



(12)发明专利申请

(10)申请公布号 CN 111454931 A

(43)申请公布日 2020.07.28

(21)申请号 202010161444.9	A61P 17/16(2006.01)
(22)申请日 2012.11.21	A61P 25/02(2006.01)
(62)分案原申请数据	A61P 11/02(2006.01)
201280075874.2 2012.11.21	A61P 25/08(2006.01)
(71)申请人 益普生生物创新有限公司	A61P 25/00(2006.01)
地址 英国牛津郡	A61P 11/04(2006.01)
(72)发明人 A·拉梅尔	A61P 15/02(2006.01)
(74)专利代理机构 北京市中咨律师事务所	A61P 17/00(2006.01)
11247	A61P 21/02(2006.01)
代理人 陈迎春 黄革生	A61P 25/14(2006.01)
(51)Int.Cl.	A61P 3/04(2006.01)
C12N 9/52(2006.01)	A61P 21/00(2006.01)
C12N 15/57(2006.01)	A61P 27/02(2006.01)
C07K 16/40(2006.01)	A61K 39/40(2006.01)
C12P 21/06(2006.01)	A61P 1/02(2006.01)
C12Q 1/37(2006.01)	A61P 1/06(2006.01)
A61K 38/48(2006.01)	A61P 13/06(2006.01)
A61K 39/395(2006.01)	A61P 13/10(2006.01)
A61P 13/08(2006.01)	

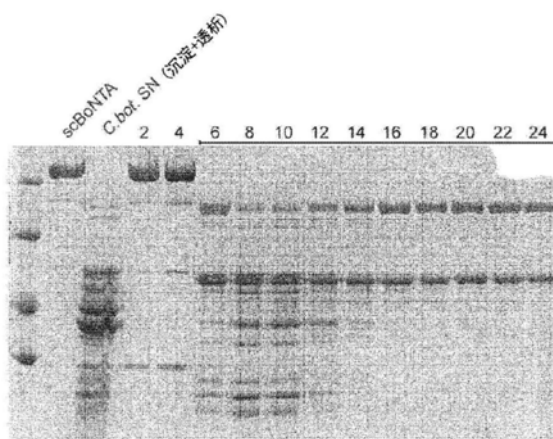
权利要求书2页 说明书28页  
序列表16页 附图13页

(54)发明名称

用于制备经蛋白水解处理的多肽的方法

(57)摘要

本发明涉及发明名称为“用于制备经蛋白水解处理的多肽的方法”的发明,其属于蛋白质工程领域,具体而言,本发明涉及一种具有蛋白水解活性的新型多肽以及所述多肽(及其它)在筛选和制备方法中的多种应用,本发明通过提供具有特定序列的蛋白酶而解决了将单链前体CNT有效转换成真正成熟的切割产物(即双链神经毒素)的技术问题。



1. 一种具有蛋白水解活性的多肽,其包含与SEQ ID NO:1的序列具有至少50%序列相同性的多肽序列。
2. 一种核酸分子,其包含编码权利要求1的多肽的核酸序列和任选的调节元件。
3. 一种载体,其包含如权利要求2所述的核酸分子。
4. 一种细胞,其包含如权利要求2所述的核酸分子或如权利要求3所述的载体。
5. 一种制备具有蛋白水解活性的多肽的方法,所述方法包括如下步骤:
  - (a.) 化学合成或从核苷酸序列翻译多肽,所述多肽包含与SEQ ID NO:1的序列具有至少50%序列相同性的多肽序列;和
  - (b.) 纯化步骤(a.)中的多肽。
6. 由权利要求5所述的方法能获得的多肽。
7. 一种抗体,其特异性结合至如权利要求1或6所述的多肽。
8. 如权利要求7所述的抗体在纯化权利要求1或6所述的多肽的方法中的应用。
9. 用于制备经蛋白水解处理的多肽的方法,所述方法包括将如下物质(a)与(b)接触的步骤:
  - (a) 第一多肽,所述第一多肽选自权利要求1或6所述的多肽、Lys-N、Lys-C、精氨酰内肽酶、胞浆素或OmpT蛋白,
  - (b) 第二多肽,所述第二多肽易被所述第一多肽蛋白水解;  
其中所述接触导致该第二多肽被蛋白水解处理成为至少两种切割产物。
10. 如权利要求9所述的方法,其特征在于,所述第二多肽包含与选自SEQ ID NO:3~25中任一条的多肽序列具有至少50%序列相同性的氨基酸序列;  
优选地,其中优选所述第一多肽在所述SEQ ID NO:3~25中任一条序列中的碱性氨基酸残基(例如,His、Lys、Arg)的直接C端位置上蛋白水解切割所述第二多肽。
11. 如权利要求9或权利要求10所述的方法,其特征在于,所述第二多肽是梭菌神经毒素(例如BoNT/A)。
12. 如权利要求11所述的方法,其特征在于,所述梭菌神经毒素选自:缺乏功能结合结构域(Hcc)并因而无法结合至天然梭菌神经毒素受体(例如,包含梭菌神经毒素的LH<sub>n</sub>片段的多肽或由该片段组成的多肽)的梭菌神经毒素多肽、具有结合至天然梭菌神经毒素受体的经修饰的梭菌神经毒素结合结构域(Hcc)的梭菌神经毒素多肽,或具有使所述梭菌神经毒素与非天然梭菌神经毒素受体结合的非梭菌结合结构域的梭菌神经毒素多肽(所述多肽任选地缺乏功能性结合结构域(Hcc),以使梭菌神经毒素与天然梭菌神经毒素受体的结合最小化)。
13. 如权利要求12所述的方法,其特征在于,所述梭菌神经毒素的L-链和H-链组分来自相同或不同梭菌神经毒素血清型和/或亚型。
14. 如权利要求9-13中任一项所述的方法,其特征在于,所述第二多肽是通过在大肠杆菌中的重组表达制备的单链梭菌神经毒素。
15. 如权利要求9-14中任一项所述的方法,其特征在于,所述第二多肽是单链梭菌神经毒素,并且其中所述第一和第二多肽的接触产生梭菌神经毒素双链多肽,所述梭菌神经毒素双链多肽包含通过二硫键共价连接至梭菌神经毒素H-链组分的梭菌神经毒素L-链组分。
16. 如权利要求9-15中任一项所述的方法,其特征在于,所述经蛋白水解处理的第二多

肽是梭菌神经毒素双链多肽,其中所述L-链的C端和所述H-链的N端与由野生型梭菌中相同单链梭菌神经毒素多肽生成的相应双链梭菌神经毒素的相应末端相同。

17.如权利要求9-16中任一项所述的方法,其特征在于,所述经蛋白水解处理的第二多肽是梭菌神经毒素双链多肽,其与由野生型梭菌中相同单链梭菌神经毒素多肽生成的相应梭菌神经毒素双链多肽具有相同的氨基酸序列。

18.如权利要求9-17中任一项所述的方法在产物质量评估或药物制备中的应用。

19.通过权利要求9-18中任一项所述的方法能获得的组合物,其特征在于,所述组合物包含经处理和未经处理的第二多肽的混合物,所述混合物包含少于5%的未处理的第二多肽。

20.一种筛选抑制剂的方法,所述方法包括如下步骤:

(a)使权利要求1或6所述的多肽与已知底物和任选的假定抑制剂接触;和

(b)检测所述假定抑制剂对底物转化为切割产物的作用

其中,切割产物的量减少指示所述假定抑制剂的抑制作用。

21.如权利要求1或6所述的具有蛋白水解活性的多肽的抑制剂,其特征在于,所述抑制剂是

(a.)包含如SEQ ID NO:4-25中任一条所示氨基酸序列的抑制剂,其中所包含的碱性氨基酸被非碱性氨基酸替代;或

(b.)如权利要求7所述的抗体。

22.一种药物组合物,其包含如权利要求1或6所述的多肽、如权利要求7所述的抗体、如权利要求19所述的组合物或如权利要求21所述的抑制剂。

## 用于制备经蛋白水解处理的多肽的方法

[0001] 本申请为2012年11月21日提交的,发明名称为“用于制备经蛋白水解处理的多肽的方法”的申请号为201280075874.2的分案申请。

[0002] 本发明涉及新型的具有蛋白水解活性的多肽以及所述多肽在筛选和制备方法中的多种应用。

[0003] 肉毒芽孢杆菌(*Clostridium botulinum*)和破伤风芽胞梭菌(破伤风芽胞梭菌)产生高度有力的神经毒素,即,分别是肉毒菌神经毒素(BoNT)和破伤风神经毒素(TeNT)。这些梭菌神经毒素(CNT)特异性地结合至神经元细胞,并且干扰神经递质的释放。肉毒芽胞梭菌分泌肉毒菌神经毒素(BoNT)的七种以A到G命名的差异抗原性血清型。所有血清型以及破伤风芽胞梭菌分泌的相关破伤风神经毒素(TeNT)破伤风芽胞梭菌均为阻断突触的胞吐的 $Zn^{2+}$ -内蛋白酶,其通过切割控制细胞膜融合的SNARE复合体的形成中涉及的蛋白质来进行上述阻断。CNT造成可见于肉毒症和破伤风的弛缓性肌肉麻痹。此外,还显示CNT活性会影响腺体分泌。CNT对肌肉和腺体活性的这些生理作用,被越来越多地用于各种治疗及美容应用。1989年,美国批准了肉毒菌神经毒素血清型A(BoNT/A)应用于人类以治疗斜视、眼睑痉挛及其它疾病。其市售可得为肉毒菌神经毒素A蛋白质制剂,例如,以商品名BOTOX(Allergan公司)和商品名DYSPOORT(Ipsen有限公司)出售。对于治疗性应用,通常将包含该神经毒素和其它细菌蛋白质的复合物直接注射进入待治疗的肌肉。在生理学pH下,该毒素从所述蛋白质复合物释放(Eisele等,2011,Toxicon 57(4):555-65.)随后发挥所需的药理作用。已可购得一种改善的不含复合蛋白的BoNT/A制剂,其以商品名XEOMIN或Bocouture(德国法兰克福的麦氏制药股份有限公司(Merz制药股份有限公司))。BoNT的作用仅是暂时性的,因此经常需要重复给予BoNT来维持治疗效应。

[0004] 各CNT初始时均被合成为无活性的单链多肽。在BoNT的情况中,该神经毒素多肽的分子量约为150kDa。该单链多肽的翻译后处理涉及在称作环体(loop)的暴露区域(参见表1)的限制性蛋白分解以及附近二硫键的形成。活化的双链神经毒素由该单链前体多肽的蛋白水解产生的两种切割产物构成:约50kDa的N末端轻链,和约100kDa的重链,这两者由二硫键连接。CNT结构上由三个结构域构成,即,催化性轻链、包含易位结构域(N末端的半个部分)和受体结合结构域(C末端的半个部分)的重链(参见Krieglstein 1990,Eur J Biochem 188:39;Krieglstein 1991,Eur J Biochem 202:41;Krieglstein 1994,J Protein Chem 13:49;Lacy等,1998,Nat.Struct.Biol.5(10):898-902)。根据该单链上存在的,位于形成该催化性结构域的氨基酸残基和形成该易位结构域的氨基酸残基之间的切割位点的数量,内肽酶活性能够产生两个大切割产物(即,轻链和重链),以及,代表前述的在该神经毒素的单链中起桥接作用的环体区域(之后形成轻链和重链)的特征性短肽(参见下表1)。

[0005] 从发酵溶液纯化CNT特别有难度,因为其中包含的神经毒素是未处理的、部分处理的和完全处理的多肽的混合物,其全部具有非常相似的生化及物理性质。如果内蛋白水解活性将所述轻链与所述环体之间的肽键水解,而所述环体与所述重链N端之间的肽键仍然完整,则通常生成部分处理的神经毒素。此外,如果所述内蛋白水解活性从所述重链释放所述环体肽,而所述环体肽与所述轻链的C端之间的肽键还未被水解,则也会生成部分处理的

神经毒素。根据发酵条件以及神经毒素的种类,不含环肽的完全处理的多肽可被5%~99%的部分处理或未处理多肽严重污染。而在一些情况下,所述神经毒素主要是未处理的,其在治疗应用之前需经内肽酶处理以变得具有生物活性。

[0006] 现有技术中描述了利用异源性蛋白酶处理芽孢梭菌神经毒素以减少未处理或部分处理的前体蛋白的量的多种方法。最广泛应用以活化芽孢梭菌神经毒素的蛋白酶是胰蛋白酶,其还可有用地活化芽孢梭菌神经毒素血清型B (BoNT/B) 和E (BoNT/E) (DasGupta和Sugiyama 1972, *Biochem. Biophys. Res. Commun.* 48:108-112; Kozaki等, 1974, *Infect. Immun.* 10:750-756), 但其似乎会产生次级产物(猜测其由BoNT/A附近的重链亚单位的C端的蛋白质水解产生), 从而似乎破坏毒素与其细胞受体的结合(Shone等, 1985, *Eur. J. Biochem.* 151:75-82)。理论上,从天然宿主(例如生产化呢个BoNT/A的肉毒芽孢梭菌)分离的内源性蛋白酶应能够产生更具特异性的切割产物。因此,人们做出多项尝试以从天然宿主细胞分离涉及芽孢梭菌神经毒素的蛋白水解活化的内源性蛋白酶。Dekleva和DasGupta (Dekleva和DasGupta, 1989, *Biochem. Biophys. Res. Commun.* 162:767-772) 从生成BoNT/A的肉毒芽孢梭菌培养物纯化了能够蛋白水解切割BoNT/A成为重链和轻链的亚单位的组分。相同作者的后续研究进一步表征了从肉毒芽孢梭菌分离的内源性蛋白酶(Dekleva和DasGupta, 1990, *J. Bact.* 172:2498-2503) 并揭示了一种62kDa的蛋白质,其由15.5kDa的多肽和48kDa的多肽构成。然而,观察到CNT在限制性暴露至Dekleva和DasGupta的62kDa蛋白质之后的大规模断裂,表明这种经分离的蛋白酶可能不是在芽孢梭菌细胞培养和感染期间负责活化CNT的尚未被识别的蛋白水解酶。事实上,其他研究者近期指出,梭菌蛋白酶(Clostripain), 也被称为梭菌肽酶B (Mitchel和Harrington, 1968, *JBC* 243:4683-4692) 可能涉及CNT的特异性活化(Sebahia等, 2007, *Genome Res.* 17 (7):1082-1092; W02009/014854)。有趣的是,该酶的结构和底物特异性使人联想到溶组织芽孢梭菌(*Clostridium histolyticum*) 分泌的 $\alpha$ 梭菌蛋白酶的那些特性(Dargatz等, 1993), 其同源物(具有74%氨基酸相同性)存在于肉毒芽孢梭菌中(CB01920)。溶组织芽孢梭菌 $\alpha$ 梭菌蛋白酶是对精胺酰基键最具严谨特异性的一种半胱氨酸内肽酶。其被合成为无活性的前酶原(prepro-enzyme), 其经自催化切割以生成15.4和43kDa的多肽,这两者相连以形成异质二聚体活性酶(Dargatz等, 1993)。溶组织芽孢梭菌 $\alpha$ 梭菌蛋白酶和肉毒芽孢梭菌62-kDa的蛋白酶均需要还原剂和钙来获得完全活性,并且对相同蛋白酶抑制剂敏感。这些数据强烈显示 $\alpha$ 梭菌蛋白酶的肉毒芽孢梭菌直接同源物(CB01920) 是负责蛋白水解切割肉毒芽孢梭菌的神经毒素的内源性蛋白酶。编码梭菌蛋白酶(CPE0846)的基因也存在于产气荚膜芽孢梭菌(*C. perfringens*) 中,并且发现该基因收到双成分系统VirR/VirS的正调节(Shimizu等, 2002b)。

[0007] 直至今日,仍然缺乏进一步的结论性实验证据,该领域仍然没有能够将单链前体CNT有效转换成真正成熟的切割产物(即双链神经毒素)的蛋白酶。本发明解决了上述一个或多个问题。

[0008] 用于减少未处理和/或部分处理的神经毒素肽,并由此改善神经毒素制剂的质量的方法高度需要但尚未获得。因此,本发明的潜在技术问题可被视为通过符合前述需求提供用于改善神经毒素多肽的制备的手段和方法。该技术问题通过如下权利要求和下文所述的实施方式来解决。

[0009] 因此,本发明的一方面涉及具有蛋白水解活性的多肽,其包含与SEQ ID NO:1具有至少50%序列相同性的多肽序列。在另一个方面,本发明涉及一种具有蛋白水解活性的多肽,其由与SEQ ID NO:1的序列具有至少50%序列相同性的多肽序列组成。在另一个方面,本发明涉及一种具有蛋白水解活性的多肽,其由如SEQ ID NO:1所示的多肽序列组成。

[0010] 本文所用术语“具有蛋白水解活性的多肽”指的是本发明的多肽的催化功能,并且表示本发明的多肽能够水解肽键。在一方面,“具有蛋白水解活性的多肽”指的是能够水解包含选择SEQ ID NO:4~25中任一条氨基酸序列的多肽的多肽。本文所用的术语“无蛋白水解活性的多肽”指的是本发明的多肽的催化功能,并且表示本发明的多肽无法水解肽键。

[0011] 经过测试多肽的蛋白水解活性,本领域技术人员能够确定根据本文所述的序列定义的多肽。一种用于检测蛋白水解活性的试验或测定系统包括:使包含与SEQ ID NO:1的序列具有至少50%序列相同性的多肽序列的多肽与测试底物接触。测试底物通常是已知可被本发明的多肽切割的多肽。优选地,所述测试底物是CNT,例如BoNT或其片段。所述测试底物可以是例如未经切割/未处理的BoNT,本文中称为“scBoNT”,并且可以是例如血清型A、B、C1、D、E、F或G(例如,“scBoNT/A”、“scBoNT/B”等)或所述测试底物可以是破伤风神经毒素。或者,所述测试底物可以是芽孢梭菌神经毒素的片段,所述片段包含选择SEQ ID NO:4~25中任一条氨基酸序列。所述片段可以是具有50个或更多个氨基酸残基的多肽,或者是具有多至49个氨基酸残基的肽。贯穿本说明书所用的术语“多肽”指的是具有50个或更多个氨基酸的分子,而术语“肽”指的是具有2~49个氨基酸残基的分子。在一方面中,所述测试底物是被称为LH<sub>N</sub>的可溶性神经毒素片段,其包含轻链多肽、暴露的环体肽区域以及重链多肽的N末端半部分(即易位结构域H<sub>N</sub>)。在另一个方面,所述测试底物是,或者包含选自SEQ ID NO:4~25中任一条的肽(参见表1)。而在另一个方面中,所述测试底物是嵌合型神经毒素,其包含源自两种或更多种血清型的氨基酸残基。

[0012] 用于测定蛋白水解活性的试验通常包括测定所述测试底物转化成其切割产物的程度的步骤。在多肽与测试底物接触之后,观察到产生一种或多种切割产物,或观察到切割产物的量增加,表示该多肽具有蛋白水解活性。所述测试步骤可能涉及底物与切割产物的比较。所述比较可能涉及测定底物的量和/或一种或多种切割产物的量。此外,用于测定蛋白水解活性的试验可能包括比较测试样品与参比样品的步骤,其中所述参比样品通常包含(a)多肽,其包含与SEQ ID NO:1的序列具有至少50%序列相同性的多肽序列,并且已知具有蛋白水解活性,和(b)测试底物,其已知可被(a)的多肽切割。在一方面,用于测定蛋白水解活性的分析包括:通过电泳或柱色谱以及(任选的)光谱分析,将底物与切割产物分离。方便的是,以一种或多种标记物标记测试底物,以更容易地检测测试底物的减少和/或产物的增加。本文所用术语“标记/标记物”,指的是可检测的标志物,并且包含,例如,放射性标记物、抗体、荧光标记物。所述测试底物和/或切割产物的量可通过例如放射显影技术或光谱法测定,包括基于至少两种标记物之间的能量共振转移的方法。或者,可利用免疫学方法,例如western印迹或ELISA来检测。用于测定本发明的多肽的蛋白水解活性的优选方法描述于下文用于说明本发明的实施例中。在本发明的特别优选的实施方式中,于37°C下,120分钟之内采用选自100mM Tris-HCl, pH 8.0或PBS (50mM Na<sub>2</sub>HPO<sub>4</sub>, 150mM NaCl, pH 7.4)的缓冲液,若多于20%,优选多于95%的测试底物被转化成切割产物(例如轻链和重链),则多肽具有蛋白水解活性。若测试底物不是全长神经毒素,而是例如该全长神经毒素的片段或该神

经毒素的衍生物,则仍适用该相同条件。明显的是,此时的切割产物会有所不同。然而,技术人员仍能定量对应的切割产物。在另一个方面,该试验通常采用100ng的具有蛋白水解活性的多肽以及相对于底物的1:100的摩尔比。而在另一个方面中,可以间隔采样以了解不同时间的催化活性。该试验可通过利用例如多个量的具有蛋白水解活性的多肽来加以改良。

[0013] SEQ ID NO:2显示衍生自肉毒芽孢梭菌菌株ATCC 3502 (GenBank登录号No:“CAL82988.1”)的不具有蛋白水解活性的多肽的多肽序列,其具有581残基的氨基酸长度。SEQ ID NO:1显示SEQ ID NO:2的具有蛋白水解活性的衍生物,其缺乏SEQ ID NO:2的氨基酸残基1~248。

[0014] 术语“包含与SEQ ID NO:1的序列具有至少50%序列相同性的多肽序列的多肽”指的是一种多肽,其与SEQ ID NO:1的序列具有至少50%的序列相同性。此外,该术语指一种多肽,其包含与SEQ ID NO:1的序列具有至少50%的序列相同性的多肽序列。所述多肽可具有其它氨基酸,例如在SEQ ID NO:1所示序列的内部位置或N或C端,或在与SEQ ID NO:1的序列具有至少50%的相同性的氨基酸序列的内部位置或N或C端,其中甲硫氨酸可能存在于该多肽的N端。此外,该术语指一种多肽,其缺失例如在SEQ ID NO:1所示序列的内部位置或N或C端的一个或多个氨基酸残基或在与SEQ ID NO:1的序列具有至少50%的相同性的序列的内部位置或N或C端的一个或多个氨基酸残基。

[0015] 本文所用的术语“序列相同性”指的是,测定参比氨基酸序列与研究序列之间的相同性,其中比对所述序列以获得最高级别的匹配,并且该相同性可利用计算机程序编码的公开技术或方法计算,例如,BLASTP、BLASTN、FASTA (Altschul 1990, J MoI Biol 215:403)。在一个方面中,相同性的百分比数值针对整个氨基酸序列计算。在另一个方面,序列相同性针对多至50aa残基、多至100aa残基、多至150aa、多至250aa、300aa、350aa、400aa、450aa、500aa或550aa残基的序列长度计算。在另一个方面,序列相同性针对至少50aa残基、至少100aa、至少150aa或至少250aa残基计算。在优选实施方式中,序列相同性针对SEQ ID NO:1或2的整体长度测定,即分别针对333aa或581aa的长度测定。基于多种算法的程序系列可被技术人员获得,以用来比较不同序列。在该情况下,Needleman和Wunsch或Smith和Waterman的算法产生特别可靠的结果。为了进行序列比对并计算本文所示序列的相同性的值,针对整体序列区域使用基于Clustal W算法的商用程序DNASTAR Lasergene MegAlign 7.1.0版,设定如下:成对比对参数:缺口罚分:10.00,缺口长度罚分:0.10,蛋白质重量矩阵:Gonnet 250,除非另有说明,一般采用上述设定作为序列比对的标准设定。

[0016] 本文中术语“至少50%序列相同性”指的是至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%或100%。

[0017] 本发明的具有蛋白水解活性的多肽可与如SEQ ID NO:1所示的参比多肽序列具有相同数量的氨基酸。本发明还包括具有额外或较少氨基酸残基的多肽。在一方面中,本发明的具有蛋白水解活性的多肽是或包含SEQ ID NO:1或2的截短突变体或与SEQ ID NO:1或2的序列具有至少50%序列相同性的多肽的截短突变体。SEQ ID NO:2的截短突变体可能缺失例如第249位氨基酸处的N端的一个或多个氨基酸残基。截短突变体可以是具有蛋白水解活性的N端或C端截短突变体和/或内部截短突变体。在一个方面中,SEQ ID NO:2的该截短突变体缺失SEQ ID NO:2的第1~248位氨基酸。在另一个方面,SEQ ID NO:2的截短突变体是C端截短突变体。在一个方面中,所示截短突变体缺失多至1、2、3、4、5、6、7、8、9、10、15、

20、50、100、150或多至170个连续氨基酸残基。在另一个方面,本发明的具有蛋白水解活性的多肽具有至少200aa残基、至少250aa残基、至少300aa残基或至少333aa残基的氨基酸长度。在另一个方面,本发明的具有蛋白水解活性的多肽具有至多333aa残基、至多350aa残基、至多573残基、至多581aa残基、至多592aa残基、至多600aa或至多617aa残基。

[0018] 在另一个方面,本发明的具有蛋白水解活性的多肽涵盖:在SEQ ID NO:1(或与SEQ ID NO:1的序列具有至少50%的序列相同性的多肽序列)的多肽链的N或C端和/或内部位置包含额外氨基酸残基的多肽。这些额外的氨基酸残基可包含至多5、至多10或甚至至多200、300或至多400个连续氨基酸残基。在一个方面中,所示额外氨基酸残基作为蛋白分解活性的抑制剂。在另一个方面,所示额外氨基酸残基可通过蛋白酶移除。在另一个方面,排除抑制本发明的多肽的蛋白分解活性的额外残基。所述额外氨基酸残基可能侧接一个或多个蛋白酶切割位点。在另一个方面,所述额外的氨基酸序列可作为可检测标签和/或允许与固体支持物结合。

[0019] 在另一个方面,SEQ ID NO:1(或与SEQ ID NO:1的序列具有至少50%的序列相同性的多肽序列)的多肽链通过交换一个或多个氨基酸残基来修饰。本文中所用的术语“交换”指的是以不同的氨基酸取代某一氨基酸。例如,在多肽序列内至多1aa、2aa、3aa、4aa、5aa、6aa、7aa、8aa、9aa、10aa、15aa、20aa或至多50aa可能被取代。交换可能涉及保守性或非保守性氨基酸的改变,目的在于,例如,增加或减少本发明的多肽的底物结合或蛋白水解活性。

[0020] 在一个方面中,本发明的具有蛋白水解活性的多肽涵盖:能将底物水解成两种或更多种天然切割产物的多肽。在另一个方面中,本发明的多肽将底物水解成与所述天然切割产物不同的两种或更多种切割产物。本文所用的术语“天然切割产物”或“天然产物”指的是,当与野生型细胞培养物中由相同底物产生的产物比较时,具有相同氨基酸序列的源自该底物的产物。在一个方面中,所述切割产物是肉毒神经毒素或破伤风神经毒素的双链神经毒素,在另一方面中,所述双链神经毒素是从肉毒芽孢梭菌血清型A、B、C1、D、E、F或G分离的神经毒素。而在其它方面中,所述双链神经毒素是天然双链神经毒素。

[0021] 表1显示所述前体,TeNT和BoNT/A-G的天然双链神经毒素,并鉴定包含被本发明的多肽切割的氨基酸序列的暴露环体。

毒素	暴露环体	LC	H <sub>N</sub>	H <sub>CN</sub>	H <sub>CC</sub>
BoNT/A1	SEQ ID NO: 4	M1-K438	A449-N872	I873-S1092	N1093-L1296
BoNT/A2	SEQ ID NO: 5	M1-K438	A449-N872	I873-S1092	N1093-L1296
BoNT/A3	SEQ ID NO: 6	M1-K434	A445-N868	I869-S1088	N1089-L1292
BoNT/A3	SEQ ID NO: 7	M1-K434	A445-N868	I869-S1088	N1089-L1292

[0022]

毒素	暴露环体	LC	H <sub>N</sub>	H <sub>CN</sub>	H <sub>CC</sub>
BoNT/A4	SEQ ID NO: 8	M1-K438	A449-N872	I873-S1092	N1093-L1296
BoNT/A5	SEQ ID NO: 9	M1-K438	A449-N872	I873-S1092	N1093-L1296
BoNT/A6	SEQ ID NO: 5	M1-K438	A449-N872	I873-S1093	N1094-L1297
BoNT/A7	SEQ ID NO: 10	M1-K438	A449-N872	I873-S1092	N1093-L1296
BoNT/B1	SEQ ID NO: 11	M1-K441	A442-I860	L861-S1079	Y1080-E1291
BoNT/B2	SEQ ID NO: 12	M1-R441	A442-I860	L861-S1079	Y1080-E1291
BoNT/B3	SEQ ID NO: 12	M1-R441	A442-I860	L861-S1079	Y1080-E1291
BoNT/B4bv	SEQ ID NO: 11	M1-K441	A442-I860	L861-S1079	Y1080-E1291
BoNT/B5nP	SEQ ID NO: 13	M1-K441	V442-I860	L861-S1079	Y1080-E1291
BoNT/B6	SEQ ID NO: 11	M1-K441	A442-I860	L861-S1079	Y1080-E1291
BoNT/C1	SEQ ID NO: 14	M1-R444	T450-I868	N869-L1092	Q1093-E1291
BoNT/CD	SEQ ID NO: 14	M1-R444	T450-I868	N869-Q1083	I1084-E1280
BoNT/D	SEQ ID NO: 15	M1-K442	D446-I864	N865-Q1079	I1080-E1276
BoNT/DC	SEQ ID NO: 16	M1-R442	D446-I864	N865-L1088	Q1089-E1285
BoNT/E1-E5	SEQ ID NO: 17	M1-K419	S424-I847	K848-P1067	N1068-K1252
BoNT/E6	SEQ ID NO: 18	M1-K419	S424-I847	K848-P1067	N1068-K1252
BoNT/F1	SEQ ID NO: 19	M1-R435	A440-I866	K867-P1085	D1086-N1278
BoNT/F2	SEQ ID NO: 20	M1-R435	Q440-I866	K867-P1088	D1089-E1280
BoNT/F3	SEQ ID NO: 20	M1-R435	Q440-I866	K867-P1088	D1089-E1279
BoNT/F4	SEQ ID NO: 21	M1-R435	A440-I866	K867-P1085	D1086-E1277

[0023]

毒素	暴露环体	LC	H <sub>N</sub>	H <sub>CN</sub>	H <sub>CC</sub>
BoNT/F5	SEQ ID NO: 22	M1-K434	P440-I863	K864-P1085	D1086-E1277
BoNT/F6	SEQ ID NO: 19	M1-R435	A440-I866	K867-P1088	D1089-E1275
BoNT/F7	SEQ ID NO: 23	M1-K427	N432-I857	I858-P1076	D1077-E1268
BoNT/G	SEQ ID NO: 24	M1-K442	S447-I865	S866-S1086	S1087-E1297
TeNT	SEQ ID NO: 25	M1-R449	T456-K883	S884-L1109	S1110-D1315

[0024] 应理解,除非另有说明,上述及下述术语的定义和解释适用于本说明书中所述的所有方面。

[0025] 本发明的具有蛋白水解活性的多肽适用于多种应用。一种市售相关的应用是其在制备治疗性神经毒素中的应用,所述神经毒素例如从肉毒芽孢梭菌分离的那些。目前,用于制备市售可得的肉毒神经毒素制剂的肉毒芽孢梭菌的细胞培养物被大量比处理或/或未处理的神经毒素污染,两者均负面影响(即减少)这些药物组合物的特定活性。在例如溶解肉毒芽孢梭菌之后采用本发明的具有蛋白水解活性或活化的多肽,将可能处理包含未处理和/或部分处理的神经毒素的组合物,从而将这些污染物转化成完全处理的神经毒素。因此,可提供具有增加神经毒素特定活性的市售产品,其中细菌蛋白质的总量可被减少,进一步降低患者形成抗体的风险。

[0026] 在另一个方面,本发明涉及一种核酸分子,该核酸分子包含编码本发明的多肽的核酸序列和可选的调节元件。本文所用术语“调节元件”指的是基因表达(包括转录和翻译)的调节元件,且包含例如tata盒、启动子、增强子、核糖体结合位点、Shine-Dalgarno序列、IRES区、聚腺苷酸化信号、末端加帽结构等。所述调节元件可包含一种或多种异源性调节元件或一种或多种同源性调节元件。“同源性调节元件”是本发明的核酸分子的衍生来源的野生型细胞的调节元件,其涉及该野生型细胞中该核酸分子或该多肽的基因表达的调节。本发明还涵盖包含异源性调节元件的核酸分子。术语“异源性调节元件”是不涉及该野生型细胞中所述核酸分子或所述多肽的基因表达的调节的调节元件。也包含诱导型表达的调节元件,例如诱导型启动子。所述核酸分子可以是,例如,hnRNA、mRNA、RNA、DNA、PNA、LNA和/或经修饰的核酸分子。所述核酸分子可以是环状、线性、纳入基因组或游离核酸。也涵盖编码融合蛋白的串联体,所述融合蛋白包含本发明的3、4、5、6、7、8、9或10种多肽。另外,所述核酸分子可包含编码用于细胞内运输的信号序列的序列,例如用于运输进入细胞内隔室或用于运输通过细胞膜的信号。

[0027] 在另一个方面,本发明涉及一种载体,所述载体包含基于本发明的核酸分子的核酸分子。载体可适用于体外和/或体内表达本发明的多肽。所述载体可以是暂时型和/或稳定型基因表达载体。在一个实施方式中,所述载体还包含调节元件和/或选择标志物。在一个实施方式中,所述载体是病毒源性载体,在另一个实施方式中,其为噬菌体源性载体,而在另一个实施方式中,其为细菌源性载体。

[0029] 在另一个方面,本发明涉及包含本发明的核酸分子或载体的细胞。本文中术语“细胞”涵盖用于表达所述核酸分子或所述载体以及特别是本发明的多肽的原核细胞和/或真核细胞。所述细胞可以是不表达本发明的多肽或其同源物的宿主细胞。本文所用术语“同源物”指的是一种多肽,其包含与SEQ ID NO:1的序列具有至少50%的序列相同性的多肽序列。然而,本发明还涵盖表达本发明的多肽或其同源物的细胞,特别是野生型细胞。在特定的方面中,本发明的细胞选自肉毒芽孢梭菌、酪酸芽孢梭菌(*C. butyricum*)、巴拉提尼芽孢梭菌(*C. baratii*)和破伤风芽孢梭菌。在一个优选方面中,所述细胞是肉毒芽孢梭菌血清型A、B或F。在另一个方面,所述细胞是肉毒芽孢梭菌的Hall菌株(ATCC 3502)。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/A菌株ATCC 19397,其也被称为NCTC 4587和NCTC 7272。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/A菌株NCTC 2916。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/A2菌株Kyoto-F和Mauritius/NCTC 9837。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/A3菌株A254 Loch Maree/NCTC 2012。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/A4和B菌株CDC657。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/A5和B3'菌株H04402 065。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/B1菌株Okra/NCTC 7273。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/B和F菌株CDC4013/NCTC 12265。在另一个方面,所述细胞是肉毒芽孢梭菌的产BoNT/F1菌株Langeland/NCTC 10281。在另一个方面,所述细胞是产芽孢梭菌(*Clostridium sporogenes*)、产气荚膜芽孢梭菌(*Clostridium perfringens*)、丙酮丁醇梭菌(*Clostridium acetobutylicum*)、蜡状杆菌(*B. cereus*)、苏力菌(*B. thuringiensis*)、蕈状芽胞杆菌(*B. mycoidis*)、嗜热溶蛋白芽孢杆菌(*B. thermoproteolyticus*)、炭疽杆菌(*B. anthracis*)、巨杆菌(*B. megaterium*)、枯草杆菌(*B. subtilis*)、大肠杆菌(*E. coli*)或酵母细胞。在一个方面中,本发明的多肽在细胞内经修饰(即,糖基化、磷酸化、经蛋白酶处理等)。修饰也包括添加非蛋白质辅因子,包括金属离子。包含如上所述的不具蛋白水解活性、任何中间多肽产物以及本文所述的最终具有蛋白水解活性的多肽的细胞均包含于本发明中。本发明还涵盖包含本发明的多肽的表达诱导子的细胞。所述表达诱导子可以是核酸分子或多肽或化学实体,包括具有增加本发明的具有蛋白水解活性的多肽于细胞培养物或其裂解物中的量或活性的作用的小化学实体。所述表达诱导子能够例如增加编码本发明的多肽的核酸分子的转录或翻译。或者,所述表达诱导子可以是能活化所述不具蛋白水解活性的多肽SEQ ID NO:2(或包含与SEQ ID NO:2的序列具有至少50%的序列相同性的多肽序列的多肽)的化合物。在一个方面中,所述细胞包含诱导子,其是能够移除该多肽的N端的抑制性氨基酸残基的具有蛋白水解活性的多肽。所述诱导子可通过例如本领域技术人员所知的重组方法来表达。或者,所述诱导子可从细胞分离,例如芽孢梭菌细胞。

[0030] 本发明还涉及本发明的核酸分子用于制备本发明的具有蛋白水解活性的多肽的应用。

[0031] 在一个相关方面中,本发明涉及一种制备具有蛋白水解活性的多肽的方法,所述方法包括如下步骤:(a)化学合成或从核苷酸序列翻译多肽,该多肽包含与SEQ ID NO:1的序列具有至少50%的序列相同性的多肽序列;和(b)纯化步骤(a)的多肽。

[0032] 术语“化学合成”指的是通过化学手段合成多肽。所述方法在例如如下文献中有综述:Nilsson等,Ann.Rev.Biophys.Biomol.Struct.2005.34:91-118。术语“纯化多肽”指的

是从包含本发明多肽的混合物移除除了该多肽以外的化合物。该用于还表示从包含除了本发明多肽以外的化合物的混合物移出该多肽。在特定的方面中,所述术语指从其不具有蛋白水解活性的前体分离该具有蛋白水解活性的多肽。

[0033] 核酸可在细胞或无细胞系统中被翻译。技术人员可采用多种无细胞翻译系统。本发明涵盖,例如,在无细胞蛋白质翻译系统中翻译包含兔网状红细胞裂解物、小麦胚芽裂解物、大肠杆菌裂解物或其它细胞裂解物,例如产自肉毒芽孢梭菌的裂解物等。本发明还涵盖从本发明的核苷酸序列或本发明的载体翻译本发明的多肽。转录可通过一种或多种异源性调节元件或同源性调节元件调节或控制。本发明的该方面还涵盖在野生型细胞中翻译,即从天然分离的细胞,例如肉毒芽孢梭菌、酪酸芽孢梭菌、巴拉提尼芽孢梭菌及破伤风芽孢梭菌。在一个特定方面中,所述细胞是肉毒芽孢梭菌的Hall菌株(ATCC 3502)。多种标准手段和方法均可被本领域技术人员所采用,以将核酸分子或载体带入细胞中并在细胞中将本发明的多肽表达为重组蛋白。此外,技术人员应知晓许多从细胞或细胞裂解物或从无细胞表达系统分离多肽的标准技术(例如,Recombinant DNA Principles and Methodologies(《重组DNA原理和方法》),J.Green,马塞尔-德克公司(Marcel Dekker Inc.),1998;The Condensed Protocols:From Molecular Cloning:A Laboratory Manual(《精选方案:来自分子克隆:实验室手册》),Sambrook等,Cold Spring Harbor Laboratory(冷泉港实验室出版社),2006;Molecular Cloning:A Laboratory Manual(《分子克隆:实验室手册》),Sambrook等,Cold Spring Harbor Laboratory(冷泉港实验室出版社),2000)。这些任何手段和方法均可用于本发明方法中。

[0034] 本发明的第一多肽可从编码所述具有蛋白水解型的多肽的核酸分子翻译。SEQ ID NO:26是该核酸分子的一个示例。或者,所述核酸分子可编码一种不具有蛋白水解活性的前体多肽,但该前体多肽可被转化成本发明的具有蛋白水解活性的多肽。SEQ ID NO:27是所述核酸分子的示例。所述不具蛋白水解活性的前体也被称为“无活性的BoNT水解酶”,简称iBH。该蛋白水解活性的多肽可在例如翻译期间或翻译后活化,或通过例如使所述不具蛋白水解活性的多肽与能移除该不具蛋白水解活性的多肽的N端的灭活氨基酸残基的蛋白酶接触。不具蛋白水解活性的多肽的一个示例是由SEQ ID NO:2所示的多肽。另一个示例是包含与SEQ ID NO:2的序列具有至少50%的序列相同性的多肽序列的多肽。在一个方面中,术语“N端的灭活氨基酸残基”涉及该多肽的前248aa残基。在另一个方面,该术语指所示多肽的至多10aa、50aa、100aa、150aa、200aa、250aa个残基的片段。任何这些多肽均可用于本发明的方法以制备具有蛋白水解活性的多肽。在一个方面中,能从此多肽的N端移除灭活氨基酸残基的蛋白酶分离自例如肉毒芽孢杆菌、酪酸芽孢梭菌、巴拉提尼芽孢梭菌和破伤风芽孢梭菌。在另一个方面,能移除所述灭活氨基酸的蛋白酶是通过提供该细胞的经分级分离的裂解物或未经分级分离的裂解物提供。灭活氨基酸残基可通过使所述不具蛋白水解活性的多肽与该裂解物接触并且培养直至该不具蛋白水解活性的多肽被转化成具有蛋白水解活性的多肽来被移除。

[0035] 在本发明方法的另一个方面,所述多肽在细胞中翻译。所述细胞可以是原核或真核细胞。在一个方面中,所述细胞选自大肠杆菌、枯草杆菌或酵母菌。本发明还涵盖在野生型细胞(即从天然分离的细胞)中翻译本发明的多肽,例如肉毒芽孢梭菌、酪酸芽孢梭菌、巴拉提尼芽孢梭菌和破伤风芽孢梭菌的任何已知的分离株。在具体方面中,所述细胞是肉毒

芽孢梭菌的Hall菌株(ATCC 3502)。在另一个具体方面中,所述细胞是如上所述的本发明的细胞。

[0036] 通过本发明的方法获得的翻译产物可通过多种手段纯化,所述手段均为本领域技术人员所知(例如,Recombinant DNA Principles and Methodologies(《重组DNA原理和方法》),J.Green,马塞尔-德克公司(Marcel Dekker Inc.),1998;The Condensed Protocols:From Molecular Cloning:A Laboratory Manual(《精选方案:来自分子克隆:实验室手册》),Sambrook等,Cold Spring Harbor Laboratory(冷泉港实验室出版社),2006;Molecular Cloning:A Laboratory Manual(《分子克隆:实验室手册》),Sambrook等,Cold Spring Harbor Laboratory(冷泉港实验室出版社),2000)。纯化本发明多肽的典型方法可包括离心细胞裂解物、硫酸铵沉淀蛋白质、重悬蛋白质、离心经重悬的蛋白质、离子交换色谱、分子排阻色谱、疏水性相互作用色谱等。所述步骤的若干不同顺序的结合可被用于纯化本发明的多肽。纯化本发明的多肽的优选方法描述于本发明的实施例中。

[0037] 在一个方面中,所述纯化步骤包括使本发明的多肽与固体支持物结合。术语“固体支持物”指的是一种基质,该基质包含例如硅石、交联葡聚糖、交联聚丙烯酰胺或交联琼脂糖等。还具体包括多肽、玻璃、聚苯乙烯、聚丙烯、聚乙烯、聚乙二醇(PEG)、葡聚糖、尼龙、淀粉酶、天然或经修饰的纤维素、聚丙烯酰胺、辉长岩和磁铁矿。在本发明的一个方面中,固体支持物是选自下组的多糖基质:琼脂糖凝胶(sepharose)、葡聚糖凝胶、琼脂糖、葡聚糖纤维素(sephacell)、微纤维素和藻酸盐珠。在另一个方面,所述固体支持物可由玻璃珠和/或多肽基质组成。

[0038] 在一个方面中,所述固体支持物连接至本发明的抗体。在一个方面中,术语“连接”指稳定连接或稳定联接。在另一个方面,连接包括例如间接或直接、非可逆或可逆、物理化学、静电和/或共价键相互作用。在一个方面中,所述抗体直接或通过接头分子共价连接至所述固体支持物。所述抗体可通过接头与该固体支持物连接,所述接头包括小分子化合物和肽(或多肽)接头分子。所述固体支持物可具有几乎任何可能的结构构型或安排,只要该经耦合的抗体能与其抗原结合即可。因此,所述基质或固体支持物可以是球状(如珠)或圆柱状(如试管内侧表面或棒杆外侧表面)。或者,所述表面可以是不规则或平面,例如薄板或测试条。

[0039] 与固体支持物连接的所述抗体可被用于例如本发明的制备方法或诊断方法中。在一个方面中,所述制备方法可包括亲和色谱的步骤,其中所述亲和色谱基于与固体支持物连接的抗体。在一个实施方式中,所述抗体是与本发明的具有蛋白水解活性的多肽特异性结合的抗体。在其它实施方式中,所述可能固体示于本发明的不具有蛋白水解活性的多肽特异性结合的抗体。

[0040] 在另一个方面,用于制备本发明的具有蛋白水解活性的多肽的方法包括:从包含额外成分的混合物纯化本发明的多肽。纯化可基于例如极性、电荷及大小。因此,所述方法在一个方面中可包括选自下组的一个或多个分离步骤:正相HPLC、逆相HPLC、亲水性相互作用色谱(HILIC)、疏水性相互作用色谱(HIC)、离子交换色谱(IEC)(包括阴离子交换色谱和阳离子交换色谱)、尺寸排阻色谱(SEC)、凝胶渗透色谱(GPC)。

[0041] 在另一个方面,所述纯化包括如下步骤:(a)通过离子交换色谱分离;(b)通过尺寸排阻色谱分离;(c)通过疏水性相互作用色谱分离;和(d)通过尺寸排阻色谱分离。

[0042] 从色谱柱收集的一种或多种组分可通过例如沉淀或超过滤方式浓缩。

[0043] 在一个方面中,本发明涉及包含本发明的具有蛋白水解活性的多肽的组合物。采用本文所述的方法,能够制备本发明的具有蛋白水解活性的多肽,其基本不含不具蛋白水解活性的多肽。换言之,本发明的方法提供具有蛋白水解活性的多肽和基本不含本发明的多肽的无活性前体蛋白污染的组合物。认为组合物基本无污染或基本不含不具蛋白水解活性的前体多肽,如果采用基于western印迹的检测方法,可检测到低于5%的不具蛋白水解活性的前体,其中所述5%指的是不具蛋白水解活性的前体相对于具有蛋白水解活性与不具活性的多肽总量的量。在另一个方面,所述组合物基本纯并且包含至少50%的本发明的具有蛋白水解活性的多肽,其中所述50%指的是具有蛋白水解活性的前体相对于所述组合物中所含蛋白质总量的量。在另一个方面,所述基本纯的组合物包含至少75%、80%、90%或至少98%的具有蛋白水解活性的多肽。

[0044] 在另一个方面,本发明还涉及可从上述及实施例中所述的制备具有蛋白水解活性的多肽的方法获得的多肽。在一个方面中,所述具有蛋白水解活性的多肽是含有SEQ ID NO:1的多肽序列的具有蛋白水解活性的多肽。在另一个方面,所述具有蛋白水解活性的多肽是与SEQ ID NO:1的序列具有至少50%的序列相同性的多肽。而在另一个方面中,所述具有蛋白水解活性的多肽是包含与SEQ ID NO:1的序列具有至少50%的序列相同性的多肽序列的多肽。本文所用的术语“可获得的多肽”在一个方面中指的是从本发明的核酸翻译的多肽。所述多肽随后可进行翻译后修饰,例如酰基化、烷基化、酰胺化、氨基酸添加、氨基酸删除、糖基化、氧化、S-谷胱甘肽化、磷酸化、硫酸化、蛋白分解处理等。此外,所述多肽可与金属离子结合,例如 $\text{Li}^+$ 、 $\text{Na}^+$ 、 $\text{K}^+$ 、 $\text{Ag}^+$ 、 $\text{Cs}^+$ 、 $\text{Mg}^{2+}$ 、 $\text{Ca}^{2+}$ 、 $\text{Co}^{2+}$ 、 $\text{Ni}^{2+}$ 、 $\text{Mn}^{2+}$ 、 $\text{Cu}^{2+}$ 或 $\text{Zn}^{2+}$ 。优选地,所述金属离子是 $\text{Zn}^{2+}$ 、 $\text{Mn}^{2+}$ 或 $\text{Co}^{2+}$ 。

[0045] 本发明在一个方面中还涉及与本发明的多肽特异性结合的抗体。本文所用的术语“抗体”涵盖单克隆抗体、多克隆抗体、单链抗体、人抗体、人化抗体、灵长类动物化抗体或嵌合抗体、双特异性抗体、合成抗体、经化学或酶修饰的衍生物、任何所述抗体的片段或由天然存在和/或化学修饰的核酸组成的适体。所述抗体的片段包括 $\text{F}(\text{ab}')_2$ 、 $\text{F}(\text{ab})$ 、 $\text{Fv}$ 或 $\text{scFv}$ 片段或任何所述片段的经化学或酶修饰的衍生物。

[0046] 在一个方面中,本发明的抗体将与本发明的具有蛋白水解活性的多肽或其不具蛋白水解活性的前体特异性结合。在一个方面中,对本发明的不具蛋白水解活性的多肽具有特异性的抗体与本文所述的不具蛋白水解活性的多肽交叉反应。在另一个方面,所述抗体能区分本发明的具有蛋白水解活性的多肽和其不具活性的前体。在另一方面中,所述抗体特异性的表位位于所述不具蛋白水解活性的多肽的氨基酸区域,但不存在于具有蛋白水解活性的多肽。例如,所述表位可以由多肽的氨基酸残基1~248所组成的多肽区域的表位,所述多肽包含与SEQ ID NO:2的序列具有至少50%的序列相同性的多肽序列。

[0047] 在另一个方面,所述表位由位于多肽的氨基酸248的N端的氨基酸残基形成,该多肽包含与SEQ ID NO:2的序列具有至少50%的序列相同性的多肽序列。在另一个方面,所述表位通过蛋白水解处理以从本文所述的不具蛋白水解活性的多肽移除。

[0048] 在另一个方面,本发明的抗体特异性的表位是位于多肽N端的表位,该多肽包含与SEQ ID NO:1的序列具有至少50%的序列相同性的多肽序列。在本发明的一个方面中所用术语“N端”指的是包含该多肽序列的N端50个氨基酸残基的多肽区域,优选地该多肽序列的

N端25个氨基酸残基。在具体方面中,该术语指的是N端14个氨基酸残基。本文所用的术语“表位”涉及被本发明的抗体所识别的抗原决定簇。在一个方面中,所述表位是线性表位,在另一方面中,所述表位是构象表位。在具体方面中,所述抗原决定簇由本发明的具有蛋白水解活性的多肽的N端的氨基酸序列的肽组成,其中所述肽可具有7~14个氨基酸长度,优选地,8、9、10、11、12、13或14个氨基酸残基。

[0049] 在一个方面中,术语“特异性结合”或“特异性结合至”指的是本发明的抗体不与本发明的多肽上的其它表位或其它一般多肽以显著程度交叉反应。表位特异性是本发明抗体的重要特征。所述抗体对于具有蛋白水解活性和不具蛋白水解活性的多肽的特异性,在一个方面中应是至少95%、至少96%、至少97%、至少98%、至少99%。特异性结合可通过多种熟知的技术测定,包括例如竞争试验。另一个重要特征是抗体的敏感性。在本发明的一个方面中,敏感性是使样品包含的至少70%、至少80%、至少90%、至少95%的表位被结合。敏感性可通过熟知技术测定。本领域技术人员将能够通过常规实验确定各相测定的操作和理想的测试条件。用于结合研究的传统技术包括放射性免疫试验、ELISA、平衡透析、恒温微量热计量法、**BIACORE®** 试验(表面等离子共振,SPR)或其它表面吸附方法。所述**BIACORE®** SPR系统检测抗体-抗原相互作用。SPR反应反映当分析物结合或解离时,在检测器表面的质量浓度的改变。基于SPR,实时**BIACORE®**测量结果直接监测即时发生的相互作用,参见BIAapplications Handbook(《BIA应用手册》),AB版(1998重印),**BIACORE®**编号:BR-1001-86;BIATEchnology Handbook(《BIA技术手册》),AB版(1998重印),**BIACORE®**编号:BR-1001-84。结合特性(例如本发明的抗体的敏感性)原则上可通过例如用存在于感测器表面的固体抗原(配体)的结合分析测定。受测抗体(分析物)将以流动相(即溶液)提供。在一些情况中,抗原通过与另一种称为捕捉分子的固定分子结合,而间接地连接至该表面。当抗体以离散型脉冲注射流过具有固定抗原的表面时,基本上可细分出三相:(i)抗体与抗原在样品注射期间结合;(ii)样品注射期间的平衡或稳定状态,其中抗体结合的速率与从该抗体-抗原复合物解离的速率平衡;(iii)在缓冲液流动期间抗体从表面解离。应理解,该分析可替代性地通过固定所需测试的抗体和使用包含抗原的溶液作为流动相来实施。所述结合和解离相提供分析物-配体相互作用的动力学资料( $k_a$ 和 $k_d$ ,复合物形成和解离的速率, $k_d/k_a=K_D$ )。平衡相提供分析物-配体相互作用的亲和性的资料( $K_D$ )。在本发明的一个方面中,本发明的抗体具有低于 $0.5\mu\text{M}$ 的 $K_D$ ,在另一个方面中,低于 $0.05\mu\text{M}$ 的 $K_D$ ,而在另一个方面中,低于 $0.02\mu\text{M}$ 的 $K_D$ 。

[0050] 本发明所述的抗体可通过使用例如Harlow和Lane,1988(Harlow和Lane,“Antibodies,A Laboratory Manual(《抗体,实验室手册》)”,CSH印刷厂,冷泉港出版公司(Cold Spring Harbor),1988)中所述的方法制造。单克隆抗体可通过最初于Kohler和Milstein,1975(Kohler和Milstein 1975,Nature 256:495)以及Galfre和Milstein,1981(Galfre和Milstein 1981,Meth Enzymol 73:3)中所述的技术制备。所述技术包括融合小鼠骨髓瘤细胞和衍生自经免疫接种哺乳动物的脾细胞。抗体可进一步通过本领域中熟知的技术改善。例如,在**BIACORE®**系统使用的表面等离子共振可被用于增加与本发明的多肽内的前述表位结合的噬菌体抗体的效率(参见Schier等,1996,Human Antibodies Hybridomas 7:97;Malmberg等,1995,J.Immunol Methods 183:7)。

[0051] 在本发明的一个方面中,所述抗体通过使用包含前述表位或由前述表位组成的肽

生成。所述肽可例如通过合成制备,或通过重组表达制备。或者,本发明的抗体可通过采用天然存在的本发明的具有蛋白水解活性或不具有蛋白水解活性的多肽制备。在后这中,应理解,所形成的抗体将进一步测试其对本发明的多肽的特异性。在本发明的其它方面中,本发明的单克隆抗体通过采用本发明的多肽制备,其中本发明的多肽可通过去污剂处理以使所述表位为免疫原性可用。然而,应理解,如果该抗体将以构象表位为目标,则不应采用所述去污剂处理。在另一方面中,免疫刺激剂例如轮孔状帽贝血蓝素(KLH)也可应用于该方法中,特别是当采用合成肽时。

[0052] 本发明的抗体可被用于例如亲和性色谱、免疫沉淀和免疫定位本发明的多肽,以及用于监测该多肽在样品或重组有机体中的存在。此外,本发明的抗体可被用于检测方法或诊断方法中。在特定的方面中,本发明的抗体用于western印迹或ELISA中。此外,本发明的抗体可被用于治疗应用。具体而言,所述抗体可被用于抑制本发明的具有蛋白水解活性的多肽的活性。因此,本发明的抗体也具有如下所述的许多治疗应用。

[0053] 本发明还涉及本发明的具有蛋白水解活性的多肽在蛋白水解处理多肽的方法中的应用。在一个方面中,本发明涉及一种制备经蛋白水解处理的多肽的方法,所述方法包括使如下(a)与(b)接触的步骤:(a)第一多肽,所述第一多肽是本发明的多肽,和(b)第二多肽,所述第二多肽易被所述第一多肽蛋白水解,其中所述接触导致该第二多肽被蛋白水解处理成为至少两种切割产物。

[0054] 本发明还涉及来自产酶溶杆菌(*Lysobacter enzymogenes*) (ATCC 29487) d Lys-N和/或Lys-C和/或精胺酰基内肽酶(内蛋白酶Arg-C, LeR)的应用(Wright DS, Graham LD, Jennings PA. *Biochim Biophys Acta*. 1998年12月22日; 1443(3):369-74)。此外,本发明还涵盖胞浆素和/或OmpT蛋白(OmpT)(一种在(Arg/Lys)-(Arg/Lys)处切割基序的膜结合丝氨酸蛋白酶)(K. Sugimura和T. Nishihara. *J. Bacteriol.* 170(1988), 第5625-5632页)在蛋白水解处理CNT(例如BoNT/A)的方法中的应用。在一个方面中,本发明涉及一种用于制备经蛋白水解处理的多肽的方法,所述方法包括使(a)与(b)接触的步骤:(a)第一多肽,所述第一多肽是Lys-C或Lys-N,和(b)第二多肽,所述第二多肽易被所述第一多肽蛋白水解,其中所述接触导致该第二多肽被蛋白水解处理成至少两种切割产物,并且其中所述第二多肽是BoNT/A的单链。术语“Lys-C”指的是特异性切割赖氨酸C端的肽键的源自产酶溶杆菌的33kDa丝氨酸内蛋白酶Lys-C(赖氨酰基内肽酶, LeK, Genbank登录号Q7M135)或其具有至少60%的序列相同性的同源物。术语“Lys-N”指的是分离自多叶奇果菌(*Grifola frondosa*)和糙皮侧耳(*Pleurotus ostreatus*)的金属内肽酶Lys-N(Nonaka T等, 1997, *J Biol Chem.* 272:30032-30039; Nonaka T等, 1998, *J Biochem.* 1998 124:157-162; Hori T等, 2001, *Acta Crystallogr D Biol Crystallogr.* 57:361-368)。该术语还涵盖所述蛋白酶的具有至少60%的序列相同性的同源物。

[0055] 该方法还可用于,例如,制备经蛋白水解处理的神经毒素(CNT)或肉毒神经毒素(BoNT)。本发明全文使用的术语“BoNT”,指的是肉毒神经毒素并且指可从肉毒芽孢梭菌获得的神经毒素,例如血清型A、B、C1、D、E、F或G的BoNT。术语“CNT”和“BoNT”也包含重组神经毒素且包含一种或多种修饰(包括化学修饰或遗传修饰)。术语“遗传修饰”指的是缺失、取代或添加一种或多种氨基酸残基。使用本发明的方法,能够获得受显著较少的未处理或部分处理的神经毒素污染的神经毒素组合物,因为这些污染物被有效地处理成为双链神经毒

素。在一个方面中,所述双链神经毒素是天然双链神经毒素,其中轻链的C端与重链的N端与从野生型芽孢梭菌分离的该对应的经完全处理的双链神经毒素一致。

[0056] 本文所用术语“接触”是指使至少两种不同的化合物物理接近以允许这些化合物之间发生物理和/或化学相互作用。根据本发明方法,在一个方面中,所述两种不同的化合物是由溶液包含的第一多肽和第二多肽。接触在足以允许第一多肽和第二多肽发生相互作用的条件和时间下进行。本文所用的术语“经蛋白水解处理的多肽”在一个方面中指的是一种多肽,其多肽链的一个或多个肽键已经被水解或切割。在另一个方面,该术语指已被内蛋白酶或内肽酶蛋白水解切割的多肽。在另一个方面,该术语指已被切割至少50%的程度的多肽。在另一个方面,所述经蛋白水解处理的多肽是第二多肽。在另一个方面,至少60%、70%、80%、90%或95%经蛋白水解处理。

[0057] 本文所用的术语“第一多肽”指的是本发明的多肽,即具有蛋白水解活性或活化的多肽,又称为“活性BoNT水解酶”。因为该活性BoNT水解酶可从肉毒芽孢梭菌的上清液获得,其最初被命名为天然BoNT水解酶,简称“nBH”。然而,术语“第一多肽”和“nBH”也指可从其它来源获得的BoNT水解酶。本文所用术语“第二多肽”指的是所述第一多肽的底物。术语“易被蛋白水解”指的是所述第二多肽的特征或条件,并且用于本文中指所述第二多肽可被所述第一多肽蛋白水解切割。换言之,术语“易被蛋白水解”指的是第二多肽包含允许其作为第一多肽的底物的蛋白酶识别和切割位点。所述“第二多肽”是第一多肽的底物,并且经蛋白水解处理成两种或多种切割产物。采用上述试验,技术人员能够根据本发明的定义测定给定的多肽是否是第一多肽的底物,即,“第二多肽”。术语“至少两种切割产物”包括,例如,多至2、3、4、5和多至6种切割产物。

[0058] 该方法可被用于例如制备包含芽孢梭菌神经毒素的药物组合物或用于生成质谱法中所用的多肽片段。所述第一多肽和第二多肽可在经蛋白水解处理的多肽的制备方法中的不同步骤接触。在一个方面中,使所述第一多肽和第二多肽接触的步骤在细胞内进行。在该实施方式的一个具体方面中,所述第一多肽和第二多肽在所述细胞内表达。

[0059] 在另一个方面,所述接触步骤在细胞裂解物或经纯化的细胞裂解物中进行。该方面涵盖添加第一多肽至所述裂解物或经纯化的裂解物。所述第一多肽可在从细胞裂解物纯化第二多肽的过程中的不同步骤添加。例如,第一多肽可在下列步骤之前或之后添加:蛋白质沉淀、离子交换色谱、疏水性相互作用色谱和/或尺寸排阻色谱。此外,还涵盖添加第一多肽至药物组合物。在该方面,本发明的多肽用于,例如,蛋白水解切割第二多肽,以例如使作为药物组合物中所含的治疗剂的第二多肽活化。还设想向对象给予第一多肽,以蛋白水解处理该对象内的第二多肽。给予还包括共同给予所述第一多肽和第二多肽。该方法还包括在足以切割所述第二多肽的条件和时间下进行孵育的步骤。在一个方面中,所述条件可包括添加选自下组的缓冲液:100mM Tris-HCl, pH 8.0或PBS (50mM Na<sub>2</sub>HPO<sub>4</sub>, 150mM NaCl, pH 7.4)。优选的缓冲液条件包括100mM Tris-HCl, pH 8.0。“足以切割的时间”可利用上述试验测定。在一个方面中,所述“足以切割的时间”取决于经蛋白水解处理的多肽或包含其的组合物应具有切割程度。在一个方面中,所述方法包括孵育第一多肽和第二多肽至少30分钟、60分钟、120分钟或至少240分钟的步骤。在另一个方面,所述第一多肽和第二多肽孵育多至30分钟、60分钟、120分钟、240分钟、480分钟或多至600分钟。在另一个方面,所述方法包括在4℃或37℃下孵育第一和第二多肽的步骤。在另一个方面,所述方法包括孵育多至1

小时、多至2小时、4小时、6小时、10小时或多至16小时的步骤。

[0060] 在一个方面中,所述第二多肽的多肽链包含选自SEQ ID NO:4~25中任一条的序列。在更具体的方面,所述第二多肽的多肽链包含选自SEQ ID NO:4~25中任一条的序列,并且其中所述第二多肽在SEQ ID NO:4~25中任一条的所述序列内的碱性氨基酸残基的C端被切割。所述序列代表本发明的具有蛋白水解活性的多肽的已知底物的氨基酸序列。如本文中所示,所述底物在序列所包含的碱性氨基酸残基的C端被切割,比较表1中的LC栏和H<sub>N</sub>栏。在一个优选方面中,所述第二多肽包含选自SEQ ID NO:4~10的序列。在另一个优选方面,所述第二多肽是BoNT/A或其衍生物,包括例如SEQ ID NO:3的多肽或其衍生物。关于本发明的该方面和其它方面所用的术语“衍生物”包含氨基酸突变例如添加、取代、缺失或截短一个或多个氨基酸残基。

[0061] 在一个方面中,所述第二多肽包含SEQ ID NO:4~25中任一条或SEQ ID NO:3的衍生物,其中所述衍生物具有一个或多个点突变和/或一个或多个额外氨基酸残基。在另一个方面,所述衍生物具有多至1、多至2、多至3、多至4、多至5、多至6、多至7、多至8、多至9、多至10、多至15个点突变。通过采用如本文所述的测定蛋白酶活性的活性试验,技术人员能够测定给定衍生物是否经本发明的具有蛋白水解活性的多肽处理。在另一个方面,所述衍生物包含将碱性氨基酸变为非碱性氨基酸残基的点突变。在另一个方面,所述衍生物与SEQ ID NO:4~25中任一条具有至少50%的序列相同性。在另一个方面,所述衍生物或包含所述衍生物的多肽是第一多肽的底物,并且可被第一多肽蛋白水解切割。典型的示例是SEQ ID NO:3的衍生物,其包含例如轻链或重链中的一个或多个点突变。

[0062] 在另一个方面,所述第二多肽包含(a)与SEQ ID NO:3的序列具有至少30%的序列相同性的多肽序列[ (ATCC 3502的BoNT/A, Genbank登录号.AAA23262) ];或(b)选自下组的多肽序列:破伤风神经毒素、凝血级联因子X或凝血酶原(因子II)的蛋白、胰消化酶(如胰蛋白酶、胰凝乳酶)、胃蛋白酶、木瓜酶。至少30%指至少30%、至少40%、至少50%、至少85%。在一个具体方面中,与SEQ ID NO:3的序列具有至少50%的序列相同性的所述第二多肽序列的序列相同性基于SEQ ID NO:3的第420~466位氨基酸来确定,在另一个方面,所述序列相同性基于SEQ ID NO:4~25中任一条来确定。换言之,所述方面指包含多肽序列的第二多肽,所述多肽序列与SEQ ID NO:3的第420~466位氨基酸的多肽序列具有例如至少30%序列相同性。基于该定义的多肽可获自例如肉毒芽孢梭菌、破伤风芽孢梭菌或产芽孢梭菌。所述第二多肽可以是,例如,天然存在的神经毒素(例如BoNT/A、B、C1、D、E、F或G)或其包含一个或多个氨基酸突变(例如添加、取代、删除或截短一个或多个氨基酸残基)的衍生物。涵盖例如缺失例如天然神经毒素H<sub>c</sub>结构域或其部分的衍生物,或具有其它取代神经毒素H<sub>c</sub>结构域的氨基酸残基的衍生物和具有BoNT的额外轻链或与轻链N端融合的另一蛋白货物(cargo)分子的衍生物。

[0063] 在另一个方面,所述第二多肽可包含在N端或C端或内部位置的额外氨基酸残基。所述额外氨基酸残基可侧接一个或多个蛋白酶切割位点。在另一个方面,所述额外的氨基酸序列作为可检测标签和/或允许与个体支持物结合。一个示例是his标签或GST标签。另一个示例是包含Strep标签的氨基酸序列VPPTPGSAWHPQFEK,优选被添加至C端。

[0064] 在一个具体方面,所述第二多肽是包含如GenBank编号CBZ04958.1、YP\_002805603.1、ZP\_02994746.1、YP\_001788403.1、YP\_001782718.1、ZP\_02616437.1、ZP\_

02614241.1、YP\_001392361.1、YP\_001255575.1所示的多肽序列的多肽或其具有至少50%的序列相同性的同源物。

[0065] 在另一个方面,所述第二多肽的生物活性由蛋白水解切割调节。技术人员熟知的是,多种多肽的功能可通过蛋白水解处理来调节。本文中所用的“调节”是指增加或减少、活化或失活化。例如,多种芽孢梭菌神经毒素的生物活性通过蛋白水解将单链神经毒素处理成双链神经毒素来增加或引发,其中所述双链神经毒素由多肽轻链和重链组成,其通过二硫键共价连接。所述神经毒素的生物活性涵盖至少三种不同活性:第一活性是位于该神经毒素的轻链中的“蛋白水解活性”,并且负责水解涉及调节细胞膜融合的一种或多种多肽的肽键。第二活性是位于经处理的神经毒素的重链N端的“翻译活性”,其涉及运输轻链通过溶酶体膜进入细胞质。第三活性是位于经处理的神经毒素的重链C端的“受体结合活性”,其涉及神经毒素和靶细胞的结合以及靶细胞对神经毒素的摄取。在一个优选方面,本文所用的术语生物活性指的是蛋白水解活性。在更优选的方面中,该术语指增加蛋白水解活性。

[0066] 芽孢梭菌神经毒素的生物活性可通过多种测试来检测,其全部为本领域技术人员已知。这些测试允许测定上述一种或多种活性。例如,小鼠LD<sub>50</sub>试验或立体小鼠膈神经半横膈(MPN)试验(如Pearce等于1994年所述(Pearce LB,Borodic GE,First ER,MacCallum RD (1994),*Toxicol Appl Pharmacol* 128:69-77)和Habermann等于1980年所述(Habermann E,Dreyer F,Bigalke H.(1980),*Naunyn Schmiedebergs Arch Pharmacol*.311:33-40))允许测定给定的神经毒素制剂对于活生物体或经分离的神经肌肉组织的毒性效应。为了建立LD<sub>50</sub>试验中的毒性效应,神经毒素必需具备上述三种活性中的各种生物活性。此外,还有许多其它试验可用于测定例如神经毒素或该神经毒素的轻链是否具有蛋白水解活性。此类试验基于例如使BoNT/A接触SNAP-25。或者,可采用代表SNAP-25的切割位点的肽,其中该肽可经标记以便于检测。在优选方面中,生物活性通过采用如上所述的MPN试验测定。

[0067] 在另一个方面,所述第一多肽通过蛋白水解处理无活性的前体多肽来活化,所述无活性的前体多肽包含与SEQ ID NO:2的序列具有至少60%序列相同性的多肽序列。该方面基于具有SEQ ID NO:2的多肽序列的多肽不具蛋白水解活性,而其N端截短体具有蛋白水解活性的观察结果。本发明还设想采用本文所述的方法中的不具蛋白水解活性的多肽。本文所述的不具蛋白水解活性的多肽可通过例如移除N端片段或移除包含SEQ ID NO:2的氨基酸残基1~248的整个N端来活化。在一个方面中,N端通过蛋白酶移除,在另一个方面,N端通过SEQ ID NO:2的自体蛋白水解移除。60%的序列相同性指与全长NT02CB1447进行序列比对。

[0068] 在另一个方面,本发明的制备经蛋白水解处理的多肽的方法包括,纯化该经蛋白水解处理的第二多肽或其至少一种或两种或更多种切割产物的步骤。肉毒芽孢梭菌表达的BoNT/A的纯化基本上可如现有技术(DasGupta 1984,*Toxicon* 22,415;Sathyamoorthy 1985,*J Biol Chemistry* 260,10461)所述的方法进行。具体而言,神经毒素的纯化可包括一个或多个沉淀和萃取步骤、一个或多个浓缩步骤以及其它不同的色谱步骤。重组单链BoNT/A及其纯化描述于现有技术(Rummel等,2004,*Mol Microbiol*.51:631-43)。

[0069] 在一个优选实施方式中,芽孢梭菌菌株是例如生成BoNT/A或其衍生物的肉毒芽孢梭菌。就发酵而言,可采用由DasGupta B.R.等在*Toxicon*(《毒素》),第22卷第3期第414~424页(1984)所述的方法。因此,将0.5%的酵母萃取物和0.6%的高压灭菌酵母糊料添加至

2%的N-Z-胺A型培养基,利用4N NaOH调节pH至7.2,以此方式制备培养基之后进行高压灭菌。可向该培养基添加另经高压灭菌的葡萄糖(每体积重量20%),以使培养基中的葡萄糖终浓度达到0.5%。孵育可在例如37°C下进行(不搅拌),其中发酵在例如96小时之后终止。在本发明的范围内,除了前述的批式发酵以外,也可进行半批式发酵、重复批式发酵或连续发酵。

[0070] 在实际发酵和分离发酵培养基和细胞之后,该发酵培养基可进行一次沉淀以移除大型蛋白质。沉淀优选是酸沉淀。所述酸沉淀的反应条件为本领域技术人员已知。通常可采用1.5M H<sub>2</sub>SO<sub>4</sub>,以使上清液酸化至pH 3.5。离心通常以2400x g于4°C下进行20分钟。离心获得的沉淀可用水清洗,优选重复清洗。随后,沉淀可用0.1M的柠檬酸-柠檬酸三钠缓冲液(pH 5.5)萃取例如1小时。随后,可进行其它离心步骤,例如以9800x g在4°C下离心20分钟。如此获得的沉淀随后可选择地再次如上所述萃取。萃取的上清液和重复萃取的二次上清液随后可经硫酸鱼精蛋白沉淀。沉淀可在例如8°C下持续过夜。随后,该沉淀物可在4°C下以12,000x g离心20分钟。离心的上清液可经沉淀(例如硫酸铵沉淀)处理,由此移除其它较大蛋白质。在硫酸铵沉淀步骤之后可增加另一离心步骤,接着所获的沉淀可被重新溶解并可任选地经透析处理。优选地,经再次透析和离心的萃取物可经连续色谱步骤处理以纯化神经毒素。各色谱步骤用于移除污染物,例如硫酸鱼精蛋白、剩余的DNA、较小蛋白质以及中型蛋白质的部分,以及肉毒神经毒素蛋白复合物的血细胞凝集素。出于该目的,优选的实施方式中可采用一个或多个色谱步骤。任选地,例如最后的色谱步骤的洗出液可经过滤以减少细菌。任选地,洗出液在过滤之前可经稀释并添加适当佐剂。在其它步骤过程中,另一灭菌过滤步骤可在添加佐剂之后进行。在一个方面中,过滤在反应容器中进行,该容器随后可进行冷冻干燥步骤。

[0071] 本发明还涉及一种组合物,所述组合物可通过本发明的制备经蛋白水解处理的多肽的方法获得。在一个方面中,所述组合物包含经处理和未经处理的第二多肽的混合物,其中所述混合物可包含低于5%、4%、3%、2%或低于1%的未经处理的第二多肽。在所述组合物的一个方面中,所述第二多肽是BoNT或其衍生物。所述BoNT可以例如选自下组:BoNT的血清型A、B、C、D、E、F和G,包括其衍生物。所述组合物可以是例如液体或固体组合物并且可包含一种或多种载体、佐剂和/或赋形剂。

[0072] 在另一个方面,本发明还涉及一种制备药物(药物组合物)的方法,所述方法包括前述方法的步骤和配制经纯化的双链神经毒素作为药物的其它步骤。在一个方面中,所述药物包含经处理和未经处理的第二多肽的混合物,其中所述混合物可包含低于5%的未经处理的第二多肽。在优选实施方式中,所述混合物包含低于4%、3%、2%或低于1%的未处理的第二多肽。

[0073] 本发明还涉及本文所述的化合物的多种医学用途:

[0074] 在一个方面中,本发明涉及用作药物或用于药物组合物中的本发明的具有蛋白水解活性的多肽。

[0075] 在另一个方面,本发明涉及用作药物或用于药物组合物中的本发明的组合物。

[0076] 而在另一个方面,本发明涉及用作药物或用于药物组合物中的本发明的抗体。

[0077] 而在另一个方面,本发明涉及用作药物或用于药物组合物中的本发明的抑制剂。

[0078] 具体而言,本发明涉及包含本发明的多肽、本发明的抗体、本发明的组合物或本发

明的抑制剂的药物组合物。

[0079] 本文所用术语“组合物”指的是任何经配制为固体、液体、气雾剂(或气体)形式等的组合物。所述组合物包含例如本发明的治疗活性化合物和任选的合适辅助化合物,例如稀释剂或载体或其它成分。在一个方面中,治疗活性化合物是本发明的具有蛋白水解活性的多肽。在另一个方面,治疗性化合物是上文所述的经蛋白水解处理的第二多肽,例如双链神经毒素。在另一个方面,所述治疗活性化合物是本发明的抗体。在另一个方面,所述治疗活性化合物是本发明的具有蛋白水解活性的多肽的抑制剂。

[0080] 本文中,本发明区分辅助化合物和其它成分,所谓辅助化合物即是为达所需目的施用组合物时不造成由本发明的化合物引发的效应的化合物,而其它成分是指造成其它效应或调节本发明的化合物的效应的化合物。合适的稀释剂和/或载体取决于所要采用该组合物和其它成分的目的。本领域技术人员能够确定该合适的稀释剂和/或载体而不需赘述。

[0081] 载体必须是可与制剂的其它成分相容且对其受者无害。所用药物载体可包括固体、凝胶或液体。示例性的固体载体有乳糖、白土、蔗糖、滑石、明胶、琼脂、果胶、阿拉伯胶、硬脂酸镁、硬脂酸等。液体载体的示例有磷酸缓冲盐水溶液、糖浆、油、水、乳液、多种类型的润湿剂等。类似地,载体或稀释剂可包括本领域熟知的延迟性材料,例如单独的甘油单硬脂酸酯或甘油二硬脂酸酯或与蜡组合。所述合适的载体包括上文所述那些以及本领域熟知的其它载体,参见,例如,Remington's Pharmaceutical Sciences(《雷明顿药物科学》),默克出版公司(Mack Publishing Company),宾夕法尼亚州伊斯顿。

[0082] 选择稀释剂以不影响所述组合的生物活性为准。所述稀释剂的示例有蒸馏水、生理盐水、林格(Ringer's)溶液、右旋糖溶液和汉氏(Hank's)溶液,此外,所述药物组合物或制剂还可包含其它载体、佐剂或非毒性、非治疗性、非免疫原性的稳定剂等。

[0083] 在一个方面中,本文所用的药物组合物包含通过本发明的方法获得的具有生物活性的神经毒素,任选地包含一种或多种药学上可接受的载体。所述具有活性的神经毒素可以液体或冻干形式存在。在一个方面中,所述化合物可与甘油、蛋白质稳定剂(例如,人血清白蛋白(HSA))或非蛋白质稳定剂(例如聚乙烯吡咯烷酮或玻糖醛酸)一起存在。在一个方面中,所述药物组合物局部给予。传统应用的药物给予是肌肉内、皮下(接近腺体)给予。然而,根据化合物的性质和作用模式,所述药物组合物也可通过其它途径给予。双链神经毒素多肽是该组合物的活性成分,在一个方面中根据传统方法通过组合该药物和标准药物载体所制备的传统剂型给予。这些方法可能涉及混合、造粒和压制或溶解该合适成分至该所需的制剂。应理解,所述药学上可接受的载体或稀释剂的形式和特征受其待组合的活性成分的量、给予途径和其它为人熟知的变化因素的影响。

[0084] 治疗有效剂量是指用以预防、改善或治疗伴随本说明书中所述的疾病或病症的症状的待用于本发明的药物组合物中的神经毒素化合物的量。化合物的治疗性疗效和毒性可通过在细胞培养物或实验动物中进行的标准药学方法来确定,例如ED<sub>50</sub>(对50%的群体有治疗效果的剂量)和LD<sub>50</sub>(使50%的群体致死的剂量)。治疗效应与毒性效应之间的剂量比即为治疗指数,其可由LD<sub>50</sub>/ED<sub>50</sub>的比表示。

[0085] 给药方案将由主治医师和其它临床因素决定。如医药领域所熟知,用于任何患者的剂量取决于多种因素,包括该患者的体型大小、身体表面积、年龄、待给予的特定化合物、性别、给予时间和途径、整体健康和同时给予的其它药物。可通过定期检查来监测进展。本

文所述的药物组合物和制剂给予至少一次以治疗或改善或预防本说明书中所述的疾病或病症。然而,所述药物组合物可给予多于一次。

[0086] 在本发明的其它方面中,前述组合物是药物或美容组合物。在一个方面中,包含所述具有生物活性的神经毒素的药物可用于预防和/或治疗下列疾病或病症中的至少一种:主动肌强直、局部肌张力障碍(包括头、颈肌张力障碍和良性自发性眼睑痉挛、半边面部痉挛和局部痉挛)、胃肠疾病、多汗症和美容除皱,在另一方面中还包括眼睑痉挛、口下颌肌张力障碍、张口型、闭口型、磨牙、梅杰氏(Meige)综合征、舌肌张力障碍、眼睑肌肉运动失调、张口型颈肌张力障碍、颈项前屈、颈项后屈、颈项侧屈、斜颈、咽部肌张力障碍、喉部肌张力障碍、痉挛性发声障碍/内收肌型、痉挛性发声障碍/外展肌型、痉挛性呼吸困难、四肢肌张力障碍、手臂肌张力障碍、特定任务性肌张力障碍、书写痉挛、音乐家指痉挛、高尔夫手痉挛、腿部肌张力障碍、大腿内收、大腿外展膝关节弯曲、膝关节伸直、踝关节弯曲、踝关节伸直、马蹄内翻足、畸形足肌张力障碍、纹状趾、趾弯曲、趾伸直、中轴肌张力障碍、比萨综合征、肚皮舞者肌张力障碍、节段性肌张力障碍、半身性肌张力障碍、全身性肌张力障碍、lubag综合征中的肌张力障碍、皮质基底核退化中的肌张力障碍、lubag综合征中的肌张力障碍、迟缓性肌张力障碍、脊髓小脑失调症中的肌张力障碍、帕金森氏症中的肌张力障碍、亨廷顿氏舞蹈症中的肌张力障碍、髓素异常表象终端肌张力障碍、多巴诱导的运动障碍/多巴诱导的肌张力障碍、迟缓性运动障碍/迟缓性肌张力障碍、阵发性运动困难/肌张力障碍、运动型非运动性动作诱导的上颚肌阵挛、肌阵挛肌纤维颤动、僵硬、良性肌肉痉挛、遗传性下巴颤抖、矛盾性下巴肌活动、半侧咀嚼肌痉挛、肥厚性腮肌病、嚼肌肥厚、胫骨前肌肥厚、眼球震颤、振动幻视核上性凝视麻痹、癫痫、持续性不全癫痫、计划痉挛性斜颈操作、外展肌声带麻痹、反抗性突变烦躁不安、上食道括约肌功能不全、声带肉芽肿、口吃性妥瑞氏症、中耳肌阵挛、保护性喉闭合、喉切除术术后、语言障碍、保护性垂睑、眼睑内翻、奥迪氏(Odi)括约肌功能异常、假性食道迟缓不能、非迟缓不能性食道运动障碍、阴道痉挛、术后制动震颤、膀胱功能失常、膀胱尿道括约肌失常、膀胱括约肌痉挛、半侧脸痉挛、神经再生性运动困难、鱼尾纹的美容使用、皱眉脸不对称、颞肌皱缩、僵硬人症候群、破伤风性前列腺增生、过胖、治疗幼儿脑性麻痹斜视、混合麻痹性共同性斜视、视网膜剥离手术之后、白内障手术之后、无晶体眼中肌炎性斜视、肌病性斜视、分离性垂直眼偏斜、作为斜视手术的辅助治疗、内斜视、外斜视、迟缓不能、肛裂、外分泌腺活性过高、弗雷氏综合征、假哭性综合征、多汗症、腋窝掌足鼻漏、中风中的相对流涎过多、帕金森氏症中的相对流涎过多、肌萎缩性侧索硬化症、横贯性脊髓炎、戴维克氏(Devic)综合征、病毒感染、细菌感染、寄生虫感染、真菌感染、遗传性痉挛性截瘫中风后综合征脑半球梗塞中、脑干梗塞、脊椎梗塞、偏头痛、中枢神经系统外伤、脑半球病灶、脑干病灶、脊椎病灶、中枢神经系统出血中、脑内出血、蜘蛛膜下腔出血、硬膜下出血、脊椎内出血、在癌中、脑半球性肿瘤、脑干肿瘤、脊椎肿瘤、打鼾(WO 2000/033863)。详细信息和症状参见例如Jost 2007, Drugs 67 (5), 669或Dressler 2000刊于Botulinum Toxin Therapy(《肉毒菌毒素治疗》), 纽约斯图加特的Thieme Verlag出版公司。

[0087] 在本发明的另一个方面,所述组合物是可如上述药物组合物配制的美容组合物。同样对于美容组合物,设想本发明的化合物在一个方面中以基本纯的形式使用。美容组合

物在另一方面中经肌肉内给予。在本发明的甚至其它方面中,包含神经毒素的美容组合物可经配制为抗皱溶液。

[0088] 在另一个方面,所述药物组合物包含本发明的抗体或抑制剂。由于本发明的多肽负责活化芽孢梭菌神经毒素,该抗体可被用于减少芽孢梭菌感染时观察到的毒性作用。因此,本发明的抗体在一方面中可被用于治疗芽孢梭菌感染,包括产气荚膜芽孢梭菌、难养芽孢梭菌、破伤风芽孢梭菌、肉毒芽孢梭菌、巴拉提尼芽孢梭菌、酪酸芽孢梭菌、产芽孢梭菌、丙酮丁醇芽孢梭菌、溶血芽孢梭菌、诺维氏芽孢梭菌和水肿芽孢梭菌。另外,本发明的抗体在另一方面中可用于治疗与该感染有关的症状。此外,所述抗体可用于治疗与病症有关的病状或症状,其中所述病状选自肉毒杆菌症、破伤风、伪膜性结肠、坏疽、食物中毒等。

[0089] 在另一个方面,所述药物组合物包含本发明的具有蛋白水解活性的多肽。所述药物组合物在在一个方面中可用于蛋白水解切割与协同凝集有关的多肽,特别是治疗协同凝集功能不足的患者。在另一个方面,所述药物组合物可被用作纤维蛋白溶解(fibrinolyticum),尤其是治疗心肌梗塞、肺栓塞、深层静脉血栓栓塞的患者,即用于移除血栓。还设想采用药物组合物治疗中风。此外,在其它方面中,所述药物组合物可用于治疗外分泌胰功能不全以取代胰蛋白酶、胰凝乳酶和胃蛋白酶。此外,在其它方面中,所述药物组合物可用于治疗有炎症反应的患者,用于治疗癌患者,尤其是用于蛋白水解切割表面暴露的肿瘤抗原。另外,在另一方面中,所述药物组合物可用于治疗乳头状瘤。

[0090] 本发明还涉及一种筛选抑制剂的方法,所述方法包括如下步骤:(a)使本发明的具有蛋白水解活性的多肽与已知底物和任选的假定抑制剂接触;和(b)测定该假定抑制剂对底物转化成切割产物的影响,其中切割产物的量的减少表示该假定抑制剂有抑制作用。在一个方面中,该假定抑制剂是包含选自SEQ ID NO:4~25中任一条的氨基酸序列的肽,其中该氨基酸序列的至少一个碱性氨基酸由非碱性氨基酸取代。在其它方面中,所述肽包含一种或多种化学修饰。在另一个方面中,所述抑制剂是所述肽的拟肽。在另一个方面,所述假定抑制剂是化学化合物微阵列(即有机化学化合物的集合)的一部分。而在其它方面中,所述抑制剂是本发明的抗体。该方法可用于鉴别能抑制本发明的具有蛋白水解活性的多肽的化合物。初步筛选可基于例如包含选自SEQ ID NO:4~25中任一条的氨基酸序列的肽。能抑制本发明的多肽的肽可经修饰以增加抑制。修饰包括氨基酸取代或化学修饰。通常,该方法通过如下方式进行:使本发明的多肽与已知底物在假定抑制剂存在和不存在的情况下接触(该方法的步骤(a))并通过比较该假定抑制剂对底物转化为切割产物的影响。在假定抑制剂存在的情况下转化率减少,说明有抑制效应。

[0091] 本发明还涉及本发明的具有蛋白水解活性的多肽的抑制剂,其中所述抑制剂是(a)包含如SEQ ID NO:4~25中任一条所示的氨基酸序列的抑制剂,其中包含于其中的碱性氨基酸由非碱性氨基酸取代;或(b)本发明的抗体。

[0092] 本说明书中所引用的全部文献以引用其全部公开内容的方式纳入本文,并且在说明书中说明性地提及其公开内容。

[0093] 附图显示:

[0094] 图1:从HiPrep 16/10Q FF收集的组分的活性测试。

[0095] 从HiPrep 16/10Q FF跑动收集的5 $\mu$ l组分,通过在37 $^{\circ}$ C下与2 $\mu$ g scBoNTA(泳道2)孵育1小时并随后进行10%SDS-PAGE来分析酶活性。泳道1:低分子量标志物(LMW):116kDa、

66kDa、45kDa、35kDa。

[0096] 图2:通过12.5% SDS-PAGE分析从SEC (HiLoad 16/60Superdex 75)收集的组分中的nBH含量(分子量约37.3kDa)。组分9~11主要包含nBH。(泳道1:LMW:116kDa、66kDa、45kDa、35kDa、25kDa、18.4kDa、14.4kDa)

[0097] 图3:通过12.5% SDS-PAGE分析测定三个nBH纯化批量的纯度和蛋白质浓度。

[0098] 泳道1,LMW (116kDa、66kDa、45kDa、35kDa、25kDa、18.4kDa、14.4kDa);泳道2,nBH批号TE311206 (192ng/ $\mu$ l成熟NT02CB1446/CB01444,Genbank登录号CAL82987.1的氨基酸254-594,MW:38.6kDa);泳道3,nBH批号TIK301009 (130ng/ $\mu$ l成熟NT02CB1447/CB01445,SEQ ID NO:1,Genbank登录号CAL82988.1的氨基酸249-581,MW:37.3kDa);泳道4,nBH批号TIK280509 (114ng/ $\mu$ l成熟NT02CB1447/CB01445,SEQ ID NO:1,Genbank登录号CAL82988.1的氨基酸249-581,MW:37.3kDa)。

[0099] 图4:ESI-MS/MS谱分析报告。

[0100] nBH批号TE311206的38.6kDa蛋白质条带被鉴定为NT02CB1446/CB01444,Mascot分数为725,肽MS/MS序列覆盖率为整个开放阅读框(ORF)的29.6%。没有鉴定出源自N端253个氨基酸的肽(灰色框鉴定为MS肽;红色方块鉴定为在MS/MS衰变之后肽的氨基酸y-/b-离子)。批号TE311206的MS/MS分析显示基于形成nBH的C端氨基酸254-594,序列覆盖率为52%。

[0101] 图5:ESI-MS/MS谱分析报告

[0102] nBH批号TIK301009的37.3kDa蛋白质条带被鉴定为NT02CB1447/CB01445,Mascot分数为555,且肽MS/MS序列覆盖率为整个开放阅读框(ORF)的28.4%。除了一个肽以外,全部肽(灰色框鉴定为MS肽;红色方块鉴定为在MS/MS衰变之后肽的氨基酸y-/b-离子)均鉴定为源自C端333个氨基酸。批号TIK301009的MS/MS分析显示基于形成nBH的C端氨基酸249-581,序列覆盖率为49.5%。

[0103] 图6:比较源自三个纯化批量的nBH的浓度依赖性蛋白水解活性。

[0104] A通过12.5% SDS-PAGE的活性测试分析源自批号TIK301009、TIK280509和TE311206的nBH,采用浓缩nBH的下列稀释液:1:10,1:30,1:100,1:300,1:1000。该试验通过在37°C下孵育1 $\mu$ g scBoNT/A和2 $\mu$ l dH<sub>2</sub>O和1 $\mu$ l的对应稀释nBH 60分钟来进行。对于SDS-PAGE分析,添加3 $\mu$ l的还原性4x SDS Laemmli缓冲液至终体积10 $\mu$ l。150kDa scBoNT/A被切割成100kDa重链和50kDa轻链。

[0105] B对重链、轻链以及scBoNT/A的蛋白条带的光学密度定量,将轻链和重链产物条带的总和除以LC、HC和scBoNT/A蛋白质条带总和。较高的稀释倍数的第一多肽减少切割率。三个不同批量的特定蛋白分解活性几乎一致。

[0106] 图7:nBH对scBoNT/A野生型及包含修饰环体的突变体的时间依赖性切割。

[0107] A环体序列修饰。在scBoNTAS Throm中,所有赖氨酸残基均被移除,取而代之地插入凝血酶识别区序列LVPRGS,而在scBoNT Res中,该环体缺失任何碱性氨基酸。缩短环体至八个小型残基或具有大型侧链的5个氨基酸分别产生scBoNTAS (GGSG)<sub>2</sub>和scBoNTAS FQWYI。在scBoNTAS CGS-C中,整个环体被删除且形成半胱氨酸的二硫键被甘氨酸和丝氨酸取代。

[0108] B scBoNT/A野生型和突变体的时间依赖性切割的SDS-PAGE分析。

[0109] C scBoNTAS野生型被nBH以时间依赖性方式在120分钟内活化成为轻链和重链。缺

失赖氨酸和插入单一精氨酸残基延长该环体的切割(scBoNTAS Throm)。缺失任何碱性残基的环体仍可被切割(scBoNTAS Res)。缩短环体至8聚体肽、导入具有大型侧链的5个氨基酸或删除整个环体产生不可切割的scBoNT/A。

[0110] 图8:用nBH消化scBoNT/A后50kDa和100kDa切割产物的MS/MS分析。

[0111] A 50kDa切割产物的分析,其被鉴定为BoNT/A的轻链,Mascot分数为1460。最C端肽覆盖氨基酸G433~K438,其对应BoNT/A LC的生理性观察C端。

[0112] B 100kDa切割产物的分析,其被鉴定为BoNT/A的重链,Mascot分数为96。最N端肽覆盖氨基酸A449~K456,其对应BoNT/A HC的生理性观察N端。

[0113] 图9:A以12.5%SDS-PAGE分析三次汇集的抗nBH-IgY的蛋白质含量(mg/ml)。B ELISA:Nunc Maxisorp F96微滴定板用不同批量PBS中的nBH(500ng/mL)在4°C下过夜被覆,接着用含有0.1%吐温-20和2%无脂脱脂奶粉的PBS封闭缓冲液封闭1小时。清洗后,添加每次汇集的IgY稀释液(10 $\mu$ g/ml于封闭缓冲液中)1小时,接着用经生物素标记的驴抗鸡IgY、抗生蛋白链霉素-辣根过氧化物酶(均来自Dianova公司,德国汉堡)和3,3',5,5'-四甲基联苯胺(西格玛公司)检测。

[0114] 图10:A重组表达和通过Talon IMAC分离无活性的BH 1-581(63kDa)。

[0115] 10%SDS-PAGE分析Talon IMAC组分(LMW:116kDa,66kDa,45kDa,35kDa,25kDa;SS34,清澈裂解物;TD,流通物;W,清洗组分;E1-E7,咪唑洗脱组分1~7)。B在37°C下孵育1小时后没有观察到有scBoNT/A被重组iBH(SEQ ID NO:2;“E”;63kDa)内蛋白酶分解成LC(50kDa)和HC(100kDa)(泳道6)(LMW:116kDa,66kDa,45kDa,35kDa,25kDa)。

[0116] 图11:使用经纯化的活性BoNT水解酶(nBH)获得经蛋白水解处理的多肽

[0117] A 200 $\mu$ g的重组纯化的scBoNT/A与350ng纯化的活性BoNT水解酶在37°C下孵育12分钟。通过SEC(柱Superdex 200 10/300GL,缓冲液:50mM NaP pH7.5,150mM NaCl,样品体积=0.3ml,流速=0.25ml/分钟)移除nBH来停止反应,切割的量通过10%SDS-PAGE分析。B包含约40%的经处理的BoNT/A的组分1(1800 $\mu$ l)与350ng纯化的活性BoNT水解酶在37°C下孵育15分钟,并通过超过滤浓缩至300 $\mu$ l。为最终停止反应,通过SEC(柱Superdex 200 10/300GL,缓冲液:50mM NaP pH 7.5,150mM NaCl,样品体积=0.3ml,流速=0.25ml/分钟)移除nBH,切割的量通过10%SDS-PAGE分析。C包含约80%的经处理BoNT/A的组分1和2(1800 $\mu$ l)被合并与120ng纯化的活性BoNT水解酶在37°C下孵育25分钟,然后通过超过滤浓缩至300 $\mu$ l。为了最终停止反应,通过SEC(柱Superdex 200 10/300GL,缓冲液:50mM NaP pH 7.5,150mM NaCl,样品体积=0.3ml,流速=0.25ml/分钟)移除nBH,切割的量通过10%SDS-PAGE分析。获得A>95%经处理的BoNT/A(SEQ ID NO.3)。

[0118] 序列列表显示:

[0119] SEQ ID NO:1:源自肉毒芽胞梭菌菌株ATCC 3502的具有蛋白水解活性的多肽,GenBank登录号:“CAL82988.1”,缺失248N端氨基酸残基

[0120] SEQ ID NO:2:源自肉毒芽胞梭菌菌株ATCC 3502的不具有蛋白水解活性的多肽,GenBank登录号:“CAL82988.1”

[0121] SEQ ID NO:3:ATCC 3502的BoNT/A,Genbank登录号“AAA23262”

[0122] SEQ ID NO:4:BoNT/A1的环体

[0123] SEQ ID NO:5:BoNT/A2/A6的环体

- [0124] SEQ ID NO:6:BoNT/A3的环体
- [0125] SEQ ID NO:7:BoNT/A3的环体
- [0126] SEQ ID NO:8:BoNT/A4的环体
- [0127] SEQ ID NO:9:BoNT/A5的环体
- [0128] SEQ ID NO:10:BoNT/A7的环体
- [0129] SEQ ID NO:11:BoNT/B1/B4bv/B6的环体
- [0130] SEQ ID NO:12:BoNT/B2/B3的环体
- [0131] SEQ ID NO:13:BoNT/B5np的环体
- [0132] SEQ ID NO:14:BoNT/C/CD的环体
- [0133] SEQ ID NO:15:BoNT/D的环体
- [0134] SEQ ID NO:16:BoNT/DC的环体
- [0135] SEQ ID NO:17:BoNT/E1-E5的环体
- [0136] SEQ ID NO:18:BoNT/E6的环体
- [0137] SEQ ID NO:19:BoNT/F1/F6的环体
- [0138] SEQ ID NO:20:BoNT/F2/F3的环体
- [0139] SEQ ID NO:21:BoNT/F4的环体
- [0140] SEQ ID NO:22:BoNT/F5的环体
- [0141] SEQ ID NO:23:BoNT/F7的环体
- [0142] SEQ ID NO:24:BoNT/G的环体
- [0143] SEQ ID NO:25:TeNT的环体
- [0144] SEQ ID NO:26:编码SEQ ID NO:1的核酸序列
- [0145] SEQ ID NO:27:编码SEQ ID NO:2的核酸序列
- [0146] 下文实施例说明本发明,并应在任何情况下均被理解为不限制其范围。

## 实施例

[0147] 实施例1:天然BoNT水解酶(nBH)的纯化和表征,其特异性地将单链BoNT/A切割成其活性双链形式

[0148] (1) 读出系统/活性测试:为了特异性地检测并纯化在肉毒芽胞梭菌的培养物上清液中和色谱步骤之间水解肉毒神经毒素A(BoNT/A)成为50kDa轻链(LC)和100kDa重链(HC)的酶活性,在大肠杆菌中表达150kDa BoNT/A为单链(sc)多肽。将该重组scBoNT/A与合适酶活性(nBH)孵育应产生可通过还原性10-13%SDS-PAGE观察到的50kDa LC和100kDa HC。

[0149] (2) 芽胞梭菌蛋白酶表达:将肉毒芽胞梭菌菌株ATCC 3502的单一菌落接种于100ml脑心浸液(BHI)培养基中并且使该培养物在37°C下于厌氧条件下孵育过夜。将10ml的0/N培养物接种至11 BHI培养基,厌氧孵育48-72小时。

[0150] (3) 硫酸铵沉淀:通过离心(4°C,6500xg,25分钟)收集11培养物上清液。添加硫酸铵至终浓度85%(此时为575g),在4°C搅拌悬浮液6小时,随后离心(4°C,6500xg,30分钟)。使成团的硫酸铵沉淀溶解于小体积(此时为5ml)的50mM NaP pH 7.5,并以50mM NaP,150mM NaCl pH 7.5透析。最后,使该透析液离心(4°C,40000xg,60分钟)并将上清液用于IEC。

[0151] (4) 离子交换色谱(IEC,柱HiPrep 16/10Q FF):将(3)的上清液(图1,泳道3)施加

至HiPrep 16/10Q FF阴离子交换柱,该柱采用包含50mM NaP pH 7.5,150mM NaCl的缓冲液平衡并流析。所述流析以1ml/分钟的速率进行。活性测试通过使5 $\mu$ l的所有其他组分和2 $\mu$ g scBoNTA在37 $^{\circ}$ C下孵育1小时进行,随后在SDS-PAGE上分析(图1)。将组分6-24合并,并且利用超过滤(Amicon-Ultra MWC0 10,000)将其体积浓缩至3.5ml。

[0152] (5) 尺寸排阻色谱(SEC,HiLoad 16/60 Superdex 200):随后,将(4)的浓缩蛋白质溶液载入HiLoad 16/60Superdex 200柱,用50mM NaP pH 7.5,150mM NaCl平衡。分离以1ml/分钟的流速进行。滞留体积为80ml~100ml的组分利用活性测试(1)分析,而含有该酶活性(nBH)的合适组分经合并(约10ml)并通过超过滤浓缩至3ml。随后添加硫酸铵至终浓度12.5%=500mM(+0.2g)。

[0153] (6) 疏水性相互作用色谱(HIC,HiTrap苯基琼脂糖):使nBH与苯基琼脂糖于缓冲液A(50mM NaP pH 7.5,500mM硫酸铵)中结合。结合的nBH通过减少硫酸铵的量洗脱,这是因为以1ml/分钟的流速线性增加缓冲液B(50mM NaP pH7.5)所致。所有含蛋白质的组分利用活性测试(1)分析,将合适组分合并并通过超过滤浓缩至3.5ml。所述溶液的缓冲液调节至50mM NaP pH 7.5;150mM NaCl。

[0154] (7) SEC(HiLoad16/60 Superdex 75):最后,通过SEC纯化nBH,使用HiLoad16/60Superdex 75柱,流速1ml/分钟50mM NaP pH 7.5,150mM NaCl。滞留体积为70ml~80ml的组分通过12.5%SDS-PAGE(图2)分析,并将移动至约37.3kDa的含有nBH的组分8~12合并(约10ml)并通过超过滤浓缩至1ml。

[0155] (8) 移动至约37.3kDa的主要蛋白(nBH)根据Edman降解程序通过N端肽测序分析。该鉴定的肽序列是V Q G Q S V K G V G并且对应SEQ ID NO:1的前十个残基。

[0156] (9) 两批nBH(NT02CB1447,37.3kDa,图3,泳道3:TIK301009,泳道4:TIK280509)根据上述方法可再现分离。经下述调整分离方法得到nBH同种型NT02CB1446(38.6kDa,图3,泳道2,批号TE311206):(i)肉毒芽胞梭菌培养物的生长:18小时,而非48~72小时;(ii)色谱步骤变化:IEC $\rightarrow$ SEC Superdex 75 $\rightarrow$ HIC苯基琼脂糖,而非IEC $\rightarrow$ SEC Superdex 200 $\rightarrow$ HIC苯基琼脂糖 $\rightarrow$ SEC Superdex 75。

[0157] 实施例2:通过质谱法(MS)从肉毒芽胞梭菌鉴定nBH的序列

[0158] (1) 胰蛋白酶消化:切除在SDS-PAGE移动至大约38kDa的蛋白质条带(nBH)以供胰蛋白酶消化,使其在37 $^{\circ}$ C下于50mM NH<sub>4</sub>HCO<sub>3</sub>,50%乙腈中轻摇30分钟以脱色。重复脱色直至凝胶斑点变为透明。添加乙腈(100%)并在3分钟后移除。随后,在真空离心干燥系统(德国的Eppendorf公司)中干燥斑点。添加内含胰蛋白酶(10ng/ $\mu$ l)的50mM NH<sub>4</sub>HCO<sub>3</sub>并在冰上孵育1小时。然后,移除剩余的胰蛋白酶溶液,加入小体积50mM NH<sub>4</sub>HCO<sub>3</sub>并在37 $^{\circ}$ C下消化过夜。收集上清液,并用5%TFA,10%乙腈萃取凝胶片两次。合并所有液体,在真空离心干燥系统中干燥,提取的肽贮存在4 $^{\circ}$ C下。

[0159] (2) 基质辅助激光解吸离子化飞行时间(MALDI-TOF/TOF)MS:样品于MALDI-TOF/TOF质谱仪(Ultraflex1 Bruker Daltonik股份有限公司)中以线性模式,加速电压25kV分析。检测从700m/z至4,500m/z的质量。样品(2 $\mu$ l)以含有50%乙腈和0.2%三氟乙酸(TFA)的2 $\mu$ l芥子酸溶液直接在不锈钢MALDI靶标板上共结晶。每个样品收集500个激光斑。

[0160] (3) 逆相色谱肽分离:肽分离通过逆相色谱进行,使用由自动采样仪和梯度泵组成的纳米HPLC系统(Agilent Technologies公司,德国瓦尔德布隆)。将该样品溶解于缓冲液A

(5%乙腈,0.1%甲酸),并且取至多10 $\mu$ l等分以流速5 $\mu$ l/分钟注射至C18柱(Zorbax SB-C18,5 $\mu$ m,300A,0.5mm内径,长度15cm)。上样后,该柱用缓冲液A清洗15分钟,肽利用洗脱液A和洗脱液B(内含70%(v/v)乙腈的0.1%(v/v)甲酸)的梯度从0%到100%的洗脱液B在75分钟内洗脱。

[0161] (4) 电子喷雾离子化 (ESI)-界面和离子阱质谱法:HPLC的出口直接与离子阱质谱仪的纳米型ESI源连接,并且使用Agilent共轴鞘液体喷雾器(Agilent Technologies)。出口毛细管由环绕式钢针固定,突出0.1~0.2mm。喷雾通过N<sub>2</sub>作为喷雾器气体(5l/分钟)来稳定。离子化电压设定在4,500V,并且干燥气体以5psi并以250 $^{\circ}$ C供应。光谱仪Esquire3000+离子阱质谱仪(Bruker Daltonik)以每秒13,000m/z的扫描速度收集。采用正离子模式的ESI,质谱数据在50~1600m/z扫描模式获得,根据数据需要可在MS和MS/MS分析之间切换。为了增加MS/MS谱的质量,仅选择来自一个谱的两种前体离子供于MS/MS分析,并且将主动排除设定于2分钟以排除已经被测量过的前体离子。

[0162] (4) 数据处理:数据处理采用数据分析(Data Analysis)(3.0版)和BioTools(3.0版)软件包(Bruker Daltonik)进行。蛋白质鉴定采用MASCOT软件(2.1版)和MSDB数据库(Matrix Science,英国伦敦)来进行。

[0163] (5) 结果:

[0164] 表2:通过MS鉴定的nBH

泳道	批号nBH	蛋白质浓度	ORF名称	Genbank 登录号	ORF的氨 基酸	MW [kDa]	Mascot 分数
2	TE311206	192 ng/ $\mu$ l	NT02CB1446	CAL82987.1	254-594	38.6	725
			CBO1444				
3	TIK301009	130 ng/ $\mu$ l	NT02CB1447	CAL82988.1	249-581	37.3	555
			CBO1445				
4	TIK280509	114 ng/ $\mu$ l	NT02CB1447	CAL82988.1	249-581	37.3	609
			CBO1445				

[0165] 泳道2的38.6kDa蛋白质条带(nBH批号TE311206)经鉴定为NT02CB1446/CBO1444,Mascot分数为725,肽MS/MS序列覆盖率为整个开放阅读框(ORF)的29.6%。没有鉴定到源自N端253个氨基酸的肽(图4)。批号TE311206的MS/MS分析显示基于形成nBH的C端氨基酸254-594,序列覆盖率为52%。

[0167] 泳道3(nBH批号TIK301009)和泳道4(nBH批号TIK280509)的37.3kDa蛋白质条带分别被鉴定为Mascot分数为555和609的NT02CB1447/CBO1445。除了一个肽以外,所有鉴定的肽均源自C端333个氨基酸(图5)。批号TIK301009的MS/MS分析显示基于形成nBH的C端氨基酸249-581,序列覆盖率为49.5%。

[0168] 实施例3:nBH酶特异性的表征

[0169] (1) 比较源自三个纯化批量的nBH的浓度依赖性蛋白质水解活性(图6)。分析源自批号TIK301009,TIK280509和TE311206的nBH的不同稀释倍数的活性测试显示较高稀释倍数降低切割率。这三种不同批号的蛋白水解活性几乎一致,表明成熟同种型NT02CB1446

(TE311206) 显示出与成熟NT02CB1447 (SEQ ID NO:1) 相似的特异性活性。

[0170] (2) 使用活性测试分析nBH对scBoNT/A野生型和突变体的时间依赖性切割(图7)。超过95%的scBoNTAS野生型被nBH以时间依赖性方式在120分钟内活化成轻链和重链。环体序列被修饰以使该切割位点特征化。在scBoNTAS Throm中,所有赖氨酸残基均被移除,并且插入凝血酶识别序列LVPRGS,这延长该切割速率。在scBoNT Res中,该环体缺失任何碱性氨基酸,这大幅延缓了完全水解,表明nBH对于在切割位点的碱性残基如赖氨酸和精氨酸具有强识别偏好。此外,nBH对于环体的可接近性受到缩短该环体至8个小型残基或具有大型侧链的5个氨基酸的阻碍(scBoNTAS (GGSG)<sub>2</sub>和scBoNTAS FQWYI)。

[0171] (3) 用nBH消化scBoNT/A后对50kDa切割产物的MS/MS分析显示最C端肽覆盖氨基酸G433至K438,其对应BoNT/A LC的生理学观察的C端(图8A)。对于被鉴定为BoNT/A的重链的100kDa的切割产物的分析证明,最N端肽覆盖氨基酸A449至K456,其对应BoNT/A HC的生理学观察的N端(图8B)。因此,经分离的nBH产生经处理的BoNT/A,并且优先水解在赖氨酸和精氨酸残基C端的肽键。

[0172] 实施例4:BoNT水解酶的进化保守性及其同种型

[0173] SEQ ID NO:2 (Genbank登录号CAL82988.1/YP\_001253958.1)的蛋白质序列分析显示三个保守性结构域。残基18-573对应锌金属蛋白酶(弹性蛋白酶)或涉及氨基酸运输和代谢的LasB,Blast得分738。残基148-212对应肽酶前肽和YPEB结构域或PepSY (Blast得分97)。残基336-573是包括嗜热菌蛋白酶、溶蛋白素、融金菌素和中性蛋白酶的肽酶M4家族的一部分(Blast得分803)。

[0174] 肉毒芽胞梭菌ATCC 3502的基因组测序显示存在编码iBH同种型的六个ORF (Sebahia等,2007,Genome Res.17(7):1082-1092)。其他基因组数据可获自10个组I肉毒芽胞梭菌菌株以及非BoNT分泌性产芽胞梭菌,其所有均包含编码iBH的5~7个ORF。nBH (SEQ ID NO:1)与其它63种同种型共享最少64%的氨基酸序列相同性。

[0175] 实施例5:产生对BoNT水解酶具有特异性的抗体

[0176] (1) IgY的产生:将16周龄的鸡[ISA布朗和罗曼选择的来亨鸡(LSL), Spreenhagener Vermehrungsbetrieb für Legehennen股份有限公司,德国贝斯滕塞]饲养在专门为鸡建造的个体饲笼中(Ebeco,德国卡斯特罗普)。食物(ssniff Legehühner-Zucht 1和2;SsniffSpezialitäten股份有限公司,德国索斯特)和水可自由采食,鸡从23~25周龄开始下蛋。每天收集、标记鸡蛋,将其存放于4℃直至进一步处理。所有动物的饲养和实验均根据柏林当地主管机关规定进行(第H0069/03号)。鸡在1年期间经肌肉注射(左右侧胸肌)免疫接种并加强免疫共10次,每次间隔4~8周。使用的间隔根据先前研究显示要到免疫后至少3周,才出现记忆细胞(Pei和Collisson,2005)。使用的抗原浓度是每次注射约20μg (nBH)。每次免疫不注射超过500μl的抗原溶液。第一次免疫采用完全弗氏(Freund's)佐剂,之后的加强注射采用FIA。IgY的纯化修改自Polson等.(1980)。简言之,蛋黄以1:2的比例用无菌PBS (pH 7.4,罗氏公司,德国曼海姆)稀释。为了去除脂肪和脂蛋白,添加3.5% (w/v)的聚乙二醇(PEG) 6000 (罗氏公司,德国卡尔斯鲁厄)。轻摇后离心(以10,000×g在4℃下离心20分钟),倒出上清液,加入固体PEG 6000至终浓度12% (w/v)。该混合物接着如上述离心。沉淀物溶解于10ml的PBS,添加PEG至12% (w/v),并将该溶液离心。最后,沉淀物溶解于1.2ml的PBS,转移至微透析装置(QuixSep,Roth,德国)并在4℃下对PBS进

行透析。蛋白质含量 (mg/ml) 通过 12.5% SDS-PAGE 分析 (图 9A) 并在 280nm 处以光度计检测, 根据 Lambert-Beer 规则计算 IgY 的消光系数为 1.33。

[0177] (2) ELISA: Nunc Maxisorp F96 微滴定板 (VWR 国际股份有限公司, 德国达姆施塔特) 以内含不同批量的 nBH (500ng/mL) 的 PBS 在 4°C 下被覆过夜, 随后以含有 0.1% 的吐温-20 和 2% 无脂脱脂奶粉 (默克公司, 德国达姆施塔特) 封闭 1 小时。在清洗后, 添加 IgY 稀释液 (10 μg/ml 于封闭缓冲液中) 1 小时, 接着利用经生物素标记的驴抗鸡 IgY、链霉亲和素-辣根过氧化物酶 (均为 Dianova, 德国汉堡) 和 3,3',5,5'-四甲基联苯胺 (西格玛公司) 检测。被检测到的 nBH 如图 9B 所示。

[0178] (3) Western 印迹: nBH 以 12.5% SDS-PAGE 分离, 使用标准免疫印迹技术转移至偏二氟乙烯膜 (英杰股份有限公司, 德国卡尔斯鲁厄) 上。膜于 4°C 封闭过夜, 与 IgY (1:5,000 于封闭缓冲液中) 一起孵育 1 小时。清洗后, 膜用经生物素标记的驴抗鸡 IgY 探测 30 分钟, 使用碱性磷酸酶和 CDP-Star (Perkin Elmer, Waltham, MA) 显色。

[0179] 实施例 6: 重组表达 BoNT 水解酶

[0180] (1) 质粒构建体: 编码天然 BH (SEQ ID NO:1) 及其前肽 (SEQ ID NO:2) 的基因部分通过 PCR 扩增, 使用合适的寡核苷酸和肉毒芽胞梭菌 ATCC 3502 的基因组 DNA, 其与编码 His6 标签的寡核苷酸融合并插入到 pQE3 (凯杰公司 (Qiagen)) 以分别产生表达质粒 pQ-BH1445H6-249-581 和 pQ-BH1445H6-1-581。核苷酸序列通过 DNA 测序验证。

[0181] (2) 重组蛋白的纯化: 与羧基端 His6 标签融合的 nBH 和 iBH 用大肠杆菌菌株 M15pREP4 (凯杰公司) 在室温下孵育 10 小时生成, 并使用 Talon-琼脂糖珠 (Clontech 公司) 根据生产商说明纯化。将包含所需蛋白质的组分汇集, 在液氮中冷冻, 保存于 -70°C。iBH 经分离为 MW 为 63kDa 的重组蛋白 (图 10A)。iBH 的不具活性利用活性测试显示, 在 37°C 下孵育 1 小时后, 没有 scBoNT/A 野生型 (wt) 被水解成 LC 和 HC (图 10B)。

[0182] 实施例 7: BoNT 水解酶的抑制

[0183] (1) 筛选 BH 的肽抑制剂: 将合成缺失一个或多个碱性残基的基于 SEQ ID NO:4~25 的肽。将各肽添加至根据活性测试的混合物。能减少经处理的 scBoNT/A 的量、延长完全处理 scBoNT/A 所需的时间或阻断处理 scBoNT/A 的肽被认为是 nBH 的抑制剂。

[0184] (2) 筛选基于抗体的抑制剂: 针对源自 nBH 的表位产生的抗体如实施例 5 的 IgY 与 nBH 一起培养, 接着进行活性测试。能减少经处理的 scBoNT/A 的量、延长完全处理 scBoNT/A 所需时间或阻断处理 scBoNT/A 的抗体被视作 nBH 的抑制剂。

[0185] 实施例 8: 使用经纯化的活性 BoNT 水解酶 (nBH) 获得经蛋白水解处理的多肽

[0186] (1) 200μg 的重组纯化的 scBoNT/A 与 350ng 的经纯化的活性 BoNT 水解酶在 37°C 下培养 12 分钟。要停止反应, 通过 SEC (柱 Superdex 200 10/300GL, 缓冲液: 50mM NaP pH 7.5, 150mM NaCl, 样品体积 = 0.3ml, 流速 = 0.25ml/分钟) 移除 nBH, 切割的量通过 10% SDS-PAGE 分析 (图 11A)。

[0187] (2) 包含约 40% 的经处理的 BoNT/A 的组分 1 (1800μl) 与 350ng 经纯化的活性 BoNT 水解酶于 37°C 下孵育 15 分钟, 并通过超过滤浓缩至 300μl。最后, 为停止反应, 通过 SEC (柱 Superdex 200 10/300GL, 缓冲液: 50mM NaP pH 7.5, 150mM NaCl, 样品体积 = 0.3ml, 流速 = 0.25ml/分钟) 移除 nBH, 切割的量通过 10% SDS-PAGE 分析 (图 11B)。

[0188] (2) 将包含约 80% 的经处理的 BoNT/A 的组分 1 和 2 (1800μl) 合并, 并与 120ng 经纯化

的活性BoNT水解酶于37℃下孵育25分钟,并通过超过滤浓缩至300 $\mu$ l。最后,为停止反应,通过SEC(柱Superdex 200 10/300GL,缓冲液:50mM NaP pH 7.5,150mM NaCl,样品体积=0.3ml,流速=0.25ml/分钟)移除nBH,切割的量通过10%SDS-PAGE分析(图11C)。获得>95%的经处理的BoNT/A(Seq ID NO.3)。若所述第二多肽在37℃下于一个步骤中处理50分钟(200 $\mu$ g scBoNT/A与350ng nBH孵育),则获得相同的经完全处理的第二多肽(>95%经处理的BoNT/A)。在37℃下孵育1小时后,超过97%的BoNT/A是经处理的。



210	215	220
Pro Ala His Met Lys Asp Tyr Lys Tyr Lys Thr Met Asn Asp Asp Asn		
225	230	235
Gly Gly Val His Thr Asn Ser Gly Ile Ile Asn His Ala Ala Tyr Leu		
	245	250
Val Ala Asp Gly Ile Glu Lys Thr Gly Ala Lys Asn Ser Lys Asp Ile		255
	260	265
Met Gly Lys Ile Phe Tyr Thr Ala Asn Cys Tyr Lys Trp Asp Glu Thr		270
	275	280
Thr Asn Phe Ala Lys Cys Arg Asn Asp Val Val Gln Val Thr Lys Glu		285
	290	300
Leu Tyr Gly Glu Asn Ser Asn Tyr Val Lys Ile Val Glu Lys Ala Phe		
305	310	315
Asp Gln Val Gly Ile Thr Ala Thr Pro Gln Leu Pro Leu		
	325	330
<210> 2		
<211> 581		
<212> PRT		
<213> 肉毒梭菌		
<400> 2		
Met Lys Ser Lys Lys Leu Leu Ala Thr Val Leu Ser Ala Val Ile Thr		
1	5	10
Phe Ser Thr Val Ser Ala Val Tyr Ala Ala Pro Val Gly Lys Glu Ser		
	20	25
Lys Val Glu Pro Lys Thr Thr Thr Ile Thr Trp Glu Lys Asn Glu Gln		
	35	40
Asn Thr Lys Lys Ala Ala Thr Asp Ile Thr Glu Lys Lys Phe Asn Asn		
	50	55
Ser Glu Glu Ile Thr Lys Phe Phe Glu Lys Asn Ile Ser Lys Phe Gly		
65	70	75
Val Gln Lys Gly Ser Leu Lys Asn Thr Lys Thr Val Lys Asp Glu Lys		
	85	90
Gly Lys Thr Asn Tyr His Met Ile Tyr Glu Val Glu Gly Ile Pro Val		
	100	105
Tyr Tyr Gly Arg Ile Val Phe Thr Thr Glu Lys Asp Ser Ser Met Asp		
	115	120
Ser Ile Asn Gly Arg Ile Asp Thr Val Phe Glu Asn Gly Asn Trp Lys		
	130	135
Asn Lys Ile Lys Leu Ser Lys Glu Asp Ala Ile Ala Lys Ala Lys Asn		

145	150	155	160
Asp Ile Lys Asp Glu Lys Ala Thr Ser Lys Lys Thr Asp Leu Tyr Leu			
	165	170	175
Tyr Asn Phe Glu Gly Lys Pro Tyr Val Val Tyr Leu Val Asp Leu Ile			
	180	185	190
Thr Asp Asn Gly Ser Trp Thr Val Phe Val Asn Ala Glu Asp Gly Ser			
	195	200	205
Ile Val Asn Lys Phe Asn Asn Thr Pro Thr Leu Ile Asp Thr Lys Asp			
	210	215	220
Gln Lys Leu Pro Asn Ala Lys Lys Ile Lys Asp Glu Ala Lys Lys Ala			
225	230	235	240
Ser Asn Ala Asn Asn Val Ile Asp Val Gln Gly Gln Ser Val Lys Gly			
	245	250	255
Val Gly Lys Thr Ser Leu Asp Gly Leu Val Asn Ile Asp Val Thr Tyr			
	260	265	270
Gly Asn Gly Lys Tyr Tyr Leu Lys Asp Ser Asn Lys Asn Ile Tyr Leu			
	275	280	285
Tyr Asp Leu Lys Asn Gln Val Asp Glu Tyr Asp Leu Tyr Asn Tyr Leu			
	290	295	300
Ser Arg Pro Asn Tyr Lys Gln Ile Leu Met Ser Lys Ser Glu Leu Ile			
305	310	315	320
Ser Asn Tyr Asn Asn Asn Phe Ile Ala Asn Asn Gln Val Asn Ser Val			
	325	330	335
Asp Ala Tyr Val Asn Thr Asn Lys Thr Tyr Asp Tyr Tyr Lys Asn Lys			
	340	345	350
Leu Asn Arg Asn Ser Ile Asp Asn Lys Gly Met Asn Ile Asn Gly Phe			
	355	360	365
Val His Val Gly Arg Asn Tyr Gly Asn Ala Phe Trp Tyr Gly Pro Tyr			
	370	375	380
Asp Gly Met Phe Phe Gly Asp Gly Asp Gly Ile Tyr Phe Ser Ser Leu			
385	390	395	400
Ala Lys Ser Leu Asp Val Val Gly His Glu Leu Ser His Gly Val Thr			
	405	410	415
Asn Lys Glu Ser Asn Leu Lys Tyr Glu Asn Glu Ser Gly Ala Leu Asn			
	420	425	430
Glu Ser Phe Ser Asp Ile Met Gly Val Ala Val Glu Gly Lys Asn Phe			
	435	440	445
Val Leu Gly Glu Asp Cys Trp Val Ala Gly Gly Val Met Arg Asp Met			
	450	455	460

Glu Asn Pro Ser Arg Gly Gly Gln Pro Ala His Met Lys Asp Tyr Lys  
 465 470 475 480  
 Tyr Lys Thr Met Asn Asp Asp Asn Gly Gly Val His Thr Asn Ser Gly  
 485 490 495  
 Ile Ile Asn His Ala Ala Tyr Leu Val Ala Asp Gly Ile Glu Lys Thr  
 500 505 510  
 Gly Ala Lys Asn Ser Lys Asp Ile Met Gly Lys Ile Phe Tyr Thr Ala  
 515 520 525  
 Asn Cys Tyr Lys Trp Asp Glu Thr Thr Asn Phe Ala Lys Cys Arg Asn  
 530 535 540  
 Asp Val Val Gln Val Thr Lys Glu Leu Tyr Gly Glu Asn Ser Asn Tyr  
 545 550 555 560  
 Val Lys Ile Val Glu Lys Ala Phe Asp Gln Val Gly Ile Thr Ala Thr  
 565 570 575  
 Pro Gln Leu Pro Leu  
 580  
 <210> 3  
 <211> 1296  
 <212> PRT  
 <213> 肉毒梭菌  
 <400> 3  
 Met Pro Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly  
 1 5 10 15  
 Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro  
 20 25 30  
 Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg  
 35 40 45  
 Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu  
 50 55 60  
 Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr  
 65 70 75 80  
 Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu  
 85 90 95  
 Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val  
 100 105 110  
 Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys  
 115 120 125  
 Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr  
 130 135 140

Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile  
 145 150 155 160  
 Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr  
 165 170 175  
 Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe  
 180 185 190  
 Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu  
 195 200 205  
 Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu  
 210 215 220  
 Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn  
 225 230 235 240  
 Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu  
 245 250 255  
 Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys  
 260 265 270  
 Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Tyr Asn  
 275 280 285  
 Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val  
 290 295 300  
 Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys  
 305 310 315 320  
 Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu  
 325 330 335  
 Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp  
 340 345 350  
 Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn  
 355 360 365  
 Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr  
 370 375 380  
 Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn  
 385 390 395 400  
 Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu  
 405 410 415  
 Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg  
 420 425 430  
 Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly Tyr Asn Lys  
 435 440 445  
 Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp Asp Leu Phe Phe

450	455	460
Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu		
465	470	475
Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu		
	485	490
Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro		
	500	505
Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu		
	515	520
Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu		
	530	535
Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu		
545	550	555
His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu		
	565	570
Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys		
	580	585
Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu		
	595	600
Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr		
610	615	620
Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala		
625	630	635
Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu		
	645	650
Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala		
	660	665
Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys		
	675	680
Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu		
690	695	700
Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys		
705	710	715
Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu		
	725	730
Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn		
	740	745
Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp		
755	760	765

Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile  
 770 775 780  
 Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met  
 785 790 795 800  
 Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys  
 805 810 815  
 Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly  
 820 825 830  
 Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp  
 835 840 845  
 Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser  
 850 855 860  
 Thr Phe Thr Glu Tyr Ile Lys Asn Ile Ile Asn Thr Ser Ile Leu Asn  
 865 870 875 880  
 Leu Arg Tyr Glu Ser Asn His Leu Ile Asp Leu Ser Arg Tyr Ala Ser  
 885 890 895  
 Lys Ile Asn Ile Gly Ser Lys Val Asn Phe Asp Pro Ile Asp Lys Asn  
 900 905 910  
 Gln Ile Gln Leu Phe Asn Leu Glu Ser Ser Lys Ile Glu Val Ile Leu  
 915 920 925  
 Lys Asn Ala Ile Val Tyr Asn Ser Met Tyr Glu Asn Phe Ser Thr Ser  
 930 935 940  
 Phe Trp Ile Arg Ile Pro Lys Tyr Phe Asn Ser Ile Ser Leu Asn Asn  
 945 950 955 960  
 Glu Tyr Thr Ile Ile Asn Cys Met Glu Asn Asn Ser Gly Trp Lys Val  
 965 970 975  
 Ser Leu Asn Tyr Gly Glu Ile Ile Trp Thr Leu Gln Asp Thr Gln Glu  
 980 985 990  
 Ile Lys Gln Arg Val Val Phe Lys Tyr Ser Gln Met Ile Asn Ile Ser  
 995 1000 1005  
 Asp Tyr Ile Asn Arg Trp Ile Phe Val Thr Ile Thr Asn Asn Arg  
 1010 1015 1020  
 Leu Asn Asn Ser Lys Ile Tyr Ile Asn Gly Arg Leu Ile Asp Gln  
 1025 1030 1035  
 Lys Pro Ile Ser Asn Leu Gly Asn Ile His Ala Ser Asn Asn Ile  
 1040 1045 1050  
 Met Phe Lys Leu Asp Gly Cys Arg Asp Thr His Arg Tyr Ile Trp  
 1055 1060 1065  
 Ile Lys Tyr Phe Asn Leu Phe Asp Lys Glu Leu Asn Glu Lys Glu

1070	1075	1080
Ile Lys Asp Leu Tyr Asp	Asn Gln Ser Asn Ser	Gly Ile Leu Lys
1085	1090	1095
Asp Phe Trp Gly Asp Tyr	Leu Gln Tyr Asp Lys	Pro Tyr Tyr Met
1100	1105	1110
Leu Asn Leu Tyr Asp Pro	Asn Lys Tyr Val Asp	Val Asn Asn Val
1115	1120	1125
Gly Ile Arg Gly Tyr Met	Tyr Leu Lys Gly Pro	Arg Gly Ser Val
1130	1135	1140
Met Thr Thr Asn Ile Tyr	Leu Asn Ser Ser Leu	Tyr Arg Gly Thr
1145	1150	1155
Lys Phe Ile Ile Lys Lys	Tyr Ala Ser Gly Asn	Lys Asp Asn Ile
1160	1165	1170
Val Arg Asn Asn Asp Arg	Val Tyr Ile Asn Val	Val Val Lys Asn
1175	1180	1185
Lys Glu Tyr Arg Leu Ala	Thr Asn Ala Ser Gln	Ala Gly Val Glu
1190	1195	1200
Lys Ile Leu Ser Ala Leu	Glu Ile Pro Asp Val	Gly Asn Leu Ser
1205	1210	1215
Gln Val Val Val Met Lys	Ser Lys Asn Asp Gln	Gly Ile Thr Asn
1220	1225	1230
Lys Cys Lys Met Asn Leu	Gln Asp Asn Asn Gly	Asn Asp Ile Gly
1235	1240	1245
Phe Ile Gly Phe His Gln	Phe Asn Asn Ile Ala	Lys Leu Val Ala
1250	1255	1260
Ser Asn Trp Tyr Asn Arg	Gln Ile Glu Arg Ser	Ser Arg Thr Leu
1265	1270	1275
Gly Cys Ser Trp Glu Phe	Ile Pro Val Asp Asp	Gly Trp Gly Glu
1280	1285	1290
Arg Pro Leu		
1295		
<210>	4	
<211>	25	
<212>	PRT	
<213>	肉毒梭菌	
<220>		
<221>	BoNT/A1环体	
<222>	(1) .. (25)	
<400>	4	

Cys Val Arg Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Lys Gly  
 1                    5                    10                    15

Tyr Asn Lys Ala Leu Asn Asp Leu Cys  
                   20                    25

<210> 5

<211> 25

<212> PRT

<213> 肉毒梭菌

<220>

<221> BoNT/A2/A6环体

<222> (1) .. (25)

<400> 5

Cys Val Arg Gly Ile Ile Pro Phe Lys Thr Lys Ser Leu Asp Glu Gly  
 1                    5                    10                    15

Tyr Asn Lys Ala Leu Asn Asp Leu Cys  
                   20                    25

<210> 6

<211> 25

<212> PRT

<213> 肉毒梭菌

<220>

<221> BoNT/A3环体

<222> (1) .. (25)

<400> 6

Cys Val Arg Gly Ile Ile Pro Phe Lys Thr Lys Ser Leu Asp Glu Gly  
 1                    5                    10                    15

Tyr Asn Lys Ala Leu Asn Asp Leu Cys  
                   20                    25

<210> 7

<211> 25

<212> PRT

<213> 肉毒梭菌

<220>

<221> BoNT/A3环体

<222> (1) .. (25)

<400> 7

Cys Val Arg Gly Ile Ile Pro Phe Lys Thr Lys Ser Leu Asp Glu Gly  
 1                    5                    10                    15

Tyr Asn Lys Ala Leu Asn Tyr Leu Cys

	20	25
<210>	8	
<211>	25	
<212>	PRT	
<213>	肉毒梭菌	
<220>		
<221>	BoNT/A4环体	
<222>	(1) .. (25)	
<400>	8	
Cys Val Arg Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Glu Gly		
1	5	10 15
Tyr Asn Lys Ala Leu Asn Glu Leu Cys		
	20	25
<210>	9	
<211>	25	
<212>	PRT	
<213>	肉毒梭菌	
<220>		
<221>	BoNT/A5环体	
<222>	(1) .. (25)	
<400>	9	
Cys Val Arg Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Asp Glu Gly		
1	5	10 15
Tyr Asn Lys Ala Leu Asn Asp Leu Cys		
	20	25
<210>	10	
<211>	25	
<212>	PRT	
<213>	肉毒梭菌	
<220>		
<221>	BoNT/A7环体	
<222>	(1) .. (25)	
<400>	10	
Trp Val Arg Gly Ile Ile Pro Phe Lys Pro Lys Ser Leu Asp Glu Gly		
1	5	10 15
Ser Asn Lys Ala Leu Asn Asp Leu Cys		
	20	25
<210>	11	
<211>	10	

<212> PRT  
 <213> 肉毒梭菌  
 <220>  
 <221> BoNT/B1/B4bv/B6环体  
 <222> (1) .. (10)  
 <400> 11  
 Cys Lys Ser Val Lys Ala Pro Gly Ile Cys  
 1                    5                    10  
 <210> 12  
 <211> 10  
 <212> PRT  
 <213> 肉毒梭菌  
 <220>  
 <221> BoNT/B2/B3环体  
 <222> (1) .. (10)  
 <400> 12  
 Cys Lys Ser Val Arg Ala Pro Gly Ile Cys  
 1                    5                    10  
 <210> 13  
 <211> 10  
 <212> PRT  
 <213> 肉毒梭菌  
 <220>  
 <221> BoNT/B5np环体  
 <222> (1) .. (10)  
 <400> 13  
 Cys Lys Ser Val Lys Val Pro Gly Ile Cys  
 1                    5                    10  
 <210> 14  
 <211> 17  
 <212> PRT  
 <213> 肉毒梭菌  
 <220>  
 <221> BoNT/C/CD环体  
 <222> (1) .. (17)  
 <400> 14  
 Cys His Lys Ala Ile Asp Gly Arg Ser Leu Tyr Asn Lys Thr Leu Asp  
 1                    5                    10                    15  
 Cys



1	5	10	15
<210>	19		
<211>	17		
<212>	PRT		
<213>	肉毒梭菌		
<220>			
<221>	BoNT/F1/F6环体		
<222>	(1) .. (17)		
<400>	19		
Cys Lys Ser Val Ile Pro Arg Lys Gly Thr Lys Ala Pro Pro Arg Leu			
1	5	10	15
Cys			
<210>	20		
<211>	17		
<212>	PRT		
<213>	肉毒梭菌		
<220>			
<221>	BoNT/F2/F3环体		
<222>	(1) .. (17)		
<400>	20		
Cys Lys Ser Ile Ile Pro Arg Lys Gly Thr Lys Gln Ser Pro Ser Leu			
1	5	10	15
Cys			
<210>	21		
<211>	17		
<212>	PRT		
<213>	肉毒梭菌		
<220>			
<221>	BoNT/F4环体		
<222>	(1) .. (17)		
<400>	21		
Cys Lys Ser Ile Ile Pro Arg Lys Gly Thr Lys Ala Pro Pro Arg Leu			
1	5	10	15
Cys			
<210>	22		
<211>	15		
<212>	PRT		
<213>	肉毒梭菌		
<220>			



<212> DNA

<213> 肉毒梭菌(*C. Botulinum*)

<400> 26

```

atggttcaag gtcaaagcgt taaaggagta ggaaaaacta gcttggatgg actagtaaat      60
attgatgtaa cttatggaaa tggaaaatac tattttaaag atagcaacaa aaatatttat      120
ctatatgact taaaaaatca agttgatgaa tatgatctat acaattatct tagtagacct      180
aactataaac aaatattaat gagcaaatct gaattaatat ctaattacaa taataatfff      240
atagccaaca atcagggttaa ttctgtagat gcttatgtaa acacaaataa aacctatgat      300
tattataaaa acaaattaaa tagaaacagt attgataata agggtagtaa tattaatggg      360
tttgttcatg taggtagaaa ttatggtaat gctttttggg acgggtccata tgatgggatg      420
ttctttggcg atggcgacgg aatatacttc tcttccttg caaatcttt agatgttgta      480
ggccacgaat taagtcattg tgtaacaaat aaagagtcta atcttaata tgaaaatgaa      540
tctggtgcc taaatgaatc tttctcagat attatgggag tagctgttga gggtaaaaac      600
tttgtaactag gtgaagattg ctgggttgct ggaggagtaa tgagagatat ggaaaatcca      660
tccagaggag gccaaccagc tcatatgaaa gattataaat acaaaactat gaatgacgat      720
aacgggtggtg ttcatacaaa ttcaggtata ataaacctg ctgcttattt agttgcagat      780
ggaatagaaa aaactggtgc aaaaaatagt aaagatatta tgggaaaaat attctataca      840
gctaattgct ataaatggga tgaaacaaca aattttgcta agtgcagaaa tgatgtagtc      900
caagttacta aagaacttta tggcgaaaat agcaactatg taaaattgt tgaaaaagct      960
tttgaccaag ttggaataac tgctacacct caattacat tataa                        1005

```

<210> 27

<211> 1746

<212> DNA

<213> 肉毒梭菌

<400> 27

```

atgaaaagta aaaaattatt agctacagtg ctaagtgccg tgatcacttt ttctactgtt      60
tctgcagttt atgctgcgcc tgtaggaaa gaaagtaaag ttgaaccaa aactacaaca      120
ataacttggg aaaaaaatga acaaaactact aaaaaagctg ctactgatat aactgaaaag      180
aaatttaaca attctgagga gataactaaa ttctttgaaa aaaatatatc taaatttggg      240
gtacaaaaaag gttctcttaa aaacaccaag actgtaaaag acgaaaaagg taaaactaac      300
tatcatatga tttatgaagt agaaggtata cctgtatact atggaagaat tgtttttaca      360
actgaaaaaag actcctccat ggattctata aacggtagaa ttgatactgt ttttgaaaat      420
gggaattgga aaaacaaaaat caaactatca aaagaagatg ctatagcaa agctaaaaat      480
gatattaaag atgaaaaagc aactagtaaa aagaccgatt tatatctgta taattttgag      540
ggcaaacctt atgtagttta tttagtagat ctaattacag acaacgggag ttggacggtt      600
ttcgtaaatg ctgaggatgg ttctatagta aataaattta ataatactcc tacttttaatt      660
gatactaaag atcaaaaatt acccaatgct aaaaaatta aagatgaagc taaaaagct      720
agtaatgcaa ataatgtaat tgatgttcaa ggtcaaagcg ttaaaggagt aggaaaaact      780
agcttggatg gactagtaaa tattgatgta acttatggaa atggaaaata ctatttataa      840

```

gatagcaaca	aaaatattta	tctatatgac	ttaaaaaatc	aagttgatga	atatgatcta	900
tacaattatc	ttagtagacc	taactataaa	caaatattaa	tgagcaaadc	tgaattaata	960
tctaattaca	ataataattt	tatagccaac	aatcaggtta	attctgtaga	tgcttatgta	1020
aacacaaata	aaacctatga	ttattataaa	aacaaattaa	atagaaacag	tattgataat	1080
aagggtatga	atattaatgg	gtttgttcat	gtaggtagaa	attatggtaa	tgctttttgg	1140
tacgggccat	atgatgggat	gttctttggc	gatggcgacg	gaatataactt	ctcttcctt	1200
gcaaaatctt	tagatgttgt	aggccacgaa	ttaagtcatg	gtgtaacaaa	taaagagtct	1260
aatcttaaat	atgaaaatga	atctggtgcc	ctaaatgaat	ctttctcaga	tattatggga	1320
gtagctgttg	agggtaaaaa	ctttgtacta	ggtgaagatt	gctgggttgc	tggaggagta	1380
atgagagata	tggaaaatcc	atccagagga	ggccaaccag	ctcatatgaa	agattataaa	1440
tacaaaacta	tgaatgacga	taacggtggt	gttcatacaa	attcaggtat	aataaacct	1500
gctgcttatt	tagttgcaga	tggaatagaa	aaaactggtg	caaaaaatag	taaagatatt	1560
atgggaaaaa	tattctatac	agctaattgc	tataaatggg	atgaaacaac	aaatdddgct	1620
aagtcagaaa	atgatgtagt	ccaagttact	aaagaacttt	atggcgaaaa	tagcaactat	1680
gtaaaaattg	ttgaaaaagc	ttttgaccaa	gttggataaa	ctgctacacc	tcaattacca	1740
ttataa						1746

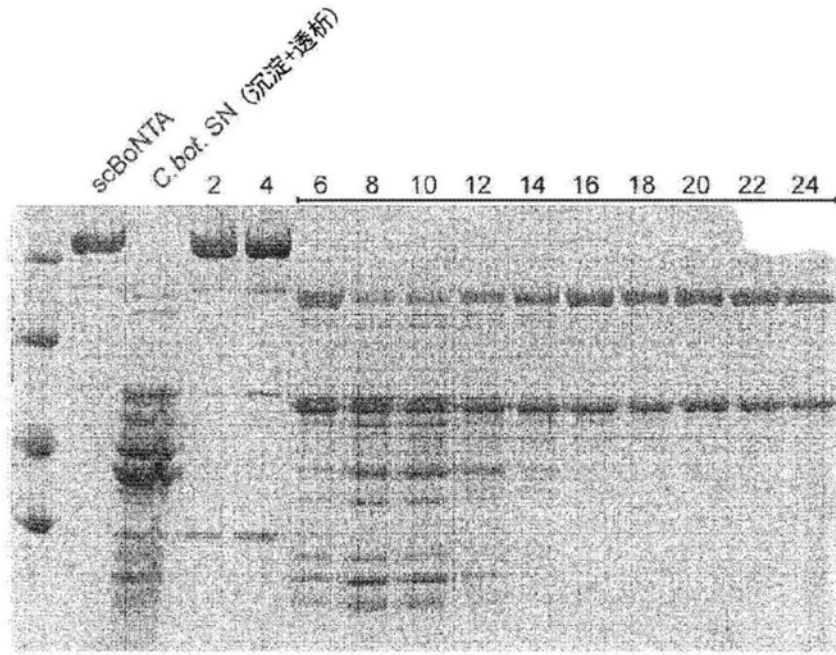


图1

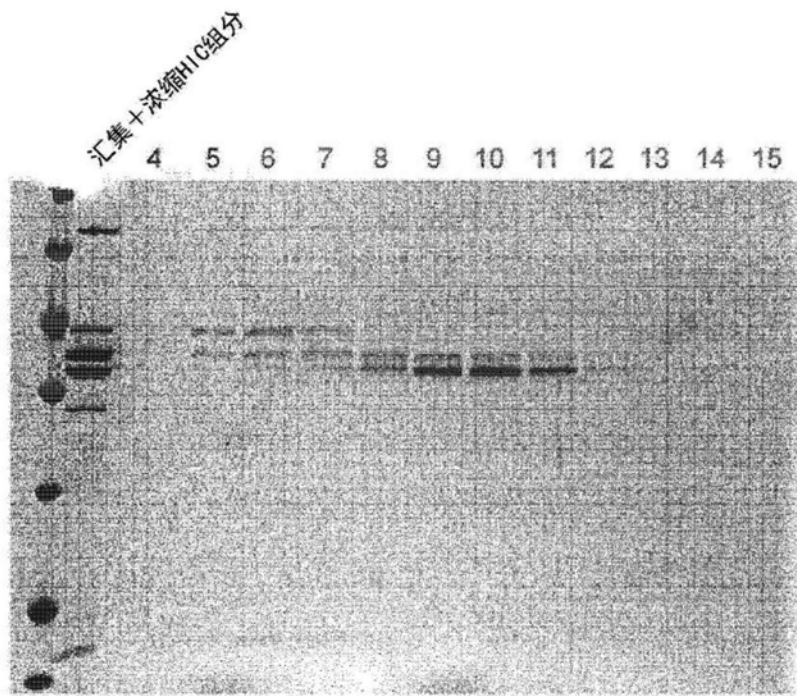


图2

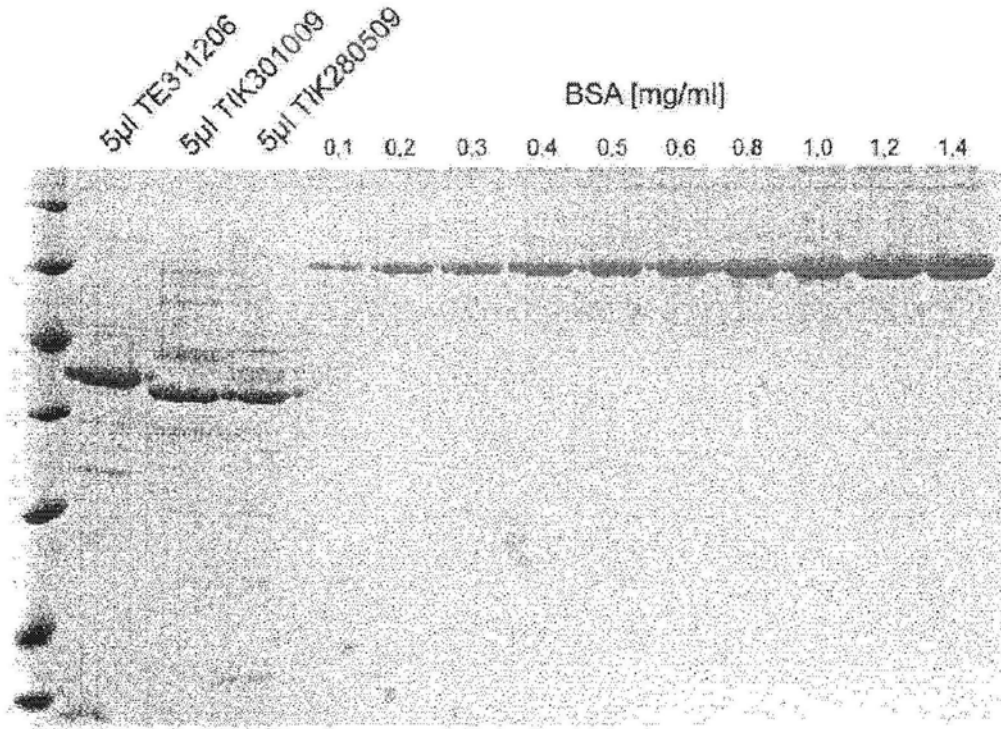


图3

10	20	30	40	50	60	70	80	90	100	110
HRKELLATV	LSAVITTSAV	SAVSRAPVGR	EKSDPKFTT	ISMDSQMA	KKATTDIKQ	KFDNAQITK	FFERNISKFG	VKRGTLMSIK	TLKDDKQKTH	YHTIYEVEGI
120	130	140	150	160	170	180	190	200	210	220
PVYGRIVFT	TEKSTHDSI	NGRVDTVFEN	GWRNKIKLS	KEDATARAKA	DIRNEESNKE	KAELYLYNFE	GRFYVYVLVN	SITUSGNEDI	FVMAEDGSIV	NKYNNTPTLL
230	240	250	260	270	280	290	300	310	320	330
DTRAEK.PNA	KKIKDEAKTV	EAKKANMNV	IIDVQOQSVK	GLGRISLQGL	VNINLTYDNG	RVTLEDNRK	IYLYDLKQV	ELDLDLDFYD	SPEGGHKEEL	MRSEILWSKS
340	350	360	370	380	390	400	410	420	430	440
NNNFVDDNQV	NSVDAYANDA	KSYDYTKNL	SPNSLDNKGH	NKGFVHFDK	NLGNAPVWGE	YDSMFTGDGD	GURLSPLAKA	LWVGHLSH	GVTNKQEDLE	YKESGALNE
450	460	470	480	490	500	510	520	530	540	550
SFSDINGTAX	ECEMFPIGED	QWPSDFYGE	IHRDHDFSR	GNQPMHKDF	RDLFVDEHD	WQGVHTWSGI	INHAAVLIAD	GFERWQZDS	KQIMAKIFYI	ANCTYUDQIT
560	570	580	590	600						
DFSKGRIBAV	KVAKDLYGDM	SEEVQIVKNA	FDQNGITATP	OLPL						

图4

10	20	30	40	50	60	70	80	90	100	110
MKSKLLATV	LSAVITFSTV	SAVYAPVUGK	ESKVEPKTTT	ITWEKMEONT	KAAATDITEK	KENNSEELTK	FEKMLSKFG	VOKGSLKNTK	TVEDEKXIN	YHMLIYEVGI
120	130	140	150	160	170	180	190	200	210	220
PVYYGRUVFT	TEKSSHDSI	NGRIDTVFEN	GNWRKIELS	KEDALAKAN	DIEDIKATSK	KTDLYLYNFE	GEPYVYVLVD	LITDNGSUTV	FVNAEDGSIV	NKFNNTPTLI
230	240	250	260	270	280	290	300	310	320	330
DTEQKLPNA	EKIKDEAKKA	SNANWIDVQ	QOSVRGVGKT	SLDGLWNIDV	TYGCKEYVLK	DSNENIYLYD	LENQVDEYDL	YAVLSRFNYK	QILMSKSELI	SNWNNFIAM
340	350	360	370	380	390	400	410	420	430	440
NOVNSVQAV	NIMKTYDYK	NKLNENSIDM	KGNNINGFVH	VGRNYGNAFN	YGPYDGEFFG	DGGGLYFSSL	AKSLDWWGHI	LSHGVTINKES	NLKYVESGA	LNESFSDING
450	460	470	480	490	500	510	520	530	540	550
VAVEGKMFVL	GEMCWAQGV	MEDMENSPOG	GQPAHMKOYK	YKTIENDDNGG	VHWSNGIINH	AAVLVADGIE	KTGAENSKDI	EGRIFYTANC	WUDEITNFA	KCPNDVWQVT
560	570	580	590	600	610	620	630	640	650	660
KELYGEMNY	VKIVKAFDQ	VGITATPOLP	L							

图5

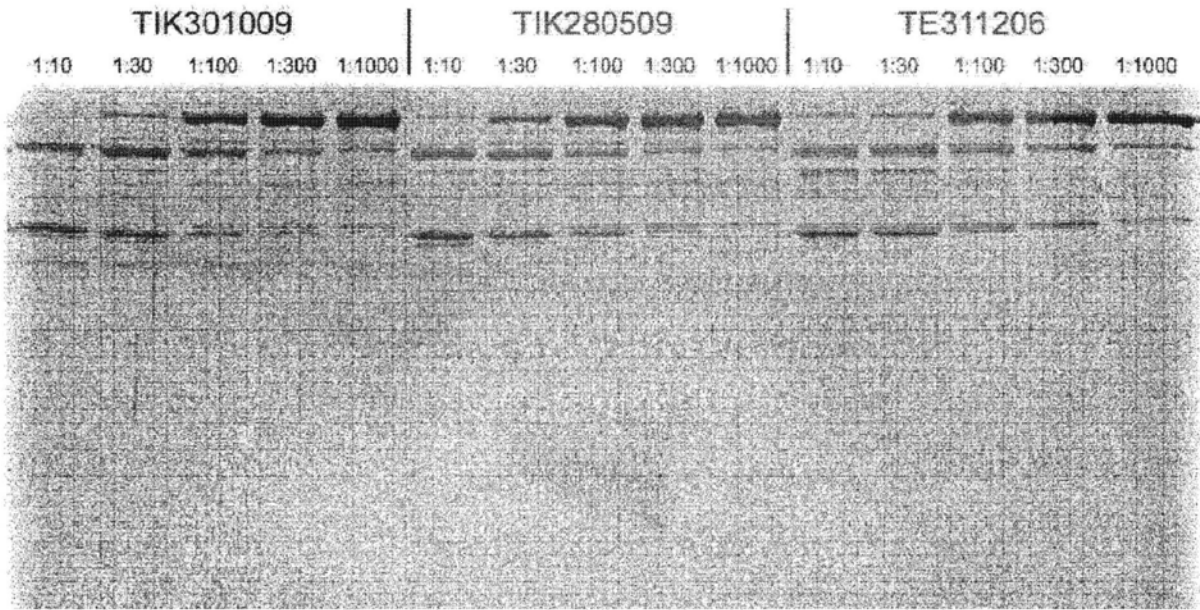


图6A

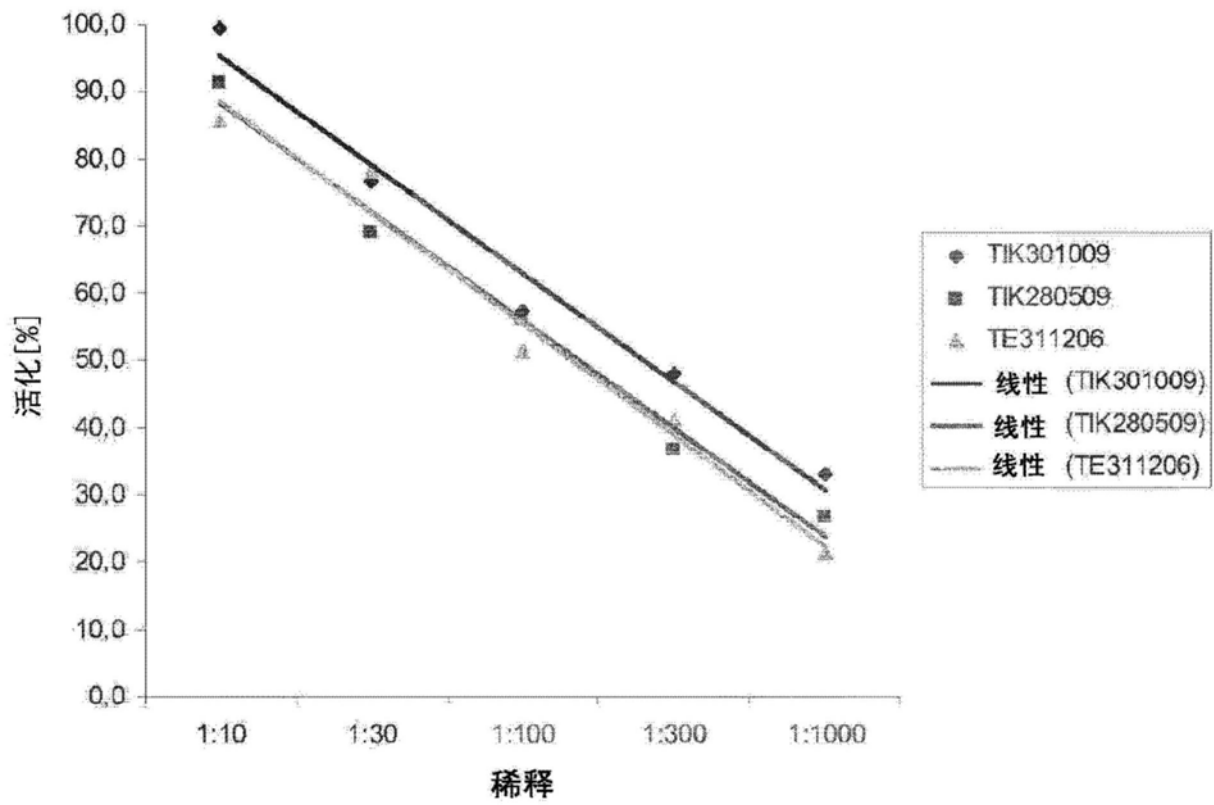


图6B

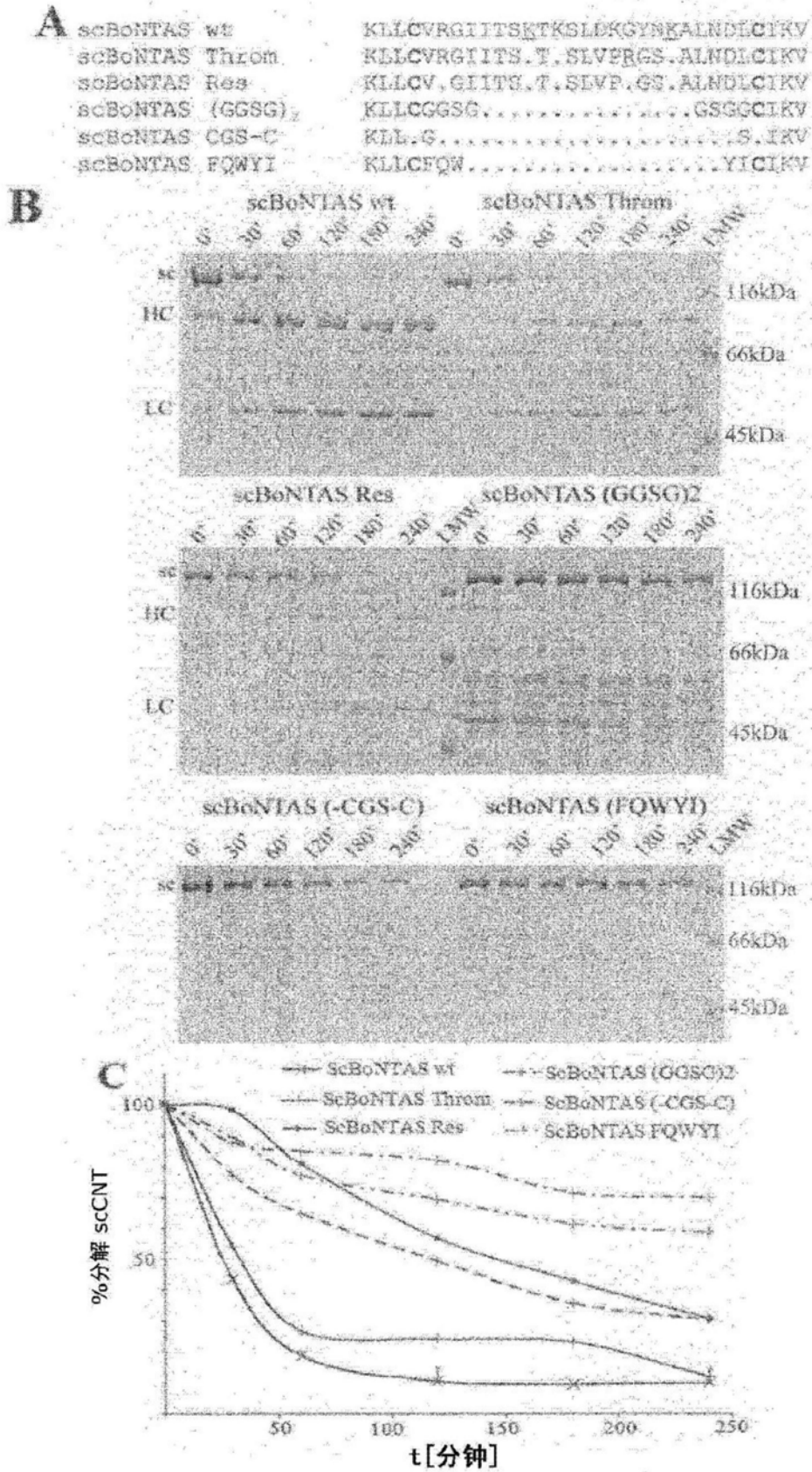


图7

10	20	30	40	50	60	70	80	90	100	110	120	130
MPVWCFYH	KDPJNGYDIA	VIKINAGOR	QVKAIAKIH	KIUVIEPBT	FIMPEBDDN	FPFPAKQPV	SYDSTELST	DRKUNILKG	VTKLPERYS	IDLGRLELTS	IVRCIFUGG	SIUDZLAKVI
140	150	160	170	180	190	200	210	220	230	240	250	260
VDICLVIOF	DGSTESEEN	LVIIGPADI	IOFEKSFCH	EVLNLRNGY	GSTOYIFFP	DFIFGFESL	EVDINPLGA	GRFATFAVI	LAHELHACH	RUTGAIKPN	KVIKATRAY	YEMSCAEVSF
270	280	290	300	310	320	330	340	350	360	370	380	390
EEBFTGGHD	AKFDSIQEN	EPRLYYINK	KDIISTLHKA	KSTVUTASL	QYENVFEEK	YLSIBDTSK	FSVDLAFDK	LYEHLTEIYT	EMFUKPFKY	LMRRTYLNPS	KAVFKINIVF	KYNTTIDGCF
400	410	420	430	440	450	460	470	480	490	500	510	520
MLRMTLAAH	FMGONTIEM	MAFTLEHPT	GLPEYKLC	VBCILTSKTK	SLDKGYNEAL	NDACIKWNU	DLPFSPEDN	FMTDINKGEE	ITSDRWIEAA	EENISIDLLIO	QVYLTNFDN	EPENISISTEM
530	540	550	560	570	580	590	600	610	620	630	640	650
SSBIIIGLEL	MPNIERTPNG	KXYELDKYIK	THYKRAQEFK	HGRSRILATN	SVAEALLNFS	RVTIFFSDY	VKKVUKATEA	ANFLCWVQOL	VYDFIDEISE	VSTIDKIADI	TIIPYIGPA	LNIGNHLYKD
660	670	680	690	700	710	720	730	740	750	760	770	780
DFUGALIFBG	AVILLETYPE	HAIPVIGTFA	LWSTIANEVL	TUQTDINALS	KERNEKQDEVT	KYIVTUNLAK	VNTQILIRK	KRKEALENQA	EATKAIINYO	YKQVTEEEK	NINYNIDDLIS	SZLMSISZKA
790	800	810	820	830	840	850	860	870	880	890	900	910
MINIMKFLNQ	CSVSYLKNSM	IPYGVHRLD	ZDASEPALL	KYIYANGTL	IGAVDRLEK	VNUTLSTIIP	FOLSKTYVNO	RELSTFIIYI	KMLINISILN	LPTVENHLID	LSRTASKINI	GSKVNFPPID
920	930	940	950	960	970	980	990	1000	1010	1020	1030	1040
KNOLQLENLE	SSEKIEVILEN	AIVVNSHYEN	ISTRFVIRIP	KIFENSISLNN	EYIILINCEN	NSGJKVYSLNY	QELIUTLODT	QEIQRVWTK	YSOMINISDY	INRNIIVITII	NRRLMRSKIY	INGELIDORF
1050	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
ISNLGNITHAS	MNIWFKLOGC	ROTHEVITIK	VNUHPEELN	ENFTZIDLVM	QSNNGILKDF	UGDVLQTXP	VYMLNLDPN	KYVDVNNVGI	PGWYMLKGRF	GSGOMINIV	NSSLTRGCTF	IIRKYASGAK
1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300
DIIVERNDRY	YINVTVCSQKE	YSLAINASQA	GVEKILSALE	IFDWCNLSOV	VWKRSHDQO	ITOKCKENLQ	DNMNDICGFI	GFHOFNRIAK	LVAASWVNEQ	IERSSRILQ	SWEIFPVDG	WGERPFL

图8A

10	20	30	40	50	60	70	80	90	100
PPFVFKQFNY	KD-PVNGWDIA	YIKIPNACQM	QPVEAKFIHW	KIWIPIERDT	FIMPERCDLN	PPPEAKQPPV	STYDSTYLST	DNEHOMYLKG	WTKCFEAIYS
110	120	130	140	150	160	170	180	190	200
TDLCERMLTS	IIVRCIPWCC	STIDTELKVI	DTNCLNVIQP	DCSYRSEELN	LVLICPSADI	IQPECKSPGH	EVLNLTMECY	CSTOYIRPSP	DFTFCFEESL
210	220	230	240	250	260	270	280	290	300
EVDINPLGA	GKFAIDPAVI	LAHLEIHACH	RLYGLAINPH	RVPKVINAY	YEHSGLEVSE	BELEIFGCHD	AKFIDSQEN	EFRLYYMKF	KOLASLANKA
310	320	330	340	350	360	370	380	390	400
KSIYGTIASL	QYERWFKK	YLSRDTISCK	FSUDELEKFK	LYRMLTEIYT	EDNFVKFFKV	LMBRTYLMFD	KAVFKINIYV	KVMYTIYDGF	NLRNTNLAAN
410	420	430	440	450	460	470	480	490	500
FNCQNTLNN	HOVTELEKFT	CLFEPYKLLC	VRCITTSRTK	SIDRCYNEAL	EDLICIKVNNW	PLFFSPSEDN	FTNDLNEGER	ITSDTNIEAA	RENISLDLIQ
510	520	530	540	550	560	570	580	590	600
QYTLERFDN	EPENISIEML	SSDIICQLEL	MPNTEPPFC	KRYELDRYTH	PHYLAQRE	HCKSRIALTM	SWEALMDS	EMYFFSSDY	VKRYRATERA
610	620	630	640	650	660	670	680	690	700
AMFLCWEQL	VYDFIDEISE	VSTIDRIADI	TIIIPYICPA	LNICMELYED	DFVCALLIFSC	AVILLEPIPE	LAIDVLCFPA	LVSYIANKVL	TVOQIDNALS
710	720	730	740	750	760	770	780	790	800
KRREKWEVY	KYIVINGLAK	VHTQDLLEK	KIKALENQA	ZATPAINQ	YMOYTEREN	NINFNIDDS	SKLINESINKA	MIMIRKFINQ	CSVSYLNNNS
810	820	830	840	850	860	870	880	890	900
IPYCKRLED	EDASLEKALL	KYIVDMGTE	IGQVLEKDK	VNNTLSTDP	FOLSKYVDMQ	RULSTFTYI	KMLINTSLN	LRYKSHLID	LSRYASKINI
910	920	930	940	950	960	970	980	990	1000
GSKVREPID	RNQIQLYNLE	SSKIEVILRN	AINVMSYEN	FSTFWIRP	KYFNSISLNN	EYTIINCHEN	NECKKVSILNY	CRITITLQDT	QELKQVWFK
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100
YSGRINIEDY	INRWIVTIT	NHRLNNSKIY	INCLDQKP	ISNLCNIRAS	NNIMFKLDEC	RQTHRYIYIK	YFNLDFEELN	EKREILDYDN	QSYSGIIFDF
1110	1120	1130	1140	1150	1160	1170	1180	1190	1200
WCDYLOQDKP	YTHEMLDPR	KYVDVNVGI	RGVTKLQPR	GSVNTINYL	NSZLYRGKRF	LIRKYASENK	DIVVRNDRV	YIMVYKMKKE	YRLATRASQA
1210	1220	1230	1240	1250	1260	1270	1280	1290	1300
CVEKILSALE	IPDVGNLSQV	WVHRSKMDQC	ITWCKNMLQ	DMNCMDICFI	CPHQFNIAK	LVASMYYNRQ	IRSSRTLCC	SWEFIPVDDG	WCRZPL

图8B

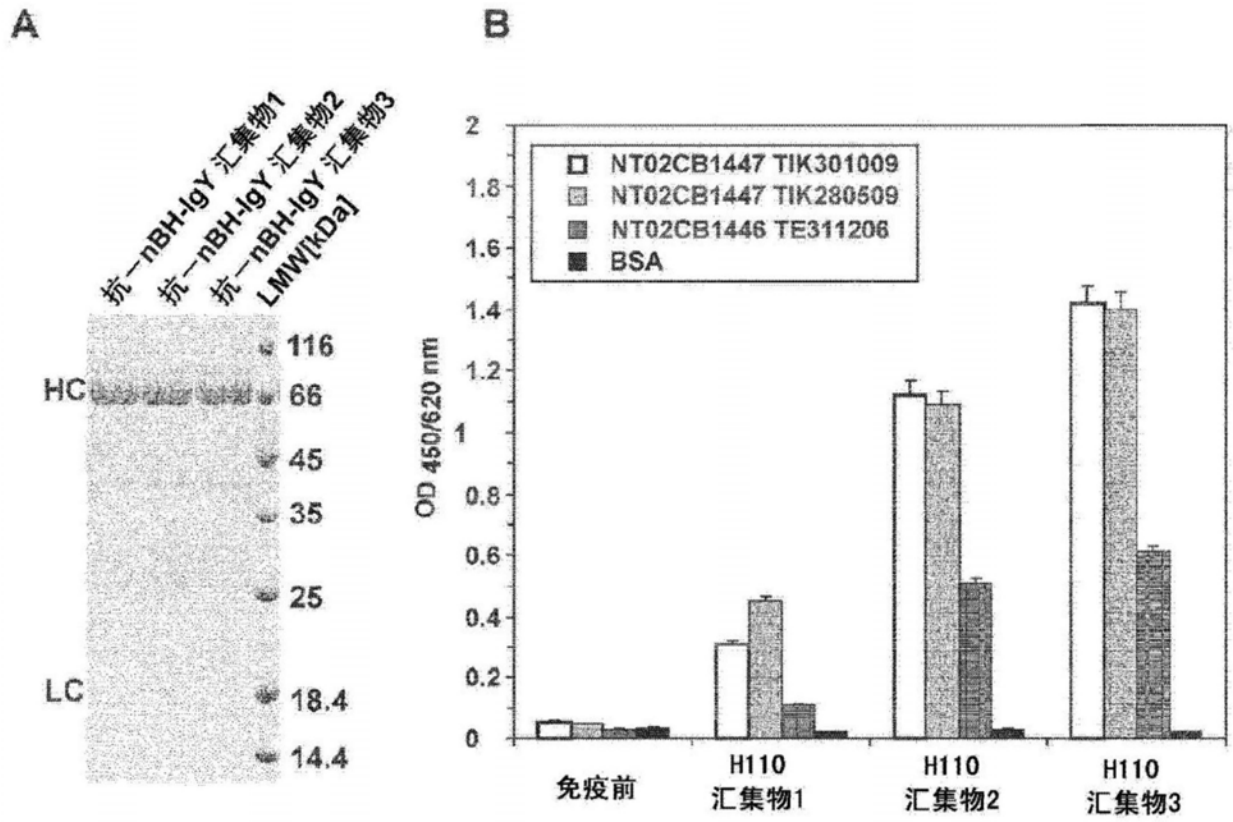


图9

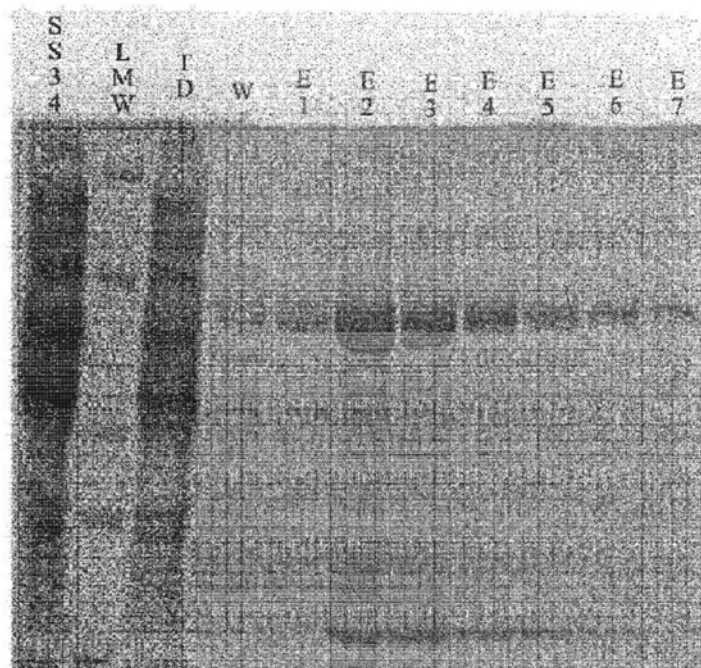


图10A

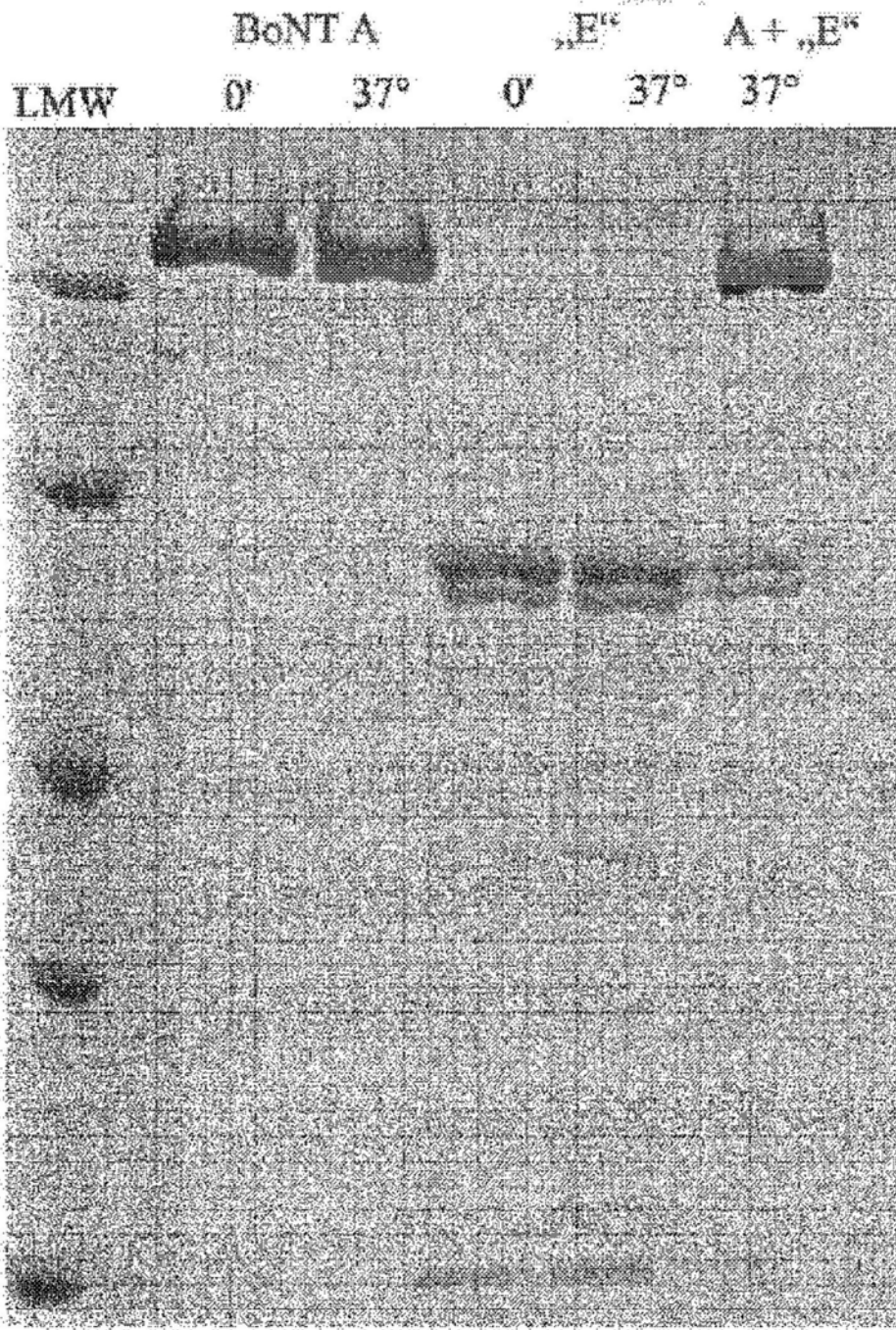


图10B

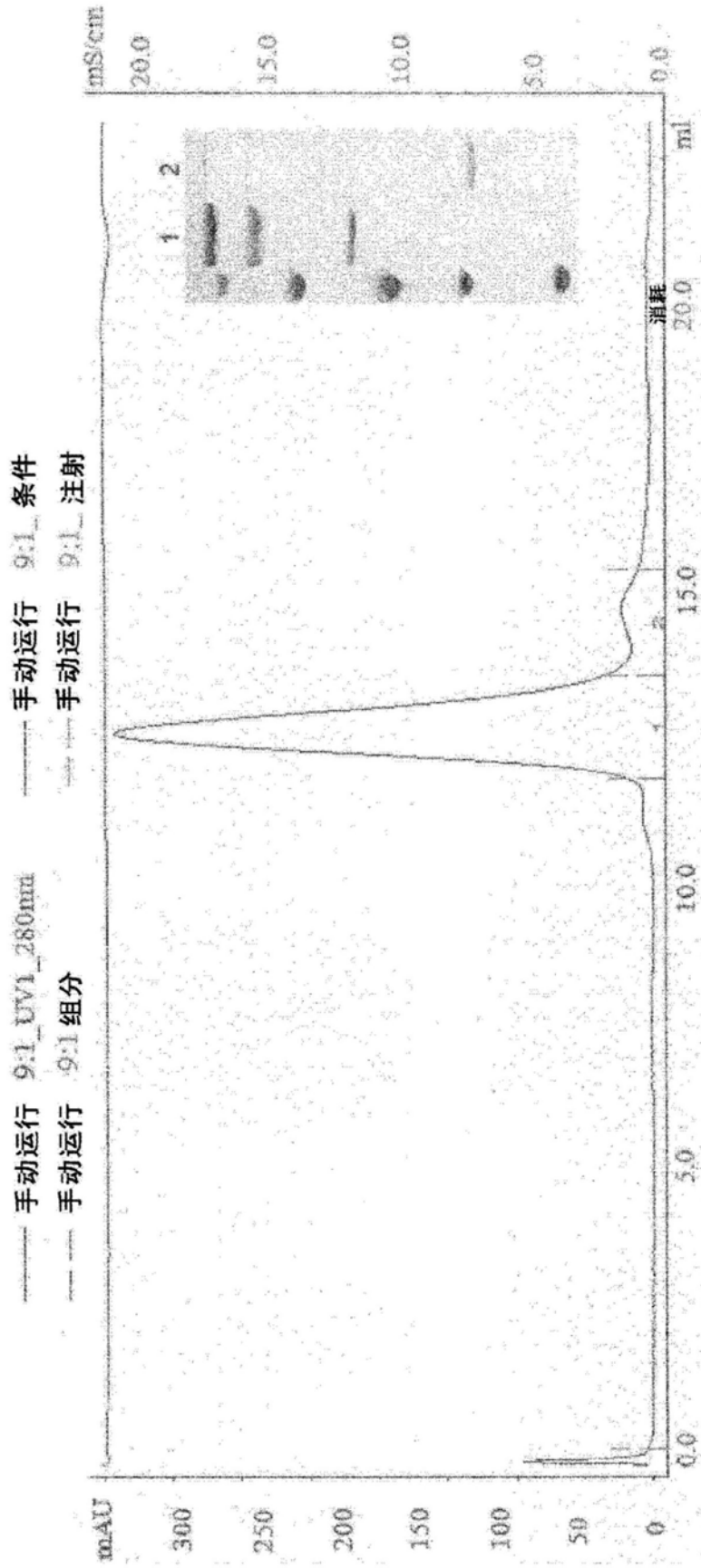


图11A

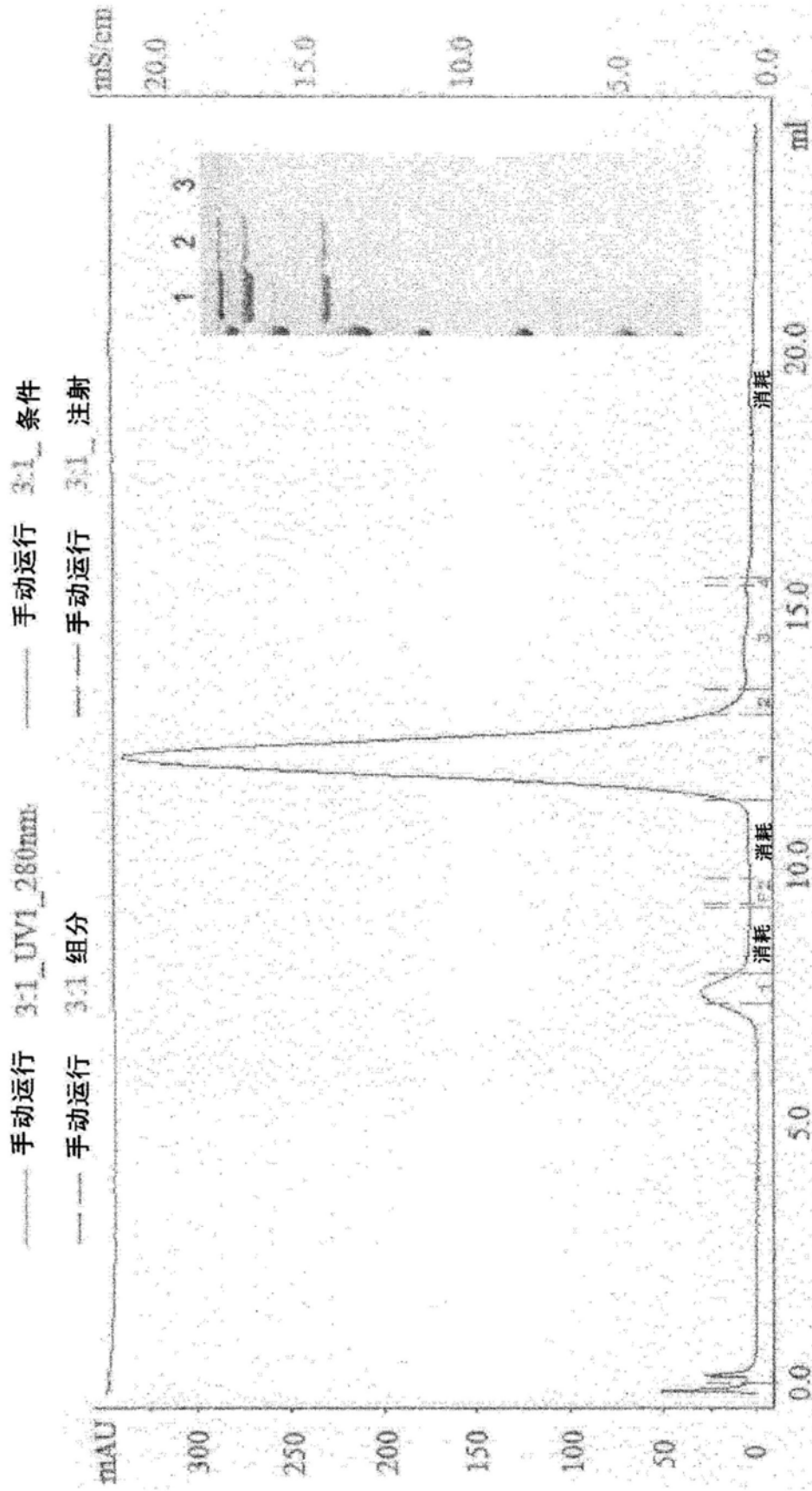


图11B

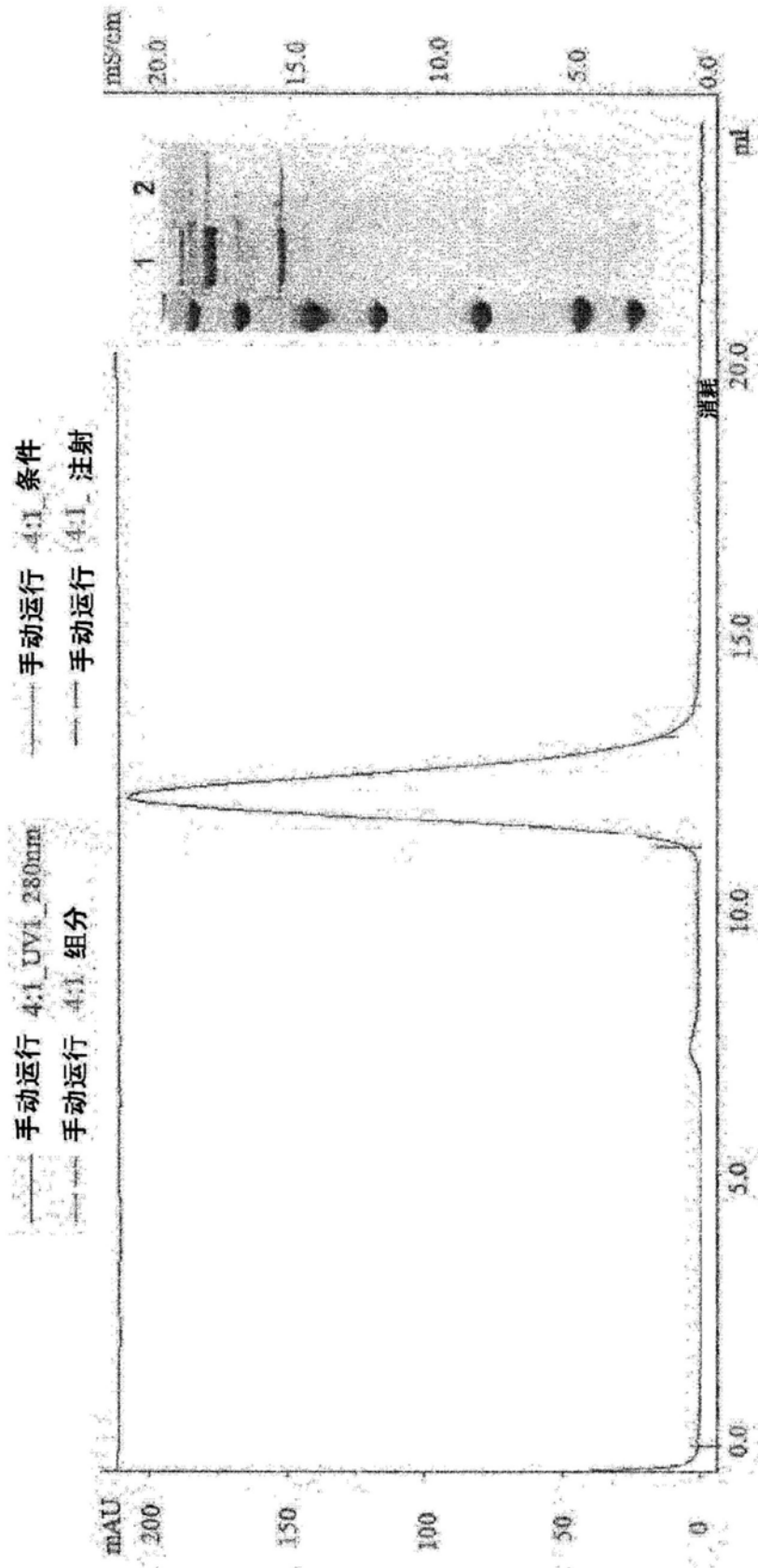


图11C