :517) (本段落中的 SEQ ID 编号是 W02007080392 (US20070003549) 中出现的)、

[0204] dAb8、dAb10、dAb36、dAb7r20、dAb7r21、dAb7r22、dAb7r23、dAb7r24、dAb7r25、

dAb7r26、dAb7r27、dAb7r28、dAb7r29、dAb7r30、dAb7r31、dAb7r32、dAb7r33、dAb7h21、dAb7h22、dAb7h23、Ab7h24、Ab7h25、Ab7h26、dAb7h27、dAb7h30、dAb7h3、dAb2、dAb4、dAb7、dAb11、dAb12、dAb13、dAb15、dAb16、dAb17、dAb18、dAb19、dAb21、dAb22、dAb23、dAb24、dAb25、dAb26、dAb27、dAb30、dAb31、dAb33、dAb34、dAb35、dAb38、dAb41、dAb46、dAb47、dAb52、dAb53、dAb54、dAb55、dAb56、dAb7m12、dAb7m16、dAb7m26、dAb7r1、dAb7r3、dAb7r4、dAb7r5、dAb7r7、dAb7r8、dAb7r13、dAb7r14、dAb7r15、dAb7r16、dAb7r17、dAb7r18、dAb7r19、dAb7h1、dAb7h2、dAb7h6、dAb7h7、dAb7h8、dAb7h9、dAb7h10、dAb7h11、dAb7h12、dAb7h13、dAb7h14、dAb7p1 和 dAb7p2。

[0205] 例如,结合人血清白蛋白的 dAb 可以包含与下述氨基酸序列有至少大约 90%,或至少大约 95%,或至少大约 96%,或至少大约 97%,或至少大约 98%,或至少大约 99%氨基酸序列同一性的氨基酸序列:DOM7h-2(SEQ ID NO:482)、DOM7h-3(SEQ ID NO:483)、DOM7h-4(SEQ ID NO:484)、DOM7h-6(SEQ ID NO:485)、DOM7h-1(SEQ ID NO:486)、DOM7h-7(SEQ ID NO:487)、DOM7h-8(SEQ ID NO:496)、DOM7r-13(SEQ ID NO:497)、DOM7r-14(SEQ ID NO:498)、DOM7h-22(SEQ ID NO:489)、DOM7h-23(SEQ ID NO:490)、DOM7h-24(SEQ ID NO:491)、DOM7h-25(SEQ ID NO:492)、DOM7h-26(SEQ ID NO:493)、DOM7h-21(SEQ ID NO:494)、DOM7h-27(SEQ ID NO:495)(本段落中的 SEQ ID 编号是 W02007080392(US20070003549)中出现的那些)、

[0206] dAb8、dAb10、dAb36、dAb7h21、dAb7h22、dAb7h23、Ab7h24、Ab7h25、Ab7h26、dAb7h27、dAb7h30、dAb7h31、dAb2、dAb4、dAb7、dAb11、dAb12、dAb13、dAb15、dAb16、dAb17、dAb18、dAb19、dAb21、dAb22、dAb23、dAb24、dAb25、dAb26、dAb27、dAb30、dAb31、dAb33、dAb34、dAb35、dAb38、dAb41、dAb46、dAb47、dAb52、dAb53、dAb54、dAb55、dAb56、dAb7h1、dAb7h2、dAb7h6、dAb7h7、dAb7h8、dAb7h9、dAb7h10、dAb7h11、dAb7h12、dAb7h13 和 dAb7h14。

[0207] 一些实施方案中,结合人血清白蛋白的 dAb 包含的氨基酸序列与选自下述的 dAb 的氨基酸序列有至少大约 80%,或至少大约 85%,或至少大约 90%,或至少大约 95%,或至少大约 96%,或至少大约 97%,或至少大约 98%,或至少大约 99%的氨基酸序列同一性:

[0208] DOM7h-2(SEQ ID NO:482)、DOM7h-6(SEQ ID NO:485)、DOM7h-1(SEQ ID NO:486)、DOM7h-7(SEQ ID NO:487)、DOM7h-8(SEQ ID NO:496)、DOM7h-22(SEQ ID NO:489)、DOM7h-23(SEQ ID NO:490)、DOM7h-24(SEQ ID NO:491)、DOM7h-25(SEQ ID NO:492)、DOM7h-26(SEQ ID NO:493)、DOM7h-21(SEQ ID NO:494)、DOM7h-27(SEQ ID NO:495)(本段落中的 SEQ ID 编号是 W02007080392(US20070003549)中出现的那些)、

[0209] dAb7h21、dAb7h22、dAb7h23、Ab7h24、Ab7h25、Ab7h26、dAb7h27、dAb7h30、dAb7h31、dAb2、dAb4、dAb7、dAb38、dAb41、dAb7h1、dAb7h2、dAb7h6、dAb7h7、dAb7h8、dAb7h9、dAb7h10、dAb7h11、dAb7h12、dAb7h13 和 dAb7h14。

[0210] 在更具体的实施方案中,dAb 是结合人血清白蛋白并且具有选自下述的氨基酸序列的 V_{κ} dAb:

[0211] DOM7h-2(SEQ ID NO:482)、DOM7h-6(SEQ ID NO:485)、DOM7h-1(SEQ ID NO:486)、DOM7h-7(SEQ ID NO:487)、DOM7h-8(SEQ ID NO:496)(本段中 SEQ ID 编号是 W02007080392(US20070003549)中出现的那些)、

[0212] dAb2、dAb4、dAb7、dAb38、dAb41、dAb54、dAb7h1、dAb7h2、dAb7h6、dAb7h7、dAb7h8、

dAb7h9、dAb7h10、dAb7h11、dAb7h12、dAb7h13 和 dAb7h14。

[0213] 在更具体的实施方案中，dAb 是结合人血清白蛋白并且具有选自 dAb7h30 和 dAb7h31 的氨基酸序列的 V_H dAb。

[0214] 在更具体的实施方案中，dAb 是 dAb7h11 或 dAb7h14。

[0215] 其他实施方案中，dAb、配体、激动剂或拮抗剂结合人血清白蛋白并且包含上述氨基酸序列中任何一个的一个、两个或三个 CDR，例如 dAb7h11 或 dAb7h14 的一个、两个或三个 CDR。

[0216] 结合血清白蛋白的合适 Camelid V_{HH} 包括 WO 2004/041862 (Ablynx N. V.) (US2009238829) 和 WO2007080392 (US20070003549) (这些 V_{HH} 序列及其对应的核酸序列通过引用并入本文，并构成本说明书公开内容的一部分) 中公开的那些，比如序列 A (SEQ ID NO :518)、序列 B (SEQ ID NO :519)、序列 C (SEQ ID NO :520)、序列 D (SEQ ID NO :521)、序列 E (SEQ ID NO :522)、序列 F (SEQ ID NO :523)、序列 G (SEQ ID NO :524)、序列 H (SEQ ID NO :525)、序列 I (SEQ ID NO :526)、序列 J (SEQ ID NO :527)、序列 K (SEQ ID NO :528)、序列 L (SEQ ID NO :529)、序列 M (SEQ ID NO :530)、序列 N (SEQ ID NO :531)、序列 O (SEQ ID NO :532)、序列 P (SEQ ID NO :533)、序列 Q (SEQ ID NO :534)，这些序列编号与 WO2007080392 或 WO 2004/041862 (Ablynx N. V.) 中引用的编号对应。一些实施方案中，Camelid V_{HH} 结合人血清白蛋白并且包含的氨基酸序列与 WO2007080392 中公开的 ALB1 或者 SEQ ID NOS :518-534 (这些序列编号与 WO2007080392 或 WO 2004/041862 中引用的对应) 中的任何一个具有至少大约 80%，或至少大约 85%，或至少大约 90%，或至少大约 95%，或至少大约 96%，或至少大约 97%，或至少大约 98%，或至少大约 99% 氨基酸序列同一性。

[0217] 一些实施方案中，配体、激动剂或拮抗剂包含抗血清白蛋白 dAb，其与本文公开的任何抗血清白蛋白 dAb 竞争结合血清白蛋白 (例如人血清白蛋白)。

[0218] 在替代实施方案中，激动剂、拮抗剂或配体包含靶抗原 (例如，人 TNFR1) 的特异结合部分，其中如同时待审的申请 PCT/GB2008/000453 (2008 年 2 月 8 日提交) 的描述，所述部分包含非免疫球蛋白序列，这些结合部分的公开、它们的制备和 (例如从各种文库的) 挑选方法以及它们的序列通过引用并入本文，并作为本说明书公开内容的一部分。

[0219] 与半衰期延长组分 (例如白蛋白) 的偶联

[0220] 在一个实施方案中，(一或多个) 半衰期延长组分 (例如，白蛋白、转铁蛋白以及它们的片段和类似物) 与本发明的靶抗原结合多肽、dAb、激动剂或拮抗剂偶联或关联。WO 2005077042 (其公开内容通过引用并入本文并构成本说明书公开内容的一部分) 中描述了适用于靶抗原结合形式的白蛋白、白蛋白片段或白蛋白变体的实例。尤其以下白蛋白、白蛋白片段或白蛋白变体可以用于本发明：

[0221] • SEQ ID NO :1 (如 WO 2005077042 所公开，该序列通过引用明确并入本文公开内容)；

[0222] • 白蛋白片段或变体，包含 WO 2005077042 中 SEQ ID NO :1 的氨基酸 1-387，或者由所述氨基酸 1-387 构成；

[0223] • 白蛋白，或者其片段或变体，包含选自下述的氨基酸序列：(a) WO2005077042 中 SEQ ID NO :1 的氨基酸 54 到 61；(b) WO 2005077042 中 SEQ ID NO :1 的氨基酸 76 到 89；(c) WO 2005077042 中 SEQ ID NO :1 的氨基酸 92 到；(d) WO 2005077042 中 SEQ ID NO :1 的氨基

酸 170 到 176 ;(e)WO2005077042 中 SEQ ID NO :1 的氨基酸 247 到 252 ;(f)WO 2005077042 中 SEQID NO :1 的氨基酸 266 到 277 ;(g)WO 2005077042 中 SEQ ID NO :1 的氨基酸 280 到 288 ;(h)WO 2005077042 中 SEQ ID NO :1 的氨基酸 362 到 368 ;(i)WO2005077042 中 SEQ ID NO :1 的氨基酸 439 到 447 ;(j)WO 2005077042 中 SEQID NO :1 的氨基酸 462 到 475 ;(k)WO 2005077042 中 SEQ ID NO :1 的氨基酸 478 到 486 ;以及 (l)WO 2005077042 中 SEQ ID NO :1 的氨基酸 560 到 566。

[0224] 用于靶抗原结合形式的合适白蛋白、片段和类似物的其他实例在 W003076567(US2008108560) 中有描述,该公开通过引用并入本文并构成本说明书公开内容的一部分。特别是以下白蛋白、片段或变体可以用于本发明:

[0225] • WO 03076567 中例如图 3 公开的人血清白蛋白(该序列信息通过引用明确并入本公开);

[0226] • 人血清白蛋白(HA),由 585 个氨基酸的单个非糖基化多肽链构成,其化学式分子量为 66,500(参见, Meloun, et al., FEBS Letters 58 :136(1975);Behrens, et al., Fed. Proc. 34 :591(1975);Lawn, et al., Nucleic Acids Research9 :6102-6114(1981);Minghetti, et al., J. Biol. Chem. 261 :6747(1986));

[0227] • 如 Weitkamp, et al., Ann. Hum. Genet. 37 :219(1973) 中描述的白蛋白多态变体或类似物或片段;

[0228] • 如 EP 322094 中描述的白蛋白片段或变体,例如 HA(1-373)、HA(1-388)、HA(1-389)、HA(1-369) 和 HA(1-419) 以及 1-369 和 1-419 之间的片段;

[0229] • 如 EP 399666 中描述的白蛋白片段或变体,例如 HA(1-177) 和 HA(1-200) 以及 HA(1-X) 之间的片段,其中 X 是 178 到 199 的任何数字。

[0230] 在利用(一或多个)半衰期延长组分(例如白蛋白、转铁蛋白以及它们的片段和类似物)来形成本发明的靶抗原结合多肽、dAbs、激动剂和拮抗剂的情况中,可以利用任何合适的方法对半衰期延长组分进行偶联,比如通过直接融合到靶抗原结合部分(例如,抗 TNFR1dAb),例如通过利用编码融合蛋白的单个核苷酸构建体,其中所述融合蛋白被编码为单个多肽链,半衰期延长组分位于靶抗原结合部分的 N- 或者 C- 末端。替代地,可以通过在部分之间使用连接肽来实现偶联,例如 WO 03076567(US2008108560) 或 W02004003019 中描述的连接肽(关于这些连接分子的公开通过引用并入本发明的公开从而提供用于本发明的实例)。在一个实施方案中,偶联可以通过螺旋连接分子,比如象本文描述的螺旋连接分子。可以理解,其他连接分子也可以用于该目的,包括比如富含甘氨酸-丝氨酸的连接分子。在一个实施方案中,连接分子可以是蛋白酶抗性连接分子。一般来说,增强体内血清半衰期的多肽是这样的多肽,它们天然存在于体内,并能抵抗那些将不需要的物质从生物体(例如人)移走的内源机制的降解或去除。例如,增强体内血清半衰期的多肽可以选自来自胞外基质的蛋白;血液中可以发现的蛋白;可见于血脑屏障或神经组织的蛋白;定位于肾、肝、肺、心脏、皮肤或骨的蛋白;应激蛋白;疾病特异性蛋白或者参与 Fc 转运的蛋白。

[0231] 本公开内容所描述的发明实施方案中,考虑了技术人员可以使用包含发明所述结合靶抗原的 dAb 的一或多个或者全部三个 CDRs(例如移植到合适的蛋白支架或骨架上(例如亲和体、SpA 支架、LDL 受体 A 类结构域或 EGF 结构域)的 CDRs)的多肽或结构域,而不是在发明的激动剂、拮抗剂或配体中使用抗靶抗原“dAb”。应当对公开内容整体进行相应

的分析以便提供对使用这种结构域代替 dAb 的激动剂或拮抗剂的公开。在这方面,参见 W02008096158,其公开内容通过引用并入本文。

[0232] 因此,在一个实施方案中,发明的激动剂或拮抗剂包含处于合适形式的免疫球蛋白单个可变域或者对靶抗原具有结合特异性的域抗体 (dAb),或者这种 dAb 的互补决定区。激动剂或拮抗剂可以是由这样的 dAb,或者基本由这样的 dAb 构成的多肽。激动剂或拮抗剂可以是包含合适形式中的 dAb (或者 dAb 的 CDRs) 的多肽,所述形式是比如抗体形式 (例如 IgG 样形式、scFv、Fab、Fab'、F(ab')₂) 或者双特异性配体,所述配体包含结合靶抗原的 dAb 和结合另一个靶蛋白、抗原或表位 (例如血清白蛋白) 的第二 dAb。

[0233] 本发明的多肽、dAbs、激动剂和拮抗剂可以制成许多种本领域已知的合适的抗体形式,比如 IgG 样形式、嵌合抗体、人源化抗体、人抗体、单链抗体、双特异性抗体、抗体重链、抗体轻链、抗体重链和 / 或轻链的同二聚体和异二聚体、任何以上分子的抗原结合片段 (例如 Fv 片段 (例如单链 Fv (scFv), 二硫键连接的 Fv)、Fab 片段、Fab' 片段、F(ab')₂ 片段)、单个可变域 (例如 V_H、V_L)、dAb, 以及任何以上片段的修饰形式 (例如通过共价附着聚烷撑二醇 (例如聚乙二醇、聚丙二醇、聚丁二醇) 或者其他合适的聚合物进行的修饰)。

[0234] 一些实施方案中,发明提供了 IgG 样形式的配体 (例如抗 TNFR1 拮抗剂)。这类形式具有 IgG 分子常规的四链结构 (两个重链和两个轻链), 其中一或多个可变区 (V_H 和 / 或 V_L) 已被本发明的 dAb 取代。在一个实施方案中,可变区中的每一个 (2 个 V_H 区和 2 个 V_L 区) 被 dAb 或单个可变域取代,其中至少一个是本发明的抗靶抗原 dAb。包含在 IgG 样形式中的 dAb(s) 或单个可变域具有相同或不同的特异性。一些实施方案中,IgG 样形式是四价的,可以具有一种 (只抗靶抗原)、两种 (例如,抗靶抗原和抗 SA)、三种或四种特异性。例如,IgG 样形式可以是单特异性的,包含 4 个具有相同特异性的 dAbs; 双特异性的,包含 3 个具有相同特异性的 dAbs 和另一个具有不同特异性的 dAb; 双特异性的,包含 2 个具有相同特异性的 dAbs 和两个具有不同的共有特异性的 dAbs; 三特异性的,包含具有相同特异性的第一和第二 dAbs, 具有不同特异性的第三 dAb, 和具有与第一、第二和第三 dAbs 不同的特异性的第四 dAb; 或者是四特异性的,包含四个具有各不相同的特异性的 dAbs。可以制备 IgG 样形式的抗原结合片段 (例如 Fab、F(ab')₂、Fab'、Fv、scFv)。在一个实施方案中,IgG 样形式或其抗原结合片段不会使靶抗原发生交联,例如该形式对于靶抗原而言可以是单价的。如果需要补体活化和 / 或抗体依赖性细胞毒性 (ADCC) 功能,配体可以是 IgG1 样形式。如果需要,IgG 样形式可以包含突变的恒定区 (变体 IgG 重链恒定区) 从而尽量减少与 Fc 受体的结合和 / 或固定补体的能力 (参见例如 Winter et al., GB 2, 209, 757B; Morrison et al., WO 89/07142; Morgan et al., WO 94/29351, 1994 年 12 月 22 日)。

[0235] 发明的配体 (多肽、dAbs、激动剂和拮抗剂) 可以制成融合蛋白,所述融合蛋白含有第一免疫球蛋白单个可变域与第二免疫球蛋白单个可变域直接融合。如果需要,这种形式还可以包含半衰期延长组分。例如,配体可以包含第一免疫球蛋白单个可变域直接融合到第二免疫球蛋白单个可变域上,后者直接融合到能够结合血清白蛋白的免疫球蛋白单个可变域上。

[0236] 通常,多肽结构域 (含有对靶分子有结合特异性的结合位点) 的方向以及配体是否包含连接分子只是涉及选择的问题的。但是,一些方向,有或者没有连接分子,可能比其他方向提供更好的结合特性。发明涵盖了所有方向 (例如 dAb1- 连接分子 -dAb2; dAb2- 连

接分子 -dAb1), 通过筛选可以容易地鉴定到含有能够提供所需结合特性的方向的配体。

[0237] 根据本发明的多肽和 dAbs, 包括 dAb 单体、二体和三体, 可以连接着抗体 Fc 区, 所述 Fc 区包含 C_H2 和 C_H3 结构域中的一个或两个, 以及任选的铰链区。例如, 可以利用编码与 Fc 区和配体连成单个核苷酸序列的载体来制备这种多肽。发明并且提供了以上提及的 dAb 单体的二体、三体和多体。

实施例

[0238] 实施例 1

[0239] 研究目的

[0240] 研究的目的是通过在文库上进行噬菌体挑选, 并结合用各种蛋白酶 (包括表达宿主中天然存在的那些蛋白酶) 对噬菌体进行处理来获得 GLP-1AlbudAbTM融合蛋白的蛋白酶抗性变体, 所述文库来源于包含 DPP IV 抗性 GLP-1 (文中称为 *GLP-1) 的 GLP 变体。正如本文描述的, AlbudAbTM是特异结合血清白蛋白的免疫球蛋白单个可变域。

[0241] GLP-1 受体

[0242] 胰高血糖素样肽 -1 受体 (GLP-1R) 属于七次跨膜 G 蛋白偶联受体的 B1 家族。受体及其天然激动剂配体 GLP-1 之间的相互作用由配体结合到受体 (ECD GLP-1R) 的胞外 N 末端而引发, 随后是配体与跨膜部分的核心之间的相互作用 (Al-Sabah et al, 2003; FEBS Lett ;553(3) :342-6)。已证明如果去掉跨膜核心, 虽然亲和力下降了, 但 GLP-1 与分离的 N 末端结构域仍能结合 (Lopez de Maturana et al, 2003; J. Biol. Chem ;278(12) : 10195-200)。由于在不使用助溶去污剂的情况下, 带有跨膜结构域的受体在溶液中的溶解性很差, 使用整个受体在溶液中进行噬菌体挑选是不可取的, 因此利用分离的胞外结构域进行噬菌体捕获来简化试验并富集展示对 ECD GLP-1R 有亲和力的分子的噬菌体。

[0243] ECD GLP-1R 带 His 标签的 Fc 单体的核苷酸和氨基酸序列如下:

[0244] 核苷酸序列 (SEQ ID NO :1) :

[0245] ATGGCCGGCG CCCCCGCCCC GCTGCGCCTT GCGCTGCTGC TGCTCGGGAT

[0246] GGTGGGCAGG GCCGGCCCCC GCCCCAGGG TGCCACTGTG TCCCTCTGGG

[0247] AGACGGTGCA GAAATGGCGA GAATACCGAC GCCAGTGCCA GCGCTCCCTG

[0248] ACTGAGGATC CACCTCCTGC CACAGACTTG TTCTGCAACC GGACCTTCGA

[0249] TGAATACGCC TGCTGGCCAG ATGGGGAGCC AGGCTCGTTC GTGAATGTCA

[0250] GCTGCCCTG GTACCTGCCC TGGGCCAGCA GTGTGCCGCA GGGCCACGTG

[0251] TACCGTTCT GCACAGCTGA AGGCTCTGG CTGCAGAAGG ACAACTCCAG

[0252] CCTGCCCTGG AGGGACTTGT CGGAGTGCGA GGAGTCCAAG CGAGGGGAGA

[0253] GAAGTCCCC GGAGGAGCAG CTCCTGTTCC TCAAGCTTGA GCCCAAATCG

[0254] GCCGACAAAA CTCACACATC ACCACCGTCA CCAGCACCTG AACTCCTGGG

[0255] GGGACCGTCA GTCTTCCTCT TCCCCCAAA ACCCAAGGAC ACCCTCATGA

[0256] TCTCCCGGAC CCCTGAGGTC ACATGCGTGG TGGTGGACGT GAGCCACGAA

[0257] GACCCTGAGG TCAAGTTCAA CTGGTACGTG GACGGCGTGG AGGTGCATAA

[0258] TGCCAAGACA AAGCCGCGGG AGGAGCAGTA CAACAGCACG TACCGGGTGG

[0259] TCAGCGTCCT CACCGTCCTG CACCAGGACT GGCTGAATGG CAAGGAGTAC

[0260] AAGTGCAAGG TCTCCAACAA AGCCCTCCCA GCCCCCATCG AGAAAACCAT
 [0261] CTCCAAAGCC AAAGGGCAGC CCCGAGAACC ACAGGTGTAC ACCCTGCCCC
 [0262] CATCCCGGGA TGAGCTGACC AAGAACCAGG TCAGCCTGAC CTGCCTGGTC
 [0263] AAAGGCTTCT ATCCCAGCGA CATCGCCGTG GAGTGGGAGA GCAATGGGCA
 [0264] GCCGGAGAAC AACTACAAGA CCACGCCTCC CGTGCTGGAC TCCGACGGCT
 [0265] CCTTCTTCTT CTACAGCAAG CTCACCGTGG ACAAGAGCAG GTGGCAGCAG
 [0266] GGGAACGTCT TCTCATGCTC CGTGATGCAT GAGGCTCTGC ACAACCACTA
 [0267] CACGCAGAAG AGCCTCTCCC TGTCTCCGGG TAAACATCAC CATCATCATC
 [0268] ACTGA

[0269] 氨基酸序列 (SEQ ID NO :2) :

[0270] MAGAPGLRL ALLLLGMVGR AGPRPQGATV SLWETVQKWR EYRRQCQRSL
 [0271] TEDPPPATDL FCNRTFDEYA CWPDPGEPGSF VNVSCPWYLP WASSVPQGHV
 [0272] YRFCTAEGLW LQKDNSSLPW RDLSECEESK RGERSSPEEQ LLFLKLEPKS
 [0273] ADKTHTSPPS PAPELLGGPS VFLFPPKPKD TLMISRTPEV TCVVVDVSHE
 [0274] DPEVKFNWYV DGVEVHNAKT KPREEQYNST YRVVSVLTVL HQDWLNGKEY
 [0275] KCKVSNKALP APIEKTISKA KGQPREPQVY TLPPSRDELT KNQVSLTCLV
 [0276] KGFYPSDIAV EWESNGQPEN NYKTTTPVLD SDGSFFLYSK LTVDKSRWQQ
 [0277] GNVFSCSVMH EALHNHYTQK SLSLSPGKHH HHHH

[0278] 还表达了带有 IgG Fc 标签的 GLP-1R ECD, 所述标签使得能够在蛋白 A 琼脂糖上对蛋白进行初步纯化。然后在噬菌体挑选过程中, 可以用包被蛋白 A 的珠子捕获可溶性受体。

[0279] 测试噬菌体的挑选

[0280] 进行测试来验证 GLP-1 受体的胞外结构域可以用于噬菌体展示文库的挑选。

[0281] 噬菌体载体

[0282] 使用了丝状噬菌体 (fd) 展示载体 -pDOM34 (它是 pDOM4 的衍生物), 这个载体基于带有 myc 标签的 fd 载体, 可以将蛋白序列克隆到其中的限制酶切位点之间来提供蛋白 - 基因 III 融合。(pDOM4, 如 WO 2007/085815 所述, 是 Fd 噬菌体载体的衍生物, 其中基因 III 信号肽序列被酵母糖脂锚定表面蛋白 (GAS) 信号肽 (WO 2005/093074) 所取代。它还含有位于前导序列和基因 III 之间的 c-myc 标签, 使得基因 III 回到读框)。

[0283] 由 pDOM4 产生 pDOM34 的改动包括:

[0284] 1.) 敲除 pDOM4 中位点 7476nt 的 NcoI 位点

[0285] 2.) 删除融合在 cpIII 的 N' 末端的 Myc 标签

[0286] 3.) 引入 NcoI 限制酶位点以协助信号肽后面的克隆。

[0287] 编码文库集合的基因以 NcoI/NotI 片段克隆。

[0288] 载体在大肠杆菌 MachI 细胞中增殖, 利用 Plasmid Mega Prep 试剂盒 (Qiagen) 分离, 采用标准技术 (Sambrook and Maniatis 1989) 经氯化铯梯度超速离心分离超螺旋部分。载体用 NcoI 和 NotI 酶切割, 然后用 PstI 切割以减少自连接比率。苯酚 / 氯仿抽提后, 将 DNA 用乙醇沉淀并在 Chromaspin TE-1000 柱子 (Clontech) 上与 NcoI 和 NotI 位点之间不需要的“填充物”DNA 片段纯化开。纯化后, 用载体 DNA 测试与多样化的 DAT-X DNA 片段的连接。

[0289] DAT-X 文库构建

[0290] 在包含 DPP IV 抗性 GLP-1 (以下称为 *GLP-1) 的 DAT-X 亲代分子的基础上构建了 18 个集合。

[0291] *GLP-1 (7-37) :

[0292] 氨基酸序列

[0293] HEGTFTSDVSSYLEGQAAKEFIAWLVKGRG (SEQ ID NO :3)

[0294] 核苷酸序列 :

[0295] CATGGTGAAGGGACCTTTACCAGTGATGTAAGTTCTTATTTGGAAGGCCAAGCTGCCAAGGAATTCATTGCTTGGCTGGTGAAGGCCGAGGA (SEQ ID NO :4)

[0296] DAT-X 亲代分子还包含与 DOM7h-14 (能够结合血清白蛋白的域抗体 (dAb) (albudab ;AlbudAbTM)) 的融合。

[0297] DOM7h-14 :

[0298] 氨基酸序列 :

[0299] DIQMTQSPSSLASVGDRTITCRASQWIGSQLSWYQQKPKAPKLLIMWRSSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCAQGAALPRTFGQGTKVEIKR (SEQ IDNO :5)

[0300] 核苷酸序列 :

[0301] GACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACCGTGTACCATCACTTGC CGGGCAAGTCAGTGGATTGGGTCTCAGTTATCTTGGTACCAGCAGAAACCAGGGAAAGCCCCTAAGCTCCTGATCAT GTGGCGTTCTCGTTGCAAAGTGGGGTCCCATCACGTTTCAGTGGCAGTGGATCTGGGACAGATTTCACTCTCACCA TCAGCAGTCTGCAACCTGAAGATTTTGCTACGTACTACTGTGCTCAGGGTGC GGCGTTGCCTAGGACGTTCGGCCAA GGGACCAAGGTGGAAATCAAACGG (SEQ ID NO :6)

[0302] DAT-X 亲代分子中 *GLP-1 和 DOM7h-14 通过螺旋连接分子接起来 :

[0303] 螺旋连接分子

[0304] 氨基酸序列 :

[0305] KEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAA (SEQ ID NO :7)

[0306] 核苷酸序列 :

[0307] AAAGAAGCGGCGGCGAAAGAAGCGGCGGCGAAAGAAGCGGCGGCGAAAGAATTGGCCGCAAAGAAGCG GCGGCGAAAGAAGCGGCGGCGAAAGAAGCGGCGGCGAAAGAATTGGCCGCA (SEQ ID NO :8)

[0308] 为了覆盖 *GLP-1 的全部序列 (除了已知对受体结合很重要的位点 (例如在 Sarrauste de Menthère et al. Eur J Med Chem. 2004 Jun ;39(6) :473-80 ;Neidigh et al. Biochemistry. 2001 Nov 6 ;40(44) :13188-200 ;Hjorth et al. J Biol Chem. 1994 Dec 2 ;269(48) :30121-4 和 Gallwitz et al. Regul Pept. 1996 May 7 ;63(1) :17-22 中描述的重要位点)), 采用组装 PCR 方案, 用 Phusion 高忠实度聚合酶 (NEB) 在 50 微升反应体积中构建了 17 个集合。通过一级 PCR 中的引物给每个文库引入四个随机化的核苷酸, 然后用生物素化的引物进行组装。用 Mutazyme II kit (Stratagene)、生物素化的引物和 5-50pg 模板 (50 μ l 反应), 经易错 PCR 在 *GLP-1 内引入随机突变。因为 *GLP-1 核苷酸序列较短, 进行两次易错 PCR 来提高突变率。

[0309] 用 NcoI 和 NotI 消化后, 利用包被链霉抗生物素蛋白的珠子从未被消化的产物中纯化插入片段。进行测试连接, 其中被消化的产物被连接到 pDOM34 中的相应位点。

[0310] 对该测试连接克隆进行测序,证实所有文库中都有预期的多样性,因此随后进行了全面的文库连接和转化。该连接是在 500 微升总体积中,使用 T4DNA 连接酶 (NEB),用 1 微克经消化的载体和插入片段以 1 : 2 比率进行。每个文库按照每 100 微升经电击的感受态大肠杆菌 TB1 细胞 10 微升的水平转化两次,回收后,100ml 培养物在 37°C 振荡培养 1 小时,将文库铺到含有 2XTY Tet 琼脂的大 (22cm) 的方形平板上。平板生长过夜,然后刮到 5ml 含有 15%甘油的 2xTY 中制备菌种。文库大小在 10^7 - 10^8 转化子范围内。

[0311] 对于噬菌体文库制备,培养开始将 100 微升甘油菌种接种到 200 毫升含有抗生素的 2xTY 培养基中,使得接种后培养物的即时终密度不超过 $OD_{600} = 0.1$ 。文库在 37°C 振荡过夜培养约 18 小时。培养物经离心沉淀,用 PEG 经二次沉淀制备噬菌体文库并重悬于 PBS 中。

[0312] 从没有经过挑选的文库中随机选择几个克隆进行测序来证实文库构建成功,并在噬菌体文库制备后进行第一轮淘选。

[0313] 除非另有说明,淘选、制备甘油储菌和扩增噬菌体的方法如下进行。

[0314] 使用 GLP-1 受体的胞外结构域进行淘选。将 100 微升噬菌体文库与含有 100nM GLP-1R 的 2% Marvell PBS 温育。温育于室温进行 1 小时,然后将噬菌体与预先封闭的 (2% Marvell PBS、1h、RT) 蛋白 A Dynabeads (Dyna1) 合并。室温下在转盘上温育 1 小时后,珠子用 0.1% Tween PBS 在 KingFisher 纯化系统 (Thermo Electron Corporation) 中洗 8 次 (通过利用磁性探针将珠子从洗液中转移到洗液,KingFisher 机器人将清洗过程自动化),通过在 500 微升 0.1M 甘氨酸 (pH 2.0) 中洗脱来回收特异噬菌体。用 100 微升 1M Tris-Cl (pH 8.0) 中和后,噬菌体用于在 37°C 感染对数期大肠杆菌 TG1 细胞 45 分钟。将被感染的细胞铺在琼脂 Tet 平板上,于 37°C 生长过夜。文库的效价、输入 / 输出以及文库大小如下表所示。

[0315]

文库	文库大小	1 st 挑选; 噬菌体 ϕ /ml	
		输入	输出
1	2.8×10^8	3.9×10^{10}	3.2×10^7
2	1.2×10^8	1.0×10^{11}	1.0×10^6
3	2.4×10^7	4.6×10^{10}	7.0×10^5
4	8.0×10^7	1.0×10^{11}	8.0×10^6
5	4.0×10^7	2.6×10^9	1.0×10^7
6	8.0×10^7	7.3×10^{10}	6.0×10^6
7	4×10^7	8.0×10^9	6.0×10^6
8	2.8×10^7	5.4×10^9	1.0×10^7
9	6.8×10^7	9.4×10^9	5.0×10^6
10	2.0×10^8	5.7×10^8	1.5×10^6

[0316]

11	8.8×10^8	4.5×10^9	2.0×10^6
12	6.6×10^8	4.0×10^9	1.6×10^6
13	1.8×10^8	6.7×10^{10}	2.0×10^5
14	4.8×10^7	6.0×10^9	3.0×10^5
15	6.0×10^7	4.2×10^{10}	1.0×10^6
16	2.4×10^8	1.6×10^{10}	4.8×10^5
17	4.2×10^8	1.3×10^{10}	1.4×10^6
18 易错	1.5×10^8	2.5×10^9	6.0×10^5
自连接	4.0×10^5		
DAT-X 对照	-	4.0×10^9	1.8×10^7

[0317] 所有文库在第一轮挑选都产生了合理的输出量。

[0318] 甘油储菌的制备是将琼脂平板上的菌落用 2ml 含有 15%甘油的 2XTY 培养液刮下来,并分装到冻存管中。

[0319] 对汇合在一起的噬菌体进行了以下挑选。将大肠杆菌甘油培养物合并在一起得到扩增的噬菌体,其中所述培养物含有全部 18 个文库第一轮挑选中的输出。培养物开始是取 50 微升每种淘选到的文库甘油储菌接种到 1 升含有抗生素的 2xTY 培养液中。将培养液分到两个 2L 的摇瓶中,每个半升,于 37°C 以 250rpm 振荡培养过夜。培养 18-20 小时后,经一次 PEG 沉淀并重悬在 PBS 中来制备噬菌体。

[0320] 用蛋白酶对 DAT-X 噬菌体展示文库进行挑选

[0321] 由第一轮挑选输出得到的扩增噬菌体用于以恒定浓度 (100nM) 的 GLP-1R 进一步挑选。此外,在用胰蛋白酶挑选前,将来自第一轮挑选输出的一批噬菌体做亚克隆,其中 AlbuAb 序列中带有 R108W 突变。这个突变使得 κ AlbuAb 克隆当展示在噬菌体上时,更抗胰蛋白酶处理。这是因为 dAb 羧基端的连接域抗体和 pIII 蛋白的精氨酸残基是胰蛋白酶敏感的。该位点的突变去掉了胰蛋白酶切割位点,使得蛋白酶挑选更加针对目标肽的所需区域。因此,含有或不含有 AlbuAb 中的 R108W 突变的两批噬菌体被用于第二轮挑选。

[0322] 噬菌体以不同浓度的胰蛋白酶或胰凝乳蛋白酶处理或者未做处理,之后与受体在室温下温育 1 小时。第二轮挑选效价(ϕ /ml)如下表所示。

[0323]

Φ	蛋白酶	输入	无蛋白酶	1 μ g/ml	10 μ g/ml
--------	-----	----	------	--------------	---------------

[0324]

DAT-X c+	α -胰凝乳蛋白酶	2×10^{10}	7×10^6	4×10^5	$< 1 \times 10^4$
R108W c+	胰蛋白酶	2×10^{10}	6×10^7	3×10^6	4×10^4
Lib 1-18	α -胰凝乳蛋白酶	2×10^{11}	3×10^8	2×10^6	3×10^4
R108W	胰蛋白酶	1×10^{12}	1×10^8	1×10^7	3×10^6

[0325] 通过取 50 微升甘油储菌接种到 50 毫升含有 Tet 的 2xTY 中,于 37°C 振荡过夜培养 20 小时将每组挑选的噬菌体输出 ($0 \mu\text{g/ml}$ - $10 \mu\text{g/ml}$) 扩增。纯化的噬菌体用于第三轮挑选,温育条件相同。

[0326] 第 3 轮挑选的效价(ϕ/ml)如下表所示。

[0327]

		α -胰凝乳蛋白酶		
第 2 轮后的 Φ 及浓度:	输入	没有蛋白酶	1 $\mu\text{g/ml}$	10 $\mu\text{g/ml}$
Lib 1-18_No 蛋白酶	1×10^{10}	7×10^8	5×10^4	$\sim 1 \times 10^2$
Lib 1-18_1 $\mu\text{g/ml}$	1×10^{10}	8×10^7	5×10^4	$\sim 1 \times 10^3$
Lib 1-18_10 $\mu\text{g/ml}$	1×10^{10}	2×10^8	1×10^6	1×10^4
对照 DAT-X	1×10^{10}	5×10^6	6×10^4	ND

[0328]

		胰蛋白酶		
第 2 轮后的 Φ 及浓度:	输入	没有蛋白酶	1 $\mu\text{g/ml}$	10 $\mu\text{g/ml}$
Lib 1-18 R108W_No 蛋白酶	2×10^{11}	2×10^9	1×10^7	1×10^6
Lib 1-18 R108W_1 $\mu\text{g/ml}$	1×10^{11}	5×10^8	1×10^7	2×10^5
Lib 1-18 R108W_10 $\mu\text{g/ml}$	1×10^{11}	7×10^7	1×10^7	1×10^6
对照 DAT-X R108W	3×10^{10}	1×10^7	1×10^7	3×10^4

[0329] 正如对代表性克隆的菌落测序证实的,克隆的多样性在第三轮挑选后已经下降,因此只有几个挑选输出的克隆被(如下所述)表达为可溶性蛋白。

[0330] DAT-X 噬菌体展示文库挑选输出

[0331] 数个 *GLP-1 变体被选中克隆为与 AlbuAb 形成的融合(带有如下所述的替代的连接分子),并被表达、纯化和在 GLP-1 受体分析法中检测。它们的氨基酸序列列为序列 1-10(见图 1)。一个 *GLP-1 序列变体(7)在用胰凝乳蛋白酶和胰蛋白酶处理进行的挑选输出中均大量存在,其融合被命名为 DMS7149;两个存在于使用胰凝乳蛋白酶时的输出中(DMS7150(8)和 51(9));噬菌体表达和分泌的过程中细胞内只有天然蛋白酶发挥作用而没有用胰蛋白酶或胰凝乳蛋白酶预处理的情况中,输出中观察到 1 个(DMS7148(6)),一个被克隆用于产生胰蛋白酶切割位点敲除(DMS7152(10))。

[0332] 含有氨基酸序列 1-4(见图 1)的蛋白经检验,相对 GLP-1 和对照 DAT-X 变体,显示较低的效力。Edman 测序提示,这些蛋白的加工有错误。

[0333] 克隆的进行是通过给 DAT-Y 克隆引入突变,所述 DAT-Y 克隆包含通过替代连接分子(氨基酸序列 PSS(SEQ ID NO:9),核苷酸序列 CCAAGCTCG(SEQ ID NO:10))与 DOM7h-14 连接的 *GLP-1。

[0334] 通过一级 PCR 中使用的引物给 *GLP-1 序列引入所选的突变,在组装 PCR 中,给融

合序列的 5' 和 3' 端分别引入 NcoI 和 BamHI 消化位点。

[0335] 组装 PCR 用 NcoI 和 BamHI 限制性核酸内切酶消化。

[0336] 制备表达载体 pDOM35 用于克隆。

[0337] 载体 pDOM35 是 pET12a 的衍生物,改动如下:

[0338] •OmpT 信号肽的最后三个残基从 SFA 变为 AWA,这样提高了大肠杆菌信号肽酶在正确位点的加工。

[0339] • 引入 NcoI 位点来协助信号肽后面的克隆。

[0340] • NcoI 和 BamHI 位点之间存在“填充物”

[0341] 用 NcoI 和 BamHI 消化 pDOM35,被切割的组装 PCR 用 Quick Ligation Kit (NEB) 连接到载体中。2 微升连接反应用于转化 MachI 细胞,回收后将细胞铺在含有羧苄青霉素的琼脂平板上,生长过夜。将菌落测序,含有正确序列的菌落用于质粒扩增和分离 (Plasmid Mini Prep kit, Qiagen)。用质粒 DNA 转化 BL21 (DE3) 细胞,得到的菌落用于接种表达培养物。

[0342] 表达的进行是通过接种 50 毫升 2xTY 培养基,该培养基中补充有 Overnight Express™自动诱导溶液 (1 微升 Solution 1 (Cat. No. 71298)、2.5 毫升 Solution 2 (Cat. No. 71299)、50 微升 Solution 3 (Cat. No. 71304), Novagen) 和 100 微克 / 毫升羧苄青霉素。培养在 37°C 过夜进行,然后通过 3700xg 离心 45 分钟将培养液上清澄清。然后用蛋白 L Streamline (GE Healthcare, Cat. No. 28-4058-03, 偶联蛋白 L 的) 从澄清的上清液中纯化所表达的蛋白,并用 0.1M 甘氨酸 pH 2.0 从蛋白 L 上洗脱,然后用 0.2 体积的 1M Tris pH 8.0 中和。

[0343] 来自 DMS7148 的 *GLP-1 变体部分也以通过替代的 PSS 连接分子相接的融合 DMS7161 (序列 11) 和有更高亲和力的 albudab 的融合 DOM7h-14-10 克隆到 pDOM35 中。

[0344] DOM7h-14-10

[0345] 氨基酸序列 (SEQ ID NO :22) :

[0346] DIQMTQSPSSLASVGDRTTITCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVPSRFSGSGSGT
DFTLTISSLQPEDFATYYCAQGLRHPKTFGQGTKVEIKR

[0347] 核苷酸序列 (SEQ ID NO :23) :

[0348] GACATCCAGATGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACCGTGTCCACCATCACTT
GCCGGCAAGTCAGTGGATTGGGTCTCAGTTATCTTGGTACCAGCAGAAACCAGGGAAAGCCCCTAAGCTCCTGA
TCATGTGGCGTTCCTCGTTGCAAAGTGGGGTCCCATCACGTTTCAGTGGCAGTGGATCTGGGACAGATTTCACTC
TCACCATCAGCAGTCTGCAACCTGAAGATTTTGCTACGTACTACTGTCTCAGGGTTTGAGGCATCCTAAGACGT
TCGGCCAAGGGACCAAGGTGGAAATCAAACGG

[0349] 包含 DMS7161 的序列的质粒 DNApDOM35 被转化到 BL21 (DE3) 中之后,由生长出来的菌落通过将菌落刮到含有葡萄糖的基础培养基中,并添加甘油至终浓度大约 15% 来制备甘油储菌。启动 DMS7161 的表达,即将甘油储菌接种到 50 毫升基础培养基 (DMS7161A) 中并补加酵母提取物至终浓度 10g/L (DMS7161B),以便得到起始培养密度 OD600 = 0.024。将培养物于 30°C 生长至 OD600 大约 1.4,然后通过加入 0.1mM 异丙基硫代-β-D-半乳糖苷进行诱导。培养在 23°C 再进行 24 小时,然后通过 3700xg 离心 45 分钟将培养物上清澄清。然后利用蛋白 L (GE Healthcare, Cat. No. 28-4058-03, 偶联蛋白 L 的) 从澄清的上清

液中纯化所表达的蛋白,并用 0.1M 甘氨酸 (pH2.0) 从蛋白 L 上洗脱,然后用 0.2 体积的 1M Tris(pH8.0) 中和。

[0350] DMS7148-52 的质量控制

[0351] 蛋白 DMS7148-52 被表达,并在非还原性 SDS-PAGE 上被观察, DMS7148 和 DMS7161 克隆在大肠杆菌中很好地表达,主要以预期的大小迁移(图 2、3 和 4)。质谱(图 5a)和 Edman 测序分析证实了序列的完整性。所有其他蛋白在 25-色氨酸的氨基和羧基位点被降解(产物分别为 24-142 和 26-142),而 DMS7148 含有 W25-D 突变,天冬氨酸不构成大肠杆菌细胞中的天然蛋白酶的切割位点(图 5a 到 f)。剩下的克隆中也观察到了降解产物 28-142。

[0352] 按以下方案在 GLP-1 受体结合检验法中测试了蛋白 DMS7148 的活性:

[0353] 背景:GLP-1R 是在 CHO 细胞上表达的 7TM G-蛋白偶联受体。该受体被 GLP-1 或这类类似物活化导致 ATP 被偶联在受体上的腺苷酸环化酶转化为 cAMP。CHO 细胞稳定转染 6CRE/luc 受体基因。在受体被 GLP-1 活化后产生 cAMP 时,启动子基因(含有 6 个拷贝的 cAMP 反应元件 -6CRE)驱动萤光素酶报告基因的表达。该酶然后催化萤光素的反应而产生可以在光度计上进行测量的光。

[0354] 方案:CHO 6CRE GLP1R 细胞(CHO K1 细胞被 6 个 cAMP 反应元件驱动的萤光素酶报告基因以及被人 GLP-1 受体稳定转染)以 2×10^5 细胞 /mL 接种悬浮培养液。悬浮培养物保持 48 小时。然后将细胞稀释到 15mM HEPES、2mM L 谷氨酰胺 (2.5×10^5 细胞 /ml) 中,并分配到含有 10u1/孔待检测化合物的 384-孔培养板中。加入检测对照后,将平板放回 37°C 保温箱在 5% CO₂ 保温 3 小时。温育后,按照试剂盒的描述向小孔中加入 Steady Glo 萤光素酶底物(Promega),将培养板用自动粘合的培养板盖(Weber Marking Systems Inc. Cat. No. 607780)密封。培养板放在读数器(Packard TopCount)中,预保温 5 分钟,然后读取荧光并对结果做图。检验中使用了不同浓度的化合物,对反应曲线进行拟合并计算 pC50。

[0355] 发现蛋白 DMS7148 有活性,但比 GLP-1 肽活性弱(图 6 和以下的总结)。

	分子	pEC50	% 最大反应
[0356]	DMS 7148	9.45	107
	GLP 7-36	11.92	97

[0357] 按照以下方案在 GLP-1 检验法中测试了 DMS7161 的活性:

[0358] 方法:

[0359] 通过将冻存管半浸到 37°C 水浴中,快速解冻 CHO 6CRE GLP1R 细胞,冻存管内容物转移到 50ml Falcon 管中,每个管中加入 10ml RPMI(不含酚红)检验培养液(Sigma, cat#R7509)+2mM L-谷氨酰胺(Gibco, cat#25030)+15mM HEPES(Sigma, cat#H0887)。计数并在 1200rpm 离心 5 分钟后,将细胞重悬于合适体积的 RPMI 检验培养液中达到 1×10^6 细胞 /ml,取 50 μ l 分配到白色 96 孔平底组织培养板(Costar 96 孔组织培养板,白色无菌, cat#3917)的每个小孔中。细胞在 37°C/5% CO₂ 培养过夜。第二天从保温箱中取出细胞,向孔中加入 50 μ l 先前制备的对照 / 样品,培养板放回保温箱于 37°C 和 5% CO₂ 保温 3 小时。

[0360] 制备 GLP-1(7-36)对照(Sigma, cat#G814)

[0361] 在 V 形底 96 孔板中,向 18 μ l RPMI 检验培养液中加入 2 μ l 1mg/mlGLP-1(7-36) 得到 30 μ M 溶液。向 298 μ l RPMI 检验培养液中加入 2 μ l 30 μ M 溶液得到 200nM 溶液(检验中终浓度 100nM)。将对照在培养板中以 1 : 10(15 μ l 对照 +135 μ l RPMI 检验培养液)系列稀释制作 8 个点的曲线。

[0362] 制备 Exendin-4 对照 (Sigma, cat#E7144)

[0363] 在 V 形底 96 孔板中,向 198 μ l RPMI 检验培养液中加入 2 μ l 1mg/ml 毒蜥外泌素得到 2.39 μ M 溶液。向 237 μ l RPMI 检验培养液中加入 2 μ l 2.39 μ M 溶液得到 20nM 溶液(检验中终浓度 10nM)。将对照在培养板中以 1 : 10(15 μ l 对照 +135 μ l RPMI 检验培养液)系列稀释制作 8 个点的曲线。

[0364] 制备未知样品

[0365] 采用制备对照的指导方针来同样地制备未知样品。制备最高浓度是检验所需终浓度的两倍,并在培养板上 1 : 10 稀释。

[0366] 制备萤光素酶 (Promega, cat#E2620)

[0367] 从冰箱中取出所需数量的 Bright-Glo 萤光素小份,允许在室温 (RT) 避光解冻。一个 5ml 小管足够一个检验培养板。

[0368] 温育时间后,向所有小孔中加入 50 μ l Bright-Glo 萤光素试剂,培养板在室温下保温 3 分钟以便细胞进行裂解。利用 M5e 微量板读数仪读取发光(每秒计数 (counts per second, CPS)),每个孔读 0.1 秒。从所有其他孔中减去只含有细胞的背景孔的 CPS。对照孔 (GLP-1(7-36) 或毒蜥外泌素 -4) 应当在最高浓度表现出最大的激发。利用 GraphPad Prism 或 ExcelFit 软件对未知样品的浓度影响曲线进行拟合,从中计算 EC50。

[0369] 如图 7 和以下总结所示, DMS7161 效力大, EC50 在 1.6 和 3.4nM 之间。

[0370]

	EC50 (M)	pEC50
ex-4	3.6E-09	8.45
DMS7161 A	3.4E-09	8.47
DMS7161 B	1.6E-09	8.79

[0371] 发现 DMS7161 活性和 DMS7148 一样。

[0372] 总结

[0373] 对于那些在大肠杆菌中表达期间对蛋白酶天然非常敏感并会降解的多样化肽融合而言,对它们进行噬菌体选择使我们能够鉴定抗天然细菌蛋白酶并能在大肠杆菌中表达的 *GLP-1 变体 - 融合物。这个克隆中被敲除的蛋白酶位点与胰蛋白酶和胰凝乳蛋白酶所识别的那些位点类似,但是该序列在用其它胰蛋白酶或胰凝乳蛋白酶处理进行的选择的结果中不存在。

序列	SEQ ID NO:
His标记的ECD GLP1R的Fc单体核苷酸序列	1
His标记的ECD GLP1R的Fc单体氨基酸序列	2
[0374] *GLP1 (737) 氨基酸序列	3
*GLP1 (737) 核苷酸序列	4
DOM7h14 氨基酸序列	5
DOM7h14核苷酸序列:	6
螺旋连接蛋白的氨基酸序列:	7
螺旋连接蛋白的核苷酸序列:	8
PSS	9
PSS核苷酸序列:	10
DMS7190	11
DMS7191	12
DMS7192	13
DMS7193	14
DMS7194	15
[0375] DMS7148	16
DMS7149	17
DMS7150	18
DMS7151	19
DMS7152	20
DMS7161	21
DOM7h1410氨基酸序列	22
DOM7h1410核苷酸序列	23

[0001]

序列表

<110> 葛兰素集团有限公司 (Glaxo Group Limited)

Enever, Carolyn

Jespers, Laurent

Pupecka, Malgorzata

Tomlinson, Ian

<120> 选出蛋白酶抗性多肽的方法

<130> DB00064W0

<150> 61/120, 135

<151> 2008-12-05

<160> 23

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1155

<212> DNA

<213> 人工序列

<400> 1

```

atggccggcg cccccggccc gctgcgcctt gcgctgctgc tgctcgggat ggtgggcagg 60
gccggcccc gcccccaggg igccacigtg tccctctggg agacggigca gaaatggcga 120
gaataccgac gccagtcca gcgctccctg actgaggatc cacctcctgc cacagacttg 180
ttctgcaacc ggaccttcca tgaatacgcc tgctggccag atggggagcc aggctcgttc 240
gtgaatgtca gctgcccctg gtacctgccc tgggccagca gtgtgccgca gggccacgtg 300
taccggttct gcacagctga aggcctctgg ctgcagaagg acaactccag cctgcccctg 360
agggacttgt cggagtgcga ggagtccaag cgaggggaga gaagctcccc ggaggagcag 420
ctcctgttcc tcaagcttga gccc aaatcg gccgacaaaa ctcacacatc accaccgtca 480
ccagcacctg aactcctggg gggaccgtca gtcttctctt tcccccaaa acccaaggac 540
accctcatga tctcccggac ccctgaggtc acatgcgttg tggiggacgt gagccacgaa 600
gacctgagg tcaagttcaa ctggtacgtg gacggcgttg aggtgcataa tgccaagaca 660
aagccgcggg aggagcagta caacagcacg taccgggtgg tcagcgtcct caccgtcctg 720
caccaggact ggctgaatgg caaggagtac aagtgcaagg tctccaaca agccctccca 780
gccccatcg agaaaacat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac 840
accctgcccc catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc 900
aaagcttct atcccagca catcgccgtg gactgggaga gcaatgggca gccggagaac 960
aactacaaga ccacgcctcc cgtgctggac tccgacggt ccttcttctc ctacagcaag 1020
ctcaccgtgg acaagagcag gtggcagcag gggaaactct tctcatgctc cgtgatgcat 1080

```

[0002]

gaggctctgc acaaccacta cagcagaag agcctctccc tgcctccggg taaacatcac 1140
 catcatcatc actga 1155

<210> 2

<211> 384

<212> PRT

<213> 人工序列

<220>

<400> 2

Met Ala Gly Ala Pro Gly Pro Leu Arg Leu Ala Leu Leu Leu Gly
 1 5 10 15
 Met Val Gly Arg Ala Gly Pro Arg Pro Gln Gly Ala Thr Val Ser Leu
 20 25 30
 Trp Glu Thr Val Gln Lys Trp Arg Glu Tyr Arg Arg Gln Cys Gln Arg
 35 40 45
 Ser Leu Thr Glu Asp Pro Pro Ala Thr Asp Leu Phe Cys Asn Arg
 50 55 60
 Thr Phe Asp Glu Tyr Ala Cys Trp Pro Asp Gly Glu Pro Gly Ser Phe
 65 70 75 80
 Val Asn Val Ser Cys Pro Trp Tyr Leu Pro Trp Ala Ser Ser Val Pro
 85 90 95
 Gln Gly His Val Tyr Arg Phe Cys Thr Ala Glu Gly Leu Trp Leu Gln
 100 105 110
 Lys Asp Asn Ser Ser Leu Pro Trp Arg Asp Leu Ser Glu Cys Glu Glu
 115 120 125
 Ser Lys Arg Gly Glu Arg Ser Ser Pro Glu Glu Gln Leu Leu Phe Leu
 130 135 140
 Lys Leu Glu Pro Lys Ser Ala Asp Lys Thr His Thr Ser Pro Pro Ser
 145 150 155 160
 Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro
 165 170 175
 Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys
 180 185 190
 Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp
 195 200 205
 Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu
 210 215 220
 Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu
 225 230 235 240
 His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn
 245 250 255
 Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly

[0003]

	260		265		270										
Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Arg	Asp	Glu
	275						280					285			
Leu	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys	Leu	Val	Lys	Gly	Phe	Tyr
	290					295					300				
Pro	Ser	Asp	Ile	Ala	Val	Glu	Trp	Glu	Ser	Asn	Gly	Gln	Pro	Glu	Asn
305					310					315					320
Asn	Tyr	Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe
			325					330						335	
Leu	Tyr	Ser	Lys	Leu	Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	Gln	Gly	Asn
			340					345						350	
Val	Phe	Ser	Cys	Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	His	Tyr	Thr
	355						360					365			
Gln	Lys	Ser	Leu	Ser	Leu	Ser	Pro	Gly	Lys	His	His	His	His	His	His
	370					375						380			

<210> 3
 <211> 31
 <212> PRT
 <213> 人工序列

<220>

<400> 3
 His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1 5 10 15
 Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
 20 25 30

<210> 4
 <211> 93
 <212> DNA
 <213> 人工序列

<220>

<400> 4
 catggggaag ggacctttac cagtgatgta agttcttatt tggaaggcca agctgccaag 60
 gaattcattg cttggctgtg gaaaggccga gga 93

<210> 5
 <211> 108

[0004]

<212> PRT

<213> 人工序列

<220>

<400> 5

```

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 Trp Ile Gly Ser Gln
           20           25           30
Leu Ser Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
           35           40           45
Met Trp Arg 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 Ala Gln Gly Ala Ala Leu Pro Arg
           85           90           95
Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
           100          105

```

<210> 6

<211> 324

<212> DNA

<213> 人工序列

<220>

<400> 6

```

gacatccaga tgaccagtc tccatcctcc ctgtctgcat ctgtaggaga ccgtgtcacc 60
atcaacttgc gggcaagtca gtggattggg tctcagttat cttggtacca gcagaaacca 120
gggaaagccc ctaagctcct gatcatgtgg cgttctctgt tgcaaagtgg ggtcccatca 180
cgtttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct 240
gaagattttg ctacgtacta ctgtgctcag ggtgcggcgt tgcctaggac gttcgcccaa 300
gggaccaagg tggaaatcaa acgg                                     324

```

<210> 7

<211> 40

<212> PRT

<213> 人工序列

<220>

[0005]

<211> 149

<212> PRT

<213> 人工序列

<220>

<400> 11

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

<210> 12

<211> 149

<212> PRT

<213> 人工序列

<220>

<400> 12

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

[0007]

```

      50              55              60
Arg Ala Ser Gln Trp Ile Gly Ser Gln Leu Ser Trp Tyr Gln Gln Lys
65              70              75              80
Pro Gly Lys Ala Pro Lys Leu Leu Ile Met Trp Arg Ser Ser Leu Gln
      85              90              95
Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe
      100             105             110
Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr
      115             120             125
Cys Ala Gln Gly Ala Ala Leu Pro Arg Thr Phe Gly Gln Gly Thr Lys
      130             135             140
Val Glu Ile Lys Arg
145

```

<210> 13

<211> 159

<212> PRT

<213> 人工序列

<220>

<400> 13

```

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

```

[0008]

<210> 17
 <211> 142
 <212> PRT
 <213> 人工序列

<220>

<400> 17
 His Gly Glu Gly Thr Phe Thr Ser Glu Phe Val Thr Tyr Leu Glu Gly
 1 5 10 15
 Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Pro
 20 25 30
 Ser Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser
 35 40 45
 Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Trp Ile Gly
 50 55 60
 Ser Gln Leu Ser Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu
 65 70 75 80
 Leu Ile Met Trp Arg Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe
 85 90 95
 Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu
 100 105 110
 Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Ala Gln Gly Ala Ala Leu
 115 120 125
 Pro Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
 130 135 140

<210> 18
 <211> 142
 <212> PRT
 <213> 人工序列

<400> 18
 His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1 5 10 15
 Met Thr Ser Arg Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly Pro
 20 25 30
 Ser Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser
 35 40 45
 Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Trp Ile Gly
 50 55 60
 Ser Gln Leu Ser Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu

[0011]


```

65             70             75             80
Leu Ile Met Trp Arg Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe
             85             90             95
Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu
             100             105             110
Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Ala Gln Gly Ala Ala Leu
             115             120             125
Pro Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
             130             135             140

```

<210> 19

<211> 142

<212> PRT

<213> 人工序列

<400> 19

```

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

```

<210> 20

<211> 142

<212> PRT

<213> 人工序列

<400> 20

```

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly

```

[0012]

```

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

```

<210> 21

<211> 142

<212> PRT

<213> 人工序列

<400> 21

```

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

```

[0013]

<210> 22

<211> 108

<212> PRT

<213> 人工序列

<400> 22

```

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 Trp Ile Gly Ser Gln
           20           25           30
Leu Ser Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
           35           40           45
Met Trp Arg 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 Ala Gln Gly Leu Arg His Pro Lys
           85           90           95
Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
           100           105

```

<210> 23

<211> 324

<212> DNA

<213> 人工序列

<400> 23

```

gacatccaga tgaccagtc tccatcctcc ctgtctgcat ctgtaggaga ccgtgtcacc 60
atcacttgcc gggcaagtca gtggattggg tctcagttat cttggtacca gcagaaacca 120
gggaaagccc ctaagtcct gatcatgtgg cgttcctcgt tgcaaagtgg ggtccatca 180
cgtttcagtg gcagtggatc tgggacagat ttcactctca ccatcagcag tctgcaacct 240
gaagattttg ctacgtacta ctgtgctcag ggtttgaggc atcctaagac gttcgccaa 300
gggaccaagg tggaaatcaa acgg                                     324

```

序列 1 (SEQ ID NO:11);

DMS7190

HGEGTFTSDVSSYSEEAAAKEFIAWLVKGRGKEAAAKELAADIQMTQSPS
SLSASVGDRVTITCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVP
SRFSGSGSGTDFTLTISSLQPEDFATYYCAQGAALPRTFGQGTKVEIKW

序列 2 (SEQ ID NO:12);

DMS7191

HGEGTFTSDGADLLEGQAAKEFIAWLVKGRGKEAAAKELAADIQMTQSPS
SLSASVGDRVTITCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVP
SRFSGSGSGTDFTLTISSLQPEDFATYYCAQGAALPRTFGQGTKVEIKR

序列 3 (SEQ ID NO:13);

DMS7192

HGEGTFTSDVATAACEGQAAKEFIACLVKGRGKEAAAKEAAAKEAAAKELA
ADIQMTQSPSSLSASVGDRVTITCRASQWIGSQLSWYQQKPGKAPKLLIM
WRSSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCAQGAALPRTFG
QGTKVEIKR

序列 4 (SEQ ID NO:14);

DMS7193

HGEGTFTSDVSSYLEGQAAKEFIAWLVTGLEREAAAKEAAAKFLAADIQM
TQSPSSLSASVGDRVTITCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSL
QSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCAQGAALPRTFGQGTKV
EIKR

图 1

序列 5 (SEQ ID NO:15);

DMS7194

HGEGTFTSEFVTYLEGQAAKEFIAWLVKGKEAAAKELAADIQMTQSPSSL
SASVGDRVTTTCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVPSR
FSGSGSGTDFLTISLQPEDFATYYCAQGAALPRTFGQGGTKVEIKW

序列 6 (SEQ ID NO:16);

DMS7148

HGEGTFTSDVSSYLEGQAAKEFIADLVEGRGPSSDIQMTQSPSSLSASVG
DRVTTTCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVPSRFSGSG
SGTDFLTISLQPEDFATYYCAQGAALPRTFGQGGTKVEIKR I

序列 7 (SEQ ID NO:17);

DMS7149

HGEGTFTSEFVTYLEGQAAKEFIAWLVKGRGPSSDIQMTQSPSSLSASVG
DRVTTTCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVPSRFSGSG
SGTDFLTISLQPEDFATYYCAQGAALPRTFGQGGTKVEIKR I

序列 8 (SEQ ID NO:18);

DMS7150

HGEGTFTSDVSSYLEGMTSREFIAWLVKGRGPSSDIQMTQSPSSLSASVG
DRVTTTCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVPSRFSGSG
SGTDFLTISLQPEDFATYYCAQGAALPRTFGQGGTKVEIKR I

序列 9 (SEQ ID NO:19);

DMS7151

HGEGTFTSDVSSYLEGQAAKEFIAWLVTGLEPSSDIQMTQSPSSLSASVG
DRVTTTCRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVPSRFSGSG
SGTDFLTISLQPEDFATYYCAQGAALPRTFGQGGTKVEIKR I

图 1(续)

序列 10 (SEQ ID NO:20);

DMS7152

HGEGTFTSDVSSYLEGQAASEFIAWLVDGGPSSDIQMTQSPSSLSASVG
DRVITICRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVPSRFGSGG
SGTDFLTITISSLPEDFATYYCAQGAALPRTFGQGGTKVEIKR I

序列 11 (SEQ ID NO:21);

DMS7161

HGEGTFTSDVSSYLEGQAAKEFIADLVEGRGPSSDIQMTQSPSSLSASVGD
DRVITICRASQWIGSQLSWYQQKPGKAPKLLIMWRSSLQSGVPSRFGSGSG
SGTDFLTITISSLPEDFATYYCAQGLRHPKTFGQGGTKVEIKR

图 1(续)

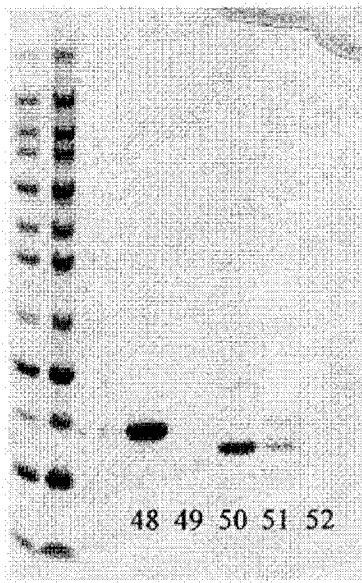


图 2

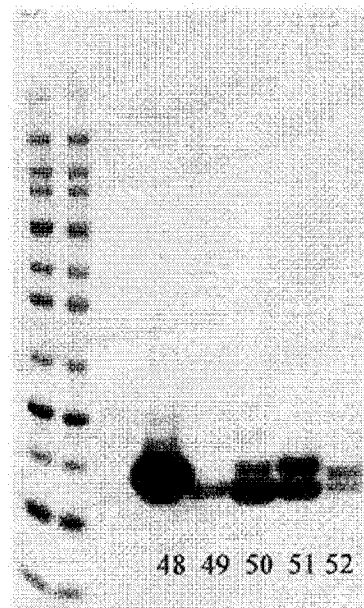


图 3

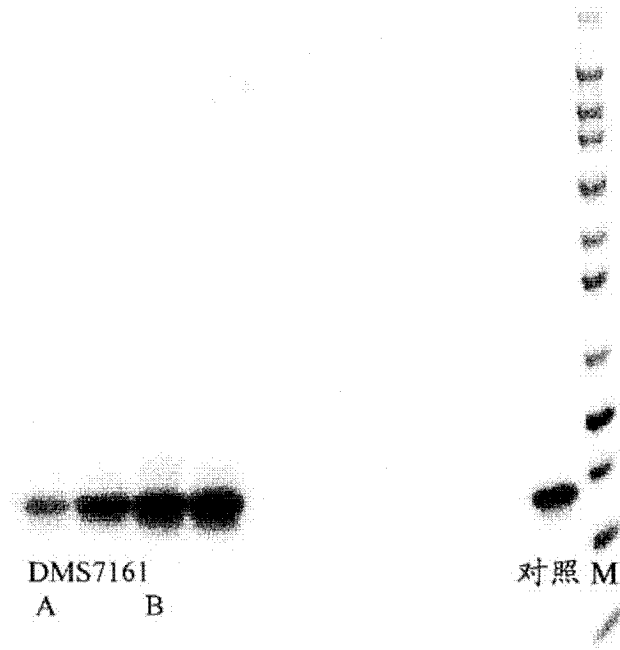


图 4

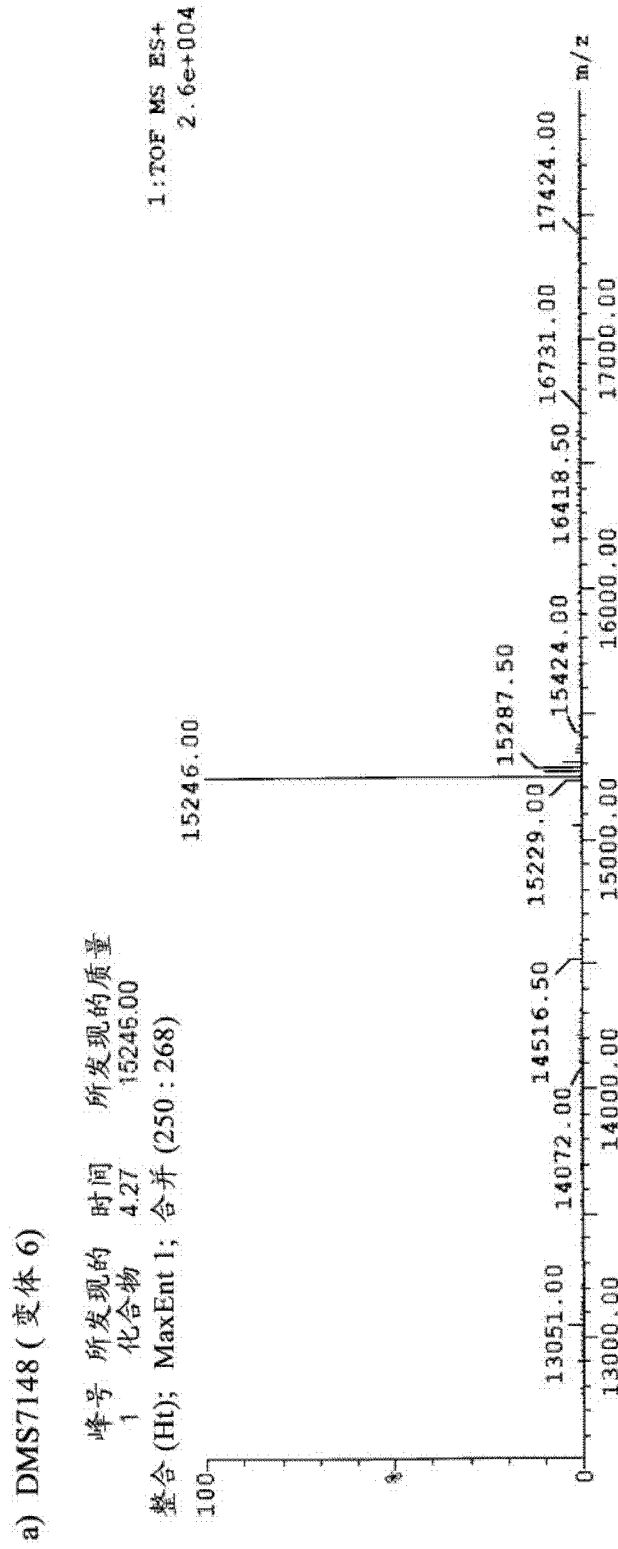


图 5

b) DMS7149 (变体 7)

峰号 所发现的 时间 所发现的质量
1 化合物 4.07 12391.00
整合 (Ht); MaxEnt 1; 合并 (238 : 256)

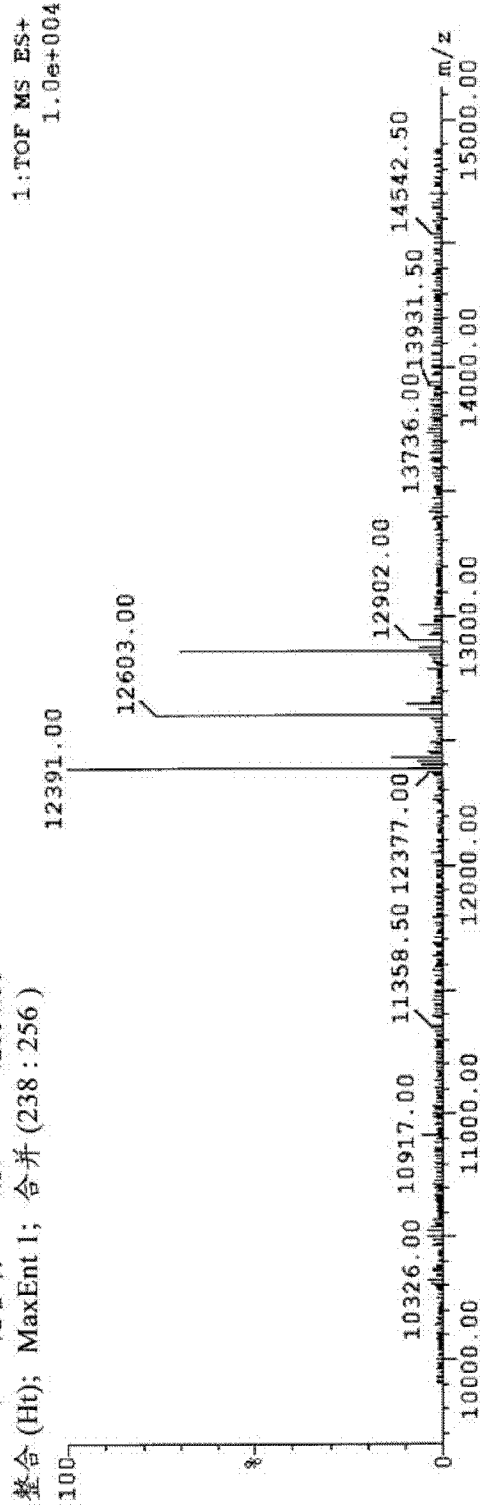


图 5(续)

c) DMS7150 (变体 8)

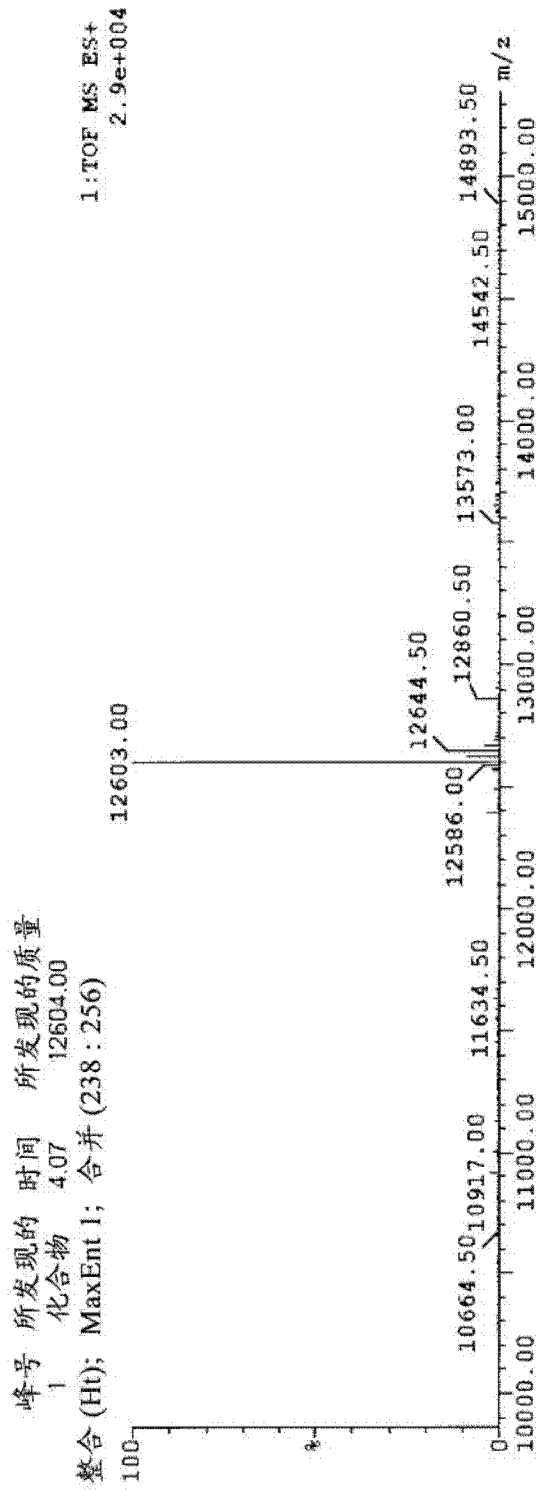


图 5(续)

d) DMS7151 (变体 9)

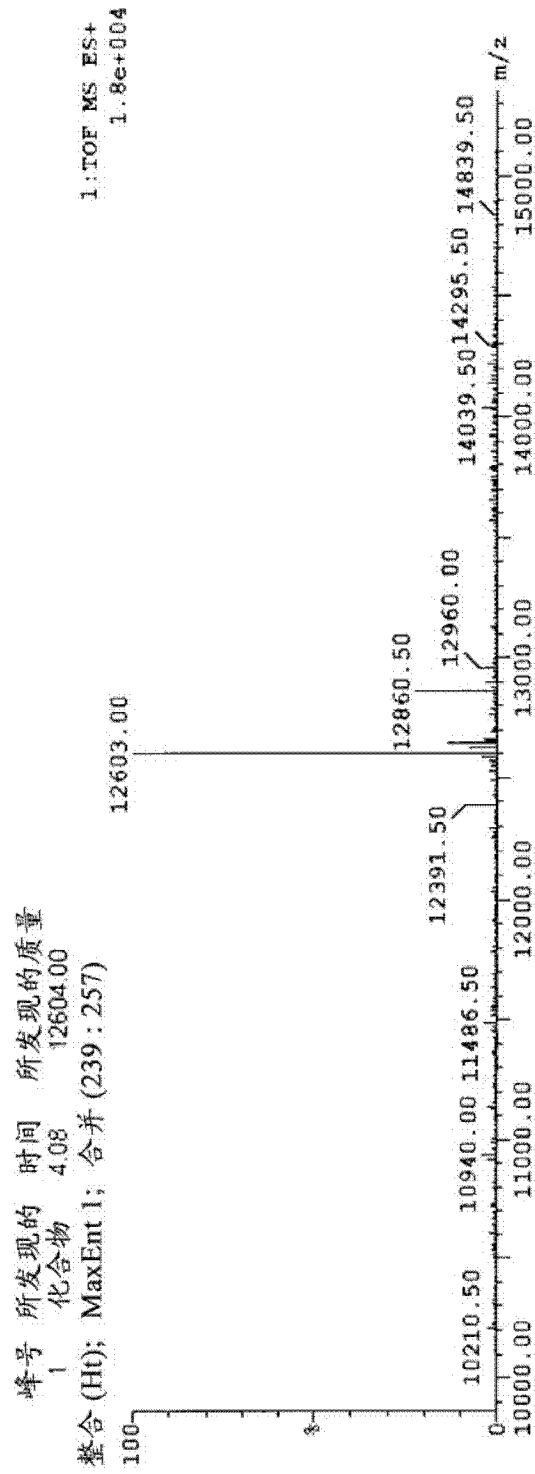


图 5(续)

e) DMS7152 (变体 10)

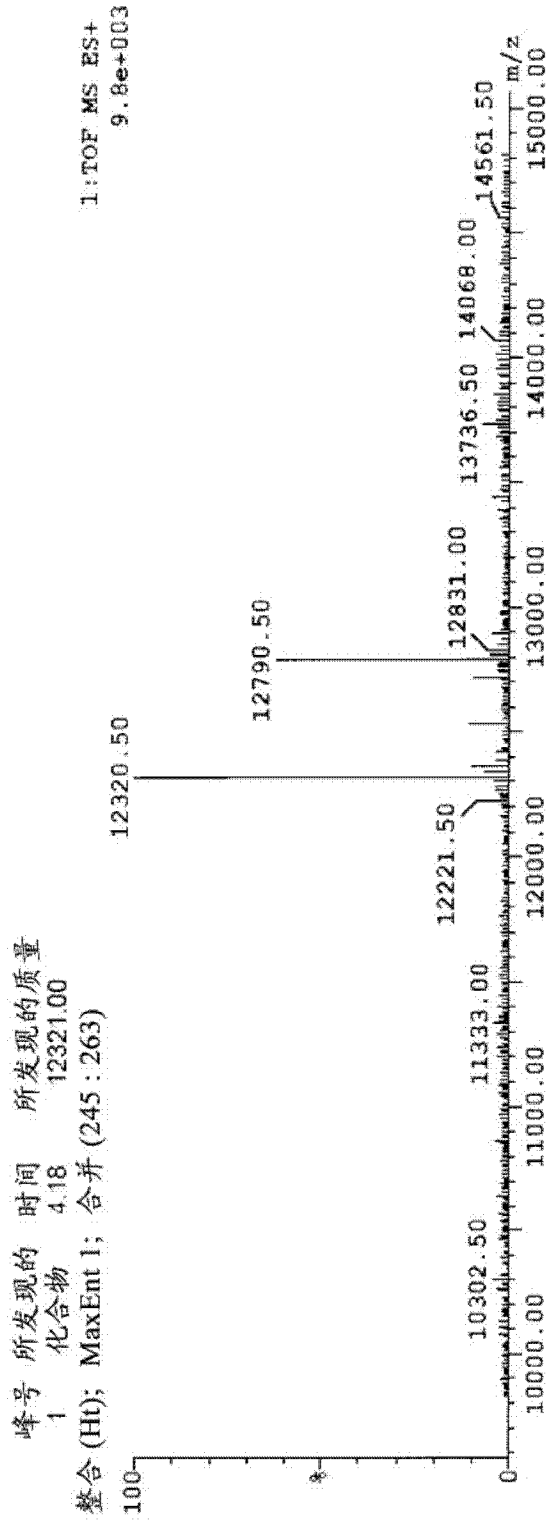


图 5(续)

f) DMS7161 (变体 II)

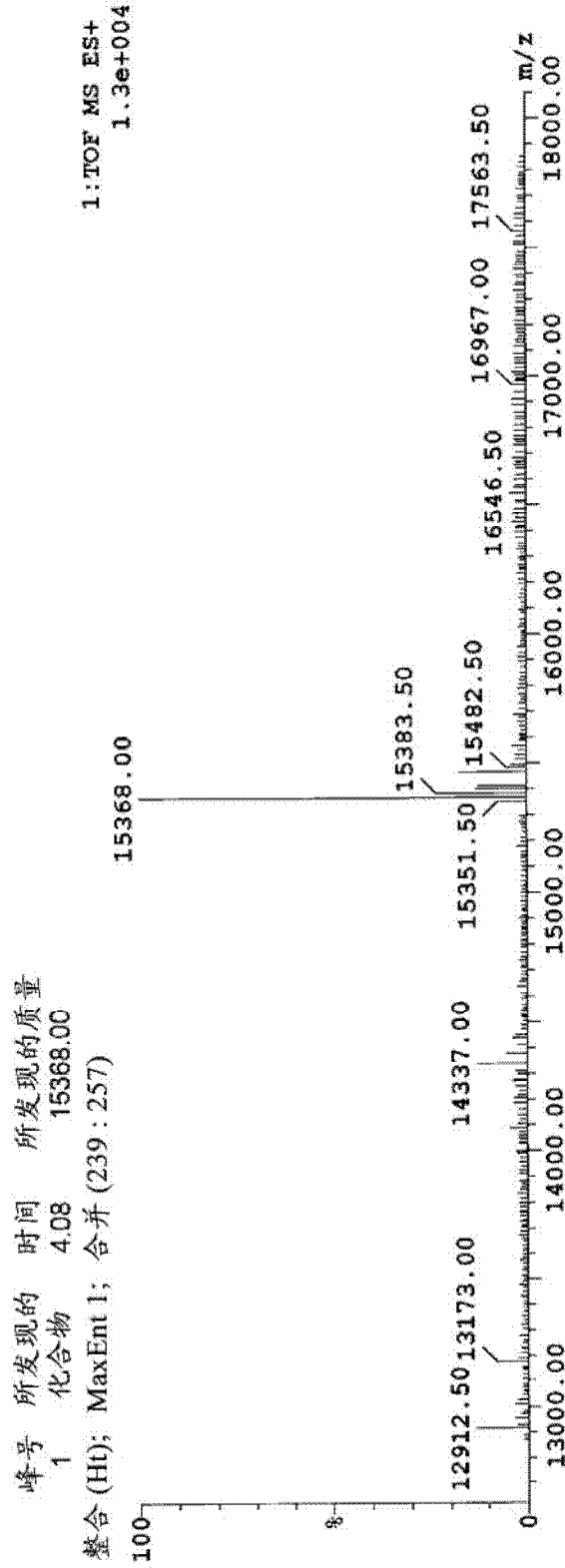


图5(续)

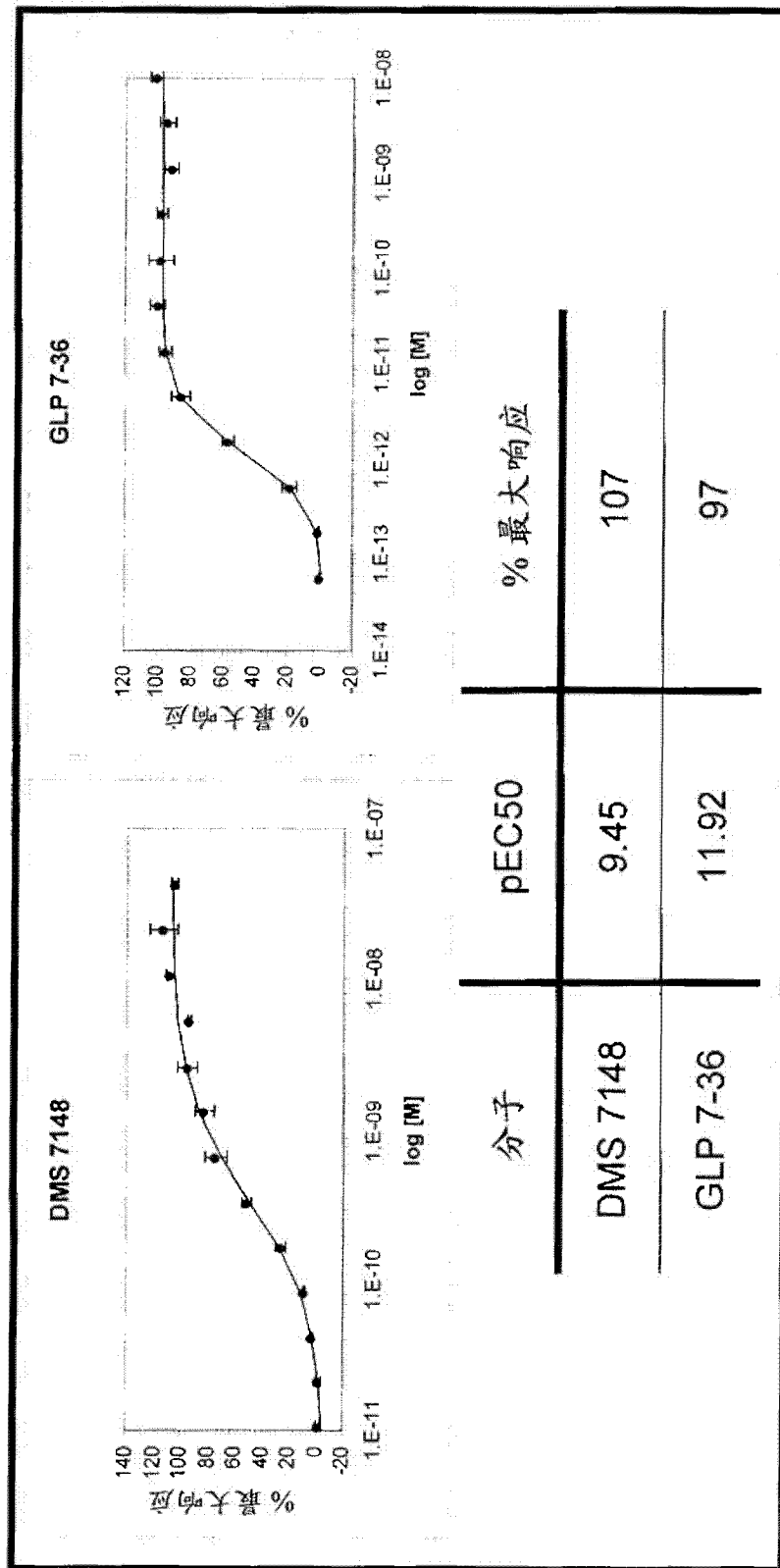


图 6

刺激 GLP-1/Ex-4 CHO 细胞
DAT01 结合试验 081128RDP

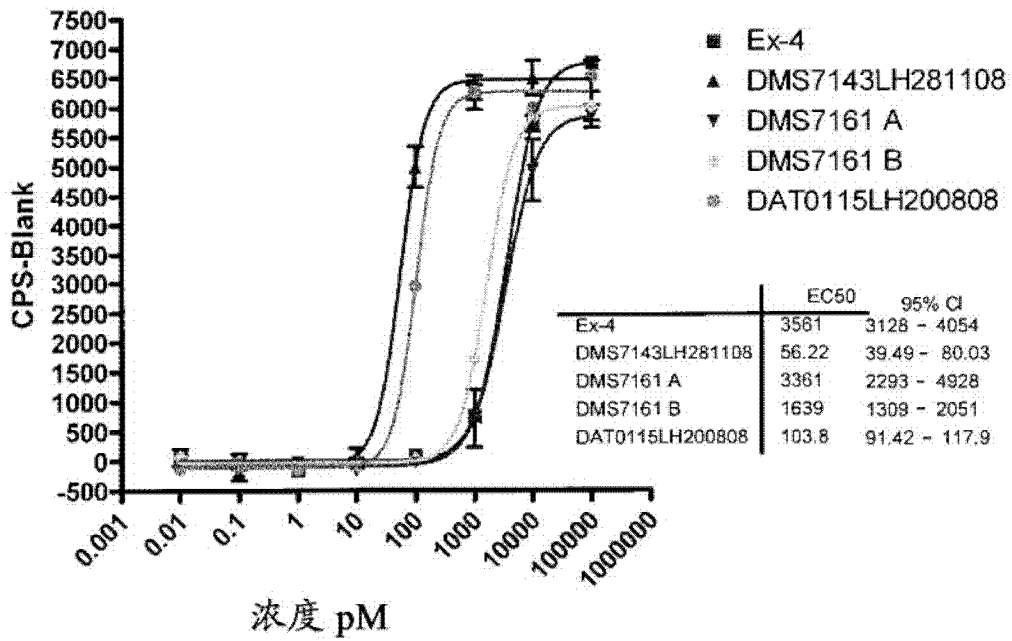


图 7