(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 104364378 A
(43)申请公布日 2015.02.18

(21)申请号 201380025049.6

(22)申请日 2013.03.19

(30)优先权数据
61/612,817 2012.03.19 US

(85)PCT国际申请进入国家阶段日
2014.11.17

(86)PCT国际申请的申请数据
PCT/AU2013/000273 2013.03.19

(87)PCT国际申请的公布数据
WO2013/138850 EN 2013.09.26

(71)申请人 麦德林制药私人有限公司
地址 澳大利亚南澳大利亚州

(72)发明人 斯坦·巴斯蒂拉斯 安吉洛·圭多林
本·亨特 赛义夫·莱希德
雷扎·扎雷尔

(74)专利代理机构 北京志霖恒远知识产权代理事务所(普通合伙) 11435
代理人 孟阿妮 郭栋梁

(51)Int.Cl.
C12N 15/63(2006.01)
C12N 15/62(2006.01)
C12N 15/70(2006.01)
C07K 19/00(2006.01)

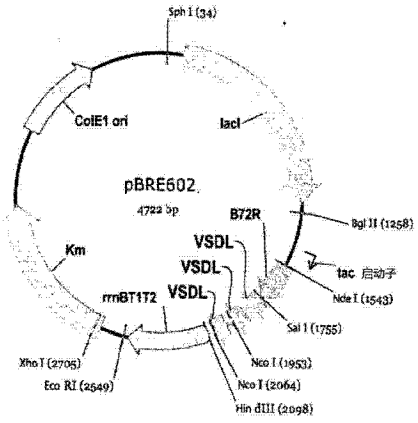
权利要求书2页 说明书27页
序列表9页 附图13页

(54)发明名称

制备重组肽的方法

(57)摘要

公开了制备重组肽如血管扩张肽(VSDL)的方法,包括涉及表达包含所述肽的串联重复的融合多肽的特殊方法,其中所述肽侧翼是肽切割位点,用裂解剂在肽切割位点切割融合多肽,使得从融合多肽释放所述肽。还公开了用于制备融合多肽的表达构建体和宿主细胞以及用于重组肽的纯化方法。



CN 104364378 A

1. 一种制备重组肽的方法，所述方法包括以下步骤：

表达包含所述肽的融合多肽，其中所述肽侧翼是一个或多个肽切割位点，并且其中所述肽另外不存在所述肽切割位点；和

用在肽切割位点切割的裂解剂在肽切割位点切割融合多肽。；

其中切割融合多肽的步骤从融合多肽释放所述肽。

2. 一种制备重组肽的方法，所述方法包括以下步骤：

表达包含所述肽的串联拷贝的融合多肽，其中所述肽侧翼是肽切割位点，并且其中所述肽特征在于所述肽另外不存在所述肽的切割位点；和

用在肽切割位点切割的裂解剂在肽切割位点切割融合多肽。；

其中切割融合多肽的步骤从融合多肽释放所述肽。

3. 如权利要求 1 或 2 所述的方法，所述肽包括根据下式所述的氨基酸序列

$X_1 \sim X_2 \sim \cdots X_{n\sim 1} \sim X_n \sim (R/K)$ (SEQ ID NO:1)；

其中 X 是除了精氨酸或赖氨酸的任何氨基酸，R/K 是精氨酸或赖氨酸，以及 n 是氨基酸的任何合适的数目。

4. 如权利要求 1 ～ 3 任一项所述的方法，其中所述肽具有天然序列。

5. 如权利要求 1 ～ 3 任一项所述的方法，其中所述肽包括血管扩张 (VSDL) 肽或者其变体或其修饰的肽。

6. 如权利要求 1 ～ 5 任一项所述的方法，其中所述 VSDL 肽包含 SEQ ID NO:2 的氨基酸序列或者由根据 SEQ ID NO:2 的氨基酸序列构成。

7. 如权利要求 1 ～ 6 任一项所述的方法，其中所述肽切割位点是胰蛋白酶切割位点。

8. 如权利要求 7 所述的方法，其中所述裂解剂是胰蛋白酶或胰蛋白酶样酶。

9. 一种制备重组血管扩张 (VSDL) 肽或者其变体或其修饰的肽的方法，所述方法包括以下步骤：

表达包含至少一种 VSDL 肽或者其变体或其修饰的肽的融合多肽，其中所述肽侧翼是一个或多个肽的切割位点，并且其中所述肽特征在于另外不存在胰蛋白酶切割位点；和

用在胰蛋白酶切割位点切割的裂解剂在胰蛋白酶切割位点切割融合多肽；

其中切割融合多肽的步骤从融合多肽释放所述肽。

10. 如权利要求 9 所述的方法，其中所述肽是包含 SEQ ID NO:2 的氨基酸序列或者由根据 SEQ ID NO:2 的氨基酸序列构成的 VSDL 肽。

11. 如权利要求 9 或 10 所述的方法，其中切割融合多肽的步骤包括用胰蛋白酶或胰蛋白酶样酶切割融合多肽。

12. 如权利要求 9 ～ 11 任一项所述的方法，进一步包括步骤：

将切割的融合多肽的 pH 调节至低于 5；和

利用至少一种阴离子交换技术从消化的融合多肽纯化所述肽。

13. 如权利要求 12 所述的方法，进一步包括喷雾或冻干干燥所述肽。

14. 如权利要求 1 ～ 13 任一项所述的方法，其中所述融合多肽包含所述肽的 2 ～ 20 个串联拷贝。

15. 如权利要求 1 ～ 13 任一项所述的方法，其中所述融合多肽包含所述肽的 3 个串联拷贝。

16. 如权利要求 1 ～ 13 任一项所述的方法，其中所述融合多肽包含所述肽的 10 个串联拷贝。

17. 如权利要求 1 ～ 15 任一项所述的方法，其中所述肽在一个末端包含一个肽切割位点。

18. 如权利要求 1 ～ 17 任一项所述的方法，其中所述肽包含一个 C －末端精氨酸或赖氨酸残基。

19. 如权利要求 1 ～ 18 任一项所述的方法，其中所述融合多肽包含具有 C －末端肽切割位点的先导融合配偶体。

20. 如权利要求 19 所述的方法，其中所述先导融合配偶体包含 C －末端精氨酸或赖氨酸残基。

21. 如权利要求 1 ～ 20 任一项所述的方法，其中表达融合多肽的步骤是使用大肠杆菌作为表达宿主进行的。

22. 一种表达构建体，包含编码与编码具有 C －末端胰蛋白酶切割位点的肽的串联拷贝的至少一个核苷酸序列操作融合的先导融合配偶体的核苷酸序列。

23. 一种表达构建体，包含编码与编码具有 C －末端胰蛋白酶切割位点的 VSDL 肽或者其变体或其修饰的肽的至少一个核苷酸序列操作融合的先导融合配偶体的核苷酸序列。

24. 如权利要求 22 或 23 所述的表达构建体，其中所述先导融合配偶体具有 C －末端胰蛋白酶切割位点。

25. 如权利要求 22 ～ 24 任一项所述的表达构建体，其中编码先导融合配偶体的核苷酸序列与编码所述肽的核苷酸序列的 2 ～ 20 个串联拷贝连接。

26. 一种重组表达肽的宿主细胞，其中所述宿主细胞包含如权利要求 22 ～ 25 任一项的表达构建体。

27. 根据权利要求 1 ～ 21 任一项所述的方法制备的重组肽。

28. 一种重组融合多肽，包含与具有 C －末端胰蛋白酶切割位点的肽的串联拷贝连接的先导融合配偶体。

29. 如权利要求 28 所述的多肽，其中所述先导融合配偶体具有 C －末端胰蛋白酶切割位点。

30. 如权利要求 28 或 29 所述的多肽，其中所述先导融合配偶体与所述肽的 2 ～ 20 个串联拷贝连接。

31. 一种重组多肽，包含或者由具有 C －末端胰蛋白酶切割位点的肽的串联拷贝构成。

32. 如权利要求 31 所述的多肽，包含所述肽的 2 ～ 20 个串联拷贝。

33. 如权利要求 32 所述的多肽，包含所述肽的 3 ～ 10 个串联拷贝。

34. 如权利要求 28 ～ 33 任一项所述的多肽，其中所述肽没有内部胰蛋白酶切割位点。

35. 如权利要求 28 ～ 34 任一项所述的多肽，其中所述肽是 VSDL 肽或者其变体或其修饰的肽。

36. 如权利要求 28 ～ 35 任一项所述的多肽，其中所述肽是包含 SEQ ID NO:2 的氨基酸序列或者由根据 SEQ ID NO:2 的氨基酸序列构成的 VSDL 肽。

制备重组肽的方法

技术领域

[0001] 本发明涉及制备肽特别是重组肽的方法。在特殊的应用中，该方法涉及包含肽的串联拷贝的融合多肽的切割。

[0002] 引用参考

[0003] 本专利申请要求 2012 年 3 月 19 日递交的美国专利申请 No 61/612817 的优选权。该申请全部内容在此引作参考。

[0004] 发明背景

[0005] 重组制备的肽对于工业和临床都越来越重要。在各种表达系统，例如细菌、酵母、昆虫、植物或哺乳动物细胞中重组产生的真核蛋白质代替合成肽制备，提供低成本。但是，重组表达是不可预期的并且经常有问题，因为很多重组表达肽不足够溶解性，不稳定，错折叠或没有活性，和 / 或以低产率表达，因此不经常以足够高的质量或有商业价值的产率制备。表达重组表达的有低分子量的肽能是特别困难的。操作上将肽与合适的肽融合配偶体融合能增强稳定性和溶解性 ；但是，一般不期望为了治疗应用的肽中存在有融合配偶体。

[0006] 通过大量的肽裂解剂包括酶促裂解剂（即蛋白酶）以及化学裂解剂能将肽裂解。许多肽裂解剂在肽内的特定氨基酸序列或基序处识别并切割肽。肽通常包含许多不同肽切割位点的多个拷贝。例如，成熟的人白细胞介素 (IL)-1β 肽 (153 个氨基酸的肽，参见 Genbank 登录号 AAA74137.1)，据预计被至少 20 种不同的肽裂解剂裂解。成熟 IL-1β 肽序列包括三个预测的 Arg-C 蛋白酶切割位点，8 个预测的 Asp-N 内肽酶切割位点，6 个预测的溴化氰切割位点，8 个预测的甲酸切割位点，3 个预测的梭菌蛋白酶切割位点，17 个预测的胰蛋白酶切割位点等。因此，在重组制备肽的过程中通常不使用肽裂解剂，因为肽常常会被裂解成更小的片段，这是不希望的。

[0007] 本发明人已经认识到，某些肽是不寻常的，在于它们缺乏特定的肽的切割位点，或包含位于肽的末端的特定的肽的切割位点的单一拷贝。例如，本发明人已经认识到天然血管扩张 (VSDL) 肽异常地包含单一的胰蛋白酶切割位点，其位于所述 VSDL 肽的 C- 末端。

[0008] 血管扩张肽 (VSDL) 是一种天然存在的 37 个氨基酸 (aa) 的肽，这是在 126 氨基酸心房钠尿肽 (ANP，也称为心房利钠因子 (ANF)) 激素原 (ANP 原的 ；Vesely，2003) 处理后在体内产生的。VSDL 由 ANP 激素原的氨基酸 31-67 组成。VSDL 的主要生物活性是通过调节排钠利尿剂、利尿剂和血液动力学效应 (Vesely，2003) 来调节血压和维持健康个体内血浆容积。通过临床前和人体临床研究两种已经对 VSDL 用于治疗心脏疾病，例如充血性心脏衰竭 (CHF) 的用途进行了研究。已经证明 VSDL 能显著提高排钠利尿剂、利尿剂和血液动力学参数而没有任何症状副作用。VDSL 被认为对于调节有益血液动力学效果是安全和有效的治疗，带有另外的有益排钠利尿剂、利尿剂和肾脏作用，同时在临床上可接受的范围内调节血浆体积和血压 (BP) 而没有严重的副作用。因此，已经提出了对急性代谢失调充血性心脏衰竭 (ADCHF) 的患者施用 VSDL。此外，VDSL 也被发现具有抗癌作用 (Skelton 等，2011)，并已被证明允许在急性肾功能衰竭的治疗中 (Vesely 等，2003 年)。因此，可以理解，VSDL 是用于各种疾病治疗的有用候选者。基于天然 VSDL 序列的合成产生的 VSDL 已被用于临床试验，

并已被有利地证明可用于治疗疾病；但是，多肽的合成制备是一个昂贵的过程，特别是对于在工业规模上所需要的肽。

[0009] 本发明人设计了一种一流的重组制备系统，其利用从融合多肽切割肽。

发明内容

[0010] 因此，在第一个方面，本发明提供了一种制备重组肽的方法，所述方法包括以下步骤：

[0011] 表达包含所述肽的融合多肽，其中所述肽侧翼是一个或多个肽切割位点，并且其中所述肽另外不存在所述肽切割位点；和

[0012] 用在肽切割位点切割的裂解剂在肽切割位点切割融合多肽；

[0013] 其中切割融合多肽的步骤从融合多肽释放所述肽。

[0014] 在第二方面，本发明提供了一种制备重组肽的方法，所述方法包括以下步骤：

[0015] 表达包含所述肽的串联拷贝的融合多肽，其中所述肽侧翼是一个或多个肽的切割位点，并且其中所述肽特征在于所述肽另外不存在所述肽的切割位点；和

[0016] 用在肽切割位点切割的裂解剂在肽切割位点切割融合多肽。；

[0017] 其中切割融合多肽的步骤从融合多肽释放所述肽。

[0018] 在一种实施方案中，所述融合多肽包含 2 至 20 个（或更多）肽的串联拷贝。在一种实施方案中，所述融合多肽包含 3 个肽的串联拷贝。在进一步的实施方案中，所述融合多肽包含 10 个肽的串联拷贝。

[0019] 在一种实施方案中，所述肽在其一个末端包含肽的肽切割位点。

[0020] 在进一步的实施方案中，肽包含 C- 末端精氨酸 (R) 残基。

[0021] 在另一种实施方案中，所述肽包含根据式 $X_1-X_2-\cdots X_{n-1}-X_n-(R/K)$(SEQ ID NO:1) 的氨基酸序列；其中 X 是除了精氨酸或赖氨酸的任何氨基酸，R/K 为精氨酸或赖氨酸，和 n 是氨基酸的任何合适的数目。在一个实施例中，所述肽具有天然序列。

[0022] 在另一种实施方案中，肽包括血管扩张 (VSDL) 肽或其变体。

[0023] 在一个具体的实施方案中，VSDL 肽包含根据 SEQ ID NO:2 的氨基酸序列或其变体。

[0024] 在一种实施方案中，所述融合多肽包括具有 C- 末端肽切割位点的先导融合配偶体。

[0025] 在一个优选实施例中，先导融合配偶体包含 C 末端精氨酸残基。

[0026] 在一个特定实施方案中，先导融合配偶体包括根据 SEQ ID NO:3 的氨基酸序列。

[0027] 在一种实施方案中，所述肽的切割位点是胰蛋白酶切割位点。在一种实施方案中，裂解剂是胰蛋白酶或胰蛋白酶样酶。

[0028] 在第三方面，本发明提供了一种制备重组血管扩张 (VSDL) 肽（或其变体或其修饰的肽）的方法，所述方法包括以下步骤：

[0029] 表达包含至少一种 VSDL 肽（或其变体或其修饰的肽）的融合多肽，其中所述肽侧翼是一个或多个肽的切割位点，并且其中所述肽特征在于另外不存在胰蛋白酶切割位点；和

[0030] 用在胰蛋白酶切割位点切割的裂解剂在胰蛋白酶切割位点切割融合多肽。；

[0031] 其中切割融合多肽的步骤从融合多肽释放所述肽。

[0032] 在另一个方面，本发明提供包含与编码具有 C- 末端胰蛋白酶切割位点的 VSDL 肽（或其变体或其修饰的肽）的至少一个核苷酸序列操作融合的先导融合配偶体（具有 C- 末端胰蛋白酶切割位点）的核苷酸序列的表达构建体。

[0033] 在一种实施方案中，编码先导融合配偶体的核苷酸序列可操作地融合到编码具有 C- 末端胰蛋白酶切割位点的 VSDL 肽（或其变体或其修饰的肽）的核苷酸序列的 2 — 20 个（或更多）串联拷贝。

[0034] 在另一个方面，本发明提供用于重组表达 VSDL 肽（或其变体或其修饰的肽）的宿主细胞，其中所述宿主细胞包含表达构建体，表达构建体包含编码与具有 C- 末端胰蛋白酶切割位点的 VSDL 肽（或其变体或其修饰的肽）的至少一个核苷酸序列操作融合的具有 C- 末端胰蛋白酶切割位点的先导融合配偶体的核苷酸序列。

附图说明

[0035] 图 1 提供了包含指定为 B72R 的先导融合配偶体的 VSDL 肽的氨基酸序列和 VSD1 — 链节和 3 — 链节融合蛋白的氨基酸序列；斜体序列代表先导融合配偶体 (B72R) 并且箭头指示胰蛋白酶切割位点；

[0036] 图 2 提供了 B72R 和 B72(Q)R-VSDL 基因融合体的核苷酸序列，包括侧翼的限制性内切酶切割位点的核苷酸序列；蛋白质编码区以粗体表示，其中斜体序列代表先导融合配偶体的核苷酸序列。每个基因融合体的长度示于括号内；

[0037] 图 3 提供了用于表达包含 3,6,8 和 10 个 VSDL 串联重复的以 B72R-VSDL 和 B72R- 为基础的融合多肽的质粒的简明示意图。质粒的功能是复制 (ColE1ori) 起源，lac 阻遏基因 (lacI)，TRC 和 tac 启动子，VSDL 基因融合体，转录终止子 (rrnBT1T2) 和卡那霉素抗性标记 (Km)；

[0038] 图 4 提供了 B72R-VSDL 和 B72R-$(VSDL)_3$ 表达的 SDS-PAGE 分析的图像；分级培养样品 (P：沉淀球 / 不溶部分，WCL: 全细胞溶胞产物和 S：上清液 / 可溶性级分在减少 4-12％的 NuPAGE MES 凝胶公司，加利福尼亚州卡尔斯巴德，美国）上分析，并用 SimplyBlue™ Safe(Invitrogen)，SeeBlue® 预先染色蛋白质标准物 (Invitrogen) 被用作分子量标记；

[0039] 图 5 提供的 B72(Q)R-VSDL 融合多肽表达的 SDS-PAGE 分析的图像；将指定培养样品进行归一化并且溶胞产物在用 SimplyBlue™ Safe 染色的减少 4-12％ NuPAGE MES 凝胶上分析，用SeeBlue®预先染色蛋白质标准物作为分子量标记；

[0040] 图 6 提供分批加料培养之后的 B72R-VSDL 和 B72R-$(VSDL)_3$ 表达的 SDS-PAGE 分析的图像；将指定培养样品进行归一化并且溶胞产物在用 SimplyBlue™ Safe 染色的减少 4-12％ NuPAGE MES 凝胶上分析，用SeeBlue®预先染色蛋白质标准物作为分子量标记；

[0041] 图 7 提供显示室温下消化过夜后对从 B72R-$(VSDL)_3$ 溶胞产物 (3g/L 融合多肽）产生的一定量 VSDL 加入TrypZean® (Sigma-Aldrich Co，圣路易斯，Mo，美国）的作用的图形结果；

[0042] 图 8 提供了一个显示来自消化的溶胞产物的重组 VSDL 的色谱分离 (Amberchrom CG300C；Rohm and Haas Company，宾夕法尼亚州费城，美国）的示踪图；实线＝UV 280 吸收

和虚线＝导电性；

[0043] 图 9 提供显示来自消化的溶胞产物的重组 VSDL 的色谱分离 (Fractogel TMAE；默克公司，新泽西州怀特豪斯站，美国）的示踪图；实线＝ UV 280 吸收和虚线＝导电性；

[0044] 图 10 显示根据加里 50 分光光度计（瓦里安公司，帕洛阿尔托加利福尼亚州，美国）测定的合成 VSDL 肽和重组 VSDL（右）的 UV 280 吸收曲线；

[0045] 图 11 提供显示来自消化的溶胞产物的重组 VSDL 的色谱分离（改进的 Amberchrom CG300C 方案）的示踪图，方框代表汇集的洗脱级分；和

[0046] 图 12 提供显示 Chromasorb 膜过滤之后的来自消化的溶胞产物的重组 VSDL 的分离的示踪图，其中方块代表不同的汇集的洗脱级分。

具体实施方式

[0047] 本发明人已经认识到，有利地，某些肽的氨基酸序列不同寻常地缺乏特定肽切割位点，除了位于所述肽的一个末端的单一肽切割位点，或能被工程化成具有这样的序列。本发明人还认识到，通过表达融合多肽前后的肽，可以在肽的另一末端提供另一个肽切割位点，以使得所表达的融合多肽的切割从融合多肽释放肽。这样的制备系统能够有利地提供足够数量和质量的治疗应用的重组产生感兴趣的肽的方法。

[0048] 因此，在第一个方面，本发明提供了一种制备重组肽的方法，所述方法包括以下步骤：

[0049] 表达包含所述肽的融合多肽，其中所述肽侧翼是一个或多个肽切割位点，并且其中所述肽另外不存在所述肽的切割位点；和

[0050] 用在肽切割位点切割的裂解剂在肽切割位点切割融合多肽。；

[0051] 其中切割融合多肽的步骤从融合多肽释放所述肽。

[0052] 本领域技术人员理解术语”重组”是指包含或由多于一种起源的遗传材料产生的生物材料，例如，重组肽通常被理解为是从编码第一来源的肽的多核苷酸分子产生，所述第一来源的肽在第二来源的合适的宿主细胞中被表达或制备。能产生重组肽的宽范围的表达系统是已知的，例如，在适合于特殊肽表达的条件下的细胞培养物。宿主细胞能包括细菌细胞（例如大肠杆菌 (Escherichia coli)，链霉菌属 (Streptomyces sp.)，芽孢杆菌属 (Bacillus sp.)，假单胞菌属 (Pseudomonas sp.)，葡萄球菌属 (Staphylococcus sp.)（例如鼠伤寒沙门氏菌 (S.typhimurium) 等）等），真菌细胞，如酵母细胞（例如酿酒酵母 (Saccharomyces cerevisiae)，巴斯德毕赤酵母 (Pichia pastoris) 等），霉菌或丝状真菌（例如曲霉 (Aspergillus sp.)），昆虫细胞（例如果蝇 S2 和草地夜蛾 (Spodoptera frugiperda)（例如夜蛾 Sf9 细胞），并包括杆状病毒辅助系统），哺乳动物细胞（适合的大量细胞系包括中国仓鼠卵巢 (CHO) 细胞，猴肾 (COS) 细胞，人胚胎肾 (HEK) 细胞，幼仓鼠肾 (BHK) 细胞，鼠黑素瘤细胞等），和植物细胞。重组多肽或者也可以在转基因植物和动物中表达。在本发明中，优选的是表达融合多肽的步骤是使用大肠杆菌作为表达宿主进行的。

[0053] 本领域技术人员理解术语”肽”是指通过肽键键合在一起的氨基酸的单一线性聚合物链，第一个氨基酸是在氨基 (N) 端和最后一个氨基酸是在羧基 (C) 末端。在本发明的上下文中，应当理解本文所用的术语”肽”是为了指感兴趣的肽，其要被表达，并在融合多肽裂解之后从融合多肽释放。肽可以是全长肽（即如本领域技术人员理解的从编码核苷酸序

列转录的全长肽）。但是肽还可以从较大的肽或多肽衍生，例如，肽可以是前蛋白原的成熟肽部分，或者是能被加工成很多较小肽的多肽的一部分，或者是较大的肽的片段，等。优选地，肽可从编码核苷酸序列翻译成它的最终形式，即，肽不经历任何附加的翻译后加工只是被根据本发明的裂解剂裂解。但是，本发明并不打算排除实施方案，其中肽是一个较大肽的一部分，如前蛋白原或多肽，其包含能够被加工成许多较小的肽的氨基酸序列（例如通过表达后处理（例如裂解）去除较大肽的部分，其中所述处理是除根据本发明的裂解之外的）。肽可以是天然肽或如下所述的其经修饰的肽。

[0054] 本领域技术人员理解的是，与肽相关联使用时的术语”氨基端”或“氨基方向”指的是朝向肽的 N- 末端方向，同样地，与肽相关联使用时的术语”羧基端”或“羧基方向”指的是朝向肽的 C- 末端方向。

[0055] 如本文所用的术语”融合多肽”意在指至少两个肽通过肽键键合在一起的氨基酸单一线性聚合物链，使得第一肽的 C- 末端残基邻接第二肽的 N- 末端残基（任选被一个短接头序列分隔）。本领域技术人员理解，一些肽能以这种方式融合在一起，在这种情况下，所述第二肽的 C- 末端残基可以邻接第三肽的 N- 末端残基，等等。

[0056] 本领域技术人员明白术语”肽的切割位点”是指在那里裂解剂切割肽主链的位点，其中那个位点与作为裂解剂的识别位点的特定氨基酸或氨基酸特定序列相关。在氨基酸的特定识别序列的情况下，在那个序列中的一个位点，在序列的 C- 末端或 N- 末端，或在序列的外部，裂解剂可以切割肽主链，即是，一些残基从序列的羧基或者氨基方向被去除。

[0057] 根据本发明，所述肽”侧翼是一个或多个肽的切割位点”。这个短语意指肽的切割位点位于融合多肽内，在肽的至少一个末端，更优选地，在肽的两个末端（例如针对肽的每个复制体位于一个 N- 末端和一个 C- 末端，使得用在肽的切割位点切割的裂解剂切割融合多肽时，从融合多肽能释放所述肽。

[0058] 裂解剂可以是以特定方式即在与特定的氨基酸或氨基酸的特定序列相关联的特定肽切割位点切割肽的本领域技术人员已知的任何合适的裂解剂。裂解剂可以是化学裂解剂或酶促裂解剂。在本发明中可以使用的多种裂解剂列在表 1 中。

[0059] 表 1. 肽裂解剂和它们各自的肽切割位点

[0060]

裂解剂	切割位点
Arg-C 蛋白酶	R↓
Asp-N 肽内切酶	↓N
Asp-N 肽内切酶+ N-末端 Glu	↓N，↓E
梭菌蛋白酶（梭菌肽酶 B）	R↓
肠肽酶（肠激酶）	DDDDK↓（SEQ ID NO: 4）
凝血酶	R↓，但是优选 LVPR↓GS
烟草蚀刻病毒蛋白酶	EXXYXQ↓（G/S）（SEQ ID NO: 5）或 XYXQ↓（G/S）（SEQ ID NO: 6），其中 X 是任何氨基酸
胰蛋白酶	K↓或 R↓
BNPS-粪臭素[2-（2-硝基苯基亚磺酰）-3-甲基吲哚]	Y↓
N-氯琥珀酰亚胺（NCS）	W↓
N-溴琥珀酰亚胺（NBS）	W↓
溴化氢	M↓
甲酸	D↓

[0061] 在这里，肽切割位点如作为裂解剂识别位点的单一的特定氨基酸的氨基侧或羧基侧箭头所示，或者如作为裂解剂识别位点的氨基酸的特定序列的氨基侧，羧基侧，或序列内箭头所示。与肽切割位点（这里的所有氨基酸序列）相关使用的三字母或单字母氨基酸代码如表 2 所示。

[0062] 表 2. 氨基酸代码

[0063]

氨基酸	三字母代码	单字母代码
丙氨酸	Ala	A
精氨酸	Arg	R
天冬酰胺	Asn	N
天冬氨酸（天冬氨酸盐）	Asp	D
半胱氨酸	Cys	C
谷氨酰胺	Gln	Q
谷氨酸（谷氨酸盐）	Glu	E
甘氨酸	Gly	G
组氨酸	His	H
异亮氨酸	Ile	I
亮氨酸	Leu	L
赖氨酸	Lys	K
甲硫氨酸	Met	M
苯丙氨酸	Phe	F
脯氨酸	Pro	P
丝氨酸	Ser	S
苏氨酸	Thr	T
色氨酸	Trp	W
酪氨酸	Tyr	Y

缬氨酸	Val	V

[0064] 在优选的实施方案中，切割位点是酪氨酸或酪氨酸样酶（即在酪氨酸切割位点裂解）。

[0065] 但是其他裂解剂也可以适合用于本发明，包括，例如，半胱天冬酶 1，半胱天冬酶 2，半胱天冬酶 3，半胱天冬酶 4，半胱天冬酶 5，半胱天冬酶 6，半胱天冬酶 7，半胱天冬酶 8，半胱天冬酶 9，半胱天冬酶 10，高特异性胰凝乳蛋白酶，低特异性胰凝乳蛋白酶，因子 Xa，甲酸，谷氨酰肽内切酶，颗粒酶 B，羟胺，亚碘酰安息香酸，LysC，LysN，NTCB(2- 硝基 -5- 氰硫基苯甲酸)，胃蛋白酶 (pH 1.3)，胃蛋白酶 (pH>2)，脯氨酸 - 肽内切酶，蛋白酶 K，葡萄球菌肽酶 I，嗜热菌蛋白酶，反素蛋白酶等。本领域技术人员将明白其他裂解剂可以根据本发明来使用，除了确定用具体的裂解剂切割要求什么条件之外，要知道确定合适的裂解剂的肽切割位点的方法。

[0066] 在一种实施方案中，在肽的两端提供肽切割位点，并且可以需要两种裂解剂从融合多肽释放肽。

[0067] 本发明人已经认识到融合多肽可有利地包含该肽的串联拷贝，其中肽的每个拷贝侧翼连着一个或多个肽切割位点。

[0068] 因此，在第二方面，本发明提供了一种制备重组肽的方法，所述方法包括以下步骤：

[0069] 表达包含肽的串联拷贝的融合多肽，其中所述肽侧翼是一个或多个肽切割位点，并且其中所述肽另外不存在所述肽切割位点；和

[0070] 用在肽切割位点切割的裂解剂在肽切割位点切割融合多肽；

[0071] 其中切割融合多肽的步骤从融合多肽释放所述肽。

[0072] 短语”肽串联拷贝”被理解为是指以串联头 - 尾排列可操作地连一个以上的肽，使得第一肽的 C- 末端残基相邻接第二肽的 N- 末端残基。在一些实施方案中，其中包括两个以上的多联体肽，所述第二肽的 C- 末端残基相邻接第三肽的 N- 末端残基，等等。优选地，一个肽的 C- 末端残基直接邻接下一个肽的 N- 末端残基。但是，本领域技术人员显而易见的是通过能潜在地用来将肽拷贝连接在一起的短连接序列能连接肽。要明白这样的实施方案落在本发明的范围之内。

[0073] 任何数量的肽的拷贝可以串联排列在融合多肽内，前提是融合多肽被表达到令人满意的水平。在一种实施方案中，融合多肽包含 2 至 20 个（或更多）的肽的串联拷贝。例如，融合多肽可以包含肽的二，三，四，五，六，七，八，九，十，十一，十二，十三，十四，十五，十六，十七，十八，十九个或二十个拷贝。在一种实施方案中，所述融合多肽包含 2 至 10 个串联肽。在一种实施方案中，所述融合多肽包含 3 个串联肽。在另一种实施方案中，所述融合多肽包含肽的 10 个串联拷贝。但是，融合多肽中有特定数目拷贝的情况下，例如，肽的 3 个串联拷贝，要理解的是，多肽中总共有三个所述肽。因此，”肽的 3 个串联拷贝”不是要指除了所述肽的 3 个”拷贝”之外有一个初始肽。包含多联体肽的融合多肽（和编码所述融合多肽的多核苷酸构建体）另外在这里被指作，例如，当存在三个串联肽时，”3- 链节”，当存在六个串联肽时，”6- 链节”，等。同样，包含单一肽的融合多肽（和编码所述融合多肽的多核苷酸构建体）另外在这里被指作”1- 链节”。

[0074] 根据本发明，所述肽”侧翼是一个或多个肽的切割位点”。在所述肽的串联拷贝前后，这个短语意指肽切割位点位于融合多肽中，在所述肽的每个拷贝的至少一个末端，更优选地，在所述肽的两端（例如就所述肽的每个拷贝来说一个位于 N- 末端和一个位于 C- 末端），使得使用在肽切割位点切割的裂解剂对融合多肽切割时所述肽从所述融合多肽被释放。

[0075] 在一个优选的实施方案中，在肽的每个拷贝的两端提供肽的切割位点，一个单一肽切割位点位于各个肽拷贝之间，如此排列，使得在所述位点由单一裂解剂对肽切割，同时产生肽的一个拷贝的 C- 末端和肽的邻接拷贝的 N- 末端。在这个实施方案中，优选单一肽切割位点是肽所有拷贝之间的相同类型（例如 R ↓ ）。

[0076] 本发明第二方面的方法可以特别适合制备特征在于没有内在肽切割位点的任何肽。

[0077] 在一种实施方案中，本发明的肽具有天然序列。上下文中术语”天然”意指肽的序列是自然界发现的，即没有经修饰的，例如去除肽切割位点或者另外导入氨基酸取代。但是，肽可以具有截短的天然序列或天然变体序列。

[0078] 在一个可替代实施方案中，肽可被修改。适用于本发明中使用的修饰肽可包括天然肽的衍生物，变体或模拟物，其包括氨基酸序列小的变化，前提是这样的变化不会导致生物活性任何实质性下降或变化（即相比于天然肽），并且没有引入另外的肽切割位点（优选地，没有其它胰蛋白酶切割位点）。这些变化可包括保守的氨基酸取代，如 :Gly，Ala，Val，Ile，Leu，Met ;Asp，Glu，Asn，Gln ;Ser，Thr ;Lys，Arg，His ;Phe，Tyr，Trp，His ;和 Pro，Nα - 烷基氨基酸 ;和非保守性氨基酸取代。表 3 详细列出普通的保守性氨基酸取代。

[0079] 表 3. 例示的保守性氨基酸取代

[0080]

	保守性氨基酸取代
Ala	Val*，Leu，Ile
Arg	Lys*
Asn	Gln*，His，Asp
Asp	Glu*，Asn
Cys	Ser
Gln	Asn*，His，
Glu	Asp*，γ - 羧基谷氨酸 (Gla)
Gly	Pro
His	Asn，Gln，

Ile	Leu*, Val, Met, Ala, Phe, 正亮氨酸 (Nle)
Leu	Nle, Ile*, Val, Met, Ala, Phe
Lys	Arg*, Gln, Asn, 鸟氨酸 (Orn)
Met	Leu*, Ile, Phe, Nle
Phe	Leu*, Val, Ile, Ala
Pro	Gly*, 羟基脯氨酸 (Hyp), Ser, Thr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp, Phe*, Thr, Ser
Val	Ile, Leu*, Met, Phe, Ala, Nle

[0081] *指优选的保守性取代

[0082] 合适的经修饰的肽可使用本领域技术人员公知的任何方法来获得。例如，肽的模拟物能使用本领域技术人员公知的任何方法根据没有二级和三级结构信息的氨基酸序列进行设计 (Kirshenbaum 等，1999)。

[0083] 此外，本领域技术人员会理解，它可以有可能工程处理肽以除去现有的内部肽切割位点（例如，通过取代或缺失适当的氨基酸）。例如，天然人胰高血糖素样肽 (GLP)-1 的氨基酸序列（来自于人胰高血糖素前原的氨基酸 98 至 127，它具有 NCBI 参考 NP_002045）包含一个 C- 末端精氨酸 (R) 残基和两个内部赖氨酸 (K) 残基。两个内部赖氨酸残基能被修饰为，例如，谷氨酰胺 (Q) 或天冬酰胺 (N)，其去除了内部胰蛋白酶切割位点。在另一个实施例中，相比细胞质蛋白，一些膜蛋白质具有低数量的精氨酸和赖氨酸残基，而这样的残基可通过修改被移除以创建无内部肽切割位点的肽。

[0084] 在一种实施方案中，肽的每个串联拷贝与所述肽的每一个其它串联拷贝是相同的。然而，在可替代实施例中，肽的不同变体可在多联体内被表达。也就是说，术语”肽的串联拷贝”意在包括与所述肽的另一个串联拷贝相比，所述肽的一个（或多个）”串联拷贝”是其天然变体或其修饰的肽。例如，所述肽的一个串联拷贝可以包括天然序列，而所述肽的另一个串联拷贝包括所述肽的天然变体，等。在另一个实施例中，所述肽的不同串联拷贝可以包括不同的天然或修饰的氨基酸序列，使得产生适合作为裂解剂的识别位点的不同的氨基酸或氨基酸序列。

[0085] 在一种实施方案中，所述肽的切割位点是一种胰蛋白酶切割位点。

[0086] 术语”胰蛋白酶切割位点”被本领域技术人员理解为是指其中所述肽主链由胰蛋白酶切割的位点，其中所述位点与适合作为胰蛋白酶识别位点的特定氨基酸或特定氨基酸

序列相关。胰蛋白酶倾向于在赖氨酸(K)或精氨酸(R)氨基酸残基的羧基侧裂解肽，除了当它们后面紧跟着脯氨酸(P)。但是，赖氨酸或精氨酸残基周围确切的氨基酸残基可以影响胰蛋白酶裂解的效率。胰蛋白酶识别位点的例子包括，例如，DR ↓ E，QR ↓ E，SK ↓ N，等，其中箭头指肽切割位点（即肽主链被裂解处的位点）。但是本领域技术人员理解大多数赖氨酸或精氨酸残基可以适合作为胰蛋白酶羧基侧切割的位点。

[0087] 切割步骤能通过，例如，用在胰蛋白酶切割位点（即在精氨酸或赖氨酸残基的羧基端）处切割的裂解剂在适合发生裂解的条件下温育融合多肽，这是本领域技术人员明白的。

[0088] 在一种实施方案中，裂解剂是胰蛋白酶。胰蛋白酶是一种丝氨酸蛋白酶并且经常在研究实验室使用以裂解肽。根据本发明使用的胰蛋白酶可从许多不同的物种（包括牛，猪，人等）分离，所有这些是可商购的和/或重组产生的胰蛋白酶（包括牛，猪等），这也是容易获得的，都可以使用。此外，胰蛋白酶能在植物细胞，例如玉米细胞中重组产生。TrypZean®（如牛 TrypZean，Sigma-Aldrich 公司）是植物中制备的重组胰蛋白酶的例子。因此，在一种实施方案中，裂解剂是植物细胞重组产生的胰蛋白酶。植物产生的重组胰蛋白酶特别有用于切割用于治疗目的的肽，以减少不希望的动物来源的污染物存在于最终肽产物的可能性。但是，以一种方式产生使得胰蛋白酶基本上没有（或者有可接受水平的）动物污染物的其他胰蛋白酶产物可以用于治疗应用而制备的肽。例如，裂解剂可以是酵母细胞产生的胰蛋白酶。优选地，裂解剂是重组胰蛋白酶。

[0089] 然而，本领域技术人员认识到，有许多具有胰蛋白酶或胰蛋白酶样活性（即在胰蛋白酶切割位点切割的胰蛋白酶样酶）的其它酶也可以适合用于在按照本发明在某些情况下在胰蛋白酶切割位点切割肽。例如，凝血酶是胰蛋白酶样丝氨酸蛋白酶，其优先识别氨基酸序列的 LVPR ↓ GS(SEQ ID NO:7) 并且在精氨酸和甘氨酸残基之间的肽切割位点（即在一个精氨酸(R)残基的羧基侧）切割。因此，凝血酶将在一些胰蛋白酶切割位点裂解肽。在另一个实施例中，梭菌蛋白酶（也称为蛋白内切酶 Arg-C 和梭状芽胞杆菌肽酶 B）主要在精氨酸残基处切割，但也可以在赖氨酸残基处切割至较轻程度。因此，梭菌蛋白酶将在胰蛋白酶切割位点切割。本领域技术人员理解，可以存在也适合于在按照本发明的胰蛋白酶切割位点裂解肽的其它裂解剂。因此，在一种实施方案中，裂解剂是胰蛋白酶或胰蛋白酶样酶。在一种实施方案中，裂解剂是重组胰蛋白酶或重组胰蛋白酶样酶。胰蛋白酶消化时间能根据需要改变。例如，融合多肽可以被胰蛋白酶或胰蛋白酶样酶消化 5 分钟，1 小时，2 小时，6 小时，或过夜等。

[0090] 在一种实施方案中，所述肽包含一个 C- 末端胰蛋白酶切割位点。在一种实施方案中，肽包含根据下式的氨基酸序列：

[0091] $X_1-X_2-\cdots X_{n-1}-X_n-(R/K)$(SEQ ID NO:1)；

[0092] 其中 X 是除了精氨酸或赖氨酸的任何氨基酸，R/K 是精氨酸或赖氨酸，和 n 是氨基酸的任何合适的数目。

[0093] 在一种实施方案中，制备用于治疗用途的肽。因此，肽可包括细胞因子或激素（包括从激素原分子再释放的激素）或表位或抗原（如用作疫苗）。

[0094] 所述肽可以是，例如，3 ～ 100 个氨基酸长度，更优选地，长度为 5 ～ 50 个氨基酸，还更优选地，长度为 10 ～ 45 个氨基酸。在优选实施方式中，所述肽是 30 至 40 个氨基酸长

度，包括例如，33，34，35，36 和 37 个氨基酸长度的肽。优选地，所述肽没有内部胰蛋白酶切割位点。

[0095] 在一种实施方案中，所述肽选自从心房钠尿肽 (ANP) 衍生的肽，如果需要的话，作为修改，除去内部的肽切割位点。更优选地，所述肽是不具有内部肽切割位点，并优选没有内部胰蛋白酶切割位点（即除 C- 末端精氨酸之外，所有的精氨酸和赖氨酸残基已经被瞄定进行氨基酸取代，优选通过用组氨酸 (H) 的保守性氨基酸取代）的血管扩张肽 (VSDL) 或其变体或其修饰的肽，或修饰的钾尿肽 (KP)。

[0096] 因此，在一种实施方案中，所述肽是修饰的 KP 肽（没有内部胰蛋白酶切割位点），并且优选地，是包含下列氨基酸序列或由下列氨基酸序列（衍生于人心房利钠激素原 (ANP 原）的残基 79-98) 组成的修饰的人 KP:

[0097] Ser-Ser-Asp-X_a-Ser-Ala-Leu-Leu-X_b-Ser-X_c-Leu-X_d-Ala-Leu-Leu-Thr-Ala-Pro-Arg(SEQ ID NO:20)；

[0098] 其中 X 是除了精氨酸或赖氨酸的任何氨基酸。

[0099] 在一个优选的实施方案中，肽是 VSDL 肽或其变体或其修饰的肽。如本文所用，VSDL 的变体可以是天然变体。另外变体 VSDL 可以是改性 VSDL 肽。修饰的 VSDL 肽不仅包括变体肽（即非天然变体），而且还有天然 VSDL 肽的衍生物和模拟物，其中包括不导致生物学活性的任何显著下降或变化（例如根据美国专利 No 5,691,310（这里引作参考）中 DLVesely 描述的体外血管扩张试验（使用主动脉条）测定的，或者根据 DL Vesely(Vesely 等 .1987) 描述的用于增强的环 GMP 水平的试验测定的，其从中衍生的 VSDL 肽的生物学活性下降或变化不超过 10％）的该氨基酸序列中小的改变，并且不包括或引入另一个被所述裂解剂识别的肽切割位点（优选没有另外的胰蛋白酶切割位点）。这些变化可包括保守的氨基酸取代，如上面表 3 中详述的。VSDL 肽内合适的氨基酸取代的一些具体的例子可包括 Pro → Gln（具体地在 proANP 的位置 41；即 VSDL 肽的位置 10)，Thr → Ala（具体地在 proANP 的位置 59；即 VSDL 肽的位置 28)，Glu → Asp（具体地在 proANP 的位置 61；即 VSDL 肽的位置 30)，和 Ser → Asn（具体地在 proANP 的位置 63；即 VSDL 肽的位置 32)。

[0100] 更优选地，所述肽是具有天然序列的 VSDL 肽，如人 VSDL 的天然氨基酸序列（从人 proANP 的残基 31-67 衍生的如下）：

[0101] Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg

[0102] (SEQ ID NO:2).

[0103] 其它合适的天然 VSDL 肽可以包括：

[0104] 婆罗洲猩猩 (Pongo pygmaeus)（普通猩猩）

[0105] Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Gln-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(SEQ ID NO:8)；

[0106] 和猫属 (Felis catus)

[0107] Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Gln-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Ala-Gly-Glu-Val-Asn-Pro-Ala-Gln-

Arg(SEQ ID NO:9).

[0108] 最优选地，所述肽是包含根据 SEQ ID NO2 的氨基酸序列或由根据 SEQ ID NO2 的氨基酸序列构成的 VSDL 肽。然而，在其它实施方案中，VSDL 肽或者可以包含根据 SEQ ID NO :8 或 9 的氨基酸序列或由根据 SEQ ID NO :8 或 9 的氨基酸序列构成。

[0109] 优选地，VSDL 肽（或其变体或其修饰的肽）包含 C- 端精氨酸残基。例如，这里所描述的天然 VSDL 肽包含的 C- 末端精氨酸残基，但有没有内部的精氨酸或赖氨酸残基。因此，C- 末端精氨酸残基可适合作为 C- 末端位置的肽切割位点。即，在精氨酸残基的羧基端切割的肽裂解剂，如胰蛋白酶和胰蛋白酶样酶，精确地在 C- 末端从融合多肽（例如，从 VSDL 肽或其变体或其修饰的相邻拷贝肽）裂解 VSDL 肽（或其变体或其修饰的肽）。

[0110] 然而，人 VSDL 肽（或其变体或其修饰的肽），除了缺失内部的精氨酸残基，另外缺失天冬氨酸，半胱氨酸，组氨酸，异亮氨酸，赖氨酸，甲硫氨酸，苯丙氨酸或酪氨酸残基。因此，有可能利用肽切割位点，其在根据本发明的这些位点中的一个处切割，也就是，例如，就 VSDL 肽（或其变体或其修饰的肽）来说在位于 N- 末端的肽切割位点，如下面所述。或者或另外，可以用另一个天门冬氨酸，半胱氨酸，组氨酸，异亮氨酸，赖氨酸，甲硫氨酸，苯丙氨酸或酪氨酸残基取代 VSDL 肽（或其变体或其修饰的肽）的 C- 末端精氨酸残基，使得所述肽能通过这些残基中的一个处裂解的裂解剂从融合多肽裂解。本领域的技术人员理解，这个实施方案作为选择可被应用到其它合适的肽，如本文中所描述的。

[0111] 在一种实施方案中，融合多肽包含位于 VSDL 肽（或其变体或其修饰的肽）N- 末端的一个或多个氨基酸，包括一个肽切割位点。在本发明的一个优选实施方案中，所述融合多肽包含具有 C- 末端肽切割位点的先导融合配偶体，其向 VSDL 肽（或其变体或其修饰的肽）提供了一个 N- 末端取位的肽切割位点。

[0112] 在一种实施方案中，VSDL 肽（或其变体或其修饰的肽）”侧翼有胰蛋白酶切割位点”。这要理解为其含义是，胰蛋白酶切割位点位于所述融合多肽内，在 VSDL 肽或其变体或修饰肽的两端（即，一个在 VSDL 肽的 N- 末端，一个在 C- 末端），使得该 VSDL 肽（或其变体或其修饰的肽）能在胰蛋白酶切割位点用酶消化时被从融合多肽释放。对本领域技术人员理解，在胰蛋白酶切割位点的精氨酸或赖氨酸残基的羧基端上发生裂解时，释放时从 VSDL 肽（或其变体或其修饰的肽）排除精氨酸或赖氨酸残基，而 C- 末端取位的胰蛋白酶切割位点意思是精氨酸或赖氨酸在释放时将包括 VSDL 肽（或其变体或其修饰的肽）中（即释放时精氨酸或赖氨酸残基会形成 VSDL 肽的 C- 末端残基）。

[0113] 优选地，从融合多肽释放的 VSDL 肽（或其变体或其修饰的肽）是完整的（即，全长 VSDL 肽）；然而，要明白，该融合多肽可通过改变胰蛋白酶切割位点的位置被设计来释放 VSDL 肽（或其变体或其修饰的肽）的截短或延伸的变体。例如，本发明包括融合多肽在某些情况下的表达，其中位于 N- 末端的胰蛋白酶切割位点位于 VSDL 肽（或其变体或其修饰的肽）的更上游（即在氨基方向）的 VSDL 肽的表达（或其变体或其修饰的肽），酶促裂解时产生 N- 末端延伸的 VSDL 肽。类似地，C- 末端延伸的 VSDL 肽（或其变体或其修饰的肽）能被设计在 C- 末端胰蛋白酶切割位点之前包括附加的氨基酸。或者，如本领域技术人员理解的，通过缺失氨基酸能设计截短的 VSDL 肽（或其变体或其修饰的肽）。

[0114] 在一种实施方案中，融合多肽包含 VSDL 肽的串联拷贝（即 VSDL 肽的多联体重复）或其变体或其修饰的肽。

[0115] 优选地，所述 VSDL 肽的每个拷贝（或其变体或其修饰的肽）具有肽切割位点（例如胰蛋白酶切割位点），在每个 VSDL 肽（或其变体或其修饰的肽）的 C- 末端，以使得肽的切割位点还适合作为多联体排列中 VSDL 肽（或其变体或其修饰的肽）之后的位于 N- 末端的肽的切割位点（除了融合多肽中最后 VSDL 肽），使得每个 VSDL 肽（或其变体或其修饰的肽）侧翼是根据本发明的肽切割位点。这种排列有利地导致各 VSDL 肽（或其变体或其修饰的肽）在用裂解剂裂解时从融合多肽分别释放。

[0116] 在融合多肽中可以串联排列任何数目的 VSDL 肽（或其变体或其修饰的肽）的拷贝，前提是融合多肽被表达至令人满意的水平。在一种实施方案中，以至少 0.01g/L 表达融合多肽，但优选地，以至少 0.5g/L 表达融合多肽，更优选地，以至少 1g/L，更优选至少 2g/L，更优选 5g/L，或更优选 10g/L。优选地，所述融合多肽包含两个至 20（或更多）的 VSDL 肽（或其变体或其修饰的肽）的串联拷贝。

[0117] 例如，融合多肽可以包含二，三，四，五，六，七，八，九，十，十一，十二，十三，十四，十五，十六，十七，十八，十九或二十个 VSDL 肽（或其变体或其修饰的肽）的串联拷贝。在一种实施方案中，所述融合多肽包含 2 至 10 个 VSDL 肽（或其变体或其修饰的肽）的多联体肽。在一个优选实施方案中，所述融合多肽包含 3 个 VSDL 肽（或其变体或其修饰的肽）的多联体肽。在另一种实施方案中，所述融合多肽包含肽的 10 个 VSDL 肽（或其变体或其修饰的肽）的串联拷贝。在一些实施方案中，VSDL 肽（或其变体或其修饰的肽）的不同变体可以被串联表达。

[0118] 包含串联 VSDL 肽（和编码所述融合多肽的多核苷酸构建体）或其变体或其修饰的肽的融合蛋白另外在这里被指作，例如，当存在三个多联体 VSDL 肽（或其变体或其修饰的肽）时，”3- 链节”，当存在六个多联体 VSDL 肽（或其变体或其修饰的肽）时，”6- 链节”，等。同样，包含单一 VSDL 肽或其变体或其修饰的肽的融合多肽（和编码所述融合多肽的多核苷酸构建体）另外在这里被指作”1- 链节”。

[0119] 在其中肽包括 C- 末端肽切割位点的本发明的实施方案中，本发明意在包括在其范围内这种情况，其中一系列串联肽中的第一个肽不从融合蛋白释放，因为它不具有 N- 末端肽切割位点。同样，在其中肽包括 N- 末端肽切割位点的本发明的实施方案中，本发明意在包括在其范围内这种情况，其中一系列串联肽中的最后一个肽不从融合蛋白释放，因为它不具有 C- 末端肽切割位点。但是，优选地，融合多肽进一步包括肽切割位点（即就包括 C- 末端肽切割位点的肽的串联拷贝来说位于 N- 末端，或者可替换的，就包括 N- 末端肽切割位点的肽的串联拷贝来说位于 C- 末端）使得肽的各个串联拷贝能从融合多肽被释放。

[0120] 在一种实施方案中，融合多肽可以包含融合配偶体。有利地，融合配偶体还可以提供具有提高的溶解度，稳定性，活性（例如，通过减少错折叠）和 / 或提高表达产量的融合多肽。优选地，融合配偶体提供与相邻肽的一端邻接的肽切割位点，使得，当肽包括或者在其另一端提供一个肽切割位点时，所述肽侧翼是根据本发明的肽切割位点。在一种实施方案中，融合多肽可以包含与所述肽相比不同的肽切割位点。在这种情况下，可能需要两种裂解剂从融合多肽释放所述肽。

[0121] 在一种实施方案中，融合多肽包含一个示踪融合配偶体，即，示踪融合配偶体位于所述肽或肽的串联拷贝的 C- 末端。示踪融合配偶体可以包括对肽提供位于 C- 末端的肽切割位点的 N- 末端肽切割位点（例如，对包括位于 N- 末端的肽切割位点的肽提供，使得肽的

侧翼是根据本发明的肽切割位点）。因此，在一种实施方案中，示踪融合肽包含一个 N- 末端肽切割位点。本领域技术人员会认识到，示踪融合配偶体的这样的 N- 末端肽切割位点可适合作为先导肽的位于 C- 末端的肽切割位点（即，位于示踪融合配偶体的 N- 末端）。

[0122] 在另一种实施方案中，融合多肽包含先导融合配偶体，先导融合配偶体位于所述肽或肽的串联拷贝的 N- 末端。先导融合配偶体可以包括对肽提供位于 N- 末端的肽切割位点的 C- 末端肽切割位点（例如，对包括位于 C- 末端的肽切割位点的肽提供，使得肽的侧翼是根据本发明的肽切割位点）。因此，在一种实施方案中，先导融合肽包含一个 C- 末端肽切割位点。本领域技术人员会认识到，先导融合配偶体的这样的 C- 末端肽切割位点可适合作为后面肽的位于 N- 末端的肽切割位点（即，位于先导融合配偶体的 C- 末端）。

[0123] 本领域技术人员公知的任何合适的肽可以是融合配偶体，前提是得到的肽以足够的质量和 / 或产率被表达。已知的融合配偶体的例子包括泛素，葡萄球菌 A 蛋白，硫氧还蛋白，麦芽糖结合蛋白，谷胱甘肽 -S- 转移酶，凝乳酶原，β - 半乳糖苷酶，得自 T4 的 gp 55，等。已知的重组标签也可以提供融合配偶体，包括 His 标签，B 标签，FLAG，c-myc 标签，蛋白质 C 标签，S- 标签，甲硫氨酸，等。融合配偶体能任选地包含附加的肽切割位点，导致融合配偶体在融合多肽裂解时被裂解为很多小的肽，其可以有助于期望的肽的纯化。

[0124] 在一种实施方案中，融合多肽包含从白介素 (IL)-2 例如牛 IL-2 衍生的先导融合多肽。更优选地，先导融合配偶体是从牛 IL-2 衍生的 72 氨基酸片段，这里称为 B72R，它具有下面氨基酸序列：

[0125] Met-Lys-Glu-Val-Lys-Ser-Leu-Leu-Leu-Asp-Leu-Gln-Leu-Leu-Leu-Glu-Lys-Val-Lys-Asn-Pro-Glu-Asn-Leu-Lys-Leu-Ser-Arg-Met-His-Thr-Phe-Asp-Phe-Tyr-Val-Pro-Lys-Val-Asn-Ala-Thr-Glu-Leu-Lys-His-Leu-Lys-Ala-Leu-Leu-Glu-Glu-Leu-Lys-Leu-Leu-Glu-Glu-Val-Leu-Asn-Leu-Ala-Pro-Ser-Lys-Asn-Leu-Asn-Val-Asp-Arg(SEQ ID NO:3).

[0126] 因此，在一种实施方案中，先导融合配偶体是 B72R。

[0127] 在一种实施方案中，B72R 融合配偶体的倒数第二个天冬氨酸残基被取代为谷氨酰胺残基（这里称为 B72(Q)R），具有下列氨基酸序列：

[0128] Met-Lys-Glu-Val-Lys-Ser-Leu-Leu-Leu-Asp-Leu-Gln-Leu-Leu-Leu-Glu-Lys-Val-Lys-Asn-Pro-Glu-Asn-Leu-Lys-Leu-Ser-Arg-Met-His-Thr-Phe-Asp-Phe-Tyr-Val-Pro-Lys-Val-Asn-Ala-Thr-Glu-Leu-Lys-His-Leu-Lys-Ala-Leu-Leu-Glu-Glu-Leu-Lys-Leu-Leu-Glu-Glu-Val-Leu-Asn-Leu-Ala-Pro-Ser-Lys-Asn-Leu-Asn-Val-Gln-Arg(SEQ ID NO:10).

[0129] 因此，在一种实施方案中，先导融合配偶体是 B72(Q)R。

[0130] 在优选的实施方案中，先导融合配偶体直接在肽之前，或者可选地，所述肽的第一串联拷贝之前。在与多联体肽相关的优选实施方案中，第一个肽直接在第二个肽之前（等等）。

[0131] 例如，包含 B72R 先导融合肽和 VSDL 肽的融合多肽的氨基酸序列可以是如下：

[0132] Met-Lys-Glu-Val-Lys-Ser-Leu-Leu-Leu-Asp-Leu-Gln-Leu-Leu-Leu-Glu-Lys-Val-Lys-Asn-Pro-Glu-Asn-Leu-Lys-Leu-Ser-Arg-Met-His-Thr-Phe-Asp-Phe-Tyr-Val-Pro-Lys-Val-Asn-Ala-Thr-Glu-Leu-Lys-His-Leu-Lys-Ala-Leu-Leu-Glu-Glu-Leu-Lys-Leu-

Leu-Glu-Glu-Val-Leu-Asn-Leu-Ala-Pro-Ser-Lys-Asn-Leu-Asn-Val-Asp-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)(SEQ ID NO:11).

[0133] 箭头指示胰蛋白酶切割位点的位置。

[0134] 在另一个实施例中，包含 B72R 先导融合肽和三个串联 VSDL 肽的融合多肽的氨基酸序列可以是如下：

[0135] Met-Lys-Glu-Val-Lys-Ser-Leu-Leu-Leu-Asp-Leu-Gln-Leu-Leu-Leu-Glu-Lys-Val-Lys-Asn-Pro-Glu-Asn-Leu-Lys-Leu-Ser-Arg-Met-His-Thr-Phe-Asp-Phe-Tyr-Val-Pro-Lys-Val-Asn-Ala-Thr-Glu-Leu-Lys-His-Leu-Lys-Ala-Leu-Leu-Glu-Glu-Leu-Lys-Leu-Leu-Glu-Glu-Val-Leu-Asn-Leu-Ala-Pro-Ser-Lys-Asn-Leu-Asn-Val-Asp-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)(SEQ ID NO:12).

[0136] 本领域技术人员明白根据这里所描述的，融合多肽中可以包括多个多联体 VSDL 肽。例如，包含 B72R 先导融合肽和 10 个串联 VSDL 肽的融合多肽的氨基酸序列如下：

[0137] Met-Lys-Glu-Val-Lys-Ser-Leu-Leu-Leu-Asp-Leu-Gln-Leu-Leu-Leu-Glu-Lys-Val-Lys-Asn-Pro-Glu-Asn-Leu-Lys-Leu-Ser-Arg-Met-His-Thr-Phe-Asp-Phe-Tyr-Val-Pro-Lys-Val-Asn-Ala-Thr-Glu-Leu-Lys-His-Leu-Lys-Ala-Leu-Leu-Glu-Glu-Leu-Lys-Leu-Leu-Glu-Glu-Val-Leu-Asn-Leu-Ala-Pro-Ser-Lys-Asn-Leu-Asn-Val-Asp-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-V

al-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(↓)Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg(SEQ ID NO:13)，

[0138] 等等。

[0139] 因此，在一种实施方案中，单一的肽切割位点可以位于先导融合配偶体和下面的肽之间（或者，在包含多联体肽的融合多肽情况下，在先导融合配偶体和所述第一肽之间，以及在肽的每个串联拷贝之间）。但是，要理解融合多肽有可能在肽的每个串联拷贝之间包含短连接序列，其中连接序列侧翼是肽切割位点，使得在用根据本发明的裂解剂裂解时从融合多肽释放完整的肽。

[0140] 或者或另外，要理解融合多肽有可能在各个肽之间包含短连接序列，其中用裂解剂裂解之后，连接序列保持连接所述肽的一个末端，使得以接头延伸形式从融合多肽释放所述肽，然后通过使用另一个肽切割位点的第二种裂解剂能从所述肽任选地切割连接序列，使得从连接序列释放肽。

[0141] 使用编码本发明的融合多肽的表达构建体和作为重组宿主细胞的大肠杆菌，本发明人能获得可溶性融合多肽的合适的高水平表达，包括根据本发明带有一个 VSDL 肽和三个，六个，八个和十个 VSDL 肽的串联拷贝的融合多肽。此外，本发明人已经证明获得的融合多肽能被完全消化得到适合的高浓度的释放的 VSDL 肽。

[0142] 在本发明的一种实施方案中，释放出的肽被纯化。

[0143] 该肽的纯化能有利地减少污染物（例如，细胞组分，包括宿主细胞蛋白质，核酸，内毒素），达到适合于治疗用途的可接受水平。

[0144] 用本领域技术人员公知的任何合适的技术能纯化肽。在一种实施方案中，在用裂解剂切割之前对融合多肽进行纯化。

[0145] 在一种实施方案中，在用裂解剂对融合多肽切割之后对肽进行纯化。如果要临床使用肽，如果纯化方法不要求在高浓度无毒、挥发性或易燃溶剂（例如甲醇）中洗脱肽，则其是有益的。

[0146] 在一种实施方案中，所述方法进一步包括纯化从切割的融合多肽释放的肽的步骤。本领域技术人员知道大范围纯化技术可以适合纯化释放的肽，例如，离子交换技术和其它分离技术。适合纯化肽的离子交换技术包括商业上可获得树脂和膜，例如 ChromaSorb 膜 (Merck-Millipore，比勒瑞卡 (Billerica)，MA，美国），Sartobind 膜 (Sartorious AG，哥廷根，德国），Mustang 膜 (Pall Corporation，华盛顿港口，纽约，美国），CIM QA Monolith 膜 (BIA Separations GmbH，非拉赫，奥地利），反相树脂，CaptoMMC（混合模式树脂；GE Healthcare，瓦克夏，WI，美国），羟基磷灰石（混合模式树脂；BioRad Laboratories，Inc，Hercules CA，美国），Q 和 DEAE 巨石（离子交换树脂；BIA Separations)，Macroprep HQ 和 DEAE（阴离子交换树脂；BioRad)，POROS HQ，PI and D（阴离子交换树脂；Biosystems Inc，马萨诸赛州弗雷明翰，美国），Nuvia Q（阴离子交换树脂；BioRad)，UNOsphere Q（阴离子交换树脂；BioRad)，Q Sepharose FF 和 BB（阴离子交换树脂；GE Healthcare)，SP Sepharose FF（阳离子交换树脂；GEHealthcare)，Fractogel TMAE（阴离子交换树脂；

Merck-Millipore),MEP Hypercel(混合模式树脂;Pall Corporation),等。适合纯化释放的肽的非离子交换分离技术的技术包括Blue Sepharose(假亲和树脂;GE Healthcare),很多树脂来源于Phenomenex或Vydac,等。

[0147] 在VSDL肽(或其变体或其修饰的肽)的情况下,在一种实施方案中,用反相色谱法从被消化的融合多肽和其它宿主细胞杂质纯化被释放的VSDL肽(或其变体或其修饰的肽)。例如,反相色谱法可利用不溶性聚苯乙烯二乙烯基苯聚合物树脂,例如Amberchrom CG300C或Amberchrom XT柱(Rohm and Haas)。在一些实施方案中,用40%至20%异丙醇(体积/体积)从这样的树脂能洗脱被释放的VSDL肽(或其变体或其修饰的肽)。优选地,所述洗脱在乙酸存在下进行。例如,在0.5%乙酸存在下用30%异丙醇可以洗脱VSDL肽(或其变体或其修饰的肽)。在可替代实施方案中,使用10～20%异丙醇能从这样的树脂洗脱被释放的VSDL肽(或其变体或其修饰的肽)。优选地,所述洗脱发生在弱碱性pH值下。例如,例如在25mM Tris pH 8存在下,在缓冲到pH 8的15%异丙醇中可以洗脱VSDL肽(或其变体或其修饰的肽)。有利地,这样低浓度的异丙醇在大规模重组蛋白质制备中被认为是可以接受的。

[0148] 在另一种实施方案中,使用阴离子交换色谱从被消化融合多肽纯化被释放的VSDL肽(或其变体或其修饰的肽)。例如,阴离子交换层析可以是利用交联的聚甲基丙烯酸酯树脂例如Fractogel TMAE柱(Merck)的强阴离子交换层析。在一些实施方案中,用醋酸钠能以高纯度从这样的树脂洗脱VSDL肽(或其变体或其修饰的肽)。优选地,所述洗脱发生在酸性pH下(例如,pH 5.0,pH4.0或pH 3.5的10mM醋酸钠)。在其它实施方案中,用低浓度盐酸(例如10mM HCl)能以高纯度从这样的树脂洗脱被释放的VSDL肽(或其变体或其修饰的肽)。

[0149] 在另一种实施方案中,被释放的VSDL肽(或其变体或其修饰的肽体)是用反相色谱法从经消化的融合多肽中纯化。使用反相色谱法从被消化融合多肽纯化被释放的VSDL肽(或其变体或其修饰的肽)。例如,反相色谱法可以利用不溶性聚苯乙烯二乙烯基苯聚合物树脂,例如Amberchrom CG300C或Amberchrom XT柱(Rohm and Haas)。在一种实施方案中,使用10～20%异丙醇能以高产率从这样的柱子洗脱VSDL肽(或其变体或其修饰的肽)。优选地,所述洗脱发生在弱碱性pH值下。例如,例如在25mM Tris pH 8存在下,在缓冲到pH 8的15%异丙醇中可以洗脱VSDL肽(或其变体或其修饰的肽)。

[0150] 另外或可替代地,使用Chromasorb膜并且用10mM盐酸洗脱能纯化被释放的VSDL肽(或其变体或其修饰的肽)。

[0151] 在另一种实施方案中,最初利用反相色谱法,例如使用不溶性聚苯乙烯二乙烯基苯聚合物树脂从被消化的融合多肽纯化被释放的VSDL肽(或其变体或其修饰的肽),然后可以利用强阴离子交换色谱法(例如,使用交联聚甲基丙烯酸酯树脂)进一步纯化。

[0152] 在一种实施方案中,在选自异丙醇,丁醇和乙腈的有机溶剂中可以洗脱所述肽。

[0153] 本领域技术人员明白,用裂解剂对融合多肽切割之后,其它纯化方法可适合用于纯化VSDL肽(或其变体或其修饰的肽)。

[0154] 在一种实施方案中,纯化的肽可以被冻干。冻干的肽随后可以重新构成。

[0155] 在一种实施方案中,纯化的多肽可以用药物或兽医学上可接受的赋形剂或载体配制。

[0156] 在第三方面，本发明提供了一种制备重组血管扩张（VSDL）肽（或其变体或其修饰的肽）的方法，所述方法包括以下步骤：

[0157] 表达包含至少一种 VSDL 肽（或其变体或其修饰的肽）融合多肽，其中所述肽侧翼是一个或多个胰蛋白酶切割位点，并且其中所述肽特征在于另外不存在胰蛋白酶切割位点；和

[0158] 用在胰蛋白酶切割位点切割的裂解剂在胰蛋白酶切割位点切割融合多肽；

[0159] 其中切割融合多肽的步骤从融合多肽释放所述肽。

[0160] 应当理解，在这个第三方面中，所述 VSDL 肽（或其变体或其修饰的肽），融合多肽，和胰蛋白酶切割位点中的每个可以和本发明的第一或第二方面如上所述的一样。例如，融合多肽可以包含先导融合配偶体；融合多肽可以包含 VSDL 肽（或其变体或其修饰的肽）的串联拷贝；和 / 或 VSDL 肽（或其变体或其修饰的肽）可以具有如上所述序列；等等。

[0161] 在第四方面，本发明提供了一种制备重组血管扩张（VSDL）肽（或其变体或其修饰的肽）的方法，所述方法包括以下步骤：

[0162] 表达包含一个先导融合配偶体和至少一个 VSDL 肽（或其变体或其修饰的肽）的融合多肽，其中所述肽侧翼是一个或多个胰蛋白酶切割位点；

[0163] 用胰蛋白酶或胰蛋白酶样酶切割融合多肽；

[0164] 将切割的融合多肽的 pH 调节至低于 5；

[0165] 利用阴离子交换技术从消化的融合多肽纯化所述肽。

[0166] 要明白，在这个第四方面中，所述 VSDL 肽（或其变体或其修饰的肽），融合多肽，先导融合配偶体，胰蛋白酶切割位点，胰蛋白酶或胰蛋白酶样酶等中的每个可以和本发明的第一或第二方面如上所述的一样。例如，融合多肽可以包含 VSDL 肽（或其变体或其修饰的肽）的串联拷贝；和 / 或 VSDL 肽（或其变体或其修饰的肽）可以具有如上所述序列；等等。

[0167] 在第五方面，本发明提供了一种表达构建体，其包含本发明第一至第四方面任一方面中定义的编码融合多肽的多核苷酸序列。

[0168] 优选地，所述表达构建体包含与至少一个编码具有 C- 末端胰蛋白酶切割位点的 VSDL 肽（或其变体或其修饰的肽）的核苷酸序列操作融合的编码具有 C- 末端胰蛋白酶切割位点的先导融合配偶体的核苷酸序列。

[0169] 如本文所用，本领域技术人员理解术语”操作融合”意思是指在不存在任何中止密码子存在下所述核苷酸序列在同一阅读框内编码整个融合多肽，使得融合多肽的肽的每一个被翻译。

[0170] 在一种实施方案中，编码具有 C- 末端的胰蛋白酶切割位点的先导融合配偶体的核苷酸序列可以操作融合到编码具有 C- 末端的胰蛋白酶切割位点的 VSDL 肽（或其变体或其修饰的肽）的核苷酸序列。在另一种实施方案中，编码具有 C- 末端胰蛋白酶切割位点的先导融合配偶体的核苷酸序列可以操作融合任何数目的核苷酸序列，所述任何数目的核苷酸序列的每一个编码具有 C- 末端的胰蛋白酶切割位点的 VSDL 肽（或其变体或其修饰的肽）。优选地，编码具有 C- 末端胰蛋白酶切割位点的先导融合配偶体的核苷酸序列可以操作融合 2 — 20 个（或更多）核苷酸序列，所述核苷酸序列的每一个编码具有 C- 末端的胰蛋白酶切割位点的 VSDL 肽（或其变体或其修饰的肽）。也就是

说，所述核苷酸序列可以优选地编码融合多肽，所述融合多肽包含具有 C- 末端的胰蛋白酶切割位点的先导融合配偶体，所述先导融合配偶体与具有 C- 末端的胰蛋白酶切割位点的 VSDL 肽（或其变体或其修饰的肽）的二，三，四，五，六，七，八，九，十，十一，十二，十三，十四，十五，十六，十七，十八，十九，或二十个串联拷贝操作融合。

[0171] 优选地，C- 末端胰蛋白酶切割位点是 C- 末端精氨酸残基。

[0172] 在一种实施方案中，所述核苷酸序列编码一种融合多肽，其包含 SEQ ID NO :11，SEQ ID NO :12 和 SEQ ID NO :13 的氨基酸序列或由选自 SEQ ID NO :11，SEQ ID NO :12 和 SEQ ID NO :13 构成的组的氨基酸序列组成。

[0173] 表达构建体可以位于表达质粒内，如本领域技术人员所熟知的。

[0174] 在第六方面，本发明提供了一种用于重组表达肽的包含第五方面的表达构建体的宿主细胞。优选地，所述肽是一种 VSDL 肽（或其变体或其修饰的肽）。能制备重组肽的宽范围的表达系统是本领域技术人员公知的。此外，本发明的宿主细胞可以是任何合适的宿主细胞，如本领域的技术人员所理解的。宿主细胞可以包括细菌细胞（例如大肠杆菌 (Escherichia coli)，链霉菌属 (Streptomyces sp.)，芽孢杆菌属 (Bacillus sp.)，假单胞菌属 (Pseudomonas sp.)，葡萄球菌属 (Staphylococcus sp.)（例如鼠伤寒沙门氏菌 (S.typhimurium) 等）等），真菌细胞，如酵母细胞（例如酿酒酵母 (Saccharomyces cerevisiae)，巴斯德毕赤酵母 (Pichia pastoris) 等），霉菌或丝状真菌（例如曲霉 (Aspergillus sp.)），昆虫细胞（例如果蝇 S2 和草地夜蛾 (Spodoptera frugiperda)（例如夜蛾 Sf9 细胞），并包括杆状病毒辅助系统），哺乳动物细胞（适合的大量细胞系包括中国仓鼠卵巢 (CHO) 细胞，猴肾 (COS) 细胞，人胚胎肾 (HEK) 细胞，幼仓鼠肾 (BHK) 细胞，鼠黑素瘤细胞等），和植物细胞。重组多肽或者也能在转基因植物和动物中表达。因此，宿主细胞可以是合适的植物或动物细胞。优选地，宿主细胞是大肠杆菌细胞。

[0175] 在进一步的方面，本发明提供了一种重组融合多肽，其包含连接于具有 C- 末端胰蛋白酶切割位点的肽的串联拷贝（例如肽的 2 到 20 个或更多串联拷贝）的先导融合配偶体。优选地，所述先导融合配偶体具有 C- 末端胰蛋白酶切割位点。优选地，所述肽没有内部胰蛋白酶切割位点。优选地，所述肽是 VSDL 肽或者其变体或者其修饰的肽，更优选地，是包括 SEQ ID NO:2 的氨基酸序列或由根据 SEQ ID NO:2 的氨基酸序列构成的 VSDL 肽。

[0176] 在进一步的方面，本发明提供了一种重组融合多肽，其包含或由具有 C- 末端胰蛋白酶切割位点的肽的串联拷贝（例如肽的 2 到 20 个或更多串联拷贝）构成。优选地，所述先导融合配偶体具有 C- 末端胰蛋白酶切割位点。优选地，所述肽没有内部胰蛋白酶切割位点。优选地，所述肽是 VSDL 肽或者其变体或者其修饰的肽，更优选地，包括或由根据 SEQ ID NO:2 的氨基酸序列构成的 VSDL 肽。优选地，所述多肽包含所述肽的 3 — 10 个串联拷贝。优选地，所述肽没有内部胰蛋白酶切割位点。优选地，所述肽是 VSDL 肽或者其变体或者其修饰的肽，更优选地，是包含 SEQ ID NO:2 的氨基酸序列或由根据 SEQ ID NO:2 的氨基酸序列构成的 VSDL 肽。

[0177] 本发明下文通过参考以下非限制性实施例和附图进行描述。

[0178] 实施例

[0179] 实施例 1 在大肠杆菌中制备 VSDL

[0180] 人 VSDL（在实施例中指定 VSDL）的氨基酸序列是 ：

[0181] Glu-Val-Val-Pro-Pro-Gln-Val-Leu-Ser-Glu-Pro-Asn-Glu-Glu-Ala-Gly-Ala-Ala-Leu-Ser-Pro-Leu-Pro-Glu-Val-Pro-Pro-Trp-Thr-Gly-Glu-Val-Ser-Pro-Ala-Gln-Arg

[0182] (SEQ ID NO:2).

[0183] 在该实施例中，利用融合多肽途径，以试图尝试多肽序列中不存在内部精氨酸 (R) 和赖氨酸 (K) 残基，以及便利地提供胰蛋白酶切割位点的存在 C- 末端精氨酸 (R) 的策略，对在大肠杆菌中重组表达 VSDL 进行研究 (图 1)。

[0184] 表达构建体的制备

[0185] 首先，设计两个融合多肽 :

[0186] (i) 与指定为 B72R 的 73 氨基酸融合配偶体融合的单一 VSDL 序列 (指定为 B72R-VSDL 并且这里还称作”1- 链节”)，和

[0187] (ii) 与 B72R 融合的三个 VSDL 重复 (指定为 B72R-$(VSDL)_3$ 并且这里还称作”3- 链节”)。这些融合多肽的氨基酸序列示于图 1。

[0188] 然后制备第二轮基因融合体，其中用谷氨酰胺残基 (Q) 置换融合配偶体的倒数第二个天冬氨酸残基 (D)(指定为 B72(Q)R)，并且编码的融合多肽包括 VSDL 肽序列的 3,6,8 和 10 个串联拷贝。预期天冬氨酸残基的取代可能导致融合配偶体 /VSDL 连接处改进的切割。

[0189] 以 pUC57 载体亚克隆 (GenScript USA Inc，新泽西州皮斯卡塔维，美国) 获得 VSDL 融合多肽基因。基因融合体的核苷酸序列如图 2 所示。通过用限制酶 NdeI 和 HindIII(Roche Diagnostics Corporation，印第安纳州波利斯，美国) 消化，从各 pUC57 亚克隆切离 VSDL 基因融合体。将质粒消化物脱盐，然后通过在 T4DNA 连接酶 (Roche) 存在下在 16℃下温育过夜直接连接 pBRE508 表达载体 (Hospira Adelaide Pty Ltd，Thebarton SA，澳大利亚)。连接反应用来转化 XL1-Blue 电子感受态细胞 (Agilent Technologies Inc，加利福尼亚州圣克拉拉，美国)，并且在含有卡那霉素 (30μg/mL)VA 平板上选择阳性转化体。使用 QIAprep Spin Miniprep 试剂盒 (Qiagen NV，Venlo，荷兰)，根据厂商说明，从各个转化体制备质粒 DNA。然后用 NdeI/HindIII 消化纯化的质粒，来鉴定成功的候选克隆。对于不同 VSDL 基因融合体表现出预期的泳带图案的质粒克隆被指定为 pBRE601 至 pBRE606。图 3a-f 描述 VSDL 表达构建体的图。

[0190] 小规模表达研究

[0191] 以 10mL 摇瓶规模进行最初的表达实验。将 VSDL 表达构建体 pBRE601(1- 链节) 和 pBRE602(3- 链节) 转化到 E.coli 宿主株 BR067(Hospira Adelaide ;MM294OmpT 缺失) 中并且在含有卡那霉素 (30μg/mL)VA 平板上选择转化体。

[0192] 通过将单一菌落从转化板无菌转移到有卡那霉素 (30μg/mL) 的 3mL C2- 定义培养基培养物中建立过夜 3mL 培养物。通过旋转转盘搅动下将培养物在 37℃温育 16 小时。通过将 0.875mL 的各培养物与 0.125mL 80％丙三醇混合来制备丙三醇原种，接着在 -80℃下贮存。

[0193] 关于表达实验，将过夜培养物 (1mL) 在含有卡那霉素 (30μg/mL) 的 10mL C2- 定义培养基中无菌继代培养。这些培养物在摇动下在 37℃温育 2.5 小时，然后用 IPTG(0.25mM 终浓度) 诱导。进一步温育 3 小时后，制备 SDS-PAGE 样品。对分级的 BR067(pBRE601) 和

BR067(pBRE602)培养物样品的SDS-PAGE分析（图4）显示BR72R-VSDL和B72R-$(VSDL)_3$融合多肽高水平可溶性表达。

[0194] 接着，用编码包含3,6,8和10个VSDL串联拷贝的B72(Q)R融合多肽的构建体进行相等表达实验。如图5所示，培养物也证明了融合多肽的高水平可溶性表达。但是，注意到，因为VSDL序列拷贝数增加，融合多肽表现出异常移动行为（表4）。这是VSDL肽序列高负电荷性质最可能的结果。

[0195] 表4.VSDL融合多肽预测和表观分子量

[0196]

表达构建体	VSDL拷贝	预测 M_r(kDa)	表观 M_r(kDa)
pBRE601	1	12.7	17
pBRE602/pBRE603	3	20.3	35
pBRE604	6	31.9	53
pBRE605	8	39.7	62
pBRE606	10	47.4	80

[0197] 补料分批发酵

[0198] 在连接BioStat Qplus生物方法控制器(Sartorius)的1L反应器中进行补料分批培养。分别从细胞系BR067(pBRE601)和BR067(pBRE602)的冷冻丙三醇原种开始试验。以批量模式在37℃生长过夜后开始输送营养物。当培养物达到OD600为40-50时，通过加入IPTG对它们诱导并且继续培养5小时。用每个细胞系实现的最后OD600是～70，并且最后生物量是～80g/L。最终培养物样品的SDS-PAGE分析，如图6所示，证明在这些条件下实现BR72-VSDL和B72R-$(VSDL)_3$的高水平表达。

[0199] 通过离心回收两个细胞系的发酵生物质，并且重新悬浮于25mM Tris HCl pH 8.0。然后将收集的细胞高压分裂（在920巴下3次）并且离心(8000rpm离心15分钟）得到澄清的溶胞产物。将溶胞产物分成等份并且在-80℃储存，这种材料用于下文描述的裂解和纯化研究。

[0200] 融合多肽的裂解

[0201] 将含有B72R-VSDL和B72R-$(VSDL)_3$融合多肽的溶胞产物解冻。以25μg/mL加入测序级改进的猪胰蛋白酶(Promega V511A;Promega Corporation，威斯康星州菲奇堡，美国）并且在室温下温育过夜。将样品1:1与1.0% H_3PO_4/0.2% TFA混合并且离心去除沉淀物。通过在Waters Alliance HPLC分离模块和UV检波器(Waters Corporation，马萨诸塞州米尔福德，美国）上进行RP-HPLC分析上清液。在Phenomenex Jupiter C4,250x 2.1mm,5μm,300A柱(#302)上用Empower 2软件(Phenomenex，Inc，加州托兰斯，美国）控制进行分析。使用5,10和20μg的VSDL合成肽（即化学合成的VSDL）三个注射量构建标准曲线，其被用来定量测定裂解溶胞产物中产生的重组VSDL的量。

[0202] 在38分钟保留时间时洗脱两种融合多肽，而在28分钟保留时间时洗脱VSDL。重

组 VSDL 和合成 VSDL 保留时间是可比较的。在试验条件下消化融合多肽。前面的研究已经揭示部分消化的 (N- 末端延伸的) 变体的存在，指定为 NLNVDR-VSDL，其通过 RP-HPLC 在较后保留时间洗脱。通过减慢的梯度分析消化产物并且证实显著量的这种变体被保留 ；与 3 －链节消化产物相比，在 1 －链节消化产物中存在更大比例的变体。

[0203] 通过 RP-HPLC 测定 B72R-VSDL and B72R-$(VSDL)_3$ 消化产物中 VSDL 浓度分别是 0.58g/L 和 1.49g/L。如果消化作用已经完全，即，直到不存在 N- 末端延伸的变体 NLNVDR-VSDL，则 1 －链节和 3 －链节消化产物中 VSDL 浓度分别增加至 0.91g/L 和 1.76g/L。

[0204] 考虑到 VSDL 和融合配偶体的相对分子量，估计 1 －链节和 3 －链节溶胞产物中融合多肽浓度分别为 2.91g/L 和 3.04g/L。但是，在样品制备期间，酸性导致显著量可见沉淀物，因为估计没有消化的溶胞产物中 RP-HPLC 测定的融合多肽的浓度显著小于上述提出的数据，认为显著量的 B72R-VSDL 和 B72R-$(VSDL)_3$ 融合多肽对粗溶胞产物中存在的其他大肠杆菌产生的蛋白质旁的沉淀敏感。

[0205] 对氨基酸序列的检查揭示，当在融合配偶体 B72R 的 C- 末端的 VDR ↓ E 切割位点，和 VSDL 序列的 C- 末端的 AQR ↓ E 切割位点 (其中 ↓ 指示肽切割位点的位置) 发生裂解时，产生”可信的”VSDL(即从任一个 1 －链节和 3 －链节融合多肽)。通过比较，N- 末端延伸变体 (NLNVDR-VSDL) 的肽切割位点是 PSK ↓ N。VDR ↓ E 处较低的裂解效率可能是由于精氨酸 (R) 切割位点之前紧接着存在天冬氨酸残基 (D)，其可能减小胰蛋白酶对于该位点的亲和力。因为 VSDL 和 NLNVDR-VSDL 之间的物理化学性质相对小的差异有可能增加从终产物去除变体的纯化难度，假设融合配偶体的 C- 末端处切割位点从 VDR ↓ E 改变为 VQR ↓ E(使得融合配偶体的 C- 末端处末端和倒数第二个氨基酸残基与 VSDL 肽的 C- 末端处的那些相同) 会提高总的裂解效率，此外，VSDL 序列增加的数目也可以减少变体的出现，而随着融合配偶体与融合多肽质量比降低而进一步提高产率。

[0206] 在制备 RP-HPLC 柱 (Phenomenex Jupiter C4，250x 10mm，5μm，300 Å) 上通过加载 4mL 溶胞产物纯化 3- 链节融合多肽 (即 B72R-$(VSDL)_3$)。在 0.1％ TFA 存在下用 5-60％乙腈梯度经 60 分钟洗脱融合多肽。用离心冻干机从蛋白质蒸发乙腈和 TFA，然后将蛋白质重新配制至 0.5mg/mL，在 50mM Tris，1mM CaCl2pH 8.0 中。将 10mg 小瓶的重组牛 TrypZean® (Sigma-Aldrich) 重新悬浮至 5mg/mL，在 50％丙三醇 /50mM 乙酸中，并且在 -20℃下储存。以 1:120 质量比向融合多肽加入 TrypZean® ，并且在室温下发生消化作用过夜。10 分钟之后用 1％乙酸酸化样品以终止消化作用。通过 LC-MS 对消化的和没有消化的样品分析。样品还通过 Zorbax C3 柱 (Agilent) 进行色谱，用 1％存在下的乙腈梯度洗脱，与 Bruker HCTultra Ion-Trap 质谱仪连接。没有消化的样品得到 20091.4Da 物质，其与 3- 链节融合多肽的理论质量 (即 20091.0Da) 相当。消化的样品得到 3877.8Da 物质，其与 VSDL 的理论质量 (即 3878.3Da) 相当。检测到的其它肽相应于融合配偶体部分消化片段。显著地，没有 VSDL 多联体或者与融合配偶体连接的证据。通过 RP-HPLC 证明没有 NLNVDR-VSDL 变体，并且指示通过使用 TrypZean® 和 / 或在消化之前纯化融合多肽提高裂解效率。用溶胞产物接下来的裂解试验表明 TrypZean® 在 1:120(w/w) 裂解比例不留残余的 NLNVDR-VSDL。

[0207] 通过向固定量的 B72R-$(VSDL)_3$ 溶胞产物 (3g/L 融合多肽) 加入不同量并且在室温

下温育过夜来确定对于TrypZean®的需求。通过RP-UPLC在Shimadzu CBM-20A(Shimadzu Corporation,京都,日本)使用Restek Pinnacle DBC8,100x 2.1mm,和1.9μm柱测定裂解的VSDL的量。使用5,10和20μg的合成VSDL肽三个注射量构建标准曲线,其被用来定量测定重组VSDL。高于6mg/L TrypZean®,1:480(w/w) TrypZean®与融合多肽比例等价产生的重组VSDL开始平稳(图7).。

[0208] 纯化

[0209] 这个实验的目的是证明纯化方法:(1)得到每升发酵物大于0.25g的纯化VSDL产率,(2)证明与合成VSDL肽相当的纯化曲线,(3)将内毒素减小至低水平,(4)在10mMHCl中配制VSDL,(5)使用与大规模制备相匹配的有机溶剂量,(6)使用低压色谱步骤,和(7)可升级。

[0210] 初始方法

[0211] 在前期工作中,开展了纯化方法,其应用下面步骤:

[0212] (1)胰蛋白酶例解,(2)通过酸化沉淀宿主细胞杂质,(3)使用酸/甲醇反相分离,(4)pH调节至8,和(5)用NaCl阴离子交换。

[0213] 通过在1mL Amberchrom XT30柱(Rohm and Haas)接着在1mL Q Sepharose HP柱(GE Healthcare,Little Chalfont,Buckinghamshire,英国)上加载15mL消化的3-链节溶胞产物重复该方法。用Q柱纯化大约10mg的VSDL,每升发酵物产率是1.23g。通过RP-HPLC使用合成VSDL肽有利地比较纯化的VSDL的洗脱位置和杂质分布图,并且不存在N-末端延伸的NLNVDR-VSDL变体。但是,这种纯化方法使用100%甲醇从柱子洗脱被结合的VSDL,因此甲醇的挥发性、易燃性和毒性意味着这种纯化方法对于大规模制备是不期望的。

[0214] 有机溶剂使用的减少

[0215] 还研究了使用Amberchrom反相树脂在纯化期间减少有机溶剂的潜在性。

[0216] 在Amberchrom CG300C(Rohm and Haas)上加载消化的3-链节溶胞产物并用甲醇梯度洗脱。对于XT30,要求用100%甲醇洗脱VSDL。但是,在下面的实验中,用异丙醇代替甲醇并且发现在0.5%乙酸存在下用30%异丙醇能以高纯度洗脱VSDL。在VSDL洗脱之前用15%异丙醇能从柱子洗脱核酸污染物。没有调节pH的溶胞产物也被加载到柱子上并且在25mM Tris pH 8而不是0.5%乙酸存在下进行异丙醇洗脱。出乎意料地,发现在15%异丙醇较高pH下洗脱VSDL。显著地,15%异丙醇的可燃性显著小于30%异丙醇,而且确实类似于20%乙醇(nb.闪点是相同的-35℃),其在大规模重组蛋白制备中普遍应用。

[0217] 用1:120质量比的TrypZean®在室温下将溶胞产物消化过夜。在第二天用1M乙酸将pH调节至4.4来沉淀宿主细胞杂质。然后离心沉积沉淀并且将13mL上清液加载到1mL Amberchrom CG300C柱上,所述柱子已经用0.5%乙酸预平衡过。用0.5%乙酸,15%异丙醇/0.5%乙酸和25mM Tris pH 8(Figure 8)冲洗柱子。用15%异丙醇/25mM Tris pH 8洗脱一个大的峰。该峰中肽含量是18.5mg,其翻译为一步骤回收率为81%。接着用25mM Tris pH 8将1mL Fractogel TMAE柱子预先平衡,并且加载6.9mg这种材料。然后用25mM Tris pH 8和10mM乙酸钠pH 4.5冲洗柱子(图9)。用10mM HCl洗脱出一个大峰,肽含量是5.8mg,其证明一步回收率是84%。每升发酵物纯化产率是1.37g。

[0218] Amberchrom CG300C 和 Fractogel TMAE 对于 VSDL 的结合能力分别是～ 20mg/mL 和～ 7mg/mL。通过 RP-UPLC 分析将大的峰与合成 VSDL 相比较并且分别测得纯度是 97.0％和 93.7％。相对于合成 VSDL 肽来说对于 VSDL 没有新的明显杂质峰。此外，合成 VSDL 肽和重组 VSDL UV 曲线几乎相同（图 10）。此外，Fractogel VSDL 峰的内毒素含量，以相同方式纯化，是低的 (5EU/mg)。

[0219] 结论

[0220] 实施例中显示的结果证明以好的产率表达作为稳定融合多肽的 VSDL 的开发的细胞系，和与大规模制备相匹配的方式纯化 VSDL 的方法。

[0221] 本发明涉及融合多肽的设计，开发作为胰蛋白酶切割位点的 VSDL 的 C- 末端精氨酸残基 (R)，胰蛋白酶是很好表征的酶，作为非来自动物的重组蛋白质TrypZean®是可获得的。最初，设计两个融合多肽：与 B72R 融合的单一 VSDL 序列 (1- 链节）和与 B72R 融合的三个 VSDL 重复序列 (3- 链节)，其以可溶形式被表达，从而避免从分离的包函体纯化表达的蛋白质的需要。进一步的工作确认 VSDL 还能以高水平被表达为 6- 链节，8- 链节和 10- 链节。

[0222] 使用实验级猪胰岛素源的初步胰蛋白酶裂解研究证明融合多肽能被完全消化得到 VSDL 肽，接着用包括下面步骤的纯化方法将其回收：(1) 用 1:120 TrypZean®：融合多肽 (w/w) 消化过夜，(2)pH 调节至 4.4，(3) 离心，(4)Amberchrom CG300C 反相色谱，和 (5) Fractogel TMAE 阴离子交换色谱。通过 RP-UPLC 分析将 Fractogel 物质与合成 VSDL 制备相比较，并且测得纯化分别是 97.0％和 93.7％。重要地，对于重组 VSDL 没有新的明显杂质峰（相对于合成 VSDL 肽），并且重组和合成 VSDL 肽的 UV 曲线几乎相同。此外，纯化的重组 VSDL 的内毒素含量估计是 5EU/mg。

[0223] 因此，发现纯化方法：(1) 每升发酵物给出纯化重组 VSDL 产率 1.2-1.4g，(2) 证明与合成 VSDL 制备相比较的纯化曲线，(3) 将内毒素降低至低水平，(4) 在 10mM HCl 中配制 VSDL，(5) 使用与大规模制备相匹配的有机溶剂量，(6) 应用低压色谱步骤，和 (7) 是可升级的。

[0224] 实施例 2 大肠杆菌中 VSDL 10 一链节的重组制备

[0225] 根据实施例 1”加料分批发酵”部分中描述的，从实施例 1 描述的包含 pBRE606 表达构建体的 BR067 大肠杆菌菌株的发酵物产生溶胞产物，其编码具有 VSDL 肽的 10 个串联拷贝的 B72R-$(VSDL)_{10}$ 融合多肽（即 10 一链节）。

[0226] 将 275mL 的 10 一链节溶胞产物解冻，并且用 1M NaOH 将 pH 调节到 pH 8.0。TrypZean®（如实施例 1 所述）以 1:120 质量比被加给融合多肽，消化作用在室温下进行 2 小时。用 1M 乙酸将消化的溶胞产物的 pH 调节至 pH 4.4 以沉淀杂质，然后用 Millistak+Pod Labscale D0HC 过滤器 (MD0HC027H2)0.027m^2 深度过滤。

[0227] 通过直接加载到反相 Amberchrom CG300C 树脂柱上纯化滤液。用大约 15 柱体积的 0.5％乙酸，然后大约 10 柱体积的 0.5％乙酸 /15％异丙醇；然后大约 15 柱体积的 25mMTris pH 8 冲洗柱子。用 25mM Tris/15％异丙醇洗脱柱子直到 UV 吸光率达到主峰的 10％。然后将洗脱物加载到 Chromasorb 膜 (Merck-Millipore) 上，用 10mM 乙酸钠 pH 5 洗脱，直到获得 UV 基线，然后用 10mM 乙酸钠 /50mM 氯化钠 pH 5 冲洗。然后用 10mM 乙酸钠 /200mM 氯化钠将膜洗提，直到获得 UV 基线。然后将 Chromasorb 洗脱物加载回到过滤器上，用 10mM 乙

酸钠 pH 5 洗涤直到获得 UV 基线，然后用 10mM HCl 洗提直到获得 UV 基线。以大约 1mg/ml 溶胞产物的产率回收 VSDL 肽。

[0228] 将 Chromasorb 洗脱物等份分成数个试管 / 小瓶，并且在 -70℃下干燥大约 30 分钟，然后转移给 Labconco 旋转式干燥器用于如下冻干。样品在 Labconco 中在 -25℃下冷冻 30 分钟，然后将样品抽真空。第一次干燥发生在 -15℃进行 24hr。第二次干燥发生在 +17℃过夜。然后将冻干的在 -20℃冰柜中储存。

[0229] 冻干物质在 0.9％ (w/v)NaCl 中重新配制并且利用标准技术通过平行试验进行分析 (表 5)。

[0230] 表 5. 再生冻干的重组 VSDL 的分析

[0231]

分析	结果
纯度 (SEC)	99.2％
HMW 杂质 (SEC)	0.1％
纯度 (RP-UPLC)	98.8％
最高单一杂质 (RP-UPLC)	0.3％
内毒素	<0.6EU/mg
宿主细胞蛋白质 (ELISA)	<12ng/mg
DNA	<2600pg/mg
N- 末端测序	正确序列

[0232] *nb SEC 是尺寸排阻色谱法的首字母，HMW 是高分子量的首字母，RP-UPLC 是反相超高压液相色谱的首字母。

[0233] 实施例 3 大肠杆菌中 VSDL 3 一链节的重组制备

[0234] 和每个实施例 2 一样获得 VSDL 溶胞产物，除了使用实施例 1 中描述的包含 pBRE602 表达构建体的 BR067 大肠杆菌菌株，其编码具有 VSDL 肽的 3 个串联拷贝的 B72R-$(VSDL)_3$ 融合多肽 (即 3 一链节)。

[0235] 将 B72R-$(VSDL)_3$ 溶胞产物 (175mL) 解冻达到 15 和 25℃之间的温度。用 1MNaOH 将溶胞产物的 pH 调节至 8.0±0.1。以 25μg/mL 溶胞产物加入重新配制的 TrypZean。将消化物轻轻混合 5 分钟，然后在室温下静置过夜。用 1M 将消化液的 pH 调节至 4.4±0.1。使用磁子搅拌器轻轻搅拌溶液。

[0236] 然后通过 Millistak+Pod Labscale D0HC 过滤器 (MD0HC027H2)0.027m^2 以 4.2mL/min 将溶胞产物深度过滤，然后直接加载到 Amberchrom CG300C 柱上 (16mm ID 柱，床高 7.0cm)。用 0.5％乙酸以 4.2mL/min 大约 50 分钟，然后 0.5％乙酸 /15％异丙醇以 4.2mL/min 大约 25 分钟，然后 25mM Tris 以 4.2mL/min 大约 35 分钟，和然后 25mM Tris/3％异丙

醇以 4.2mL/min 冲洗柱子。当 UV 吸光率从最大开始降低时，冲洗变为 25mM Tris/15％异丙醇。用 25mM Tris/15％异丙醇以 4.2mL/min 洗提柱子直到 UV 吸光率达到主峰的 10％。图 11 图示从消化的溶胞产物分离 VSDL。

[0237] 实施例 4 B72QR 作为先导融合配偶体的新的构建体发酵分析

[0238] 对用谷氨酰胺残基对 B72R 融合配偶体的倒数第二个天冬氨酸残基的取代研究先导肽的 C- 末端精氨酸残基处增强的胰蛋白酶裂解。有所述取代作用的先导融合肽在这里被指作 B72(Q)R（指作 SEQ ID NO:10）。利用与实施例 1 所述相似的技术制备与 B72(Q)R 先导肽融合的新的 VSDL 6- 链节，8- 链节和 10- 链节构建体。得到的构建体被转化到大肠杆菌菌株 BR067 中并且如实施例 1 所述进行筛选。

[0239] 根据实施例 1 所述培养 B72R-$(VSDL)_3$，B72QR-$(VSDL)_6$，B72QR-$(VSDL)_8$，和 B72QR-$(VSDL)_{10}$ 菌株。使用 1:120 摩尔比的 TypZean 室温下将细胞溶胞产物温育过夜。如实施例 1 所述通过 RP-UPLC 定量测定 VSDL 肽。如表 6 所示，以 B72QR 为先导融合配偶体的 VSDL 8 —链节 10 —链节记录高浓度的 VSDL(2.5mg/ml)。

[0240] 表 6. 与 B72R 3- 链节构建体相比较的 B72QR 6- 链节，8- 链节或 10- 链节构建体的 RP-UPLC 分析

[0241]

VSDL 拷贝数目	VSDL 浓度 (mg/mL)
3	1.96
6	1.53
8	2.54
10	2.54

[0242] 预言实施例 1 GLP 的重组制备

[0243] 对编码天热人胰高血糖素样肽 (GLP，从人胰高血糖素前蛋白质原的氨基酸 98 至 127 衍生，其具有 NCBI 参考号 NP_002045) 的核苷酸序列进行修饰，从相应的氨基酸序列去除两个内部赖氨酸 (K) 残基，使得赖氨酸残基被转化为谷氨酰胺 (Q) 残基。天然氨基酸序列天然具有一个 C- 末端精氨酸 (R) 残基。在标准大肠杆菌载体中构建编码与修饰的 GLP 核苷酸序列的三个或八个串联拷贝操作连接的后面接着一个精氨酸残基（提供胰蛋白酶切割位点）的 N- 末端 His 标记的表达构建体。表达载体被转化到大肠杆菌中并且选择转化体，然后根据实施例 1 所述进行培养。

[0244] 通过用 1:120 摩尔比的 Trypzean 过夜培养裂解溶胞产物。根据实施例 1 — 4 所述能评价从融合肽分离重组 GLP。应用标准技术能确定 GLP 的最佳纯化。

[0245] 贯穿本说明书，词汇”包括”，或者变化形式”包含”或”含有”，被理解意味着包含陈述元件，整数或步骤，或者元件组，整数或步骤，但是不排除任何其它元件，整数或步骤，或者元件组，整数或步骤。

[0246] 本说明书中提到的所有的出版物在这里引作参考。本申请说明书中已经包括的文献、行为、材料、装置、物品等等的任何讨论只是为了提供本发明上下文的目的。不要承认因

为其在本申请各项权利要求的优先权日之前在澳大利亚或其他地方存在就认为这些技术方案中的任何或全部构成现有技术基础的一部分并且是与本发明相关领域中的公知常识。

[0247] 本领域技术人员明白对具体实施方案中所示的本发明可以进行许多改变和/或修改，而不脱离广泛描述的本发明的精神或范围。本实施例实施方案因此被认为在所有方面是说明性的而不是限制性。

[0248] 参考文献

[0249] De Palo EF 等，《临床化学》(Clin Chem)46:843-847(2000)

[0250] Franz M 等，《肾脏诠释》(Kidney Int)58:374-378(2000)

[0251] Franz M 等，《肾脏诠释》(Kidney Int)59:1928-1934(2001)

[0252] Hunter EFM 等，《扫描临床实验室投资》(Scan J Clin Lab Invest)58:205-216(1998)

[0253] Kirshenbaum K 等，《结构生物学最新见解》(Curr Opin Struct Biol)9:530-535(1999)

[0254] Skelton WP 4th 等，《抗癌研究》(Anticancer Res)(2)395-402(2011)(Abstract)

[0255] Vesely DL 等，《生物化学和生物物理研究通讯》(Biochem Biophys Res Commun)148:1540-1548(1987)

[0256] Vesely DL 等，《生物医学快报》(Proc Soc Exp Biol Med)192:230-235(1989)

[0257] Vesely DL，《美国生理学杂志》(Am.J.Physiology)285:F167-F177(2003)

[0258] Winters CJ，《循环》(Circulation)80:438-449(1989)

[0001]

序列表

<110> 麦德林制药私人有限公司
<120> 制备重组肽的方法
<130> 39357PCT
<160> 20
<170> PatentIn 版本 3.5

<210> 1
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 其中 Xaa 是除了精氨酸或赖氨酸的任意氨基酸；XBB 为精氨酸或赖氨酸。
<220>
<221> misc_特征
<222> (1)..(3)
<223> Xaa 可以是任意天然存在的氨基酸。
<220>
<221> misc_特征
<222> (5)..(5)
<223> Xaa 可以是任意天然存在的氨基酸。
<400> 1
Xaa Xaa Xaa Asn Xaa Asn Xbb
1 5

<210> 2
<211> 37
<212> PRT
<213> 智人
<400> 2
Glu Val Val Pro Pro Gln Val Leu Ser Glu Pro Asn Glu Glu Ala Gly
1 5 10 15
Ala Ala Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Thr Gly Glu Val
20 25 30
Ser Pro Ala Gln Arg
35

<210> 3
<211> 73
<212> PRT
<213> 原牛
<400> 3
Met Lys Glu Val Lys Ser Leu Leu Leu Asp Leu Gln Leu Leu Leu Glu
1 5 10 15
Lys Val Lys Asn Pro Glu Asn Leu Lys Leu Ser Arg Met His Thr Phe
20 25 30
Asp Phe Tyr Val Pro Lys Val Asn Ala Thr Glu Leu Lys His Leu Lys
35 40 45
Ala Leu Leu Glu Glu Leu Lys Leu Leu Glu Glu Val Leu Asn Leu Ala
50 55 60
Pro Ser Lys Asn Leu Asn Val Asp Arg
65 70

[0002]

<210> 4
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 肠激酶（肠激酶）的肽的切割位点。
<400> 4
Asp Asp Asp Asp Lys
1 5

<210> 5
<211> 8
<212> PRT
<213> 人工序列
<220>
<223> 烟草蚀纹病毒蛋白酶肽切割位点。其中 X 是任意氨基酸。
<220>
<221> misc_特征
<222> (2)..(3)
<223> Xaa 可以是任意天然存在的氨基酸。
<220>
<221> misc_特征
<222> (5)..(5)
<223> Xaa 可以是任意天然存在的氨基酸。XBB 为 Gly 或 Ser。
<400> 5
Glu Xaa Xaa Tyr Xaa Gln Xbb
1 5

<210> 6
<211> 6
<212> PRT
<213> 人工序列
<220>
<223> 烟草蚀纹病毒蛋白酶肽切割位点。其中 X 是任何氨基酸。
<220>
<221> misc_特征
<222> (1)..(1)
<223> Xaa 可以是任意天然存在的氨基酸。
<220>
<221> misc_特征
<222> (3)..(3)
<223> Xaa 可以是任意天然存在的氨基酸。. XBB 为 Gly 或 Ser。
<400> 6
Xaa Tyr Xaa Gln Xbb
1 5

<210> 7
<211> 6
<212> PRT
<213> 人工序列
<220>

[0003]

<223> 凝血酶切割位点。
<400> 7
Leu Val Pro Arg Gly Ser
1 5

<210> 8
<211> 37
<212> PRT
<213> 猩猩
<400> 8
Glu Val Val Pro Pro Gln Val Leu Ser Glu Gln Asn Glu Glu Ala Gly
1 5 10 15
Ala Ala Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Thr Gly Glu Val
20 25 30
Ser Pro Ala Gln Arg
35

<210> 9
<211> 37
<212> PRT
<213> Felis catus
<400> 9
Glu Val Val Pro Pro Gln Val Leu Ser Glu Gln Asn Glu Glu Ala Gly
1 5 10 15
Ala Ala Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Ala Gly Glu Val
20 25 30
Asn Pro Ala Gln Arg
35

<210> 10
<211> 73
<212> PRT
<213> 人工序列
<220>
<223>
<400> 10
Met Lys Glu Val Lys Ser Leu Leu Leu Asp Leu Gln Leu Leu Leu Glu
1 5 10 15
Lys Val Lys Asn Pro Glu Asn Leu Lys Leu Ser Arg Met His Thr Phe
20 25 30
Asp Phe Tyr Val Pro Lys Val Asn Ala Thr Glu Leu Lys His Leu Lys
35 40 45
Ala Leu Leu Glu Glu Leu Lys Leu Leu Glu Glu Val Leu Asn Leu Ala
50 55 60
Pro Ser Lys Asn Leu Asn Val Gln Arg
65 70

<210> 11
<211> 110
<212> PRT
<213> 人工序列
<220>

[0004]

<223>
<400> 11
Met Lys Glu Val Lys Ser Leu Leu Leu Asp Leu Gln Leu Leu Leu Glu
1 5 10 15
Lys Val Lys Asn Pro Glu Asn Leu Lys Leu Ser Arg Met His Thr Phe
20 25 30
Asp Phe Tyr Val Pro Lys Val Asn Ala Thr Glu Leu Lys His Leu Lys
35 40 45
Ala Leu Leu Glu Glu Leu Lys Leu Leu Glu Glu Val Leu Asn Leu Ala
50 55 60
Pro Ser Lys Asn Leu Asn Val Asp Arg Glu Val Val Pro Pro Gln Val
65 70 75 80
Leu Ser Glu Pro Asn Glu Glu Ala Gly Ala Ala Leu Ser Pro Leu Pro
85 90 95
Glu Val Pro Pro Trp Thr Gly Glu Val Ser Pro Ala Gln Arg
100 105 110

<210> 12
<211> 184
<212> PRT
<213> 人工序列
<220>
<223>
<400> 12
Met Lys Glu Val Lys Ser Leu Leu Leu Asp Leu Gln Leu Leu Leu Glu
1 5 10 15
Lys Val Lys Asn Pro Glu Asn Leu Lys Leu Ser Arg Met His Thr Phe
20 25 30
Asp Phe Tyr Val Pro Lys Val Asn Ala Thr Glu Leu Lys His Leu Lys
35 40 45
Ala Leu Leu Glu Glu Leu Lys Leu Leu Glu Glu Val Leu Asn Leu Ala
50 55 60
Pro Ser Lys Asn Leu Asn Val Asp Arg Glu Val Val Pro Pro Gln Val
65 70 75 80
Leu Ser Glu Pro Asn Glu Glu Ala Gly Ala Ala Leu Ser Pro Leu Pro
85 90 95
Glu Val Pro Pro Trp Thr Gly Glu Val Ser Pro Ala Gln Arg Glu Val
100 105 110
Val Pro Pro Gln Val Leu Ser Glu Pro Asn Glu Glu Ala Gly Ala Ala
115 120 125
Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Thr Gly Glu Val Ser Pro
130 135 140
Ala Gln Arg Glu Val Val Pro Pro Gln Val Leu Ser Glu Pro Asn Glu
145 150 155 160
Glu Ala Gly Ala Ala Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Thr
165 170 175
Gly Glu Val Ser Pro Ala Gln Arg
180

<210> 13
<211> 443
<212> PRT

[0005]

<213> 人工序列
<220>
<223>
<400> 13

```
Met Lys Glu Val Lys Ser Leu Leu Leu Asp Leu Gln Leu Leu Leu Glu
1               5                  10                  15
Lys Val Lys Asn Pro Glu Asn Leu Lys Leu Ser Arg Met His Thr Phe
            20                  25                  30
Asp Phe Tyr Val Pro Lys Val Asn Ala Thr Glu Leu Lys His Leu Lys
        35                  40                  45
Ala Leu Leu Glu Glu Leu Lys Leu Leu Glu Glu Val Leu Asn Leu Ala
    50                  55                  60
Pro Ser Lys Asn Leu Asn Val Asp Arg Glu Val Val Pro Pro Gln Val
65                  70                  75                  80
Leu Ser Glu Pro Asn Glu Glu Ala Gly Ala Ala Leu Ser Pro Leu Pro
                85                  90                  95
Glu Val Pro Pro Trp Thr Gly Glu Val Ser Pro Ala Gln Arg Glu Val
            100                 105                 110
Val Pro Pro Gln Val Leu Ser Glu Pro Asn Glu Glu Ala Gly Ala Ala
        115                 120                 125
Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Thr Gly Glu Val Ser Pro
    130                 135                 140
Ala Gln Arg Glu Val Val Pro Pro Gln Val Leu Ser Glu Pro Asn Glu
145                 150                 155                 160
Glu Ala Gly Ala Ala Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Thr
                165                 170                 175
Gly Glu Val Ser Pro Ala Gln Arg Glu Val Val Pro Pro Gln Val Leu
            180                 185                 190
Ser Glu Pro Asn Glu Glu Ala Gly Ala Ala Leu Ser Pro Leu Pro Glu
        195                 200                 205
Val Pro Pro Trp Thr Gly Glu Val Ser Pro Ala Gln Arg Glu Val Val
    210                 215                 220
Pro Pro Gln Val Leu Ser Glu Pro Asn Glu Glu Ala Gly Ala Ala Leu
225                 230                 235                 240
Ser Pro Leu Pro Glu Val Pro Pro Trp Thr Gly Glu Val Ser Pro Ala
                245                 250                 255
Gln Arg Glu Val Val Pro Pro Gln Val Leu Ser Glu Pro Asn Glu Glu
            260                 265                 270
Ala Gly Ala Ala Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Thr Gly
        275                 280                 285
Glu Val Ser Pro Ala Gln Arg Glu Val Val Pro Pro Gln Val Leu Ser
    290                 295                 300
Glu Pro Asn Glu Glu Ala Gly Ala Ala Leu Ser Pro Leu Pro Glu Val
305                 310                 315                 320
Pro Pro Trp Thr Gly Glu Val Ser Pro Ala Gln Arg Glu Val Val Pro
                325                 330                 335
Pro Gln Val Leu Ser Glu Pro Asn Glu Glu Ala Gly Ala Ala Leu Ser
            340                 345                 350
Pro Leu Pro Glu Val Pro Pro Trp Thr Gly Glu Val Ser Pro Ala Gln
        355                 360                 365
Arg Glu Val Val Pro Pro Gln Val Leu Ser Glu Pro Asn Glu Glu Ala
    370                 375                 380
```

[0006]

Gly Ala Ala Leu Ser Pro Leu Pro Glu Val Pro Pro Trp Thr Gly Glu
385 390 395 400
Val Ser Pro Ala Gln Arg Glu Val Val Pro Pro Gln Val Leu Ser Glu
405 410 415
Pro Asn Glu Glu Ala Gly Ala Ala Leu Ser Pro Leu Pro Glu Val Pro
420 425 430
Pro Trp Thr Gly Glu Val Ser Pro Ala Gln Arg
435 440

<210> 14
<211> 340
<212> DNA
<213> 人工序列
<220>
<223> B72R-VSDL (340 bp)
<400> 14
catatgaaag aagtgaagtc attgctgctg gatttacagt tgcttttgga gaaagttaaa 60
aatcctgaga acctcaagct ctccaggatg catacatttg acttttacgt gcccaaggtt 120
aacgctacag aattgaaaca tcttaaggcg ttactagaag aactcaaact tctagaggaa 180
gtactaaatt tagctccaag caaaaacctg aacgtcgacc gtgaagttgt accgccgcag 240
gttctgtctg aaccgaacga agaagcgggt gcggcgctgt ctccgctgcc ggaagttccg 300
ccgtggaccg gtgaagtgag cccagcgcag cgttaagctt 340

<210> 15
<211> 562
<212> DNA
<213> 人工序列
<220>
<223> B72R-(VSDL)3 (562 bp)
<400> 15
catatgaaag aagtgaagtc attgctgctg gatttacagt tgcttttgga gaaagttaaa 60
aatcctgaga acctcaagct ctccaggatg catacatttg acttttacgt gcccaaggtt 120
aacgctacag aattgaaaca tcttaaggcg ttactagaag aactcaaact tctagaggaa 180
gtactaaatt tagctccaag caaaaacctg aacgtcgacc gtgaagttgt accgccgcag 240
gttctgtctg aaccgaacga agaagcgggt gcggcgctgt ctccgctgcc ggaagttccg 300
ccgtggaccg gtgaggtgag cccagcgcag cgcgaagttg taccgccgca ggttctgtct 360
gaaccgaacg aagaagcggg tgcggcgctg tctccgctgc cggaagttcc gccatggacc 420
ggcgaagtga gcccagcgca gcgcgaagtt gtaccgccgc aggttctgtc tgaaccgaac 480
gaagaagcgg gtgcggcgct gtctccgctg ccggaagttc cgccatggac cggtgaagtg 540
agcccagcgc agcgttaagc tt 562

<210> 16
<211> 562
<212> DNA
<213> 人工序列
<220>
<223> B72(Q)R-(VSDL)3 (562 bp)
<400> 16
catatgaaag aagtgaagtc attgctgctg gatttacagt tgcttttgga gaaagttaaa 60
aatcctgaga acctcaagct ctccaggatg catacatttg acttttacgt gcccaaggtt 120
aacgctacag aattgaaaca tcttaaggcg ttactagaag aactcaaact tctagaggaa 180
gtactaaatt tagctccaag caaaaacctg aacgtccagc gtgaagttgt accgccgcag 240

[0007]

gttctgtctg aaccgaacga agaagcgggt gcggcgctgt ctccgctgcc ggaagttccg 300
ccgtggaccg gtgaggtgag cccagcgcag cgcgaagttg taccgccgca ggttctgtct 360
gaaccgaacg aagaagcggg tgcggcgctg tctccgctgc cggaagttcc gccatggacc 420
ggcgaagtga gcccagcgca gcgcgaagtt gtaccgccgc aggttctgtc tgaaccgaac 480
gaagaagcgg gtgcggcgct gtctccgctg ccggaagttc cgccatggac cggtgaagtg 540
agcccagcgc agcgttaagc tt 562

<210> 17
<211> 895
<212> DNA
<213> 人工序列
<220>
<223> B72(Q)R-(VSDL)6 (895 bp)
<400> 17
catatgaaag aagtgaagtc attgctgctg gatttacagt tgcttttgga gaaagttaaa 60
aatcctgaga acctcaagct ctccaggatg catacatttg acttttacgt gcccaaggtt 120
aacgctacag aattgaaaca tcttaaggcg ttactagaag aactcaaact tctagaggaa 180
gtactaaatt tagctccaag caaaaacctg aacgtccagc gtgaagtggt tccgccgcag 240
gttctgagcg aaccgaacga agaagcaggt gcagcactgt ctccgctgcc ggaagtcccg 300
ccgtggaccg gtgaagttag tccggcgcaa cgtgaagtgg ttccgccgca ggttctgagc 360
gaaccgaacg aagaagcagg tgcagcactg tctccgctgc cggaagtccc gccgtggacc 420
ggtgaagtta gtccggcgca acgtgaagtg gttccgccgc aggttctgag cgaaccgaac 480
gaagaagcag gtgcagcact gtctccgctg ccggaagtcc cgccgtggac cggtgaagtt 540
agtccggcgc aacgtgaagt ggttccgccg caggttctga gcgaaccgaa cgaagaagca 600
ggtgcagcac tgtctccgct gccggaagtc ccgccgtgga ccggtgaagt tagtccggcg 660
caacgtgaag tggttccgcc gcaggttctg agcgaaccga acgaagaagc aggtgcagca 720
ctgtctccgc tgccggaagt cccgccgtgg accggtgaag ttagtccggc gcaacgtgaa 780
gtggttccgc cgcaggttct gagcgaaccg aacgaagaag caggtgcagc actgtctccg 840
ctgccggaag tcccgccgtg gaccggtgaa gttagtccgg cgcaacgtta agctt 895

<210> 18
<211> 1117
<212> DNA
<213> 人工序列
<220>
<223> B72(Q)R-(VSDL)8 (1117 bp)
<400> 18
catatgaaag aagtgaagtc attgctgctg gatttacagt tgcttttgga gaaagttaaa 60
aatcctgaga acctcaagct ctccaggatg catacatttg acttttacgt gcccaaggtt 120
aacgctacag aattgaaaca tcttaaggcg ttactagaag aactcaaact tctagaggaa 180
gtactaaatt tagctccaag caaaaacctg aacgtccagc gtgaagtggt tccgccgcag 240
gttctgagcg aaccgaacga agaagcaggt gcagcactgt ctccgctgcc ggaagtcccg 300
ccgtggaccg gtgaagttag tccggcgcaa cgtgaagtgg ttccgccgca ggttctgagc 360
gaaccgaacg aagaagcagg tgcagcactg tctccgctgc cggaagtccc gccgtggacc 420
ggtgaagtta gtccggcgca acgtgaagtg gttccgccgc aggttctgag cgaaccgaac 480
gaagaagcag gtgcagcact gtctccgctg ccggaagtcc cgccgtggac cggtgaagtt 540
agtccggcgc aacgtgaagt ggttccgccg caggttctga gcgaaccgaa cgaagaagca 600
ggtgcagcac tgtctccgct gccggaagtc ccgccgtgga ccggtgaagt tagtccggcg 660
caacgtgaag tggttccgcc gcaggttctg agcgaaccga acgaagaagc aggtgcagca 720
ctgtctccgc tgccggaagt cccgccgtgg accggtgaag ttagtccggc gcaacgtgaa 780
gtggttccgc cgcaggttct gagcgaaccg aacgaagaag caggtgcagc actgtctccg 840
ctgccggaag tcccgccgtg gaccggtgaa gttagtccgg cgcaacgtga agtggttccg 900

[0008]

ccgcaggttc tgagcgaacc gaacgaagaa gcaggtgcag cactgtctcc gctgccggaa 960
gtcccgccgt ggaccggtga agttagtccg gcgcaacgtg aagtggttcc gccgcaggtt 1020
ctgagcgaac cgaacgaaga agcaggtgca gcactgtctc cgctgccgga agtcccgccg 1080
tggaccggtg aagttagtcc ggcgcaacgt taagctt 1117

<210> 19
<211> 1339
<212> DNA
<213> 人工序列
<220>
<223> B72(Q)R-(VSDL)10 (1339 bp)
<400> 19
catatgaaag aagtgaagtc attgctgctg gatttacagt tgcttttgga gaaagttaaa 60
aatcctgaga acctcaagct ctccaggatg catacatttg acttttacgt gcccaaggtt 120
aacgctacag aattgaaaca tcttaaggcg ttactagaag aactcaaact tctagaggaa 180
gtactaaatt tagctccaag caaaaacctg aacgtccagc gtgaagtggt tccgccgcag 240
gttctgagcg aaccgaacga agaagcaggt gcagcactgt ctccgctgcc ggaagtcccg 300
ccgtggaccg gtgaagttag tccggcgcaa cgtgaagtgg ttccgccgca ggttctgagc 360
gaaccgaacg aagaagcagg tgcagcactg tctccgctgc cggaagtccc gccgtggacc 420
ggtgaagtta gtccggcgca acgtgaagtg gttccgccgc aggttctgag cgaaccgaac 480
gaagaagcag gtgcagcact gtctccgctg ccggaagtcc cgccgtggac cggtgaagtt 540
agtccggcgc aacgtgaagt ggttccgccg caggttctga gcgaaccgaa cgaagaagca 600
ggtgcagcac tgtctccgct gccggaagtc ccgccgtgga ccggtgaagt tagtccggcg 660
caacgtgaag tggttccgcc gcaggttctg agcgaaccga acgaagaagc aggtgcagca 720
ctgtctccgc tgccggaagt cccgccgtgg accggtgaag ttagtccggc gcaacgtgaa 780
gtggttccgc cgcaggttct gagcgaaccg aacgaagaag caggtgcagc actgtctccg 840
ctgccggaag tcccgccgtg gaccggtgaa gttagtccgg cgcaacgtga agtggttccg 900
ccgcaggttc tgagcgaacc gaacgaagaa gcaggtgcag cactgtctcc gctgccggaa 960
gtcccgccgt ggaccggtga agttagtccg gcgcaacgtg aagtggttcc gccgcaggtt 1020
ctgagcgaac cgaacgaaga agcaggtgca gcactgtctc cgctgccgga agtcccgccg 1080
tggaccggtg aagttagtcc ggcgcaacgt gaagtggttc cgccgcaggt tctgagcgaa 1140
ccgaacgaag aagcaggtgc agcactgtct ccgctgccgg aagtcccgcc gtggaccggt 1200
gaagttagtc cggcgcaacg tgaagtggtt ccgccgcagg ttctgagcga accgaacgaa 1260
gaagcaggtg cagcactgtc tccgctgccg gaagtcccgc cgtggaccgg tgaagttagt 1320
ccggcgcaac gttaagctt 1339

<210> 20
<211> 20
<212> PRT
<213> 人工序列
<220>
<223> wherein X is any amino acid except arginine or lysine
<220>
<221> misc_特征
<222> (4)..(4)
<223> Xaa 可以是任意天然存在的氨基酸。
<220>
<221> misc_特征
<222> (9)..(9)
<223> Xaa 可以是任意天然存在的氨基酸。
<220>
<221> misc_特征

[0009]

<222> (11)..(11)
<223> Xaa 可以是任意天然存在的氨基酸。
<220>
<221> misc_特征
<222> (13)..(13)
<223> Xaa 可以是任意天然存在的氨基酸。
<400> 20
Ser Ser Asp Xaa Ser Ala Leu Leu Xaa Ser Xaa Leu Xaa Ala Leu Leu
1 5 10 15
Thr Ala Pro Arg
20

VSDL
EVVPPQVLSEPNEEAGAALSPLPEVPPWTGEVSPAQR (SEQ ID No:2)

B72R-VSDL
MKEVKSLLLDLQLLLEKVKNPENLKLSRMHTFDFYVPK
VNATELKHLKALLEELKLLEEVLNLAPSKNLNVDR↓
EVVPPQVLSEPNEEAGAALSPLPEVPPWTGEVSPAQR (SEQ ID No:11)

B72R-(VSDL)$_3$
MKEVKSLLLDLQLLLEKVKNPENLKLSRMHTFDFYVPK
VNATELKHLKALLEELKLLEEVLNLAPSKNLNVDR↓
EVVPPQVLSEPNEEAGAALSPLPEVPPWTGEVSPAQR↓
EVVPPQVLSEPNEEAGAALSPLPEVPPWTGEVSPAQR↓
EVVPPQVLSEPNEEAGAALSPLPEVPPWTGEVSPAQR↓(SEQ ID No:12)

图 1

B72R-VSDL (340 bp)

CATATGAAAGAAGTGAAGTCATTGCTGCTGGATTTACAGTTGCTTTTGGAGAAAGTTAAAAATCCTGAGAACCTCAAGCTCTCCAGGATGCATACATTTG
ACTTTTACGTGCCCAAGGTTAACGCTACAGAATTGAAACATCTTAAGGCGTTACTAGAAGAACTCAAACTTCTAGAGGAAGTACTAAATTTAGCTCCAAG
CAAAAACCTGAACGTCGACCGTGAAGTTGTACCGCCGCAGGTTCTGTCTGAACCGAACGAAGAAGCGGGTGCGGCGCTGTCTCCGCTGCCGGAAGTTCCG
CCGTGGACCGGTGAAGTGAGCCCAGCGCAGCGTTAAGCTT (SEQ ID No:14)

B72R-(VSDL)$_3$ (562 bp)

CATATGAAAGAAGTGAAGTCATTGCTGCTGGATTTACAGTTGCTTTTGGAGAAAGTTAAAAATCCTGAGAACCTCAAGCTCTCCAGGATGCATACATTTG
ACTTTTACGTGCCCAAGGTTAACGCTACAGAATTGAAACATCTTAAGGCGTTACTAGAAGAACTCAAACTTCTAGAGGAAGTACTAAATTTAGCTCCAAG
CAAAAACCTGAACGTCGACCGTGAAGTTGTACCGCCGCAGGTTCTGTCTGAACCGAACGAAGAAGCGGGTGCGGCGCTGTCTCCGCTGCCGGAAGTTCCG
CCGTGGACCGGTGAGGTGAGCCCAGCGCAGCGCGAAGTTGTACCGCCGCAGGTTCTGTCTGAACCGAACGAAGAAGCGGGTGCGGCGCTGTCTCCGCTGC
CGGAAGTTCCGCCATGGACCGGCGAAGTGAGCCCAGCGCAGCGCGAAGTTGTACCGCCGCAGGTTCTGTCTGAACCGAACGAAGAAGCGGGTGCGGCGCT
GTCTCCGCTGCCGGAAGTTCCGCCATGGACCGGTGAAGTGAGCCCAGCGCAGCGTTAAGCTT (SEQ ID No:15)

B72(Q)R-(VSDL)$_3$ (562 bp)

CATATGAAAGAAGTGAAGTCATTGCTGCTGGATTTACAGTTGCTTTTGGAGAAAGTTAAAAATCCTGAGAACCTCAAGCTCTCCAGGATGCATACATTTG
ACTTTTACGTGCCCAAGGTTAACGCTACAGAATTGAAACATCTTAAGGCGTTACTAGAAGAACTCAAACTTCTAGAGGAAGTACTAAATTTAGCTCCAAG
CAAAAACCTGAACGTCCAGCGTGAAGTTGTACCGCCGCAGGTTCTGTCTGAACCGAACGAAGAAGCGGGTGCGGCGCTGTCTCCGCTGCCGGAAGTTCCG
CCGTGGACCGGTGAGGTGAGCCCAGCGCAGCGCGAAGTTGTACCGCCGCAGGTTCTGTCTGAACCGAACGAAGAAGCGGGTGCGGCGCTGTCTCCGCTGC
CGGAAGTTCCGCCATGGACCGGCGAAGTGAGCCCAGCGCAGCGCGAAGTTGTACCGCCGCAGGTTCTGTCTGAACCGAACGAAGAAGCGGGTGCGGCGCT
GTCTCCGCTGCCGGAAGTTCCGCCATGGACCGGTGAAGTGAGCCCAGCGCAGCGTTAAGCTT (SEQ ID No:16)

B72(Q)R-(VSDL)$_6$ (895 bp)

CATATGAAAGAAGTGAAGTCATTGCTGCTGGATTTACAGTTGCTTTTGGAGAAAGTTAAAAATCCTGAGAACCTCAAGCTCTCCAGGATGCATACATTTG
ACTTTTACGTGCCCAAGGTTAACGCTACAGAATTGAAACATCTTAAGGCGTTACTAGAAGAACTCAAACTTCTAGAGGAAGTACTAAATTTAGCTCCAAG
CAAAAACCTGAACGTCCAGCGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCG
CCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGC
CGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACT
GTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCA
GGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGA
ACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCT
GAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTTAAGCTT
(SEQ ID No:17)

B72(Q)R-(VSDL)$_8$ (1117 bp)

CATATGAAAGAAGTGAAGTCATTGCTGCTGGATTTACAGTTGCTTTTGGAGAAAGTTAAAAATCCTGAGAACCTCAAGCTCTCCAGGATGCATACATTTG
ACTTTTACGTGCCCAAGGTTAACGCTACAGAATTGAAACATCTTAAGGCGTTACTAGAAGAACTCAAACTTCTAGAGGAAGTACTAAATTTAGCTCCAAG
CAAAAACCTGAACGTCCAGCGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCG
CCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGC
CGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACT
GTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCA
GGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGA
ACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCT
GAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCG
CCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTG
AAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCC
GGCGCAACGTTAAGCTT (SEQ ID No:18)

B72(Q)R-(VSDL)$_{10}$ (1339 bp)

CATATGAAAGAAGTGAAGTCATTGCTGCTGGATTTACAGTTGCTTTTGGAGAAAGTTAAAAATCCTGAGAACCTCAAGCTCTCCAGGATGCATACATTTG
ACTTTTACGTGCCCAAGGTTAACGCTACAGAATTGAAACATCTTAAGGCGTTACTAGAAGAACTCAAACTTCTAGAGGAAGTACTAAATTTAGCTCCAAG
CAAAAACCTGAACGTCCAGCGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCG
CCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGC
CGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACT
GTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCA
GGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGA
ACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCT
GAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCG
CCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCCGGCGCAACGTG
AAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGTGAAGTTAGTCC
GGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGCCGTGGACCGGT
GAAGTTAGTCCGGCGCAACGTGAAGTGGTTCCGCCGCAGGTTCTGAGCGAACCGAACGAAGAAGCAGGTGCAGCACTGTCTCCGCTGCCGGAAGTCCCGC
CGTGGACCGGTGAAGTTAGTCCGGCGCAACGTTAAGCTT (SEQ ID No:19)

图 2

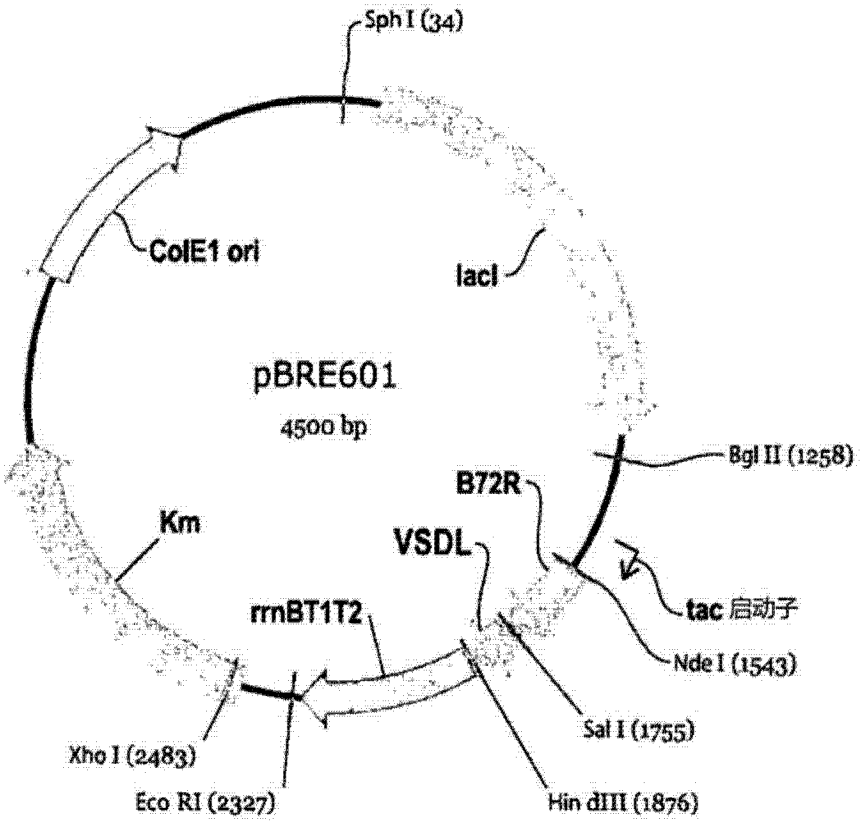

Sph I (34)
ColE1 ori
lacI
pBRE601
4500 bp
Bgl II (1258)
B72R
Km
VSDL
tac 启动子
rrnBT1T2
Nde I (1543)
Sal I (1755)
Xho I (2483)
Eco RI (2327)
Hin dIII (1876)

图 3A

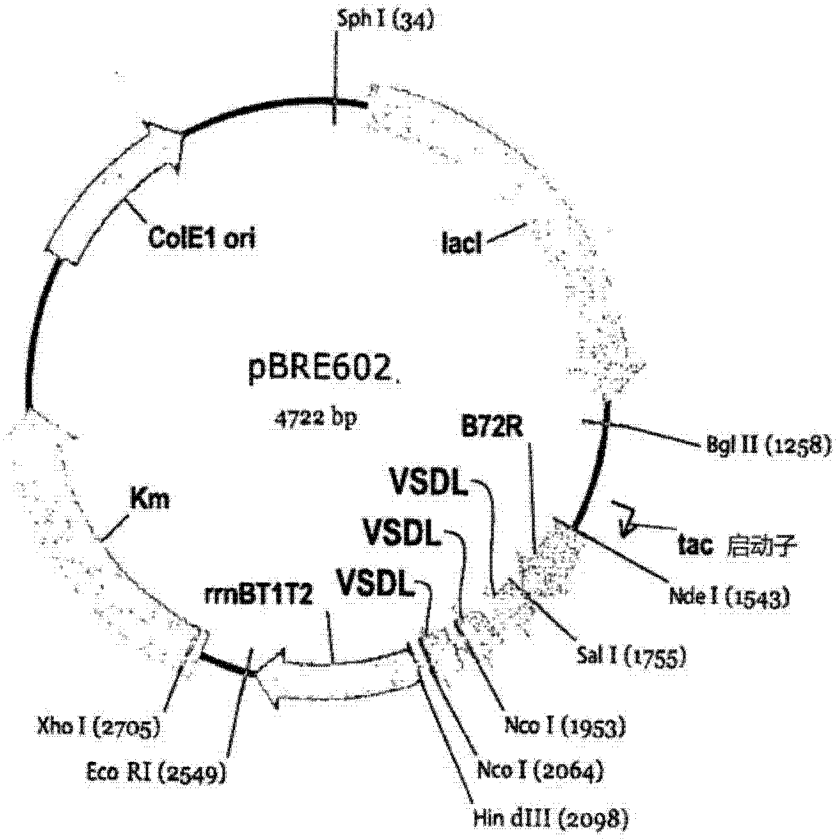

Sph I (34)
ColE1 ori
lacI
pBRE602
4722 bp
B72R
Bgl II (1258)
VSDL
Km
VSDL
tac 启动子
rrnBT1T2
VSDL
Nde I (1543)
Sal I (1755)
Nco I (1953)
Xho I (2705)
Eco RI (2549)
Nco I (2064)
Hin dIII (2098)

图 3B

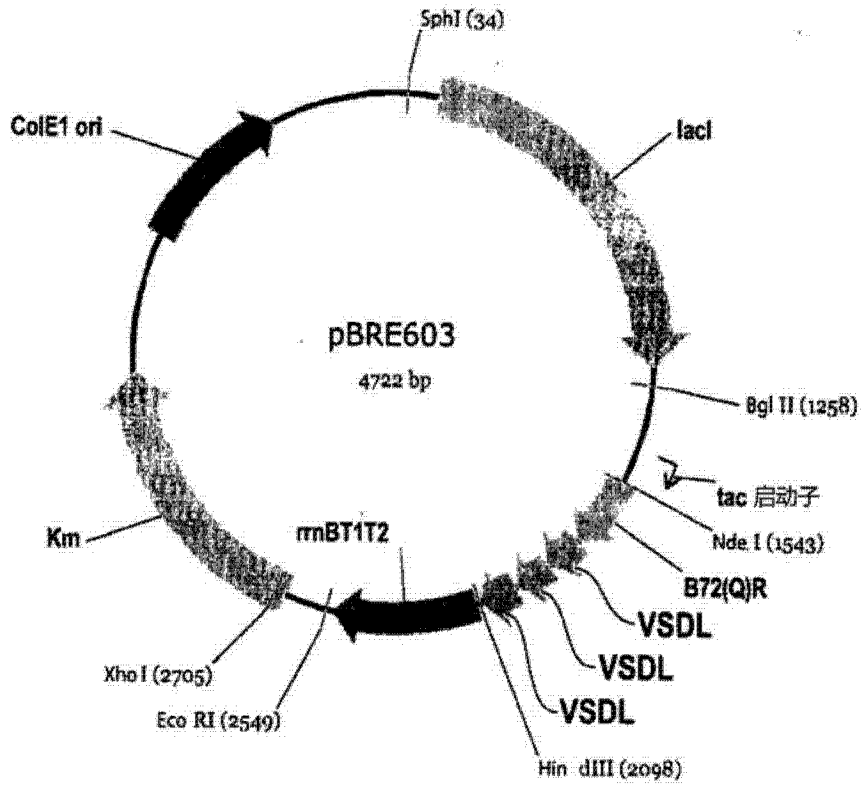

SphI (34)
lacI
ColE1 ori
pBRE603
4722 bp
Bgl II (1258)
tac 启动子
Nde I (1543)
B72(Q)R
VSDL
VSDL
VSDL
Hin dIII (2098)
Eco RI (2549)
Xho I (2705)
Km
rrnBT1T2

图 3C

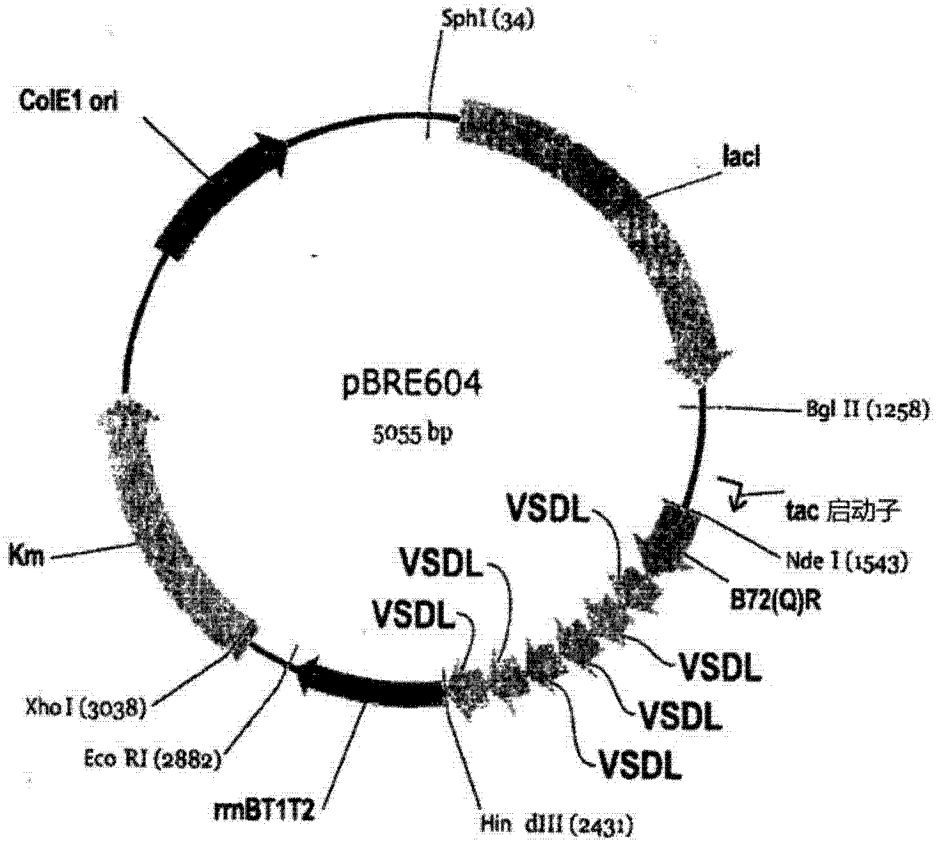

SphI (34)
ColE1 ori
lacI
pBRE604
5055 bp
Bgl II (1258)
tac 启动子
VSDL
Nde I (1543)
Km
VSDL
B72(Q)R
VSDL
VSDL
VSDL
VSDL
Xho I (3038)
Eco RI (2882)
rrnBT1T2
Hin dIII (2431)

图 3D

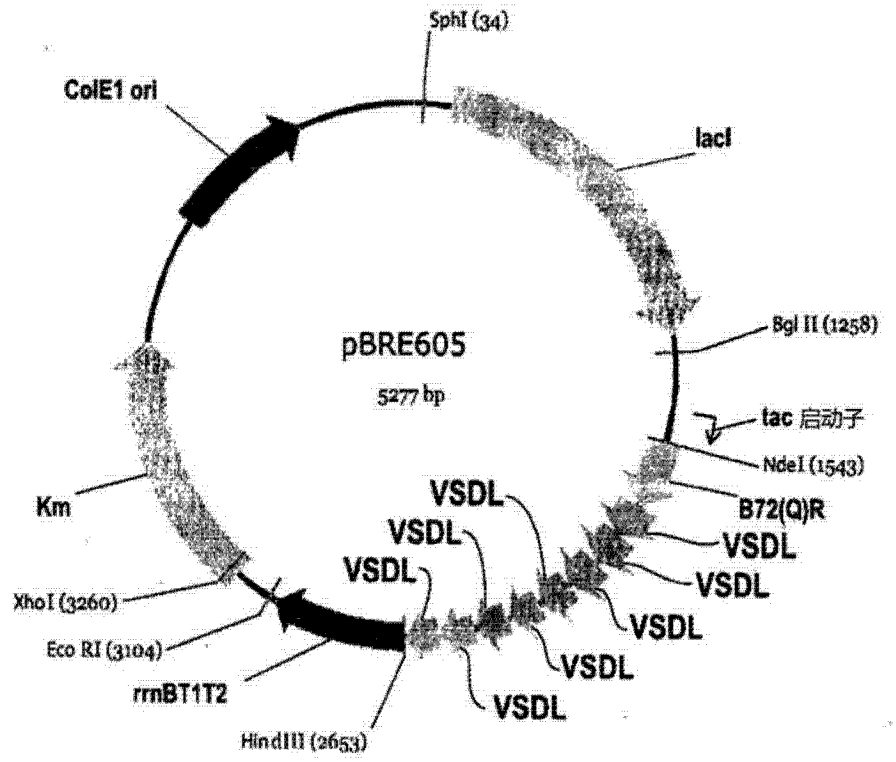

SphI (34)
ColE1 ori
lacI
pBRE605
5277 bp
Bgl II (1258)
tac 启动子
NdeI (1543)
B72(Q)R
VSDL
VSDL
VSDL
VSDL
VSDL
VSDL
VSDL
Km
XhoI (3260)
Eco RI (3104)
rrnBT1T2
HindIII (2653)

图 3E

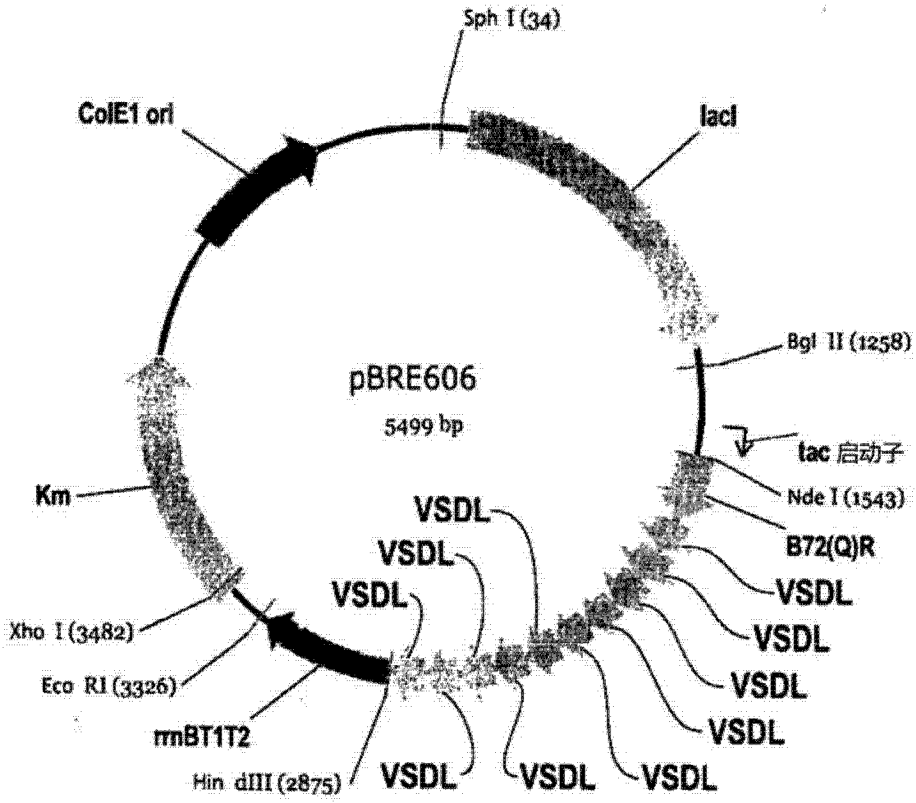

Sph I (34)
ColE1 ori
lacI
pBRE606
5499 bp
Bgl II (1258)
tac 启动子
Nde I (1543)
B72(Q)R
VSDL
VSDL
VSDL
VSDL
VSDL
VSDL
VSDL
VSDL
VSDL
Km
Xho I (3482)
Eco RI (3326)
rrnBT1T2
Hin dIII (2875)

图 3F

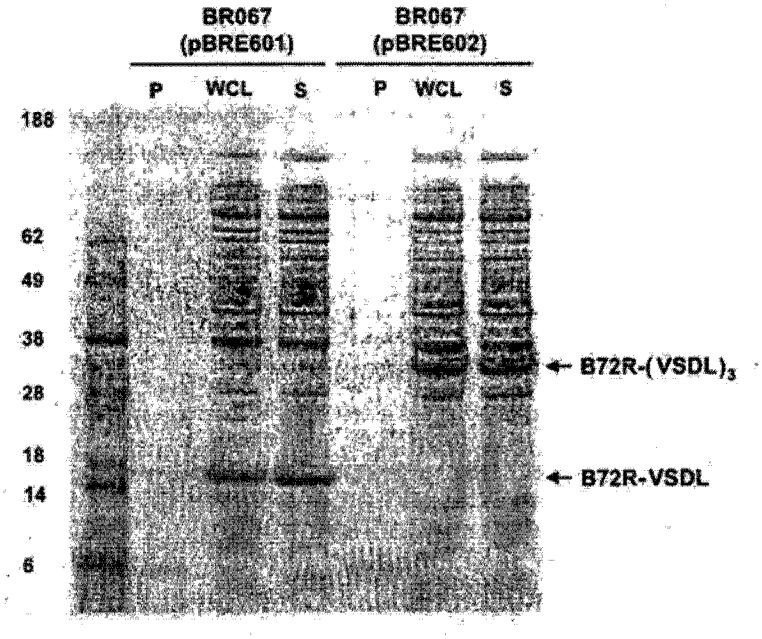
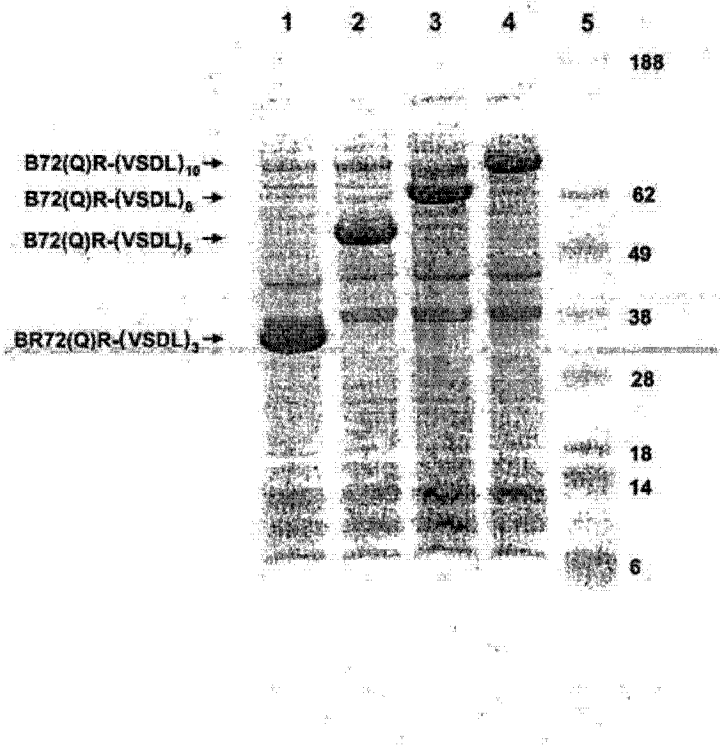

BR067
(pBRE601)
BR067
(pBRE602)
P
WCL
S
P
WCL
S
186
62
49
38
28
18
14
6
B72R-(VSDL)3
B72R-VSDL

图 4



泳道	样 品
1	BR067(pBRE603) + IPTG
2	BR067(pBRE604) + IPTG
3	BR067(pBRE605) + IPTG
4	BR067(pBRE606) + IPTG
5	见蓝色标记

图 5

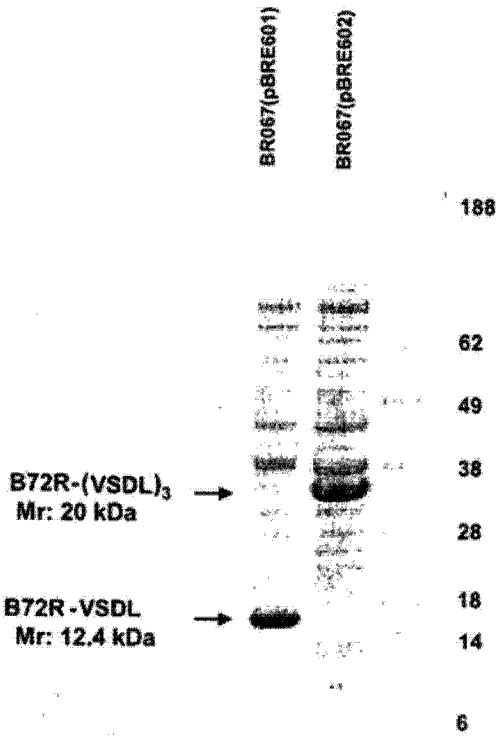

BR067(pBRE601)
BR067(pBRE602)
188
62
49
38
28
18
14
6
B72R-(VSDL)3
Mr: 20 kDa
B72R-VSDL
Mr: 12.4 kDa

图 6

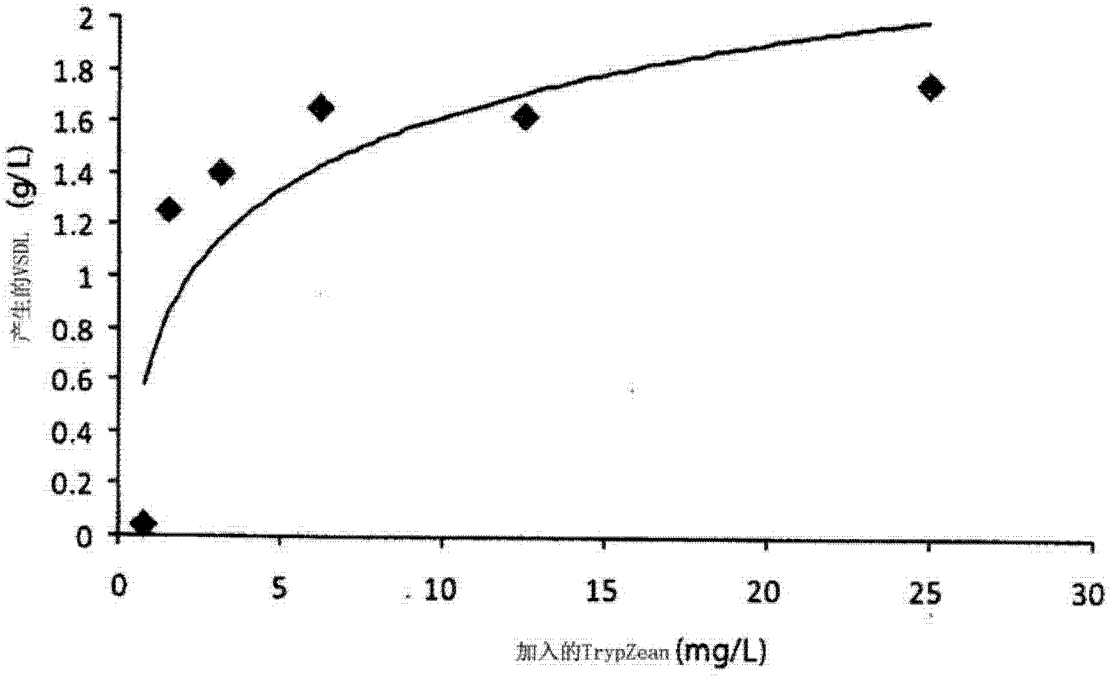

产生的VSDL (g/L)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
加入的TrypZean (mg/L)

图 7

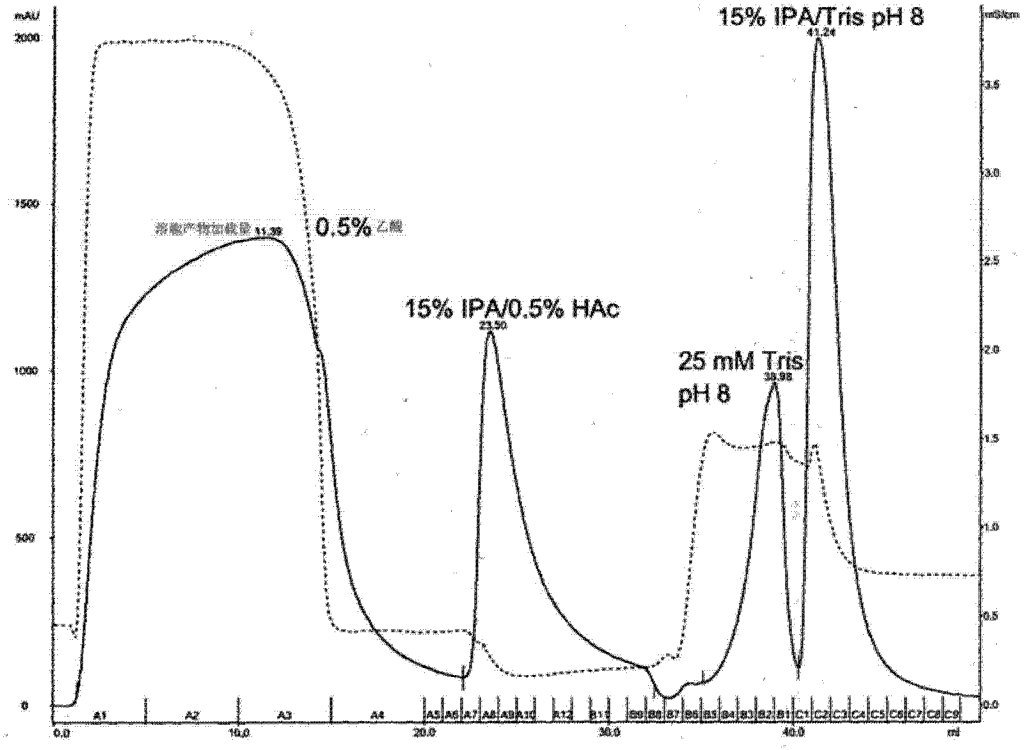

15% IPA/Tris pH 8
0.5%
15% IPA/0.5% HAc
25 mM Tris
pH 8

图 8

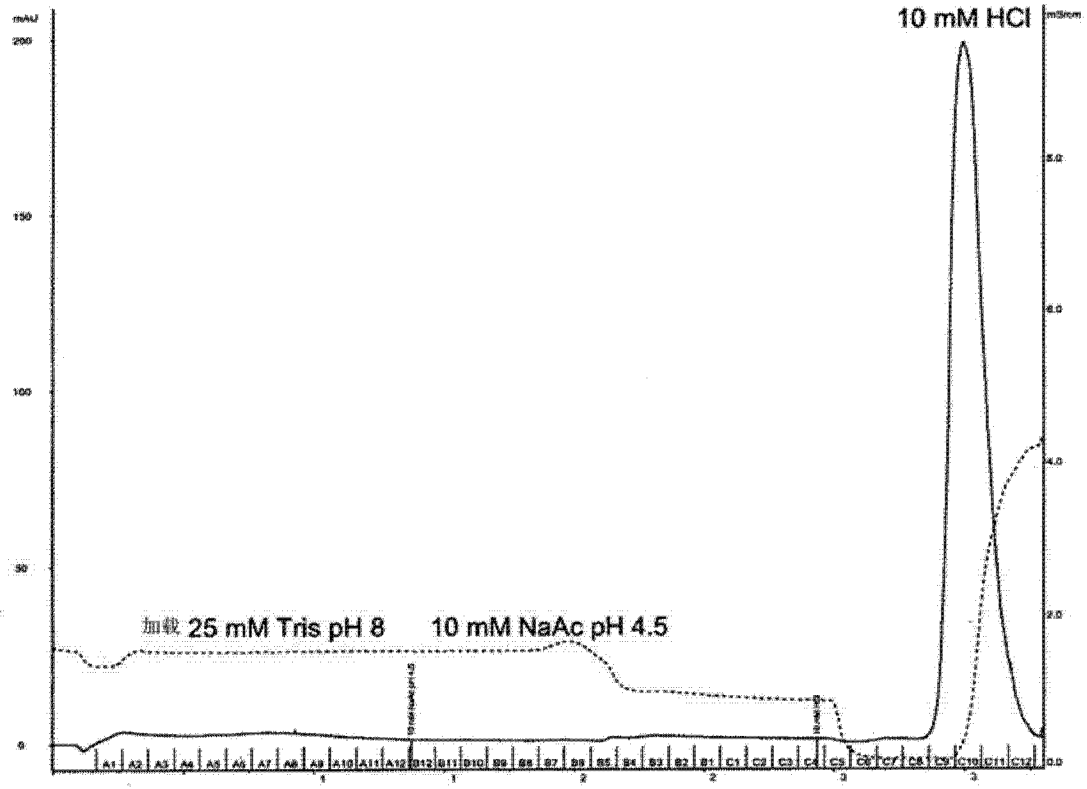

10 mM HCl
加载 25 mM Tris pH 8
10 mM NaAc pH 4.5

图 9

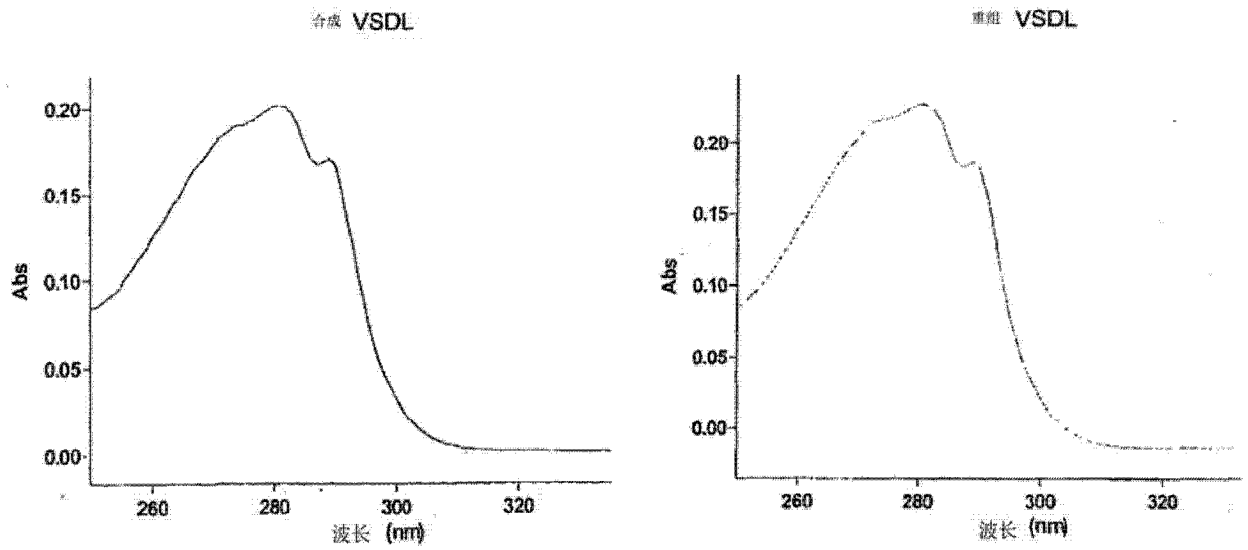

合成 VSDL
Abs
0.20
0.15
0.10
0.05
0.00
260
280
300
320
波长 (nm)
重组 VSDL
Abs
0.20
0.15
0.10
0.05
0.00
260
280
300
320
波长 (nm)

图 10

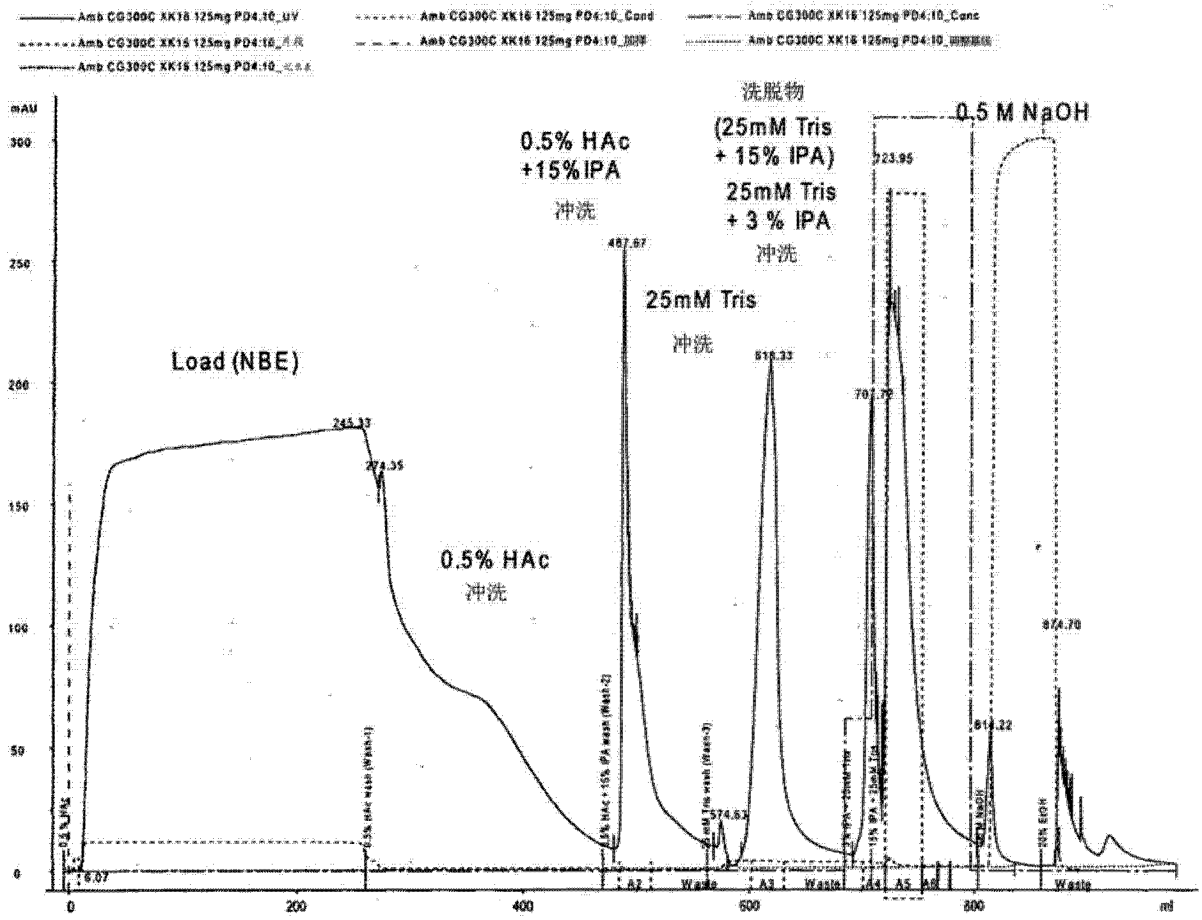

mAU
300
250
200
150
100
50
0
Load (NBE)
0.5% HAc
冲洗
0.5% HAc
+15%IPA
冲洗
洗脱物
(25mM Tris
+ 15% IPA)
25mM Tris
+ 3 % IPA
冲洗
25mM Tris
冲洗
0.5 M NaOH
245.33
274.35
487.67
574.63
616.33
701.72
723.95
814.22
874.70
6.07
A2
Waste
A3
Waste
A4
A5
Waste
0
200
400
600
800
ml

图 11

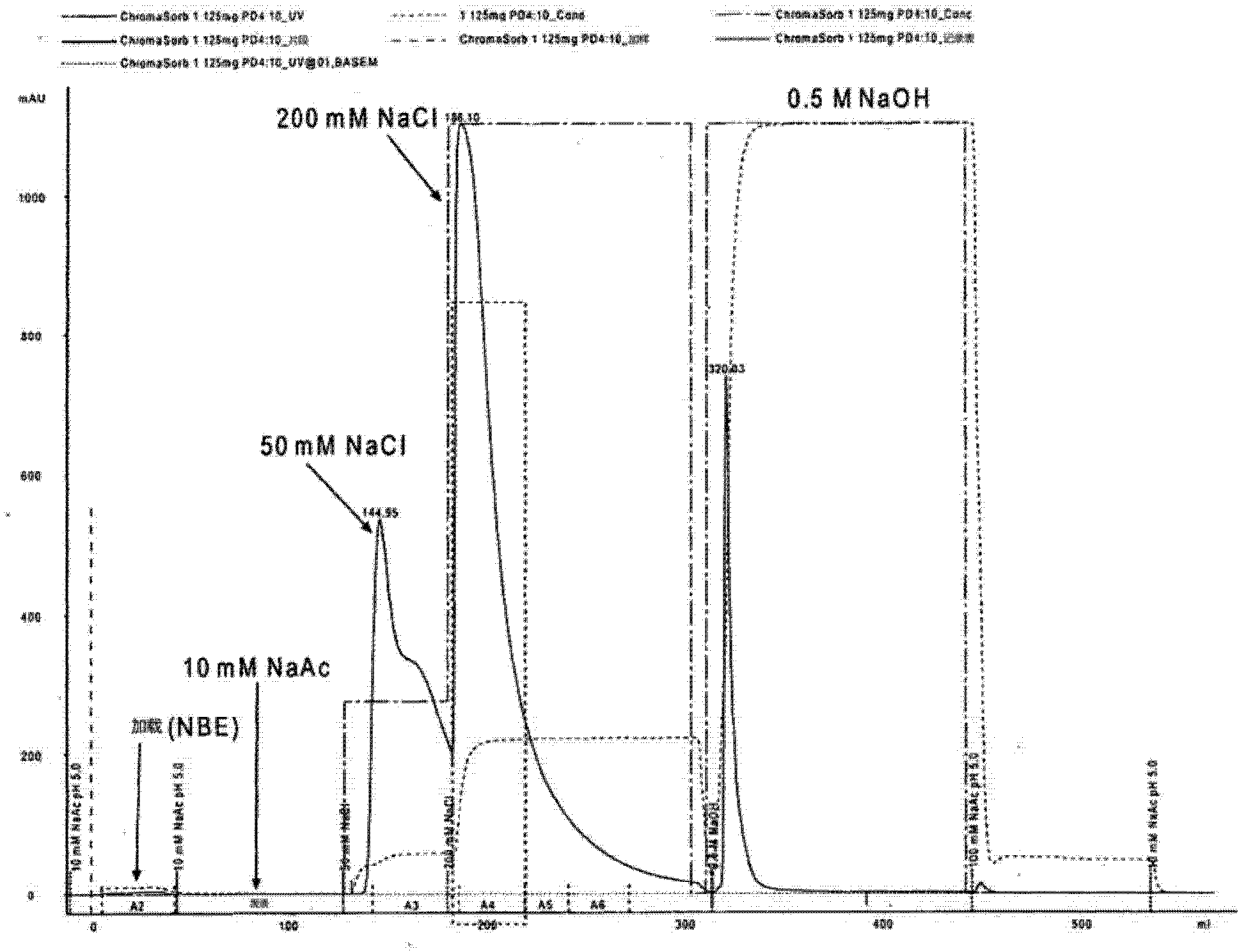

ChromaSorb 1 125mg PD4:10_UV
ChromaSorb 1 125mg PD4:10_UV@01,BASEM
mAU
1000
800
600
400
200
0
200 mM NaCl
50 mM NaCl
10 mM NaAc
加载 (NBE)
0.5 M NaOH
144.55
320.83
10 mM NaAc pH 5.0
A2
A3
A4
A5
A6
0
100
200
300
400
500
ml

图 12