



(12) 发明专利

(10) 授权公告号 CN 107108703 B

(45) 授权公告日 2022.09.23

(21) 申请号 201580071947.4
 (22) 申请日 2015.01.09
 (65) 同一申请的已公布的文献号
 申请公布号 CN 107108703 A
 (43) 申请公布日 2017.08.29
 (85) PCT国际申请进入国家阶段日
 2017.06.30
 (86) PCT国际申请的申请数据
 PCT/GB2015/050043 2015.01.09
 (87) PCT国际申请的公布数据
 W02016/110662 EN 2016.07.14
 (73) 专利权人 益普生生物创新有限公司
 地址 英国牛津郡
 (72) 发明人 D·B·安德森 G·S·哈克特
 S·M·刘

(74) 专利代理机构 北京市中咨律师事务所
 11247
 专利代理师 沈晓书 黄革生

(51) Int.Cl.
 C07K 14/33 (2006.01)
 A61K 49/00 (2006.01)
 A61K 38/48 (2006.01)
 A61Q 19/08 (2006.01)
 A61K 39/08 (2006.01)

(56) 对比文件
 CN 103974973 A, 2014.08.06
 CN 101107361 A, 2008.01.16
 CN 102481351 A, 2012.05.30
 US 2014242110 A1, 2014.08.28
 WO 2006094269 A3, 2007.01.11
 审查员 徐乐

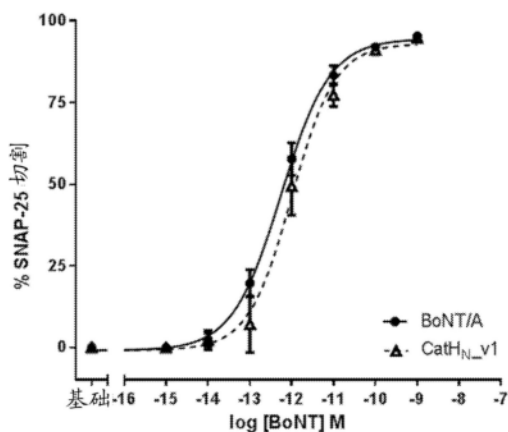
权利要求书1页 说明书39页
 序列表22页 附图12页

(54) 发明名称

阳离子神经毒素

(57) 摘要

本发明提供了包含至少一个氨基酸修饰的工程化的梭菌毒素,其中所述的至少一个氨基酸修饰把工程化的梭菌毒素的等电点(pI)增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.2个pI单位,并且其中所述的至少一个氨基酸修饰不位于梭菌毒素的结合结构域(H_C结构域)。也提供了工程化的梭菌毒素的医疗用途。



	pEC ₅₀
BoNT/A	12.26 ± 0.13
CatH _N _v1	12.00 ± 0.09

1. 工程化的肉毒梭菌神经毒素A (BoNT/A), 其氨基酸序列为SEQ ID No:1、3或5。

2. 核酸, 其编码权利要求1的工程化的BoNT/A。

3. 制备权利要求1的工程化的BoNT/A的方法, 该蛋白质具有轻链和重链, 所述的方法包括在合适的宿主细胞中表达权利要求2的核酸, 裂解宿主细胞以提供包含单链工程化的BoNT/A的宿主细胞匀浆, 以及分离单链工程化的BoNT/A。

4. 权利要求3的方法, 其进一步包括使所述单链工程化的BoNT/A与在位于轻链和重链之间的识别位点处切割所述单链工程化的BoNT/A的蛋白酶接触, 以及把所述单链工程化的BoNT/A转化为双链BoNT/A, 其中轻链和重链通过二硫键连接在一起。

5. 权利要求1的工程化的BoNT/A或由权利要求4的方法获得的工程化的BoNT/A在制备用于预防或治疗选自下列的疾病或病症的药物中的用途: 斜视、斜眼、张力失常、斜颈、神经肌肉障碍或眼能动性病症、夜磨牙症、肝豆状核变性、震颤、抽搐、节段性肌阵挛、痉挛、由慢性多发性硬化症引起的痉挛状态、导致异常的膀胱控制的痉挛状态、提肌骨盆综合征、脊柱裂、迟发性运动障碍、帕金森氏病、口吃、眼睑障碍、脑性瘫痪、病灶性痉挛状态、痉挛性结肠炎、神经源性膀胱、四肢痉挛状态、肛裂、弛缓不能、吞咽困难、流泪、多汗症、过度流涎、过度胃肠分泌、肌肉疼痛、头痛、骨肿瘤、子宫障碍、泌尿生殖器障碍、慢性神经原性炎症、平滑肌障碍以及皮肤皱纹。

6. 权利要求5所述的用途, 其中:

所述张力失常是痉挛性张力失常、口颞肌张力失常、病灶性张力失常、迟发性张力失常、喉部张力失常、四肢张力失常或子宫颈张力失常; 或

所述斜颈是痉挛性斜颈; 或

所述神经肌肉障碍或眼能动性病症是共同性斜视、垂直斜视、外直肌瘫痪、眼球震颤、甲状腺机能障碍性肌病; 或

所述肌肉疼痛是来自肌肉痉挛的疼痛或者四头肌僵痛; 或

所述皮肤皱纹是额皱纹; 或

所述痉挛是眼睑痉挛、指痉挛、背部痉挛、半面痉挛或肛门痉挛; 或所述泌尿生殖器障碍是泌尿生殖器-神经障碍; 或

所述头痛是紧张性头痛。

阳离子神经毒素

[0001] 本发明涉及包含至少一个氨基酸修饰的工程化的梭菌毒素,以及这样的工程化的梭菌毒素在医学和治疗中的用途。

[0002] 梭菌属(Clostridia)中的细菌产生高度强大的并且特异性的蛋白毒素,它们能够毒害它们被递送之处的神经元和其它细胞。这样的梭菌毒素的实例包括由破伤风梭菌(C.tetani)(TeNT)和肉毒梭菌(C.botulinum)(BoNT)血清型A-G产生的神经毒素以及由巴氏梭菌(C.baratii)和丁酸梭菌(C.butyricum)产生的那些神经毒素。

[0003] 在梭菌毒素中有一些是已知最强大的毒素。例如,肉毒梭菌神经毒素对于小鼠的半数致死剂量(LD₅₀)值范围为0.5至5ng/kg,这取决于血清型。破伤风梭菌毒素和肉毒梭菌毒素两者都通过抑制受影响的神经元的功能(特别是神经递质的释放)而起作用。虽然肉毒梭菌毒素在神经肌肉接头处起作用并且抑制周围神经系统中胆碱能的传导,但是破伤风梭菌毒素却在中枢神经系统中起作用。

[0004] 在自然界中,梭菌毒素是以单链多肽形式被合成的,通过一种蛋白酶解切割事件对该多肽进行翻译后的修饰而形成通过二硫键连接在一起的两条多肽链。切割发生在特异性的切割位点(常常被称为活化位点)处,其位于提供链间二硫键的半胱氨酸残基之间。正是这种双链形式成为了毒素的活性形式。这两条链被称为重链(H-链)(具有大约100kDa的分子量)和轻链(L-链)(具有大约50kDa的分子量)。H-链包含N-末端易位元件(H_N结构域)和C-末端靶向元件(H_C结构域)。切割位点位于L-链和易位结构域元件之间。在H_C结构域与它的靶标神经元结合以及所结合的毒素经由内体内化到细胞中以后,H_N结构域使L-链易位跨越内体膜并且进入胞质中,并且L-链提供蛋白酶功能(也被称为无细胞毒性的蛋白酶)。无细胞毒性的蛋白酶通过蛋白酶解切割被称为SNARE蛋白的细胞内转运蛋白(例如SNAP-25、VAMP或突触融合蛋白)而起作用-参见Gerald K(2002)“Cell and Molecular Biology”(第4版)John Wiley&Sons, Inc.首字母缩略词SNARE源于术语可溶性的NSF连接受体,其中NSF意指N-乙基马来酰亚胺-敏感因子。SNARE蛋白是细胞内的小泡融合体所必需的,因此是经由小泡转运的来自细胞的分子分泌所必需的。蛋白酶功能是锌依赖性的内肽酶活性并且显示对SNARE蛋白的高度底物特异性。因此,一旦被递送到所要的靶细胞,无细胞毒性的蛋白酶就能够抑制来自靶细胞的细胞分泌。梭菌毒素的L-链蛋白酶是切割SNARE蛋白的无细胞毒性的蛋白酶。

[0005] 鉴于SNARE蛋白的遍在本性,梭菌毒素例如肉毒梭菌毒素已经被成功地应用在广泛的治疗中。

[0006] 举例来说,我们参考William J.Lipham,Cosmetic and Clinical Applications of Botulinum Toxin(Slack, Inc., 2004),它描述了梭菌毒素,例如肉毒梭菌神经毒素(BoNTs)、BoNT/A、BoNT/B、BoNT/C₁、BoNT/D、BoNT/E、BoNT/F和BoNT/G以及破伤风梭菌神经毒素(TeNT)的用途,用于在许多治疗和美容化妆或美学应用中抑制神经元传导-例如,市售的肉毒梭菌毒素产品目前被批准作为下列适应证的治疗剂,包括局灶性痉挛、上肢痉挛、下肢痉挛、颈肌张力障碍、眼睑痉挛、半面痉挛、腋窝多汗症、慢性偏头痛、神经源性逼尿肌过度活动、眉间线和严重的外眦线。另外,描述了梭菌毒素治疗用于治疗神经肌肉障碍(参见

US 6,872,397);用于治疗子宫障碍(参见US 2004/0175399);用于治疗溃疡和胃食管反流病(参见US 2004/0086531);用于治疗肌张力障碍(参见US 6,319,505);用于治疗眼障碍(参见US 2004/0234532);用于治疗眼睑痉挛(参见US 2004/0151740);用于治疗斜视(参见US 2004/0126396);用于治疗疼痛(参见US 6,869,610、US 6,641,820、US 6,464,986和US 6,113,915);用于治疗纤维肌痛(参见US 6,623,742、US 2004/0062776);用于治疗下背痛(参见US 2004/0037852);用于治疗肌肉损伤(参见US 6,423,319);用于治疗窦性头痛(参见US 6,838,434);用于治疗紧张性头痛(参见US 6,776,992);用于治疗头痛(参见US 6,458,365);用于治疗减少偏头痛性头痛(参见US 5,714,469);用于治疗心血管疾病(参见US 6,767,544);用于治疗神经障碍例如帕金森病(参见US 6,620,415、US 6,306,403);用于治疗神经精神障碍(参见US 2004/0180061、US 2003/0211121);用于治疗内分泌障碍(参见US 6,827,931);用于治疗甲状腺障碍(参见US 6,740,321);用于治疗胆碱能影响的汗腺障碍(参见US 6,683,049);用于治疗糖尿病(参见US 6,337,075、US 6,416,765);用于治疗胰腺障碍(参见US 6,261,572、US 6,143,306);用于治疗癌症例如骨肿瘤(参见US 6,565,870、US 6,368,605、US 6,139,845、US 2005/0031648);用于治疗耳障碍(参见US 6,358,926、US 6,265,379);用于治疗自主性障碍例如胃肠肌肉障碍和其它平滑肌功能障碍(参见US 5,437,291);用于治疗与皮肤细胞增殖障碍相关的皮肤损伤(参见US 5,670,484);用于控制神经原性炎症障碍(参见US 6,063,768);用于减少脱发和刺激毛发生长(参见US 6,299,893);用于治疗嘴角朝下(参见US 6,358,917);用于减少食欲(参见US 2004/40253274);用于牙齿的治疗和牙科手术(参见US 2004/0115139);用于治疗神经肌肉障碍和病症(参见US 2002/0010138);用于治疗多种障碍和病症以及相关的疼痛(参见US 2004/0013692);用于治疗由粘液分泌过多引起的病症例如哮喘和COPD(参见WO 00/10598);以及用于治疗非神经元病症例如炎症、内分泌病症、外分泌病症、免疫病症、心血管病症、骨病症(参见WO 01/21213)。将全部以上出版物的全部内容并入本文作为参考。

[0007] 预期无细胞毒性的蛋白酶例如梭菌毒素(例如BoNT和TeNT)在治疗性和美容化妆性处理人和其它哺乳动物中的用途将要扩大到范围日益宽广的能够受益于这些毒素性质的疾病和不适。

[0008] 为了避免系统性的神经作用,许多梭菌毒素疗法利用向给定的靶位点(例如靶组织)直接施用梭菌毒素治疗剂。当以这种方式施用基于梭菌毒素的治疗剂时的一个问题是毒素远离施用位点而进入周围组织或系统性循环的传播。毒素远离靶组织的扩散被认为对不想要的副作用负责,这些副作用在极端情况下可能会是威胁生命的。当以高的剂量、浓度和注射体积使用梭菌毒素治疗剂(例如BoNT治疗剂)时,这可能是一个特别的关注。对于商业化的BoNT/A治疗剂,已经报道的与这个问题相关的副作用包括:无力、广泛性肌肉虚弱、复视、下垂、吞咽困难、发声困难、构音困难、尿失禁和呼吸困难。吞咽和呼吸困难可能是威胁生命的并且已经报道了与毒素作用的传播相关的死亡。

[0009] 因此,在本领域中存在对这样的梭菌毒素的需要:与已知的梭菌毒素相比,它在施用位点具有增大的组织保留性质并且因此显示出远离施用位点的扩散的减少。

[0010] 本发明通过提供工程化的梭菌毒素解决了上述问题,正如权利要求书中具体说明的那样。

[0011] 一个方面,本发明提供了包含至少一个(例如至少一个、两个或三个)氨基酸修饰

的工程化的梭菌毒素,其中所述的至少一个氨基酸修饰把工程化的梭菌毒素的等电点(pI)增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.2个(例如至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位,并且其中所述的至少一个氨基酸修饰不位于梭菌毒素结合结构域(H_C结构域)中。

[0012] 在一个实施方案中,“不位于梭菌毒素结合结构域(H_C)结构域中”是指所述的至少一个氨基酸修饰位于梭菌毒素的H_N结构域或梭菌毒素的轻链中。

[0013] 在一个实施方案中,本发明提供了包含至少一个(例如至少一个、两个或三个)氨基酸修饰的工程化的梭菌毒素,其中所述的至少一个氨基酸修饰把工程化的梭菌毒素的等电点(pI)增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.2个(例如至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位,并且其中所述的至少一个氨基酸修饰不位于梭菌毒素的易位结构域(H_N结构域)中。

[0014] 在另一个实施方案中,本发明提供了包含至少一个(例如至少一个、两个或三个)氨基酸修饰的工程化的梭菌毒素,其中所述的至少一个氨基酸修饰把工程化的梭菌毒素的等电点(pI)增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.2个(例如至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位,并且其中所述至少一个氨基酸修饰位于梭菌毒素的轻链中。

[0015] 在一个实施方案中,其中所述的至少一个氨基酸修饰位于梭菌毒素轻链中,所述的至少一个氨基酸修饰不向梭菌毒素的轻链引入E3连接酶识别基序。因此,在一个实施方案中,本发明的工程化的梭菌毒素的轻链不包含E3连接酶识别基序。

[0016] 如上所使用的,术语“E3连接酶识别基序”是指轻链的修饰,其导致神经毒素多肽被应用神经毒素的个体中存在的内源性蛋白酶体降解途径加速降解。“E3连接酶识别”基序是允许通过E3连接酶(也称为E3泛素连接酶;因此,“E3连接酶识别基序”也可以称为“E3泛素连接酶识别基序”)识别基序并且结合基序的结构基序。E3连接酶识别基序是本领域技术人员熟悉的。

[0017] E3连接酶识别基序的实例包括以下序列(其中“X”可以表示任何天然存在的氨基酸):

E3 泛素连接酶	识别基序 (共有)
VBCCul2	ALAPYIP (SEQ ID NO: 9)
MDM2	XXFXXXWXXLXX (SEQ ID NO: 10)
MNM2	RFMDYWEGL (SEQ ID NO: 11) FXXXLWXXL (SEQ ID NO: 12)
[0018] Smurf2	ELESPPPPYSRYPM (SEQ ID NO: 13)
RN181	KVGFFKR (SEQ ID NO: 14)
E3alpha	LLVRGRTLTVV (SEQ ID NO: 15)
SCF	DRHDSGLDSM (SEQ ID NO: 16)
Siah	PXAXVXP (SEQ ID NO: 17)
Itch	PPXYXXM (SEQ ID NO: 18)
Nedd4-2	PPXY (SEQ ID NO: 19)

[0019] E3连接酶识别基序的其它实例包括: ETFSDLWKLLE (SEQ ID NO: 20), TSFAEYWLLSP (SEQ ID NO: 21), LTFEHYWAQLTS (SEQ ID NO: 22), LTFEHWWAQLTS (SEQ ID NO: 23), LTFEHSWAQLTS (SEQ ID NO: 24), ETFEHNWAQLTS (SEQ ID NO: 25), LTFEHNWAQLTS (SEQ ID NO: 26), LTFEHWASLTS (SEQ ID NO: 27), LTFEHWSSLTS (SEQ ID NO: 28), LTFTHWAQLTS (SEQ ID NO: 29), ETFEHWWAQLTS (SEQ ID NO: 30), LTFEHWSQLTS (SEQ ID NO: 31), LTFEHWWAQLLS (SEQ ID NO: 32), ETFEHWSQLLS (SEQ ID NO: 33), RFMDYWEGL (SEQ ID NO: 34), MPRFMDYWEGLN (SEQ ID NO: 35), SQETFSDLWKLLEPEN (SEQ ID NO: 36), 和LTFEHNWAQLEN (SEQ ID NO: 37)。

[0020] 在一个实施方案中,其中所述的至少一个氨基酸修饰位于梭菌毒素的轻链中,所述的至少一个氨基酸修饰不向梭菌毒素的轻链引入MDM2 E3连接酶识别基序。因此,在一个实施方案中,本发明的工程化的梭菌毒素的轻链不包含MDM2 E3连接酶识别基序。

[0021] 在一个实施方案中,其中所述的至少一个氨基酸修饰位于梭菌毒素的轻链中,工程化的梭菌毒素在N-末端脯氨酸不包含氨基酸修饰。

[0022] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,并且其中所述的至少一个氨基酸修饰位于梭菌毒素的轻链中,所述工程化的BoNT/E在以下氨基酸位置中的任何一个上不包含用赖氨酸的取代:Q53、N72、N378、N379、R394、T400。

[0023] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,并且其中所述的至少一个氨基酸修饰位于梭菌毒素的轻链中,所述的工程化的BoNT/E在以下氨基酸位置中的任何一个上不包含用赖氨酸的取代:Q53、N72、N378、N379、T400。

[0024] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,并且其中所述的至少一个氨基酸修饰位于梭菌毒素的轻链中,所述的工程化的BoNT/E在以下氨基酸位置中的任何三个上不包含用赖氨酸的取代:Q53、N72、N378、N379、R394、T400。

[0025] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,并且其中所述的至少一个

氨基酸修饰位于梭菌毒素的轻链中,所述的工程化的BoNT/E在以下氨基酸位置中的任何三个上不包含用赖氨酸的取代:Q53、N72、N378、N379、T400。

[0026] 在一个实施方案中,任选地其中至少一个氨基酸修饰位于梭菌毒素的轻链中,工程化的梭菌毒素不是BoNT/E。

[0027] 本发明的工程化的梭菌毒素不包含位于梭菌毒素H_c结构域中的任何氨基酸修饰。因此,在本发明的工程化的梭菌毒素中,所述的至少一个氨基酸修饰不位于梭菌毒素的H_c结构域中。

[0028] 在一个实施方案中,其中所述的至少一个氨基酸修饰位于梭菌毒素的轻链中,所述的至少一个氨基酸修饰不包括用赖氨酸残基取代氨基酸残基。

[0029] 在一个实施方案中,其中工程化的梭菌毒素是如上所述的工程化的梭菌毒素,所述的至少一个氨基酸修饰包括用赖氨酸或精氨酸残基取代酸性氨基酸残基或不带电荷的氨基酸残基。

[0030] 在一个实施方案中,其中工程化的梭菌毒素是如上所述的工程化的梭菌毒素,所述的至少一个氨基酸修饰包括用精氨酸残基取代酸性氨基酸残基或不带电荷的氨基酸残基。

[0031] 在一个实施方案中,其中工程化的梭菌毒素是如上所述的工程化的梭菌毒素,其中所述的至少一个氨基酸修饰把工程化的梭菌毒素的pI增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.4个pI单位。在一个实施方案中,所述的至少一个氨基酸修饰把工程化的梭菌毒素的pI增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.5个pI单位。在一个实施方案中,所述的至少一个氨基酸修饰把工程化的梭菌毒素的pI增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.6个pI单位。在一个实施方案中,所述的至少一个氨基酸修饰把工程化的梭菌毒素的pI增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.8个pI单位。在一个实施方案中,所述的至少一个氨基酸修饰把工程化的梭菌毒素的pI增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少1个pI单位。

[0032] 在某些实施方案中,工程化的梭菌毒素包含至少2、3、4、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85或90个氨基酸修饰。

[0033] 在某些实施方案中,所述的至少一个氨基酸修饰把工程化的梭菌毒素的pI增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少2、3、4或5个pI单位。

[0034] 在某些实施方案中,工程化的梭菌毒素包含至少3个氨基酸修饰,并且所述的至少3个氨基酸修饰把工程化的梭菌毒素的pI增大到一个值,该值比缺乏所述的至少3个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.2个pI单位。

[0035] 在某些实施方案中,工程化的梭菌毒素包含至少5个氨基酸修饰,并且所述至少5个氨基酸修饰把工程化的梭菌毒素的pI增大到一个值,该值比缺乏所述的至少5个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.5个pI单位。

[0036] 本发明人已经发现,通过增大梭菌毒素的pI,例如增大至少0.2个pI单位或0.5个pI单位或1个pI单位(通过向梭菌毒素蛋白中引入至少一个氨基酸修饰),所得的工程化的

梭菌毒素有利地显示了增大的组织保留性质和减少的远离施用位点的扩散,同时保留了靶细胞结合、易位以及靶SNARE蛋白切割的能力。因此,与缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素相比,所述的梭菌毒素从施用位点的传播被显著地减少了。

[0037] 本发明的工程化的梭菌毒素适用于任何上述的治疗,并且与使用已知的梭菌毒素治疗剂相比,可以有利地显示副作用的减少或不存在。

[0038] 本发明的工程化的梭菌毒素的增大的组织保留性质还提供了增加的作用效力和/或持续时间,并且与已知的梭菌毒素治疗剂相比,能够允许使用减少的剂量(或增大的剂量而没有任何额外的副作用),从而提供了进一步的优点。

[0039] 因此,在一个实施方案中,与相应的未修饰的梭菌毒素相比,本发明的工程化的梭菌毒素具有增大的效力、增大的组织保留和/或增长的作用持续时间。

[0040] 正如下面更详细地论述的那样,由至少一个氨基酸修饰所提供的pI的增大意思是本发明的工程化的梭菌毒素在给定的pH下具有的净电荷比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素上的净电荷更正。不希望受任何一种理论限制,本发明人相信,这种增大的正电荷允许本发明的工程化的梭菌毒素在施用位点处展示更长的组织保留时间,这归因于在施用位点处,在工程化的梭菌毒素与阴离子细胞外组分(例如细胞膜和硫酸肝素蛋白聚糖)之间有利的静电相互作用。这些改善的静电相互作用用于减少了工程化的梭菌毒素远离施用位点的扩散,从而改善了组织保留。

[0041] 举例来说,本发明的工程化的梭菌毒素的改善的组织保留性质可以允许(i)更高的剂量进入个别的肌肉,例如胸锁乳突肌,而不会传播到附近的颈部肌肉中而引起吞咽困难,以及(ii)在单次治疗中更高的总剂量(对于所有的肌肉),而不会传播到循环中,引起系统性的效应例如呼吸困难。对于患者的益处可以包括对大肌肉(例如胸锁乳突肌)的更有效的治疗、增大的在每次治疗期间注射几块不同的肌肉的机会以及由于更高的给药而导致的可能更长的有效治疗持续时间(再次治疗之前更长的持续时间是必要的)。

[0042] 在一个实施方案中,本发明的工程化的梭菌毒素在使用中具有正的净电荷(例如当工程化的梭菌毒素在使用中位于组织中的所要施用位点时)。

[0043] 等电点(pI)是给定蛋白质的特性。正如本领域中众所周知的那样,蛋白质是由特定的氨基酸(当在一种蛋白质中时,也被称为氨基酸残基)序列组成。20种氨基酸的标准集合中的每一种氨基酸都具有不同的侧链(或R基团),意味着在一种蛋白质中,每一种氨基酸残基显示不同的化学性质例如电荷和疏水性。这些性质可以受周围化学环境(例如温度和pH)的影响。一种蛋白质的总的化学特征将会取决于这些多种因素的总和。

[0044] 某些氨基酸残基(下面详述)具有能够电离的侧链,该侧链可能显示电荷,取决于周围的pH。在给定的pH下,这样的侧链是否带电取决于相关的能够电离的部分的pKa,其中pKa是酸的一个指定质子从一个共轭碱分离的解离常数(Ka)的负对数。

[0045] 例如酸性的残基例如天冬氨酸和谷氨酸具有pKa值大约为4.1(精确的pKa值可能取决于温度、离子强度和能够电离的基团的微环境)的侧链羧酸基团。因此,这些侧链在pH7.4下(经常被称为“生理pH”)显示负的电荷。在低pH值下,这些侧链将会变成质子化的并且失去它们的电荷。

[0046] 相反地,碱性的残基例如赖氨酸和精氨酸具有pKa值大约为10-12的含氮的侧链基团。这些侧链因此在pH7.4下显示正的电荷。在高pH值下,这些侧链将会变成去质子化的并

且失去它们的电荷。

[0047] 因此,蛋白质分子的总的(净)电荷取决于存在于蛋白质中的酸性残基和碱性残基的数目(以及它们的表面暴露程度)和环境的pH。改变环境的pH,就改变了蛋白质上的总的电荷。因此,对于每一种蛋白质,存在着一个给定的pH,在该pH下正电荷和负电荷的数目是相等的并且蛋白质显示总体上无净电荷。这个点被称为等电点(pI)。等电点是蛋白质生物化学中的一个标准的概念,技术人员将会对它很熟悉。

[0048] 因此,等电点(pI)被定义为蛋白质显示零净电荷时的pH值。pI的增大意味着蛋白质显示零净电荷需要更高的pH值。因此,pI的增大表示在给定的pH下,蛋白质的净正电荷的增加。相反地,pI的减小意味着蛋白质显示零净电荷需要更低的pH值。因此,pI的减小表示在给定的pH下,蛋白质的净正电荷的减少。

[0049] 在本领域中确定蛋白质的pI的方法是已知的并且对于技术人员来说将会是熟悉的。举例来说,从存在于蛋白质中的每一种氨基酸的平均pKa值能够把蛋白质的pI计算出来(“计算的pI”)。这样的计算可以使用本领域已知的计算机程序进行;用于计算pI值的优选的示例性计算机程序包括来自Scripps Research Institute的Protein Calculator和来自ExPASy的Compute pI/MW工具。应该使用相同的计算技术/程序来比较不同分子之间的pI值。

[0050] 在适当的情况下,蛋白质的计算的pI可以使用等电聚焦技术在试验中测定出来(“观察到的pI”)。这种技术根据蛋白质的pI使用电泳来分离蛋白质。典型地使用具有固定化pH梯度的凝胶来进行等电聚焦。当应用电场时,蛋白质迁移穿过pH梯度,直到它达到它具有零净电荷时的pH,这个点就是蛋白质的pI。通过等电聚焦提供的结果本质上通常是相对较低的分辨率,因此本发明人认为由计算的pI(如上所述)提供的结果更适合使用。

[0051] 在本说明书中,除非另有说明,否则“pI”是指“计算的pI”。

[0052] 通过改变显示在蛋白质表面上的碱性基团和/或酸性基团的数目,可以增大或减小蛋白质的pI。这能够通过修饰蛋白质的一个或多个氨基酸来实现。例如,可以通过减少酸性残基的数目或通过增大碱性残基的数目来提供pI的增大。下面更详细地讨论这样的氨基酸修饰。

[0053] 天然(未修饰的)梭菌毒素具有大约5-6的pI。因此,在pH7.4下,天然肉毒梭菌毒素具有负净电荷。举例来说,BoNT/A的pI是6.4,并且BoNT/A分子在pH7.4下具有-8的净电荷。如上所述计算出了这些pI值。

[0054] 表1

[0055]

梭菌毒素	pI
BoNT/A	6.4
BoNT/B	5.3
BoNT/C ₁	5.5
BoNT/D	5.5
BoNT/E	6.0
BoNT/F	5.6
BoNT/G	5.2
TeNT	5.8

[0056] 如上所述,在一个实施方案中,本发明的工程化的梭菌毒素包含至少一个氨基酸修饰,其中所述的至少一个氨基酸修饰将工程化的梭菌毒素的等电点(pI)增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少0.2个pI单位。

[0057] 因此,在本发明的上下文中,在工程化的BoNT/A梭菌毒素的上下文中pI增大0.2个单位将会是pI从6.4增大到6.6。

[0058] 如上所述,在一个实施方案中,本发明的工程化的梭菌毒素包含至少一个氨基酸修饰,其中所述的至少一个氨基酸修饰将工程化的梭菌毒素的等电点(pI)增大到一个值,该值比缺乏所述的至少一个氨基酸修饰、其它方面相同的梭菌毒素的pI高至少1个pI单位。

[0059] 因此,在本发明的上下文中,在工程化的BoNT/A梭菌毒素的上下文中pI增大1个单位将会是pI从6.4增大到7.4。

[0060] 在一个实施方案中,工程化的梭菌毒素具有至少为5.5的pI。

[0061] 在一个实施方案中,工程化的梭菌毒素具有至少为6(例如至少6、至少7、至少8或至少9)的pI。

[0062] 在一个实施方案中,工程化的梭菌毒素具有至少为6.5的pI。

[0063] 在一个实施方案中,工程化的梭菌毒素具有至少为7的pI。

[0064] 在一个实施方案中,工程化的梭菌毒素具有6.5至9.5的pI(例如6.5至7.5的pI)。

[0065] 如上所述,本发明的工程化的梭菌毒素具有增大的组织保留性质,该性质也提供了增大的效力和/或作用持续时间,并且与已知的梭菌毒素治疗剂相比能够允许减少使用的剂量(或者允许增大的剂量而没有任何额外的效应)。其中可以定义这些有利的性质(它们代表治疗指数的增大)的一种方式是根据工程化的梭菌毒素的安全比。在这个方面,通过测量相关的动物模型中体重减少的百分数(例如小鼠,其中在7天的施用内检测体重减少),能够从试验上评估梭菌毒素的不希望的效应(由毒素远离所述施用位点的扩散引起的)。相反地,通过趾外展评分(DAS)分析(肌肉瘫痪量度),能够从试验上评估梭菌毒素的希望的中靶效应。可以通过将20 μ L配制在明胶磷酸盐缓冲液中的梭菌毒素注射到小鼠腓肠肌/比目鱼肌复合体中,然后通过使用Aoki的方法评估趾外展评分(Aoki KR, Toxicicon 39:1815-1820;2001)来进行DAS分析。在DAS分析中,通过尾部短暂地悬挂小鼠以便引出特征性的惊恐反应,其中小鼠伸展它的后肢并且外展它的后趾。在注射梭菌毒素后,按照五点量表(0=正常至4=趾外展和腿伸展减少最大)对不同程度的趾外展进行评分。

[0066] 然后将梭菌毒素的安全比表示成在体重下降10%所需的毒素的量(在小鼠中给药后第一个7天内在峰值效应下测得的)与DAS得分为2所需的毒素的量之间的比率。因此,高安全比得分是希望的,并且指示能够有效地使靶肌肉瘫痪并且具有微小的不希望的脱靶效应的毒素。本发明的工程化的毒素具有的安全比高于相当的未修饰的(天然)肉毒梭菌毒素具有的安全比。

[0067] 因此,在一个实施方案中,本发明的工程化的梭菌毒素具有至少为8(例如至少8、9、10、15、20、25、30、35、40、45或50)的安全比,其中安全比是这样计算的:体重变化-10%所需的毒素的剂量(pg/小鼠)除以DASED₅₀(pg/小鼠)[ED₅₀=产生为2的DAS得分所需的剂量]。

[0068] 在一个实施方案中,本发明的工程化的梭菌毒素具有至少为10的安全比。在一个实施方案中,本发明的工程化的梭菌毒素具有至少为15的安全比。

[0069] 本发明的工程化的梭菌毒素包含至少一个氨基酸修饰。所述的至少一个氨基酸修饰增大梭菌毒素的pI,如上所述。在本发明的上下文中,氨基酸修饰是梭菌毒素的氨基酸序列的修饰。这样的修饰可以通过把序列中的一个氨基酸替换为另一个氨基酸(即取代)、通过把新的氨基酸插入序列中或者通过缺失序列的氨基酸来实施。被插入蛋白质的氨基酸序列中的氨基酸也被称为氨基酸残基。

[0070] 在蛋白质中发现的20种标准氨基酸如下所示:

[0071] 表2:

氨基酸			侧链
天冬氨酸	Asp	D	带电荷的(酸性的)
谷氨酸	Glu	E	带电荷的(酸性的)
精氨酸	Arg	R	带电荷的(碱性的)
[0072] 赖氨酸	Lys	K	带电荷的(碱性的)
组氨酸	His	H	不带电荷的(极性的)
天冬酰胺	Asn	N	不带电荷的(极性的)
谷氨酰胺	Gln	Q	不带电荷的(极性的)
丝氨酸	Ser	S	不带电荷的(极性的)
苏氨酸	Thr	T	不带电荷的(极性的)
酪氨酸	Tyr	Y	不带电荷的(极性的)
甲硫氨酸	Met	M	不带电荷的(极性的)
色氨酸	Trp	W	不带电荷的(极性的)
半胱氨酸	Cys	C	不带电荷的(极性的)
[0073] 丙氨酸	Ala	A	不带电荷的(疏水的)
甘氨酸	Gly	G	不带电荷的(疏水的)
缬氨酸	Val	V	不带电荷的(疏水的)
亮氨酸	Leu	L	不带电荷的(疏水的)
异亮氨酸	Ile	I	不带电荷的(疏水的)
脯氨酸	Pro	P	不带电荷的(疏水的)
苯丙氨酸	Phe	F	不带电荷的(疏水的)

[0074] 下列氨基酸被认为是带电荷的氨基酸:天冬氨酸(负的)、谷氨酸(负的)、精氨酸(正的)和赖氨酸(正的)。

[0075] 在pH7.4下,天冬氨酸(pKa 3.1)和谷氨酸(pKa 4.1)的侧链具有负电荷,同时精氨酸(pKa 12.5)和赖氨酸(pKa 10.8)的侧链具有正电荷。天冬氨酸和谷氨酸被称为酸性氨基

酸残基。精氨酸和赖氨酸被称为碱性氨基酸残基。

[0076] 下列氨基酸被认为是不带电荷的极性的(意思是它们能够参与氢键键合)氨基酸:天冬酰胺、谷氨酰胺、组氨酸、丝氨酸、苏氨酸、酪氨酸、半胱氨酸、甲硫氨酸、色氨酸。

[0077] 下列氨基酸被认为是不带电荷的疏水的氨基酸:丙氨酸、缬氨酸、亮氨酸、异亮氨酸、苯丙氨酸、脯氨酸和甘氨酸。

[0078] 通过向梭菌毒素中引入一个或多个增大梭菌毒素中正电荷与负电荷的比率的氨基酸修饰,能够实现梭菌毒素的pI的增大。

[0079] 在一个实施方案中,至少一个氨基酸修饰选自:氨基酸取代、氨基酸插入和氨基酸缺失。

[0080] 在氨基酸取代中,形成梭菌毒素氨基酸序列的一部分的氨基酸残基被不同的氨基酸残基所替换。替换氨基酸残基可以是如上所述的20种标准氨基酸中的一种。

[0081] 可选择的是,在氨基酸取代中,替换的氨基酸可以是一种非标准氨基酸(不是上述20种氨基酸的标准组的一部分的氨基酸)。举例来说,替换氨基酸可以是碱性非标准氨基酸,例如L-鸟氨酸、L-2-氨基-3-胍基丙酸,或赖氨酸、精氨酸和鸟氨酸的D-异构体)。把非标准氨基酸引入蛋白质中的方法是本领域中已知的,并且包括使用大肠杆菌(E.coli)营养缺陷型表达宿主的重组蛋白质合成。

[0082] 在氨基酸插入中,把另外的氨基酸残基(正常不存在的一个氨基酸)掺入梭菌毒素的氨基酸序列中,从而增大了所述的序列中氨基酸残基的总数。在氨基酸缺失中,从梭菌毒素氨基酸序列中除去氨基酸残基,从而减少了所述的序列中氨基酸残基的总数。

[0083] 通过取代、插入或缺失氨基酸残基来修饰蛋白质的方法是本领域中已知的。举例来说,可以通过对编码梭菌毒素的DNA序列的修饰来引入氨基酸修饰。这能够使用标准的分子克隆技术来实现,例如通过定点诱变,其中采用聚合酶,使用为所要的氨基酸编码的DNA的短链(寡核苷酸)来替换原编码序列,或者通过采用多种酶(例如连接酶和限制内切核酸酶)插入/缺失基因的部分。可选择地,能够化学合成修饰的基因序列。

[0084] 在一个实施方案中,至少一个氨基酸修饰选自:酸性氨基酸残基被碱性氨基酸残基取代;酸性氨基酸残基被不带电荷的氨基酸残基取代;不带电荷的氨基酸残基被碱性氨基酸残基取代;插入碱性氨基酸残基;以及缺失酸性氨基酸残基。

[0085] 在优选的实施方案中,至少一个氨基酸修饰是取代,该取代有利地保持了梭菌毒素中氨基酸残基数相同。在一个实施方案中,取代选自:酸性氨基酸残基被碱性氨基酸残基取代、酸性氨基酸残基被不带电荷的氨基酸残基取代以及不带电荷的氨基酸残基被碱性氨基酸残基取代。在一个实施方案中,碱性氨基酸残基是赖氨酸残基或精氨酸残基。在一个实施方案中,碱性氨基酸残基是赖氨酸残基。在一个实施方案中,碱性的氨基酸残基是精氨酸残基。在一个实施方案中,其中取代是酸性氨基酸残基被不带电荷的氨基酸残基取代,酸性氨基酸残基被它的相应的不带电荷的酰胺氨基酸残基替换(即天冬氨酸被天冬酰胺替换,以及谷氨酸被谷氨酰胺替换)。

[0086] 在另一个优选的实施方案中,至少一个氨基酸修饰是酸性氨基酸残基被碱性氨基酸残基取代或者不带电荷的氨基酸残基被碱性氨基酸残基取代。在一个实施方案中,碱性氨基酸残基是赖氨酸残基。在一个实施方案中,碱性氨基酸残基是精氨酸残基。

[0087] 本发明的工程化的梭菌毒素可以包含多于一个的氨基酸修饰。因此,在一个实施

方案中,工程化的梭菌毒素(如上所述)包含1至90个氨基酸修饰(例如1至80个、1至70个、1至60个、1至50个、1至40个、1至30个、1至20个、1至10个、3至50个、3至40个、3至30个、4至40个、4至30个、5至40个、5至30个或10至25个氨基酸修饰)。在一个实施方案中,工程化的梭菌毒素(如上所述)包含1至20个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素(如上所述)包含1至10个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素(如上所述)包含2至20个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素(如上所述)包含2至15个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素(如上所述)包含2至10个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素(如上所述)包含3至20个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素(如上所述)包含3至15个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素(如上所述)包含3至10个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素(如上所述)包含4至40个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素包含至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个或至少10个氨基酸修饰。在一个实施方案中,工程化的梭菌毒素包含至少3个氨基酸修饰(例如至少3个氨基酸取代)。在一个实施方案中,工程化的梭菌毒素包含至少4个氨基酸修饰(例如至少4个氨基酸取代)。每一个所述的氨基酸修饰是如上所述的氨基酸修饰。因此,每一个所述的氨基酸修饰都有助于工程化的梭菌毒素的pI的增大(与缺乏所述的氨基酸修饰而其它方面相同的梭菌毒素的pI相比)。

[0088] 任何不位于梭菌毒素结合结构域(H_c 结构域)的梭菌毒素的氨基酸(即氨基酸残基)都能够如上所述被修饰,只要所述的修饰的结果是增大梭菌毒素的pI(如上所述)。然而,本发明人已经鉴定了特别适合用于修饰的靶的梭菌毒素氨基酸的亚组。

[0089] 优选的靶氨基酸可以具有某些性质。举例来说,优选的靶氨基酸可以:(i)是表面暴露的氨基酸;(ii)位于梭菌毒素蛋白二级结构的外面;(iii)位于梭菌毒素蛋白的不是蛋白质功能所必需的区域中;(iv)是其同一性在梭菌毒素的类型、亚型或血清型之间没有被保留的氨基酸;(iv)是其修饰没有产生预期的泛素化位点的氨基酸;或(v)是前述的任意组合。

[0090] 如上所述,本发明的工程化的梭菌毒素的特点是位于梭菌毒素 H_N 易位结构域或位于梭菌毒素轻链的一个或多个氨基酸修饰。

[0091] 梭菌毒素轻链参考序列的实例包括:

[0092] 肉毒梭菌A型神经毒素:氨基酸残基1-448

[0093] 肉毒梭菌B型神经毒素:氨基酸残基1-440

[0094] 肉毒梭菌 C_1 型神经毒素:氨基酸残基1-441

[0095] 肉毒梭菌D型神经毒素:氨基酸残基1-445

[0096] 肉毒梭菌E型神经毒素:氨基酸残基1-422

[0097] 肉毒梭菌F型神经毒素:氨基酸残基1-439

[0098] 肉毒梭菌G型神经毒素:氨基酸残基1-441

[0099] 破伤风梭菌神经毒素:氨基酸残基1-457。

[0100] 梭菌毒素 H_N 结构域参考序列的实例包括:

[0101] 肉毒梭菌A型神经毒素:氨基酸残基449-871

[0102] 肉毒梭菌B型神经毒素:氨基酸残基443-862

- [0103] 肉毒梭菌C₁型神经毒素:氨基酸残基450-866
- [0104] 肉毒梭菌D型神经毒素:氨基酸残基449-871
- [0105] 肉毒梭菌E型神经毒素:氨基酸残基455-845
- [0106] 肉毒梭菌F型神经毒素:氨基酸残基450-864
- [0107] 肉毒梭菌G型神经毒素:氨基酸残基449-871
- [0108] 破伤风梭菌神经毒素:氨基酸残基456-879。
- [0109] 梭菌毒素H_c结构域参考序列的实例包括:
- [0110] 肉毒梭菌A型神经毒素:氨基酸残基872-1278
- [0111] 肉毒梭菌B型神经毒素:氨基酸残基863-1291
- [0112] 肉毒梭菌C₁型神经毒素:氨基酸残基867-1291
- [0113] 肉毒梭菌D型神经毒素:氨基酸残基872-1276
- [0114] 肉毒梭菌E型神经毒素:氨基酸残基846-1252
- [0115] 肉毒梭菌F型神经毒素:氨基酸残基865-1278
- [0116] 肉毒梭菌G型神经毒素:氨基酸残基872-1297
- [0117] 破伤风梭菌神经毒素:氨基酸残基880-1315。
- [0118] 应将上述参考序列视为指南,因为根据亚血清型可能会发生轻微的变化。举例来说,US 2007/0166332(将其全部内容并入本文作为参考)列举了轻微不同的梭菌序列:
- [0119] 轻链:
- [0120] 肉毒梭菌A型神经毒素:氨基酸残基M1-K448
- [0121] 肉毒梭菌B型神经毒素:氨基酸残基M1-K441
- [0122] 肉毒梭菌C₁型神经毒素:氨基酸残基M1-K449
- [0123] 肉毒梭菌D型神经毒素:氨基酸残基M1-R445
- [0124] 肉毒梭菌E型神经毒素:氨基酸残基M1-R422
- [0125] 肉毒梭菌F型神经毒素:氨基酸残基M1-K439
- [0126] 肉毒梭菌G型神经毒素:氨基酸残基M1-K446
- [0127] 破伤风梭菌神经毒素:氨基酸残基M1-A457。
- [0128] H_n结构域:
- [0129] 肉毒梭菌A型神经毒素:氨基酸残基A449-K871
- [0130] 肉毒梭菌B型神经毒素:氨基酸残基A442-S858
- [0131] 肉毒梭菌C₁型神经毒素:氨基酸残基T450-N866
- [0132] 肉毒梭菌D型神经毒素:氨基酸残基D446-N862
- [0133] 肉毒梭菌E型神经毒素:氨基酸残基K423-K845
- [0134] 肉毒梭菌F型神经毒素:氨基酸残基A440-K864
- [0135] 肉毒梭菌G型神经毒素:氨基酸残基S447-S863
- [0136] 破伤风梭菌神经毒素:氨基酸残基S458-V879。
- [0137] H_c结构域:
- [0138] 肉毒梭菌A型神经毒素:氨基酸残基N872-L1296
- [0139] 肉毒梭菌B型神经毒素:氨基酸残基E859-E1291
- [0140] 肉毒梭菌C₁型神经毒素:氨基酸残基N867-E1291

[0141] 肉毒梭菌D型神经毒素:氨基酸残基S863-E1276

[0142] 肉毒梭菌E型神经毒素:氨基酸残基R846-K1252

[0143] 肉毒梭菌F型神经毒素:氨基酸残基K865-E1274

[0144] 肉毒梭菌G型神经毒素:氨基酸残基N864-E1297

[0145] 破伤风梭菌神经毒素:氨基酸残基I880-D1315。

[0146] 在一个实施方案中,其中所述的至少一个氨基酸修饰位于梭菌毒素的易位结构域(H_N 结构域)中,所述的至少一个氨基酸修饰不位于梭菌毒素带区域(belt region)中。梭菌毒素带区域(通过结构和模型的目视检查确定)定义如下:

[0147] 肉毒梭菌A型神经毒素:氨基酸残基492-545

[0148] 肉毒梭菌B型神经毒素:氨基酸残基472-532

[0149] 肉毒梭菌C₁型神经毒素:氨基酸残基494-543

[0150] 肉毒梭菌D型神经毒素:氨基酸残基489-539

[0151] 肉毒梭菌E型神经毒素:氨基酸残基466-515

[0152] 肉毒梭菌F型神经毒素:氨基酸残基485-536

[0153] 肉毒梭菌G型神经毒素:氨基酸残基489-538

[0154] 破伤风梭菌神经毒素:氨基酸残基506-556。

[0155] 在一个实施方案中,至少一个氨基酸修饰(如上所述)是表面暴露的氨基酸残基的修饰。表面暴露的氨基酸残基是存在于折叠的蛋白质的外部上的那些,并且因此易于接近周围溶剂,与位于折叠的蛋白质内部中的那些氨基酸残基相反。氨基酸残基的表面暴露程度以及因此它暴露于周围溶剂的程度取决于它在折叠的蛋白质中的位置,也取决于蛋白质采用的构象。因此,与具有低表面暴露程度的氨基酸残基的修饰相比,具有高表面暴露程度的氨基酸残基的修饰可能对蛋白质的等电点具有更大的影响。测定氨基酸残基的表面暴露程度的方法是本领域中已知的。举例来说,能够使用计算机程序ArealMol (CCP4计算机程序套件的一部分)来计算给定的蛋白质中氨基酸残基的表面暴露程度。也可以通过目视检查蛋白质晶体结构(例如由X射线晶体学提供的)来鉴定表面暴露的氨基酸残基。在一个实施方案中,表面暴露的氨基酸残基具有至少40的总ArealMol值。

[0156] 在一个实施方案中,至少一个氨基酸修饰包括选自以下的氨基酸残基的修饰:天冬氨酸残基、谷氨酸残基、组氨酸残基、天冬酰胺残基、谷氨酰胺残基、丝氨酸残基、苏氨酸残基、丙氨酸残基、甘氨酸残基、缬氨酸残基、亮氨酸残基和异亮氨酸残基。本发明人已经鉴定该组中的氨基酸残基代表用于根据本发明的修饰的特别合适的靶。

[0157] 在一个实施方案中,其中氨基酸修饰包括选自以下的氨基酸残基的修饰:天冬氨酸残基、谷氨酸残基、组氨酸残基、天冬酰胺残基、谷氨酰胺残基、丝氨酸残基、苏氨酸残基、丙氨酸残基、甘氨酸残基、缬氨酸残基、亮氨酸残基和异亮氨酸残基(如上所述),氨基酸残基被赖氨酸残基或精氨酸残基取代。在一个实施方案中,氨基酸残基被精氨酸残基取代。因此,在一个实施方案中,带负电荷的残基、极性残基或不带电荷的残基被带正电荷的残基取代,从而增加了正电荷与负电荷的比例并且增加了梭菌毒素的pI。

[0158] 在一个实施方案中,至少一个氨基酸修饰(如上所述)包括天冬酰胺氨基酸残基或谷氨酰胺氨基酸残基(二者均为不带电荷、极性残基)的修饰。在一个实施方案中,天冬酰胺或谷氨酰胺氨基酸残基被赖氨酸残基或精氨酸残基(二者均为带正电荷的残基)取代。在一

个实施方案中,天冬酰胺或谷氨酰胺氨基酸残基被赖氨酸残基取代。在一个实施方案中,天冬酰胺或谷氨酰胺氨基酸残基被精氨酸残基取代。

[0159] 在一个实施方案中,工程化的梭菌毒素是BoNT/A。参考BoNT/A序列具有UniProtKB登录号P10845。在一个实施方案中,其中工程化的梭菌毒素是BoNT/A,工程化的梭菌毒素是BoNT/A1。

[0160] 本发明人已经鉴定了代表BoNT/A梭菌毒素H_N结构域中的氨基酸修饰的优选靶的某些氨基酸。

[0161] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/A,所述的工程化的BoNT/A包含至少一个(例如,至少1、2、3、4、5、10、15、20、25、30、35、40、45、50或全部55个)选自以下的氨基酸的修饰:D474、N476、D484、N486、I487、E488、A489、A490、E491、D546、E558、E560、H561、I566、L568、N570、S571、L577、N578、A597、E599、A601、E620、V621、T623、D625、T631、N645、L647、D650、D651、I668、E670、A672、V675、S683、I685、A686、N687、N752、Q753、T755、E756、E757、E758、N760、N761、I762、N763、D825、I831、G832、T847、D848和D858;并且所述的一个或多个氨基酸修饰把工程化的BoNT/A的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/A的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0162] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/A,所述的工程化的BoNT/A包含至少一个(例如,至少1、2、3、4、5、10、15或全部17个)选自以下的氨基酸的修饰:N476、S564、N578、E599、L647、D650、D651、V675、I685、N687、T755、E757、N761、N763、I831、T847和I849;并且所述的一个或多个氨基酸修饰把工程化的BoNT/A的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/A的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0163] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/A,所述的工程化的BoNT/A包含至少一个(例如,至少1、2、3、4、5或全部6个)选自以下的氨基酸的修饰:S564、L647、D650、D651、T847和I849;并且所述的一个或多个氨基酸修饰把工程化的BoNT/A的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/A的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0164] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/A,所述的工程化的BoNT/A包含至少一个(例如,至少1、2、3、4、5或全部6个)选自以下的氨基酸的修饰:N476、N763、N687、E599、I831和N761;并且所述的一个或多个氨基酸修饰把工程化的BoNT/A的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/A的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案

中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0165] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/A,所述的工程化的BoNT/A包含至少一个(例如,至少1、2、3、4或全部5个)选自以下的氨基酸的修饰:N578、V675、I685、T755和E757;并且所述的一个或多个氨基酸修饰把工程化的BoNT/A的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/A的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0166] 在一个实施方案中,工程化的梭菌毒素是BoNT/B。参考BoNT/B序列具有UniProtKB登录号P10844。

[0167] 本发明人已经鉴定了代表BoNT/B梭菌毒素H_N结构域中的氨基酸修饰的优选靶的某些氨基酸。

[0168] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/B,所述的工程化的BoNT/B包含至少一个(例如,至少1、2、3、4、5、10、15、20、25、30、35或全部40个)选自以下的氨基酸的修饰:V443、G444、D453、S468、D533、E534、N535、T545、L548、D549、I550、D552、S557、L564、S566、N582、V584、N609、L619、N632、E633、G637、A646、I655、E657、V662、E669、S670、I672、D673、N739、I740、N748、N750、I818、G819、T834、I842、N845和S858;并且所述的一个或多个氨基酸修饰把工程化的BoNT/B的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/B的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0169] 在一个实施方案中,工程化的梭菌毒素是BoNT/C₁。参考BoNT/C₁序列具有UniProtKB登录号P18640。

[0170] 本发明人已经鉴定了代表BoNT/C₁梭菌毒素H_N结构域中的氨基酸修饰的优选靶的某些氨基酸。

[0171] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/C₁,所述的工程化的BoNT/C₁包含至少一个(例如,至少1、2、3、4、5、10、15、20、25、30、35、40、45或全部50个)选自以下的氨基酸的修饰:L451、D452、C453、E455、V472、T474、D475、L478、N483、E484、E485、E487、I489、L555、S556、D557、N558、E560、D561、E569、N574、S575、T584、G592、Q594、G596、D617、N640、S641、V642、G645、N646、E661、E665、T667、A670、S678、V680、Q681、E682、S750、G751、S759、Q760、V826、G827、N842、T843、N847和N853;并且所述的一个或多个氨基酸修饰把工程化的BoNT/C₁的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/C₁的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0172] 在一个实施方案中,工程化的梭菌毒素是BoNT/D。参考BoNT/D序列具有UniProtKB登录号P19321。

[0173] 本发明人已经鉴定了代表BoNT/D梭菌毒素H_N结构域中的氨基酸修饰的优选靶的某些氨基酸。

[0174] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/D,所述的工程化的BoNT/D包含至少一个(例如,至少1、2、3、4、5、10、15、20、25、30、35、40、45、50、55、60或全部62个)选自以下的氨基酸的修饰:Q469、E470、E473、N474、D479、E480、N482、V483、Q484、N485、S487、D488、S552、N553、N554、V555、E556、N557、I558、L560、T562、S563、V564、G569、S571、N572、G588、Q590、T614、D616、S619、S622、N636、S637、L639、G641、N642、E657、E661、T663、A666、V669、S674、I676、Q677、E678、S746、G747、D749、E751、N752、I753、Q756、N818、V822、G823、E837、N838、T839、N843、N849和N850;并且所述的一个或多个氨基酸修饰把工程化的BoNT/D的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/B的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0175] 在一个实施方案中,工程化的梭菌毒素是BoNT/E。参考BoNT/E序列具有UniProtKB登录号Q00496。在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,工程化的梭菌毒素是BoNT/E3。

[0176] 本发明人已经鉴定了代表BoNT/E梭菌毒素H_N结构域中的氨基酸修饰的优选靶的某些氨基酸。

[0177] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,所述的工程化的BoNT/E包含至少一个(例如,至少1、2、3、4、5、10、15、20、25、30、35、40、45、50或全部52个)选自以下的氨基酸的修饰:D474、N476、E479、E480、D484、N486、I487、E488、A489、A490、E491、E492、L496、D497、Q500、Q501、L504、N507、D509、N510、N514、S516、E518、Q527、L530、N533、I534、E535、N539、Y548、I566、L568、D589、A597、E599、A601、L604、Y612、E620、N645、L647、Y648、D651、E737、E741、Y803、Y824、D825、G828、I831、G832和D835;并且所述的一个或多个氨基酸修饰把工程化的BoNT/E的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/E的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0178] 在一个实施方案中,工程化的梭菌毒素是BoNT/F。参考BoNT/F序列具有UniProtKB登录号YP_001390123。

[0179] 本发明人已经鉴定了代表BoNT/F梭菌毒素H_N结构域中的氨基酸修饰的优选靶的某些氨基酸。

[0180] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/F,所述的工程化的BoNT/F包含至少一个(例如,至少1、2、3、4、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85或全部86个)选自以下的氨基酸的修饰:N463、E464、N468、T469、D474、D475、T476、T477、

N478、N482、N485、N495、I499、Q501、I502、Q505、T506、N508、T509、V511、D513、D521、S522、S526、E527、I528、E529、V534、D535、L536、E549、G550、T552、N553、S558、E566、E567、S568、V586、H587、Q608、D613、A616、D617、S619、N630、N633、N639、E654、V656、E658、L660、T663、L665、V666、S671、I673、G674、S675、S676、E677、N678、T746、N751、L753、E754、E756、N758、I759、N760、N761、S799、S821、I822、N840、S841、E845、L846、S847、S848、T850、N851、D852、I854、L855和I856；并且所述的一个或多个氨基酸修饰把工程化的BoNT/F的等电点(pI)增大到一个值，该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/F的pI高至少0.2个(例如，至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中，所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中，所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中，所述的修饰包括用精氨酸残基取代氨基酸。

[0181] 在一个实施方案中，工程化的梭菌毒素是BoNT/G。参考BoNT/G序列具有UniProtKB登录号Q60393。

[0182] 本发明人已经鉴定了代表BoNT/G梭菌毒素H_N结构域中的氨基酸修饰的优选靶的某些氨基酸。

[0183] 在一个实施方案中，其中工程化的梭菌毒素是BoNT/G，所述的工程化的BoNT/G包含至少一个(例如，至少1、2、3、4、5、10、15、20、25、30、35或全部36个)选自以下的氨基酸的修饰：N480、Q482、N483、N484、T485、E487、D540、N562、N570、N571、N572、T588、V589、T615、D621、N637、E638、E642、N643、I660、E662、I667、E674、S675、V677、G678、N679、S747、N755、D757、L823、D839、I841、D844、S846和L847；并且所述的一个或多个氨基酸修饰把工程化的BoNT/G的等电点(pI)增大到一个值，该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/G的pI高至少0.2个(例如，至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中，所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中，所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中，所述的修饰包括用精氨酸残基取代氨基酸。

[0184] 在一个实施方案中，工程化的梭菌毒素是TeNT。参考TeNT序列具有UniProtKB登录号P04958。

[0185] 本发明人已经鉴定了代表TeNT梭菌毒素H_N结构域中的氨基酸修饰的优选靶的某些氨基酸。

[0186] 在一个实施方案中，其中工程化的梭菌毒素是TeNT，所述的工程化的TeNT包含至少一个(例如，至少1、2、3、4、5、10、15、20、25、30、35、40、45或全部49个)选自以下的氨基酸的修饰：A457、S458、L459、D461、L462、E486、E487、Q490、D491、N497、N504、D557、T571、T572、L573、Q574、N580、S581、N588、S589、T590、S598、Q605、G606、Q608、T631、I633、S640、Q655、E658、G659、N660、E675、I677、E679、T681、V684、A691、E692、S694、T695、Q696、A772、D773、E774、S862、N866、L867和D868；并且所述的一个或多个氨基酸修饰把工程化的TeNT的等电点(pI)增大到一个值，该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的TeNT的pI高至少0.2个(例如，至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中，所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中，所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中，所述的修饰包括用精氨酸残基

取代氨基酸。

[0187] 本发明人已经鉴定了代表BoNT/A梭菌毒素轻链中的氨基酸修饰的优选靶的某些氨基酸。

[0188] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/A,所述的工程化的BoNT/A包含至少一个(例如,至少1、2、3、4、5、10、15、20、25或全部28个)选自以下的氨基酸的修饰:N5、Q7、N9、D12、N15、Q31、D58、N60、D74、N82、T122、D124、E126、Q139、D141、E281、L284、S295、Q311、D326、D334、N377、TYR387、N394、N396、N410、M411和N418;并且所述的一个或多个氨基酸修饰把工程化的BoNT/A的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/A的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0189] 本发明人已经鉴定了代表BoNT/B梭菌毒素轻链中的氨基酸修饰的优选靶的某些氨基酸。

[0190] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/B,所述的工程化的BoNT/B包含至少一个(例如,至少1、2、3、4、5、10、15、20、25、30、35、40或全部41个)选自以下的氨基酸的修饰:N6、N7、N9、N11、D12、N16、N17、N18、D41、E48、E57、N60、D75、D77、N80、E127、N130、N144、E147、E149、E185、N216、D245、E253、N316、D333、E335、D341、N385、D388、N389、E390、E395、E396、D402、D404、E406、E408、Q419、E423和E427;并且所述的一个或多个氨基酸修饰把工程化的BoNT/B的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/B的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0191] 本发明人已经鉴定了代表BoNT/C₁梭菌毒素轻链中的氨基酸修饰的优选靶的某些氨基酸。

[0192] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/C₁,所述的工程化的BoNT/C₁包含至少一个(例如,至少1、2、3、4、5、10、15、20、30、40或全部41个)选自以下的氨基酸的修饰:N6、N7、N9、D12、D15、N18、N31、E32、N55、N59、N75、N120、N121、N125、D128、Q142、N145、N177、N178、Q183、E184、D208、E211、Q247、N255、N311、E335、E339、N343、N368、N386、D389、D390、N391、Q396、N405、N407、N425、E427、D442和N448;并且所述的一个或多个氨基酸修饰把工程化的BoNT/C₁的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/C₁的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0193] 本发明人已经鉴定了代表BoNT/D梭菌毒素轻链中的氨基酸修饰的优选靶的某些氨基酸。

[0194] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/D,所述的工程化的BoNT/D包

含至少一个(例如,至少1、2、3、4、5、10、15、20、30或全部34个)选自以下的氨基酸的修饰:D7、N9、D12、N15、D16、N17、D53、D73、D119、E124、E139、E142、N143、Q177、Q178、N180、E184、E255、N308、D335、N336、N339、N343、N368、N386、D389、D390、N391、D397、N403、N407、E409、E416和N443;并且所述的一个或多个氨基酸修饰把工程化的BoNT/D的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/D的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0195] 本发明人已经鉴定了代表BoNT/E梭菌毒素轻链中的氨基酸修饰的优选靶的某些氨基酸。

[0196] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,所述的工程化的BoNT/E包含至少一个(例如,至少1、2、3、4、5、10、15、20、30、35或全部37个)选自以下的氨基酸的修饰:N5、N8、N10、D11、N14、D15、Q27、E28、Q53、N72、Q75、D117、N118、D121、N122、Q123、N138、N169、N170、N195、Q237、ILE244、Q290、N293、N297、D312、Q344、N362、N365、D366、N370、E373、N378、N379、N383、N390和T397;并且所述的一个或多个氨基酸修饰把工程化的BoNT/E的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/E的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0197] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,所述的工程化的BoNT/E包含至少一个(例如,至少1、2、3、4、5、10、15或20个)选自以下的氨基酸的修饰:N5、N8、N10、D11、N14、D15、Q27、E28、N72、Q75、N118、D121、N122、Q123、N138、Q237、Q290、N297、N362、N365、D366、N378和N379;并且所述的一个或多个氨基酸修饰把工程化的BoNT/E的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/E的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0198] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,所述的工程化的BoNT/E包含至少一个(例如,至少1、2、3、4、5、10、15或全部19个)选自以下的氨基酸的修饰:N5、N8、N10、D11、N14、D15、N72、Q75、N118、N122、Q123、N138、Q237、Q290、Q297、N362、D366、N378和N379;并且所述的一个或多个氨基酸修饰把工程化的BoNT/E的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/E的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0199] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,所述的工程化的BoNT/E包含至少一个(例如,至少1、2、3、4、5或全部6个)选自以下的氨基酸的修饰:N8、N10、Q75、

Q123、N138和Q237；并且所述的一个或多个氨基酸修饰把工程化的BoNT/E的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/E的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0200] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/E,所述的工程化的BoNT/E包含至少一个(例如,至少1、2或全部3个)选自以下的氨基酸的修饰:Q123、N138和Q237；并且所述的一个或多个氨基酸修饰把工程化的BoNT/E的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/E的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0201] 本发明人已经鉴定了代表BoNT/F梭菌毒素轻链中的氨基酸修饰的优选靶的某些氨基酸。

[0202] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/F,所述的工程化的BoNT/F包含至少一个(例如,至少1、2、3、4、5、10、15、20、30或全部35个)选自以下的氨基酸的修饰:N6、N9、N11、D12、N15、D16、D17、E28、D55、D60、D74、N76、E105、E121、N126、E127、N144、D185、N211、Q252、N305、E310、D312、N314、N329、D331、N379、D382、D383、D384、E390、N396、N400、D414和D418；并且所述的一个或多个氨基酸修饰把工程化的BoNT/F的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/F的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0203] 本发明人已经鉴定了代表BoNT/G梭菌毒素轻链中的氨基酸修饰的优选靶的某些氨基酸。

[0204] 在一个实施方案中,其中工程化的梭菌毒素是BoNT/G,所述的工程化的BoNT/G包含至少一个(例如,至少1、2、3、4、5、10、15、20、30、35或全部38个)选自以下的氨基酸的修饰:N4、N7、N9、N11、D12、N15、D17、E48、Q55、D57、N60、D75、D127、Q144、E148、D149、Q150、N178、E185、E208、D211、E255、D315、D332、N334、D340、E383、D387、N388、Q393、N394、E395、N403、E407、E418、E422、E426和N443；并且所述的一个或多个氨基酸修饰把工程化的BoNT/G的等电点(pI)增大到一个值,该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的BoNT/G的pI高至少0.2个(例如,至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中,所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中,所述的修饰包括用精氨酸残基取代氨基酸。

[0205] 本发明人已经鉴定了代表TeNT轻链中的氨基酸修饰的优选靶的某些氨基酸。

[0206] 在一个实施方案中,其中工程化的梭菌毒素是TeNT,所述的工程化的TeNT包含至少一个(例如,至少1、2、3、4、5、10、15、20、30、35或全部36个)选自以下的氨基酸的修饰:N6、

N7、N15、N16、D17、D31、E51、E57、N60、N76、N101、D126、D143、N167、D179、N180、E251、Q257、N313、N316、D318、D335、N337、Q339、N368、N387、D390、D391、N395、D396、E403、D406、E410、N421、D427和E450；并且所述的一个或多个氨基酸修饰把工程化的TeNT的等电点(pI)增大到一个值，该值比缺乏所述的一个或多个氨基酸修饰、其它方面相同的TeNT的pI高至少0.2个(例如，至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9或1个)pI单位。在一个实施方案中，所述的修饰包括用赖氨酸残基或精氨酸残基取代氨基酸。在一个实施方案中，所述的修饰包括用赖氨酸残基取代氨基酸。在一个实施方案中，所述的修饰包括用精氨酸残基取代氨基酸。

[0207] 本发明适用于许多不同种类的梭菌毒素。因此，在本发明的上下文中，术语“梭菌毒素”包括由肉毒梭菌(肉毒梭菌神经毒素血清型A、B、C₁、D、E、F和G)、破伤风梭菌(破伤风神经毒素)、丁酸梭菌(肉毒梭菌神经毒素血清型E)和巴氏梭菌(肉毒梭菌神经毒素血清型F)以及修饰的梭菌毒素或衍生自任何前述的衍生物制备的毒素。术语“梭菌毒素”也包括肉毒梭菌神经毒素血清型H。

[0208] 肉毒梭菌神经毒素(BoNT)是由肉毒梭菌以大的蛋白质复合物的形式制备的，该复合物由BoNT自身与多个辅助蛋白复合组成。目前有8种不同的类别的肉毒梭菌神经毒素，即：肉毒梭菌神经毒素血清型A、B、C₁、D、E、F、G和H，所有这些共有相似的结构和作用模式。基于被特异性的中和抗血清失活能够区分不同的BoNT血清型，这种通过血清型的分类与在氨基酸水平上序列同一性百分比相关。基于氨基酸的序列同一性百分比，进一步把给定的血清型BoNT蛋白质划分成不同的亚型。

[0209] BoNTs在胃肠道中被吸收，并且在进入全身循环以后，与胆碱能神经末端的突出前膜结合并且阻止它们的神经递质乙酰胆碱的释放。BoNT/B、BoNT/D、BoNT/F和BoNT/G切割突触小泡蛋白/小泡缔合性膜蛋白(VAMP)；BoNT/C₁、BoNT/A和BoNT/E切割突触小体缔合性25kDa的蛋白(SNAP-25)；而BoNT/C₁切割突触融合蛋白。

[0210] 破伤风梭菌毒素是由破伤风梭菌以单一的血清型制备。丁酸梭菌制备BoNT/E，而巴氏梭菌制备BoNT/F。

[0211] 术语“梭菌毒素”也旨在包括修饰的梭菌毒素及其衍生物，包括但不限于下面描述的那些。修饰的梭菌毒素或衍生物可以包含与梭菌毒素的天然(未修饰的)形式相比已经被修饰的一个或多个氨基酸，或者可以包含不存在于梭菌毒素的天然(未修饰的)形式中的插入的一个或多个氨基酸。举例来说，与天然(未修饰的)梭菌毒素序列相比，修饰的梭菌毒素可以在一个或多个结构域中含有修饰的氨基酸序列。这样的修饰可以改变毒素的功能方面，例如生物学活性或持久性。因此，在一个实施方案中，本发明的工程化的梭菌毒素是工程化的修饰的梭菌毒素，或工程化的修饰的梭菌毒素衍生物，或工程化的梭菌毒素衍生物。

[0212] 修饰的梭菌毒素可以在重链的氨基酸序列(例如修饰的H_C结构域)中具有一个或多个修饰，其中所述修饰的重链以比天然(未修饰的)梭菌毒素更高或更低的亲和性与靶神经细胞结合。在H_C结构域中这样的修饰能够包括修饰H_C结构域的神经节苷脂结合位点中或蛋白质(SV2或突触结合蛋白)结合位点中的残基，该修饰改变了与靶神经细胞的神经节苷脂受体和/或蛋白质受体的结合。这样的修饰的梭菌毒素的实例被描述在WO 2006/027207和WO 2006/114308中，将其二者的全部内容并入本文作为参考。

[0213] 修饰的梭菌毒素可以在轻链的氨基酸序列中具有一个或多个修饰，例如在底物结合或催化结构域中的修饰，该修饰可以改变或修饰所述修饰的LC的SNARE蛋白特异性。这样

的修饰的梭菌毒素的实例被描述在WO 2010/120766和US 2011/0318385,将其二者的全部内容并入本文作为参考。

[0214] 修饰的梭菌毒素可以包含增大或减小修饰的梭菌毒素的生物学活性和/或生物学持久性的一个或多个修饰。例如,修饰的梭菌毒素可以包含基于亮氨酸或酪氨酸的基序,其中所述的基序增大或减小修饰的梭菌毒素的生物学活性和/或生物学持久性。合适的基于亮氨酸的基序包括xDxxxLL、xExxxLL、xExxxIL和xExxxLM(其中x是任何氨基酸)。合适的基于酪氨酸的基序包括Y-x-x-Hy(其中Hy是疏水的氨基酸)。包含基于亮氨酸和酪氨酸的基序的修饰的梭菌毒素的实例被描述在WO 2002/08268,将其全部内容并入本文作为参考。

[0215] 术语“梭菌毒素”旨在包括杂合体和嵌合体梭菌毒素。杂合体梭菌毒素包含来自一种梭菌毒素或其亚型的轻链的至少一部分和来自另一种梭菌毒素或梭菌毒素亚型的重链的至少一部分。在一个实施方案中,杂合体梭菌毒素可以包含来自一种梭菌毒素亚型的完整轻链和来自另一种梭菌毒素亚型的重链。在另一个实施方案中,嵌合体梭菌毒素可以包含一种梭菌毒素亚型的重链的一部分(例如结合结构域),重链的另一部分来自另一种梭菌毒素亚型。类似地或可选择地,治疗性元素可以包含来自不同的梭菌毒素的轻链部分。这样的杂合体或嵌合体梭菌毒素是有用的,例如作为递送手段,把这样的梭菌毒素的治疗益处递送到对给定的梭菌毒素亚型有免疫抗性的患者,递送到可以具有低于平均的给定的梭菌毒素重链结合结构域的受体浓度的患者,或递送到可以具有膜或小泡毒素底物的抗蛋白酶的变体(例如SNAP-25、VAMP和突触融合蛋白)的患者。杂合体和嵌合体梭菌毒素被描述在US 8,071,110中,将该公开的全部内容并入本文作为参考。因此,在一个实施方案中,本发明的工程化的梭菌毒素是工程化的杂合体梭菌毒素或工程化的嵌合体梭菌毒素。

[0216] 术语“梭菌毒素”旨在包括再靶向的梭菌毒素。在再靶向的梭菌毒素中,梭菌毒素被修饰为包含称为靶向部分(TM)的外源配体。选择TM以提供对所需的靶细胞的结合特异性,并且作为再靶向过程的一部分,梭菌毒素的天然结合部分(例如H_C结构域或H_{CC}结构域)可以被除去。再靶向技术例如被描述在:EP-B-0689459;WO 1994/021300;EP-B-0939818;US 6,461,617;US 7,192,596;WO 1998/007864;EP-B-0826051;US 5,989,545;US 6,395,513;US 6,962,703;WO 1996/033273;EP-B-0996468;US 7,052,702;WO 1999/017806;EP-B-1107794;US 6,632,440;WO 2000/010598;WO 2001/21213;WO 2006/059093;WO 2000/62814;WO 2000/04926;WO 1993/15766;WO 2000/61192;和WO 1999/58571中,将其全部内容并入本文作为参考。因此,在一个实施方案中,本发明的工程化的梭菌毒素是工程化的再靶向的梭菌毒素。

[0217] 本发明还包括具有非天然蛋白酶切割位点的梭菌毒素。在这样的梭菌毒素中,天然的蛋白酶切割位点(也被称为活化位点,如上所述)被修饰或被不是该梭菌毒素天然有的蛋白酶切割位点(即外源性切割位点)替换。这样的位点将会需要外源性蛋白酶进行切割,这允许改善的对切割事件的时机和位置的控制。可以被用在梭菌毒素中的非天然蛋白酶切割位点包括:

	肠激酶	(DDDDK ↓)
	因子 Xa	(IEGR ↓ /IDGR ↓)
[0218]	TEV(烟草蚀纹病毒)	(ENLYFQ ↓ G)
	凝血酶	(LVPR ↓ GS)
	PreScission	(LEVLFQ ↓ GP).

[0219] 另外的蛋白酶切割位点包括被无细胞毒性的蛋白酶(例如被梭菌神经毒素的轻链)切割的识别序列。这些包括被无细胞毒性的蛋白酶(例如梭菌神经毒素的轻链)切割的SNARE(例如SNAP-25、突触融合蛋白、VAMP)蛋白质识别序列。包含非天然蛋白酶切割位点的梭菌毒素被描述在US 7,132,259、EP 1206554-B2和US 2007/0166332中,将其全部内容并入本文作为参考。术语蛋白酶切割位点也包括内含肽,它是自切割序列。自剪接反应是能够控制的,例如通过改变存在的还原剂的浓度。

[0220] 本发明还包括包含“破坏性切割位点”的梭菌毒素。在所述的梭菌毒素中,在选定的位置,非天然的蛋白酶切割位点被掺入梭菌毒素中,使得在所述位点的切割将会减小梭菌毒素的活性或失活。在施用后,梭菌毒素迁移到非靶向的位置的事件中,破坏性的蛋白酶切割位点可能易受局部蛋白酶切割。合适的非天然蛋白酶切割位点包括上述的那些。包含破坏性切割位点的梭菌毒素被描述在WO 2010/094905和WO 2002/044199中,将其二者的全部内容并入本文作为参考。

[0221] 本发明的工程化的梭菌毒素,特别是其轻链组件,可以是聚乙二醇化的一这可以帮助增大稳定性,例如轻链组件的作用持续时间。当轻链包含BoNT/A、B或C₁蛋白酶时,聚乙二醇化是特别优选的。聚乙二醇化优选包括把PEG添加到轻链组件的N-末端。举例来说,轻链的N-末端可以是被一个或多个氨基酸(例如半胱氨酸)残基延长,该残基可以是相同的或不同的。一个或多个所述的氨基酸残基可以具有它自己的与其相连的(例如共价连接的)PEG分子。这种技术的实例被描述在WO2007/104567中,将其全部内容并入本文作为参考。

[0222] 本发明的工程化的梭菌毒素可以没有复合蛋白,该复合蛋白存在于天然存在的梭菌毒素复合物中。

[0223] 本发明的工程化的梭菌毒素还可以包含数量有限的非标准氨基酸。因此,除了20种标准氨基酸以外,还可以用非标准氨基酸(例如4-羟基脯氨酸、6-N-甲基赖氨酸、2-氨基异丁酸、异缬氨酸和 α -甲基丝氨酸)来取代本发明的工程化的梭菌毒素的氨基酸残基。数量有限的非保守性的氨基酸、没有被遗传密码编码的氨基酸和非天然的氨基酸可以被用来取代梭菌多肽的氨基酸残基。本发明的工程化的梭菌毒素还可以包含非天然存在的氨基酸残基。

[0224] 非天然存在的氨基酸包括但不限于反式-3-甲基脯氨酸、2,4-亚甲基-脯氨酸、顺式-4-羟基脯氨酸、反式-4-羟基-脯氨酸、N-甲基甘氨酸、别苏氨酸、甲基-苏氨酸、羟基-乙基半胱氨酸、羟基乙基高半胱氨酸、硝基谷氨酰胺、高谷氨酰胺、派可酸、叔亮氨酸、正缬氨酸、2-氮杂苯丙氨酸、3-氮杂苯基-丙氨酸、4-氮杂苯基-丙氨酸和4-氟苯丙氨酸。本领域中已知有几种方法用于把非天然存在的氨基酸残基掺入蛋白质中。例如可以采用体外系统,其中使用以化学方法氨酰化的抑制子tRNAs抑制无义突变。用于合成氨基酸和氨酰化tRNA

的方法是本领域中已知的。包含无义突变的质粒的转录和翻译是在包含大肠杆菌S30提取物和商购可获得的酶和其它试剂的不含有细胞的体系中的。通过层析法纯化蛋白质。参见,例如Robertson等人,J. Am. Chem. Soc. 113:2722, 1991; Ellman等人,Methods Enzymol. 202:301, 1991; Chung等人,Science 259:806-9, 1993; 和Chung等人,Proc. Natl. Acad. Sci. USA 90:10145-9, 1993)。在第二种方法中,翻译是通过显微注射突变的mRNA和以化学方法氨酰化的抑制子tRNAs在非洲爪蟾卵母细胞中进行的(Turcatti等人,J. Biol. Chem. 271:19991-8, 1996)。在第三种方法中,在不存在将要被替换的天然氨基酸(例如苯丙氨酸)并且存在所需的非天然存在的氨基酸(例如2-氮杂苯丙氨酸、3-氮杂苯丙氨酸、4-氮杂苯丙氨酸或4-氟苯丙氨酸)的情况下培养大肠杆菌细胞。将非天然存在的氨基酸掺入多肽中,替换它的天然对应物。参见Koide等人,Biochem. 33:7470-6, 1994。

[0225] 使用重组核酸技术能够制备本发明的工程化的梭菌毒素。因此,在一个实施方案中,工程化的梭菌毒素(如上所述)是重组的工程化的梭菌毒素。

[0226] 在另一个方面中,本发明提供了核酸(例如DNA),其包含编码如上所述的工程化的梭菌毒素的核酸序列。在一个实施方案中,核酸序列是作为包含启动子和终止子的DNA载体的一部分而被制备的。

[0227] 在优选的实施方案中,载体具有选自下列的启动子:

	<u>启动子</u>	<u>诱导剂</u>	<u>典型的诱导条件</u>
[0228]	Tac (杂合体)	IPTG	0.2 mM (0.05-2.0 mM)
	AraBAD	L-阿拉伯糖	0.2%(0.002-0.4%)
	T7-lac 操纵基因	IPTG	0.2 mM (0.05-2.0 mM)

[0229] 在另一个优选的实施方案中,载体具有选自下列的启动子:

	<u>启动子</u>	<u>诱导剂</u>	<u>典型的诱导条件</u>
[0230]	Tac (杂合体)	IPTG	0.2 mM (0.05-2.0 mM)
	AraBAD	L-阿拉伯糖	0.2%(0.002-0.4%)
	T7-lac 操纵基因	IPTG	0.2 mM (0.05-2.0 mM)
	T5-lac 操纵基因	IPTG	0.2 mM (0.05-2.0 mM)

[0231] 可以使用本领域中已知的任何合适的方法来制备本发明的核酸分子。因此,可以使用化学合成技术来制备核酸分子。可选择的是,可以使用分子生物学技术来制备本发明的核酸分子。

[0232] 本发明的DNA构建体优选是在电脑中(in silico)被设计,然后通过常规的DNA合成技术被合成的。

[0233] 根据将要被采用的最终的宿主细胞(例如大肠杆菌)表达系统,任选地针对密码子偏倚修饰上述的核酸序列信息。

[0234] 在一个实施方案中,编码如上所述的工程化的梭菌毒素的核酸序列是核酸序列,它与选自SEQ ID NOs:2、4和6的核酸序列具有至少70%(例如至少75、80、85、90、95、97、98或99%)的序列同一性。在一个实施方案中,核酸序列与选自SEQ ID NO:2、4和6的核酸序列

具有至少90%的序列同一性。

[0235] 本发明还提供了由如上所述的核酸序列编码的多肽。因此，在一个方面中，本发明提供了包含氨基酸序列的多肽，该氨基酸序列与选自SEQ ID NOs:1、3和5的氨基酸序列具有至少70%（例如至少75、80、85、90、95、97、98或99%）的序列同一性。在一个实施方案中，氨基酸序列与选自SEQ ID NOs:1、3和5的氨基酸序列具有至少90%的序列同一性。

[0236] 在一个实施方案中，本发明的工程化的梭菌毒素是如上所述的工程化的BoNT/A，并且所述的工程化的BoNT/A包含氨基酸序列（或由其组成），该氨基酸序列与选自SEQ ID NOs:1、3和5的氨基酸序列具有至少70%（例如至少75、80、85、90、95、97、98、99、99.5或99.9%）的序列同一性。

[0237] 在一个实施方案中，本发明的工程化的梭菌毒素是如上所述的工程化的BoNT/A，并且所述的工程化的BoNT/A包含SEQ ID NO:1、3或5的氨基酸序列（或由其组成）。

[0238] 在一个方面中，本发明提供了多肽，它包含SEQ ID NO:1、3或5的氨基酸序列（或由其组成）。

[0239] 在一个方面中，本发明提供了编码如上所述的工程化的梭菌毒素的核酸，其中所述的核酸包含核酸序列，该核酸序列与选自SEQ ID NOs:2、4和6的核酸序列具有至少70%（例如至少75、80、85、90、95、97、98、99、99.5或99.9%）的序列同一性。在一个实施方案中，核酸包含SEQ ID NO:2、4或6的核酸序列（或由其组成）。

[0240] 在一个方面中，本发明提供了核酸，它包含SEQ ID NO:2、4或6的核酸序列（或由其组成）。

[0241] 在一个实施方案中，本发明的工程化的梭菌毒素是如上所述的工程化的BoNT/E，并且所述的工程化的BoNT/E包含氨基酸序列，该氨基酸序列与SEQ ID NO:7具有至少70%（例如至少75、80、85、90、95、97、98、99、99.5或99.9%）的序列同一性。

[0242] 在一个方面，本发明提供了编码如上所述的工程化的梭菌毒素的核酸，其中所述的核酸包含核酸序列，该核酸序列与SEQ ID NO:8具有至少70%（例如至少75、80、85、90、95、97、98、99、99.5或99.9%）的序列同一性。

[0243] 在两个或更多个核酸或氨基酸序列之间的“序列同一性百分数”是序列共有的相同位置的数量的函数。因此，可以这样计算同一性%：相同的核苷酸/氨基酸的数量除以核苷酸/氨基酸的总数，乘以100。计算序列同一性%也可以考虑需要引入以优化两个或更多个序列的比对的空位的数量以及每个空位的长度。在两个或更多个序列之间的序列比较和同一性百分比的确定可以使用特定的数学算法（例如BLAST）来进行，该算法对于技术人员来说是熟悉的。

[0244] 在一个方面中，本发明提供了制备单链工程化的梭菌毒素蛋白的方法，该蛋白具有轻链和重链，所述的方法包括在合适的宿主细胞中表达核酸（所述的核酸如上所述），裂解宿主细胞以提供包含单链工程化的梭菌毒素蛋白的宿主细胞匀浆液，并且分离单链工程化的梭菌毒素蛋白。

[0245] 在另一个方面中，本发明提供了活化工程化的梭菌毒素的方法，所述的方法包括提供能够通过如上所述的制备单链工程化的梭菌毒素蛋白的方法得到的单链工程化的梭菌毒素蛋白，使多肽与切割多肽的蛋白酶在位于轻链和重链之间的识别位点（切割位点）处接触，从而把多肽转化为双链多肽，其中通过二硫键把轻链和重链连接在一起。

[0246] 本发明的工程化的梭菌毒素可以被用来预防或治疗某些医学或美容化妆性疾病和病症。因此,在另一个方面中,本发明提供了如上所述的工程化的梭菌毒素,其用于医药中。

[0247] 在一个相关的方面中,本发明提供了如上所述的工程化的梭菌毒素,其用于预防或治疗选自下列的疾病或病症:斜视、眼睑痉挛、斜眼、张力失常(例如痉挛性张力失常、口颌肌张力失常、病灶性张力失常、迟发性张力失常、喉部张力失常、四肢张力失常、子宫颈张力失常)、斜颈(例如痉挛性斜颈)、受益于细胞/肌肉失能的美容治疗(化妆美容性)应用(经由SNARE下调或失活)、神经肌肉障碍或眼能动性病症(例如共同性斜视、垂直斜视、外直肌瘫痪、眼球震颤、甲状腺机能障碍性肌病)、指痉挛、眼睑痉挛、夜磨牙症、肝豆状核变性、震颤、抽搐、节段性肌阵挛、痉挛、由慢性多发性硬化症引起的痉挛状态、导致异常的膀胱控制的痉挛状态、男性意像、背部痉挛、四头肌僵痛、紧张性头痛、提肌骨盆综合征、脊柱裂、迟发性运动障碍、帕金森氏病、口吃、半面痉挛、眼睑障碍、脑性瘫痪、病灶性痉挛状态、痉挛性结肠炎、神经源性膀胱、肛门痉挛、四肢痉挛状态、抽搐、震颤、夜磨牙症、肛裂、弛缓不能、吞咽困难、流泪、多汗症、过度流涎、过度胃肠分泌、肌肉疼痛(例如来自肌肉痉挛的疼痛)、头痛(例如紧张性头痛)、额皱纹、皮肤皱纹、癌症、子宫障碍、泌尿生殖器障碍、泌尿生殖器-神经障碍、慢性神经原性炎症和平滑肌障碍。

[0248] 在使用中,本发明使用药物组合物,其包含工程化的梭菌毒素和至少一种选自下列的组分:药学上可接受的载体、赋形剂、佐剂、抛射剂和/或盐。

[0249] 可以配制本发明的工程化的梭菌毒素用于口服、非肠道、连续输注、吸入或局部应用。适用于注射的组合物可以是溶液剂、混悬剂或乳剂的形式或者是干燥粉末形式,在使用前把干燥粉末溶解或混悬在合适的载体中。

[0250] 在将要被局部递送的工程化的梭菌毒素的情形中,工程化的梭菌毒素可以被配制成乳膏剂(例如用于局部应用),或用于皮下注射。

[0251] 局部递送手段可以包括气雾剂或其它喷雾剂(例如喷雾器)。在这个方面,工程化的梭菌毒素的气雾剂制剂能够递送到肺和/或其它鼻的和/或支气管的或气道的通道。

[0252] 可以在参与受影响的器官的神经支配的脊髓节段的水平,通过鞘内或硬膜外注射到脊柱中,来给患者施用本发明的工程化的梭菌毒素。

[0253] 优选的施用途径是经由腹腔镜检查和/或局部化的,特别肌内注射。

[0254] 施用本发明的工程化的梭菌毒素的剂量范围是产生所需的治疗效果的那些剂量范围。应当理解所需的剂量范围取决于工程化的梭菌毒素或组合物的精确性质、施途径、制剂的性质、患者的年龄、患者病症的性质、程度或严重程度、禁忌症(如果有的话)和主治医师的判断。能够使用用于最优化的标准经验途径来调节这些剂量水平的变化。

[0255] 合适的每日剂量(每kg患者的体重)在0.0001-1ng/kg,优选在0.0001-0.5ng/kg,更优选在0.002-0.5ng/kg,并且特别优选在0.004-0.5ng/kg范围内。单位剂量能够在少于1皮克至30ng变化,但是典型地将会在0.01至1ng/剂量的区域内,可以每日施用或优选不那么频繁,例如每星期或每6个月。

[0256] 特别优选的给药方案基于0.05ng工程化的梭菌毒素作为1×剂量。在这个方面,优选的剂量在1×-100×(即0.05-5ng)范围内。

[0257] 流体剂型典型地是利用工程化的梭菌毒素和不含热原的无菌载体制备的。取决于

所用的载体和浓度,能够把工程化的梭菌毒素溶解或混悬在载体中。在制备溶液时,可以将工程化的梭菌毒素溶解于载体中,如果必要的话,通过添加氯化钠把溶液制成等渗的,并且通过使用无菌技术通过无菌过滤器过滤来灭菌,然后填充到合适的无菌小瓶或安瓿中并且密封。可选择的是,如果溶液稳定性是足够的,则可以通过高压灭菌法把处于密封容器中的溶液灭菌。有利地,可以把添加剂例如缓冲剂、增溶剂、稳定剂、防腐剂或杀菌剂、助悬剂或乳化剂和或局部麻醉剂溶解在载体中。

[0258] 干燥粉末(在使用前将其溶解或混悬在合适的载体中)可以通过在无菌区使用无菌技术把预先灭菌的成分填充在无菌容器中来制备。可选择的是,可以在无菌区使用无菌技术把成分溶解在合适的容器中。然后把产品冷冻干燥,并且把容器无菌密封。

[0259] 除了把无菌组分混悬而不是溶解在无菌载体中并且灭菌不能够通过过滤完成外,以基本上相同的方式制备适用于肌肉、皮下或皮内注射的非肠道混悬剂。可以以无菌状态分离组分,或可选择的是,可以在分离后把组分灭菌,例如通过伽马照射。

[0260] 有利地,把助悬剂例如聚乙烯吡咯烷酮包含在组合物中以促进组分的均匀分布。

[0261] 根据本发明的施用可以利用多种递送技术,包括微粒包囊、病毒递送体系或高压气溶胶冲击。

附图说明

[0262] 图1

[0263] CatH_N_v1 (图1A)、CatH_N_v2 (图1B) 和CatH_N_v3 (图1C) 纯化的SDS-PAGE纯化。

[0264] 图2

[0265] 对于CatH_N_v1的大鼠胚胎脊髓神经元(eSCN)中SNAP-25切割的百分比。将大鼠胚胎脊髓神经元培养3周,并且用CatH_N_v1处理24小时,然后用SNAP-25特异性抗体进行Western印迹。数据是来自三次独立试验(一式三份)的平均值±SEM。

[0266] 图3

[0267] 在小鼠膈神经偏侧膈测定法(mPNHD)中BoNT/A和CatH_N_v1的功效(t_{50})。数据点是单个偏侧膈制备物和平均值±SEM。CatH_N_v1在统计学上显著慢于参考蛋白BoNT/A(List Biological Laboratories)。单因素ANOVA和Dunnett多重比较检验。 $**p<0.01$, $***p<0.001$, $****p<0.0001$ (单因素ANOVA和Dunnett多重比较检验)。

[0268] 图4

[0269] 对于CatH_N_v2的大鼠胚胎脊髓神经元(eSCN)中SNAP-25切割的百分比。将大鼠胚胎脊髓神经元培养3周,并且用CatH_N_v2处理24小时,然后用SNAP-25特异性抗体进行Western印迹。数据是来自三次独立试验(一式三份)的平均值±SEM。

[0270] 图5

[0271] 在小鼠膈神经偏侧膈测定法(mPNHD)中BoNT/A和CatH_N_v2的功效(t_{50})。数据点是单个偏侧膈制备物和平均值±SEM。CatH_N_v2在统计学上等同于参考蛋白BoNT/A(List Biological Laboratories)。单因素ANOVA和Dunnett多重比较检验。 $**p<0.01$, $***p<0.001$, $****p<0.0001$ (单因素ANOVA和Dunnett多重比较检验)。

[0272] 图6

[0273] 对于CatH_N_v3的大鼠胚胎脊髓神经元(eSCN)中SNAP-25切割的百分比。将大鼠胚

胎脊髓神经元培养3周,并且用CatH_N_v3处理24小时,然后用SNAP-25特异性抗体进行Western印迹。数据是来自三次独立试验(一式三份)的平均值±SEM。

[0274] 图7

[0275] 在小鼠膈神经偏侧膈测定法(mPNHD)中BoNT/A和CatH_N_v3的功效(t_{50})。数据点是单个偏侧膈制备物和平均值±SEM。CatH_N_v3在统计学上显著慢于参考蛋白BoNT/A(List Biological Laboratories)。单因素ANOVA和Dunnett多重比较检验。 $**p<0.01$, $***p<0.001$, $****p<0.0001$ (单因素ANOVA和Dunnett多重比较检验)。

[0276] 图8

[0277] 等电点聚焦分析。与未修饰的BoNT/A相比,所有三个CatH_N构建体具有增加观察到的pI。

[0278] 图9

[0279] CatLC构建体的SDS-PAGE纯化。

[0280] 图10

[0281] 与BoNT/E LC参考相比,CatLC的催化活性,其中按照制造商的说明书,使用BoTest A/E BoNT检测试剂盒(BioSentinal Cat#A1004)中获得的 pEC_{50} 值。数据显示一次独立试验(一式三份)的平均值±标准差。

[0282] 序列

[0283] SEQ ID NO:1.工程化的BoNT/A,“CatH_N_v1”,氨基酸序列。

[0284] SEQ ID NO:2.工程化的BoNT/A,“CatH_N_v1”,核酸序列。

[0285] SEQ ID NO:3.工程化的BoNT/A,“CatH_N_v2”,氨基酸序列。

[0286] SEQ ID NO:4.工程化的BoNT/A,“CatH_N_v2”,核酸序列。

[0287] SEQ ID NO:5.工程化的BoNT/A,“CatH_N_v3”,氨基酸序列。

[0288] SEQ ID NO:6.工程化的BoNT/A,“CatH_N_v3”,核酸序列。

[0289] SEQ ID NO:7.工程化的BoNT/E轻链,“CatLC”,氨基酸序列。

[0290] SEQ ID NO:8.工程化的BoNT/E轻链,“CatLC”,核酸序列。

[0291] SEQ ID NO:1.工程化的BoNT/A,“CatH_N_v1”,氨基酸序列。

[0292] MPFVNKQFNKDPVNGVDIAYIKIPNAGQMOPVKAFKIHNKIWIWIPERDTFTNPEEGDLNPPPEAKQVP
 VSYYDSTYLSTDNEKDNLYLKGVTCLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCINVIQPDGSYRS
 EELNLVIIIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPLLGAGKFATDPAVTLAH
 ELIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQENEFRLYYYNKFKDIASLTKA
 KSIIVGTTASLQYMKNVFKKYLKLLSEDTSGKFSVDKLFKDKLYKMLTEIYTEDNFVKKFVLRKTYLNFDKAVFKIN
 IVPKVNYTIYDGFNLRNTNLAANFNGQNTNINNMFTKLKNTGLFEFYKLLCVRGIITSKTKSLDKGYNKALNDLC
 IKVNNWDLFFSPSEDNFTNDLNKGEEITSNTNIEAAEENISLDLIQYYLTFNFDNEPENISLENLSSDIIGQLELM
 PNIERFPNGKKYELDKYTMFHYLRAQEFEHGKRRIALTNSVNEALLNPSRVYTFSSDYVKKVKNKATEAMFLGWVE
 QLVYDFTDETSEVSTTDKIADITIIIPYIGPALNIGNMRYKRRFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYI
 ANKVLTVQTIIDNALSKRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAEATKAIINYQYNQYTEEEKNNI
 NFNIDDLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLDFDASLKDALLKYIYDNRGTLIGQVDRLKD
 KVNNTLSRDRPFQLSKYVDNQRLSTFTTEYIKNIINTSILNRYESNHLIDLSRYASKINIGSKVNFDPIDKNQIQ
 FNLESSKIEVILKNAIVYNSMYENFSTSFWIRIPKYFNISLNNYEYTIINCMENNSGWKVSLENYGEEIWTLQDTQEI

KQRVVKYSQMINISDYINRWIFVTITNNRLNNSKIYINGRLIDQKPI SNLGNIHASNNIMFKLDGCRDTHRYIWIWIK
YFNLFDKELNEKEIKDLYDNQNSNGILKDFWGDYLDQDKPYMLNLYDPNKYVDVNNVVGIRGYMYLKGPRGSVMTTN
IYLNSSLYRGTKFI IKKYASGNKDNIVRNNDRVYINVVVKNKEYRLATNASQAGVEKILSALEIPDVGNSLQVVMK
SKNDQGITNKCKMNLQDNNGNDIGFIGFHQFNNAKLVASNWNQRQIERSSRTLGCSEWEPVDDGWGERPL

[0293] SEQ ID NO:2. 工程化的BoNT/A, “CatH_N_v1”, 核酸序列。

[0294] ATGCCATTCGTCAACAAGCAATTCAACTACAAAGACCCAGTCAACGGCGTCGACATCGCATACATCAA
GATTCCGAACGCCGGTCAAATGCAGCCGGTTAAGGCTTTTAAAGATCCACAACAAGATTTGGGTTATCCCGGAGCGT
GACACCTTCACGAACCCGGAAGAAGGCGATCTGAACCCGCCACCGGAAGCGAAGCAAGTCCCTGTCAGCTACTACG
ATTCGACGTACCTGAGCACGGATAACGAAAAAGATAACTACCTGAAAGGTGTGACCAAGCTGTTCGAACGTATCTA
CAGCACGGATCTGGGTGCGATGCTGCTGACTAGCATTGTTTCGCGGTATCCCGTTCTGGGGTGGTAGCACGATTGAC
ACCGAACTGAAGGTTATCGACACTAAGTGCATTAACGTTATTCAACCGGATGGTAGCTATCGTAGCGAAGAGCTGA
ATCTGGTCATCATTGGCCCGAGCGCAGACATTATCCAATTCGAGTGCAAGAGCTTTGGTCACGAGGTTCTGAATCT
GACCCGCAATGGCTATGGTAGCACCCAGTACATTCGTTTTTCGCCGGATTTTACCTTCGGCTTTGAAGAGAGCCTG
GAGGTTGATACCAATCCGTTGCTGGGTGCGGGCAAATTCGCTACCGATCCGGCTGTCACGCTGGCCCATGAACTGA
TCCACGCAGGCCACCGCCTGTACGGCATTGCCATCAACCCAAACCGTGTGTTCAAGGTTAATACGAATGCATACTA
CGAGATGAGCGGCCTGGAAGTCAGCTTCAAGAAGTGCACACCTTCGGTGGCCATGACGCTAAATTCATTGACAGC
TTGCAAGAGAATGAGTTCCGTCTGTACTACTATAACAAATTCAAAGACATTGCAAGCACGTTGAACAAGGCCAAAA
GCATCGTTGGTACTACCGCGTCGTTGCAGTATATGAAGAATGTGTTTAAAGAGAAGTACCTGCTGTCCGAGGATAC
CTCCGGCAAGTTTAGCGTTGATAAGCTGAAGTTTGACAAACTGTACAAGATGCTGACCGAGATTTACACCGAGGAC
AACTTTGTGAAATTCCTCAAAGTGTGAATCGTAAAACCTATCTGAATTTTGACAAAGCGGTTTTCAAGATTAACA
TCGTGCCGAAGGTGAACTACACCATCTATGACGGTTTTAACCTGCGTAACACCAACCTGGCGGCGAACTTTAACGG
TCAGAATACGGAAATCAACAACATGAATTCACGAAGTTGAAGAAGTTCACGGGTCTGTTGAGTTCTATAAGCTG
CTGTGCGTGCGGGTATCATCACCAGCAAACCAAAAGCCTGGACAAAGGCTACAACAAGGCGCTGAATGACCTGT
GCATTAAGGTAAACAATTGGGATCTGTTCTTTTCGCCATCCGAAGATAATTTTACCAACGACCTGAACAAGGGTGA
AGAAATCACCAGCGATACGAATATTGAAGCAGCGGAAGAGAATATCAGCCTGGATCTGATCCAGCAGTACTATCTG
ACCTTTAACTTCGACAATGAACCGGAGAACATTAGCATTGAGAATCTGAGCAGCGACATTATCGGTCAGCTGGAAC
TGATGCCGAATATCGAACGTTTCCCGAACGGCAAAAAGTACGAGCTGGACAAGTACACTATGTTCCATTACCTGCG
TGCACAGGAGTTTGAACACGGTAAAcgtCGTATCGCGCTGACCAACAGCGTTAACGAGGCCCTGCTGAACCCGAGC
CGTGTCTATACCTTCTTCAGCAGCGACTATGTTAAGAAAGTGAACAAAGCCACTGAGGCCGCGATGTTCTGGGCT
GGGTGGAACAGCTGGTATATGACTTCACGGACGAGACGAGCGAAGTGAGCACTACCGACAAAATTGCTGATATTAC
CATCATTATCCCGTATATTGGTCCGGCACTGAACATTGGCAACATGCgtTACAAAcgtcgTTTTGTGGGTGCCCTG
ATCTTCTCCGGTGCCGTGATTCTGCTGGAGTTCATTCCGGAGATTGCGATCCCGGTGTTGGGTACCTTCGCGCTGG
TGTCCTACATCGGAATAAGGTTCTGACGGTTCAGACCATCGATAACGCGCTGTCGAAACGTAATGAAAAATGGGA
CGAGGTTTACAAATACATTGTTACGAATTGGCTGGCGAAAGTCAATACCCAGATCGACCTGATCCGTAAGAAAATG
AAAGAGGCGCTGGAGAATCAGGCGGAGGCCACCAAGCAATTATCAACTACCAATACAACCAGTACACGGAAGAAG
AGAAGAATAACATTAACCTCAATATCGATGATTTGAGCAGCAAGCTGAATGAATCTATCAACAAAGCGATGATCAA
TATCAACAAGTTTTTGAATCAGTGTAGCGTTTCGTACCTGATGAATAGCATGATTCCGTATGGCGTCAAACGTCTG
GAGGACTTCGACGCCAGCCTGAAAGATGCGTTGCTGAAATACATTTACGACAATCGTGGTACGCTGATTGGCCAAG
TTGACCGCTTGAAAGACAAAGTTAACAATACCCTGAGCcgTgACcgTCCATTTCAACTGAGCAAGTATGTTGATAA

TCAACGTCTGTTGAGCACTTTCACCGAGTATATCAAAAACATCATCAACTAGCATTCTGAACCTGCGTTACGAGA
 GCAATCATCTGATTGATCTGAGCCGTTATGCAAGCAAGATCAACATCGGTAGCAAGGTCAATTTTGACCCGATCGAT
 AAGAACCAGATCCAGCTGTTTAATCTGGAATCGAGCAAAATTGAGGTTATCCTGAAAAACGCCATTGTCTACAACCTC
 CATGTACGAGAATTTCTCCACCAGCTTCTGGATTTCGCATCCCGAAATACTTCAACAGCATTAGCCTGAACAACGAGT
 ATACTATCATCAACTGTATGGAGAACAACAGCGGTTGGAAGGTGTCTCTGAACTATGGTGAGATCATTTGGACCTTG
 CAGGACACCCAAGAGATCAAGCAGCGCGTGTGTTCAAGTACTCTCAAATGATCAACATTTCCGATTACATTAATCG
 TTGGATCTTCGTGACCATTACGAATAACCGTCTGAATAACAGCAAGATTTACATCAATGGTCGCTTGATCGATCAGA
 AACCGATTAGCAACCTGGGTAATATCCACGCAAGCAACAACATTATGTTCAAATTGGACGTTGCCGCGATACCCAT
 CGTTATATCTGGATCAAGTATTTCAACCTGTTTGATAAAGAACTGAATGAGAAGGAGATCAAAGATTTGTATGACAA
 CCAATCTAACAGCGGCATTTTGAAGGACTTCTGGGGCGATTATCTGCAATACGATAAGCCGTACTATATGCTGAACC
 TGTATGATCCGAACAAATATGTGGATGTCAATAATGTGGGTATTCGTGGTTACATGTATTTGAAGGGTCCGCGTGGC
 AGCGTTATGACGACCAACATTTACCTGAACTCTAGCCTGTACCGTGGTACGAAATTCATCATTAAAGAAATATGCCAG
 CGGCAACAAAGATAACATTGTGCGTAATAACGATCGTGTCTACATCAACGTGGTCGTGAAGAATAAAGAGTACCGTC
 TGGCGACCAACGCTTCGCAGGCGGGTGTGAGAAAAATTCTGAGCGCGTTGGAGATCCCTGATGTCGGTAATCTGAGC
 CAAGTCGTGGTTATGAAGAGCAAGAACGACCAGGGTATCACTAACAAGTGAAGATGAACCTGCAAGACAACAATGG
 TAACGACATCGGCTTTATTGGTTTCCACCAGTTCAACAATATTGCTAAACTGGTAGCGAGCAATTGGTACAATCGTC
 AGATTGAGCGCAGCAGCCGTACTTTGGGCTGTAGCTGGGAGTTATCCCGTTCGATGATGGTTGGGGCGAACGTCCG
 CTG

[0295] SEQ ID NO:3.工程化的BoNT/A,“CatH_N-v2”,氨基酸序列。

[0296] MPFVNKQFNYPKDPVNGVDIAYIKIPNAGQMQPVKAFKIHNKIWIWIPERDTFTNPEEGDLNPPPEAKQVP
 VSYDSTYLSTDNEKDNLYLKGVTKLFERIYSTDLGRMLLTSIVRGIPIFWGGSTIDTELKVIDTNCINVIQPDGSYRS
 EELNLVIGPSADIIQFECKSFGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESLEVDTNPLLGAAGKFDPAVTLAH
 ELIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQENEFRLYYNKFKDIASLTKA
 KSIVGTTASLQYMKNVFKEKYLLEDTSKGFVSKLFDKLYKMLTEIYTEDNFVKFFKVLNRKTYLNFDAVFKIN
 IVPKVNYYTIDGFNLRNTNLAANFNGQNTENNMNFTKLKNFTGLFEFYKLLCVRGIITSKTKSLDKGYNKALNDLC
 IKVNNWDLFFSPSEDNFTNDLKKGEEITSDTNIEAAEENISLDLIQQYYLTFNFDNEPENISIENLSSDIIGQLELM
 PNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNEALLNPSRVYTFSSDYVKKVNKATKAAMFLGWVE
 QLVYDFTDETSEVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPVLGTFALVSYI
 AKKVLTVQTIIDNLSKRNEKWDEVYKYIVTNWLAKVNTQIDLIRKKMKEALENQAATKAIINYQYNQYTEEEKNKI
 KFNIDDLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLDFDASLKDALLKYIYDNRGTLKGQVDRLKD
 KVNNTLSTDIPFQLSKYVDNQRLSTFTEYIKNIINTSILNLYESNHLIDLSRYASKINIGSKVNFDPIDKNQIQL
 FNLESSKIEVILKNAIVYNSMYENFSTSFWRIPKYFNSISLNNEYTIINCMENNSGWKVSLSNYGEIITWLQDTQEI
 KQRVVFYKYSQMINISDYINRWIFVTITNRLNNSKIYINGRLIDQKPI SNLGNIHASNNIMFKLDGCRDTHRYIWI
 YFNLFDKELNEKEIKDLYDNQNSNGILKDFWGDYLYQDKPYMLNLYDPNKYVDVNNVIRGYMYLKGPRGSVMTTN
 IYLNSSLYRGTKFIKIKYASGNKDNIVRNDRVYINVVKNKEYRLATNASQAGVEKILSALEIPDVGNSLQVVMK
 SKNDQGITNKCKMNLQDNNNGNDIGFIFGHQFNNAIKLVASNWYNRQIERSSRTLGCSEWEPVDDGWGERPL

[0297] SEQ ID NO:4.工程化的BoNT/A,“CatH_N-v2”,核酸序列。

[0298] ATGCCATTTCGTCACAAGCAATTCAACTACAAAGACCCAGTCAACGGCGTCGACATCGCATAACATCAA
 GATTCCGAACGCCGGTCAAATGCAGCCGGTTAAGGCTTTTAAGATCCACAACAAGATTTGGGTTATCCCGGAGCGT

GACACCTTCACGAACCCGGAAGAAGGCGATCTGAACCCGCCACCGGAAGCGAAGCAAGTCCCTGTCAGCTACTACG
ATTCGACGTACCTGAGCACGGATAACGAAAAAGATAACTACCTGAAAGGTGTGACCAAGCTGTTTGAACGTATCTA
CAGCACGGATCTGGGTCGCATGCTGCTGACTAGCATTGTTTCGCGGTATCCCGTTCTGGGGTGGTAGCACGATTGAC
ACCGAACTGAAGGTTATCGACACTAACTGCATTAACGTTATTCAACCGGATGGTAGCTATCGTAGCGAAGAGCTGA
ATCTGGTCATCATTGGCCCGAGCGCAGACATTATCCAATTCGAGTGCAAGAGCTTTGGTCACGAGGTTCTGAATCT
GACCCGCAATGGCTATGGTAGCACCCAGTACATTCGTTTTTCGCCGGATTTTACCTTCGGCTTTGAAGAGAGCCTG
GAGGTTGATACCAATCCGTTGCTGGGTGCGGGCAAATTCGCTACCGATCCGGCTGTCACGCTGGCCCATGAACTGA
TCCACGCAGGCCACCGCCTGTACGGCATTGCCATCAACCCAAACCGTGTGTTCAAGGTTAATACGAATGCATACTA
CGAGATGAGCGGCCTGGAAGTCAGCTTCAAGAAGTGCACACCTTCGGTGGCCATGACGCTAAATTCATTGACAGC
TTGCAAGAGAATGAGTTCCGTTCTGTACTACTATAACAAATTCAAAGACATTGCAAGCACGTTGAACAAGGCCAAAA
GCATCGTTGGTACTACCGCGTCTGTCAGTATATGAAGAATGTGTTTAAAGAGAAGTACCTGCTGTCCGAGGATAC
CTCCGGCAAGTTTAGCGTTGATAAGCTGAAGTTTGACAACTGTACAAGATGCTGACCGAGATTTACACCGAGGAC
AACTTTGTGAAATTTCTCAAAGTGTGAATCGTAAAACCTATCTGAATTTTGACAAAGCGGTTTTCAAGATTAACA
TCGTGCCGAAGGTGAACTACACCATCTATGACGGTTTTAACCTGCGTAACACCAACCTGGCGGCGAACTTTAACGG
TCAGAATACGGAAATCAACAACATGAATTTACGAAGTTGAAGAAGTTCACGGGTCTGTTTCGAGTTCTATAAGCTG
CTGTGCGTGCGCGGTATCATCACCAGCAAAACCAAAAGCCTGGACAAAGGCTACAACAAGGCGCTGAATGACCTGT
GCATTAAGGTAAACAATGGGATCTGTTCTTTTCGCCATCCGAAGATAATTTTACCAACGACCTGAAgAAGGTGA
AGAAATCACCGAGGATACGAATATTGAAGCAGCGGAAGAGAATATCAGCCTGGATCTGATCCAGCAGTACTATCTG
ACCTTTAACTTCGACAATGAACCGGAGAACATTAGCATTGAGAATCTGAGCAGCGACATTATCGGTCAGCTGGAAC
TGATGCCGAATATCGAACGTTTCCCGAACGGCAAAAAGTACGAGCTGGACAAGTACACTATGTTCCATTACCTGCG
TGCACAGGAGTTTGAACACGGTAAAAGCCGTATCGCGCTGACCAACAGCGTTAACGAGGCCCTGCTGAACCCGAGC
CGTGTCTATACCTTCTTCAGCAGCGACTATGTTAAGAAAGTGAACAAAGCCACTaAGGCCGCGATGTTCTGGGCT
GGGTGGAACAGCTGGTATATGACTTCACGGACGAGACGAGCGAAGTGAAGCACTACCGACAAAATTGCTGATATTAC
CATCATTATCCCGTATATTGGTCCGGCACTGAACATTGGCAACATGCTGTACAAAGACGATTTTGTGGGTGCCCTG
ATCTTCTCCGGTGCCGTGATTCTGTGCTGGAGTTCATTCCGGAGATTGCGATCCCGGTGTTGGGTACCTTCGCGCTGG
TGTCCTACATCGCGAAgAAGGTTCTGACGGTTCAGACCATCGATAACGCGCTGTCGAAACGTAATGAAAAATGGGA
CGAGGTTTACAAATACATTGTTACGAATTGGCTGGCGAAAGTCAATACCCAGATCGACCTGATCCGTAAGAAAATG
AAAGAGGCGCTGGAGAATCAGGCGGAGGCCACCAAGCAATTATCAACTACCAATACAACCAGTACACGGAAGAAG
AGAAGAATAAgATTAAgTTCAATATCGATGATTTGAGCAGCAAGCTGAATGAATCTATCAACAAAGCGATGATCAA
TATCAACAAGTTTTTGAATCAGTGTAGCGTTTTCGTACCTGATGAATAGCATGATTCGGTATGGCGTCAAACGTCTG
GAGGACTTCGACGCCAGCCTGAAAGATGCGTTGCTGAAATACATTTACGACAATCGTGGTACGCTGAagGGCCAAG
TTGACCGCTTGAAAGACAAAGTTAACAATACCTGAGCACCGACATCCCATTTCAACTGAGCAAGTATGTTGATAAT
CAACGTCTGTTGAGCACTTTCACCGAGTATATCAAAAACATCATCAATACTAGCATTCTGAACCTGCGTTACGAGAG
CAATCATCTGATTGATCTGAGCCGTTATGCAAGCAAGATCAACATCGGTAGCAAGGTCAATTTTGACCCGATCGATA
AGAACCAGATCCAGCTGTTAATCTGGAATCGAGCAAAAATTGAGGTTATCCTGAAAAACGCCATTGTCTACAACCTCC
ATGTACGAGAATTTCTCCACCAGCTTCTGGATTTCGCATCCCGAAATACTTCAACAGCATTAGCCTGAACAACGAGTA
TACTATCATCAACTGTATGGAGAACAACAGCGGTTGGAAGGTGTCTCTGAACTATGGTGAGATCATTGACCTTGC
AGGACACCCAAGAGATCAAGCAGCGCGTCTGTTCAAGTACTCTCAAATGATCAACATTTCCGATTACATTAATCGT
TGGATCTTCGTGACCATTACGAATAACCGTCTGAATAACAGCAAGATTTACATCAATGGTCGCTTGATCGATCAGAA

ACCGATTAGCAACCTGGGTAATATCCACGCAAGCAACAACATTATGTTCAAATTGGACGGTTGCCGCGATACCCATC
 GTTATATCTGGATCAAGTATTTCAACCTGTTTGATAAAGAACTGAATGAGAAGGAGATCAAAGATTTGTATGACAAC
 CAATCTAACAGCGGCATTTTGAAGACTTCTGGGGCGATTATCTGCAATACGATAAGCCGTAATATGCTGAACCT
 GTATGATCCGAACAAATATGTGGATGTCAATAATGTGGGTATTCTGTGGTTACATGTATTTGAAGGGTCCGCGTGGCA
 GCGTTATGACGACCAACATTTACCTGAACTCTAGCCTGTACCGTGGTACGAAATTCATCATTAAAGAAATATGCCAGC
 GGCAACAAAGATAACATTGTGCGTAATAACGATCGTGTCTACATCAACGTGGTTCGTGAAGAATAAAGAGTACCGTCT
 GCGCAGCAACGCTTCGAGGCGGGTGTGAGAAAAATTCTGAGCGCGTTGGAGATCCCTGATGTCGGTAATCTGAGCC
 AAGTCGTGGTTATGAAGAGCAAGAACGACCAGGGTATCACTAACAAGTGCAAGATGAACCTGCAAGACAACAATGGT
 AACGACATCGGCTTTATTGGTTTTCCACCAGTTCAACAATATTGCTAAACTGGTAGCGAGCAATTGGTACAATCGTCA
 GATTGAGCGCAGCAGCCGTACTTTGGGCTGTAGCTGGGAGTTATCCCGTTCGATGATGGTTGGGGCGAACGTCCGC
 TG

[0299] SEQ ID NO:5. 工程化的BoNT/A, “CatH_N_v3”, 氨基酸序列。

[0300] MPFVNKQFNKYKDPVNGVDIAYIKIPNAGQMPPVKAFKIHNKIWWIPERDTFTNPEEGDLNPPPEAKQVP
 VSYDSTYLSTDNEKDNLYLKGVTKLFERIYSTDLGRMLLTSIVRGIPFWGGSTIDTELKVIDTNCINVIQPDGSYRS
 EELNLVIGPSADIIQFECKSFSGHEVLNLTRNGYGSTQYIRFSPDFTFGFEESEVDTNPLLGAGKFATDPAVTLAH
 ELIHAGHRLYGIAINPNRVFKVNTNAYYEMSGLEVSFEELRTFGGHDAKFIDSLQENEFRLYYNKFKDIASTLNKA
 KSIVGTTASLQYMKNVFKKEYLLSEDTSGKFSVDKLFKDKLYKMLTEIYTEDNFVKFFKVLNRKTYLNFDKAVFKIN
 IVPKVNYTIYDGFNLRNTNLAANFNGQNTENNMFNFKLKNFTGLFEFYKLLCVRGIITSKTKSLDKGYNKALNDLC
 IKVNNWDLFFSPSEDNFTNDLNKGEETSNTNIEAAEENISLDLIQQYYLTFNFDNEPENISLENLSSDIIGQLELM
 PNIERFPNGKKYELDKYTMFHYLRAQEFEHGKSRIALTNSVNEALLKPSRVYTFSSDYVKKVKNKATEAAMFLGWVE
 QLVYDFTDETSEVSTTDKIADITIIIPYIGPALNIGNMLYKDDFVGALIFSGAVILLEFIPEIAIPKLGTFALVSYK
 ANKVLTVQITDNALSKRNEKWDEVYKIVTNWLAKVNTQIDLIRKKMKEALENQAEATKAIINYQYNQYKEKEKNNI
 NFNIDDLSSKLNESINKAMININKFLNQCSVSYLMNSMIPYGVKRLDFDASLKDALLKYIYDNRGTLIGQVDRLKD
 KVNNTLSTDIPFQLSKYVDNQRLSTFTEYIKNIINTSILNLRYESNHLIDLRYASKINIGSKVNFDPIDKNQIQ
 FNLESSKIEVILKNAIVYNSMYENFSTSFWRIPKYFNSISLNNEYTIINCMENNSGWKVSLSNYGEIITWLTQDTQEI
 KQRVVFYKYSQMINISDYINRWIFVTITNNRLNNSKIYINGRLIDQKPIISNLGNIHASNNIMFKLDGCRDTHRYIWI
 YFNLFDKELNEKEIKDLYDNQNSGILKDFWGDYLYQYDKPYMLNLYDPNKYVDVNVGIRGYMYLKGPRGSVMTTN
 IYLNSSLYRGTKFIIKKYASGNKDNIVRNDRVYINVVKNKEYRLATNASQAGVEKILSALEIPDVGNSLQVVMK
 SKNDQGITNKCKMNLQDNNNDIGFIFGHQFNNAIKLVASNWYNRQIERSSRTLGCSEWEIFVDDGWGERPL

[0301] SEQ ID NO:6. 工程化的BoNT/A, “CatH_N_v3”, 核酸序列。

[0302] ATGCCATTCGTCAACAAGCAATTCAACTACAAAGACCCAGTCAACGGCGTCGACATCGCATAACATCAA
 GATTCCGAACGCCGGTCAAATGCAGCCGGTTAAGGCTTTTAAGATCCACAACAAGATTTGGGTTATCCCGGAGCGT
 GACACCTTACGAACCCGGAAGAAGGCGATCTGAACCCGCCACCGGAAGCGAAGCAAGTCCCTGTCAGCTACTACG
 ATTCGACGTACCTGAGCACGGATAACGAAAAAGATAACTACCTGAAAGGTGTGACCAAGCTGTTGCAACGTATCTA
 CAGCACGGATCTGGGTCGCATGCTGCTGACTAGCATTGTTTCGCGGTATCCCGTTCTGGGGTGGTAGCACGATTGAC
 ACCGAACTGAAGGTTATCGACACTAAGTGCATTAACGTTATTCAACCGGATGGTAGCTATCGTAGCGAAGAGCTGA
 ATCTGGTCATCATTGGCCCGAGCGCAGACATTATCCAATTCGAGTGCAAGAGCTTTGGTCACGAGGTTCTGAATCT
 GACCCGCAATGGCTATGGTAGCACCCAGTACATTCGTTTTTCGCCGGATTTTACCTTCGGCTTTGAAGAGAGCCTG
 GAGGTTGATACCAATCCGTTGCTGGGTGCGGGCAAATTCGCTACCGATCCGGCTGTCACGCTGGCCCATGAACTGA

TCCACGCAGGCCACCGCCTGTACGGCATTGCCATCAACCCAAACCGTGTGTTCAAGGTTAATACGAATGCATACTA
CGAGATGAGCGGCCTGGAAGTCAGCTTCGAAGAAGTGCACACCTTCGGTGGCCATGACGCTAAATTCATTGACAGC
TTGCAAGAGAATGAGTTCCGTCTGTACTACTATAACAAATTCAAAGACATTGCAAGCACGTTGAACAAGGCCAAAA
GCATCGTTGGTACTACCGCGTTCGTTGCAGTATATGAAGAATGTGTTTAAAGAGAAGTACCTGCTGTCCGAGGATAC
CTCCGGCAAGTTTAGCGTTGATAAGCTGAAGTTTGACAACTGTACAAGATGCTGACCGAGATTTACACCGAGGAC
AACTTTGTGAAATTCTTCAAAGTGTGAATCGTAAAACCTATCTGAATTTTGACAAAGCGGTTTTCAAGATTAACA
TCGTGCCGAAGGTGAACTACCCATCTATGACGGTTTTAACCTGCGTAACACCAACCTGGCGGCGAACTTTAACGG
TCAGAATACGGAAATCAACAACATGAATTTACGAAGTTGAAGAAGTTACGGGTCTGTTTCGAGTTCTATAAGCTG
CTGTGCGTGCAGGATCATCACCAGCAAACCAAAAGCCTGGACAAAGGCTACAACAAGGCGCTGAATGACCTGT
GCATTAAGGTAAACAATTGGGATCTGTTCTTTTCGCCATCCGAAGATAATTTTACCAACGACCTGAACAAGGGTGA
AGAAATCACCAGC_gA_tACGAATATTGAAGCAGCGGAAGAGAATATCAGCCTGGATCTGATCCAGCAGTACTATCTG
ACCTTTAACTTCGACAATGAACCGGAGAACATTAGCATTGAGAATCTGAGCAGCGACATTATCGGTCAGCTGGAAC
TGATGCCGAATATCGAACGTTCCCGAACGGCAAAAAGTACGAGCTGGACAAGTACACTATGTTCCATTACCTGCG
TGCACAGGAGTTTGAACACGGTAAAAGCCGTATCGCGCTGACCAACAGCGTTAACGAGGCCCTGCTGAA_aCCGAGC
CGTGTCTATACCTTCTCAGCAGCGACTATGTTAAGAAAGTGAACAAAGCCACTGAGGCCGCGATGTTCTGGGCT
GGGTGGAACAGCTGGTATATGACTTCACGGACGAGACGAGCGAAGTGAGCACTACCGACAAAATTGCTGATATTAC
CATCATTATCCCGTATATTGGTCCGGCACTGAACATTGGCAACATGCTGTACAAAGACGATTTTGTGGGTGCCCTG
ATCTTCTCCGGTGCCGTGATTCTGCTGGAGTTCATTCCGGAGATTGCGATCCCG_{aa}GTTGGGTACCTTCGCGCTGG
TGTCCTACA_{ag}GCGAATAAGGTTCTGACGGTTCAGACCATCGATAACGCGCTGTCGAAACGTAATGAAAAATGGGA
CGAGGTTTACAAATACATTGTTACGAATTGGCTGGCGAAAGTCAATACCCAGATCGACCTGATCCGTAAGAAAATG
AAAGAGGCGCTGGAGAATCAGGCGGAGGCCACCAAGCAATTATCAACTACCAATACAACCAGTACA_aGGAA_aAAG
AGAAGAATAACATTAACCTTCAATATCGATGATTTGAGCAGCAAGCTGAATGAATCTATCAACAAAGCGATGATCAAT
ATCAACAAGTTTTTGAATCAGTGTAGCGTTTTCGTACCTGATGAATAGCATGATTCGGTATGGCGTCAAACGTCTGGA
GGACTTCGACGCCAGCCTGAAAGATGCGTTGCTGAAATACATTTACGACAATCGTGGTACGCTGATTGGCCAAGTTG
ACCGCTTGAAAGACAAAGTTAAACAATACCCTGAGCACCGACATCCCATTTCAACTGAGCAAGTATGTTGATAATCAA
CGTCTGTTGAGCACTTTACCCGAGTATATCAAAAACATCATCAACTAGCATTCTGAACCTGCGTTACGAGAGCAA
TCATCTGATTGATCTGAGCCGTTATGCAAGCAAGATCAACATCGGTAGCAAGGTCAATTTTGACCCGATCGATAAGA
ACCAGATCCAGCTGTTTAACTGGAATCGAGCAAAATTGAGGTTATCCTGAAAAACGCCATTGTCTACAACCTCCATG
TACGAGAATTTCTCCACCAGCTTCTGGATTCGCATCCCGAAATACTTCAACAGCATTAGCCTGAACAACGAGTATAC
TATCATCAACTGTATGGAGAACAACAGCGTTGGAAGGTGCTCTGAACTATGGTGAGATCATTTGGACCTTGCAGG
ACACCCAAGAGATCAAGCAGCGCGTTCGTTCAAGTACTCTCAAATGATCAACATTTCCGATTACATTAATCGTTGG
ATCTTCGTGACCATTACGAATAACCGTCTGAATAACAGCAAGATTTACATCAATGGTTCGCTTGATCGATCAGAAACC
GATTAGCAACCTGGGTAATATCCACGCAAGCAACAACATTATGTTCAAATTGGACGGTTGCCGCGATACCCATCGTT
ATATCTGGATCAAGTATTTCAACCTGTTTGATAAAGAAGTGAATGAGAAGGAGATCAAAGATTTGTATGACAACCA
TCTAACAGCGGCATTTTGAAGGACTTCTGGGGCGATTATCTGCAATACGATAAGCCGTACTATATGCTGAACCTGTA
TGATCCGAACAAATATGTGGATGTCAATAATGTGGGTATTCGTGGTTACATGTATTTGAAGGGTCCGCGTGGCAGCG
TTATGACGACCAACATTTACCTGAACTCTAGCCTGTACCGTGGTACGAAATTCATCATTAAAGAAATATGCCAGCGGC
AACAAAGATAACATTGTGCGTAATAACGATCGTGTCTACATCAACGTGGTTCGTAAGAATAAAGAGTACCGTCTGGC
GACCAACGCTTCGCAGGCGGGTGTGAGAAAATTCTGAGCGCGTTGGAGATCCCTGATGTCGGTAATCTGAGCCAAG

TCGTGGTTATGAAGAGCAAGAACGACCAGGGTATCACTAACAAGTGCAAGATGAACCTGCAAGACAACAATGGTAAC
GACATCGGCTTTATTGGTTCCACCAGTTCAACAATATTGCTAAACTGGTAGCGAGCAATTGGTACAATCGTCAGAT
TGAGCGCAGCAGCCGTACTIONTTGGGCTGTAGCTGGGAGTTATCCCGTTCGATGATGGTTGGGGCGAACGTCGCTG

[0303] SEQ ID NO:7.工程化的BoNT/E轻链,“CatLC”,氨基酸序列。

[0304] MKIEEGKLVIIWINGDKGYNGLAEVGGKFEKDTGKIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRF
GGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKG
KSALMFNLQEPYFTWPLIAADGGYAFKYENKDYDIKDVGVNAGAKAGLTFLVDL IKNKHMNADTDYSIAEAFNK
GETAMTINGPWAWSNIDTSKVNIGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNK
DKPLGAVALKSYEEELAKDPRIAATMENAQGEIMPNIQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSN
NNNNNNNNLGI EGR ISEFGSMPKINSFNNDPVNDRTILYIKPGGCQEFYKSFNIMKNIWIIPERNVIGTTPQDF
HPPTSLKNGDSSYYDPNYLQSDEEKDRFLKIVTKIFNRINNNLSGGILLEELSKANPYLGNNDTPDNKFHIGDASA
VEIKFSKGSQHILLPNVIIMGAEPDLFETNSSNISLRNNYMPSNHGFGSIAIVTFSPEYSFRFNDNSINEFIQDPA
LTLMHELIHLHGLYGAKEITTCIITQQKNPLITNRKGINIEEFLTFGGNDLNIITVAQYNDIYTNLLNDYRKI
ASKLSKVQVSNPQLNPKYKDFQEKYGLDKDASGIYSVNINKFDDILKKLYSFTEFDLTKFQVKRETYIGQYKYFK
LSNLLNDSIYNISEGYNINNLKVNFRGQANLNPRIIKPIITGRGLVKKIIRFAVDKLAALAEHHHHHH

[0305] SEQ ID NO:8.工程化的BoNT/E轻链,“CatLC”,氨基酸序列。

[0306] ATGAAAATCGAAGAAGGTAACCTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAA
GTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAGTCACCGTTGAGCATCCGATAAACTGGAAGAGAAATCCC
ACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTGGTGGCTACGCTCAATCTG
GCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTAC
AACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCC
GCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACC
TGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTAC
GACATTAAGACGTGGGCGTGGATAACGCTGGCGGAAAGCGGGTCTGACCTTCTGGTTGACCTGATTA AAAACAA
ACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTTAATAAAGGCGAAACAGCGATGACCATCAACG
GCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAAACGGTACTGCCGACCTTCAAGGGTCAA
CCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAAGAGTT
CCTCGAAAACCTATCTGCTGACTGATGAAGGTCTGGAAGCGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGA
AGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACTATGGAAAACGCCAGAAAGGTGAAATCATG
CCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTGCTCAGAC
TGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCGAACAACAACAATAACAATAACAACAACCTCG
GGATCGAGGGAAGGATTTT CAGAATTCGGATCCATGCCAAAAATCAACAGCTTTAATTACAATGACCCTGTAAACGAT
CGTACCATCCTATACATAAAGCCGGGTGGGTGTCAAGAGTTCTACAAATCTTTCAATATTATGAAGAATATATGGAT
TATACCTGAGCGTAACGTTATTGGTACGACACC GCAAGATTTTCATCCACCTACTTCGTTGAAGAACGGTGACTCTT
CCTATTACGACCCCAATTATCTCCAGTCGGATGAAGAGAAGGACAGATTCCTTAAAATAGTAACCAAAATCTTTAAC
AGGATTAATAACAATCTATCCGAGGATTTTGGCTTGAAGAGCTTAGTAAAGCTAATCCTTACCTAGGTAACGATAA
TACACCAGACAACAAGTTTCATATAGGCGATGCATCCGCCGTGAAAATCAAATTTAGCAAGGGATCACAGCATATTC
TCTTGCCCAACGTTATTATAATGGGGGCGGAACCAGATTTATTGAAACAAATTCGAGTAATATTAGCCTGAGAAAT
AACTATATGCCGTCAAACCATGGGTTCCGGTAGCATAGCGATCGTACTTTTTCTCCGAATACAGTTTTCGCTTCAA

TGATAATAGTATAAATGAGTTTATCCAAGACCCCGCACTCACGCTTATGCACGAACCTACACTCTTTACACGGCC
TGTATGGCGCTAAGGGGATAAACCCTACGTGTATCATTACTCAGCAAAGAACCATTGATAACGAACAGGAAGGGC
ATTAACATCGAGGAATTTCTTACATTTGGAGGCAACGATCTGAACATTATAACTGTCGCACAGTACAATGACATCTA
TACCAACTTACTAAATGATTATAGAAAAATCGCTTCTAAGTTATCCAAGGTTCAAGTCTCAAACCCTCAACTGAATC
CGTATAAGGACATATTCCAAGAGAAATATGGATTAGACAAAGACGCGTCAGGAATCTATTCGGTAAACATTAACAAA
TTCGACGATATTTGAAGAACTTTACAGCTTACGGAGTTCGACTTGGCCACCAAATTCAGGTCAAATGCCGAGA
GACATACATCGGACAGTATAAGTATTTCAAGCTGTGCAATCTCCTGAATGATTCCATATACAACATTAGTGAGGGTT
ACAATATAAATAACCTAAAGGTGAATTTCCGAGGCCAAAACGCCAACCTAAATCCGCGCATCATTAACCCATCACA
GGACGGGGGTTAGTGAAGAAAATAATCCGGTTTGGGTCGACAAGCTTGGCGCCGCACTCGAGCACCACCACCACCA
CCAC

实施例

[0307] 以下实施例用于说明本发明的具体实施方案,并不以任何方式限制权利要求中限定的本发明的范围。

[0308] 实施例1

[0309] 鉴定用于修饰的优选的梭菌毒素氨基酸。

[0310] 使用多种不同的标准选择鉴定为适合用于修饰(突变位点)的候选物的氨基酸。

[0311] 1. BoNT分子内残基的位置(H_N 内,不包括带区域);

[0312] 2. 二级/三级结构的位置;

[0313] 3. 残基的类型;

[0314] 4. 表面暴露度。

[0315] 对于选择,考虑酸性、中性、极性和疏水性残基。

[0316] 使用CCP4套件中的AreaIMol测定暴露的残基。通过AreaIMol分析每个结构,并且将暴露的残基鉴定为具有大于55的和值。

[0317] 使用二级结构分配程序(Stride Web Interface)鉴别每个亚型和TeNT的 H_N 内的二级结构。从选择中排除分配为形成 α -螺旋、 β -链或 3_{10} -螺旋的区域。

[0318] 所使用的序列

[0319] 登录号:

[0320] BoNT/A:P10845

[0321] BoNT/B:P10844

[0322] BoNT/C₁:P18640

[0323] BoNT/D:P19321

[0324] BoNT/E:Q00496

[0325] BoNT/F:YP_001390123

[0326] BoNT/G:Q60393

[0327] 结构数据源

[0328] 从RCSB获得的BoNT/A(3BTA.pdb)、BoNT/B(1EPW)和BoNT/E(3FFZ.pdb)的晶体结构。

[0329] 使用LOOPP和以下序列分别进行BoNT/C₁、BoNT/D、BoNT/F和BoNT/G的同源性建模:

P18640、P19321、YP_001390123和Q60393。

[0330] 在梭菌毒素H_N结构域中优选的用于修饰的梭菌毒素氨基酸残基:

[0331] BoNT/A:

[0332] D474, N476, D484, N486, I487, E488, A489, A490, E491, D546, E558, E560, H561, I566, L568, N570, S571, L577, N578, A597, E599, A601, E620, V621, T623, D625, T631, N645, L647, D650, D651, I668, E670, A672, V675, S683, I685, A686, N687, N752, Q753, T755, E756, E757, E758, N760, N761, I762, N763, D825, I831, G832, T847, D848, 和D858

[0333] BoNT/B:

[0334] V443, G444, D453, S468, D533, E534, N535, T545, L548, D549, I550, D552, S557, L564, S566, N582, V584, N609, L619, N632, E633, G637, A646, I655, E657, V662, E669, S670, I672, D673, N739, I740, N748, N750, I818, G819, T834, I842, N845, 和S858

[0335] BoNT/C₁:

[0336] L451, D452, C453, E455, V472, T474, D475, L478, N483, E484, E485, E487, I489, L555, S556, D557, N558, E560, D561, E569, N574, S575, T584, G592, Q594, G596, D617, N640, S641, V642, G645, N646, E661, E665, T667, A670, S678, V680, Q681, E682, S750, G751, S759, Q760, V826, G827, N842, T843, N847, 和N853

[0337] BoNT/D:

[0338] Q469, E470, E473, N474, D479, E480, N482, V483, Q484, N485, S487, D488, S552, N553, N554, V555, E556, N557, I558, L560, T562, S563, V564, G569, S571, N572, G588, Q590, T614, D616, S619, S622, N636, S637, L639, G641, N642, E657, E661, T663, A666, V669, S674, I676, Q677, E678, S746, G747, D749, E751, N752, I753, Q756, N818, V822, G823, E837, N838, T839, N843, N849, 和N850

[0339] BoNT/E:

[0340] D474, N476, E479, E480, D484, N486, I487, E488, A489, A490, E491, E492, L496, D497, Q500, Q501, L504, N507, D509, N510, N514, S516, E518, Q527, L530, N533, I534, E535, N539, Y548, I566, L568, D589, A597, E599, A601, L604, Y612, E620, N645, L647, Y648, D651, E737, E741, Y803, Y824, D825, G828, I831, G832, 和D835

[0341] BoNT/F:

[0342] N463, E464, N468, T469, D474, D475, T476, T477, N478, N482, N485, N495, I499, Q501, I502, Q505, T506, N508, T509, V511, D513, D521, S522, S526, E527, I528, E529, V534, D535, L536, E549, G550, T552, N553, S558, E566, E567, S568, V586, H587, Q608, D613, A616, D617, S619, N630, N633, N639, E654, V656, E658, L660, T663, L665, V666, S671, I673, G674, S675, S676, E677, N678, T746, N751, L753, E754, E756, N758, I759, N760, N761, S799, S821, I822, N840, S841, E845, L846, S847, S848, T850, N851, D852, I854, L855, 和I856

[0343] BoNT/G:

[0344] N480, Q482, N483, N484, T485, E487, D540, N562, N570, N571, N572, T588, V589, T615, D621, N637, E638, E642, N643, I660, E662, I667, E674, S675, V677, G678, N679, S747, N755, D757, L823, D839, I841, D844, S846, 和L847

[0345] TeNT:

[0346] A457, S458, L459, D461, L462, E486, E487, Q490, D491, N497, N504, D557, T571, T572, L573, Q574, N580, S581, N588, S589, T590, S598, Q605, G606, Q608, T631, I633, S640, Q655, E658, G659, N660, E675, I677, E679, T681, V684, A691, E692, S694, T695, Q696, A772, D773, E774, S862, N866, L867和D868

[0347] 在BoNT/E轻链中优选的用于修饰的梭菌毒素氨基酸残基:

[0348] N5, N8, N10, D11, N14, D15, Q27, E28, N72, Q75, N118, D121, N122, Q123, N138, Q237, Q290, N297, N362, N365, D366, N378, 和N379.

[0349] 实施例2

[0350] 工程化的BoNT/A分子的设计

[0351] 制备了根据本发明的工程化的BoNT/A分子的三个不同实例。

[0352] 使用实施例1中描述的方法, 将总共55个残基鉴定为BoNT/A H_N结构域中用于突变的候选物。通过BoNT/A晶体结构的目视检查进一步评估残基的适用性, 列出11个优选的候选物 (N476、N763、N687、E599、I831、N761、N578、V675、I685、T755、E757)。基于功能数据选择另外6个残基, 该功能数据显示这些残基适于突变而不会不利地影响蛋白质功能。这些残基中的4个在候选列表中 (L647、D650、D651、T847), 2个不在候选列表中 (S564、I849)。候选列表中的11个残基加上功能数据中的6个残基, 共选择17个残基用于突变。

[0353] 由所选择的17个残基, 制备3个构建体: CatH_N_v1、CatH_N_v2和CatH_N_v3。CatH_N构建体的突变显示在下表3中:

[0354] 表3: 具有所列出的突变的CatH_N构建体和计算的pI。

	突变	突变的数目	计算的 pI	相对于 BoNT/A (pI 6.4)的计算的 ΔpI
[0355]	CatH_N_v1 S564R, L647R, D650R, D651R, T847R,	6	7.4	1.0

	I849R			
[0356]	CatH_N_v2 N476K, N763K, N687K, E599K, I831K, N761K	6	7.3	0.9
	CatH_N_v3 N578K, V675K, I685K, T755K, E757K	5	7.1	0.7

[0357] [Δ pI = 等电点的改变]

[0358] CatH_N_v1、CatH_N_v2和CatH_N_v3的纯化分别如图1A、1B和1C所示。

[0359] 实施例3

[0360] 克隆、表达和纯化

[0361] 合成实施例2所述的编码工程化的BoNT/A分子的DNA构建体，克隆到pJ401表达载体中，然后转化到BL21 (DE3) 大肠杆菌中。这允许重组工程化的BoNT/A分子在大肠杆菌中可溶性的过表达。

[0362] 使用传统的色谱技术从大肠杆菌裂解物中纯化重组工程化的BoNTs。应用使用阳离子交换树脂的初始纯化步骤，然后是使用疏水相互作用树脂的中间纯化步骤。然后通过蛋白水解切割重组工程化的BoNT单链，得到活化的双链工程化的BoNT。然后采用最终纯化步骤去除剩余的污染物。

[0363] 实施例4

[0364] 纯化的工程化的BoNTs的表征

[0365] 用试验对上述实施例2中描述的工程化的BoNTs进行表征。

[0366] 使用大鼠胚胎脊髓神经元 (eSCN) 评估工程化的BoNTs进入神经元并切割SNAP-25 (BoNT/A的靶) 的能力。使用小鼠膈神经偏侧膈测定法 (mPNHD) 进一步评估工程化的BoNTs的效力。

[0367] CatH_N_v1:

[0368] 加入的第一组突变取代为精氨酸:

[0369] S564R、L647R、D650R、D651R、T847R、I849R

[0370] 在大鼠胚胎脊髓神经元 (eSCN) 的SNAP-25切割测定法中测试CatH_N_v1分子，并且发现其与BoNT/A (BoNT/A) 具有同等效力 (图2)。

[0371] 在小鼠膈神经偏侧膈 (mPNHD) 测定法中也证实了阳性结果 (图3)。

[0372] CatH_N_v2:

[0373] 第二组突变取代为赖氨酸:

[0374] N476K、N763K、N687K、E599K、I831K、N761K

[0375] 在eSCN SNAP-25切割测定法中测试了CatH_N_v2蛋白,并且发现其保留进入细胞并切割SNAP-25的能力。在mPNHD测定法中,CatH_N_v2与BoNT/A等效(图4和5)。

[0376] CatH_N_v3:

[0377] 第三组突变取代为赖氨酸:

[0378] N578K、V675K、I685K、T755K、E757K

[0379] 在eSCN SNAP-25切割测定法中测试CatH_N_v3分子,并且发现其保留进入细胞并切割SNAP-25的能力。类似地,在mPNHD测定法中也证实了阳性结果(图6和7)。

[0380] 等电点聚焦

[0381] 与未修饰的BoNT/A相比,所有三种CatH_N构建体具有增加的pI(图8)。

[0382] 实施例5

[0383] BoNT/E轻链中的修饰

[0384] 由于肉毒梭菌毒素的模块化,使用具有N-末端麦芽糖结合蛋白(MBP)标签和C末端6组氨酸标签(6HT)的BoNT/E轻链构建体作为替换物来测定突变和表征时的BoNT/E活性。

[0385] 制备具有下表所示突变的BoNT/E轻链构建体("CatLC")。

[0386] 表4:具有所列突变的构建体、计算的pI和计算的 ΔpI 。

	突变	突变的数目	计算的 pI	相对于 MBP-LC/E (pI 6.3)计算的 ΔpI
[0387]	Q123K N138K Q237K	3	6.6	0.3

[0388] 该构建体在BL21DE3细胞中表达并且用亲和层析纯化。纯化如图9所示。

[0389] 对构建体催化活性的评估表明,修饰的轻链保留了未修饰轻链的催化活性(图10)。

序列表

<110> 益普生物创新有限公司
 <120> 阳离子神经毒素
 <130> P44007WO
 <160> 37
 <170> PatentIn 版本 3.5
 <210> 1
 <211> 1296
 <212> PRT
 <213> 人工序列
 <220>
 <223> 工程化的BoNT/A
 <400> 1
 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

[0001]

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
 465 470 475 480

Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu
 485 490 495

Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro
 500 505 510

Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu
 515 520 525

Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu
 530 535 540

[0002] Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu
 545 550 555 560

His Gly Lys Arg Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu
 565 570 575

Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys
 580 585 590

Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu
 595 600 605

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
 625 630 635 640

Leu Asn Ile Gly Asn Met Arg Tyr Lys Arg Arg Phe Val Gly Ala Leu
 645 650 655

Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala
 660 665 670

Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys
 675 680 685

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
 705 710 715 720

Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu
 725 730 735

Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn
 740 745 750

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 Arg Asp
 835 840 845
 Arg 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

[0003]

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> 2
 <211> 3888
 <212> DNA
 <213> 人工序列

<220>
 <223> 工程化的BoNT/A

[0004]

<400> 2
 atgccattcg tcaacaagca attcaactac aaagaccag tcaacggcgt cgacatcgca 60
 tacatcaaga ttccgaacgc cggtcanaatg cagccgggta aggcittttaa gatccacaac 120
 aagatttggg ttatcccgga gcgtagacacc ttacagaacc cgaagaagaag cgaatctgaac 180
 ccgccaccgg aagcgaagca agtccctgtc agctactacg attcgaagta cctgagcagc 240
 gataacgaaa aagataacta cctgaaaggc gtgaccaagc tgttcgaacg tatctacagc 300
 ccggatctgg gtcgatgct gctgactagc attgttcgag gtatcccggt ctgggggtgg 360
 agcacgattg acaccgaact gaaggttatc gacactaact gcattaacgt tattcaaccg 420
 gatggtagct atcgtagcga agagctgaat ctggctatca ttggcccag cgcagacatt 480
 atccaattcg agtcaagag ctttggctac gaggttcga atctgacccc caatggctat 540
 ggtagcacc cgtacattcg tttttcggc gattttacct tcggcttga agagagcctg 600
 gaggttgata ccaatccggt gctgggtgag ggcaattcg ctaccgatcc ggctgtcagc 660
 ctggcccagc aactgatcca cgcaggccac cgcctgtacg gcattgccat caaccacaac 720
 cgtgtgttca aggttaatac gaatgcatac tacgagatga cggccctgga agtcagcttc 780
 gaagaactgc gcaccttogg tggccatgac gctaaattca ttgacagctt gcaagagaat 840
 ggttccgtc tgtactacta tsacaaatc aaagacatg caagcagctt gaacaaggcc 900
 aaaagcatcg ttggtactac cgcgtcgttg cagtatatga agaatgtgtt taaagagaag 960
 tacctgtgt cccaggatac ctccggcaag tttagcgttg ataagctgaa gtttgacaaa 1020
 ctgtacaaga tctgacoga gatttacacc gaggacaact ttgtgaaat cttcaaatgt 1080
 ttgaatcgt aaacctatct gaattttgac aaagcggitt tcaagattaa catcgtgccg 1140
 aaggtagaact acaccatctt tgacggtttt aacctgcgta acaccaacct ggccggcaac 1200
 tttaacggtc agaatacga aatcaacaac atgaattca cgaagttaa gaacttcagc 1260
 ggtctgttc agttctataa gctgctgtgc gtgcgggta tcatcaccag caaaacaaa 1320
 agcctggaca aaggctacaa caaggcgtc aatgacctgt gcattaaggt aaacaattgg 1380
 gatctgttct ttccccatc cgaagataat ttaccaacg acctgaacaa gggtagaaga 1440
 atcaccagcg atacgaata tgaagcagcg gaagagaata tcagccttga tctgatccag 1500
 cagctatctc tgacctttaa ctccgacaat gaaccggaga acattagcat tgagaatctg 1560
 agcagcgaca ttatcggta gctgggaactg atccgaata tcgaaccttt cccgaacggc 1620
 aaaaagtacg agctggacaa gtacactatg ttccattacc tgcgtgcaca ggagttagaa 1680
 caccgtaaac gtcgtatcgc gctgaccnac agcgttaacg aggccctgct gaacccgagc 1740
 cgtgtctata ccttcttcag cagcagctat gttaaagaag tgaacaaagc cactgaggcc 1800
 gcgatgttcc tgggctgggt ggaacagctg gtatatgact tcacggacga gacgagcga 1860

```

gtgagcacta cgcacaaaat tgcctgatatt accatcatta tcccgtatat tggtcggcca 1920
ctgaacattg gcaacaatcg ttacaacagt cgtttgtggt gfgccctgat cttctccggt 1980
gccgtgatcc tgcctggagt cttccggag attgcgatcc cgggtgtggg taccttcgag 2040
ctgggttctt acatcgcgaa taaggttctg acggttcaga ccatcgataa cgcgctgctg 2100
aaacgtaatg aaaaatggga cgggtttac aaatacattg ttacgaattg gctggcgaaa 2160
gtcaataccc agatcgacct gatccgtaag aanaatgaaag aggcgctgga gaatcaggcg 2220
gaggccacca aagcaattat caactaccna tacaaccagt acacggaaga agagaagaat 2280
aacattaaat tcaatataca tgattigagc agcaagctga atgaatctat caacaaagcg 2340
atgataanta tcaacaagtt ttgaaatcag tgtagcgttt cgtacctgat gaatagcatg 2400
attccgtatg gctgcaaacg tctggaggac ttccagccca gctgaaaga tgcgttctctg 2460
aaatacattt acgacaatcg tggtagcctg attggccaag ttgacccgtt gaaagacaaa 2520
gttaacaata ccttgagccg tgaccgtcca ttcaactgca gcaagtaigt tgataatcaa 2580
cgtctgttga gcactttcac cgagtatac aaaaacatca tcaatactag cattctgaa 2640
ctgctgttag agagcaatca tctgatfgat ctgagccgtt atgcaagcaa gatcaacatc 2700
ggtagcaagg tcaatttga cccgatcgat aagaaccaga tccagctgtt taatctggaa 2760
tcgagcaaaa ttgaggttat cctgaaaaac gccattgtct acaactccat gtagcagaaat 2820
ttctccacca gctctcggat tgcgatcccg aaatacttca acagcattag cctgaaacac 2880
gagfatacta tcaatcaactg taiggagasc aacagcgggt ggaaggtgtc tctgaactat 2940
ggtagatca tttggacctt gcaggacacc caagagatca agcagcgcgt cgtgttcaag 3000
tactctaaa tgatcaacat ttccgattac ataatcgtt ggatcttctg gaccattacg 3060
aataaccgtc tgaataacag caagatttac atcaatggtc gcttgatcga tcagaaaccg 3120
atagcaacc tgggtaatat ccacgcaagc aacaacatta tgttcaaat ggacggttgc 3180
cgcgataacc atcgttatai ctggatcaag tatttcaacc tgttigataa agaactgaat 3240
gagaaggaga tcaaaagatt gtatgacaac caatctaaac gcggcatttt gaaggacttc 3300
tggggcgatt atctgcaata cgataagccg tactatatgc tgaacctgta tgatccgaac 3360
aaatatgtgt atgtcaataa tgtgggtatt cgtggttaca tgtatttga gggtcgccgt 3420
ggcagcgata tgacgaccaa catttacctg aactctagcc tgtaccgtgg tacgaaatc 3480
atcattaaag aatatgccag cggcaacaaa gataacattg tgcgtaataa cgatcgtgtc 3540
tacctcaacg tggctgtgaa gaataaagag taccgtctgg cgaccaacgc ttccgaggcg 3600
gggtttgaga aaattctgag cgcgttggag atccctgatg tccgtaatct gagccaagtc 3660
gtggttatga agagcaagaa cgaccagggt atcactaaca agtgcaagat gaacctgcaa 3720
gacaacaatg gtaacgacat cggctttatt ggtttccacc agttcaacaa tattgctaaa 3780
ctggtagcga gcaattggta caatcgtcag attgagcgca gcagccgtac tttggcgctg 3840
agctgggagt ttatcccggt cgatgatggt tggggcgaaac gtccgctg 3888

```

[0005]

```

<210> 3
<211> 1296
<212> PRT
<213> 人工序列

<220>
<223> 工程化的BoNT/A

<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 Lys Lys Gly Glu Glu
 465 470 475 480

Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu
 485 490 495

Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro
 500 505 510

Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu
 515 520 525

[0006]

[0007]

Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu
 530 535 540
 Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu
 545 550 555 560
 His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu
 565 570 575
 Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys
 580 585 590
 Lys Val Asn Lys Ala Thr Lys Ala Ala Met Phe Leu Gly Trp Val Glu
 595 600 605
 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
 625 630 635 640
 Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu
 645 650 655
 Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala
 660 665 670
 Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala Lys Lys
 675 680 685
 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
 705 710 715 720
 Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu
 725 730 735
 Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn
 740 745 750
 Gln Tyr Thr Glu Glu Glu Lys Asn Lys Ile Lys 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 Lys 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

[0008]

<210> 4
 <211> 3888
 <212> DNA
 <213> 人工序列

<220>
 <223> 工程化的BoNT/A

<400> 4
 atgccattcg tcaacaagca attcaactac aaagaccag tcasggcggt cgacatcgca 60
 tacatcaaga ttccgaagcg cggtcasatg cagccgggta aggtttttaa gatccacaac 120

[0009]

aaagatttggg ttatcccggg gcgtgacacc ttcacgaacc cggagaagg cgaatctaac	180
ccgccaccgg aagcgaagca aglccctgtc agctactacg attcgacgta cctgagcacg	240
gataacgaaa aagataacta cctgaaaggt gtagccaagc tgttcgaacg taictacagc	300
acggatctgg gtgcgatgct gctgactagc attgttcgcg gtatcccgtt ctgggggtgt	360
agcacgattg acaccgaact gaaggttatc gacactaact gcattaacgt taatcaaccg	420
gatgtagct atcgtagcga agagctgaat ctggatca ca ttggcccag cgcagacatt	480
atccaattcg agtgcgaagc ctltggctac gaggttctga atctgacccg caatggctat	540
ggtagcacc caglacattcg tttttccgcg gattttacct tggccttga agagagcctg	600
gaggttgata ccaatccgtt gctgggtgcg ggcaaatcgc ctaccgatcc ggctgtcacg	660
ctggcccatt aactgatcca cgcaggccac gcctgtacg gcattgccat caacccaac	720
cgtgtgttca aggttaatac gaatgcatac tacgagatga cggccctgga agtcagcttc	780
gaagaactgc gcaccttcgg tggccatgac gctaaatca ttgacagctt gcaagagaat	840
gagttccgtc tctactacta taacaaatc aaagacattg caagcagctt gaacaaggcc	900
aaaagcctcg ttggtactac cgcctcgttg cagtatatga agaattgttt taaagagaag	960
tacctgcgt cggaggatca ctccggcaag tttagcgttg ataagctgaa gtttgacaaa	1020
ctgtacaaga tctgacccga gatttacacc gaggacaact ttgtgaaatt cttaaaagt	1080
ttgaatcgtt aaacctatct gaattttgac aaagcggttt tcaagattaa catcgtgccg	1140
aaagtgaaat acacactcta tgacggtttt aacctgcgta acaccaacct ggccggcaac	1200
tttaacgctc agaatacggg aatcaacaac atgaatttca cgaagtgaa gaacttcacg	1260
ggctctttcg agttctataa cgtcctgtgc gtgcgcggta tcatcaccag caaaacaaa	1320
agcctggaca aaggctacaa caaggcgtcg aatgacctgt gcattaaagt aaacnaattg	1380
gatctgttct tttcgccatc cgaagataat ttaccacag acctgaagaa gggtagaaga	1440
atcaccagcg atacgaatat tgaagcagcg gaagagata tcagcctgga tctgattccag	1500
cagtactatc tgacctttaa cttcgacaat gaaccggaga acattagcat tgagaatctg	1560
agcagcagca ttatcggta cctggactcg atgcccgaat tcgaactgtt cccgaacggc	1620
aaaagtagcg agctggacaa gtacactatg ttccattacc tgcgtgcaca ggagttgaa	1680
caacgtaaaa gccgtatcgc gctgaccaac agcgttaacg agccctgct gaacccgagc	1740
cgtctctata ccttcttcag cagcactat gttagaaga tgaacaaagc cactaaggcc	1800
gcgattgttc tggcctgggt ggaacagctg gtaatgact tcacggacga gacgagcga	1860
gtgagcacta cagacaaaat tgcgtatatt acctatcatt tccgtatatt tggctcggca	1920
ctgaacattg gcaacatgct gtacaasgac gatitgttgg gtgccctgat ctctccgggt	1980
gccgtgattc tgcgtgagtt cactccggag attgccatcc cgggttggg taccttcgcg	2040
ctgggttctt acatcgcgaa gaaggttcig acggttcaga ccatcgataa cgcgctgtcg	2100
aaacgtaatg aaaaaaggga cgaagtttac aaatacattg ttacgaattg gctggcga	2160
gtcaatacc cagatcgacct gatccgtaag aaatgaaag aggcctgga gaatcaggcg	2220
gagccacca aagcaattat caactacca tacaaccagt acacggaaga agagaagaat	2280
aaagtaagt tcaatctga tgaattgagc agcaagctga atgaatctat caacaagcg	2340
atgataata tcaacaagtt ttgaaatcag tctagcgttt cgtacctgat gaatagcatg	2400
attccgtatg gcgtcaaacg tctggaggac ttgacgcca gccgaaaga tgcgttctcg	2460
aaatacattt acgacaatcg tggtagctcg aaggcccaag ttgacccctt gaaagacaaa	2520
gttaacaata cctgagcac cgcacatcca ttcaactga gcaaglatgt tgataataa	2580
cgtctgttga gcactttcac caglatatc aaaaacatca tcaatctag caictgaa	2640
ctgcttacg agagcaatca tctgattgat ctgagccgtt atgcaagca gaatcaatc	2700
ggtagcaagg tcaattttga cccgatcgat aagaaccaga tccagctgtt taactggaa	2760
tcgagcaaaa ttgagttat cctgaaaaac gccattgtct acaactcat gtacgagaat	2820
ttctccacca gctctggat tgcacatccg aaacttca acagcattag cctgaaac	2880
gagtalacta tcaatcaactg taaggagaac aacagcgtt ggaagggtc tctgaactat	2940
ggtagatca tttggacctt gcaggacacc caagagatca agcagcgcgt cgtgttcaag	3000
tactctaaa tgaatcaact ttccgattac ataanctgtt ggatctctgt gacattacg	3060
aataaccgtc tgaataacag caagatttac atcaatggct ccttgatcga tcagaaaccg	3120
attagcaacc tgggttaatat ccacgcaagc aacaacatta tgttcaaat ggacggttc	3180
cggatacc cagcttatat ctggatcaag tattcaacc tgtttgataa gaaactgaat	3240
gagaaggaga tcaaaagatt gtatgacaac caatctaaca cggccatttt gaaggacttc	3300

```

tggggcgatt atctgcaata cgataagccg tactatatagc tgaacctgta tgaatccgaac 3360
aaafatgtgg atgtcaataa tgggggtatt cgtgggttaca tgtattttaa gggtcgcgct 3420
ggcagcgfta tgaacgacaa catttaccctg aactctagcc tgtaccgtgg tacgaaatc 3480
atcattaaga aatattgccag cggcaacaaa gataacattg tgcgtaataa cgatecgtgc 3540
tacctcaacg tggtcgtgaa gaataaagag taccgtctgg cgaccaacgc ttcgcaggcg 3600
gggtttgaga aaattctgag cgcgttggag atccctgatg tcggtaatct gagccaagtc 3660
gtggttatga agagcaagaa cgaccagggt atcactaaca agtcaagat gaacctgcaa 3720
gacaacaatg gtaacgacat cggctttatt ggtttccacc agttcaacaa tattgctaaa 3780
ctggtagcga gcaattggta caatcgtcag attgagcga gcagccgtac tttggcgtgt 3840
agctgggagt ttatcccggt cgateatggt tggggcgaac gtccgctg 3888

```

```

<210> 5
<211> 1296
<212> PRT
<213> 人工序列

<220>
<223> 工程化的BoNT/A

<400> 5

```

```

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

```

[0010]

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
 465 470 475 480
 Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu Asn Ile Ser Leu
 485 490 495
 Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe Asp Asn Glu Pro
 500 505 510
 Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu
 515 520 525
 Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu
 530 535 540
 Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu
 545 550 555 560
 His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu
 565 570 575
 Leu Lys Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys
 580 585 590
 Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu
 595 600 605
 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
 625 630 635 640
 Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu
 645 650 655
 Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala
 660 665 670
 Ile Pro Lys Leu Gly Thr Phe Ala Leu Val Ser Tyr Lys Ala Asn Lys
 675 680 685
 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
 705 710 715 720

[0011]

[0012]

Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu
 725 730 735
 Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn
 740 745 750
 Gln Tyr Lys Glu Lys 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> 6
 <211> 3888
 <212> DNA
 <213> 人工序列

<220>
 <223> 工程化的BoNT/A

[0013]

<400> 6
 atgccattcg tcaacaagca attcaactac aaagaccag tcaacggcgt cgacatcgca 60
 tacatcaaga ttccgaagc cggtcanaat gagccggtta aggcctttaa gatccacaac 120
 aagatttggg ttatcccga gcgtgacacc ttcacgaacc cggagaagg cgatctgaac 180
 ccgccaccgg aagcgaagca agtccctgtc agctactacg attcgacgta cctgagcagc 240
 gataacgaaa aagataacta cctgaaaggt gtgaccaagc tgttcgaacg tatctacagc 300
 acggatctgg gtcgcatgct gctgactagc attgttcgcg gtatcccgtt ctgggttggg 360
 agcacgattg acaccgaact gaagttate gacactaact gcattaacgt tattcaaccg 420
 gatgtagct atcgtagcga agagctgaat ctggatca tggccccgag cgcagacatt 480
 atccaattcg agtgcagag ctttgggtac gaggttctga atctgaccgg caatggctat 540
 ggtagcacc cagtacattc ttttcgccc gatittacct tcggctttga agagagcctg 600
 gaggttgata ccaatccgtt gcctgggtgc ggcaaatcg ctaccgatcc ggctgtcagc 660
 ctggccccat aactgatcca cgcaggccac cgcctgtacg gcattgccat caaccacaac 720
 cgtgtgttca aggttaatac gaatgcatac tacgagatga gcggcctgga agtcagcttc 780
 gaagaactgc gcacctcgg tggccatgac gctaaattca ttgacagctt gcagagaaat 840
 gagttccgtc tgtactacta taacaantc aaagacattg caagcacgtt gaacaaggcc 900
 aaaagcattc ttgtactac cgcgtcgttg cagtatatga agaatgtgtt taagagaaag 960
 tacctcctgt ccgaggatac ctccggcaag tttagcgttg ataagctgaa gtttgacaaa 1020
 ctgtacaaga tctgaccgca gatttacacc gaggacaact ttgtgaant ctccaagtg 1080
 ttgaatcgta aaacctatct gaattttgac aaagcgggtt tcaagattaa catcgtgccc 1140
 aaggtgaact acaccatcta tgacggtttt aacctgctga acaccaacct ggccggcaac 1200
 tttacggctc agaatagcga aatcaacaac atgaattca cgaagtga gaacttcacg 1260
 ggtctgttcg agttctataa gctgctgtgc gtgcgcggta tcatcaccag caaaaacaaa 1320
 agccttgaca aaggctacaa caaggcgcgt aatgacctgt gcattaaggt naacaattgg 1380
 gatctgttct ttctccatc cgaagataat ttaccacag accigaacaa gggtagaana 1440
 atcaccagcg atacgaatat tgaagcagcg gaagagaata tcagcctgga tctgatccag 1500
 cagtactatc tgacctttaa ctctgacaat gaaccggaga acattagcat tgagaatctg 1560

[0014]

```

agcagcgaca ttatcggta gctgggaactg atgccgaata tcgaacgttt cccgaacggc 1620
aaaaagtacg agciggacaa gtacactatg ttccattacc tgcgtgcaca ggagtttgaa 1680
cacggtaaaa gccgtatcgc gctgaccaac agcgttaacg aggccctgct gaaaccgagc 1740
cgtgctata cctcttcag cagcactat gttaaagaag tgaacaagc cactgaggcc 1800
ggatgttcc tggcctgggt ggaacagctg gtatatgact tcacggacga gacgagcgaa 1860
gtgagcacta ccgacaaaat tctgatatt accatcatta tcccgtatat tggtcggca 1920
ctgaacattg gcaacatgct gtacaaagac gatitgtgg gtgccctgat cttctcgggt 1980
gccgtgatcc tctgggagt cttccggag atfgcgaacc cgaagtggg tacctcggc 2040
ctgggtcctt acaaggcgaa taaggtctg acggttcaga ccatcgataa cgcgctgctg 2100
aaacgtaatg aaaaaaggga cgaagtttac aatacattg ttacgaattg gctggcgaaa 2160
gtcaataccc agatcgacct gatccgtaag aaaaagaag aggcgctgga gaatcaggcg 2220
gagccacca aagcaattat caactacca tacaaccagt acaaggaaa agagaagaat 2280
aacattasct tcaatatcga tgattgagc agcaagctga atgaatctat caacaagcg 2340
atgatcaata tcaacaagtt ttgaaacag tctagcgttt cgtacctgat gaatagcatg 2400
atccgtatg gcgtcaaacg tctggaggac ttcgacgcca gccgaaga tgctgtctg 2460
aatacattt acgacaatg tggtagctg atfgccaaq ttgacgctt gaaagcaaa 2520
gttaacaata ccttgagcac cgacatcca ttcaactga gcaagtatgt tgataatcaa 2580
cgtctgttga gcactttcac cgagtatac aaaaaatca tcaactactag cactctgaac 2640
ctgcttacg agagcaatca tctgatgat ctgagccgtt atgcaagcaa gatcaacatc 2700
ggtagcaagg tcaatttga cccgatcgt aagaaccaga tccagctggt taatctggaa 2760
tcgagcaaaa ttgaggttat cctgaaaaac gccattgtct acaactccat gtacgagaat 2820
ttccacca gctctggat tgcctcccg aatacttca acagcattag cctgaaaca 2880
gagtacta tcatcaactg tatggagaac aacagcgtt ggaagggtc tctgaactat 2940
ggtagatac ttggacctt gcaggacacc caagatca agcagcgcgt cgtgttcaag 3000
tacttcaaa tgatcaaat tccgattac attaactgtt ggatctcgt gaccattacg 3060
aataaccgtc tgaataacag caagattac atcaatggtc gctgatcga tcagaaaccg 3120
attagcaacc tgggtaatat ccacgcaagc acaacatta tgttcaait ggaagggtgc 3180
cgcgatcccc atcgttata ctggatcaag tattcaacc tgttgataa agaactgaat 3240
gagaaggaga tcaaaagatt gtagacaac caatcaaca cggcatttt gaaggacttc 3300
tggggcgatt atctgcaata cgataagccg tactataatc tgaaccgta tgaaccgaac 3360
aaatatgtgg atgtcaataa tgtgggtatt cgtgggttaca tgtattgaa gggtcgcgt 3420
ggcagcgtta tgaagcaaa caatfacctg aactctagcc tgaaccgtg tacgaattc 3480
atcaataaga aatatgccag cggcaacaaa gataacattg tgcgtaataa cgtatctgtc 3540
tacaacaac tggctgtgaa gaataaagag taccgtctg gcaccaacgc ttccaggcg 3600
gggtttgaga aaattctgag cgcgttggag atccctgat tccgtaatct gagcaagtc 3660
gtggttatga agagcaagaa cgaccagggt atcactaaca agtcaagat gaacctgcaa 3720
gacaanaatg gtaacgacat cggctttatt gggttccacc agtcaacaa tattgtcaa 3780
ctggtagcga gcaatfggta caatcgtcag atfgagcga gcagccgtac tttggcgtg 3840
agctgggagt ttatccgggt cgaatgatgt tggggcgaac gtccgctg 3888

```

```

<210> 7
<211> 820
<212> PRT
<213> 人工序列

<220>
<223> 工程化的BoNT/E轻链

<400> 7
Met Lys Ile Glu Glu Gly Lys Leu Val Ile Trp Ile Asn Gly Asp Lys
1 5 10 15
Gly Tyr Asn Gly Leu Ala Glu Val Gly Lys Lys Phe Glu Lys Asp Thr
20 25 30
Gly Ile Lys Val Thr Val Glu His Pro Asp Lys Leu Glu Glu Lys Phe
35 40 45
Pro Gln Val Ala Ala Thr Gly Asp Gly Pro Asp Ile Ile Phe Trp Ala
50 55 60

```

His Asp Arg Phe Gly Gly Tyr Ala Gln Ser Gly Leu Leu Ala Glu Ile
 65 70 75 80
 Thr Pro Asp Lys Ala Phe Gln Asp Lys Leu Tyr Pro Phe Thr Trp Asp
 85 90 95
 Ala Val Arg Tyr Asn Gly Lys Leu Ile Ala Tyr Pro Ile Ala Val Glu
 100 105 110
 Ala Leu Ser Leu Ile Tyr Asn Lys Asp Leu Leu Pro Asn Pro Pro Lys
 115 120 125
 Thr Trp Glu Glu Ile Pro Ala Leu Asp Lys Glu Leu Lys Ala Lys Gly
 130 135 140
 Lys Ser Ala Leu Met Phe Asn Leu Gln Glu Pro Tyr Phe Thr Trp Pro
 145 150 155 160
 Leu Ile Ala Ala Asp Gly Gly Tyr Ala Phe Lys Tyr Glu Asn Gly Lys
 165 170 175
 Tyr Asp Ile Lys Asp Val Gly Val Asp Asn Ala Gly Ala Lys Ala Gly
 180 185 190
 Leu Thr Phe Leu Val Asp Leu Ile Lys Asn Lys His Met Asn Ala Asp
 195 200 205
 Thr Asp Tyr Ser Ile Ala Glu Ala Ala Phe Asn Lys Gly Glu Thr Ala
 210 215 220
 Met Thr Ile Asn Gly Pro Trp Ala Trp Ser Asn Ile Asp Thr Ser Lys
 225 230 235 240
 Val Asn Tyr Gly Val Thr Val Leu Pro Thr Phe Lys Gly Gln Pro Ser
 245 250 255
 Lys Pro Phe Val Gly Val Leu Ser Ala Gly Ile Asn Ala Ala Ser Pro
 260 265 270
 Asn Lys Glu Leu Ala Lys Glu Phe Leu Glu Asn Tyr Leu Leu Thr Asp
 275 280 285
 Glu Gly Leu Glu Ala Val Asn Lys Asp Lys Pro Leu Gly Ala Val Ala
 290 295 300
 Leu Lys Ser Tyr Glu Glu Glu Leu Ala Lys Asp Pro Arg Ile Ala Ala
 305 310 315 320
 Thr Met Glu Asn Ala Gln Lys Gly Glu Ile Met Pro Asn Ile Pro Gln
 325 330 335
 Met Ser Ala Phe Trp Tyr Ala Val Arg Thr Ala Val Ile Asn Ala Ala
 340 345 350
 Ser Gly Arg Gln Thr Val Asp Glu Ala Leu Lys Asp Ala Gln Thr Asn
 355 360 365
 Ser Ser Ser Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn Leu Gly Ile
 370 375 380
 Glu Gly Arg Ile Ser Glu Phe Gly Ser Met Pro Lys Ile Asn Ser Phe
 385 390 395 400
 Asn Tyr Asn Asp Pro Val Asn Asp Arg Thr Ile Leu Tyr Ile Lys Pro
 405 410 415
 Gly Gly Cys Gln Glu Phe Tyr Lys Ser Phe Asn Ile Met Lys Asn Ile
 420 425 430
 Trp Ile Ile Pro Glu Arg Asn Val Ile Gly Thr Thr Pro Gln Asp Phe
 435 440 445
 His Pro Pro Thr Ser Leu Lys Asn Gly Asp Ser Ser Tyr Tyr Asp Pro
 450 455 460
 Asn Tyr Leu Gln Ser Asp Glu Glu Lys Asp Arg Phe Leu Lys Ile Val
 465 470 475 480
 Thr Lys Ile Phe Asn Arg Ile Asn Asn Asn Leu Ser Gly Gly Ile Leu
 485 490 495

[0015]

Leu Glu Glu Leu Ser Lys Ala Asn Pro Tyr Leu Gly Asn Asp Asn Thr
500 505 510

Pro Asp Asn Lys Phe His Ile Gly Asp Ala Ser Ala Val Glu Ile Lys
515 520 525

Phe Ser Lys Gly Ser Gln His Ile Leu Leu Pro Asn Val Ile Ile Met
530 535 540

Gly Ala Glu Pro Asp Leu Phe Glu Thr Asn Ser Ser Asn Ile Ser Leu
545 550 555 560

Arg Asn Asn Tyr Met Pro Ser Asn His Gly Phe Gly Ser Ile Ala Ile
565 570 575

Val Thr Phe Ser Pro Glu Tyr Ser Phe Arg Phe Asn Asp Asn Ser Ile
580 585 590

Asn Glu Phe Ile Gln Asp Pro Ala Leu Thr Leu Met His Glu Leu Ile
595 600 605

His Ser Leu His Gly Leu Tyr Gly Ala Lys Gly Ile Thr Thr Thr Cys
610 615 620

Ile Ile Thr Gln Gln Lys Asn Pro Leu Ile Thr Asn Arg Lys Gly Ile
625 630 635 640

Asn Ile Glu Glu Phe Leu Thr Phe Gly Gly Asn Asp Leu Asn Ile Ile
645 650 655

Thr Val Ala Gln Tyr Asn Asp Ile Tyr Thr Asn Leu Leu Asn Asp Tyr
660 665 670

Arg Lys Ile Ala Ser Lys Leu Ser Lys Val Gln Val Ser Asn Pro Gln
675 680 685

Leu Asn Pro Tyr Lys Asp Ile Phe Gln Glu Lys Tyr Gly Leu Asp Lys
690 695 700

Asp Ala Ser Gly Ile Tyr Ser Val Asn Ile Asn Lys Phe Asp Asp Ile
705 710 715 720

Leu Lys Lys Leu Tyr Ser Phe Thr Glu Phe Asp Leu Ala Thr Lys Phe
725 730 735

Gln Val Lys Cys Arg Glu Thr Tyr Ile Gly Gln Tyr Lys Tyr Phe Lys
740 745 750

Leu Ser Asn Leu Leu Asn Asp Ser Ile Tyr Asn Ile Ser Glu Gly Tyr
755 760 765

Asn Ile Asn Asn Leu Lys Val Asn Phe Arg Gly Gln Asn Ala Asn Leu
770 775 780

Asn Pro Arg Ile Ile Lys Pro Ile Thr Gly Arg Gly Leu Val Lys Lys
785 790 795 800

Ile Ile Arg Phe Ala Val Asp Lys Leu Ala Ala Ala Leu Glu His His
805 810 815

His His His His
820

[0016]

<210> 8
<211> 2460
<212> DNA
<213> 人工序列

<220>
<223> 工程化的BoNT/E轻链

<400> 8
atganaatcg aagaangtaa actggtaate tggattaacg gcgataaagg ctataacggt 60
ctcgcigaag tccgtangaa attcgagaaa gataccggaa ttaaagtcac cgttgagcat 120
cgggataaac tggaaagaaa attcccacag gttgcggcaa cggcgatgg cccfgacatt 180
atctctggg cacacgaccg ctttggggc tacgctcaat cggcctggt ggctgaanac 240
accccgaca aagcgttcca ggacaagctg tatccgitta cctgggatgc cgtacgttac 300
aacggcaagc tgattgctta cccgatcgt gttgaagcgt taccgctgat ttatacaaa 360

```

gatctgctgc cgaaccgcc aaaaacctgg gaagagatcc cggcgtgga taaagaactg 420
aaagcgaaag gtaagagcgc gctgatgtc aacctgcaag aaccgtactt cacctggccg 480
ctgattgctg ctgacggggg ttatgcttc aagtagaaa acggcaagta cgaccattaa 540
gacgtgggcg tggataacgc tggcgcgaaa gcgggtctga ccttccctgt tgacctgatt 600
aaaaacaaac acatgaaatg agacaccgat tactccatcg cagaagctgc ctttaataaa 660
ggcgaacacg cgaigaccat caacggcccg tgggcatggt ccaacatcga caccagcaaa 720
gtgaattatg gtgtaacggt actcccgacc ttcaagggtc aaccatcaa accgttcgtt 780
ggcgtgctga gcgagggat taacgccgcc agtccgaaca aagagctggc aaaagagttc 840
ctcgaaaact atctgctgac tgatgaaggt ctggaagcgg ttaataaaga caaacccctg 900
ggtgccgtag cgtgaaatc ttacgaggaa gactggcga agatccacg tatggccgcc 960
actatggaaa acgcccagaa aggtgaaatc atgccgaaca tcccgcagat gtccgcttcc 1020
tggfatgccg tgcgtactgc ggtgatcaac gccgccagcg gtcgtcagac tgcctgatga 1080
gccctgaaag acgcgcaagc taatcagagc tcgaacaaca scaacaafaa caataacaac 1140
aacctgggga tcgaggggag gatttcagaa ttccgatcca tgcacaaaat caacagcttt 1200
aattacaatg accctgtaaa cgatcgtacc atcctataca taangccggg tgggtgtcaa 1260
gagttctaca aatctttcaa tattatgag aatataatga ttatacctga gcgtaacgtt 1320
attggtacga caccgcaaga tttcatcca cctactctgt tgaagaacgg tgactcttcc 1380
tattacgacc ccaattatct ccagtcggat gaagagaagg acagattcct taaaatagta 1440
accaaatct ttaacagat taatacaat ctatccggag gtattttgct tgaagagctt 1500
agtaaaacta atccttacct aggtaacgat aatcacccag scaacaagtt tcatataggg 1560
gatgcatccg ccgtggaaat caaatttagc aagggatcac agcatattct ctggcccaac 1620
gttattataa tggggcgaga accagattta ttgaaacaa atcagatga tattagcctg 1680
agaaataact atatgccctc aaaccatggg ttccgtagca tagcgtcgt tactttttct 1740
cccgaataca gttttcgtt caatgataat agtataaag agtttatcca agaccccga 1800
ctcacgctta tgcacgaact catacactct ttacacggcc tgtatggcgc taaggggata 1860
accactaegt gtatcattac tcagcaaaag aacccttga taacgaacag gaagggcatt 1920
aacatcgagg aatttcttac atttggaggc aacgatctga acattatac tgtcgccag 1980
tacaatgaca tctatacca ctfactaaat gattatgaa aaatcgctc taagtattcc 2040
aaggttcaag tcicaaaacc tcactgaat ccgtataagg acatattcca agagaaata 2100
ggattagaca aagacgcgtc aggaatctat tcggttaaca ttaacaatc cgacgatatt 2160
ttgagaaac ttacagctt cacggagtic gacttggcca ccaaatcca ggtcaaatgc 2220
cgagagacat acatcggaca gtataagtat ttcaagctgt cgaatctcct gaatgatcc 2280
atafacaaca ttatggagg ttacaatata aataacctaa aggtgaattt ccgagggcaa 2340
aacgccaaac taaatccgcg catcattasa cccatcacag gacgggggtt agtgaagaaa 2400
ataatccggt ttgcggtcga caagcttgcg gccgcactcg agcaccacca ccaccaccac 2460

```

[0017]

```

<210> 9
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> VBCCu12识别基序

<400> 9
Ala Leu Ala Pro Tyr Ile Pro
1 5

```

```

<210> 10
<211> 12
<212> PRT
<213> 人工序列

```

```

<220>
<223> MDM2识别基序

```

```

<220>
<221> 其它特征(misc.feature)
<222> (1)..(2)
<223> Xaa可以是任何天然存在的氨基酸

```

```

<220>
<221> 其它特征
<222> (4)..(6)
<223> Xaa可以是任何天然存在的氨基酸

```

<220>

<221> 其它特征
 <222> (8)..(9)
 <223> Xaa可以是任何天然存在的氨基酸

 <220>
 <221> 其它特征
 <222> (11)..(12)
 <223> Xaa可以是任何天然存在的氨基酸

 <400> 10
 Xaa Xaa Phe Xaa Xaa Xaa Trp Xaa Xaa Leu Xaa Xaa
 1 5 10

 <210> 11
 <211> 9
 <212> PRT
 <213> 人工序列

 <220>
 <223> MNM2识别序列

 <400> 11
 Arg Phe Met Asp Tyr Trp Glu Gly Leu
 1 5

 <210> 12
 <211> 9
 <212> PRT
 <213> 人工序列

 <220>
 <223> MNM2识别序列

 <220>
 <221> 其它特征
 <222> (2)..(4)
 <223> Xaa可以是任何天然存在的氨基酸

 <220>
 <221> 其它特征
 <222> (7)..(8)
 <223> Xaa可以是任何天然存在的氨基酸

 <400> 12
 Phe Xaa Xaa Xaa Leu Trp Xaa Xaa Leu
 1 5

 <210> 13
 <211> 14
 <212> PRT
 <213> 人工序列

 <220>
 <223> Smurf2识别序列

 <400> 13
 Glu Leu Glu Ser Pro Pro Pro Pro Tyr Ser Arg Tyr Pro Met
 1 5 10

 <210> 14
 <211> 7
 <212> PRT
 <213> 人工序列

 <220>
 <223> RN181识别序列

 <400> 14
 Lys Val Gly Phe Phe Lys Arg
 1 5

 <210> 15
 <211> 10
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3alpha识别序列

 <400> 15
 Leu Leu Val Arg Gly Arg Thr Leu Val Val
 1 5 10

 <210> 16
 <211> 10
 <212> PRT
 <213> 人工序列

 <220>
 <223> SCF识别序列

 <400> 16

[0018]

[0019]

```

Asp Arg His Asp Ser Gly Leu Asp Ser Met
1           5           10

<210> 17
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> Siah识别序列

<220>
<221> 其它特征
<222> (2)..(2)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 其它特征
<222> (4)..(4)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 其它特征
<222> (6)..(6)
<223> Xaa可以是任何天然存在的氨基酸

<400> 17
Pro Xaa Ala Xaa Val Xaa Pro
1           5

<210> 18
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> Itch识别序列

<220>
<221> 其它特征
<222> (3)..(3)
<223> Xaa可以是任何天然存在的氨基酸

<220>
<221> 其它特征
<222> (5)..(6)
<223> Xaa可以是任何天然存在的氨基酸

<400> 18
Pro Pro Xaa Tyr Xaa Xaa Met
1           5

<210> 19
<211> 4
<212> PRT
<213> 人工序列

<220>
<223> Nedd4-2识别序列

<220>
<221> 其它特征
<222> (3)..(3)
<223> Xaa可以是任何天然存在的氨基酸

<400> 19
Pro Pro Xaa Tyr
1

<210> 20
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 20
Glu Thr Phe Ser Asp Leu Trp Lys Leu Leu Pro Glu
1           5           10

<210> 21
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 21
Thr Ser Phe Ala Glu Tyr Trp Asn Leu Leu Ser Pro
1           5           10
    
```

<210> 22
 <211> 12
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3连接酶识别基序

 <400> 22
 Leu Thr Phe Glu His Tyr Trp Ala Gln Leu Thr Ser
 1 5 10

<210> 23
 <211> 12
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3连接酶识别基序

 <400> 23
 Leu Thr Phe Glu His Trp Trp Ala Gln Leu Thr Ser
 1 5 10

<210> 24
 <211> 12
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3连接酶识别基序

 <400> 24
 Leu Thr Phe Glu His Ser Trp Ala Gln Leu Thr Ser
 1 5 10

<210> 25
 <211> 12
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3连接酶识别基序

 <400> 25
 Glu Thr Phe Glu His Asn Trp Ala Gln Leu Thr Ser
 1 5 10

[0020]

<210> 26
 <211> 12
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3连接酶识别基序

 <400> 26
 Leu Thr Phe Glu His Asn Trp Ala Gln Leu Thr Ser
 1 5 10

<210> 27
 <211> 12
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3连接酶识别基序

 <400> 27
 Leu Thr Phe Glu His Trp Trp Ala Ser Leu Thr Ser
 1 5 10

<210> 28
 <211> 12
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3连接酶识别基序

 <400> 28
 Leu Thr Phe Glu His Trp Trp Ser Ser Leu Thr Ser
 1 5 10

<210> 29
 <211> 12
 <212> PRT
 <213> 人工序列

 <220>
 <223> E3连接酶识别基序

[0021]

```

<400> 29
Leu Thr Phe Thr His Trp Trp Ala Gln Leu Thr Ser
1           5           10

<210> 30
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 30
Glu Thr Phe Glu His Trp Trp Ala Gln Leu Thr Ser
1           5           10

<210> 31
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 31
Leu Thr Phe Glu His Trp Trp Ser Gln Leu Thr Ser
1           5           10

<210> 32
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 32
Leu Thr Phe Glu His Trp Trp Ala Gln Leu Ser
1           5           10

<210> 33
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 33
Glu Thr Phe Glu His Trp Trp Ser Gln Leu Ser
1           5           10

<210> 34
<211> 9
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 34
Arg Phe Met Asp Tyr Trp Glu Gly Leu
1           5

<210> 35
<211> 12
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 35
Met Pro Arg Phe Met Asp Tyr Trp Glu Gly Leu Asn
1           5           10

<210> 36
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> E3连接酶识别基序

<400> 36
Ser Gln Glu Thr Phe Ser Asp Leu Trp Lys Leu Leu Pro Glu Asn
1           5           10           15

<210> 37

```

<211> 12
<212> PRT
<213> 人工序列

[0022]

<220>
<223> E3连接酶识别基序

<400> 37

Leu Thr Phe Glu His Asn Trp Ala Gln Leu Glu Asn
1 5 10

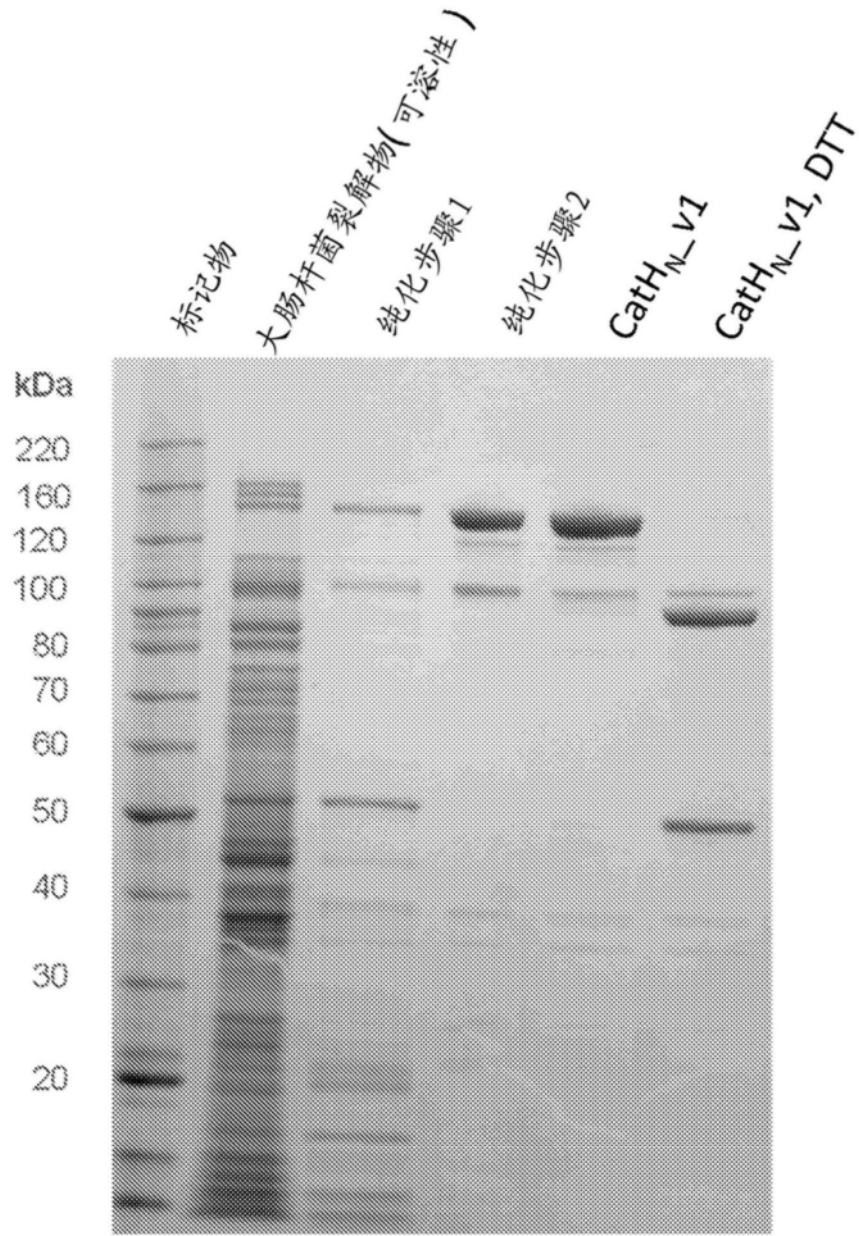


图1

图1A

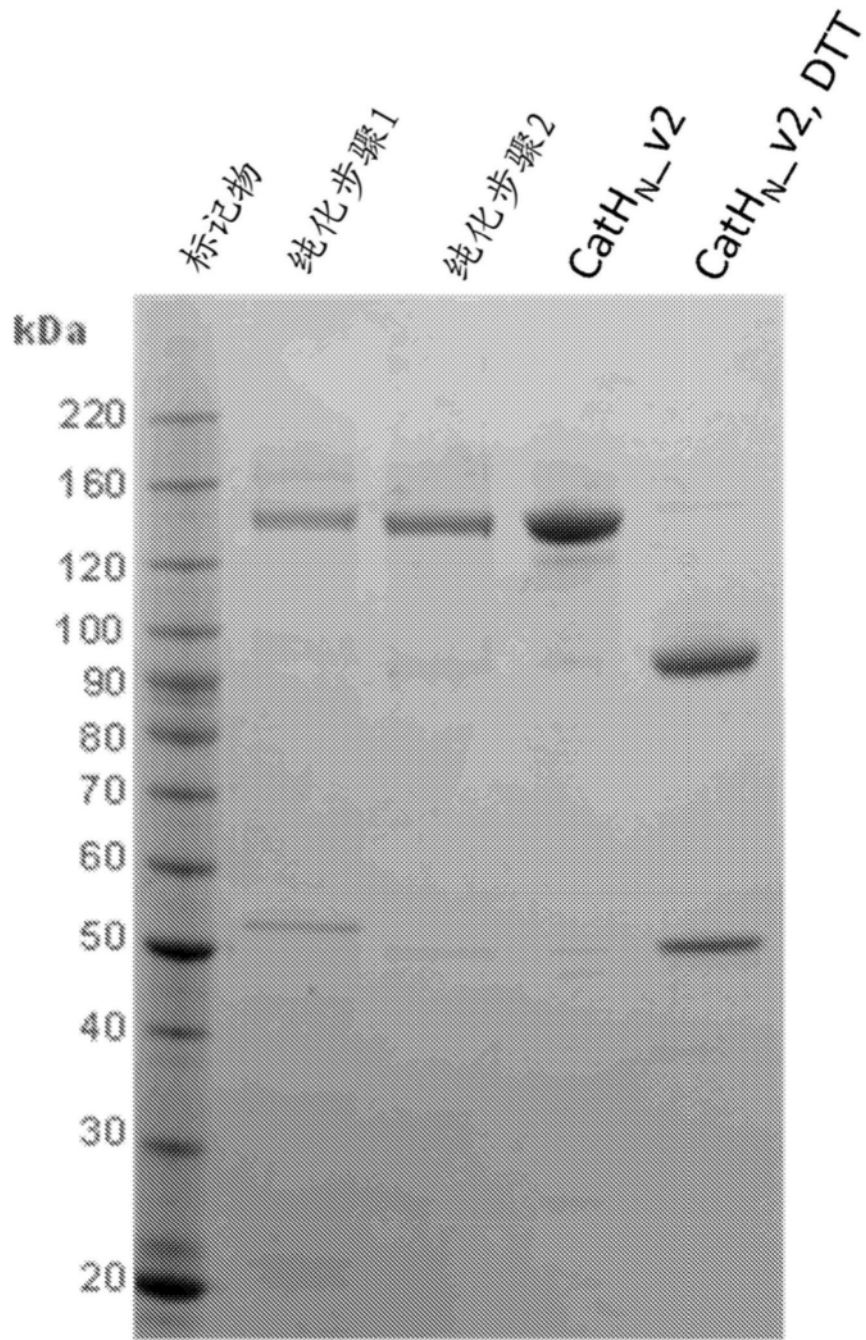


图1B

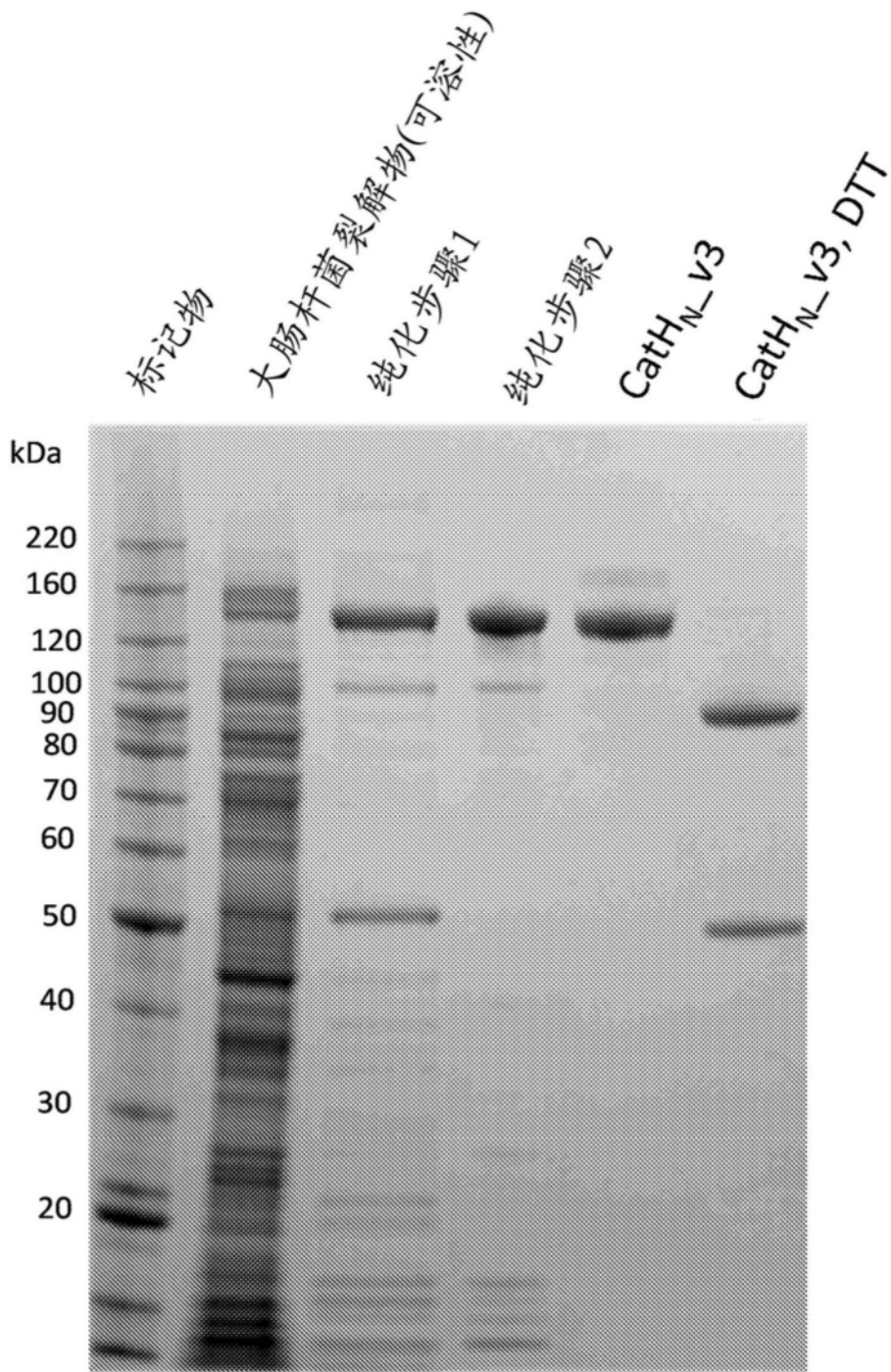
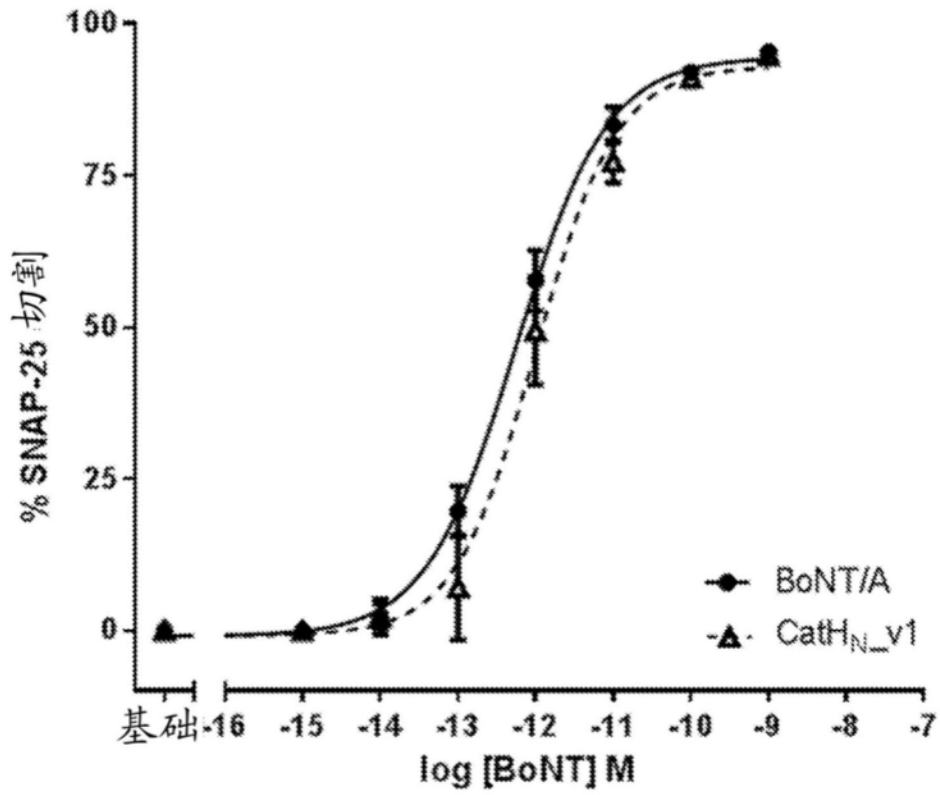


图1C



	pEC ₅₀
BoNT/A	12.26 ± 0.13
CatH _{N_v1}	12.00 ± 0.09

图2

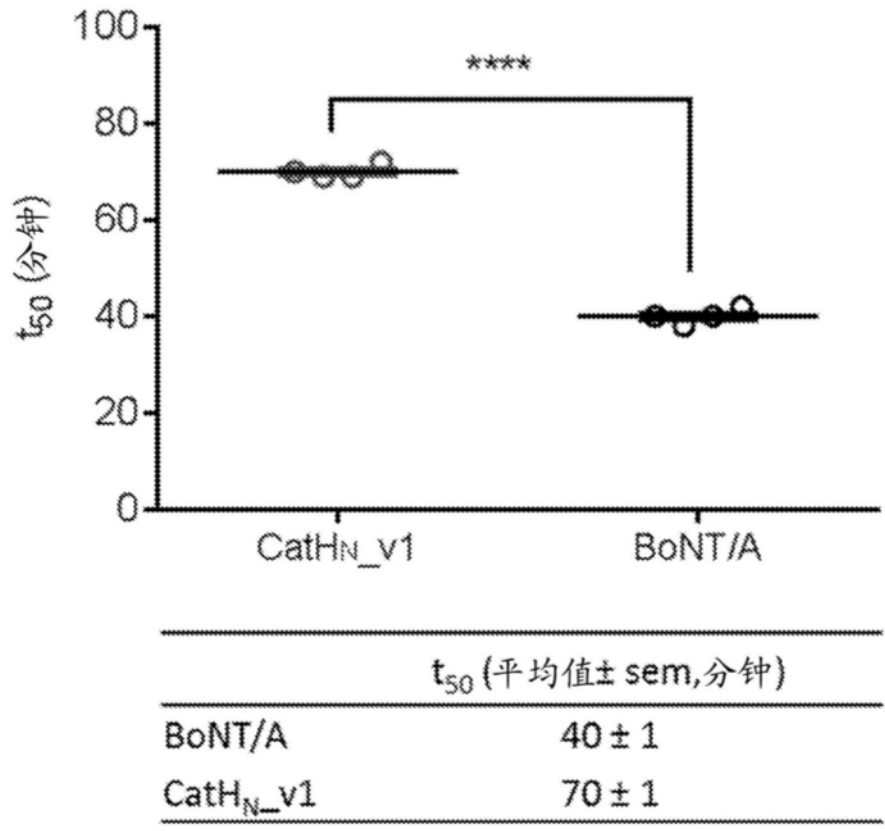


图3

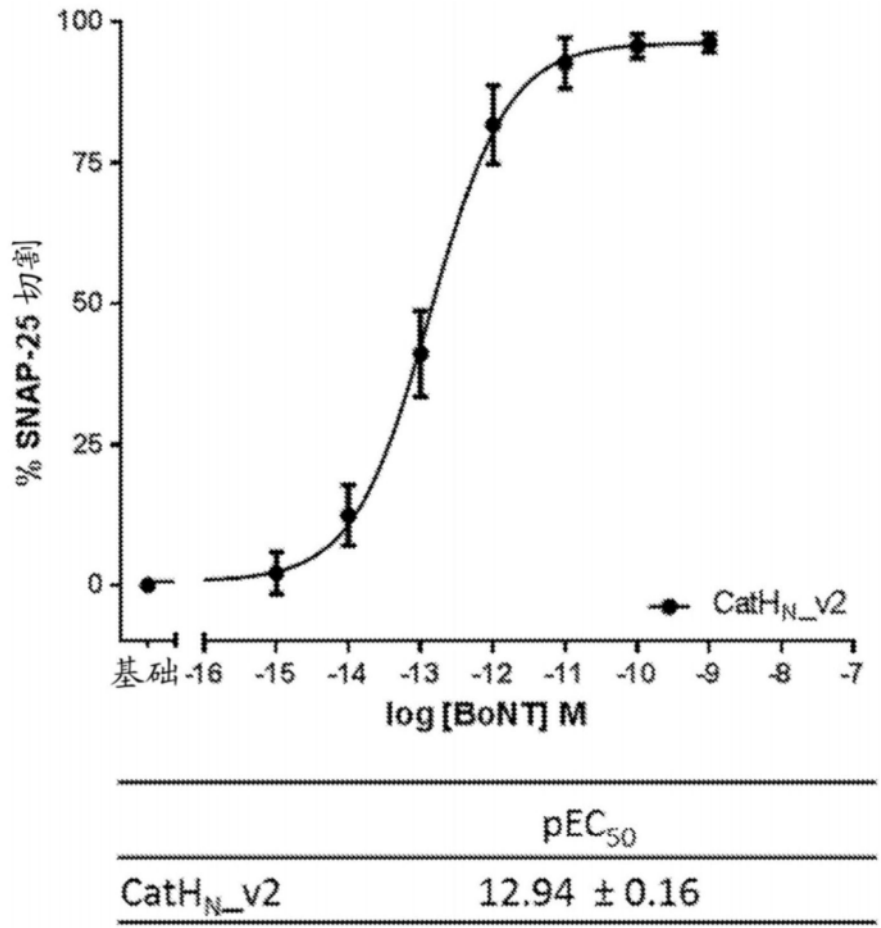


图4

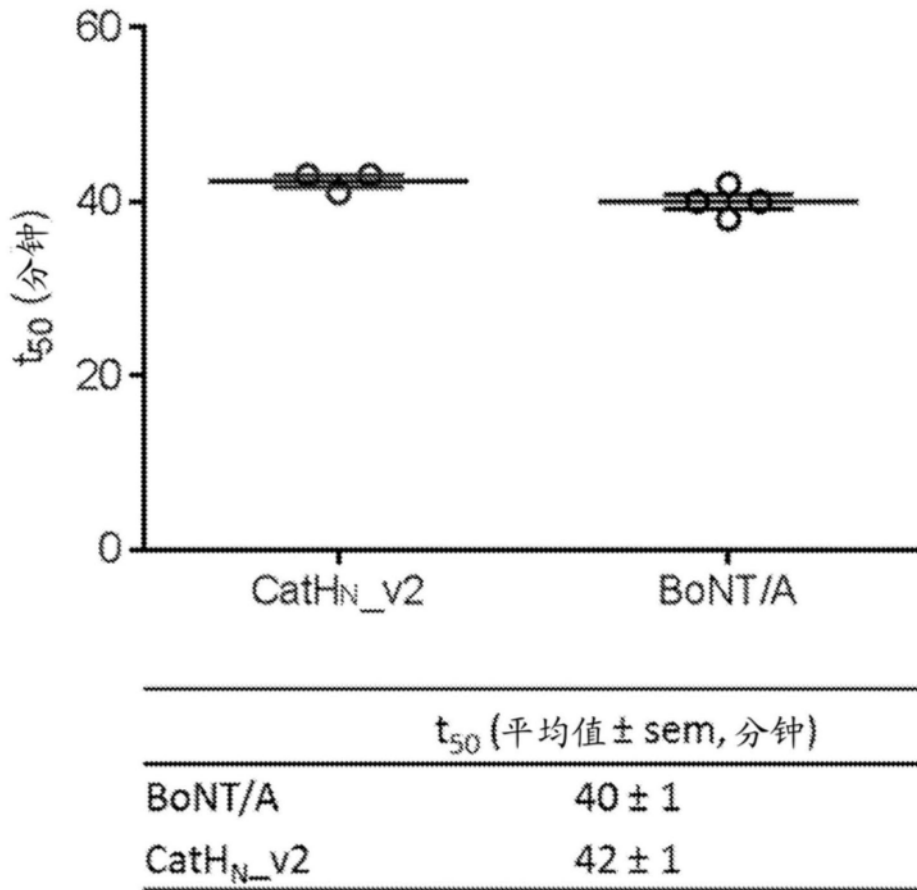
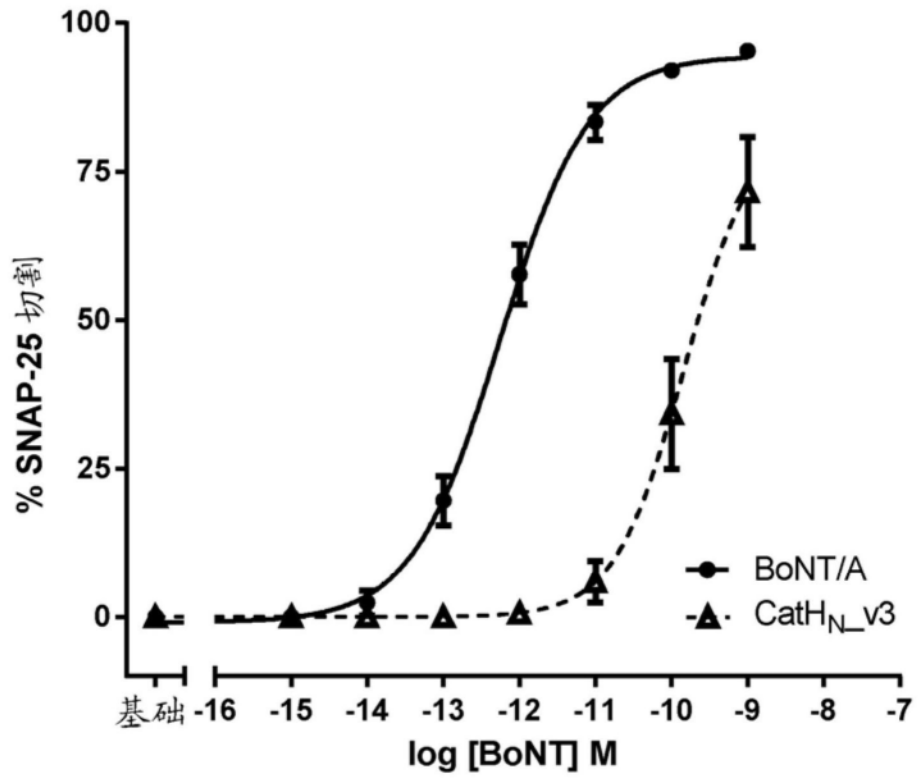


图5



	pEC ₅₀
BoNT/A	11.94 ± 0.06
CatH _N _v3	9.84 ± 0.04

图6

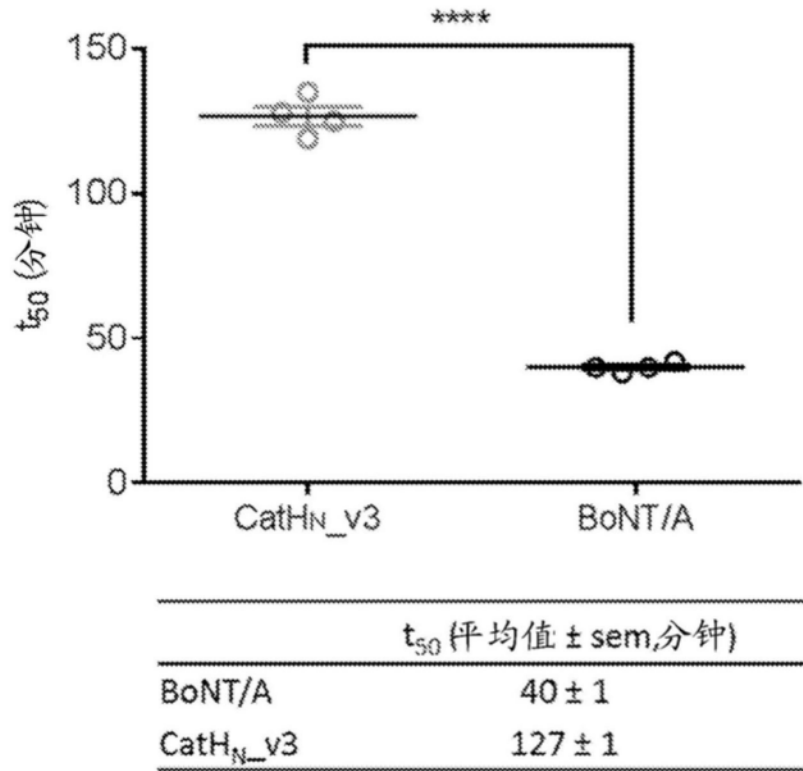
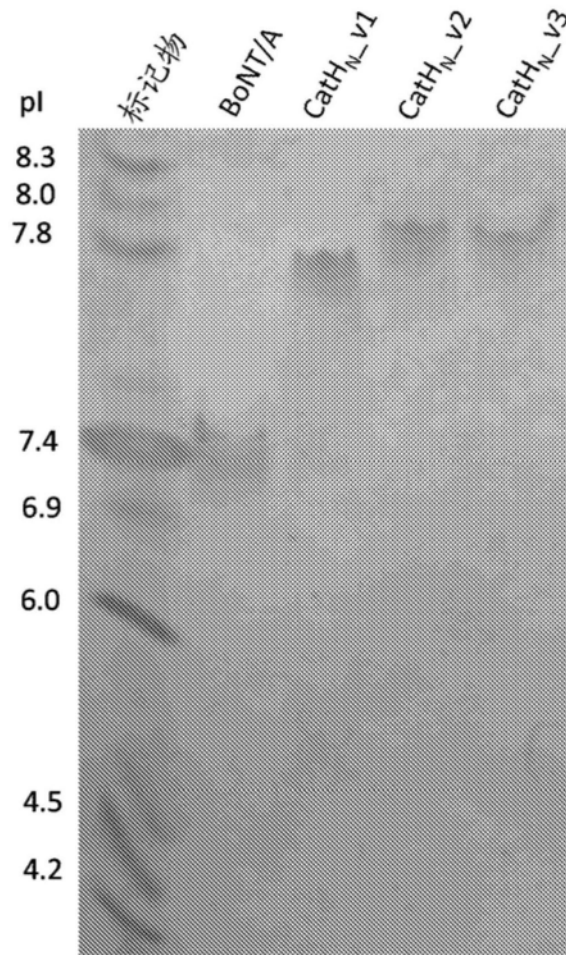


图7



泳道	构建体	计算的 pI	观察的 pI
1	BoNT/A	6.4	7.4
2	CatH _N _v1	7.4	7.8
3	CatH _N _v2	7.3	7.8-8.0
4	CatH _N _v3	7.1	7.8-8.0

图8

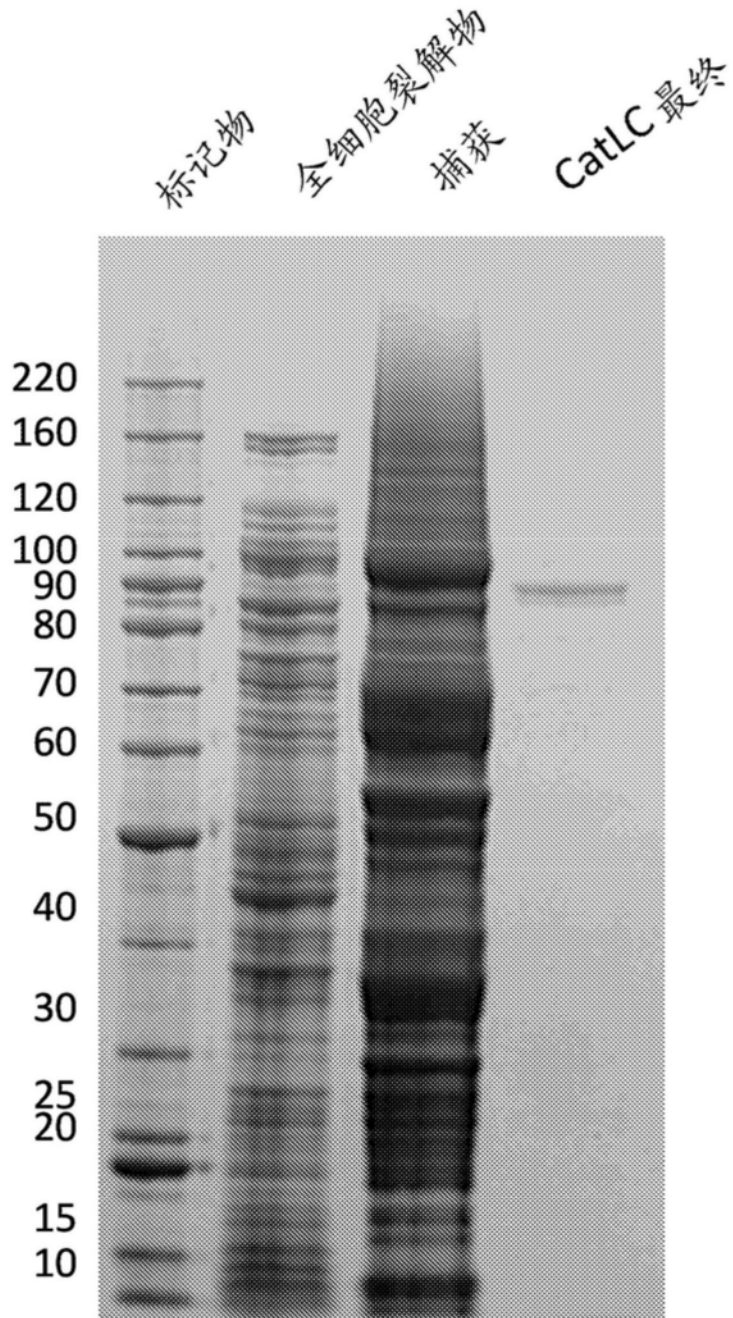


图9

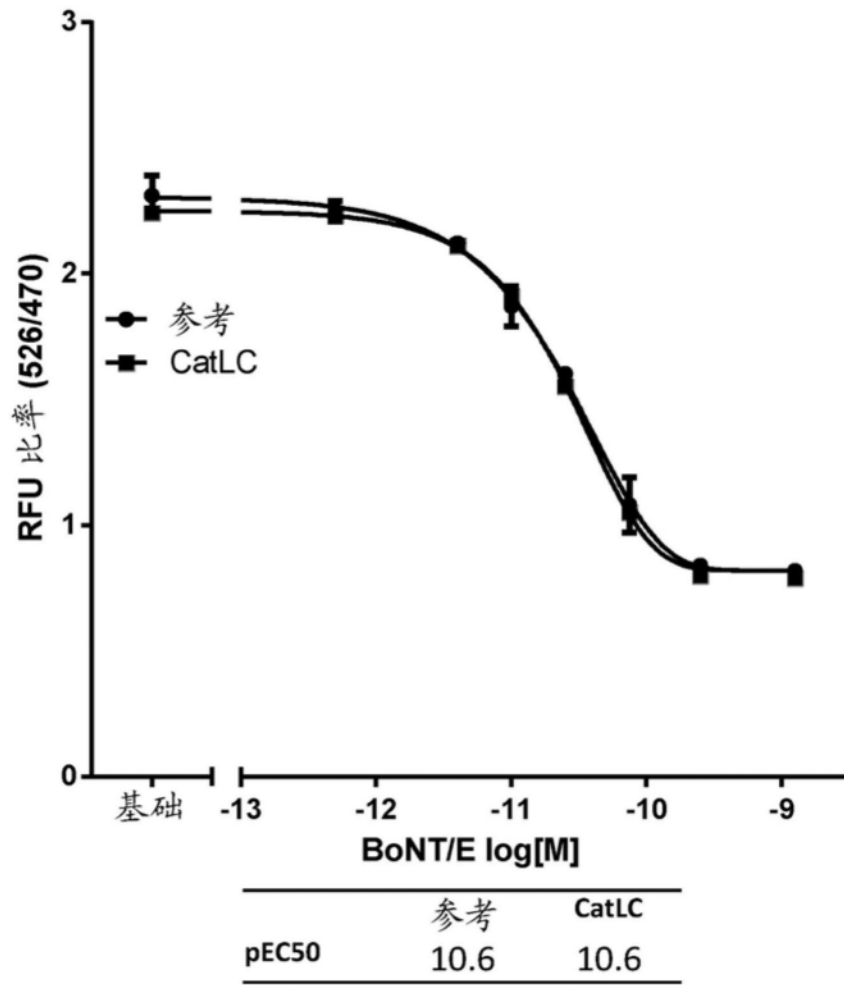


图10