



(12) 发明专利申请

(10) 申请公布号 CN 111819189 A

(43) 申请公布日 2020.10.23

(21) 申请号 201980015353.X

(22) 申请日 2019.02.21

(30) 优先权数据

1850213-8 2018.02.26 SE

(85) PCT国际申请进入国家阶段日

2020.08.25

(86) PCT国际申请的申请数据

PCT/EP2019/054310 2019.02.21

(87) PCT国际申请的公布数据

W02019/162376 EN 2019.08.29

(71) 申请人 托克斯沃科技公司

地址 瑞典耶尔费拉

(72) 发明人 帕尔·斯滕马克 杰弗里·马叙耶

(74) 专利代理机构 北京康信知识产权代理有限公司 11240

代理人 陈知宇

(51) Int.Cl.

G07K 14/33 (2006.01)

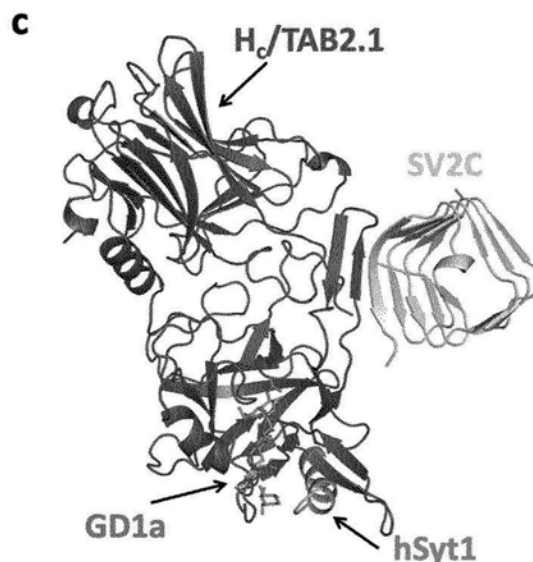
权利要求书3页 说明书27页  
序列表28页 附图19页

(54) 发明名称

肉毒神经毒素生物杂化体

(57) 摘要

本发明涉及适合于协同结合至突触结合蛋白(Syt)受体、突触相关囊泡蛋白2(SV2)受体和神经节糖苷(Gang)受体的新的肉毒神经毒素(BoNT)重链结合结构域(H<sub>c</sub>/TAB)以及包含所述新的H<sub>c</sub>/TAB的多肽、编码所述多肽的载体及其用途。



1. 具有N末端 (H<sub>CN</sub>) 和C末端 (H<sub>CC</sub>) 的肉毒神经毒素 (BoNT) 重链结合结构域 (H<sub>C</sub>/TAB), 其中所述H<sub>C</sub>/TAB包含:

- a) 突触结合蛋白 (Syt) 受体结合位点, 和
- b) 突触相关囊泡蛋白2 (SV2) 受体结合位点, 和
- c) 神经节糖苷 (Gang) 受体结合位点,

并且其中所述H<sub>C</sub>/TAB适合于协同结合至突触结合蛋白 (Syt) 受体、突触相关囊泡蛋白2 (SV2) 受体和神经节糖苷 (Gang) 受体。

2. 根据权利要求1所述的H<sub>C</sub>/TAB, 其中形成Gang受体结合位点的序列来源于任何Gang受体结合BoNT血清型和它们的亚型。

3. 根据权利要求1或2中任一项所述的H<sub>C</sub>/TAB, 其中形成Syt受体结合位点的序列来源于任何Syt受体结合BoNT血清型和它们的亚型。

4. 根据权利要求1至3中任一项所述的H<sub>C</sub>/TAB, 其中所述形成SV2受体结合的位点的序列来源于任何SV2受体结合BoNT血清型和它们的亚型。

5. 根据权利要求1至4中任一项所述的H<sub>C</sub>/TAB, 其中所述H<sub>CN</sub>序列来源于任何SV2受体结合BoNT血清型和它们的亚型。

6. 根据权利要求1至5中任一项所述的H<sub>C</sub>/TAB, 其特征在于所述H<sub>CC</sub>结构域可互换地由来自BoNT血清A型 (BoNT/A) 和BoNT血清B型 (BoNT/B) 的序列组成。

7. 根据权利要求1至6中任一项所述的H<sub>C</sub>/TAB, 其特征在于根据序列A<sub>1</sub>B<sub>1</sub>A<sub>2</sub>B<sub>2</sub>A<sub>3</sub>组成所述H<sub>CC</sub>端, 其中A表示来自BoNT/A的序列, B表示来自BoNT/B的序列。

8. 根据权利要求7所述的H<sub>C</sub>/TAB, 其中B<sub>1</sub>、A<sub>2</sub>和B<sub>2</sub>的序列包含突变和/或缺失以对于整个H<sub>C</sub>/TAB产生稳定的分子内界面。

9. 根据权利要求1至8中任一项所述的H<sub>C</sub>/TAB, 其中所述形成Gang受体结合位点的序列来源于BoNT/B。

10. 根据权利要求7至8中任一项所述的H<sub>C</sub>/TAB, 其中所述形成Gang受体结合位点的序列位于B<sub>2</sub>中。

11. 根据权利要求1至10中任一项所述的H<sub>C</sub>/TAB, 其中所述形成Syt受体结合位点的序列来源于BoNT B、DC或G。

12. 根据权利要求7至8或10中任一项所述的H<sub>C</sub>/TAB, 其中所述形成Syt受体结合位点的序列位于B<sub>1</sub>和B<sub>2</sub>中。

13. 根据权利要求1至12中任一项所述的H<sub>C</sub>/TAB, 其中所述H<sub>CN</sub>序列来源于BoNT/A。

14. 根据权利要求7至8、10或12中任一项所述的H<sub>C</sub>/TAB, 其中所述形成SV2受体结合位点的序列位于H<sub>CN</sub>中和H<sub>CC</sub>中的A<sub>1</sub>和A<sub>3</sub>中。

15. 根据权利要求1至14中任一项所述的H<sub>C</sub>/TAB, 其具有与SEQ. ID. No. 1所示的序列具有至少60%、65%、70%、75%、80%、85%、90%或95%的同一性的氨基酸序列。

16. 一种包含根据权利要求1至15中任一项所述的H<sub>C</sub>/TAB的多肽, 直接或通过接头结合至任何一个或多个其他蛋白、多肽、氨基酸序列或荧光探针。

17. 根据权利要求16所述的多肽, 其中所述多肽是BoNT多肽 (BoNT/TAB), 其特征在于除所述H<sub>C</sub>/TAB之外, 所述BoNT/TAB在所述多肽序列中包含重链移位结构域 (H<sub>N</sub>)、轻链 (LC) 和位于所述LC和H<sub>N</sub>之间的蛋白酶位点, 其中所述H<sub>N</sub>和所述LC分别且彼此独立地来源于任何BoNT

血清A、B、C、D、DC、E、En、F、G或X型和它们的亚型。

18. 根据权利要求17所述的多肽,还包含直接或通过接头与之连接的任何其他蛋白、多肽、氨基酸序列或荧光探针。

19. 根据权利要求17或18中任一项所述的多肽,其中所述蛋白酶位点是外切蛋白酶位点。

20. 根据权利要求16至19中任一项所述的多肽,其具有与SEQ. ID. No. 3、SEQ. ID. No. 5、SEQ. ID. No. 6、SEQ. ID. No. 8、SEQ. ID. No. 10或SEQ. ID. No. 12中任一项所示的序列具有至少60%、65%、70%、75%、80%、85%、90%或95%的同一性的氨基酸序列。

21. 一种载体,包含编码根据权利要求1至15中任一项所述的Hc/TAB或根据权利要求16至20中任一项所述的多肽的核酸序列。

22. 根据权利要求21所述的载体,其包含与SEQ. ID. No. 4、SEQ. ID. No. 7、SEQ. ID. No. 9、SEQ. ID. No. 11或SEQ. ID. No. 13中任一项所示的序列具有至少60%、65%、70%、75%、80%、85%、90%或95%的同一性的核酸序列。

23. 根据权利要求1至15中任一项所述的Hc/TAB或者根据权利要求16至20中任一项所述的多肽,用于在治疗方法或在美容方法中的使用。

24. 根据权利要求23所述的用于使用的Hc/TAB或多肽,其中所述治疗方法或美容方法是对无力和/或僵化肌肉的治疗。

25. 根据权利要求23或24中任一项所述的用于使用的Hc/TAB或多肽,其中所述治疗方法是对选自下列的病症的治疗和/或预防,其包括神经肌肉失调和肌肉痉挛病症。

26. 用于根据权利要求23至25中任一项所述的使用的Hc/TAB或多肽,其中所述病症选自:痉挛性发声障碍、痉挛性斜颈、喉肌张力障碍、口下颌发音障碍、舌肌张力障碍、颈肌张力障碍、局灶性手肌张力障碍、眼睑痉挛、斜视、半面痉挛、眼睑障碍、大脑性麻痹、局灶性痉挛及其他语言障碍、痉挛性结肠炎、神经原性膀胱、肛门痉挛、肢体痉挛、抽搐、震颤、磨牙症、肛裂、弛缓不能、吞咽困难及其他肌张力障碍以及特征为肌群不自主运动的其他障碍、流泪、多汗、过度流涎、过度胃肠分泌、分泌紊乱、肌肉痉挛引起的疼痛、头痛、运动损伤和抑郁症。

27. 用于在药理学测试中使用的根据权利要求16所述的多肽,其研究所述蛋白、多肽、氨基酸序列或荧光探针在突触过程中的作用。

28. 根据权利要求1至15中任一项所述的Hc/TAB,用于作为将与之结合的任何蛋白、多肽、氨基酸序列或荧光探针有效转运至神经元表面的媒介物使用。

29. 根据权利要求17至20中任一项所述的BoNT/TAB作为使用毒素移位系统将任何蛋白、多肽、氨基酸序列或荧光探针有效转运至神经元胞液的媒介物的使用。

30. 一种药物或美容组合物,其包含根据权利要求1至15中任一项所述的Hc/TAB或根据权利要求16至20中任一项所述的多肽。

31. 一种成套试剂盒,其包含根据权利要求30所述的组合物和用于所述组合物的治疗性施用的说明书。

32. 一种治疗与不希望的神元活动有关的病况的方法,所述方法包括向受试者施用治疗有效量的根据权利要求1至15中任一项所述的Hc/TAB或根据权利要求16至20中任一项所述的多肽或根据权利要求30所述的药物组合物,由此治疗所述病况,其中所述病况选自

由以下各项组成的组：痉挛性发声障碍、痉挛性斜颈、喉肌张力障碍、口下颌发音障碍、舌肌张力障碍、颈肌张力障碍、局灶性手肌张力障碍、眼睑痉挛、斜视、半面痉挛、眼睑障碍、大脑性麻痹、局灶性痉挛及其他语言障碍、痉挛性结肠炎、神经原性膀胱、肛门痉挛、肢体痉挛、抽搐、震颤、磨牙症、肛裂、弛缓不能、吞咽困难及其他肌张力障碍以及特征为肌群不自主运动的其他障碍、流泪、多汗、过度流涎、过度胃肠分泌、分泌紊乱、肌肉痉挛引起的疼痛、头痛、运动损伤和抑郁症以及皮肤学或美学/美容病况。

## 肉毒神经毒素生物杂化体

### 技术领域

[0001] 本发明涉及肉毒神经毒素多肽 (Botulinum neurotoxin) 并且具体地涉及嵌合肉毒神经毒素重链。

### 背景技术

[0002] 肉毒神经毒素 (BoNT) 是人类已知的最有效的蛋白毒素, 并且是罕见的麻痹病, 肉毒中毒的病因。该细菌毒素家族由八种血清型, BoNT/A-G 以及最近描述的 BoNT/X 组成 (Montal, 2010; Zhang 等人, 2017)。它们全部共有共同的结构, 并且作为 150kDa 的蛋白表达, 所述蛋白在翻译后切割成由通过单个二硫键连接至重链 (HC, 100kDa) 的轻链 (LC, 50kDa) 所组成的双链分子。HC 具有两个功能域, 即 N 末端移位结构域 (H<sub>N</sub>) 和 C 末端结合结构域 (H<sub>C</sub>), 而 LC 负责胞内催化活性。BoNT 首先通过特异性细胞表面受体识别胆碱能神经末端, 然后在囊泡中内吞。酸性核内体环境导致使得胞液内的 LC 移位 (也称为毒素移位) 的构象变化。然后, 游离的催化结构域, 锌-蛋白酶可以特异性靶向三种神经元 SNARE (可溶性 N-乙基马来酰亚胺敏感因子连接蛋白受体) 之一: BoNT/A、/C 和 /E 切割 SNAP-25; BoNT/B、/D、/F、/G 和 /X 靶向 VAMP (小突触泡蛋白); BoNT/C 切割突触融合蛋白 (Schiavo 等人, 2000; Zhang 等人, 2017)。这三种蛋白形成了介导突触囊泡与质膜融合的复合物 (Sudhof and Rothman, 2009)。任何 SNARE 的蛋白水解抑制胞吐作用并因此抑制神经递质的释放, 从而有效地导致肉毒中毒的弛缓性麻痹症状 (Rossetto 等人, 2014)。先前已描述了三种功能域的序列 (Lacy DB 等人, 1999)。催化结构域由氨基酸 1-437 组成, 移位结构域由氨基酸 448-872 组成并且结合结构域由氨基酸 873-1295 组成, 其在 Lacy DB 等人中称为 BoNT/A 序列。由于所有 BoNT 血清型和它们的亚型在很大程度上是同源的, 因此任何其他血清型或亚型中相应的域的位置将非常类似。

[0003] 这些毒素的高效力使它们在不断增大的范围的神经肌肉失调, 如斜视、颈肌张力障碍和眼睑痉挛以及涉及乙酰胆碱释放的其他病况, 如多汗的治疗中成为极其有用的治疗剂 (Chen, 2012)。BoNT/A 和 /B 是仅有的作为治疗剂批准并且可商购的血清型。通常认为 BoNT/A 在人中具有更高的效力, 并因此在大多数情况下是所选的血清型 (Bentivoglio 等人, 2015)。然而, 通过 BoNT 的治疗通常需要重复注射, 因为毒素的治疗作用仅是短暂的。据报道这导致在对 BoNT/A 产生免疫应答的小型患者亚组中出现了耐受性 (Lange 等人, 2009; Naumann 等人, 2013)。尽管 BoNT/B 代表了替代选择, 但其较低的效力意味着需要更高的剂量, 并因此代表了更大的免疫原性风险 (Dressier and Bigalke, 2005)。另外, BoNT/B 也与几种不利的结局有关, 如注射疼痛、短作用持续时间以及更常见的副作用 (Bentivoglio 等人, 2015)。主要的副作用经常还与治疗肌肉痉挛, 而不是美容应用有关。这是因为副作用基本上是由于毒素向身体其他区域的扩散所造成的, 并且毒素扩散的可能性与注射剂量直接相关。副作用从短暂的非严重性事件, 如上睑下垂和复视至危及生命的事件, 甚至死亡。

[0004] 已详细表征了 BoNT/A 和 /B 与神经元的结合, 并且所述结合基于双受体机制, 包括锚定在神经元膜上的突触囊泡蛋白和神经节糖苷。将 BoNT/A 的蛋白受体鉴别为 SV2 (Dong 等

人,2006;Mahrhold等人,2006)。更准确地,BoNT/A可以结合至几种人SV2同工型A、B和C,尽管毒素仅识别由SV2A和SV2B所形成的N-糖基化形式(Yao等人,2016)。BoNT/B的蛋白受体是突触结合蛋白(Syt)(Nishiki等人,1994,1996;Dong等人,2003),其中在人中相对于Syt2,优选Syt1(Strotmeier等人,2012)。对于所有的BoNT,神经节糖苷识别是中毒过程的第一步(Binz and Rummel,2009),并且其通过它们的序列中集中在保守基序H...SxWY...G中的共有结合机制介导。BoNT/A偏好地结合至GT1b和GD1a的末端N-乙酰半乳糖胺-半乳糖部分(Takamizawa等人,1986;Schengrund等人,1991),而对于BoNT/B的数据表明优选GD1b和GT1b的双唾液基基序。不同的血清型在它们的碳水化合物特异性和亲合力方面不同(Rummel,2013)。

[0005] 多种BoNT血清型的模块化布置和独特的性质使得毒素成为蛋白质工程的所选靶标。具体地,几项研究已显示有可能在血清型之间交换整个域(Masuyer等人,2014)并因此获得具有独特药物潜力的新的毒素。例如,已产生了由与移位有关的BoNT/B结合结构域和BoNT/A催化结构域所组成的几种分子(Rummel等人,2011;Wang等人,2012;Kutschenko等人,2017)。就效力和活性持续时间来说,这些所谓的嵌合毒素显示出有吸引力的药理学性质,与SV2相比,这与BoNT/B对突触结合蛋白的高亲合力以及该受体在神经元上的高表达有关(Takamori等人,2006;Wilhelm等人,2014)。

## 发明内容

[0006] 由于中和抗体的产生和毒素扩散两者均与注射剂量直接相关,因此降低毒素剂量,同时维持相同水平的毒素活性是非常期望的,这意味着必需提高各个毒素分子的效力。因此,本发明的目标是提高具有改善的持续时间和效力并且具有较低的从注射位点扩散的风险的BoNT多肽。本发明人已认识到了以上在工程设计的嵌合BoNT多肽中所提及的先前的工作中的关键问题。先前的工作均未考虑多肽的结构方面。

[0007] 使用基于结构的方法和当前有关BoNT/A和/B的受体结合机制的知识,本发明人工程设计了新的分子TriRecABTox(BoNT/TAB),其包含能够识别SV2C受体、突触结合蛋白受体和神经节糖苷受体的特异性工程设计的Hc域(Hc/TAB)。本发明人显示可以重组表达和纯化BoNT/TAB。使用X-射线晶体衍射,本发明人进一步证实BoNT/TAB可以同时结合至其三个受体。因此,BoNT/TAB应以高亲合力识别神经元细胞并且具有成为BoNT/A治疗的高效替代的潜力。

[0008] 因此,通过在第一个方面提供肉毒神经毒素(BoNT)重链结合结构域(Hc/TAB)实现了以上目标,其中Hc/TAB包含a)突触结合蛋白(Syt)受体结合位点和b)突触相关囊泡蛋白2(SV2)受体结合位点和c)神经节糖苷(Gang)受体结合位点,并且其中所述Hc/TAB适合于协同结合至突触结合蛋白(Syt)受体、突触相关囊泡蛋白2(SV2)受体和神经节糖苷(Gang)受体。

[0009] Hc/TAB具有N末端(Hc<sub>N</sub>)和C末端(Hc<sub>C</sub>)。根据一个实施方式,Hc<sub>C</sub>结构域可互换地由来自BoNT血清A型(BoNT/A)和BoNT血清B型(BoNT/B)的序列组成。

[0010] 根据其他实施方式,根据序列A<sub>1</sub>B<sub>1</sub>A<sub>2</sub>B<sub>2</sub>A<sub>3</sub>组成所述Hc<sub>C</sub>末端,其中A表示来自BoNT/A的序列并且B表示来自BoNT/B的序列。

[0011] 根据其他实施方式,B<sub>1</sub>、A<sub>2</sub>和B<sub>2</sub>的序列包含突变和/或缺失以对于整个Hc/TAB产生

稳定的分子内界面。

[0012] 根据其他实施方式,形成Gang受体结合位点的序列来源于任何Gang受体结合BoNT血清型和它们的亚型。

[0013] 根据其他实施方式,形成Gang受体结合位点的序列来源于BoNT/B。

[0014] 根据其他实施方式,形成Gang受体结合位点的序列位于B<sub>2</sub>中。

[0015] 根据其他实施方式,形成Syt受体结合位点的序列来源于任何Syt受体结合BoNT血清型和它们的亚型。

[0016] 根据其他实施方式,形成Syt受体结合位点的序列来源于BoNT B、DC或G。

[0017] 根据其他实施方式,形成Syt受体结合位点的序列位于B<sub>1</sub>和B<sub>2</sub>中。

[0018] 根据其他实施方式,H<sub>CN</sub>序列来源于任何SV2受体结合的BoNT血清型和它们的亚型。

[0019] 根据其他实施方式,H<sub>CN</sub>序列来源于BoNT/A。

[0020] 根据其他实施方式,形成SV2受体结合位点的序列位于H<sub>CN</sub>中并且位于H<sub>CC</sub>中的A<sub>1</sub>和A<sub>3</sub>中。

[0021] 根据其他实施方式,H<sub>C</sub>/TAB具有与SEQ. ID. No. 1、3、5、6、8、10或12中任一个所示的序列具有至少60%、65%、70%、75%、80%、85%、90%或95%的同一性的氨基酸序列。

[0022] 根据第二方面,提供了包含直接和通过接头与任何其他蛋白、多肽、氨基酸序列或荧光探针结合的根据第一方面和第一方面的任何实施方式的H<sub>C</sub>/TAB。

[0023] 根据第二方面的实施方式,所述多肽是BoNT多肽(BoNT/TAB),其特征在于所述BoNT/TAB除H<sub>C</sub>/TAB之外,在多肽序列中包含重链移位结构域(H<sub>N</sub>)、轻链(LC)和在LC和H<sub>N</sub>之间布置的蛋白酶位点,其中所述H<sub>N</sub>和LC分别且彼此独立地来源于BoNT血清A、B、C、D、DC、E、En、F、G或X型和它们的亚型中的任何,和BoNT样多肽。

[0024] 根据其他实施方式,所述多肽可以包含直接或通过接头与之连接的任何其他蛋白、多肽、氨基酸序列或荧光探针。

[0025] 根据其他实施方式,所述蛋白酶位点是外切蛋白酶位点。根据其他实施方式,所述外切蛋白酶位点是因子Xa位点。

[0026] 根据其他实施方式,根据第二方面的多肽具有与SEQ. ID. No. 1、3、5、6、8、10或12中任一个所示的序列具有至少60%、65%、70%、75%、80%、85%、90%或95%的同一性的氨基酸序列。

[0027] 根据第三方面,提供了包含编码根据第一方面和第一方面的任何实施方式的H<sub>C</sub>/TAB,或者根据第二方面和第二方面的任何实施方式的多肽的核酸序列的载体。

[0028] 根据第四方面,提供了根据第一方面和第一方面的任何实施方式的H<sub>C</sub>/TAB或者根据第二方面和第二方面的任何实施方式的多肽在治疗方法或在美容方法中的使用。

[0029] 根据第四方面的一个实施方式,所述治疗方法或美容方法是对无力和/或僵化肌肉(dampen and/or inactivate muscles)的治疗。

[0030] 根据第四方面的其他实施方式,所述治疗方法是对选自下列的病症的治疗和/或预防,其包括神经肌肉失调、涉及乙酰胆碱释放的病况和肌肉痉挛病症。

[0031] 根据其他实施方式,所述病症选自:痉挛性发声障碍、痉挛性斜颈、喉肌张力障碍、口下颌发音障碍(oramandibular dysphonia)、舌肌张力障碍、颈肌张力障碍、局灶性手肌张力障碍(focal hand dystonia)、眼睑痉挛、斜视、半面痉挛、眼睑障碍、大脑性麻痹、局灶

性痉挛及其他语言障碍、痉挛性结肠炎、神经原性膀胱、肛门痉挛、肢体痉挛、抽搐(tics)、震颤、磨牙症、肛裂、弛缓不能、吞咽困难及其他肌张力障碍以及特征为肌群不自主运动的其他障碍、流泪、多汗、过度流涎、过度胃肠分泌、分泌紊乱、肌肉痉挛引起的疼痛、头痛、运动损伤和抑郁症。

[0032] 根据其他实施方式,根据第一方面和第一方面的任何实施方式的Hc/TAB或者根据第二方面和第二方面的任何实施方式的多肽可以在药理学测试中使用以研究所述蛋白、多肽、氨基酸序列或荧光探针在突触过程中的作用。

[0033] 根据其他实施方式,根据第一方面和第一方面的任何实施方式的Hc/TAB或者根据第二方面和第二方面的任何实施方式的多肽可以用作将与之结合的任何蛋白、多肽、氨基酸序列或荧光探针有效转运至神经元表面的媒介物。

[0034] 根据其他实施方式,根据第一方面和第一方面的任何实施方式的Hc/TAB或者根据第二方面和第二方面的任何实施方式的多肽可以用作使用毒素移位系统将任何蛋白、多肽、氨基酸序列或荧光探针有效转运至神经元胞液中的媒介物。

[0035] 根据第五方面,提供了包含根据第一方面和第一方面的任何实施方式的Hc/TAB或者根据第二方面和第二方面的任何实施方式的多肽的药物或美容组合物。

[0036] 根据第五方面的一个实施方式,所述组合物还可以包含药物和/或美容可用的赋形剂、载体或其他添加剂。

[0037] 根据第六方面,提供了包含根据第五方面的组合物和所述组合物的治疗性施用的说明书的成套部件。

[0038] 根据第七方面,提供了治疗与不希望的神元活动有关的病况的方法,所述方法包括向受试者施用治疗有效量的根据第一方面和第一方面的任何实施方式的Hc/TAB,或者根据第二方面和第二方面的任何实施方式的多肽,或者根据第五方面所述的组合物,由此治疗所述病况,其中所述病况选自以下各项组成的组:痉挛性发声障碍、痉挛性斜颈、喉肌张力障碍、口下颌发音障碍(romandibular dysphonia)、舌肌张力障碍、颈肌张力障碍、局灶性手肌张力障碍(focal hand dystonia)、眼睑痉挛、斜视、半面痉挛、眼睑障碍、大脑性麻痹、局灶性痉挛及其他语言障碍、痉挛性结肠炎、神经原性膀胱、肛门痉挛、肢体痉挛、抽搐(tics)、震颤、磨牙症、肛裂、弛缓不能、吞咽困难及其他肌张力障碍以及特征为肌群不自主运动的其他障碍、流泪、多汗、过度流涎、过度胃肠分泌、分泌紊乱、肌肉痉挛引起的疼痛、头痛、运动损伤和抑郁症以及皮肤学或美学/美容病况。

## 附图说明

[0039] 图1:通过BoNT/A和/B的受体结合的结构信息。(a)与GT1b(PDB2VU)和与人糖基化的SV2C(PDB 5JLV)复合的BoNT/A的结合结构域的晶体结构的叠加。(b)与GD1a和大鼠突触结合蛋白2(PDB 4KBB)复合的BoNT/B的结合结构域的晶体结构。以缎带模式表示蛋白,并且以棍表示碳水化合物,(c)Hc/A(Uniprot P10845)和/B(Uniprot P10844)的序列对比,其中还提供了二级结构元件(通过ESPrict3.0作图;Robert and Gouet,2014)。对于每个域,突出显示了直接参与受体结合的区域,其中对于SV2受体,以Hc/A序列上方的线突出显示,并且对于Syt受体,以Hc/B序列下方的线突出显示;以条纹灰色线对神经节糖苷受体结合位点加下划线。

[0040] 图2:具有通过Hc/A和/B的受体结合的Hc/TAB的序列对比。通过ClustalO比对蛋白序列(Sievers等人,2011)。分别以黑色(白字)和浅灰色(黑字)突出显示了在Hc/TAB设计中使用的Hc/A和Hc/B的节段。以深灰色(短划线)显示了其中包括缺失的位置。

[0041] 图3:Hc/TAB的表征。(a)纯化的Hc/TAB的SDS-PAGE分析以及与Hc/A和Hc/B对照的比较,(b)使用聚组氨酸探针,与(a)中相同的样品的免疫印迹分析。“M”表示分子量标志物。

[0042] 图4:与SV2C、人突触结合蛋白1和GD1a复合的TriRecABTox的结合结构域的X-射线晶体结构。(a)Hc/TAB与SV2C、hSyt1和GD1a的缎带表示。(b-d)SV2C受体结合位点(b)、GD1a(c)和hSyt1(d)周围 $2\sigma$ 处的 $2F_0-F_c$ 电子密度图(网格(mesh))的实例。

[0043] 图5:与SV2受体的结合。(a)与hSV2C复合的Hc/TAB和Hc/A的晶体结构(PDB 4JRA)的叠加。(b)与糖基化的hSV2复合的Hc/TAB和Hc/A的晶体结构(PDB 5JLV)的叠加。作为棍显示了参与结合的残基(Benoit等人,2014),并根据相应Hc/A位置加标签。

[0044] 图6:与突触结合蛋白的结合。分别与人Syt1和大鼠Syt2复合的Hc/TAB和Hc/B的晶体结构(PDB 4KBB)的叠加。作为棍显示了参与结合的残基(Jin等人,2006;Chai等人,2006),并根据相应Hc/B位置加标签。

[0045] 图7:与GD1a的结合。与GD1a复合的Hc/TAB和Hc/B的晶体结构(PDB 4KBB)的叠加(分别显示为深灰色和浅灰色)。作为棍显示了参与结合的残基(Berntsson等人,2013),并根据相应Hc/B位置加标签。

[0046] 图8:BoNT/TAB的表征。(a)纯化的BoNT/TAB的SDS-PAGE分析以及与Hc/A和Hc/B对照的比较。(b-d)使用聚组氨酸探针(b);Hc/A(c)和Hc/B(d)抗血清的免疫印迹分析。样品与(a)中相同,“M”表示分子量标志物。

[0047] 图9:BoNT/TAB的激活。(a)描述功能域组织的BoNT/TAB构建体的示意图。作为黑色短划线显示了工程设计的蛋白酶激活位点。将轻链和重链之间的天然二硫键表示为黑色平直线。(b)BoNT/TAB激活测定的SDS-PAGE分析。分别地,非还原(NR)和还原(R)的未活化的BoNT/TAB(左图)和因子Xa-激活的BoNT/TAB(右图)。注释所关心的片段;“M”表示分子量标志物。

[0048] 图10:Hc/TAB的扩展使用。(a)与Hc/TAB有关的潜在功能性BoNT衍生物的示意图。所述构建体将包括来自任何血清型或亚型的功能性BoNT域(“n”)。还应包括蛋白酶激活位点(黑色短划线)。(b)使用Hc/TAB将载物蛋白(cargo protein)转运至神经元细胞表面的潜在构建体的示意图。

[0049] 图11:Hc/TAB的纯化。(a)使用5ml HisTrap FF柱,通过亲合色谱纯化的色谱图( $A_{280}$ 迹线)。(b)使用Superdex200柱,通过尺寸排除色谱纯化的色谱图( $A_{280}$ 迹线)。突出显示了纯化过程的阶段和具有Hc/TAB的级分。

[0050] 图12:与SV2C、hSyt1和GD1a复合的Hc/TAB的晶体。(a)在20%v/v聚乙二醇6000、0.1M柠檬酸盐pH 5.0中生长的晶体。(b)用于在Diamond 104-1工作站进行数据收集的固定在低温环(cryo-loop)上的晶体。(c)晶体的X-射线衍射图。

[0051] 图13:BoNT/TAB的纯化。(a)使用5ml HisTrap FF柱,通过亲合色谱纯化的色谱图( $A_{280}$ 迹线)。(b)使用Superdex200柱,通过尺寸排除色谱纯化的色谱图( $A_{280}$ 迹线)。突出显示了纯化过程的阶段和具有BoNT/TAB的级分。

[0052] 图14:与SV2C、人突触结合蛋白1和GD1a复合的Hc/TAB的结合结构域的X-射线晶体

结构。(a)和(b)Hc/TAB(a)和Hc/TAB2.1(b)的晶体结构的温度断裂分析-Putty半径表示,其中所述半径与B-因子成正比。显示了环“360”,以黑色突出显示了位置360和362,(c)Hc/TAB2.1与SV2C、hSyt1和GD1a的复合物的X-射线晶体结构(棍模型表示),(d)与(c)相同,其中标记了脂质结合环并作为棍模型显示了疏水性残基。

[0053] 图15:Hc/TAB2.1的纯化。(a)使用5ml HisTrap FF柱,通过亲和色谱纯化的色谱图(A<sub>280</sub>迹线),(b)使用Superdex200柱,通过凝胶过滤的色谱图(A<sub>280</sub>迹线)。显示了具有Hc/TAB2.1的级分,(c)和(d)Hc/TAB2.1的表征。来自纯化的Hc/TAB2.1,来自(c)中的亲和色谱法和凝胶过滤(d)的级分的SDS-PAGE分析;左侧第一泳道显示了分子量标志物。显示了对应于Hc/TAB2.1的条带。

[0054] 图16:Hc/TAB2.1.1和Hc/TAB2.1.3的纯化。(a)和(b)分别来自Hc/TAB2.1.1的亲合色谱纯化和凝胶过滤的色谱图(A<sub>280</sub>迹线)。显示了具有Hc/TAB2.1.1的级分,(c)和(d)分别来自Hc/TAB2.1.3的亲合色谱纯化和凝胶过滤的色谱图(A<sub>280</sub>迹线)。显示了具有Hc/TAB2.1.3的级分,(d)Hc/TAB2.1.1和Hc/TAB2.1.3的表征。纯化样品的SDS-PAGE分析;左侧第一泳道显示了分子量标志物。显示了对应于Hc/TAB2.1.1和Hc/TAB2.1.3的条带。

[0055] 图17:图X4:BoNT/TAB2.1.3的纯化。(a)和(b)来自BoNT/TAB2.1.3纯化的级分的SDS-PAGE分析。来自亲和色谱(a)和凝胶过滤(b)的级分;左侧第一泳道显示了分子量标志物。显示了对应于BoNT/TAB2.1.3的条带,(c)纯化的BoNT/TAB2.1.3样品的SDS-PAGE分析。在泳道1中:凝血酶激活前的样品,在泳道2中:最终激活的样品(凝血酶处理后)。显示了对应于全长(单链)HC和LC的条带。右侧泳道显示了分子量标志物,(d)来自使用Superdex200柱的最终凝胶过滤色谱(凝血酶切割后)的色谱图(A<sub>280</sub>迹线)。显示了具有BoNT/TAB2.1.3的级分。

[0056] 定义

[0057] 如本文所使用的,术语肉毒神经毒素“BoNT”涵盖了来自肉毒神经毒素的任何多肽或片段。术语BoNT可以表示全长BoNT。术语BoNT可以表示可以实施整体细胞机制,由此BoNT进入神经元并抑制神经递质释放的BoNT片段。术语BoNT可以简单地表示BoNT片段,而不要求所述片段具有任何特定功能或活性。

[0058] 如本文所使用的,术语“移位结构域”或“H<sub>N</sub>”表示可以实施介导BoNT轻链移位的中毒过程的移位步骤的BoNT域。因此,H<sub>N</sub>有利于BoNT轻链跨膜进入细胞的细胞质中的移动。

[0059] 如本文所使用的,术语“结合结构域”与“H<sub>C</sub>域”同义并且表示可以实施中毒过程的细胞结合步骤的任何天然存在的BoNT受体结合结构域,其包括(例如)BoNT与位于靶细胞质膜表面上的BoNT-特异性受体系统的结合。

[0060] 在本发明公开中,术语“核酸”和“基因”是可互换使用的以描述编码多肽的核苷酸序列或者多核苷酸。

## 具体实施方式

[0061] 如以上在背景部分中所说明的,BoNT包含通过单个二硫键与重链(HC)连接的轻链(LC)。重链(HC)具有两个功能域,即N末端移位结构域(H<sub>N</sub>)和C末端结合结构域(H<sub>C</sub>),而LC负责胞内催化活性。因此,H<sub>C</sub>包含能够特异性且不可逆地结合至在敏感神经元上表达的特异性受体的受体结合结构域,然而H<sub>N</sub>形成了允许所连接的LC从胞内体样膜囊移位至胞液的通

道。不同的BoNT血清型具有位于Hc上的不同的受体结合位点组,通常两个受体结合位点。本发明人已利用该知识工程设计了包含三个不同受体的结合位点的新的BoNT Hc结合结构域(Hc/TAB)。

[0062] 本发明人通过工程设计Hc/TAB域实现了该目标,所述Hc/TAB域包含:

[0063] a) 突触结合蛋白(Syt)受体结合位点,和

[0064] b) 突触相关囊泡蛋白2(SV2)受体结合位点,和

[0065] c) 神经节糖苷(Gang)受体结合位点。

[0066] 工程设计的Hc/TAB域的结构允许Hc/TAB协同结合至突触结合蛋白(Syt)受体、突触相关囊泡蛋白2(SV2)受体和神经节糖苷(Gang)受体。因此,实现了与神经元细胞上三个受体的协同结合,从而导致与其他BoNT Hc域相比,新的Hc/TAB域具有提高的亲合力。因此,改善了与神经元的整体结合,并因此改善了毒素的效力。

[0067] Hc还包含N末端(Hc<sub>N</sub>)和C末端(Hc<sub>C</sub>)。本发明的重要特征在于Hc/TAB的Hc<sub>C</sub>端的结构,它是受体结合结构域位于BoNT中的位置。

[0068] 在Hc/TAB的一个实施方式中,Hc<sub>C</sub>端可互换地由来自BoNT血清A型(BoNT/A)和BoNT血清B型(BoNT/B)的序列组成。通过工程设计该可互换结构,本发明人已能够优化与所有三个受体的协同结合。

[0069] 在本发明的其他实施方式中,根据序列A<sub>1</sub>B<sub>1</sub>A<sub>2</sub>B<sub>2</sub>A<sub>3</sub>组成Hc<sub>C</sub>端,其中A表示来自BoNT/A的序列,B表示来自BoNT/B的序列,参见,图2。这进一步优化了Hc/TAB的结构,以允许三个受体结合结构域至少协同结合至所有三个所述受体,可能地甚至同时结合至所有三个所述受体。本发明人已显示通过该A<sub>1</sub>B<sub>1</sub>A<sub>2</sub>B<sub>2</sub>A<sub>3</sub>序列,体外发生了与所有三个受体的同时结合。在SEQ. ID. No. 1中描述了根据该具体实施方式的工程设计的A<sub>1</sub>B<sub>1</sub>A<sub>2</sub>B<sub>2</sub>A<sub>3</sub>序列。

[0070] 为了进一步优化根据以上所述的Hc/TAB,已引入了突变和缺失以产生稳定的分子内界面,参见,图2。在SEQ. ID. No. 1中,在位置306、360和362中产生了替换并且与原始序列相比,在位置265/266和360/361之间产生了缺失。然而,技术人员将理解在距以上所指明的位置的+1、+2、+3、+4、+5或-1、-2、-3、-4或-5位置的氨基酸的突变和/或缺失可以具有相同效果。因此,在距所指明的氨基酸位置+/-5个氨基酸的位置的任何这种修饰在本发明公开的范围之内。

[0071] 根据以上具体实施方式和所有以下实施例,神经节糖苷受体结合位点来源于BoNT/B,但是可设想它可以来源于任何Gang受体结合的BoNT血清型和它们的亚型,如BoNT血清A、B、C、D、DC、E、En、F、G或X型或其亚型,因为所有血清型具有神经节糖苷受体结合位点。

[0072] 根据本发明的优选实施方式,形成Gang受体结合位点的序列位于B<sub>2</sub>中。

[0073] SV2受体结合结构域通常可以来源于任何SV2受体结合BoNT血清型和它们的亚型,并且具体地来源于BoNT血清A、D、E和F型。在以上具体实施方式和所有以下实施例中,SV2受体结合结构域来源于BoNT/A,但是如技术人员将理解的,根据所附权利要求的Hc/TAB的目的和预期用途,包含SV2受体结合结构域的任何血清型可以用作所述域的来源。

[0074] SV2受体结合结构域的部分存在于Hc<sub>N</sub>端。因此,Hc<sub>N</sub>序列可以来源于任何SV2受体结合的BoNT血清型和它们的亚型。在以上具体实施方式和所有以下实施例中,Hc<sub>N</sub>端来源于BoNT/A。然而,如技术人员将理解的,只要SV2受体结合结构域是功能性的,则Hc<sub>N</sub>序列还可以

来源于任何BoNT血清C、D、E、F或G型。

[0075] 此外,根据本发明的优选实施方式,形成SV2受体结合位点的序列位于H<sub>CN</sub>中和H<sub>CC</sub>中的A<sub>1</sub>和A<sub>3</sub>中。

[0076] Syt受体结合位点可以来源于任何Syt受体结合BoNT血清型和它们的亚型。具体地,Syt受体结合位点可以来源于BoNT血清B型、血清嵌合DC型或血清G型。根据本发明的优选实施方式,形成Syt受体结合位点的序列位于B<sub>1</sub>和B<sub>2</sub>中。

[0077] 本发明还提供了包含根据以上所述的H<sub>C</sub>/TAB的多肽。因此,所述多肽可以包含直接或通过接头与H<sub>C</sub>/TAB结合的任何其他蛋白、多肽、氨基酸序列或荧光探针。在下文中,与H<sub>C</sub>/TAB结合的蛋白、多肽或氨基酸序列被称为“蛋白”。

[0078] 根据一个优选实施方式,所述多肽是重组BoNT多肽(BoNT/TAB),其进一步在多肽序列中包含H<sub>N</sub>和LC以及位于LC和H<sub>N</sub>之间的外切蛋白酶位点。

[0079] 外切蛋白酶位点使得单链多肽能够被切割成双链分子,从而使所述分子称为活性毒素。根据本发明的实施方式,所述外切蛋白酶位点是因子Xa位点,尽管这不是根据本发明的多肽的限制性特征。

[0080] 根据一个实施方式,处于其活性形式的BoNT/TAB根据SEQ. ID.No.5。根据另一个实施方式,处于其活性形式的BoNT/TAB根据SEQ. ID.No.6、8、10或12所示序列中的任一种。优选地,处于其活性形式的BoNT/TAB根据SEQ. ID.No.12。

[0081] 所述H<sub>N</sub>和LC两者可以分别且独立地来源于任何BoNT血清A、B、C、D、DC、E、En、F、G或X型和它们的亚型以及BoNT样多肽。出现了与BoNT类似,即具有类似的领结构和不同程度的序列同一性,但通过肉毒杆菌(*C. botulinum*)以外的其他生物所产生的新的蛋白。因此,技术人员将能够选择来自任何BoNT血清型,它们的亚型或者BoNT样多肽的H<sub>N</sub>和/或LC。

[0082] 如以上所指明的,引入H<sub>C</sub>/TAB中的突变和缺失进一步确保可以作为具有正确结构和所需活性的可溶性蛋白产生工程设计的BoNT/TAB。

[0083] 优选地,作为需要重组产生的H<sub>C</sub>/TAB,重组产生了根据以上所述的多肽。

[0084] 因此,本发明公开还提供了编码任何根据以上所述的H<sub>C</sub>/TAB或多肽的分离和/或重组的核酸。编码本发明公开所述的H<sub>C</sub>/TAB或多肽的核酸可以是双链或单链DNA或RNA。在某些方面,编码分离的多肽片段的受试核酸还应理解为包含编码作为任何本文所述的H<sub>C</sub>/TAB或多肽的变体的多肽的核酸。变体核苷酸序列包括相差一个或多个核苷酸替换、添加或缺失的序列,如等位变体。

[0085] 本发明还提供了包含编码根据以上所述的H<sub>C</sub>/TAB的核酸序列的载体。所述载体还可以包含编码与H<sub>C</sub>/TAB一起重组产生的任何其他蛋白或探针的核酸序列,从而在一个多肽中获得与H<sub>C</sub>/TAB结合的所述蛋白或探针。所述载体优选地为表达载体。所述载体可以包含可操作性地连接至所述核酸的启动子。多种启动子可以用于本文所述的多肽的表达,并且它们是技术领域中的技术人员已知的。

[0086] 可以通过常规方法(例如,电穿孔、脂质体转染和磷酸钙沉淀)将包含核酸的表达载体转移至宿主细胞,然后通过常规方法培养转染的细胞以产生本文所述的多肽。在一些实施方式中,通过组成型、诱导型或组织-特异型启动子调控本文所述的多肽的表达。

[0087] 可以在任何真核或原核细胞,或者在酵母中产生多肽。还可以在无细胞系统中产生根据本发明所述的多肽。技术人员将易于能够将所选表达系统应用于人。用于产生本发

明所述的多肽的表达系统不受限于本发明的范围。

[0088] 重组蛋白的纯化和修饰在本领域中是熟知的,从而多聚蛋白前体的设计可以包括技术人员所易于理解的一些实施方式。

[0089] 要包括在所述多肽中的蛋白可以是要转运至神经元细胞和/或内化至神经元细胞中的任何多关心的蛋白。

[0090] 如果蛋白的内化是所期望的,可以有利地在多肽中包含根据以上所述的H<sub>N</sub>以及H<sub>C</sub>/TAB,并用所关心的蛋白替换LC,因为H<sub>N</sub>则将提供允许蛋白移位至神经元细胞中的通道。如果神经元细胞表面是蛋白靶标,则可以有利地将所关心的蛋白直接结合至H<sub>C</sub>/TAB。因此,基于目标递送,可以获得以下组合:

[0091] i) 蛋白-H<sub>C</sub>/TAB

[0092] ii) 蛋白-H<sub>N</sub>-H<sub>C</sub>/TAB

[0093] iii) 蛋白-LC-H<sub>N</sub>-H<sub>C</sub>/TAB

[0094] 根据以上i),通过将载物蛋白与H<sub>C</sub>/TAB的结合,可以将载物蛋白靶向至神经元表面。将可能发生通过常规细胞表面再循环程序的一些内化,但是神经元表面将是该方法的主要靶标。

[0095] 通过将载物蛋白结合至根据以上ii)与H<sub>C</sub>/TAB结合的H<sub>N</sub>或者结合至根据以上iii)的BoNT/TAB,使用毒素移位系统,所述载物蛋白可以更有效地转运至神经元内部。如背景中所述,一旦将BoNT毒素在囊泡中的神经元细胞中内化,则囊泡中的酸性内体环境导致了允许LC从囊泡移位至细胞胞液的构象变化。因此,通过利用BoNT/TAB,作为将BoNT的LC从内化囊泡移位至胞液中的机制的所述毒素移位系统可以用于将上述载物蛋白移位至神经元的胞液中。载物蛋白可以结合至H<sub>N</sub>而不是LC,且如以上所公开的,具有位于载物蛋白和H<sub>N</sub>之间的外切蛋白酶位点,或者载物蛋白可以结合至LC。两种变体将使得载物蛋白能够转运至神经元细胞的胞液中。

[0096] 因此,H<sub>C</sub>/TAB和BoNT/TAB两者可以用作将任何蛋白转运至神经元和/或转运到神经元中的媒介物。这也提供了在药理学测试中使用H<sub>C</sub>/TAB和/或BoNT/TAB的可能性以研究蛋白在(例如)突触过程中的作用。

[0097] 载物蛋白可以(例如)是任何蛋白标签,如亲合力或荧光标签或探针。因此,可以在以上公开的载体中包括任何与这种蛋白标签对应的核酸。技术人员将能够使用标准克隆方法以在载体中包含任何所关心的基因,以及用于蛋白表达的标准规程。

[0098] 可以通过允许它们翻译后重组的分选酶系统,分开表达BoNT的结合结构域和载物蛋白。因此,分选酶的转肽酶活性可以用作体外产生融合蛋白的工具,并且其很好地处于该技术领域内的技术人员知识范围内。简言之,将识别基序(LPXTG)添加至所关心的蛋白的C-末端,而将寡甘氨酸基序添加至要连接的第二蛋白的N-末端。一旦向所述蛋白混合物添加分选酶,则通过天然肽键共价连接所述两个肽。该方法可以用于产生根据本发明的多肽。在本发明的情况下,这将意味着将识别基序添加至所关心的蛋白的C-末端,而将寡甘氨酸基序添加至H<sub>C</sub>/TAB或BoNT/TAB的N-末端。

[0099] 另外,可以在治疗方法或美容方法中使用H<sub>C</sub>/TAB和/或BoNT/TAB。通常,H<sub>C</sub>/TAB和/或BoNT/TAB的使用可以非常类似于已对于BoNT/A和/或BoNT/B产物的使用。这些包括其中所述方法和治疗的目的是使肌肉无力和/或僵化的方法和治疗。

[0100] 根据本发明的Hc/TAB使得能够向细胞注射具有更高亲合力的BoNT/TAB并因此具有更高的效率。因此,需要低剂量并且更长的作用时间是可能的。因此,与BoNT/A或BoNT/B相比,对于相同的效果,可以注射较少量的BoNT/TAB,这降低了副作用,因此较少的BoNT/TAB将从注射位点扩散。通过更高的效率,更强且更有效的结合以及较低的所需剂量,多余的可用于扩散至注射位点以外的BoNT/TAB的量较少。此外,可以在具有持续效果的情况下不太频繁地施用BoNT,这也将最大程度降低免疫应答和不良反应及其后果的风险。

[0101] 可以通过根据以上所述的Hc/TAB和/或BoNT/TAB治疗和/或预防的典型的医学病况为选自下列的病况:神经肌肉失调、涉及乙酰胆碱的释放的病况和肌肉痉挛病症。更具体地,它可以涉及选自下列的病况:痉挛性发声障碍、痉挛性斜颈、喉肌张力障碍、口下颌发音障碍(romandibular dysphonia)、舌肌张力障碍、颈肌张力障碍、局灶性手肌张力障碍(focal hand dystonia)、眼睑痉挛、斜视、半面痉挛、眼睑障碍、大脑性麻痹、局灶性痉挛及其他语言障碍、痉挛性结肠炎、神经原性膀胱、肛门痉挛、肢体痉挛、抽搐(tics)、震颤、磨牙症、肛裂、弛缓不能、吞咽困难及其他肌张力障碍以及特征为肌群不自主运动的其他障碍、流泪、多汗、过度流涎、过度胃肠分泌、分泌紊乱、肌肉痉挛引起的疼痛、头痛、运动损伤和抑郁症。

[0102] 对于美容方法,Hc/TAB和/或BoNT/TAB可以优选地用于预防和/或治疗皱纹、额皱纹或不希望的纹路,以减少所述皱纹、额皱纹和纹路。

[0103] 根据以上所述的Hc/TAB和/或BoNT/TAB可以配制成任何适合的药物或美容组合物。包含Hc/TAB和/或BoNT/TAB的药物组合物还可以包含药物可用的赋形剂、载体或其他添加剂。包含Hc/TAB和/或BoNT/TAB的美容组合物还可以包含美容可用的赋形剂、载体或其他添加剂。

[0104] 药物或美容组合物的施用可以通过注射,其中在存在不希望的神元活动的身体位点施用所述注射。通常,用于通过注射施用的组合物是无菌等渗水性缓冲液中的溶液。必要时,所述组合物还可以包含增溶剂和局部麻醉剂以缓解注射位点的疼痛。

[0105] 此外,药物或美容组合物可以包含在具有所述组合物的治疗施用的说明书的试剂盒中。在该试剂盒中,可以在指明活性剂的量的密闭容器(如安瓿瓶或袋(sachette))中单独提供组合物的成分或者将其以单位剂量形式混合在一起,例如,作为干燥的冷冻干燥粉末或者无水浓缩物。可以通过输注施用所述组合物,并且在该情况下,可以通过含无菌药物级水或盐水的输液瓶进行分配。当通过注射施用所述组合物时,可以提供无菌注射用水或盐水的安瓿瓶,从而可以在施用前混合所述成分。用于全身施用的组合物可以是液体,例如,无菌盐水、乳酸盐林格溶液或汉克溶液。另外,所述组合物可以处于固体形式并在使用前立即再溶解或混悬。还考虑了冷冻干燥形式。所述组合物可以包含在脂质颗粒或囊泡内,如脂质体或微晶,其还适合于肠胃外施用。

[0106] 因此,本发明人已开发了适合于同时结合至SV2C受体、突触结合蛋白受体和神经节糖苷受体中的全部三个的工程设计的BoNT生物杂化体。由此,提供了比现有技术的BoNT多肽具有更高效能、效力和持续时间的BoNT生物杂化体。由此,本发明的生物杂化体的使用使得能够施用比根据现有技术更低剂量的毒素,同时维持相同的效果。此外,本发明的生物杂化体的使用使得能够比先前使用的BoNT更比频繁地施用。因此,使用本发明的BoNT生物杂化体对患者的治疗将更舒适,其中不必如现有技术一样频繁发生施用。

[0107] 实验部分

[0108] 材料和方法

[0109] 构建体。将编码Hc和全长(无活性)TriRecABTox(分别为Hc/TAB和BoNT/TAB)的cDNA对大肠杆菌(*E. coli*)表达(参见DNA序列的补充资料)进行密码子优化,合成并克隆至具有N末端6×His标签的pET-28a(+)载体中(GenScript, NJ, USA)。在我们的研究中使用的TriRecABTox构建体在催化位点具有三个突变以避免任何安全性问题(E224Q/R363A/Y366F)(Rossetto等人,2001;Binz等人,2002)。BoNT/TAB基因编码1311个氨基酸,并且Hc/TAB基因对应于残基[875-1311]。

[0110] 蛋白表达和纯化。将具有所关心的基因的质粒转化至大肠杆菌(*E. coli*)BL21(DE3)细胞(New England BioLabs, USA)。将类似的规程用于两种蛋白。在LEX系统(Epyphite3, Canada)中,通过将细胞在具有50μg/ml卡那霉素的超级肉汤培养基中,在37℃生长约3小时,然后用终浓度1mM的IPTG诱导,并在18℃保持过夜来进行表达。收获细胞并保存在-80℃。通过Emulsiflex-C3(Avestin, Germany),在20kPsi,在具有200mM NaCl、25mM咪唑和5%(v/v)甘油的25mM HEPES pH 7.2中实施用于蛋白质提取的细胞胞溶。将细胞碎片通过超速离心,在4℃,267,000g离心沉淀45min。首先通过亲和色谱纯化蛋白:将上清液加载至5ml HisTrap FF柱(GE Healthcare, Sweden)上,用25mM HEPES pH 7.2,200mM NaCl,25mM咪唑和5%(v/v)甘油清洗,并用25mM HEPES pH7.2,200mM NaCl,250mM咪唑和5%(v/v)甘油洗脱蛋白。然后,将样品对25mM HEPES pH 7.2,200mM NaCl和5%(v/v)甘油透析过夜,然后使用Superdex200柱,以类似的缓冲液(GE Healthcare, Sweden)进行最终的尺寸排阻纯化步骤。将Hc/TAB以4.5mg/ml并且将BoNT/TAB以7.3mg/ml保持在具有200mM NaCl,0.025mM TCEP和5%甘油的25mM HEPES pH 7.2中。

[0111] 蛋白表征。使用NuPAGE 4-12%Bis-Tris凝胶,通过凝胶电泳以及通过在PVDF膜(ThermoFisher, Sweden)上实施的免疫印迹分析蛋白样品。内部制备(在兔中)抗Hc/A和Hc/B的第一抗体,并用抗兔IgG-过氧化物酶抗体(产品目录#SAB3700852, Sigma, Sweden)检测。使用HRP-缀合的单克隆抗体(ADI.1.10, 产品目录#MA1-80218, ThermoFisher, Sweden)检测聚-组氨酸标签。将TMB底物(Promega, Sweden)用于检测。包括了与Hc/TAB类似地纯化并且由加His标签的Hc/A和Hc/B所组成的内部对照以用于比较。

[0112] BoNT/TAB的激活。通过位于轻链和重链之间的因子Xa切割位点(IEGR)设计了全长(无活性)TriRecABTox以用于激活成双链形式。通过将100μg BoNT/TAB与2μg因子Xa(New England BioLabs, USA)在4℃培育过夜来进行激活。通过凝胶电泳(如上)分析激活结果。

[0113] SV2C-L4的克隆、表达和纯化。从cDNA扩增突触囊泡糖蛋白2C的第四腔内结构域的相互作用部分(SV2C-L4, 残基474-567, Uniprot ID Q.496J9)并使用LIC克隆将其克隆至pNIC28-Bsa4(具有TEV位点的N末端His6标签)。使用类似于如上所述的规程,在大肠杆菌(*E. coli*)BL21(DE3)(New England BioLabs, USA)中表达SV2CL4。通过亲和色谱,在2mL HisTrap HP柱(GE Healthcare, Sweden)上纯化加His标签的SV2C-L4,用20mM HEPES, pH 7.5、500mM NaCl、10%(v/v)甘油、50mM咪唑和0.5mM TCEP清洗。用20mM HEPES, pH 7.5、500mM NaCl、10%(v/v)甘油、500mM咪唑和0.5mM TCEP洗脱蛋白。然后,使用Superdex75HiLoad 16/60柱(GE Healthcare, Sweden),在20mM HEPES, pH 7.5、300mM NaCl、10%(v/v)甘油和0.5mM TCEP中,通过尺寸排除色谱进一步纯化SV2CL4。

[0114] X-射线晶体学。通过在室温下,将3.6mg/ml的Hc/TAB与1mg/ml的SV2C-L4(重组人SV2C胞外环-4[残基475-565])、1mM hSyt1肽(GEGKEDAFSKLKEKFMNELHK,通过GenScript, USA合成)和4mM GD1a寡糖(Elicityl, France)预培育15min制备用于结晶的样品。

[0115] 使用沉滴(sitting drop)装置,通过与100nl由20%v/v聚乙二醇6000、0.1M柠檬酸盐pH 5.0组成的储液(JCSG-plus screen B9, Molecular Dimensions, United Kingdom)混合的200nl样品使晶体生长,并在21°C培育。2周内出现晶体,并将其转移至低温环并在液氮中冷冻。

[0116] 在配备有PILATUS-6M检测器(Dectris, Switzerland)的Diamond光源(Didcot, UK)的工作站104-1收集衍射数据。在100°K,从单晶收集了1.5Å的完整数据集。处理原始数据图像并使用CCP4 suite 7.0(CCP4, 1994),通过DIALS(Gildea等人, 2014)和AIMLESS(Evans, 2006)进行缩放。

[0117] 通过与SV2C-L4复合的Hc/A(PDB代码4JRA)和与小鼠SytII和GD1a复合的Hc/B(PDB代码4KBB)的坐标制备的模型进行分子置换以在PHASER(McCoy等人, 2007)中确定用于结构解析的初始相位。使用REFMAC5(Murshudov等人, 2011)精修工作模型并通过COOT(Emsley等人, 2010)手动调整。在其中Fo-Fc电子密度峰大于3σ的位置处添加水分子,并且可以添加潜在的氢键。通过MOLPROBITY(Chen等人, 2010)进行验证。Ramachandran统计显示全部残基中的97.0%处于最有利的区域中,仅有一个异常值处于不接受的区域中。表1中总结了结晶学数据统计。通过PyMOL(Schrödinger, LLC, USA)绘图。

[0118] 结果

[0119] TriRecABTox的设计:具有三受体结合位点的工程设计的肉毒毒素

[0120] 为了实现三受体毒素的概念,本发明人首先分析了对于BoNT/A和/B与它们的受体分子相互作用可得的结构信息。Yao等人(2016)和Benoit等人(2014)的近期工作分别提供了具有(PDB 5JLV)和不具有翻译后修饰(PDB 4JRA)的与SV2C复合的BoNT/A的受体结合结构域的X-射线晶体结构。SV2C的腔内结构域(环4)形成了主要通过主链-与-主链相互作用而与Hc/A结合的四边形β-螺旋,其中在两个子域的界面处具有开放的β-链,同时SV2C的N-聚糖朝HcN延伸(图1)。总体上,这些结构表明与所述两种SV2形式的共同结合模式还应延伸至糖基化SV2A和SV2B(Yao等人, 2016)。这些研究突出显示了因此在TriRecABTox的设计中应保留的参与毒素-SV2相互作用的关键残基和多个位点(图1)。这些包括BoNT/A的节段[949-953]、[1062-1066]、[1138-1157]和[1287-1296]。残基编号基于BoNT/A1序列(Uniprot-P10845)。

[0121] 还描述了与突触结合蛋白复合的BoNT/B的几种晶体结构,并且其帮助限定了毒素与其受体的相互作用(Chai等人, 2006; Jin等人, 2006; Berntsson等人, 2013)(图1)。一旦结合, Syt肽采取沿C末端子域的远端上的沟结合的短螺旋结构,其直接涉及BoNT/B的节段[1113-1118]和[1183-1205]。残基编号基于BoNT/B1序列(Uniprot-P10844)。因此,认为这些区域对于在TriRecABTox构建体中的包含是必要的。

[0122] 另外,与它们的神经节糖苷受体复合的晶体结构BoNT/A和/B(Stenmark等人, 2008; Hamark等人, 2017; Berntsson等人, 2013)提供了对于每种血清型的碳水化合物结合位点的详细说明。所述位点在肉毒神经毒素家族中是高度保守的并且由HcC子域上的浅口袋

组成(图1),所述口袋由中间SxWY基序(/A中的1264-1267;/B中的1260-1263)和周围的环区组成。值得注意的是,该口袋接近BoNT/B中的Syt受体结合位点,其通过环[1244-1253]分开,然而对于两种受体的同时结合尚未报道过变构效应(Bertnsson等人,2013)。为了最大程度降低对Syt受体结合位点的任何结构改变,认为在TriRecABTox的设计中更适合引入BoNT/B,而不是BoNT/A的神经节糖苷受体-结合位点。

[0123] 在识别来自对于与三个不同受体的结合所必需的两种血清型的组分之后,进行了进一步的结构分析以将它们整合成单一分子。在这种程度上,通过Clustal0对BoNT/A(Uniprot P10845)和BoNT/B(Uniprot P10844)的一级序列进行比对(Sievers等人,2011),并将它们的结合结构域的立体结构叠加(图1)。两种血清型共有40%的整体序列同一性,然而对于作为负责受体识别的主要区域的Hc的C末端子域,相似性降低至34%。结合结构域的核心折叠在全部梭菌神经毒素中是保守的(Swaminathan,2011;Rummel等人,2011),但是连接环的长度显著不同。因此,重要的是还考虑二级结构(图1),以保持主要域结构完整。因此,新设计的分子的模板表现为BoNT/A和/B元件之间的多次交替,从而产生了可能不相容的新的非天然分子内界面。对Hc/A和Hc/B的叠加的晶体结构的检查使得本发明人能够通过关键位置中的单个氨基酸替换或缺失,通过修正潜在冲突来优化设计(图2)。具体地,检查冲突区内每个残基的侧链,从而导致在BoNT/B中,用等价BoNT/A氨基酸的三个替换:N1180、G1234、N1236(SEQ.ID.No.3)。另外,除去几个氨基酸(图2)以匹配二级结构元件和补偿BoNT/A和/B之间在过渡界面(transition interfaces)的环区中的长度差异。与BoNT/A和BoNT/B序列相比,在L1139和G1140之间以及G1234和T1235之间造成缺失(参考SEQ.ID.No.3)(图2)。

[0124] 将所得分子命名为TriRecABTox,其应能够结合至三种受体:SV2、突触结合蛋白和神经节糖苷。在SEQ.ID.No.3(无活性形式)和SEQ.ID.No.5(活性形式)中提供了其蛋白序列。

[0125] TriRecABTox结合结构域的产生和表征。

[0126] TriRecABTox表征的第一步是重组产生结合结构域(Hc/TAB)以分析其生物化学性质。为此目的,将蛋白序列对于在大肠杆菌(E.coli)中的表达进行密码子优化。将所得基因克隆至pET-28a(+)载体以包含N末端聚-组氨酸标签并且有利于蛋白纯化过程,在方法部分提供了详细信息。本发明人显示可以表达并使用亲和色谱和尺寸排阻技术(图11)部分纯化(图3)Hc/TAB。原始样品存在一些可能对应于残余宿主细胞蛋白的低分子量污染物。使用如离子交换或疏水相互作用色谱的方法的其他纯化步骤应帮助获得高纯度样品。通过免疫印迹确认加His标签的Hc/TAB的存在,其中在预期大小(约53kDa)观察到了单一条带(图3)。

[0127] 与其三个受体复合的TriRecABTox结合结构域的晶体结构。

[0128] 为了评价Hc/TAB与其三个受体结合的能力,建立了共结晶试验,其包括具有人SV2C腔内结构域[残基475-565]、人Syt1肽[残基34-53]和GD1a碳水化合物的Hc/TAB。获得了高分辨率衍射(1.5Å)的晶体(图12),并且可以收集完成数据集(表1)。使用具有所有潜在组分的输入模型,通过分子置换解析结构。解析结构确认晶体结构含有全部四个元件:同时与其三个受体结合的Hc/TAB(称为Hc/TAB-3R)(图4)。该结果提供了TriRecABTox可以体外实现其目的首个实验证据,并且其还使得能够以原子级水平详细地全面分析受体结合机制。使用新确定的结构信息,我们可以直接比较Hc/TAB、Hc/A、Hc/B和它们各个受体之间的相

相互作用。

[0129] 表1.X-射线晶体学:数据收集和精修统计

<b>Hc/TAB – SV2C – hSytI – GD1a复合物</b>	
<b>数据采集</b>	
空间群	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
晶胞大小	
<i>a, b, c</i> (Å)	43.7, 115.9, 141.4
<i>a, b, g</i> (°)	90.0, 90.0, 90.0
[0130] 分辨率 (Å)	1.5-60.4 (1.51-1.53)*
全/独特反射数	3,583,212 / 113,806
<i>R</i> <sub>合并</sub>	0.119 (1.926)*
<i>R</i> <sub>pim</sub>	0.021 (0.501)*
CC <sub>1/2</sub>	1.00 (0.839)*
<i>I</i> / <i>σI</i>	13.0 (1.1)*

	完整度 (%)	99.9 (97.2)*
	冗余度	31.5 (15.1)*
<hr/>		
	精修	
	$R_{\text{工作}} / R_{\text{游离}} (\%)$	17.4 / 22.1
	原子数	
	Hc/TAB	3,682
	SV2C	761
	hSytl	143
	GD1a	56
[0131]	水	446
	B- 因子	
	Hc/TAB	27.4
	SV2C	44.4
	hSytl	35.8
	GD1a	36.8
	水	40.2
	均方根误差	
	键长 (Å)	0.009
	键角 (°)	1.37

[0132] \*括号中的值是最高分辨率壳层的值。

[0133] 首先,新设计的BoNT/TAB的结合结构域显示出与其两个子域:凝集素样H<sub>CN</sub>和H<sub>CC</sub>的β-三叶草折叠的预期折叠(图4)。所产生的多个新的分子内界面不会干扰整体结构,如通过与Hc/A叠加时0.69Å的低均方根误差(rmsd)(对于364个Ca)和当与Hc/B叠加时0.81Å的低均方根误差(对于370个Ca)所显示的。除了异常的N末端聚-组氨酸标签和环[1169-1173]之外,对完整的Hc/TAB建模[876-1311]。可以通过以下事实解释电子密度对于这些部分的缺少:这些区域不参与任何相互作用并且位于晶体的溶剂可接近的区域内。

[0134] 将Hc/TAB-3R结构与和SV2C复合的Hc/A的结构相比较。SV2C腔内结构域的结构在两种复合物中是相同的,其rmsd为**0.483Å**(相对于88个Ca)。基于Hc域,三维比对两种结构,并且如本发明人的设计所预期的,显示SV2C处于相同位置(图5)。具体地,已指定对于SV2受体结合所必需的且包含在Hc/TAB中的来自Hc/A的区域是完全保守的。通过PISA分析Hc/A和SV2C之间的界面(Kissinel,2015)并且所述界面对应于主要包括静电相互作用的**540Å<sup>2</sup>**的表面区域,其中来自两种蛋白的开放的链形成了互补的β-折叠结构(Benoit等人,2014)。使用Hc/TAB的相应分析显示了**630Å<sup>2</sup>**的具有SV2C的表面区域,并且确认了具有相当氢键数目的结合机制。另外,本发明人还通过将Hc/TAB-3R与Hc/A-gSV2C复合物相比较,考虑了与糖基化的SV2的潜在结合(图5)。最近,已显示N559的N-糖基化对于受体识别是必要的并且在SV2同工型中是保守的(Yao等人,2016)。值得注意地,在具有或不具有糖基化的情况下,Hc/A和SV2C之间的蛋白-蛋白相互作用是高度类似的。碳水化合物链朝Hc<sub>N</sub>子域延伸。参与蛋白-聚糖相互作用的Hc/A残基的分析显示它们的位置在Hc/TAB-3R中是完全保守的,因此Hc/TAB应能识别SV2的N-糖基化同工型。

[0135] 然后,本发明人将Hc/TAB-3R的结构与和rSyt2复合的Hc/B的结构进行了比较。预期BoNT/B以与其啮齿类同源物类似的方式结合至人突触结合蛋白,尽管亲和力不同(Tao等人,2017)。在本文所提供的晶体结构中,hSyt1也采取了位于与rSyt2在Hc/B中相同的结合沟内的α-螺旋结构(图6)。结合至它们各自Hc域的hSyt1与rSyt2的叠加确认了保守的肽构型,其rmsd为**0.560Å**(相对于13个Ca)。另外,在Hc/TAB中,受体-结合口袋是完全保守的,尽管参与结合的所有残基在两种结构中显示出类似的构型(图6)。这通过PISA分析得到确认,其中对于还包含11个静电键的Hc/TAB:hSyt1相互作用计算界面为**861Å<sup>2</sup>**,并且其与具有7个静电键的**712Å<sup>2</sup>**的Hc/B:rSyt2界面(PDB 4KBB)相当。识别机制主要基于强蛋白-蛋白疏水相互作用。可以通过hSyt1和rSyt2之间的序列变化,具体地对于所述肽的C末端一半部分的序列变化来解释接触表面积和静电相互作用个数的细小差异。

[0136] Hc/TAB-3R结构中所含的第三受体对应于GD1a碳水化合物,对此从Gal2至Sia5观察到了清楚的电子密度(图4)。对于Glc1和Sia6,未见电子密度,如可以根据无相互作用的柔性碳水化合物部分所预期的。已详细研究了神经节糖苷-受体结合位点,并且与GD1a复合的Hc/B的晶体结构已确认该血清型对于末端Sia(α2-3)Gal部分是优选的(Bertnsson等人,2013;Rummel,2016)。设计TriRecABTox以整合Hc/B结合口袋,并且根据天然毒素,两种结构的比较(图7)显示结合口袋的关键残基(S1260、W1262、Y1263)是完全保守的并且与GD1a相互作用。当与GD1a-结合的Hc/B相比时,大部分结合位点保持无变化,但是存在少数明显的例外。在Hc/TAB-3R中,N1122的侧链避开配体,而其Hc/B等价物N1105与Sia5形成直接氢键。这通过位置I1257在一定程度上得到补偿,其与Hc/B:GD1a结构中的I1240(其中它们相隔**7Å**)相比,在Hc/TAB-3R中显示出与Sia5更强的疏水相互作用(**4.3Å**的距离)。

[0137] 整体上,得自Hc/TAB-3R晶体结构的结果确认单个TriRecABTox分子能够以重复母体BoNT/A和/B的结合机制的方式同时结合至SV2受体、突触结合蛋白受体及其神经节糖苷受体。

[0138] 全长、无活性的TriRecABTox的产生和表征。

[0139] 通过确立Hc/TAB的结合能力,本发明人接着表达和纯化了全长、无催化活性的TriRecABTox (BoNT/TAB; SEQ. ID. No. 3)。为此目的,本发明人设计了编码1311个氨基酸并且含有三个BoNT功能域的合成基因,其中LC和H<sub>N</sub>对应于BoNT/A域,所述功能域与Hc/TAB结合。出于安全考虑,包含了催化位点处的三个突变(E224Q/R363A/Y366F) (Rossetto等人,2001; Binz等人,2002)。根据如上所述的Hc/TAB构建体,将蛋白序列对于在大肠杆菌(E. coli)中的表达进行密码子优化并克隆至具有N末端聚-组氨酸标签的pET-28a(+)中。在方法部分中提供了详细信息。本发明人显示BoNT/TAB可以作为约152kDa的可溶性蛋白表达。用于纯化的初始方法获得了有限的量的非均一的材料(图8;图13),但是使用如离子交换或疏水相互作用色谱的方法的进一步纯化应帮助获得更纯的材料,并且消除了通过凝胶电泳可见的残余的宿主细胞蛋白。最近使用这种方法产生了纯度大于80%的重组BoNT/B构建体(Elliott等人,2017)。

[0140] 实施其他表征并确认了组氨酸标签的存在,并且尽管与对照相比,与探针抗体的反应非常弱(图8B),但是在正确的大小存在可辨别的弱条带。测定还显示了与约70kDa的污染物的交叉反应。此外,如所预期的,BoNT/TAB确定与内部抗Hc/A(图8C)和Hc/B(图8C)抗血清反应,因为其应含有来自两个结合结构域的表位。

[0141] TriRecABTox的控制激活。

[0142] 设计了具有位于轻链和重链之间的因子Xa切割位点,IEGR[442-445](图9A)的BoNT/TAB,因为向双链形式的激活是获得完全活性的毒素所必需的。将如上所述的全长BoNT/TAB样品(SEQ. ID. No. 5)用于实施激活测定。尽管样品不均一,但是在BoNT/TAB与因子Xa以1μg蛋白酶比50μg毒素的比例在4℃培育过夜之后,实现了完全激活(图9B)。凝胶电泳显示了当在存在还原剂的情况下进行时,BoNT/TAB向约100和50kDa的两个片段的分离,其很可能分别对应于HC和LC。这两条链通过C430和C458之间的二硫键保持在一起,从而解释了非还原条件下,约150kDa的单一一条带。对应于HC和LC的条带在非还原条件下也是可见的,并且可以由样品制备期间二硫键的一定水平的还原引起,然而这些条带在未激活的对照中不是清楚可见的。

[0143] 总的来说,激活测定首先提供了所产生的蛋白对应于工程设计的BoNT/TAB的证据,并且其次提供了可以成功控制激活为双链分子的步骤。因此,该步骤可以包括在活性全长TriRecABTox的产生中。

[0144] BoNT/TAB的优化

[0145] 材料和方法

[0146] 构建体。通过GenScript (NJ, USA) 在如上所述的pET28(a)载体中克隆了编码Hc/TAB的cDNA和变体。通过Toxogen GmbH (Hannover, Germany) 在pET29(a)载体中克隆了BoNT/TAB2.1.3。

[0147] 蛋白表达和纯化。如以上对于Hc/TAB变体所述的。

[0148] 通过与用于Hc/TAB(亲和色谱和凝胶过滤)所使用的规程类似的规程,通过Toxogen GmbH (Hannover, Germany) 产生了BoNT/TAB2.1.3。另外,通过浓度为0.05U/μg的凝血酶进行BoNT/TAB2.1.3的激活和标签去除,并通过凝胶过滤进一步纯化BoNT/TAB2.1.3。将样品保存在具有200mM NaCl和5%甘油的25mM HEPES pH 7.2中。

[0149] 蛋白表征。如上所述(使用NuPAGE 4-12%Bis-Tris凝胶的凝胶电泳)。

[0150] X-射线晶体学。通过在室温下,将6.5mg/ml的Hc/TAB2.1与1mg/ml的SV2C-L4(重组人SV2C胞外环-4[残基475-565])、1mM hSyt1肽(GEGKEDAFSKLKEKFMNELHK,通过GenScript, USA合成)和4mM GD1a寡糖(Elicityl, France)预培育15min制备用于结晶的样品。使用沉滴装置,通过与100nl由20%v/v聚乙二醇3350、0.2M柠檬酸钾组成的储液(JCSG-plus screen B12, Molecular Dimensions, United Kingdom)混合的200nl样品使晶体生长,并在21°C培育。1周内出现晶体,并将其转移至低温环并在液氮中冷冻。在配备有PILATUS-6M检测器(Dectris, Switzerland)的Diamond光源(Didcot, UK)的工作站104-1收集衍射数据。在100°K,从单晶收集了1.4Å的完整数据组。处理原始数据图像并使用CCP4 suite 7.0 (CCP4, 1994),通过DIALS (Gildea等人, 2014)和AIMLESS (Evans, 2006)进行缩放。

[0151] 通过先前在PHASER中确定的Hc/TAB的结构进行分子置换(McCoy等人, 2007)。使用REFMAC5 (Murshudov等人, 2011)精修工作模型并通过COOT (Emsley等人, 2010)手动调整。在其中 $F_o - F_c$ 电子密度峰大于 $3\sigma$ 的位置处添加水分子,并且可以添加潜在的氢键。通过MOLPROBITY进行验证(Chen等人, 2010)。

[0152] Ramachandran统计显示全部残基中的97.0%处于最有利的区域中,仅有一个异常值处于不接受的区域中。在表X1中总结了结晶学数据统计。

[0153] 优化的Hc/TAB、Hc/TAB2.1的产生

[0154] 分析了结合至其三个受体的Hc/TAB的晶体结构以鉴别可以修饰以改善分子的稳定性和功能的潜在位点。

[0155] 具体地,可以将晶体结构内的局部温度因子(B-因子)的分析理解为蛋白局部稳定性的指示,其中高B-因子表示区域无序。根据该分析,认为由残基D357至N362 (SEQ. ID.No.6)所组成的Hc/TAB的两个子域之间界面处的环(标记为“环360”)是最优化的(参见图14)。将残基G360和N362 (SEQ ID.No.1)修饰为它们在BoNT/B中的等价残基,并分别突变为P360和Y362以引入标记为Hc/TAB2.1的新构建体的序列中 (SEQ ID.No.6)。

[0156] 通过定点突变(GenScript, USA)制备了该新构建体的质粒并将其用于Hc/TAB2.1在大肠杆菌(E. coli)中的重组表达。所使用的规程与对于Hc/TAB的产生所使用的相同(参见对于表达和纯化的原始方法部分)。我们显示可以表达并使用亲和色谱和尺寸排阻技术部分纯化Hc/TAB2.1 (图15)。样品存在一些可能对应于残余宿主细胞蛋白的低分子量污染物。使用如离子交换或疏水相互作用色谱的方法的其他纯化步骤应帮助获得高纯度样品。

[0157] 在与人SV2C腔内结构域[残基475-565]、人Syt1肽[残基34-53]和GD1a碳水化合物的共结晶试验中使用了纯化的Hc/TAB2.1 (SEQ. ID.No.6)。获得了高分辨率衍射(1.4Å)的晶体,并且可以收集完整数据组(表2)。使用结合至其三个受体的Hc/TAB的晶体结构(Hc/TAB-3R),通过分子置换解析结构。新的结构显示所有元件在Hc/TAB-3R中可见并且提供了根据Hc/TAB, Hc/TAB2.1可以同时结合至三个受体的实验证据。B-因子的分析显示了对于环“360”改善的稳定性(D357至Y362;图14)。整体上, Hc/TAB2.1的行为类似于HC/TAB,其中两个构建体显示了就得率和纯度而言相当的谱图。

[0158] 表X1. X-射线结晶学:数据收集和精修统计

## Hc/TAB2.1 – SV2C – hSytI – GD1a 复合物

数据采集	
空间群	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
晶胞大小	
<i>a, b, c</i> (Å)	43.8, 117.5, 141.5
$\alpha, \beta, \gamma$ (°)	90.0, 90.0, 90.0
分辨率 (Å)	1.4-60.6 (1.40-1.56)*
全/独特反射数	1,055,932 / 79,953
<i>R</i> <sub>合并</sub>	0.084 (1.070)*
<i>R</i> <sub>pim</sub>	0.024 (0.364)*
CC <sub>1/2</sub>	0.99 (0.716)*
<i>I</i> / $\sigma$ <i>I</i>	16.0 (1.8)*
完整度 (%)	95.1 (70.1)*
冗余度	13.2 (9.3)*
精修	
<i>R</i> <sub>工作</sub> / <i>R</i> <sub>游离</sub> (%)	17.7 / 20.4
均方根误差	
键长 (Å)	0.118
键角 (°)	1.60

[0160] \*括号中的值是最高分辨率壳层的值。

[0161] 更可溶的变体Hc/TAB2.1.3的产生

[0162] 为了准备对Hc/TAB变体未来的功能分析,使Hc/TAB2.1适合于与近期所述的分选酶连接实验相容(Zhang等人,2017)。该实验允许可以用于测试活性的全长、活性BoNT的安全且控制的重构。该构建体对应于具有可切割N末端His标签的N末端截短的Hc/TAB2.1,并将

其标记为Hc/TAB2.1.1 (SEQ. ID.No.8)。制备了Hc/TAB2.1.1的克隆 (GenScript), 并如上所述将其用于表达和纯化(图16)。我们显示可以表达并使用亲和色谱和尺寸排阻技术部分纯化Hc/TAB2.1.1。样品存在一些可能对应于残余宿主细胞蛋白的低分子量污染物。使用如离子交换或疏水相互作用色谱的方法的其他纯化步骤应帮助获得高纯度样品。

[0163] Hc/TAB2.1的结构特征的进一步分析突出显示了从蛋白的其余部分中伸出的表面暴露的疏水性环的存在(残基389-393, SEQ. ID.No.6;图14d)。另外,最近将该环鉴别为BoNT/B及其他血清型中的脂质-结合元件(Stern等人,2018)。我们假设该疏水性区域可以妨碍Hc/TAB的溶解度,因此设计了新的构建体,其中该环是截短的并且被双-天门冬酰胺基序替换以提高溶解度。将该构建体标记为Hc/TAB2.1.3 (SEQ. ID.No.10)。制备了Hc/TAB2.1.3的克隆 (GenScript), 并如上所述将其用于表达和纯化(图16)。我们显示可以表达并使用亲和色谱和尺寸排阻技术部分纯化Hc/TAB2.1.3。样品存在一些可能对应于残余宿主细胞蛋白的低分子量污染物。使用如离子交换或疏水相互作用色谱的方法的其他纯化步骤应帮助获得高纯度样品。值得注意地,与Hc/TAB2.1.1相比,Hc/TAB2.1.3显示出更好的表达得率和溶解度(图16)。

#### [0164] 全长、活性BoNT/TAB2.1.3的产生

[0165] 为了准备BoNT/TAB的未来功能分析,产生了基于Hc/TAB2.1.3构建体的全长、活性变体并将其标记为BoNT/TAB2.1.3 (SEQ. ID.No.12)。在许可的机构中,按照合同协议,在Toxogen GmbH (Hannover, Germany) 进行所有生产步骤。在pET29 (a) 载体中克隆了BoNT/TAB2.1.3, 并且其包含可切割的C末端Strep标签和聚-组氨酸标签,以及如上所述,在HC和LC域之间用于产物激活的工程设计的凝血酶切割位点(SEQ. ID.No.13)。可以作为可溶性蛋白表达BoNT/TAB2.1.3, 纯化并通过凝血酶激活(图17)。用于纯化的方法包括亲和色谱和凝胶过滤,并且其导致产生了>90%纯度的BoNT/TAB2.1.3产物。

#### [0166] 未来的实验

#### [0167] 受体结合测定

[0168] 将进行期中BoNT/TAB的受体-结合性质将与BoNT/A和/或BoNT/B相比的测定。

[0169] 例如,将实施根据先前所述方法修改的神经节糖苷受体-结合测定。简要地,在该ELISA中,将所关心的神经节糖苷受体(GT1b、GD1b、GD1a或者GM1a)固定在96-孔微板中(Chen等人,2008;Willjes等人,2013),然后,应用毒素(或它们的结合结构域),并用缀合至辣根过氧化物酶(HRP)的单克隆抗聚-组氨酸抗体检测所结合的材料。这种定性方法应提供足够的信息以确认BoNT/TAB的神经节糖苷受体-结合特征与BoNT/B的类似。

[0170] 神经节糖苷受体结合ELISA。神经节糖苷GT1b、GD1b、GD1a和GM1a购自Carbosynth (Compton, UK)。在甲醇中稀释神经节糖苷以达到2.5 $\mu$ g/ml的最终浓度;将100 $\mu$ L (0.25 $\mu$ g) 应用于96-孔PVC测定板中的每个孔中。在溶剂在21 $^{\circ}$ C蒸发(过夜)后,用200 $\mu$ L PBS/0.1% (w/v) BSA清洗(3 $\times$ )孔。通过在21 $^{\circ}$ C,在200 $\mu$ L PBS/2% (w/v) BSA中培育2h,封闭非特异性结合位点。在100 $\mu$ L含有样品(6 $\mu$ M至0.003 $\mu$ M的范围内的连续3-倍稀释)的PBS/0.1% (w/v) BSA/孔中,在4 $^{\circ}$ C进行2h的结合测定。培育后,用PBS/0.1% (w/v) BSA清洗孔3 $\times$ ,然后在4 $^{\circ}$ C,与以1:2000稀释的HRP-抗His抗体(ThermoFisher#MA1-80218) (100 $\mu$ L/孔) 培育1h。最终,在用PBS/0.1% (w/v) BSA的三个清洗步骤后,使用Ultra TMB (100 $\mu$ L/孔) 检测结合的样品。在21 $^{\circ}$ C培育5min之后,通过添加100 $\mu$ L 1M硫酸,终止反应。使用Tecan Infinite 200 (

Männedorf, Switzerland) 测量450nm的吸光度。使用非线性结合拟合,通过Prism (GraphPad, La Jolla, CA, USA) 分析结果。

[0171] 为了评价与突触结合蛋白受体的结合性质,与Berntsson等人(2013)所述的测定类似,实施等温滴定量热法(ITC)。将测量hSyt肽与毒素的结合并且应提供确认与BoNT/B类似,BoNT/TAB可以结合至受体的亲和力值( $K_d$ )。

[0172] 等温滴定量热法。通过其他尺寸排阻色谱步骤(Superdex200, GE Healthcare, Sweden),在20mM磷酸钾pH 7.0、0.15M NaCl中制备样品。在ITC200 (GE Healthcare, Sweden) 上,在25°C和750rpm测量了Syt肽与BoNT或它们的结合结构域的结合。将200 $\mu$ L蛋白溶液(20 $\mu$ M)添加至小池中。通过以200 $\mu$ M的浓度,以16次逐步注射,每次注射2.5 $\mu$ L添加肽(GenScript, USA)来测量结合。第一次滴定设置为0.5 $\mu$ L,并随后在数据分析中删除。通过生产商所提供的Origin软件分析数据

[0173] 使用如通过Benoit等人(2014)所述的下拉测定(Pull-down assay),评价与SV2C的结合。简要地,加标签的毒素和非加标签的受体(或相反地)将在一起培育并加载至Ni-琼脂糖上,然后清洗并洗脱。通过SDS-PAGE显示结果。

[0174] 后趾外展评分(DAS)测定

[0175] 可以使用小鼠后趾外展评分(DAS)测定评价BoNT制剂的效力(Broide等人,2013)。该测定体内测量了在向小鼠或大鼠下肢骨骼肌肌肉注射后,毒素的局部肌肉-弱化效力。毒素引起动物产生特征性下肢惊吓反应的能力的可测量的剂量-依赖性降低。这种非致死方法已经常用于估计不同BoNT血清型或衍生物的药理学性质,如最近描述的重组BoNT/B分子(Elliott等人,2017)。类似的方法将用于评价与BoNT/A或/B相比,BoNT/TAB的效力和持续时间。

[0176] 讨论

[0177] 在该研究中,本发明人描述了如何使用BoNT/A和/B的结合机制的结构和分子细节来工程设计具有提高的细胞识别能力的新的分子TriRecABTox。BoNT/A和/B结构的严格多尺度比较使得本发明人能够鉴别保留在单个分子中在其上整合了SV2、突触结合蛋白和神经节糖苷的受体结合位点的完整毒素骨架所需的关键元件。新产生的设计包括BoNT/A和/B元件的交替,并且通过保护适应性突变或缺失来补偿新产生的非天然分子内界面来进行优化。这些修饰被视为是确保可以作为具有正确结构和所需活性的可溶性蛋白产生工程设计的毒素BoNT/TAB所必需的。

[0178] 本发明人首先评价了通过在其Hc/TAB自身上产生结合结构域的设计的稳定性,所述Hc/TAB通过在大肠杆菌(*E. coli*)中的重组表达而具有修饰的受体识别功能。作为可以部分纯化的可溶性蛋白,表达了具有N末端聚-组氨酸标签的Hc/TAB,因此表明了工程设计的构建体的存活力。在第二步中,本发明人继续进行处于无催化活性形式的全长BoNT/TAB构建体的产生。再次,本发明人显示其可以作为153kDa的可溶性蛋白表达,并通过标准液相色谱技术部分纯化。Hc/TAB和BoNT/TAB两者上的聚-组氨酸标签的存在使其能够通过亲和色谱,通过Ni-琼脂糖基质进行纯化。可以使用其他亲和方法并且其包含应优先位于蛋白N末端上的亲和标签以避免干扰受体结合。尽管初始制剂显示出不均一的样品纯度,但是纯化方法的优化应导致产生药物标准的产品。应补充BoNT/TAB的活性形式将具有与本研究中所使用的无活性分子类似的整体结构和结合性质。本发明人与Toxogen GmbH (Hannover,

Germany) 签订协议来生产通过可除去的C末端标签成功纯化从而不防碍受体结合的BoNT/TAB的活性形式(BoNT/TAB2.1.3)。

[0179] 另外,单链BoNT翻译后切割为双链分子是毒素活性的基本步骤(DasGupta and Sathyamoorthy, 1985; Shone等人, 1985)。尽管通常通过宿主蛋白酶激活天然毒素,但是任何重组BoNT产物需要外肽酶的加工。对毒素的早期工作显示胰蛋白酶可以将BoNT/A非特异性切割成活性双链形式(Shone等人, 1985),然而这可能导致毒素不希望的其他降解。最近,重组技术已使得能够在所关心的蛋白内工程设计特异性蛋白酶识别基序,因此对BoNT的激活策略提供了更好的控制(Sutton等人, 2005)。在本文中,本发明人在LC和HC之间包含了因子Xa位点并且观察到了毒素的完全激活,因此显示了该酶的有效性。未来的BoNT/TAB的生产应引入纯化平台,其使得能够激活毒素,然后从最终产物除去外切蛋白酶。尽管因子Xa显示是胜任的,但是可以测试其他酶并在实现可接受的激活得率方面证明是成功的。本发明人与Toxogen GmbH (Hannover, Germany) 签订协议来生产通过凝血酶外切蛋白酶成功激活并纯化至均一的BoNT/TAB的活性形式(BoNT/TAB2.1.3)。

[0180] 作为验证Hc/TAB的结构完整性和确认其提高的功能性的方式,本发明人将与人SV2C、人Syt1和GD1a碳水化合物复合的纯化样品共结晶。高分辨率(1.5Å)解析复合物的X-射线晶体结构并提供了单个Hc/TAB分子可以同时结合至全部三个受体的结论性实验证据。此外,与其各自受体的Hc/A和Hc/B的已知结构的比较显示Hc/TAB遵循了几乎相同的结合机制。

[0181] 尽管晶体结构表明Hc/TAB至少可以体外满足其目的,但是需要实施其他生物化学实验来完整表征其受体结合性质。这些将包括使用蛋白受体的下拉测定和ITC测定以及神经节糖苷受体结合ELISA。预期BoNT/TAB对于SV2受体结合与BoNT/A类似,并且对于神经节糖苷受体和突触结合蛋白受体结合与BoNT/B类似。另外,体内实验将提供BoNT/TAB作为治疗剂的真正潜力的主要指示。小鼠DAS测定已经典地用于评价BoNT制剂(Broide等人, 2013)并且应使得发明人能够确定我们的分子与现用产品相比的效力和作用持续时间。

[0182] 另外,可以通过修饰一些序列元件来进一步优化BoNT/TAB的设计以改善其生物化学性质和稳定性。这些改变可以包括导致可溶性BoNT仍能够同时结合至三个受体的缺失或突变。本发明人成功产生了更稳定的变体(Hc/TAB2.1)并且以更高的产率产生了更可溶的变体(Hc/TAB2.1.3)。

[0183] 应补充从安全性角度,BoNT/TAB不代表新的威胁,因为它来源于两个已有的血清型。预期它被现用抗毒素识别,如肉毒抗毒素七价BAT或用于BoNT/A和/B的其他批准的解毒剂。

[0184] 血清A型和B型是唯一批准的上市的BoNT。尽管BoNT/A是主要治疗使用的毒素,但是具有低免疫原性和高效力的分子将提供更安全的替代选择(Naumann等人, 2013)。已多次尝试改善BoNT的性质以提高其药理学潜力(Masuyer等人, 2014)。最近成功的实例包括Tao等人(2017)的研究,其中在BoNT/B关键位置工程设计的突变(E1191M/S1199Y)为毒素提供了对人突触结合蛋白2受体更高的亲合力,并且与野生型相比,在阻断神经传递中显示出约11-倍更高的效力。Elliott等人(2017)采取了另一种改善BoNT效力的方法,其中他们分析了已知提高BoNT/B的其底物的催化活性的单个突变(S201P)的作用。在这种情况下,BoNT/B突变体在多细胞-基测定和体内中未显示出优于野生型的任何优势。总体上,对于BoNT/B的

这两项研究表明毒素效力的限制步骤在于初始神经元识别,而不是随后的胞内活性。

[0185] 旨在将一种血清型的结合性质与另一种的催化活性组合的更早期的研究导致设计了其中交换完整域的嵌合分子(Wang等人,2008,2012;Rummel等人,2011)。更具体地,Rummel等人(2011)和Wang等人(2012)设计并测试了由与BoNT/A的H<sub>N</sub>+LC域结合的Hc/B域所组成的类似分子。据报道与野生型BoNT/A相比,这些重组毒素在小鼠中显示出提高的效力并诱导了更长的效果(Kutschenko等人,2017)。当评价由与血清A型的互补域结合的BoNT/B的C末端子域(H<sub>CC</sub>) (即LC+H<sub>N</sub>+H<sub>CN</sub>) 所组成并且显示出比野生型高4倍的效力的构建体时,获得了类似的观察结果(Rummel等人,2011)。所有上述分子共同具有以下事实:它们将仅识别BoNT/B、突触结合蛋白和神经节糖苷中的两个受体。这些结果表明由于神经元上更普遍的BoNT/B受体所允许的更大的LC/A的进入,可以获得延长的作用和更高的效力。另外,这些嵌合分子未考虑可能由来自不同血清型的域的组合所引起的且可能影响这些产物的潜能的可能的域内分子内冲突。

[0186] 考虑到最近对BoNT工程设计的研究结果,很清楚修饰初始细胞识别是提高治疗产品的药理学性质的最有效方法之一。因此,作为成功工程设计的识别SV2受体以及BoNT/B受体、突触结合蛋白和神经节糖苷的单一产物,BoNT/TAB显示出大潜力,并且可以比野生型BoNT/A和/B更有效。

[0187] BoNT/TAB的主要创新在于允许多受体相互作用的结合结构域的设计。当前证据表明Hc/TAB与BoNT/A的移位和催化结构域的结合应提供具有最强效力的分子(如BoNT/TAB中所设计的)。然而,当与其他血清型的功能域组合时,Hc/TAB仍可以是所关心的(图10a)。另外,Hc/TAB还可以与作为药理学工具使用的其他所关心的蛋白结合(图10b)来研究突触过程。通过BoNT/TAB所实施的体内测定应阐明其对于这种目的应用。

[0188] 参考文献

[0189] Benoit RM,Frey D,Hilbert M,Kevenaer JT,Wieser MM,Stirnimann CU,McMillan D,Ceska T,Lebon F,Jaussi R,Steinmetz MO,Schertler GF,Hoogenraad CC,Capitani G,Kammerer RA.2014.Structural basis for recognition of synaptic vesicle protein 2C by botulinum neurotoxin A.Nature.505:108-111.

[0190] Bentivoglio AR,Del Grande A,Petracca M,lalongo T,Ricciardi L.2015.Clinical differences between botulinum neurotoxin type A and B.Toxicon.107:77-84.

[0191] Berntsson RP,Peng L,Dong M,Stenmark P.2013.Structure of dual receptor binding to botulinum neurotoxin B.Nat Commun.4:2058.

[0192] Binz T,Bade S,Rummel A,Kollewe A,Alves J.2002.Arg(362) and Tyr(365) of the botulinum neurotoxin type a light chain are involved in transition state stabilization.Biochemistry 41:1717-1723.

[0193] Binz T,Rummel A.2009.Cell entry strategy of clostridial neurotoxins.J Neurochem.109:1584-1595.

[0194] Broide RS,Rubino J,Nicholson GS,Ardila MC,Brown MS,Aoki KR,Francis J.2013.The rat Digit Abduction Score (DAS) assay:a physiological model for assessing botulinum neurotoxin-induced skeletal muscle paralysis.Toxicon.71:

18-24.

[0195] Chai Q, Arndt JW, Dong M, Tepp WH, Johnson EA, Chapman ER, Stevens RC. 2006. Structural basis of cell surface receptor recognition by botulinum neurotoxin B. *Nature*. 444:1096-1100.

[0196] Chen C, Baldwin MR, Barbieri JT. 2008. Molecular basis for tetanus toxin coreceptor interactions. *Biochemistry*. 47:7179-7186.

[0197] Chen S. 2012. Clinical uses of botulinum neurotoxins: current indications, limitations and future developments. *Toxins (Basel)*. 4:913-39.

[0198] Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* 66:12-21.

[0199] Collaborative Computational Project, Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* 50:760-763.

[0200] DasGupta BR, Sathyamoorthy VS. 1985. Separation, purification, partial characterization and comparison of the heavy and light chains of botulinum neurotoxin types A, B, and E. *J Biol Chem* 260:10461-10466.

[0201] Dong M, Richards DA, Goodnough MC, Tepp WH, Johnson EA, Chapman ER. 2003.

[0202] Synaptotagmins I and II mediate entry of botulinum neurotoxin B into cells. *J Cell Biol*. 162:1293-303.

[0203] Dong M, Yeh F, Tepp WH, Dean C, Johnson EA, Janz R, Chapman ER. 2006. SV2 is the protein receptor for botulinum neurotoxin A. *Science*. 312:592-596.

[0204] Dressier D, Bigalke H. 2005. Botulinum toxin type B de novo therapy of cervical dystonia: frequency of antibody induced therapy failure. *J Neurol*. 252:904-907.

[0205] Elliott M, Maignel J, Liu SM, Favre-Guilmand C, Mir I, Farrow P, Hornby F, Marlin S, Palan S, Beard M, Krupp J. 2017. Augmentation of VAMP-catalytic activity of botulinum neurotoxin serotype B does not result in increased potency in physiological systems. *PLoS One*. 12:e0185628.

[0206] Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D* 66:486-501.

[0207] Evans P. 2006. Scaling and assessment of data quality. *Acta Crystallogr D* 62:72-82.

[0208] Gildea RJ, Waterman DG, Parkhurst JM, Axford D, Sutton G, Stuart DL, Sauter NK, Evans G, Winter G. 2014. New methods for indexing multi-lattice diffraction data. *Acta Crystallogr D* 70:2652-2666.

[0209] Hamark C, Berntsson RP, Masuyer G, Henriksson LM, Gustafsson R, Stenmark P, Widmalm G. 2017. Glycans Confer Specificity to the Recognition of Ganglioside Receptors by Botulinum Neurotoxin A. *J Am Chem Soc*. 139:218-230.

[0210] Hatheway CL. 1990. Toxigenic Clostridia. *Clin. Microbiol. Rev.* 3:66-98.

- [0211] Jin R, Rummel A, Binz T, Brunger AT. 2006. Botulinum neurotoxin B recognizes its protein receptor with high affinity and specificity. *Nature*. 444:1092-1095.
- [0212] Krissinel E. 2015. Stock-based detection of protein oligomeric states in jsPISA. *Nucl. Acids Res.* 43:W314-9.
- [0213] Kutschenko A, Reinert MC, Krez N, Liebetanz D, Rummel A. 2017. BoNT/AB hybrid maintains similar duration of paresis as BoNT/A wild-type in murine running wheel assay. *Neurotoxicology*. 59:1-8.
- [0214] Lacy DB, Stevens RC. 1999. Sequence homology and structural analysis of the clostridial neurotoxins. *J Mol Biol.* 291:1091-104.
- [0215] Lange O, Bigalke H, Dengler R, Wegner F, deGroot M, Wohlfarth K. 2009. Neutralizing antibodies and secondary therapy failure after treatment with botulinum toxin type A: much ado about nothing? *Clin. Neuropharmacol.* 32: 213-218.
- [0216] Masuyer G, Chaddock JA, Foster KA, Acharya KR. 2014. Engineered botulinum neurotoxins as new therapeutics. *Annu Rev Pharmacol Toxicol.* 54:27-51.
- [0217] Mahrhold S, Rummel A, Bigalke H, Davletov B, Binz T. 2006. The synaptic vesicle protein 2C mediates the uptake of botulinum neurotoxin A into phrenic nerves. *FEBS Lett.* 580:2011-2014.
- [0218] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* 40:658-674.
- [0219] Montal M. 2010. Botulinum neurotoxin: a marvel of protein design. *Annu Rev Biochem.* 79:591-617.
- [0220] Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. Refmac5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D.* 67:355-367.
- [0221] Naumann M, Boo LM, Ackerman AH, Gallagher CJ. 2013. Immunogenicity of botulinum toxins. *J Neural. Transm (Vienna)*. 120:275-290.
- [0222] Nishiki T, Kamata Y, Nemoto Y, Omori A, Ito T, Takahashi M, Kozaki S. 1994. Identification of protein receptor for Clostridium botulinum type B neurotoxin in rat brain synaptosomes. *J. Biol. Chem.* 269:10498-10503.
- [0223] Nishiki T, Tokuyama Y, Kamata Y, Nemoto Y, Yoshida A, Sato K, Sekiguchi M, Takahashi M, Kozaki S. 1996. The high-affinity binding of Clostridium botulinum type B neurotoxin to synaptotagmin II associated with gangliosides GT1b/GD1a. *FEBS Lett.* 378:253-257.
- [0224] Robert X, Gouet, P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucl. Acids Res.* 42, W320-W324.
- [0225] Rossetto O, Caccin P, Rigoni M, Tonello F, Bortoletto N, Stevens RC, Montecucco C. 2001. Active-site mutagenesis of tetanus neurotoxin implicates

TYR-375 and GLU-271 in metalloproteolytic activity. *Toxicon* 39:1151-1159.

[0226] Rossetto O, Pirazzini M, Montecucco C. 2014. Botulinum neurotoxins: genetic, structural and mechanistic insights. *Nat Rev Microbiol.* 2014;12:535-549.

[0227] Rummel A, Mahrhold S, Bigalke H, Binz T. 2011. Exchange of the H(CC) domain mediating double receptor recognition improves the pharmacodynamic properties of botulinum neurotoxin. *FEBS J.* 278:4506-4515.

[0228] Rummel A. 2013. Double receptor anchorage of botulinum neurotoxins accounts for their exquisite neurospecificity. *Curr Top Microbiol Immunol.* 364: 61-90.

[0229] Rummel A. 2016. Two Feet on the Membrane: Uptake of Clostridial Neurotoxins. *Curr Top Microbiol Immunol.* Springer, Berlin, Heidelberg.

[0230] Schengrund C.L, DasGupta B.R. and Ringler N.J. 1991. Binding of botulinum and tetanus neurotoxins to ganglioside GT1b and derivatives thereof. *J. Neurochem.* 57, 1024-1032.

[0231] Schiavo G, Matteoli M, Montecucco C. 2000. Neurotoxins affecting exocytosis. *Physiol. Rev.* 80:717-766.

[0232] Shone CC, Hambleton P, Melling J. 1985. Inactivation of Clostridium botulinum type A neurotoxin by trypsin and purification of two tryptic fragments. Proteolytic action near the COOH-terminus of the heavy subunit destroys toxin-binding activity. *Eur J Biochem* 151, 75-82.

[0233] Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7:539.

[0234] Stenmark P, Dupuy J, Imamura A, Kiso M, Stevens RC. 2008. Crystal structure of botulinum neurotoxin type A in complex with the cell surface co-receptor GT1b-insight into the toxin-neuron interaction. *PLoS Pathog.* 4:e1000129.

[0235] Strotmeier, J., Willjes, G., Binz, T. & Rummel, A. 2012. Human synaptotagmin-11 is not a high affinity receptor for botulinum neurotoxin B and G: increased therapeutic dosage and immunogenicity. *FEBS Lett.* 586, 310-313.

[0236] Sudhof TC, Rothman JE. 2009. Membrane fusion: grappling with SNARE and SM proteins. *Science* 323, 474-477.

[0237] Sutton JM, Wayne J, Scott-Tucker A, O'Brien SM, Marks PM, Alexander FC, Shone CC & Chaddock JA. 2005. Preparation of specifically activatable endopeptidase derivatives of Clostridium botulinum toxins type A, B, and C and their applications. *Protein Expr Purif* 40, 31-41.

[0238] Swaminathan S. 2011. Molecular structures and functional relationships in clostridial neurotoxins. *FEBS J.* 278:4467-4485.

- [0239] Takamizawa K.,Iwamori M.,Kozaki S.,Sakaguchi G.,Tanaka R.,Takayama H.and Nagai Y.1986.TLC immunostaining characterization of Clostridium botulinum type A neurotoxin binding to gangliosides and free fatty acids.FEBS Lett.201:229-232.
- [0240] Takamori S,Holt M,Stenius K,Lemke EA,GrOnborg M,Riedel D,Urlaub H,Schenck S,Brilgger B,Ringler P,Muller SA,Rammner B,Grater F,Hub JS,De Groot BL,Mieskes G,Moriyama Y,Klingauf J,Grubmuller H,Heuser J,Wieland F,Jahn R.2006.Molecular anatomy of a trafficking organelle.Cell 127:831-846.
- [0241] Tao L,Peng L,Berntsson RP,Liu SM,Park S,Yu F,Boone C,Palan S,Beard M,Chabrier PE,Stenmark P,Krupp J,Dong M.2017.Engineered botulinum neurotoxin B with improved efficacy for targeting human receptors.Nat Commun.8:53.
- [0242] Wang J,Meng J,Lawrence GW,Zurawski TH,Sasse A,Bodeker MO,Gilmore MA,Fernandez-Salas E,Francis J,Steward LE,Aoki KR,Dolly JO.2008.Novel chimeras of botulinum neurotoxins A and E unveil contributions from the binding, translocation,and protease domains to their functional characteristics.J Biol Chem.283:16993-17002.
- [0243] Wang J,Zurawski TH,Bodeker MO,Meng J,Boddul S,Aoki KR,Dolly JO.2012.Longer-acting and highly potent chimaeric inhibitors of excessive exocytosis created with domains from botulinum neurotoxin A and B.Biochem J.444:59-67.
- [0244] Wilhelm BG,Mandad S,Truckenbrodt S,Krohnert K,Schafer C,Rammner B,Koo SJ,ClaRen GA,Krauss M,Haucke V,Urlaub H,Rizzoli SO.2014.Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins.Science 344:1023-1028.
- [0245] Willjes G,Mahrhold S,Strotmeier J,Eichner T,Rummel A,Binz T.2013.Botulinum neurotoxin G binds synaptotagmin-II in a mode similar to that of serotype B:tyrosine 1186 and lysine 1191 cause its lower affinity.Biochemistry.2013 52:3930-3938.
- [0246] Yao G,Zhang S,Mahrhold S,Lam KH,Stern D,Bagramyan K,Perry K,Kalkum M,Rummel A,Dong M,Jin R.2016.N-linked glycosylation of SV2is required for binding and uptake of botulinum neurotoxin A.Nat Struct Mol Biol.23:656-662.
- [0247] Zhang S,Masuyer G,Zhang J,Shen Y,Lundin D,Henriksson L,Miyashita SI,Martinez-Carranza M,Dong M,Stenmark P.2017.Identification and characterization of a novel botulinum neurotoxin.Nat Commun.8:14130.
- [0248] Stern D,Weisemann J,Le Blanc A,von Berg L,Mahrhold S,Piesker J,Laue M,Luppa PB,Dorner MB,Dorner BG,Rummel A.2018.A lipid-binding loop of botulinum neurotoxin serotypes B,DC and G is an essential feature to confer their exquisite potency.PLoS Pathog.14(5):e1007048.



195	200	205
Leu Asn Glu Lys Glu Ile Lys Asp Leu Tyr Asp Asn Gln Ser Asn Ser		
210	215	220
Gly Ile Leu Lys Asp Phe Trp Gly Asp Tyr Leu Gln Tyr Asp Lys Pro		
225	230	235
240		240
Tyr Tyr Met Phe Asn Ala Gly Asn Lys Asn Ser Tyr Ile Lys Leu Lys		
245	250	255
Lys Asp Ser Pro Val Gly Glu Ile Leu Gly Pro Arg Gly Ser Val Met		
260	265	270
Thr Thr Asn Ile Tyr Leu Asn Ser Ser Leu Tyr Arg Gly Glu Lys Phe		
275	280	285
Ile Ile Arg Arg Lys Ser Asn Ser Gln Ser Ile Asn Asp Asp Ile Val		
290	295	300
Arg Asn Glu Asp Tyr Ile Tyr Leu Asp Phe Phe Asn Leu Asn Gln Glu		
305	310	315
320		320
Trp Arg Val Tyr Thr Tyr Lys Tyr Phe Lys Lys Glu Glu Glu Lys Leu		
325	330	335
Phe Leu Ala Pro Ile Ser Asp Ser Asp Glu Phe Tyr Asn Thr Ile Gln		
340	345	350
Ile Lys Glu Tyr Asp Glu Gln Gly Thr Asn Ser Cys Gln Leu Leu Phe		
355	360	365
Lys Lys Asp Glu Glu Ser Thr Asp Glu Ile Gly Leu Ile Gly Ile His		
370	375	380
Arg Phe Tyr Glu Ser Gly Ile Val Phe Glu Glu Tyr Lys Asp Tyr Phe		
385	390	395
400		400
Cys Ile Ser Lys Trp Tyr Leu Lys Glu Val Lys Arg Lys Pro Tyr Asn		
405	410	415
Leu Lys Leu Gly Cys Asn Trp Gln Phe Ile Pro Val Asp Asp Gly Trp		
420	425	430
Gly Glu Arg Pro Leu		
435		
<210> 2		
<211> 1314		
<212> DNA		
<213> 人工序列		
<220>		
<223> 对于大肠杆菌(E. coli)中的表达密码子优化的Hc/TriRecABTox		
<400> 2		
aagaacatta tcaacaccag catcctgaac ctgcgctacg agagcaacca cctgatcgac 60		

ctgagccgct acgcgagcaa gattaacatc ggtagcaagg tgaactttga cccgattgat 120  
 aaaaaccaga tccaactgtt caacctggaa agcagcaaga tcgaagtgat tctgaaaaac 180  
 gcgattgttt ataacagcat gtacgaaaac ttcagcacca gcttttggat ccgtattccg 240  
 aagtatttta acagcatcag cctgaacaac gaatacacca tcattaactg catggagaac 300  
 aacagcgggt ggaaagtgag cctgaactac ggcgaaatca tttggaccct gcaggacacc 360  
 caagagatca agcagcgtgt ggttttcaag tacagccaaa tgatcaacat cagcgattac 420  
 atcaaccggt ggattttcgt taccatcacc aacaaccgtc tgaacaacag caagatctac 480  
 attaacggtc gtctgattga ccagaaaccg atcagcaacc tgggcaacat tcacgcgagc 540  
 aacaacatca tgttcaagct ggacggttgc cgtgatacc accgttatat ctggattaag 600  
 tacttcaacc tgtttgataa agagctgaac gaaaaggaga ttaaagacct gtatgataac 660  
 cagagcaaca gcggtatcct gaaggacttt tggggcgatt atctgcaata cgacaaaccg 720  
 tactatatgt tcaacgcggg taacaagaac agctacatta aactgaagaa agatagcccc 780  
 gtgggtgaaa tcctgggtcc gcgtggcagc gttatgacca ccaacatcta tctgaacagc 840  
 agcctgtacc gtggcgagaa gttcatcatt cgtcgtaaaa gcaacagcca gagcattaac 900  
 gacgatatcg tgcgtaacga agactacatt tatctggatt tctttaacct gaaccaagag 960  
 tggcgtgttt acacctacaa gtacttcaag aaagaggaag agaagctgtt cctggcgccg 1020  
 atcagcgaca gcgatgaatt ctacaacacc atccaaatca aggaatacga cgagcagggt 1080  
 accaacagct gccaaactgct gttcaagaaa gacgaagaga gcaccgatga aatcggctctg 1140  
 atcggcattc accgtttcta cgagagcggc atcgtgttcg aagagtacaa ggattacttc 1200  
 tgcacagca agtggtatct gaaagagggt aagcgtaaac cgtacaacct gaaactgggc 1260  
 tgcaactggc aatttatcc ggtggatgat ggctggggtg aacgtccgct gtaa 1314

<210> 3

<211> 1311

<212> PRT

<213> 人工序列

<220>

<223> 具有工程设计的激活位点的全长无活性TriRecABTox

<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 Gln			
	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 Ala Lys Thr Phe 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 Ile Glu Gly Arg Asp Lys  
 435 440 445  
 Gly Tyr Asn Lys Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp  
 450 455 460  
 Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn  
 465 470 475 480  
 Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu  
 485 490 495  
 Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe  
 500 505 510  
 Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile  
 515 520 525  
 Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly  
 530 535 540  
 Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala  
 545 550 555 560  
 Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val  
 565 570 575  
 Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser  
 580 585 590  
 Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu  
 595 600 605  
 Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu  
 610 615 620  
 Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr  
 625 630 635 640  
 Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe  
 645 650 655  
 Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile  
 660 665 670  
 Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr  
 675 680 685  
 Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser  
 690 695 700  
 Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn

705	710	715	720
Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met			
	725	730	735
Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn			
	740	745	750
Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe			
	755	760	765
Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala			
	770	775	780
Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu			
785	790	795	800
Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp			
	805	810	815
Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly			
	820	825	830
Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr			
	835	840	845
Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln			
	850	855	860
Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Lys Asn Ile Ile Asn Thr			
865	870	875	880
Ser Ile Leu Asn Leu Arg Tyr Glu Ser Asn His Leu Ile Asp Leu Ser			
	885	890	895
Arg Tyr Ala Ser Lys Ile Asn Ile Gly Ser Lys Val Asn Phe Asp Pro			
	900	905	910
Ile Asp Lys Asn Gln Ile Gln Leu Phe Asn Leu Glu Ser Ser Lys Ile			
	915	920	925
Glu Val Ile Leu Lys Asn Ala Ile Val Tyr Asn Ser Met Tyr Glu Asn			
	930	935	940
Phe Ser Thr Ser Phe Trp Ile Arg Ile Pro Lys Tyr Phe Asn Ser Ile			
945	950	955	960
Ser Leu Asn Asn Glu Tyr Thr Ile Ile Asn Cys Met Glu Asn Asn Ser			
	965	970	975
Gly Trp Lys Val Ser Leu Asn Tyr Gly Glu Ile Ile Trp Thr Leu Gln			
	980	985	990
Asp Thr Gln Glu Ile Lys Gln Arg Val Val Phe Lys Tyr Ser Gln Met			
	995	1000	1005
Ile Asn Ile Ser Asp Tyr Ile Asn Arg Trp Ile Phe Val Thr Ile			
1010	1015	1020	

Thr Asn Asn Arg Leu Asn Asn Ser Lys Ile Tyr Ile Asn Gly Arg	1025	1030	1035
Leu Ile Asp Gln Lys Pro Ile Ser Asn Leu Gly Asn Ile His Ala	1040	1045	1050
Ser Asn Asn Ile Met Phe Lys Leu Asp Gly Cys Arg Asp Thr His	1055	1060	1065
Arg Tyr Ile Trp Ile Lys Tyr Phe Asn Leu Phe Asp Lys Glu Leu	1070	1075	1080
Asn Glu Lys Glu Ile Lys Asp Leu Tyr Asp Asn Gln Ser Asn Ser	1085	1090	1095
Gly Ile Leu Lys Asp Phe Trp Gly Asp Tyr Leu Gln Tyr Asp Lys	1100	1105	1110
Pro Tyr Tyr Met Phe Asn Ala Gly Asn Lys Asn Ser Tyr Ile Lys	1115	1120	1125
Leu Lys Lys Asp Ser Pro Val Gly Glu Ile Leu Gly Pro Arg Gly	1130	1135	1140
Ser Val Met Thr Thr Asn Ile Tyr Leu Asn Ser Ser Leu Tyr Arg	1145	1150	1155
Gly Glu Lys Phe Ile Ile Arg Arg Lys Ser Asn Ser Gln Ser Ile	1160	1165	1170
Asn Asp Asp Ile Val Arg Asn Glu Asp Tyr Ile Tyr Leu Asp Phe	1175	1180	1185
Phe Asn Leu Asn Gln Glu Trp Arg Val Tyr Thr Tyr Lys Tyr Phe	1190	1195	1200
Lys Lys Glu Glu Glu Lys Leu Phe Leu Ala Pro Ile Ser Asp Ser	1205	1210	1215
Asp Glu Phe Tyr Asn Thr Ile Gln Ile Lys Glu Tyr Asp Glu Gln	1220	1225	1230
Gly Thr Asn Ser Cys Gln Leu Leu Phe Lys Lys Asp Glu Glu Ser	1235	1240	1245
Thr Asp Glu Ile Gly Leu Ile Gly Ile His Arg Phe Tyr Glu Ser	1250	1255	1260
Gly Ile Val Phe Glu Glu Tyr Lys Asp Tyr Phe Cys Ile Ser Lys	1265	1270	1275
Trp Tyr Leu Lys Glu Val Lys Arg Lys Pro Tyr Asn Leu Lys Leu	1280	1285	1290
Gly Cys Asn Trp Gln Phe Ile Pro Val Asp Asp Gly Trp Gly Glu	1295	1300	1305
Arg Pro Leu			

1310  
<210> 4  
<211> 3936  
<212> DNA  
<213> 人工序列  
<220>  
<223> 具有工程设计的激活位点的全长无活性TriRecABTox,  
对于在大肠杆菌(E. coli)中表达密码子优化的序列  
<400> 4  
atgccgttcg tgaataagca gttcaactac aaagatccgg ttaatggcgt ggacatcgcg 60  
tacatcaaaa tcccgaatgc gggtcagatg cagccggtga aggcgttcaa aatccacaac 120  
aaaatttggg ttatcccgga gcgtgacacc tttaccaacc cggaggaagg tgatctgaac 180  
ccgccgccgg aagcgaaaca agtgccggtt agctactatg acagcaccta tctgagcacc 240  
gacaacgaga aggataacta cctgaagggc gtgaccaaac tgttcgaacg tatctacagc 300  
accgatctgg gtcgtatget gctgaccagc attgttcgtg gcaccccggt ttgggggtggc 360  
agcaccatcg acaccgaact gaaagtgatt gataccaact gcattaacgt tatccagccg 420  
gatggtagct accgtagcga ggaactgaac ctggtgatca ttggcccagag cgcggacatc 480  
attcagtttg agtgcaagag cttcggtcac gaagttctga acctgaccgg taacggttac 540  
ggcagcacc aatataatccg tttcagcccg gatttcacct ttggcttcga ggaaagcctg 600  
gaagtggaca ccaaccgct gctgggtgcg ggcaagtttg cgaccgacc ggcggttacc 660  
ctggcgcacc agctgatcca tgcgggtcac cgtctgtacg gcattgcgat caaccggaac 720  
cgtgtgttca aagttaacac caacgcgtac tatgagatga gcggtctgga agtgagcttt 780  
gaggaactgc gtacctcgg tggccacgac gcgaagttta tcgatagcct gcaggagaac 840  
gaattccgctc tgtactacta caacaagttc aaggacatcg cgagcaccct gaacaaggcg 900  
aaaagcattg tgggtaccac cgcgagcctg caatacatga agaacgtttt caaggagaag 960  
tacctgctga gcgaagatac cagcggcaag tttagcgtgg acaagctgaa attcgataag 1020  
ctgtataaaa tgctgaccga gatctacacc gaagataact tcgtgaagtt ctttaaagtt 1080  
ctgaacgcga aaacctttct gaacttcgac aaggcggttt ttaaattaa catcgtgccg 1140  
aaggtttaact acaccatcta tgatggtttc aacctgcgta acaccaacct ggcggcgaac 1200  
tttaacggcc agaacaccga gattaacaac atgaacttta ccaagctgaa aaacttcacc 1260  
ggtctgtttg aattctataa actgctgtgc gtgcgtggca tcattaccag caagaccaa 1320  
agcctgatcg aaggtcgtga caagggttac acaaagcgc tgaacgatct gtgcattaaa 1380  
gttaacaact gggacctgtt ctttagcccc agcaggaca acttcaccaa cgatctgaac 1440  
aagggcgagg aatcaccag cgacaccaac attgaagcgg cggaggaaaa catcagcctg 1500  
gatctgattc agcaatatta cctgaccttt aacttcgaca acgagccgga aaacattagc 1560  
atcgagaacc tgagcagcga catcattggt cagctggagc tgatgccgaa catcgaacgt 1620  
ttcccgaacg gcaagaaata cgaactggat aatatacca tgttccacta cctgcgtgcg 1680  
caagagtttg aacacggcaa gagccgtatt gcgctgacca acagcgtgaa cgaggcgtg 1740  
ctgaaccgca gccgtgttta taccttcttt agcagcgact acgtgaagaa agttaacaaa 1800

gcgaccgagg cggcgatgtt cctgggttgg gtggaacagc tggtttacga ctttaccgat 1860  
gaaaccagcg aggtgagcac caccgacaaa attgcggata tcaccatcat tatcccgtat 1920  
atcgggccgg cgctgaacat tggcaacatg ctgtacaagg acgattttgt ggggtgcgctg 1980  
atcttcagcg gcgcggttat cctgctggag ttattccgg aaattgcgat cccggtgctg 2040  
ggtacctttg cgctggttag ctacatcgcg aacaaggcgc tgaccgttca aaccattgat 2100  
aacgcgctga gcaagcgtaa cgagaaatgg gacgaagtgt ataaatacat cgttaccaac 2160  
tggctggcga aggttaacac ccagattgac ctgatccgta agaaaatgaa agaggcgctg 2220  
gaaaaccaag cggaggcgac caaggcgatt atcaactatc agtacaacca atacaccgag 2280  
gaagagaaaa acaacattaa cttcaacatc gacgatctga gcagcaagct gaacgaaagc 2340  
atcaacaaaag cgatgattaa catcaacaag tttctgaacc agtgcagcgt gagctatctg 2400  
atgaacagca tgattccgta cgggtttaag cgtctggagg acttcgatgc gagcctgaag 2460  
gacgcgctgc tgaatatata ctacgataac cgtgggtacc tgattggcca agtggaccgt 2520  
ctgaaggata aagttaacaa caccctgagc accgatatcc cgttccagct gagcaaatat 2580  
gtggacaacc aacgtctgct gagcaccttt accgagtaca tcaagaacat tatcaacacc 2640  
agcattctga acctgcgta tgaagcaac cacctgatcg acctgagccg ttacgcgagc 2700  
aagattaaca tcggtagcaa agttaacttc gaccgatcg ataaaaacca gattcaactg 2760  
tttaacctgg agagcagcaa gattgaagtg atcctgaaaa acgcgatcgt ttacaacagc 2820  
atgtatgaga actttagcac cagcttctgg attcgtatcc cgaaatattt caacagcatt 2880  
agcctgaaca acgagtacac cattatcaac tgcatggaaa acaacagcgg ttggaaggtg 2940  
agcctgaact acggcgagat tatctggacc ctgcaggaca cccaagaaat caagcagcgt 3000  
gtggttttca agtacagcca aatgatcaac atcagcgatt acattaaccg ttgatcttt 3060  
gttaccatta ccaacaaccg tctgaacaac agcaaaattt acatcaacgg tcgtctgatc 3120  
gaccagaagc cgattagcaa cctgggcaac atccacgcga gcaacaacat tatgttcaag 3180  
ctggacgggt gccgtgatac ccaccgttat atttgatca agtacttcaa cctgttcgat 3240  
aaggagctga acgagaagga aatcaaagac ctgtatgata accagagcaa cagcggattt 3300  
ctgaaagact tctggggcga ttacctgcaa tatgacaagc cgtattacat gtttaacgcg 3360  
ggtaacaaga acagctacat caaactgaag aaagatagcc cgggtgggtga aattctgggt 3420  
ccgcgtggca gcgttatgac caccaacatc tatctgaaca gcagcctgta ccgtggcgaa 3480  
aagttcatta tccgtcgtaa aagcaacagc cagagcatca acgacgatat tgtgcgtaac 3540  
gaggactata tctacctgga tttctttaac ctgaaccaag aatggcgtgt ttacacctac 3600  
aagtacttca agaaagaaga ggaaaagctg tttctggcgc cgattagcga cagcgatgaa 3660  
ttctataaca ccattcagat caaagagtac gacgaacagg gtaccaacag ctgccaactg 3720  
ctgtttaaga aagacgagga aagcaccgat gagatcggtc tgattggcat ccaccgtttt 3780  
tacgaaagcg gcatcgtgtt cgaggaatac aaggattact tctgcatcag caagtggat 3840  
ctgaaagagg ttaagcgtaa accgtacaac ctgaaactgg gctgcaactg gcaatttatt 3900  
ccggtggatg atggctgggg tgaacgtccg ctgtaa 3936

<210> 5

<211> 1311

<212> PRT

<213> 人工序列

<220>

<223> 具有工程设计的激活位点的全长无活性TreRecABTox

<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
Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys		
305	310	315
Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu		
	325	330
Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp		
	340	350
Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr Leu Asn		
	355	365
Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr		
	375	380
Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn		
385	390	395
Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu		
	405	410
Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg		
	420	430
Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Ile Glu Gly Arg Asp Lys		
	435	445
Gly Tyr Asn Lys Ala Leu Asn Asp Leu Cys Ile Lys Val Asn Asn Trp		
	450	460
Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn		
465	470	480
Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala Glu Glu		
	485	495
Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe Asn Phe		
	500	510
Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser Asp Ile		
	515	525
Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro Asn Gly		
	530	540
Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu Arg Ala		
545	550	560
Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn Ser Val		
	565	575
Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser		
	580	590

Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu  
 595 600 605  
 Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu  
 610 615 620  
 Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr  
 625 630 635 640  
 Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe  
 645 650 655  
 Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile  
 660 665 670  
 Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr  
 675 680 685  
 Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser  
 690 695 700  
 Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn  
 705 710 715 720  
 Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met  
 725 730 735  
 Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn  
 740 745 750  
 Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe  
 755 760 765  
 Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala  
 770 775 780  
 Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu  
 785 790 795 800  
 Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp  
 805 810 815  
 Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly  
 820 825 830  
 Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr  
 835 840 845  
 Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln  
 850 855 860  
 Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Lys Asn Ile Ile Asn Thr  
 865 870 875 880  
 Ser Ile Leu Asn Leu Arg Tyr Glu Ser Asn His Leu Ile Asp Leu Ser  
 885 890 895  
 Arg Tyr Ala Ser Lys Ile Asn Ile Gly Ser Lys Val Asn Phe Asp Pro

900					905					910					
Ile	Asp	Lys	Asn	Gln	Ile	Gln	Leu	Phe	Asn	Leu	Glu	Ser	Ser	Lys	Ile
915					920					925					
Glu	Val	Ile	Leu	Lys	Asn	Ala	Ile	Val	Tyr	Asn	Ser	Met	Tyr	Glu	Asn
930					935					940					
Phe	Ser	Thr	Ser	Phe	Trp	Ile	Arg	Ile	Pro	Lys	Tyr	Phe	Asn	Ser	Ile
945					950					955					960
Ser	Leu	Asn	Asn	Glu	Tyr	Thr	Ile	Ile	Asn	Cys	Met	Glu	Asn	Asn	Ser
965					970					975					
Gly	Trp	Lys	Val	Ser	Leu	Asn	Tyr	Gly	Glu	Ile	Ile	Trp	Thr	Leu	Gln
980					985					990					
Asp	Thr	Gln	Glu	Ile	Lys	Gln	Arg	Val	Val	Phe	Lys	Tyr	Ser	Gln	Met
995					1000					1005					
Ile	Asn	Ile	Ser	Asp	Tyr	Ile	Asn	Arg	Trp	Ile	Phe	Val	Thr	Ile	
1010					1015					1020					
Thr	Asn	Asn	Arg	Leu	Asn	Asn	Ser	Lys	Ile	Tyr	Ile	Asn	Gly	Arg	
1025					1030					1035					
Leu	Ile	Asp	Gln	Lys	Pro	Ile	Ser	Asn	Leu	Gly	Asn	Ile	His	Ala	
1040					1045					1050					
Ser	Asn	Asn	Ile	Met	Phe	Lys	Leu	Asp	Gly	Cys	Arg	Asp	Thr	His	
1055					1060					1065					
Arg	Tyr	Ile	Trp	Ile	Lys	Tyr	Phe	Asn	Leu	Phe	Asp	Lys	Glu	Leu	
1070					1075					1080					
Asn	Glu	Lys	Glu	Ile	Lys	Asp	Leu	Tyr	Asp	Asn	Gln	Ser	Asn	Ser	
1085					1090					1095					
Gly	Ile	Leu	Lys	Asp	Phe	Trp	Gly	Asp	Tyr	Leu	Gln	Tyr	Asp	Lys	
1100					1105					1110					
Pro	Tyr	Tyr	Met	Phe	Asn	Ala	Gly	Asn	Lys	Asn	Ser	Tyr	Ile	Lys	
1115					1120					1125					
Leu	Lys	Lys	Asp	Ser	Pro	Val	Gly	Glu	Ile	Leu	Gly	Pro	Arg	Gly	
1130					1135					1140					
Ser	Val	Met	Thr	Thr	Asn	Ile	Tyr	Leu	Asn	Ser	Ser	Leu	Tyr	Arg	
1145					1150					1155					
Gly	Glu	Lys	Phe	Ile	Ile	Arg	Arg	Lys	Ser	Asn	Ser	Gln	Ser	Ile	
1160					1165					1170					
Asn	Asp	Asp	Ile	Val	Arg	Asn	Glu	Asp	Tyr	Ile	Tyr	Leu	Asp	Phe	
1175					1180					1185					
Phe	Asn	Leu	Asn	Gln	Glu	Trp	Arg	Val	Tyr	Thr	Tyr	Lys	Tyr	Phe	
1190					1195					1200					

Lys Lys Glu Glu Glu Lys Leu Phe Leu Ala Pro Ile Ser Asp Ser  
 1205 1210 1215  
 Asp Glu Phe Tyr Asn Thr Ile Gln Ile Lys Glu Tyr Asp Glu Gln  
 1220 1225 1230  
 Gly Thr Asn Ser Cys Gln Leu Leu Phe Lys Lys Asp Glu Glu Ser  
 1235 1240 1245  
 Thr Asp Glu Ile Gly Leu Ile Gly Ile His Arg Phe Tyr Glu Ser  
 1250 1255 1260  
 Gly Ile Val Phe Glu Glu Tyr Lys Asp Tyr Phe Cys Ile Ser Lys  
 1265 1270 1275  
 Trp Tyr Leu Lys Glu Val Lys Arg Lys Pro Tyr Asn Leu Lys Leu  
 1280 1285 1290  
 Gly Cys Asn Trp Gln Phe Ile Pro Val Asp Asp Gly Trp Gly Glu  
 1295 1300 1305

Arg Pro Leu  
 1310

<210> 6

<211> 437

<212> PRT

<213> 人工序列

<220>

<223> Hc/TAB2.1, 优化的环360

<400> 6

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

Phe Lys Tyr Ser Gln Met Ile Asn Ile Ser Asp Tyr Ile Asn Arg Trp  
 130 135 140  
 Ile Phe Val Thr Ile Thr Asn Asn Arg Leu Asn Asn Ser Lys Ile Tyr  
 145 150 155 160  
 Ile Asn Gly Arg Leu Ile Asp Gln Lys Pro Ile Ser Asn Leu Gly Asn  
 165 170 175  
 Ile His Ala Ser Asn Asn Ile Met Phe Lys Leu Asp Gly Cys Arg Asp  
 180 185 190  
 Thr His Arg Tyr Ile Trp Ile Lys Tyr Phe Asn Leu Phe Asp Lys Glu  
 195 200 205  
 Leu Asn Glu Lys Glu Ile Lys Asp Leu Tyr Asp Asn Gln Ser Asn Ser  
 210 215 220  
 Gly Ile Leu Lys Asp Phe Trp Gly Asp Tyr Leu Gln Tyr Asp Lys Pro  
 225 230 235 240  
 Tyr Tyr Met Phe Asn Ala Gly Asn Lys Asn Ser Tyr Ile Lys Leu Lys  
 245 250 255  
 Lys Asp Ser Pro Val Gly Glu Ile Leu Gly Pro Arg Gly Ser Val Met  
 260 265 270  
 Thr Thr Asn Ile Tyr Leu Asn Ser Ser Leu Tyr Arg Gly Glu Lys Phe  
 275 280 285  
 Ile Ile Arg Arg Lys Ser Asn Ser Gln Ser Ile Asn Asp Asp Ile Val  
 290 295 300  
 Arg Asn Glu Asp Tyr Ile Tyr Leu Asp Phe Phe Asn Leu Asn Gln Glu  
 305 310 315 320  
 Trp Arg Val Tyr Thr Tyr Lys Tyr Phe Lys Lys Glu Glu Glu Lys Leu  
 325 330 335  
 Phe Leu Ala Pro Ile Ser Asp Ser Asp Glu Phe Tyr Asn Thr Ile Gln  
 340 345 350  
 Ile Lys Glu Tyr Asp Glu Gln Pro Thr Tyr Ser Cys Gln Leu Leu Phe  
 355 360 365  
 Lys Lys Asp Glu Glu Ser Thr Asp Glu Ile Gly Leu Ile Gly Ile His  
 370 375 380  
 Arg Phe Tyr Glu Ser Gly Ile Val Phe Glu Glu Tyr Lys Asp Tyr Phe  
 385 390 395 400  
 Cys Ile Ser Lys Trp Tyr Leu Lys Glu Val Lys Arg Lys Pro Tyr Asn  
 405 410 415  
 Leu Lys Leu Gly Cys Asn Trp Gln Phe Ile Pro Val Asp Asp Gly Trp  
 420 425 430  
 Gly Glu Arg Pro Leu

435

&lt;210&gt; 7

&lt;211&gt; 1314

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; Hc/TAB2.1, 对于大肠杆菌(E. coli)中的表达进行密码子优化

&lt;400&gt; 7

```

aagaacatta tcaacaccag catcctgaac ctgcgctacg agagcaacca cctgatcgac 60
ctgagccgct acgcgagcaa gattaacatc ggtagcaagg tgaactttga cccgattgat 120
aaaaaccaga tccaactggt caacctggaa agcagcaaga tcgaagtgat tctgaaaaac 180
gcgattgttt ataacagcat gtacgaaaac ttcagcacca gcttttggat ccgtattccg 240
aagtatttta acagcatcag cctgaacaac gaatacacca tcattaactg catggagaac 300
aacagcgggt ggaaagtgag cctgaactac ggcgaaatca tttggaccct gcaggacacc 360
caagagatca agcagcgtgt ggttttcaag tacagccaaa tgatcaacat cagcgattac 420
atcaaccggt ggattttcgt taccatcacc aacaaccgtc tgaacaacag caagatctac 480
attaacggtc gtctgattga ccagaaaccg atcagcaacc tgggcaacat tcacgcgagc 540
aacaacatca tgttcaagct ggacggttgc cgtgataccc accgttatat ctggattaag 600
tacttcaacc tgtttgataa agagctgaac gaaaaggaga ttaaagacct gtatgataac 660
cagagcaaca gcggtatcct gaaggacttt tggggcgatt atctgcaata cgacaaaccg 720
tactatatgt tcaacgcggg taacaagaac agctacatta aactgaagaa agatagcccc 780
gtgggtgaaa tcctgggtcc gcgtggcagc gttatgacca ccaacatcta tctgaacagc 840
agcctgtacc gtggcgagaa gttcatcatt cgtcgtaaaa gcaacagcca gagcattaac 900
gacgatatcg tgcgtaacga agactacatt tatctggatt tctttaacct gaaccaagag 960
tggcgtgttt acacctacia gtacttcaag aaagaggaag agaagctggt cctggcgccc 1020
atcagcgaca gcgatgaatt ctacaacacc atccaaatca aggaatacga cgagcagccg 1080
acctatagct gccaactgct gttcaagaaa gacgaagaga gcaccgatga aatcggctctg 1140
atcggcattc accgtttcta cgagagcggc atcgtgttcg aagagtacaa ggattacttc 1200
tgcatcagca agtggtatct gaaagagggt aagcgtaaac cgtacaacct gaaactgggc 1260
tgcaactggc aatttatcc ggtggatgat ggctggggtg aacgtccgct gtaa 1314

```

&lt;210&gt; 8

&lt;211&gt; 432

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; Hc/TAB2.1.1

&lt;400&gt; 8

Thr Ser Ile Leu Asn Leu Arg Tyr Glu Ser Asn His Leu Ile Asp Leu

1

5

10

15

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

	325	330	335
Ser Asp Ser Asp Glu Phe Tyr Asn Thr Ile Gln Ile Lys Glu Tyr Asp			
	340	345	350
Glu Gln Pro Thr Tyr Ser Cys Gln Leu Leu Phe Lys Lys Asp Glu Glu			
	355	360	365
Ser Thr Asp Glu Ile Gly Leu Ile Gly Ile His Arg Phe Tyr Glu Ser			
	370	375	380
Gly Ile Val Phe Glu Glu Tyr Lys Asp Tyr Phe Cys Ile Ser Lys Trp			
385	390	395	400
Tyr Leu Lys Glu Val Lys Arg Lys Pro Tyr Asn Leu Lys Leu Gly Cys			
	405	410	415
Asn Trp Gln Phe Ile Pro Val Asp Asp Gly Trp Gly Glu Arg Pro Leu			
	420	425	430

<210> 9

<211> 1299

<212> DNA

<213> 人工序列

<220>

<223> Hc/TAB2.1.1, 对于大肠杆菌(E. coli)中的表达进行密码子优化

<400> 9

```

accagcatcc tgaacctgcg ctacgagagc aaccacctga tgcacctgag ccgctacgcg 60
agcaagatta acatcggtag caaggtgaac tttgaccgga ttgataaaaa ccagatccaa 120
ctgttcaacc tggaaagcag caagatcgaa gtgattctga aaaacgcatg tgtttataac 180
agcatgtacg aaaacttcag caccagcttt tggatccgta ttccgaagta ttttaacagc 240
atcagcctga acaacgaata caccatcatt aactgcatgg agaacaacag cggttgaaa 300
gtgagcctga actacggcga aatcatttgg accctgcagg acaccaaga gatcaagcag 360
cgtgtggttt tcaagtacag ccaaatgatc aacatcagcg attacatcaa ccgttggatt 420
ttcgttacca tcaccaacaa ccgtctgaac aacagcaaga tctacattaa cggtcgtctg 480
attgaccaga aaccgatcag caacctgggc aacattcag cgagcaacaa catcatgttc 540
aagctggacg gttgccgtga taccaccgt tatactctgga ttaagtactt caacctgttt 600
gataaagagc tgaacgaaaa ggagattaaa gacctgtatg ataaccagag caacagcgg 660
atcctgaagg acttttgggg cgattatctg caatacgaca aaccgtacta tatgttcaac 720
gcgggtaaca agaacagcta cattaaactg aagaaagata gcccggtggg tgaatcctg 780
gggccgcgtg gcagcgttat gaccaccaac atctatctga acagcgcct gtaccgtggc 840
gagaagtcca tcattcgtcg taaaagcaac agccagagca ttaacgacga tctcgtcgt 900
aacgaagact acatttatct ggatttcttt aacctgaacc aagagtggcg tgtttacacc 960
tacaagtact tcaagaaaga ggaagagaag ctgttctctg cgccgatcag cgacagcgat 1020
gaattctaca acaccatcca aatcaaggaa tacgacgagc agccgacct tagctgcca 1080
ctgctgttca agaaagacga agagagcacc gatgaaatcg gtctgatcgg cattcacctg 1140

```

ttctacgaga gcgcatcgt gttcgaagag tacaaggatt acttctgcat cagcaagtgg 1200  
 tatctgaaag aggttaagcg taaacctac aacctgaaac tgggctgcaa ctggcaattt 1260  
 attccgggtg atgatggctg gggatgaactg ccgctgtaa 1299

<210> 10

<211> 424

<212> PRT

<213> 人工序列

<220>

<223> Hc/TAB2.1.3, 提高的溶解度

<400> 10

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

225	230	235	240
Ala Gly Asn Lys Asn Ser Tyr Ile Lys Leu Lys Lys Asp Ser Pro Val			
	245	250	255
Gly Glu Ile Leu Gly Pro Arg Gly Ser Val Met Thr Thr Asn Ile Tyr			
	260	265	270
Leu Asn Ser Ser Leu Tyr Arg Gly Glu Lys Phe Ile Ile Arg Arg Lys			
	275	280	285
Ser Asn Ser Gln Ser Ile Asn Asp Asp Ile Val Arg Asn Glu Asp Tyr			
	290	295	300
Ile Tyr Leu Asp Phe Phe Asn Leu Asn Gln Glu Trp Arg Val Tyr Thr			
305	310	315	320
Tyr Lys Tyr Phe Lys Lys Glu Glu Glu Lys Leu Phe Leu Ala Pro Ile			
	325	330	335
Ser Asp Ser Asp Glu Phe Tyr Asn Thr Ile Gln Ile Lys Glu Tyr Asp			
	340	345	350
Glu Gln Pro Thr Tyr Ser Cys Gln Leu Leu Phe Lys Lys Asp Glu Glu			
	355	360	365
Ser Thr Asp Glu Ile Gly Leu Ile Gly Ile His Arg Phe Asn Asn Lys			
	370	375	380
Asp Tyr Phe Cys Ile Ser Lys Trp Tyr Leu Lys Glu Val Lys Arg Lys			
385	390	395	400
Pro Tyr Asn Leu Lys Leu Gly Cys Asn Trp Gln Phe Ile Pro Val Asp			
	405	410	415
Asp Gly Trp Gly Glu Arg Pro Leu			
	420		

<210> 11

<211> 1275

<212> DNA

<213> 人工序列

<220>

<223> Hc/TAB2.1.3, 对于大肠杆菌(E. coli)中的表达进行密码子优化

<400> 11

```

accagcatcc tgaacctgcg ctacgagagc aaccacctga tgcacctgag ccgctacgcg 60
agcaagatta acatcggtag caagtgtaac tttgaccgga ttgataaaaa ccagatccaa 120
ctgttcaacc tggaaagcag caagatcgaa gtgattctga aaaacgcgat tgtttataac 180
agcatgtacg aaaacttcag caccagcttt tggatccgta ttccgaagta ttttaacagc 240
atcagcctga acaacgaata caccatcatt aactgcatgg agaacaacag cggttgaaa 300
gtgagcctga actacggcga aatcatttgg accctgcagg acaccaaga gatcaagcag 360
cgtgtggttt tcaagtacag ccaaattgat aacatcagcg attacatcaa ccgtttgatt 420

```

ttcgttacca tcaccaacaa ccgtctgaac aacagcaaga tctacattaa cggtcgtctg 480  
 attgaccaga aaccgatcag caacctgggc aacattcacg cgagcaacaa catcatgttc 540  
 aagctggacg gttgccgtga taccaccgt tatactgga ttaagtactt caacctgttt 600  
 gataaagagc tgaacgaaaa ggagattaaa gacctgtatg ataaccagag caacagcgg 660  
 atcctgaagg acttttgggg cgattatctg caatacgaca aaccgtacta tatgttcaac 720  
 gcgggtaaca agaacagcta cattaactg aagaaagata gcccggggg tgaatcctg 780  
 ggtccgcgtg gcagcggtat gaccaccaac atctatctga acagcagcct gtaccgtggc 840  
 gagaagttca tcattcgtcg taaaagcaac agccagagca ttaacgacga tatcgtgcgt 900  
 aacgaagact acatttatct ggatttctt aacctgaacc aagagtggcg tgtttacacc 960  
 tacaagtact tcaagaaaga ggaagagaag ctgttctgg cgccgatcag cgacagcgat 1020  
 gaattctaca acaccatcca aatcaaggaa tacgacgagc agccgacct tagctgcca 1080  
 ctgctgttca agaaagacga agagagcacc gatgaaatcg gtctgatcgg cattcacct 1140  
 ttcaacaaca aggattactt ctgcatcagc aagtggatc tgaagaggt taagcgtaaa 1200  
 ccgtacaacc tgaactggg ctgcaactgg caatttattc cggtgatga tggctgggg 1260  
 gaactccgc tgtaa 1275

<210> 12

<211> 1343

<212> PRT

<213> 人工序列

<220>

<223> BoNT/TAB2.1.3, 全长活性蛋白

<400> 12

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
Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr		
	165	170
Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe		
	180	185
Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu		
	195	200
Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Ala		
	210	220
Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn		
225	230	235
Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu		
	245	250
Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys		
	260	265
Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Tyr Asn		
	275	280
Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val		
	290	295
Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys		
305	310	315
Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp Lys Leu		
	325	330
Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr Glu Asp		
	340	345
Asn Phe Val Lys Phe Phe Lys Val Leu Asn Ala Lys Thr Phe Leu Asn		
	355	360
Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val Asn Tyr		
	370	375
Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala Ala Asn		
385	390	395
Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr Lys Leu		
	405	410
Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys Val Arg		
	420	425
Gly Ile Ile Thr Ser Lys Ala Gly Ala Gly Lys Ser Leu Val Pro Arg		
	435	440
		445

Gly Ser Ala Gly Ala Gly Ala Leu Asn Asp Leu Cys Ile Lys Val Asn  
 450 455 460  
 Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp  
 465 470 475 480  
 Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala Ala  
 485 490 495  
 Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu Thr Phe  
 500 505 510  
 Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn Leu Ser Ser  
 515 520 525  
 Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile Glu Arg Phe Pro  
 530 535 540  
 Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr Met Phe His Tyr Leu  
 545 550 555 560  
 Arg Ala Gln Glu Phe Glu His Gly Lys Ser Arg Ile Ala Leu Thr Asn  
 565 570 575  
 Ser Val Asn Glu Ala Leu Leu Asn Pro Ser Arg Val Tyr Thr Phe Phe  
 580 585 590  
 Ser Ser Asp Tyr Val Lys Lys Val Asn Lys Ala Thr Glu Ala Ala Met  
 595 600 605  
 Phe Leu Gly Trp Val Glu Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr  
 610 615 620  
 Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile  
 625 630 635 640  
 Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp  
 645 650 655  
 Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu  
 660 665 670  
 Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val  
 675 680 685  
 Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala  
 690 695 700  
 Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val  
 705 710 715 720  
 Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys  
 725 730 735  
 Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile  
 740 745 750  
 Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile

755	760	765
Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn		
770	775	780
Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser		
785	790	800
Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp		
805	810	815
Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn		
820	825	830
Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn		
835	840	845
Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp		
850	855	860
Asn Gln Arg Leu Leu Ser Thr Phe Thr Glu Tyr Ile Lys Asn Ile Ile		
865	870	875
Asn Thr Ser Ile Leu Asn Leu Arg Tyr Glu Ser Asn His Leu Ile Asp		
885	890	895
Leu Ser Arg Tyr Ala Ser Lys Ile Asn Ile Gly Ser Lys Val Asn Phe		
900	905	910
Asp Pro Ile Asp Lys Asn Gln Ile Gln Leu Phe Asn Leu Glu Ser Ser		
915	920	925
Lys Ile Glu Val Ile Leu Lys Asn Ala Ile Val Tyr Asn Ser Met Tyr		
930	935	940
Glu Asn Phe Ser Thr Ser Phe Trp Ile Arg Ile Pro Lys Tyr Phe Asn		
945	950	955
Ser Ile Ser Leu Asn Asn Glu Tyr Thr Ile Ile Asn Cys Met Glu Asn		
965	970	975
Asn Ser Gly Trp Lys Val Ser Leu Asn Tyr Gly Glu Ile Ile Trp Thr		
980	985	990
Leu Gln Asp Thr Gln Glu Ile Lys Gln Arg Val Val Phe Lys Tyr Ser		
995	1000	1005
Gln Met Ile Asn Ile Ser Asp Tyr Ile Asn Arg Trp Ile Phe Val		
1010	1015	1020
Thr Ile Thr Asn Asn Arg Leu Asn Asn Ser Lys Ile Tyr Ile Asn		
1025	1030	1035
Gly Arg Leu Ile Asp Gln Lys Pro Ile Ser Asn Leu Gly Asn Ile		
1040	1045	1050
His Ala Ser Asn Asn Ile Met Phe Lys Leu Asp Gly Cys Arg Asp		
1055	1060	1065

Thr His Arg Tyr Ile Trp	Ile Lys Tyr Phe Asn	Leu Phe Asp Lys
1070	1075	1080
Glu Leu Asn Glu Lys Glu	Ile Lys Asp Leu Tyr	Asp Asn Gln Ser
1085	1090	1095
Asn Ser Gly Ile Leu Lys	Asp Phe Trp Gly Asp	Tyr Leu Gln Tyr
1100	1105	1110
Asp Lys Pro Tyr Tyr Met	Phe Asn Ala Gly Asn	Lys Asn Ser Tyr
1115	1120	1125
Ile Lys Leu Lys Lys Asp	Ser Pro Val Gly Glu	Ile Leu Gly Pro
1130	1135	1140
Arg Gly Ser Val Met Thr	Thr Asn Ile Tyr Leu	Asn Ser Ser Leu
1145	1150	1155
Tyr Arg Gly Glu Lys Phe	Ile Ile Arg Arg Lys	Ser Asn Ser Gln
1160	1165	1170
Ser Ile Asn Asp Asp Ile	Val Arg Asn Glu Asp	Tyr Ile Tyr Leu
1175	1180	1185
Asp Phe Phe Asn Leu Asn	Gln Glu Trp Arg Val	Tyr Thr Tyr Lys
1190	1195	1200
Tyr Phe Lys Lys Glu Glu	Glu Lys Leu Phe Leu	Ala Pro Ile Ser
1205	1210	1215
Asp Ser Asp Glu Phe Tyr	Asn Thr Ile Gln Ile	Lys Glu Tyr Asp
1220	1225	1230
Glu Gln Gly Thr Asn Ser	Cys Gln Leu Leu Phe	Lys Lys Asp Glu
1235	1240	1245
Glu Ser Thr Asp Glu Ile	Gly Leu Ile Gly Ile	His Arg Phe Tyr
1250	1255	1260
Glu Ser Gly Ile Val Phe	Glu Glu Tyr Lys Asp	Tyr Phe Cys Ile
1265	1270	1275
Ser Lys Trp Tyr Leu Lys	Glu Val Lys Arg Lys	Pro Tyr Asn Leu
1280	1285	1290
Lys Leu Gly Cys Asn Trp	Gln Phe Ile Pro Val	Asp Asp Gly Trp
1295	1300	1305
Gly Glu Arg Pro Leu Val	Pro Arg Gly Ser Ala	Asn Ser Ser Ser
1310	1315	1320
Val Asp Lys Leu Trp Ser	His Pro Gln Phe Glu	Lys Leu Glu His
1325	1330	1335
His His His His His		
1340		

&lt;210&gt; 13

<211> 4032

<212> DNA

<213> 人工序列

<220>

<223> BoNT/TAB2.1.3, 用于在大肠杆菌 (E. coli) 中的表达的包含在pEt29 (a) 载体中的序列

<400> 13

```

atgccatttg tgaacaagca gtttaactat aaggacccgg tgaacggtgt ggatatcgcg 60
tatatcaaaa tcccgaatgc gggccagatg caaccagtcaggcgttcaa gattcataac 120
aagatttggg ttattccgga acgtgatacc ttcaccaatc cggagaagg cgacttaaac 180
ccgccgccag aagccaaaca agtgccggtg agctactatg atagcacgta tcttagcacc 240
gataatgaaa aagacaatta cctgaagggc gtgaccaagt tgttcgagcg catctacagt 300
accgacttag gccgcatggt gttgacgagc atcgttcgcg gtatcccgtt ctggggcggc 360
tcgaccattg ataccgagtt gaaagtcatt gacacgaact gtatcaatgt tatccaaccg 420
gacggcagtt atcgcagcga ggagttaaatt ttggtcatca tcggtccaag cgcagatatt 480
attcagttcg aatgcaagag cttcggccat gaggtcttga atttgacgcg caacggttac 540
ggcagcacc aatacatccg ctttagcccc gatttcacct ttggcttcga ggagagcttg 600
gaggtggaca ccaaccgct gttaggtgcc ggcaaattcg caaccgacc ggcagtgacg 660
ttggcgcacg cgttgattca tgcgggtcac cgcttatacg gtatcgcgat caatccgaat 720
cgcgtcttta aagtcaatac caacgcgtac tacgaaatga gcggcttaga ggttagcttt 780
gaagaattac gcaccttcgg tggccacgac gccaaattca tcgacagcct gcaggaaaat 840
gagttccgct tgtactatta caataaattc aaggacatcg cgagcacctt aaataaagca 900
aagagcattg tgggcaccac cgcaagcttg cagtacatga agaacgtatt taaggaaaa 960
tatttgttgt cggaggatac cagcgggaaa ttcagcgtcg ataagctgaa attcgacaaa 1020
ttgtataaaa tgctgaccga gatttacacc gaggataact tcgtcaagtt ttttaaggtg 1080
ttaaattgca agacctttt aaactttgat aaagcgggtg ttaaattaa tatcgtgccg 1140
aaggtgaatt acaccatcta cgatggtttc aatttacgca acacgaatct ggcggcgaat 1200
tttaattggc aaaacaccga aattaacaac atgaacttta cgaagttaa gaatttcacg 1260
ggcttattcg aattctacaa gttattatgc gtgcgcgca tcattaccag caaggcaggt 1320
gcgggcaagt ccttggttc gcgtggcagc gccggcgccg gcgctcaa tgatctgtgt 1380
attaaagtca ataactggga cctgttcttc agcccagcg aggataactt taccaacgac 1440
ttaaacaag gcgaggagat cacgagcgat acgaacatcg aggcggcgga ggaaaatatt 1500
agcctggacc tcattcagca gtactatctg acgttcaatt ttgacaatga gccggagaac 1560
atcagcattg aaaatctcag cagcgacatc atcggtcagt tggaaactgat gccgaacatt 1620
gaacgctttc cgaacggcaa aaaatatgaa ctggacaagt ataccatggt ccattactta 1680
cgcgcacagg aatttgagca cggcaagagc cgcattgcgc tgaccaatag cgttaacgag 1740
gccttgtaa atccgagccg tgtctacacg ttcttcagca gcgattatgt caaaaaagtg 1800
aacaaggcga ccgaagccg gatgtttttg ggctgggtcg agcaattggt ttacgatttt 1860
accgacgaaa ccagcgaggt gagcacgacc gacaaaattg cagatatcac catcatcatt 1920

```

ccgtacatcg gtccggcgct caatatcggc aatatgttat acaaggacga ctttgtgggc 1980  
 gcgctgatct ttagcggcgc ggttatctta ttagaattca tcccggagat cgcaatcccg 2040  
 gtcttgggca cctttgcggt ggtgagctat atcgcgaata aagtgctcac ggtccaaacc 2100  
 atcgataacg cgctcagcaa gcgtaatgag aaatgggacg aggtttataa gtatatcgtg 2160  
 accaactggg tagcaaaaagt caatacgcag atcgatctca tccgcaaaaa aatgaaagaa 2220  
 gccttggaaa atcaagcggg ggcaaccaa gccatcatta attaccagta taaccaatat 2280  
 accgaagaag aaaaaaacia tatcaacttc aatatcgatg atttgagcag caaactgaac 2340  
 gagagcatta acaaagcgat gattaacatc aacaagttct tgaatcaatg cagcgtgagc 2400  
 tatctcatga acagcatgat cccgtatggc gtcaaacgct tggaagattt tgacgccagc 2460  
 ctgaaagatg cgctcctcaa gtatatattat gacaaccgcg gcaccctcat tggccaggtg 2520  
 gaccgcttga aggataaagt gaacaatacgc ctacgacgg atatcccggt ccagctgagc 2580  
 aagtacgtcg acaaccagcg cttactgagc acctttaccg agtatatcaa gaacatcatt 2640  
 aataccagca tcctcaactt gcgctatgag agcaatcacc tgatcgacct cagccgctac 2700  
 gccagcaaga tcaacatcgg cagcaaggtc aatttcgacc cgatcgataa gaatcagatc 2760  
 caattgttta acctggaaag cagcaagatc gaggttatct tgaagaacgc gattgtgtac 2820  
 aacagcatgt atgagaactt tagcaccagc ttctggattc gtatcccgaa atatttcaac 2880  
 agcattagcc tgaacaacga gtacaccatt atcaactgca tggaaaacia cagcggttgg 2940  
 aaggtgagcc tgaactacgg cgagattatc tggaccctgc aggacacca agaaatcaag 3000  
 cagcgtgtgg ttttcaagta cagccaaatg atcaacatca gcgattacat taaccgttgg 3060  
 atctttgtta ccattaccaa caaccgtctg aacaacagca aaatttacet caacggctctg 3120  
 ctgatcgacc agaagccgat tagcaacctg ggcaacatcc acgcgagcaa caacattatg 3180  
 ttcaagctgg acggttgccg tgataccac cgttatattt ggatcaagta cttcaacctg 3240  
 ttcgataagg agctgaacga gaaggaaatc aaagacctgt atgataacca gagcaacagc 3300  
 ggtattctga aagacttctg gggcgattac ctgcaatatg acaagccgta ttacatgttt 3360  
 aacgcgggta acaagaacag ctacatcaa ctgaagaaag atagcccggg gggtgaaatt 3420  
 ctgggtccgc gtggcagcgt tatgaccacc aacatctatc tgaacagcag cctgtacctg 3480  
 ggcgaaaagt tcattatccg tcgtaaaagc aacagccaga gcatcaacga cgatattgtg 3540  
 cgtaacgagg actatatcta cctggatttc tttaacctga accaagaatg gcgtgtttac 3600  
 acctacaagt acttcaagaa agaagaggaa aagctgtttc tggcgccgat tagcgacagc 3660  
 gatgaattct ataacacat tcagatcaaa gactacgacg aacagggtac caacagctgc 3720  
 caactgctgt ttaagaaaga cgaggaaagc accgatgaga tcggtctgat tggcatccac 3780  
 cgtttttacg aaagcggcat cgtgttcgag gaatacaagg attacttctg catcagcaag 3840  
 tggatctga aagaggttaa gcgtaaaccg tacaacctga aactgggctg caactggcaa 3900  
 tttattccgg tggatgatgg ctggggtgaa cgtccactag tgccacgagg ttccgcaat 3960  
 tcgagctccg tcgacaagct ttggagccac ccgcagttcg aaaaactcga gcaccaccac 4020  
 caccaccact ga 4032

<210> 14

<211> 21

<212> PRT

<213> 人工序列

<220>

<223> hSytI肽

<400> 14

Gly Glu Gly Lys Glu Asp Ala Phe Ser Lys Leu Lys Glu Lys Phe Met

1

5

10

15

Asn Glu Leu His Lys

20

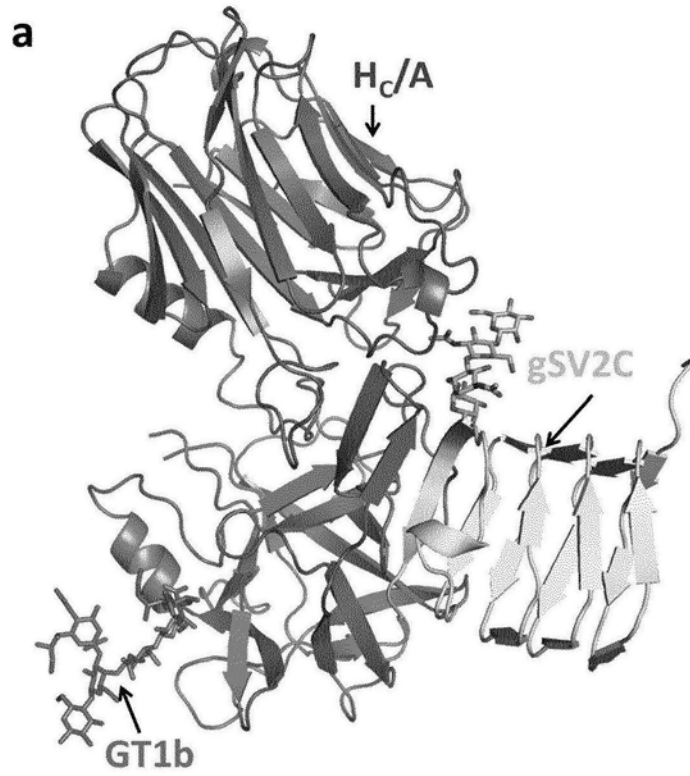


图1a

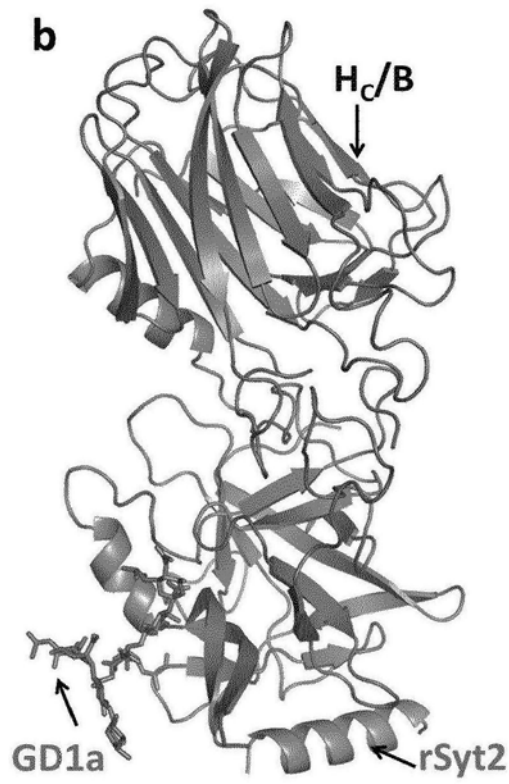


图1b





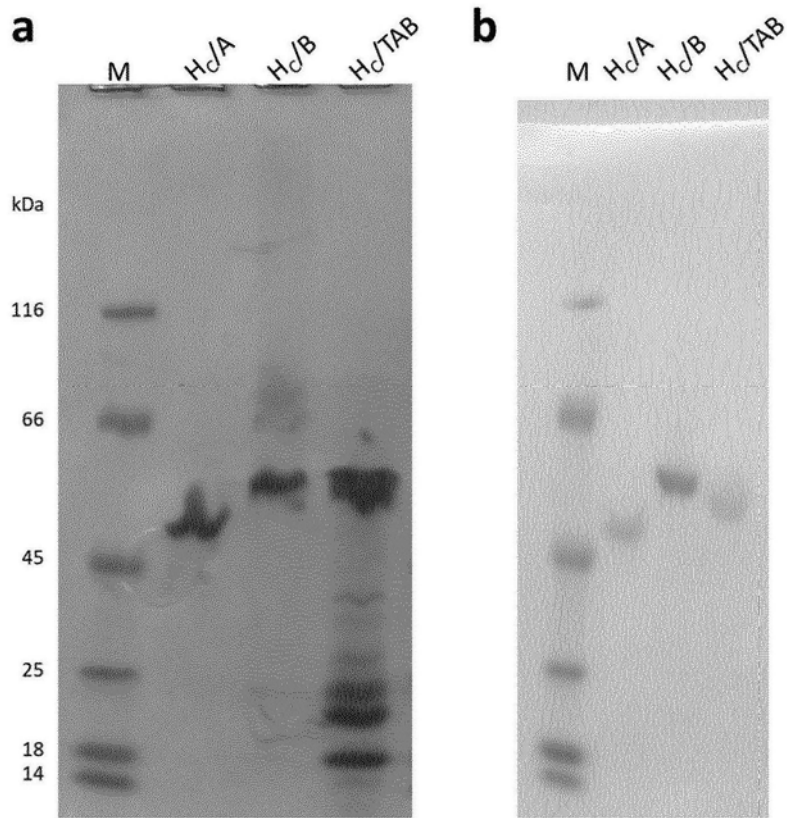


图3

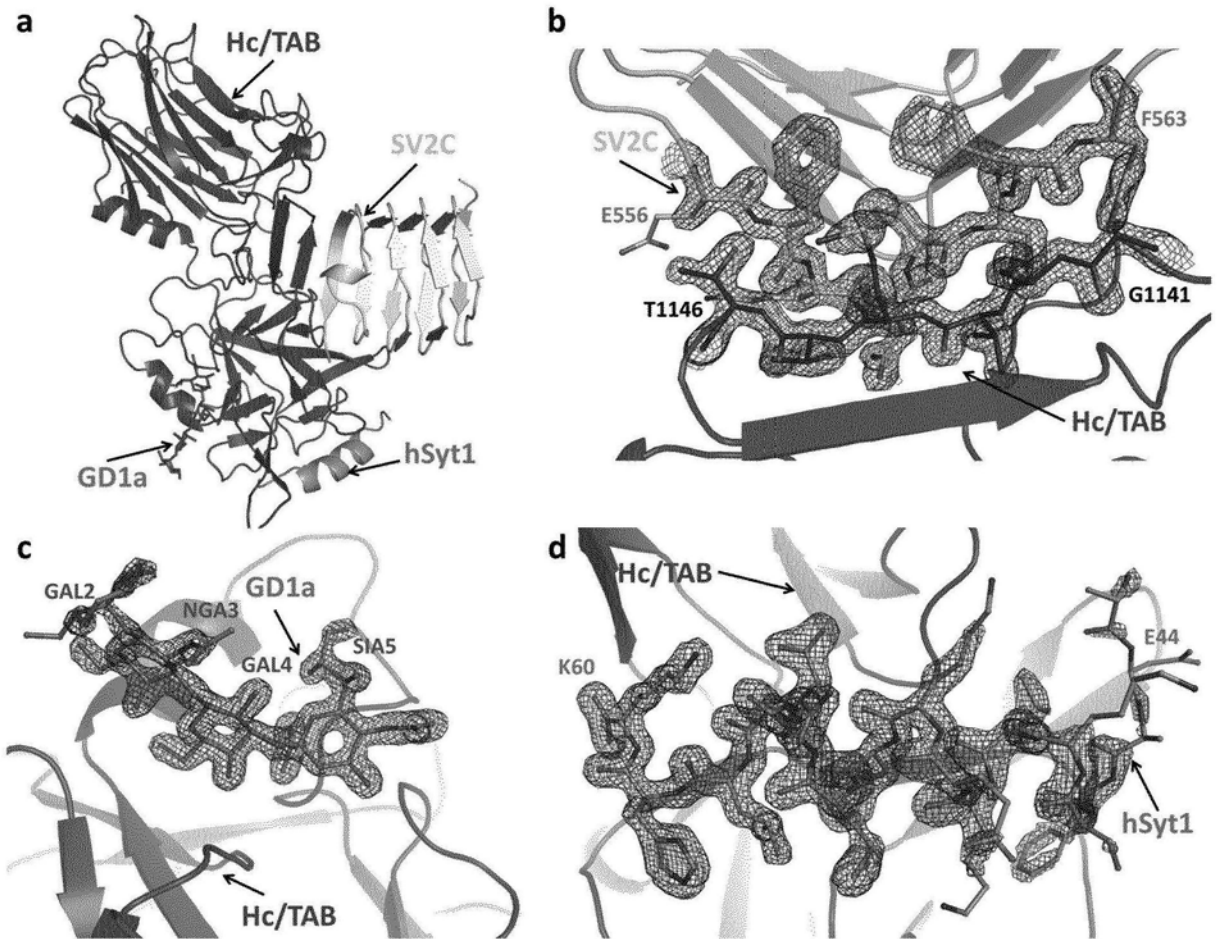


图4

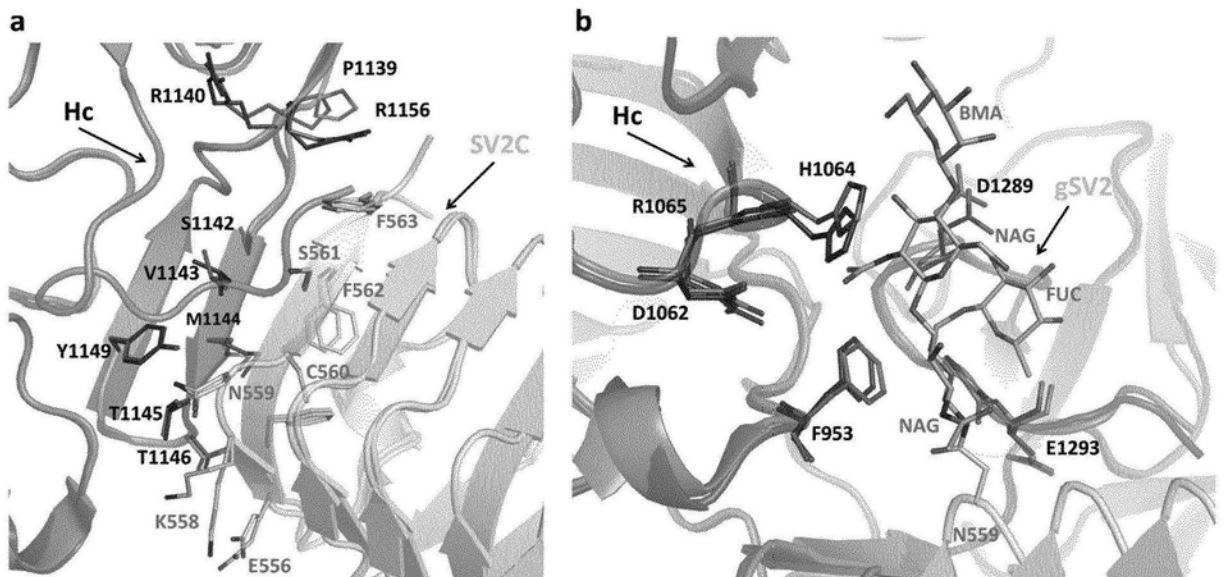


图5

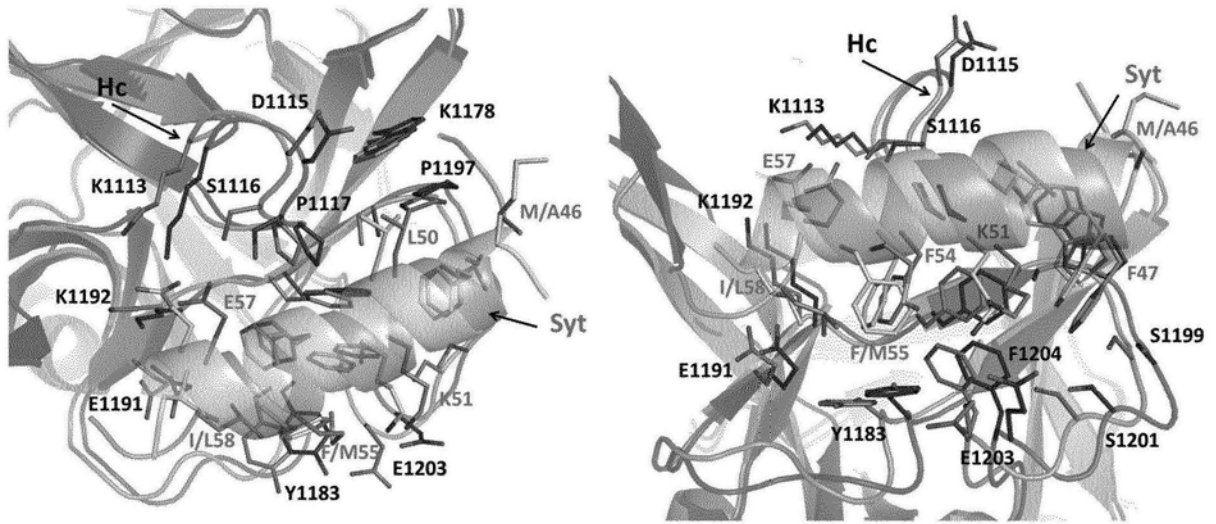


图6

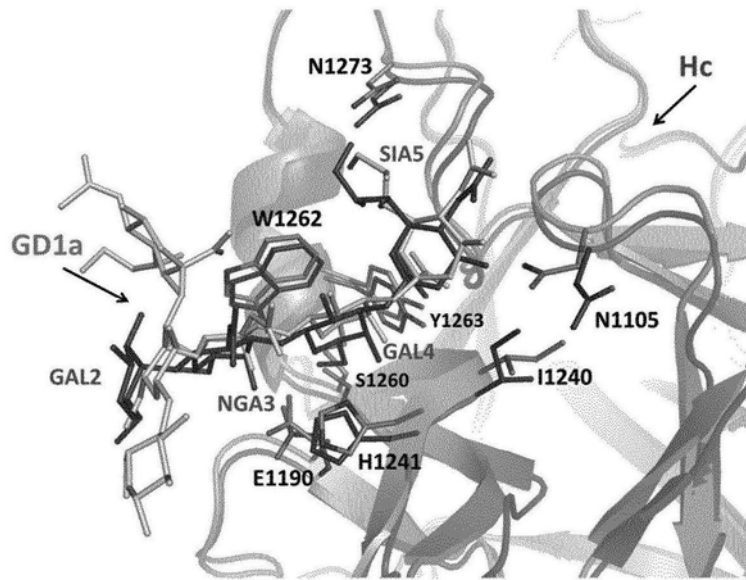


图7

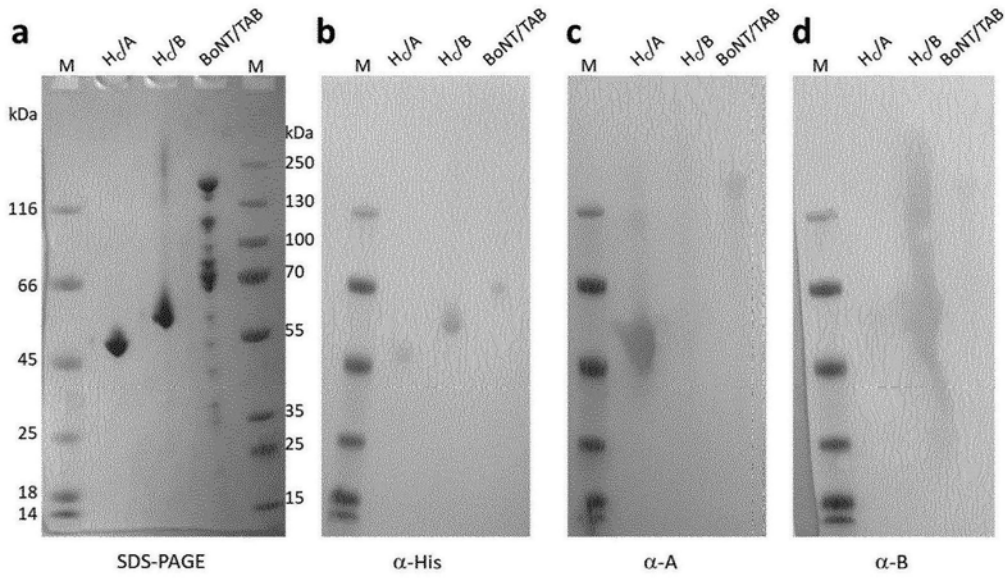


图8

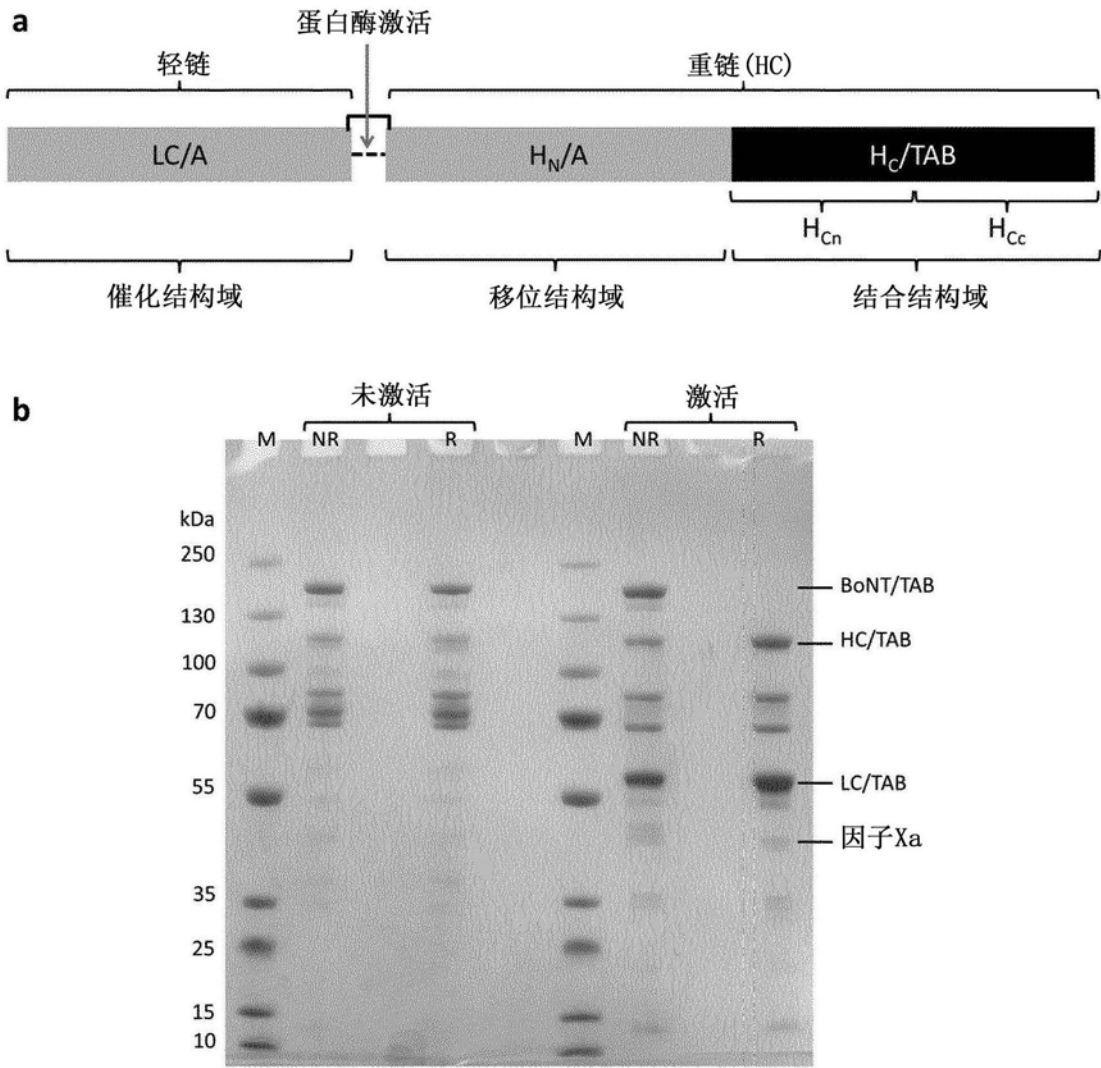


图9

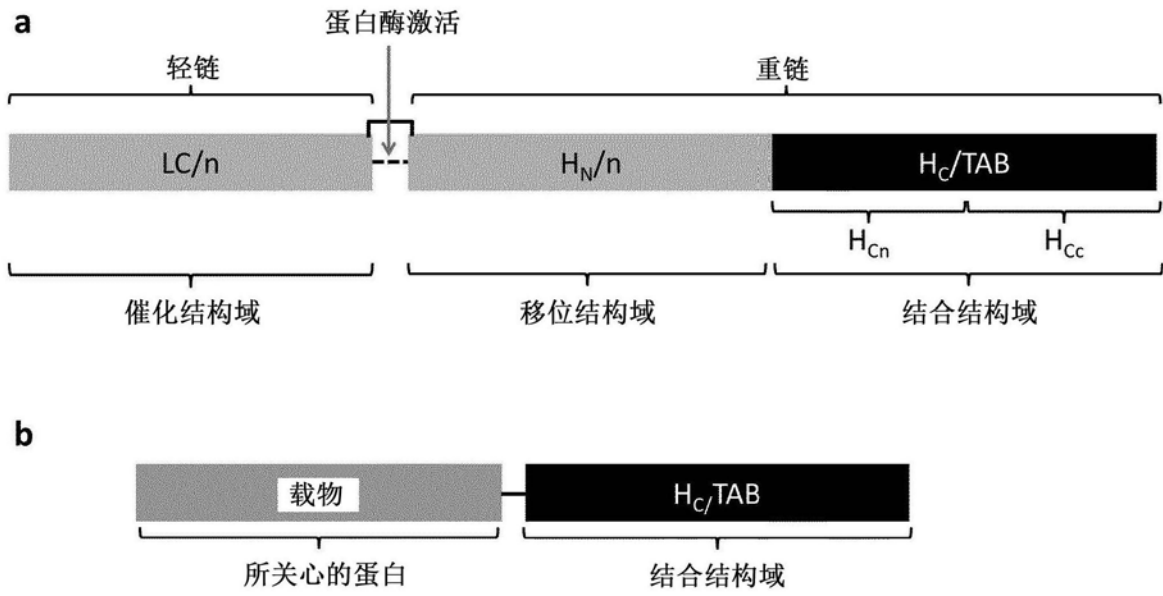


图10

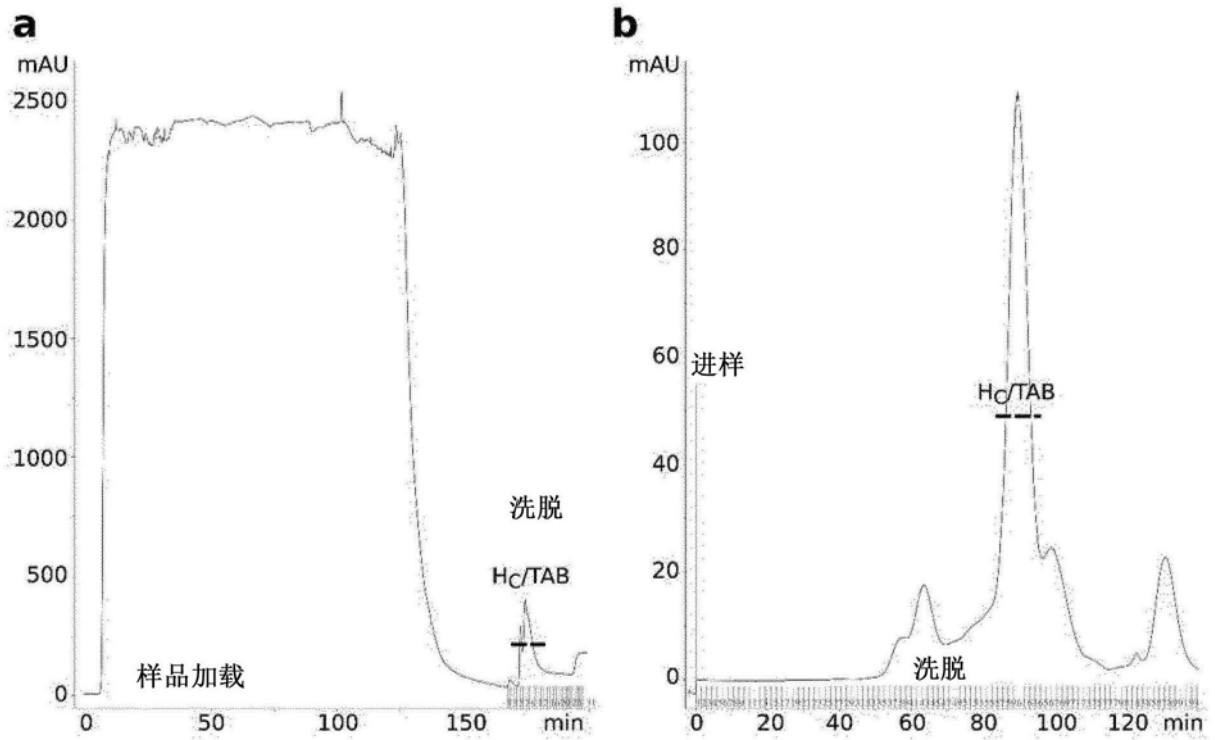


图11

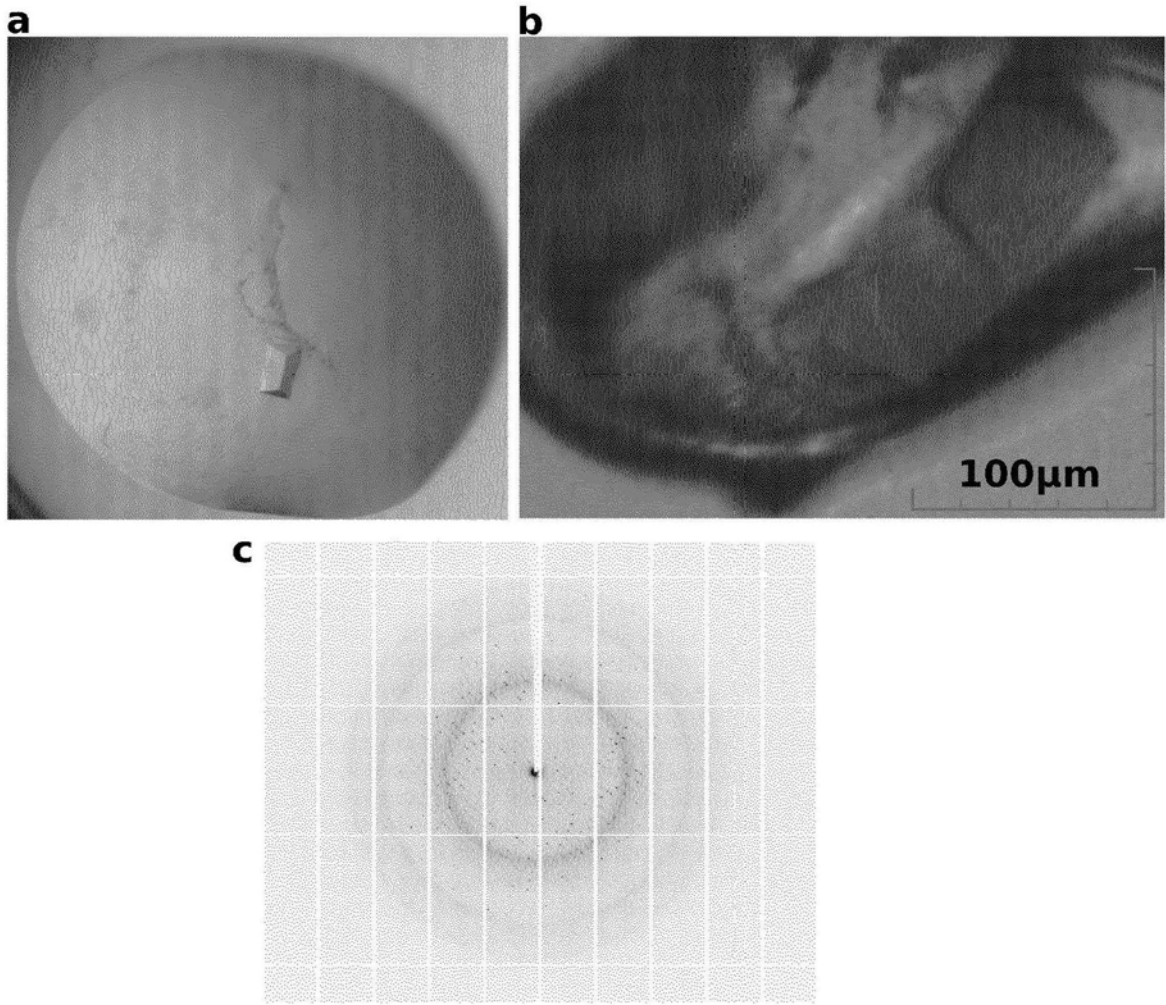


图12

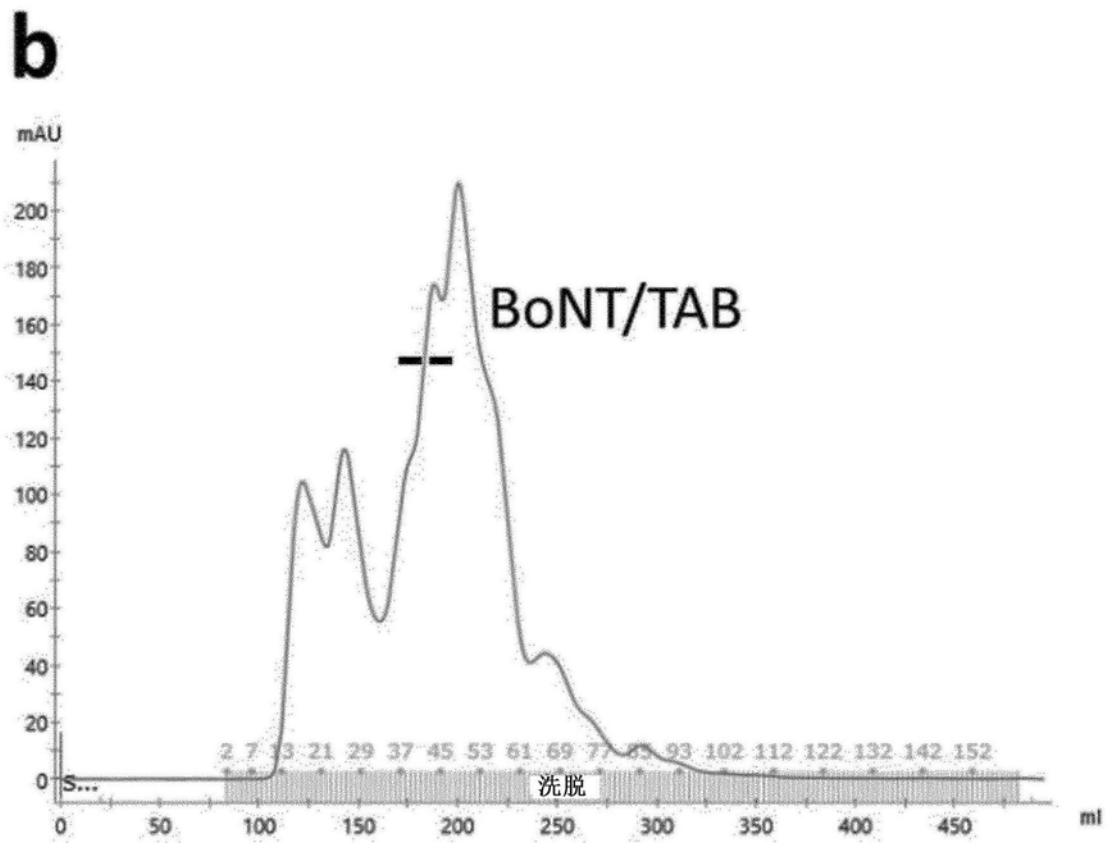
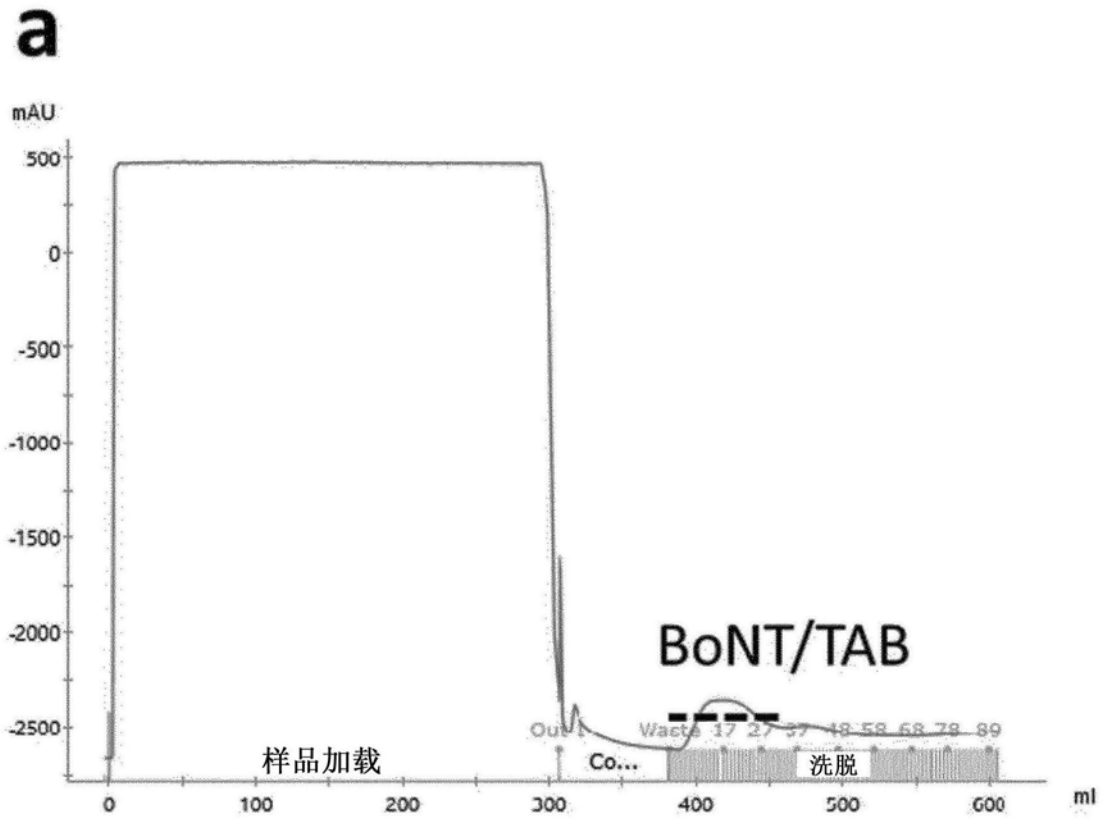


图13

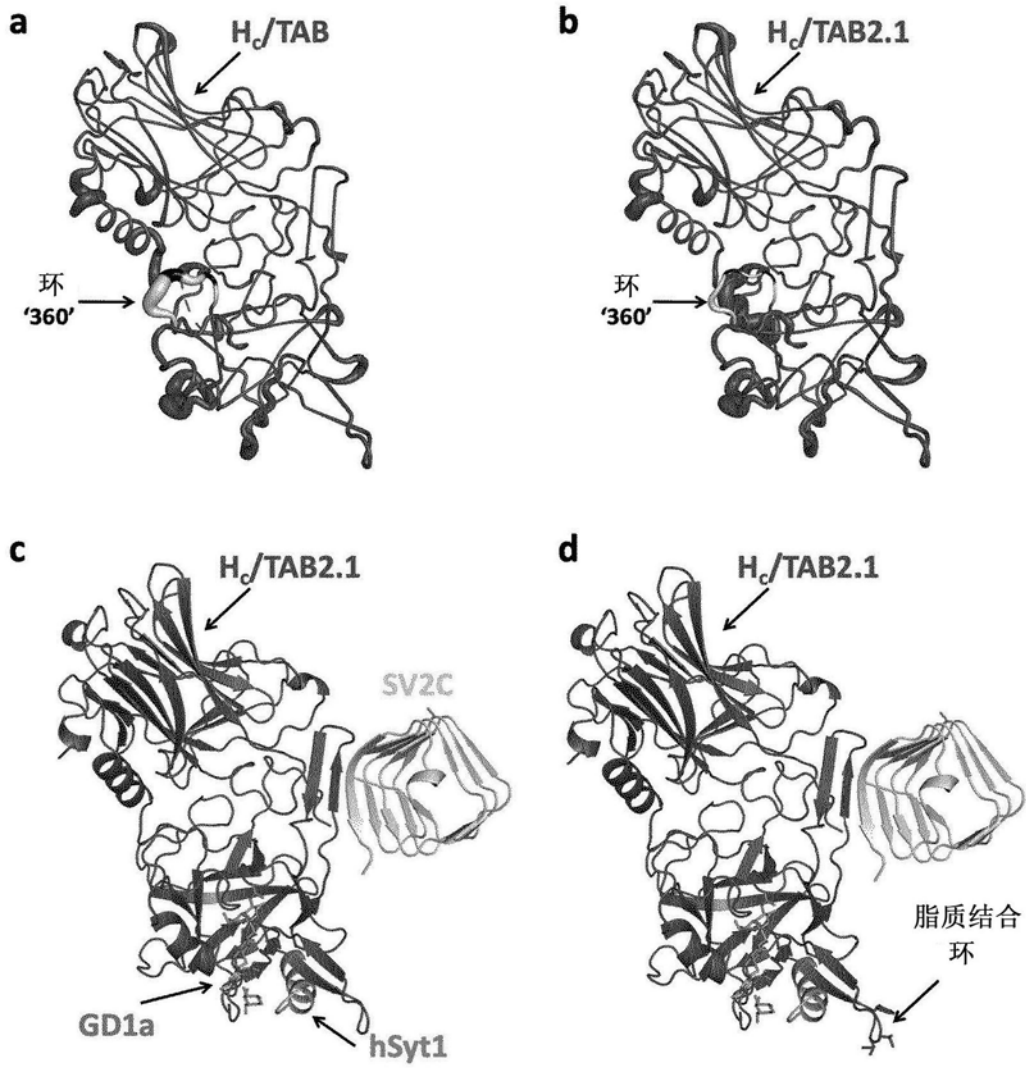


图14

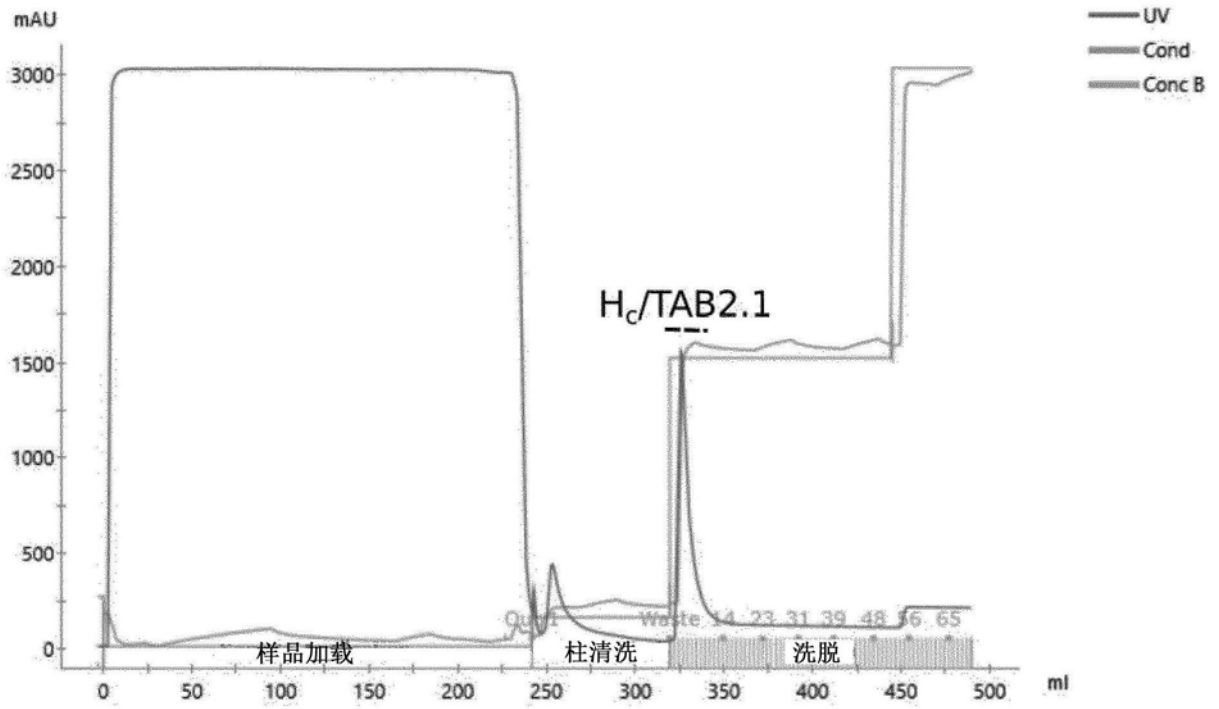


图15A

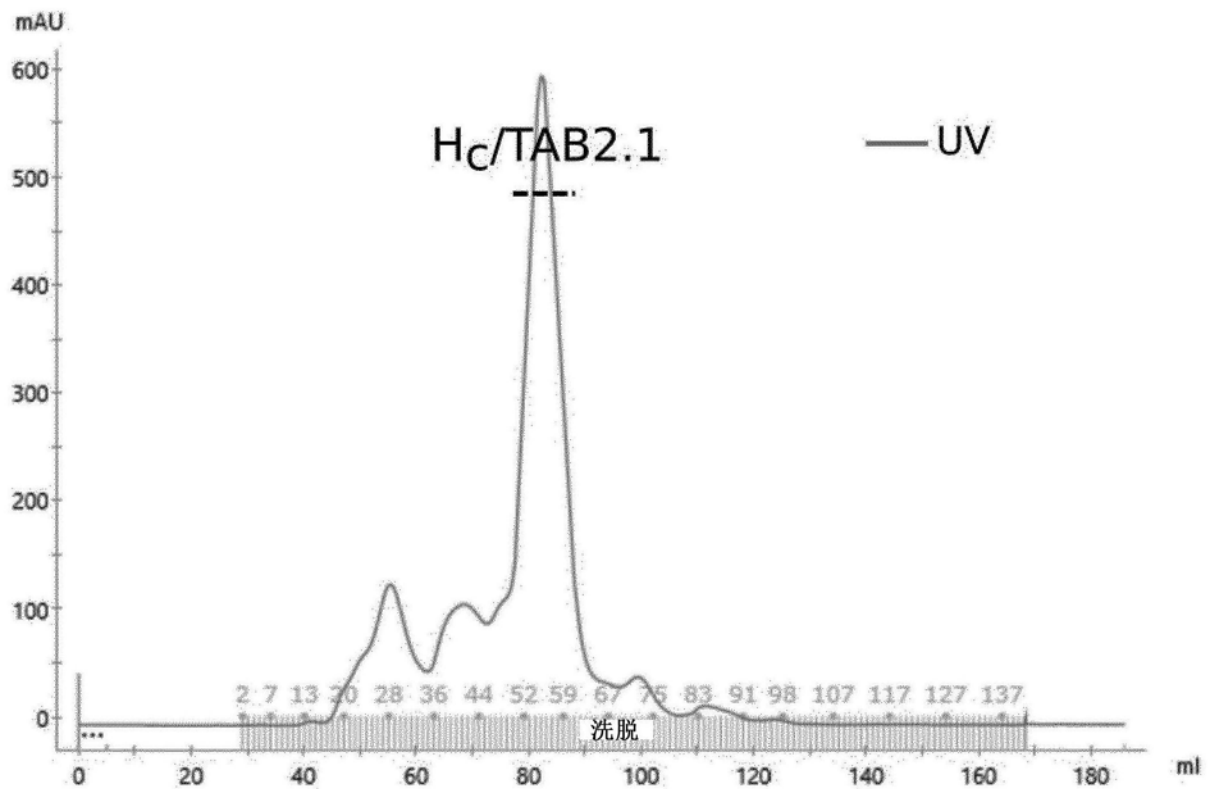


图15B

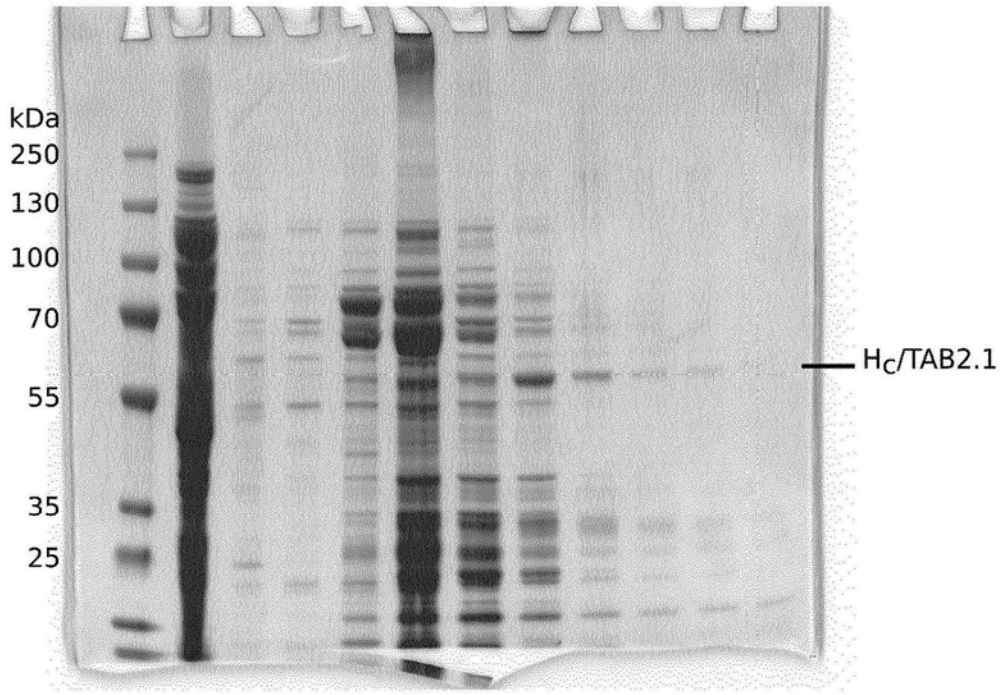


图15C

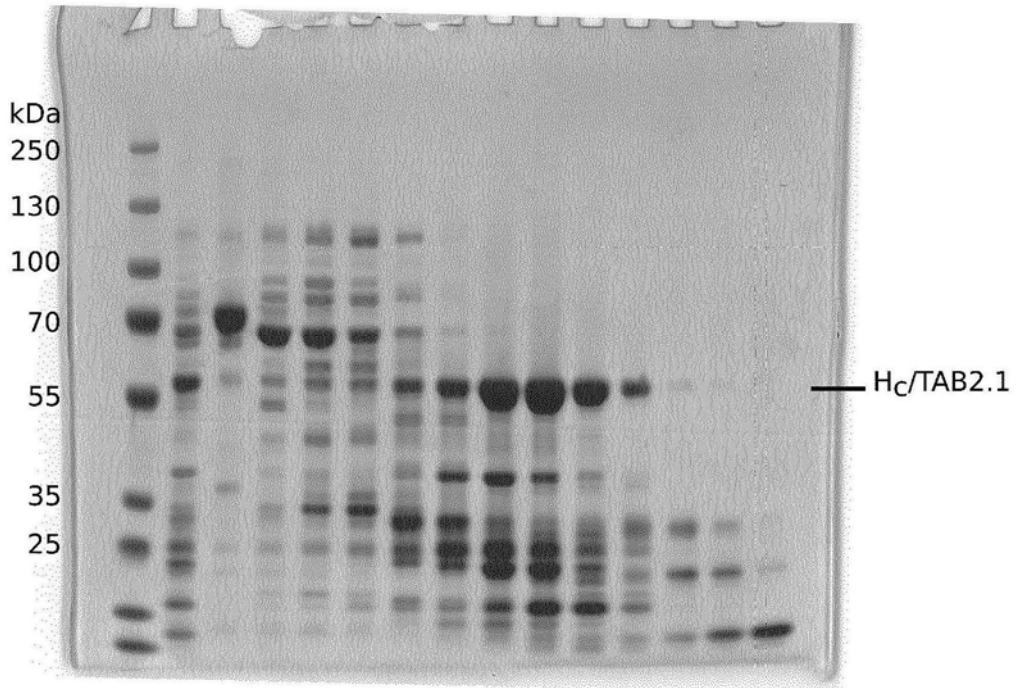


图15D

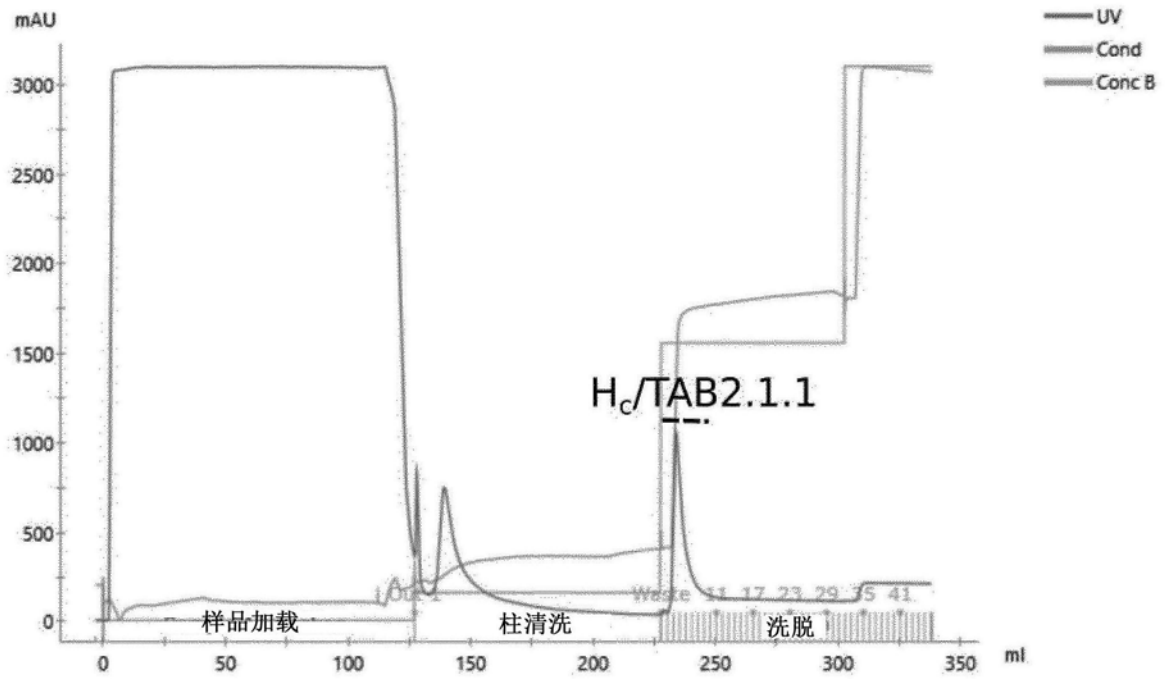


图16A

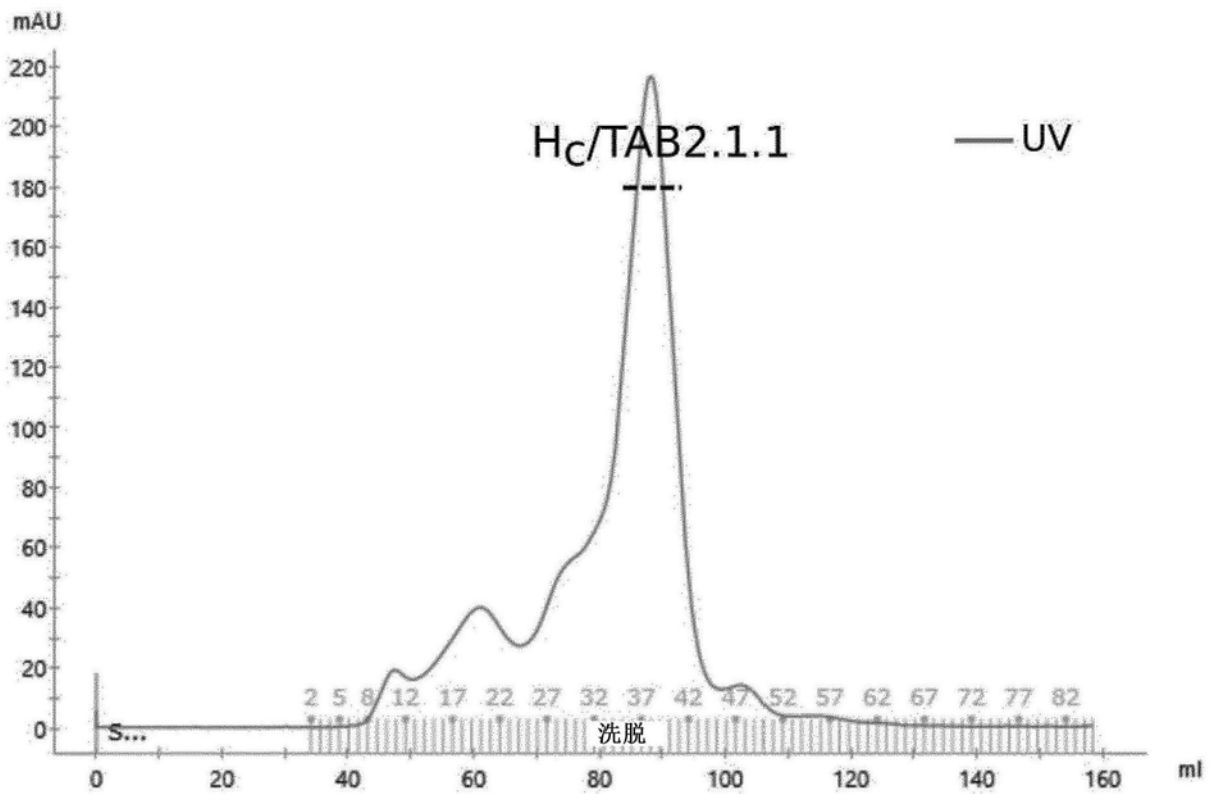


图16B

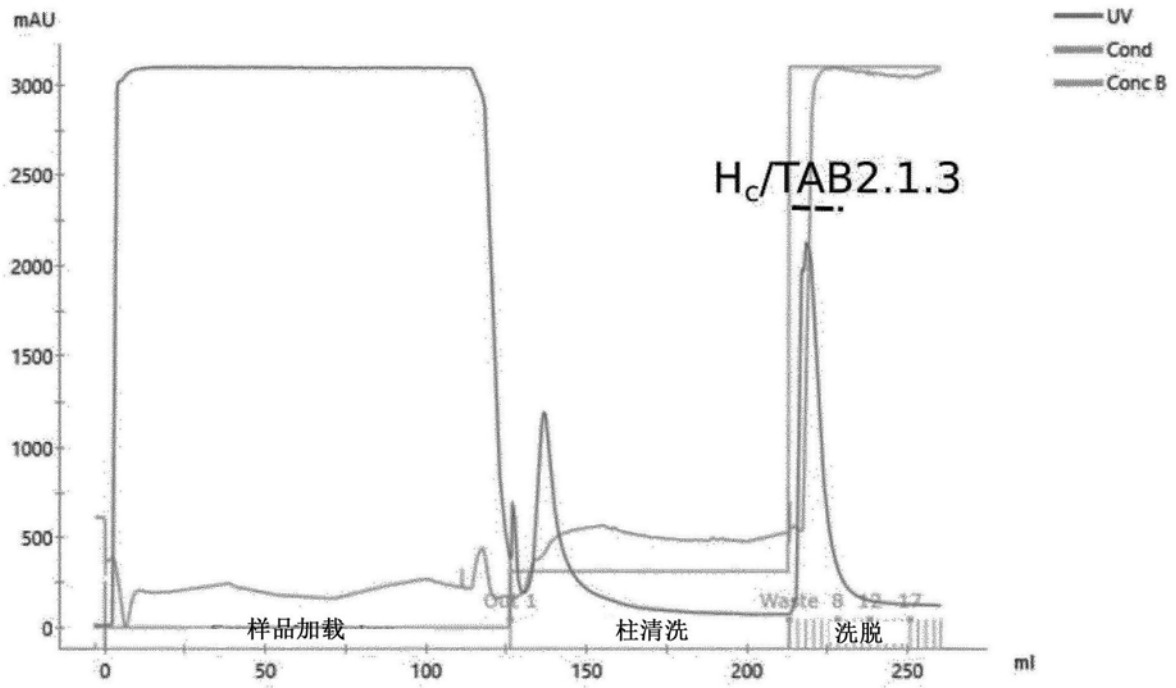


图16C

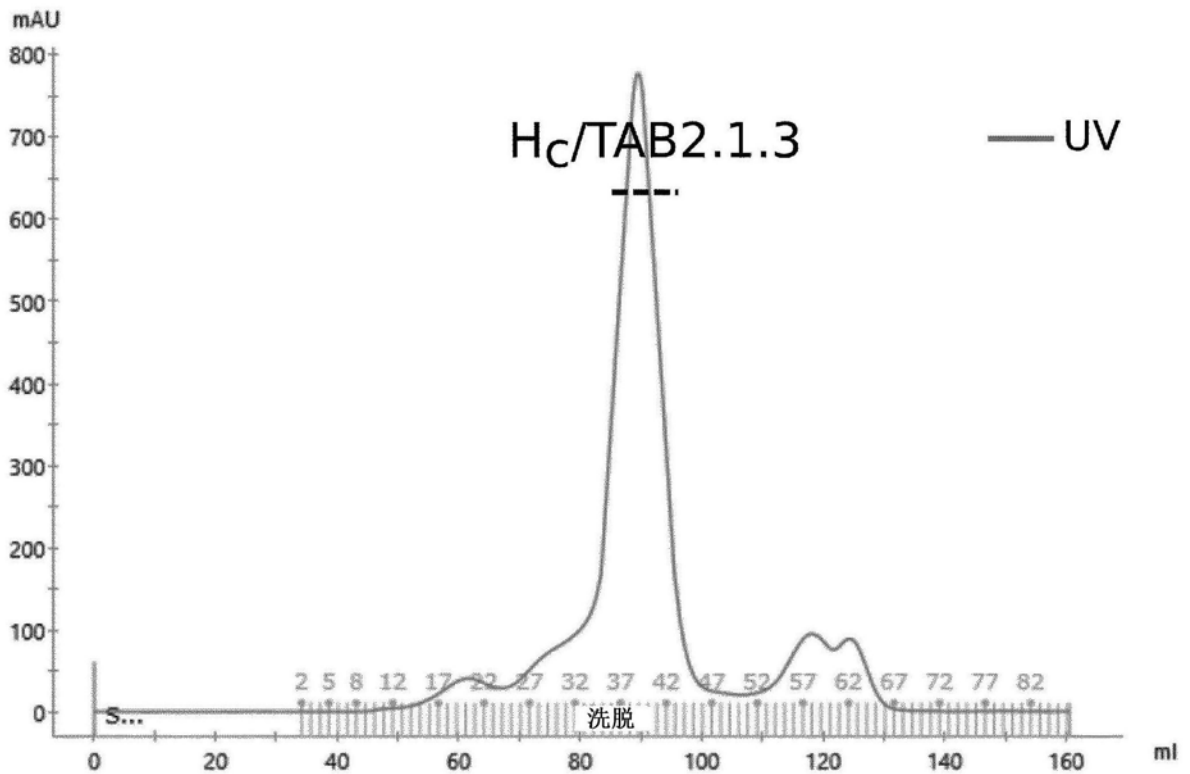


图16D

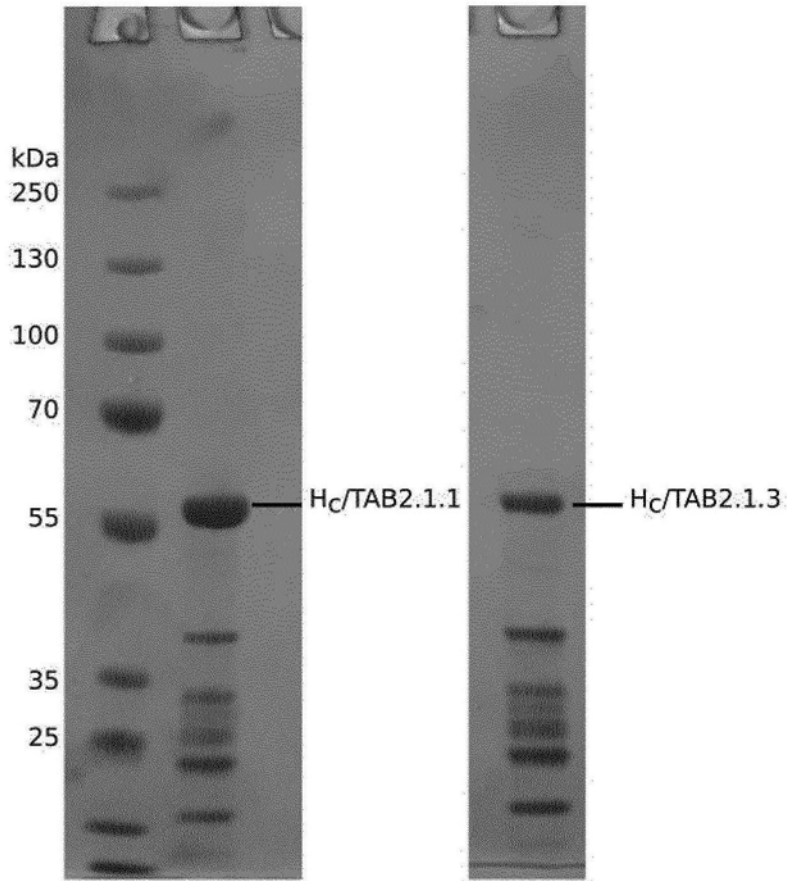


图16E

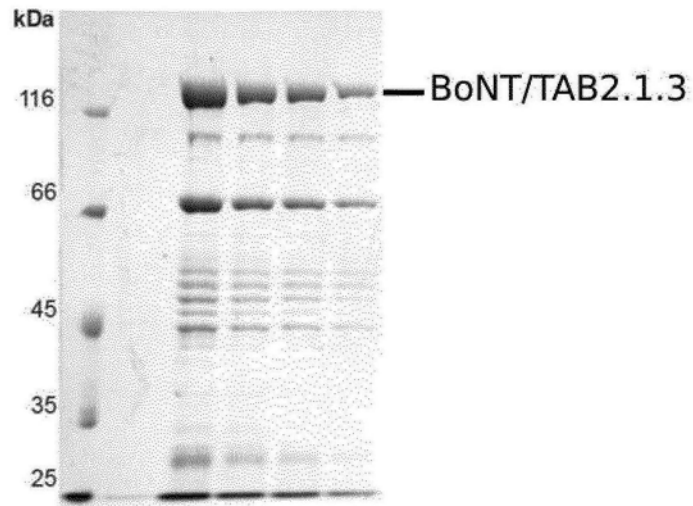


图17A

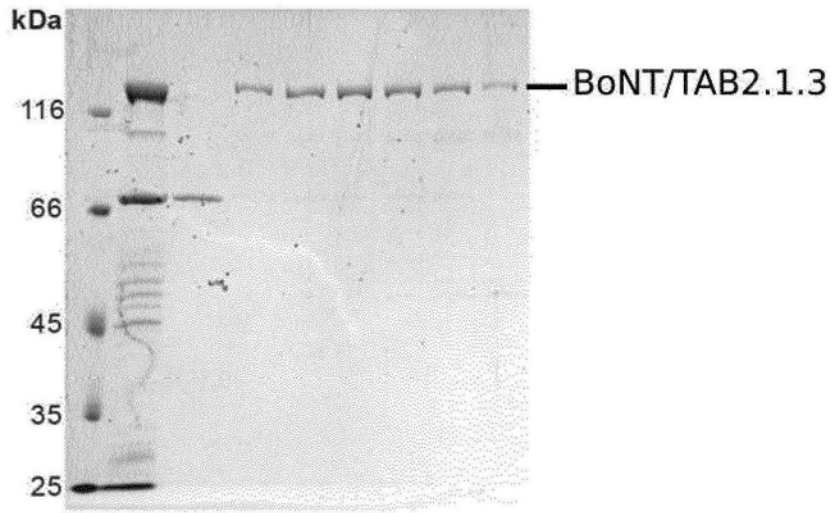


图17B

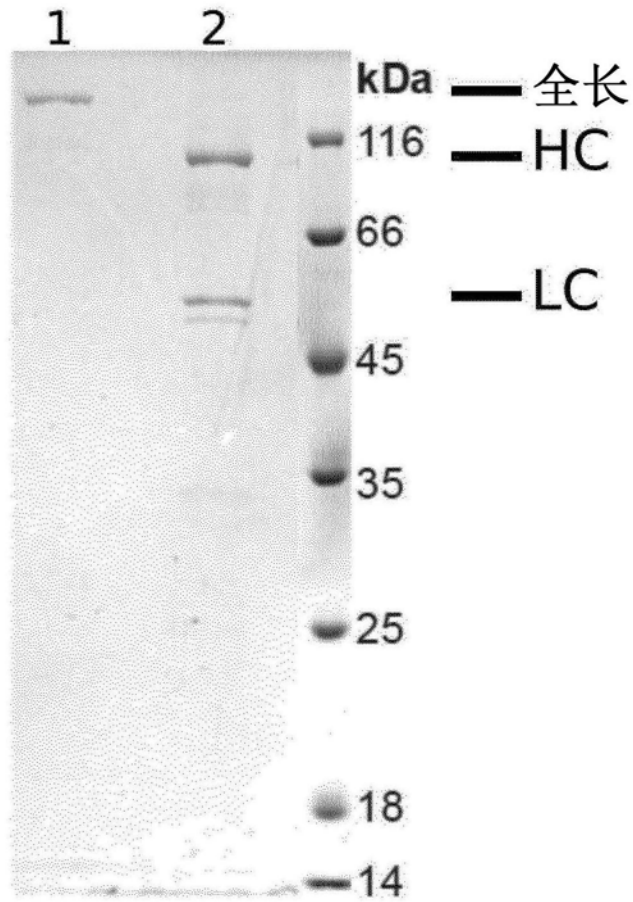


图17C

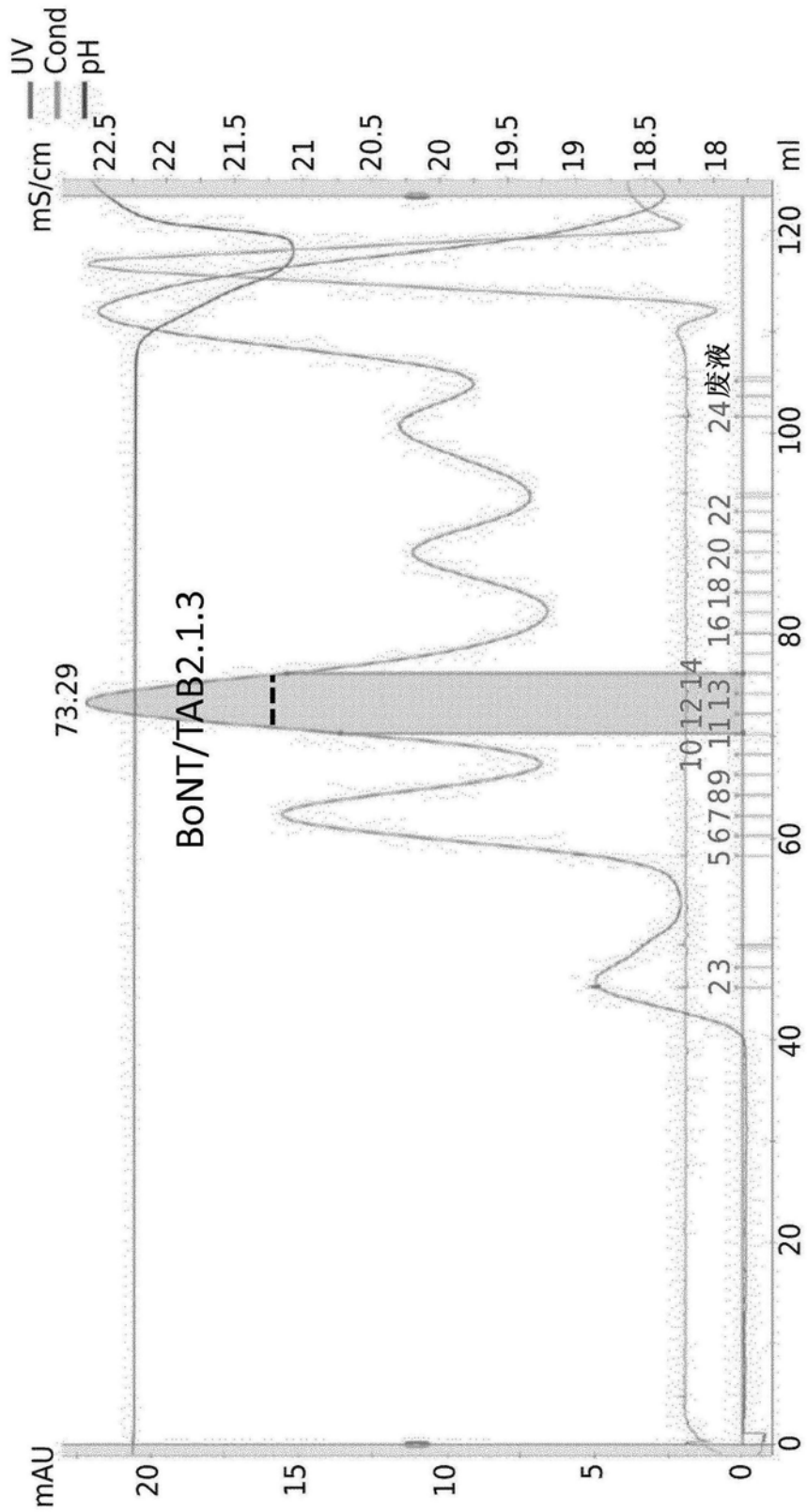


图17D