



(12) 发明专利申请

(10) 申请公布号 CN 114126644 A

(43) 申请公布日 2022. 03. 01

(21) 申请号 202080049940.3

(22) 申请日 2020.05.21

(30) 优先权数据

62/851,040 2019.05.21 US

(85) PCT国际申请进入国家阶段日

2022.01.07

(86) PCT国际申请的申请数据

PCT/US2020/034070 2020.05.21

(87) PCT国际申请的公布数据

W02020/237090 EN 2020.11.26

(71) 申请人 加利福尼亚大学董事会

地址 美国加利福尼亚州

(72) 发明人 金榕生 P·费尔格纳 陈朋

(74) 专利代理机构 上海弼兴律师事务所 31283

代理人 吕学辰 王卫彬

(51) Int.Cl.

A61K 39/08 (2006.01)

C07K 14/33 (2006.01)

C12R 1/145 (2006.01)

权利要求书2页 说明书19页

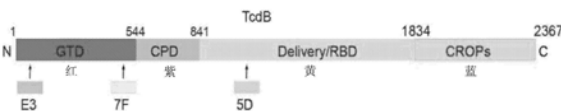
序列表34页 附图8页

(54) 发明名称

艰难梭菌的疫苗组合物

(57) 摘要

本发明提供了通过TcdB或TcdA全毒素治疗或预防艰难梭菌(C.difficile)感染(CDI)的方法和组合物。该组合物的特征在于基于TcdB或TcdA的中和表位中的一种或其组合的免疫原或结合剂,诸如抗体、纳米抗体(VHH)、单结构域抗体(sdAb)等。其中在胞内体pH条件下该免疫原抑制TcdB成孔所必需的构象变化。另外,该免疫原抑制易切键向CPD切割侧移动以及GTD相对于CPD正确定位,从而抑制激活毒素所需要的GTD切割。本发明还描述了用于治疗CDI的疫苗,例如靶向TcdB或TcdA的疫苗。



1. 一种组合物, 包含一种或多种中和艰难梭菌 (*Clostridium difficile*) 的全毒素的分离的多肽, 其中所述分离的多肽包含序列, 所述序列选自与所述全毒素结合并抑制其毒性的序列组成的组。

2. 根据权利要求1所述的组合物, 其中所述分离的多肽包含SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21或SEQ ID NO:22的序列。

3. 根据权利要求1所述的组合物, 其中所述全毒素是TcdB或TcdA。

4. 根据权利要求1所述的组合物, 其中所述分离的多肽序列通过促进TcdB在细胞外活化使其在攻击细胞之前失活来抑制其毒性, 或抑制TcdB成孔所必需的构象变化, 或抑制TcdB活化, 或抑制Rho葡萄糖基化, 从而中和TcdB。

5. 一种中和艰难梭菌全毒素的方法, 所述方法包括制备艰难梭菌的全毒素的免疫原, 并将所述免疫原引入宿主从而引发针对所述免疫原的免疫应答, 其中所述宿主基于所述免疫原产生对所述全毒素具有特异性的抗体。

6. 根据权利要求5所述的方法, 其中所述全毒素是TcdB或TcdA。

7. 根据权利要求5所述的方法, 其中所述免疫原包含SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21或SEQ ID NO:22的序列。

8. 根据权利要求5所述的方法, 其中所述免疫原是小分子。

9. 根据权利要求5所述的方法, 其中所述免疫原是结合剂。

10. 根据权利要求5所述的方法, 其中所述免疫原是抗体。

11. 根据权利要求5所述的方法, 其中所述免疫原是抗体片段。

12. 根据权利要求5所述的方法, 其中所述免疫原是纳米抗体。

13. 根据权利要求5所述的方法, 其中所述免疫原是对两种多肽具有特异性的二价免疫原。

14. 根据权利要求13所述的方法, 其中所述两种多肽相互混合。

15. 根据权利要求13所述的方法, 其中所述两种多肽共价结合。

16. 根据权利要求5所述的方法, 其中所述免疫原是对三种多肽具有特异性的三价免疫原。

17. 根据权利要求16所述的方法, 其中所述三种多肽相互混合。

18. 根据权利要求16所述的方法, 其中所述多肽中的两种或三种共价结合。

19. 根据权利要求5所述的方法, 其中所述免疫原是对四种多肽具有特异性的四价免疫原。

20. 根据权利要求19所述的免疫原, 其中所述四种多肽相互混合。

21. 根据权利要求19所述的免疫原, 其中所述多肽中的两种、三种或四种共价结合。

22. 一种设计和制备对艰难梭菌的全毒素具有特异性的疫苗的方法, 所述方法包括:

- a) 从含有免疫原序列的质粒表达带标签的蛋白
 - b) 将所述带标签的蛋白捕获到微球上
 - c) 向宿主施用所述缀合微球以产生免疫应答。
23. 根据权利要求22所述的方法, 其中所述全毒素是TcdB或TcdA。
24. 根据权利要求22所述的方法, 其中所述带标签的蛋白是带His标签的蛋白。
25. 根据权利要求22所述的方法, 其中所述带标签的蛋白通过体外转录翻译 (IVTT) 系统表达。
26. 根据权利要求22所述的方法, 其中所述带标签的蛋白在大肠杆菌 (*Escherichia coli*) 体内表达。
27. 根据权利要求22所述的方法, 其中所述免疫原包含SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:20、SEQ ID NO:21或SEQ ID NO:22的序列。
28. 根据权利要求22所述的方法, 其中所述免疫原是小分子。
29. 根据权利要求22所述的方法, 其中所述免疫原是抗体。
30. 根据权利要求22所述的方法, 其中所述免疫原是抗体片段。
31. 根据权利要求22所述的方法, 其中所述免疫原是纳米抗体。

艰难梭菌的疫苗组合物

[0001] 相关申请的交叉引用

[0002] 本申请要求2019年5月21日提交的美国临时申请第62/851,040号的权益,其说明书全文以引用方式全文并入本文。

[0003] 关于联邦资助研究或开发的声明

[0004] 本发明在美国国立卫生研究院(National Institutes of Health)办法的R01AI125704、R01AI139087号拨款以及美国DOD/DTRA颁发的HDTRA1-16-C-0009、HDTRA1-18-1-0035号拨款的政府支持下完成。政府对该发明享有一定权利。

[0005] 序列表参考

[0006] 申请人主张,根据细则第13条之三第1(a)款提交、题为UCI_19_16_PCT_Sequencing_Listing_ST25附件、以附录C/ST.25文本文件形式记录的信息与形成所提交国际申请的一部分的信息相同。该序列表的全部内容以引用方式并入本文。

技术领域

[0007] 本发明涉及中和艰难梭菌(*Clostridium difficile*)的全毒素,更具体地涉及一种用于治疗艰难梭菌感染的治疗组合物和方法。

背景技术

[0008] 艰难梭菌被美国疾病控制和预防中心(CDC)列为三大抗生素耐药性紧急威胁之一,并且在发达国家,艰难梭菌感染(CDI)已经成为导致抗生素相关性腹泻和胃肠炎相关性死亡的最常见的原因。CDI的病理主要由两种同源外毒素TcdA和TcdB介导,两者靶向并破坏结肠上皮,导致腹泻和结肠炎。虽然在CDI的发病机理中这两种毒素的相对作用尚不完全清楚,但最近的研究表明TcdB的毒性远大于TcdA,而且对诱导宿主炎症和先天性免疫应答来说更重要。

[0009] TcdA(约308kDa)和TcdB(约270kDa)含有四个功能结构域:N-末端葡糖基转移酶结构域(GTD)、半胱氨酸蛋白酶结构域(CPD)、中心跨膜转运和受体结合结构域(Delivery/RBD)和C-末端组合重复寡肽(CROP)结构域(图1A)。目前普遍认为,毒素借由Delivery/RBD和CROP与细胞表面受体结合,并通过内吞作用进入细胞。胞内体中的酸化作用引发毒素的构象变化,促使Delivery/RBD形成孔隙并跨胞内体膜转运GTD和CPD。

[0010] 胞质溶胶中,CPD由对真核生物具有特异性的肌醇六磷酸(InsP6,也称为植酸)活化,随后发生自体溶解,释放GTD。然后,GTD对Rho家族的小GTP酶(包括Rho、Rac和CDC42)进行葡糖基化。Rho蛋白质的葡糖基化抑制了它们的功能,导致肌动蛋白细胞骨架发生改变、细胞变圆,并且最终导致细胞凋亡。已经报道了TcdA和TcdB片段的许多结构,这为深入了解这些毒素结构域的功能提供了很大的帮助。但是,目前依然不清楚各结构域是如何在全毒素的超三级结构内相互作用,以及全毒素是如何精准、逐步且动态地响应环境和细胞信号(诸如低pH值和InsP6)从而导致中毒。

[0011] 美国食品和药物监督管理局(FDA)最近批准了一种抗TcdB中和抗体

(bezolotoxumab) 作为预防复发性感染的药物,原因是有高达35%的CDI患者病情复发,且很多患者可能需要多轮治疗。但该抗体并未指明适用于治疗CDI或预防CDI。

[0012] 本发明的简要概述

[0013] 本发明的目的是提供一种中和艰难梭菌的全毒素(即TcdB或TcdA)的治疗组合物和方法,如独立权利要求中所述。从属权利要求中给出了本发明的实施例。如果本发明的实施例不相互排斥,则它们可以彼此自由组合。

[0014] 本发明描述了一种治疗组合物,包含一种或多种用于中和艰难梭菌的主要全毒素(TcdB或TcdA)的分离的多肽。在一些实施方案中,分离的多肽由与全毒素结合并抑制其毒性(功能)从而使其中和的分离的多肽的组组成。

[0015] 另外,本发明的特征可在于一种中和艰难梭菌的主要全毒素(TcdB或TcdA)的方法。在一些实施方案中,该方法包括生产艰难梭菌的全毒素(TcdB或TcdA)的免疫原,以及将该免疫原引入宿主从而引发针对该免疫原的免疫应答。在另一个实施方案中,宿主基于免疫原产生对全毒素具有特异性的抗体。

[0016] 此外,本发明的特征可在于一种设计和制备对艰难梭菌的全毒素(TcdB或TcdA)具有特异性的疫苗的方法。

[0017] 在不将本发明限制于任何理论或机制的前提下,发明人认为本发明的疫苗可能具有优势(例如,与用于CDI的类毒素疫苗相比),因为本发明的免疫原是无毒的,从而使它们可能更安全;可在大肠杆菌(E.coli)中以高产率和高纯度生产本发明的免疫原,使它们的生产、配制和贮存成本更低(疫苗的生产可能非常困难);本发明的免疫原保留了它们的天然3D结构(与类毒素中被破坏的抗原结构相比),因此作为疫苗可更有效地引发免疫应答;并且本发明的免疫原较小且含有已知的中和表位,因此免疫原可更有效地引发中和抗体的产生。进一步地,由于这些免疫原是TcdB的更小(与完整毒素相比)、更具特异性的区域,所以它可以形成更好的免疫应答。本发明提供了一种比整个全毒素小但比小分子(例如,15聚)肽大的多肽,即一种具有明确的3D结构的中型肽。

[0018] 根据上下文、本说明书和本领域普通技术人员知识将显而易见的是,本文所述的任何特征或特征组合均包括在本发明的范围内,前提是包括在任何此类组合中的特征不相互矛盾。在以下详细描述和权利要求中,本发明的其它优点和方面是显而易见的。

[0019] 附图中几个图的简要说明

[0020] 结合附图以及下文的具体实施,本发明的特征和优点将变得显而易见,其中:

[0021] 图1A示出了全长TcdB全毒素的示意图,显示了TcdB的结构域排列:GTD(红色);CPD(紫色);Delivery/RBD(黄色);CROP(蓝色)以及近似VHH结合区。

[0022] 图1B示出了TcdB全毒素的3D结构的图示。为了清楚起见,省略了用于促进结晶的3个VHH(GTD;CPD;递送/RBD;CROP)。

[0023] 图2A示出了CROP结构域的示意图,显示了短重复片段(SR,细蓝条)和长重复片段(LR,粗黑条)的排列。虚线表示四个CROP单元(I-IV)的边界。

[0024] 图2B示出了进入CROP结构域的特写视图,而TcdB的其余部分以表面显示。

[0025] 图2C示出了4个CROP单元的叠加。每个CROP单元中的LR导致约132-146°的扭结。

[0026] 图2D和图2E示出了铰链区,该铰链区将CROP结构域连接到毒素的其余部分,位于TcdB的中心并被GTD、CPD和Delivery/RBD包围。

[0027] 图3A示出了SAXS研究中的曲线拟合分析,显示了CROP结构域发生pH依赖性构象变化。基于TcdB全毒素结构的理论Kratky图与pH 5.0(上图)的实验散射图像几乎相同,但与pH 7.4的实验散射图像(下图)不同。

[0028] 图3B示出了经XL-MS鉴定出的不同TcdB结构域之间的交联肽。

[0029] 图3C示出了XL-MS的结果,表明TcdB在中性pH条件下可采用“闭合”构象,其中CROP结构域的中心部分和C-末端尖端在Delivery/RBD约30Å内进行移动。

[0030] 图3D示出了TcdB全毒素的两个极限结构状态的模型。显示了与GTD结合的7F上的受体染料和与CROP结构域结合的B39(星形)上的供体染料(六边形)。

[0031] 图3E示出了pH 5.0 (n=498) 和pH 7.0 (n=594) 时染料标记VHH和TcdB的复合物的非平均FRET效率的群体直方图。

[0032] 图4A示出了TcdB的成孔中间态。5D与Delivery/RBD结合并直接与成孔区域相互作用。成孔区域显示为带状模型,而毒素的其余部分显示为表面模型。

[0033] 图4B示出了1.0σ处的成孔区域的一部分的代表性2Fo-Fc电子密度图,其被最终的细化模型覆盖。

[0034] 图4C示出了大梭菌葡糖基化毒素(LCGT)家族中不同成员之间的成孔区域的氨基酸序列比对。TcdB*、TcdB和TcdB2分别由M68株、VPI10463株和BI/NAP1/027株制备。TcdB*、TcdA的二级结构分别显示在顶部和底部。TcdB全毒素中无可见电子密度的第1032-1047位残基用“x”表示。

[0035] 图4D示出了在酸性pH条件下的TcdB(紫色)和在中性pH条件下的TcdA(橙色)在成孔区域中采用完全不同的构象。这两个结构基于Delivery/RBD进行叠加。

[0036] 图4E示出了钙黄绿素染料释放试验。在存在或不存在5D或7F的情况下,pH 4.6时用负载有50mM钙黄绿素的脂质体测试TcdB(0-25nM)。

[0037] 图4F示出了膜去极化试验。在存在跨膜KCl梯度的情况下,通过加入缬氨霉素使脂质体在正的内部电压下极化。使用电压敏感性荧光染料ANS(8-苯胺基萘-1-磺酸)测量膜电位。3分钟后,加入含不同浓度的5D或7F的TcdB。数据表示为均值±标准差,n=3。

[0038] 图5A示出了β-瓣(β-flap)、3螺旋束(3-HB)和铰链在TcdB的一级序列中的位置的示意图。

[0039] 图5B示出了TcdB全毒素中的apo CPD(灰色卷曲)和与InsP6结合的CPD片段的叠加。apo CPD中的锌原子显示为球形,而InsP6为棒状模型。

[0040] 图5C和图5D示出了β-瓣、3-HB和铰链共同定位于TcdB的中心。

[0041] 图6示出了各种纳米微珠亚单位疫苗的抗体滴度。

[0042] 本发明的详细说明

[0043] 在公开和描述本发明的化合物、组合物和/或方法之前,应当理解的是,本发明不限于具体的合成方法或具体的组合物,因为这些当然可以变化。还应理解的是,本文所用的术语仅是为了描述具体的实施方案,并非限制本发明。

[0044] 现在参考图1-6,本发明的特征在于一种用于中和艰难梭菌的主要全毒素(即TcdB和TcdA)以潜在地治疗和预防CDI的治疗组合物和方法。另外,本发明的特征在于一种生产艰难梭菌全毒素疫苗的方法。

[0045] 如本文所用,艰难梭菌的TcdB的序列来自M68株(WP_003426838.1,见下表1)。所有

氨基酸编号均参照该序列。

[0046] 表1:

描述:	序列:
艰难梭菌 M68 株的 TcdB (SEQ ID NO: 1)	MSLVNRKQLEKMANVRFVRVQEDEYVAILDALEEYHNMSENTVVEKYLKLDIN SLTDYIDTYKKSGRNKALKKFKEYLVIEILELKNSNLTPEKNLHFIWIGGQIND TAINYINQWKDVNSDYNVNVFYDSNAFLINTLKKTIIESASNDTLESFRENLDPE FNHTAFFRKRMQIIYDKQQNFINYKKAQKEENPDLIIDDIVKTYLSNEYSKDIDEL NAYIEESLNKVTENSGNDVRNFEEFKTGEVFNLYEQELVERWNLGASDILRVIL KNIGGVYLDVDMPLGIHPDLFKDINKPDSVKTAVDWEEMQLEAIMKHKEYIPEY TSKHFDTLDEEVQSSFESVLASKSDKSEIFLPLGDIEVSPLEVKIAFAKGSINQALIS AKDSYCSDLLIKIQNRYKILNDTLGPIISQGNDFTTMNNFGESLGAIANEENISFI AKIGSYLRVGFYPEANTTITLSGPTIYAGAYKDLLTFKEMSIDTSILSSELRNFEPK VNISQATEQEKNSLWQFNEERAKIQFEEYKKNYFEGALGEDDNLDFSQNTVTDK EYLLKISSSTKSSERGYVHYIVQLQGDKISYEAACNLFAKNPYDSILFQKNIEDSE VAYYYNPTDSEIQEIDKYRIPDRISDRPKIKLTFIGHGKAENFDIFAGLDVDSLSE IETAIGLAKEDISPKSIEINLLGCNMFYSVNVVEETYPGKLLLRVKDKVSELMPSMS QDSIIVSANQYEVRLINSEGRRELLDHSGEWINKESIIKDISSKEYISFNPKENKIIVK SKNLPSTLLQEIRNNSNSSDIELEEKVMLAECEINVISNIETQVVEERIEEAKSLT SDSINYIKNEFKLIESISDALCDLKQQNELEDSEHIFISFEDISETDEGFSIRFINKETGES IFVETEKTIFFSEYANHITTEISKIKGTIFDVTNGKLVKKVNLDTTHEVNTLNAAFFIQ SLIEYNSSKESLSNLSVAMKVQVYAQLFSTGLNTITDAAKVVELVSTALDETIDLL PTLSEGLPIIATIDGVS LGAAIKELSETSDPLL RQEIEAKIGIMAVNLTTATTAITSS LGIASGFSILLVPLAGISAGIPSLVNNELVLRDKATKVVDYFKHVS LVETEGVFTLL DDKVMMPQDDL VISEIDFNNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSIT YREPHLSIYDVLEVQKEELDLSKDLMLPNAPNRVFAWETGWTPGLRSLENDGT KLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTNIRINLDSNTRSFIVPIITTE YIREKLSYSFYGSGGTYALSLSQYNMGINIELSES DVWIIDVNDVVRDVTIESDKIK KGD LIEGILSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAII EVDLLSK SYKLLISGELKILMLNSNHIQQKIDYIGFNSSELQKNIPYSFVDSEGKENG FINGSTK EGLFVSEL PDVVLISKVYMDDSKPSFGYYSNNLKDVKVITKDNVNILTGY YLKDD IKISLSLTLQDEKTIKLSVHLDSEGV AEILKFMNRKGSTNTSDSLMSFLES MNIS IFVNFLQSNIKFILDANFIISGTTSIGQFEFICDENNNI QPYFIKFNLTNYTLYVGN RQNMIVEPNYDLDDSGDISSTVINFSQKYL YGIDSCVNKVVISPNITDEINITPVY ETNNTYPEVIVLDANYINEKINVNINDLSIRYVWSNDGNDFILMSTSEENKVSQVK IRFVN VF KDKTLANKLSFNFSKQDVPVSEILSFTPSYEDGLIGYDLGLVSLYNE KFYINNFGMMVSGLIYINDSLY YFKPPVNNLITGFVTVGDDKY YFNPINGGAASIG ETIIDDKNYYFNQSGVLQTGVFSTEDGFKYFAPANTLDENLEGEAIDFTGKLIIDE NIYYFEDNYRGAVEWKELDGEMHYFSPETGKAFAKGLNQIGDDKY YFNSDGVMQ KGFVSINDNKHYFDDSGVMKVGYTEIDGKH FYFAENGEMQIGVFNTEDGFKYFA HHNEDLGNEEGEEISYSGILNFNNKIYYFDDSF TAVVGWKDLEDGSKYYFDEDTA EAYIGLSLINDGQYYFNDDGIMQVGFVTINDKV FYFSDSGIIESGVQNIDDNYFYI DDNGIVQIGVFDTSDGYKYFAPANTVNDNIY GQAVEYSGLVRVGEDVYYFGETY TIETGWIYDMENESDKYYFDPETKKACKGINLIDDIKYYFDEKGIMRTGLISFENN
[0047]	
[0048]	NIYYFNENGEMQFGYINIEDKMFYFGEDGVMQIGVFNTPDGFKYFAHQNTLDENF EGESINYTGWLDLDEKRYYFTDEYIAATGSVIIDGEEYFDPDTAQLVISE

[0049] 参见表1和图1A, TcdB全毒素具有第1-544位氨基酸的N-末端葡糖基转移酶结构域

(GTD)、第545-841位氨基酸的半胱氨酸蛋白酶结构域(CPD)、第842-1834位氨基酸的转运结构域/受体结合结构域(Delivery/RBD)和第1835-2367位氨基酸的C-末端组合重复寡肽(CROP)结构域。另外,还有三种中和表位:E3(在GTD中,包含第23-63位氨基酸)、7F(GTD的C-末端紧邻切割位点,包含第147-538位氨基酸)和5D(Delivery/RBD的一部分,包含第1105-1358位氨基酸)。尽管中和表位包含的区域在一级氨基酸序列中不是线性的,但它们在3D结构中聚集在一起形成表位。

[0050] 在一些实施方案中,本发明的特征在于一种治疗组合物,包含一种或多种用于中和艰难梭菌的主要全毒素的分离的多肽。在一些实施方案中,分离的多肽包含与全毒素结合并抑制其毒性/功能从而使其中和的序列。在一些实施方案中,多肽序列可以用作免疫原、或结合剂或其它药物的靶标。

[0051] 制备艰难梭菌TcdB的各种免疫原(见表2):TcdB-FL(全长TcdB)、GTD(第1-543位氨基酸,SEQ ID NO:2)、TD(第798-1805位氨基酸,序列未示出)、TD3(第1286-1805位氨基酸,序列未示出)、CROP4(第2235-2367位氨基酸,序列未示出)和TD1(第1072-1452位氨基酸,孔隙B表位,SEQ ID NO:3)。TD指易位结构域。

[0052] 下表2描述了可以用作免疫原或用作结合剂或其它药物的靶标的多肽序列的非限制性实例。

[0053] 表2:

[0054]

描述	序列	SEQ ID NO:
GTD (第1-543位氨基酸)	MSLVNRKQLEKMANVRFRVQEDEYVAILDALEEYHNMSSENT VVEKYLKLDINSLTDITYIDTYKKSGRNKALKKFKEYLVIEIL ELKNSNLTPVEKNLHFIWIGGQINDTAINYINQWKDVNSDYN VNVFYDSNAFLINTLKKTIIESASNDTLESFRENLDPEFNHTA FFRKRMQIIYDKQQNFINYKKAQKEENPDLIIDDIVKTYLSNE YSKDIDELNAYIEESLNKVTENSGNDVRNFEEFKTGEVFNLYE QELVERWNLGASDILRVAILKNIGGVYLDVDMPLPGIHPDLF KDINKPDSVKTAVDWEEMQLEAIMKHKEYIPEYTSKHFDTLTLD EEVQSSFESVLASKSDKSEIFLPLGDIEVSPLEVKIAFAKGSIIIN QALISAKDSYCSDLLIKQIQNRYKILNDTLGPIISQGNDFNNTM NNFGESLGAIANEENISFIAKIGSYLRVGFYPEANTTITLSGPTI	2

[0055]

	YAGAYKDLLTFKEMSIDTSILSSELRNFEFPKVNISQATEQEKN SLWQFNEERAKIQFEEYKKNYFEGA	
TD1 (第1072-1452 位氨基酸)	LTTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLRDK ATKVVDYFKHVSLVETEGVFTLLDDKVMMPQDDL VISEIDFN NNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSI YDVLEVQKEELDLSKDLMVLPNAPNRVFAWETGWTPGLRSL ENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTN IRINLDSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYN MGINIELSESDVWIIDVDNVVRDVTIESDKIKKGDLIEGILSTLS IEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIEVDLLS KSYKLLISGELKILMLNSNHIQQKIDYIG	3
TD1免疫原 (第1072-1452 位氨基酸+连接 肽(下划线) +10× His标签 (粗体))	SLTTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNELVLRD KATKVVDYFKHVSLVETEGVFTLLDDKVMMPQDDL VISEIDF NNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHL SIYDVLEVQKEELDLSKDLMVLPNAPNRVFAWETGWTPGLRS LENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDT NIRINLDSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALSLSQYN MGINIELSESDVWIIDVDNVVRDVTIESDKIKKGDLIEGILSTLS IEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIEVDLLS KSYKLLISGELKILMLNSNHIQQKIDYIG <u>EFSSGHIDDDDSHML</u> <u>EH</u> HHHHHHHHHH HGM	4
TcdB的第1052- 1472位氨基酸	TSDPLL RQEIEAKIGIMAVNLTTATTAIITSSLGIASGFSILLVPL AGISAGIPSLVNNELVLRDKATKVVDYFKHVSLVETEGVFTLL DDKVMMPQDDL VISEIDFNNSIVLGKCEIWRMEGGSGHTVT DDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSKDLMVLPN APNRVFAWETGWTPGLRSLLENDGTKLLDRIRDNYEGEFYWR YFAFIADALITTLKPRYEDTNIRINLDSNTRSFIVPIITTEYIREK LSYSFYGSGGTYALSLSQYNMGINIELSESDVWIIDVDNVVRD VTIESDKIKKGDLIEGILSTLSIEENKIILNSHEINFSGEVNGSNG FVSLTFSILEGINAIEVDLLSKSYKLLISGELKILMLNSNHIQQK IDYIGFNSELQKNIPYSFVDSEGKE	5
TcdB的第1022- 1502位氨基酸	LLPTLSEGLPIIATIIDGVSLGAAIKELSETSDPLL RQEIEAKIGI MAVNLTTATTAIITSSLGIASGFSILLVPLAGISAGIPSLVNNEL VLRDKATKVVDYFKHVSLVETEGVFTLLDDKVMMPQDDLVI SEIDFNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYR EPHLSIYDVLEVQKEELDLSKDLMVLPNAPNRVFAWETGWTP GLRSLLENDGTKLLDRIRDNYEGEFYWRYFAFIADALITTLKPR YEDTNIRINLDSNTRSFIVPIITTEYIREKLSYSFYGSGGTYALS	6

[0056]

	SQYNMGINIELSESDVWIIDVDNVVRDVTIESDKIKKGDLIEGI LSTLSIEENKIILNSHEINFSGEVNGSNGFVSLTFSILEGINAIEV DLLSKSYKLLISGELKILMLNSNHIQQKIDYIGFNSELQKNIPYS FVDSEKENGFGINGSTKEGLFVSELDPDVLISKVYMDD	
TcdB的第1-533 位氨基酸	MSLVNRKQLEKMANVRFRVQEDEYVAILDALEEYHNMSSENT VVEKYLKCLKDINSLTDYIDTYKKSGRNKALKKFKEYLVIEIL ELKNSNLTPVEKNLHFIWIGGQINDTAINYINQWKDVNSDYN VNVFYDSNAFLINTLKKTIIESASNDTLESFRENLDPEFNHTA FFRKRMQIIYDKQQNFINYKKAQKEENPDLIIDDIVKTYLSNE YSKDIDELNAYIEESLNKVTENSGNDVRNFEEFKTGEVFNLVE QELVERWNLGASDILRVAILKNIGGVYLDVDMPLPGIHPDLF KDINKPDSVKTAVDWEEMQLEAIMKHKEYIPEYTSKHFDTL EEVQSSFESVLASKSDKSEIFLPLGDIEVSPLEVKIAFAKGSIN QALISAKDSYCSDLLIKQIQNRYKILNDTLGPIISQGNDFNTM NNFGESLGAIANEENISFIAKIGSYLRVGFYPEANTTITLSGPTI YAGAYKDLLTFKEMSIDTSILSSELRNFEFPKVNISQATEQEKN SLWQFNEERAKI QFE	7
TcdB的第1-593 位氨基酸	MSLVNRKQLEKMANVRFRVQEDEYVAILDALEEYHNMSSENT VVEKYLKCLKDINSLTDYIDTYKKSGRNKALKKFKEYLVIEIL ELKNSNLTPVEKNLHFIWIGGQINDTAINYINQWKDVNSDYN VNVFYDSNAFLINTLKKTIIESASNDTLESFRENLDPEFNHTA FFRKRMQIIYDKQQNFINYKKAQKEENPDLIIDDIVKTYLSNE YSKDIDELNAYIEESLNKVTENSGNDVRNFEEFKTGEVFNLVE QELVERWNLGASDILRVAILKNIGGVYLDVDMPLPGIHPDLF KDINKPDSVKTAVDWEEMQLEAIMKHKEYIPEYTSKHFDTL EEVQSSFESVLASKSDKSEIFLPLGDIEVSPLEVKIAFAKGSIN QALISAKDSYCSDLLIKQIQNRYKILNDTLGPIISQGNDFNTM NNFGESLGAIANEENISFIAKIGSYLRVGFYPEANTTITLSGPTI YAGAYKDLLTFKEMSIDTSILSSELRNFEFPKVNISQATEQEKN SLWQFNEERAKIQFEEYKKNYFEGALGEDDNLDFSQNTVTDK EYLLEKISSSTKSSERGYVHYIVQLQGDKI SYE	8
TcdB的第1-573 位氨基酸	MSLVNRKQLEKMANVRFRVQEDEYVAILDALEEYHNMSSENT VVEKYLKCLKDINSLTDYIDTYKKSGRNKALKKFKEYLVIEIL ELKNSNLTPVEKNLHFIWIGGQINDTAINYINQWKDVNSDYN VNVFYDSNAFLINTLKKTIIESASNDTLESFRENLDPEFNHTA FFRKRMQIIYDKQQNFINYKKAQKEENPDLIIDDIVKTYLSNE YSKDIDELNAYIEESLNKVTENSGNDVRNFEEFKTGEVFNLVE QELVERWNLGASDILRVAILKNIGGVYLDVDMPLPGIHPDLF	9

[0057]

	KDINKPDSVKTAVDWEEMLQLEAIMKHKEYIPEYTSKHFDTLDEEVQSSFESVLASKSDKSEIFLPLGDIEVSPLEVKIAFAKGSINQALISAKDSYCSDLLIKQIQNRYKILNDTLGPIISQGDNDFNTTMNFGESLGAIANEENISFIAKIGSYLRVGFYPEANTTITLSGPTIYAGAYKDLLTFKEMSIDTSILSSELRNFEFPKVNISQATEQEKNSLWQFNEERAKIQFEEYKKNYFEGALGEDDNLDFSQNTVTDK EYLLEKISSS TKS	
TcdB (5D) 的第1105-1358位氨基酸	PSLVNNELVLRDKATKVVDYFKHVSLVETEGVFTLLDDKVMMPQDDL VISEIDFNNSIVLGKCEIWRMEGGSGHTVTDDIDHFFSAPSITYREPHLSIYDVLEVQKEELDLSKDLMVLPNAPNRVFAWETGWTPGLRSLNDGTCLLDRIRDNYEGEFYWRYFAFIADALITTLKPRYEDTNIRINLDSNTRSFIVPIITTEYIREKLSYSFYSGGTYALSLSQYNMGINIELSESDVWIIDVDNVVRDVTI	10
TcdB (E3) 的第23-63位氨基酸	EYVAILDALEEYHNMSENTVVEKYLKLDINSLTDYI DTY	11
TcdB (F7) 的第147-538位氨基酸	ESASNDTLESFRENLDPEFNHTAFFRKRMQIIYDKQQNFINYKAQKEENPDLIIDDIVKTYLSNEYSKDIDELNAYIEESLNKVTENSGNDVRNFEFETGEVFNLYEQELVERWNLGADILRVAILKNIGGVYLDVDMPLGHPDLFKDINKPDSVKTAVDWEEMLQLEAIMKHKEYIPEYTSKHFDTLDEEVQSSFESVLASKSDKSEIFLPLGDIEVSPLEVKIAFAKGSINQALISAKDSYCSDLLIKQIQNRYKILNDTLGPIISQGDNDFNTTMNFGESLGAIANEENISFIAKIGSYLRVGFYPEANTTITLSGPTIYAGAYKDLLTFKEMSIDTSILSSELRNFEFPKVNISQATEQEKNSLWQFNEERAKIQFEEYKKN	12
TcdB的第1792-1843位氨基酸 (铰链区)	PVSEIILSFTPSYYEDGLIGYDLGLVSLYNEKFYINNFGMMVSGLIYINDSL	13
TcdB的第666-841位氨基酸 (3-HB)	GLDVDSLSSEIETAIGLAKEDISPKSIEINLLGCNMFSYSVNVEETYPGKLLLRVKDKVSELMPSMSQDSIIVSANQYEVRLINSEGRRELLDHSGEWINKEESIHKDISKEYISFNPKENKIIVKSKNLPeltsTLLQEIRNNSNSDIELEEKVMLAECEINVISNIETQVVEER	14
TcdB的第742-841位氨基酸 (β-瓣)	QYEVRLINSEGRRELLDHSGEWINKEESIHKDISKEYISFNPKENKIIVKSKNLPelstLLQEIRNNSNSDIELEEKVMLAECEINVISNIETQVVEER	15

[0058]

TcdA的第1-541位氨基酸	MSLISKEELIKLAYSIRPRENEYKTILTNLDEYNKLTNNNNENKYLQLKKLNESIDVFMNKYKTSSRNRAISNLKKDILKEVILIKNSNTSPVEKNLHFVWIGGEVSDIALEYIKQWADINA EYNIKLWYDSEAFLVNTLKKAI VESSSTEALQLLEEEIQNPQFDNMKFYKKRMEFIYDRQKRFINYYKSQINKPTVPTIDDIKSHLVSEYNRDET VLESYRTNSLRKINSNHGIDIISRPSSIGLDRWEMIKLEAIMKYKKYINNYTSENFDKLDQQLKDNFKLIIESKSEKSEIFSKLENLNVSDLEIKIAFALGSVINQALISKQGSYLTNLVIEQVKNRYQFLNQHLNPAIESDNNFTDTTKIFHDSL FNSATAENSMFLTKIAPYLQVGFMPEARSTISLSGPGAYASAYYRANSLFTEQELLNIYSQELLRGNLAAASDIVRLLALKNF GG VYLDVDM L PGIHSDLFKTDFINLQENTIEKTLKASDLIEFKFPENNLSQLTEQEINSLWSFDQASAKYQFEKYVRDYTGGS	16
TcdA的第1073-1452位氨基酸	MSLSIAATVASIVGIGAEVTIFLLPIAGISAGIPSLVNNELILHDKATSVVNYFNHLSSESKYGPLKTEDDKILVPIDDLVISEIDFNNNSIKLGT CNILAMEGGSGHTVTGNIDHFFSSPSISSHIPSLSIYSAIGIETENLDFS KIMMLPNAPSRVFWWETGAVPGLRSLENDGTRLLDSIRDLYPGKFYWRFYAFFDYAITTLKPVYEDTNIKIKLDKDTRNFIMPTITTNEIRNKLSYSFDGAGGTYSLLLSSYPISTNINLSKDDLWIFNIDNEVREISIENGTIKKGKLIKDVLSKIDINKNKL IIGNQTIDFSGDIDNKDRYIFLTCELDDKISLIEINLVAKSYSLLLSGDKNYLISNLSNTIEKINTLG	17
TcdA的第22-62位氨基酸	EYKTILTNLDEYNKLTNNNNENKYLQLKKLNESIDVFMNKY	18
TcdA的第146-536位氨基酸	SSTTEALQLLEEEIQNPQFDNMKFYKKRMEFIYDRQKRFINYYKSQINKPTVPTIDDIKSHLVSEYNRDET VLESYRTNSLRKINSNHGIDIISRPSSIGLDRWEMIKLEAIMKYKKYINNYTSENFDKLDQQLKDNFKLIIESKSEKSEIFSKLENLNVSDLEIKIAFALGSVINQALISKQGSYLTNLVIEQVKNRYQFLNQHLNPAIESDNNFTDTTKIFHDSL FNSATAENSMFLTKIAPYLQVGFMPEARSTISLSGPGAYASAYYRANSLFTEQELLNIYSQELLNRGNLAAASDIVRLLALKNF GG VYLDVDM L PGIHSDLFKTDFINLQENTIEKTLKASDLIEFKFPENNLSQLTEQEINSLWSFDQASAKYQFEKYVRD	19
TcdA的第1789-1840位氨基酸	SLGYIMSNFKSFNSENELDRDHLGFKIIDNKTYYYDEDSKLVKGLININNSL	20
TcdA的第664-842位	NKIPSNNVEEAGSKNYVHYIIQLQGDDISYEATCNLF SKNPKN SIIQRNMNESAKSYFLSDDGESILELNKYRIPERLKNKEKVKV	21

[0059]	氨基酸	TFIGHGKDEFNTSEFARLSVDSLSNEISSFLDTIKLDISPKNVEV NLLGCNMFSDNFVEETYPGKLLLSIMDKITSTLPDVNKNISITI GANQYEVVRINSEGRKELLAHSGKWINKEEAIMSDLSSKEYIFF DSIDNKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLK LNISSIGDYIYYEK	
	TcdA的第743-842位氨基酸	QYEVVRINSEGRKELLAHSGKWINKEEAIMSDLSSKEYIFFDSID NKLKAKSKNIPGLASISEDIKTLLLDASVSPDTKFILNNLKLNIE SSIGDYIYYEK	22

[0060] 现参见表2,本发明的特征在于一种分离的多肽,不限于此处列出的序列。SEQ ID NO:2为艰难梭菌的TcdB的第1-543位氨基酸。SEQ ID NO:3为艰难梭菌的TcdB的第1072-1452位氨基酸,并且第1072-1452位氨基酸是成孔所必需的易位结构域的一部分。SEQ ID NO:5为艰难梭菌的TcdB的第1052-1472位氨基酸。SEQ ID NO:6为艰难梭菌的TcdB的第1022-1502位氨基酸。SEQ ID NO:7为艰难梭菌的TcdB的第1-533位氨基酸。SEQ ID NO:8为艰难梭菌的TcdB的第1-593位氨基酸。SEQ ID NO:9为艰难梭菌的TcdB的第1-573位氨基酸。SEQ ID NO:10为艰难梭菌的TcdB的第1105-1358位氨基酸,并且是包含5D表位的区域。SEQ ID NO:11为艰难梭菌的TcdB的第23-63位氨基酸,并且是包含E3表位的区域。SEQ ID NO:12为艰难梭菌的TcdB的第147-538位氨基酸,并且包含F7表位。SEQ ID NO:13为艰难梭菌的TcdB的第1792-1845位氨基酸,对应于铰链区。SEQ ID NO:14为艰难梭菌的TcdB的第666-841位氨基酸,对应于3-HB区。SEQ ID NO:15为艰难梭菌的TcdB的第741-841位氨基酸,对应于 β -flap区域。SEQ ID NO:16为艰难梭菌的TcdA的第1-541位氨基酸。SEQ ID NO:17为艰难梭菌的TcdA的第1073-1452位氨基酸。SEQ ID NO:18为艰难梭菌的TcdA的第22-62位氨基酸。SEQ ID NO:19为艰难梭菌的TcdA的第146-536位氨基酸。SEQ ID NO:20为艰难梭菌的TcdA的第1789-1840位氨基酸。SEQ ID NO:21为艰难梭菌的TcdA的第664-842位氨基酸。SEQ ID NO:22为艰难梭菌的TcdA的第743-842位氨基酸。

[0061] 在一些实施方案中,可以靶向铰链表位。如本文所用,铰链表位包含以下中的一个、两个或三个:铰链(第1792-1834位氨基酸)、3-HB(第766-841位氨基酸)和 β -瓣(第742-765位氨基酸)。这三个结构单元在氨基酸序列中是分开的,但在3D结构中聚集在一起。

[0062] 在一些实施方案中,分离的多肽包括与其序列具有至少50%相似性的肽。在一些实施方案中,分离的多肽包括与其序列具有至少60%相似性的肽。在一些实施方案中,分离的多肽包括与其序列具有至少75%相似性的肽。在一些实施方案中,分离的多肽包括与其序列具有至少90%相似性的肽。在一些实施方案中,分离的多肽包括与其序列具有至少98%相似性的肽。

[0063] 本发明的特征还在于一种免疫原,该免疫原包含至少一种根据本发明的多肽。在一些实施方案中,免疫原是对根据本发明的两种多肽具有特异性的二价免疫原。在一些实施方案中,这两种多肽相互混合。在一些实施方案中,这两种多肽共价结合。在一些实施方案中,免疫原是对根据本发明的三种多肽具有特异性的三价免疫原。在一些实施方案中,这三种多肽相互混合。在一些实施方案中,这些多肽中两种或三种共价结合。在一些实施方案中,免疫原是对根据本发明的四种多肽具有特异性的四价免疫原。在一些实施方案中,这四

种多肽相互混合。在一些实施方案中,这些多肽中的两种、三种或四种共价结合。

[0064] 在一些实施方案中,本发明的特征在于一种中和艰难梭菌的主要全毒素的方法。在一些实施方案中,该方法包括制备艰难梭菌的全毒素的免疫原,并将该免疫原引入宿主以便引发针对该免疫原的免疫应答,其中宿主基于该免疫原产生对全毒素具有特异性的抗体。

[0065] 如本文所用,“免疫原”可以指能够在宿主中引发免疫应答的任何化合物。免疫原的非限制性示例可包括结合剂、仅含重链的抗体(称为VHH或纳米抗体)的抗原结合区(V_H)、抗体、抗体片段、小分子或药物。也可以考虑使用任何其它合适的免疫原。如本文所用,“宿主”可以是指哺乳动物,例如但不限于小鼠或人。

[0066] 在一些实施方案中,本发明的特征在于一种设计和制备对艰难梭菌全毒素具有特异性的疫苗的方法。在一些实施方案中,疫苗可包含但不限于上表2中列出的任何序列的免疫原。在某些实施方案中,疫苗包含与上表2中所列序列相似的免疫原或疫苗,例如截短型、扩大型或同源型。本发明提供了使用孔隙B表位(SEQ ID NO:3)的第一小鼠CDI疫苗。本发明不限于小鼠疫苗,且包括用于其它动物如人的疫苗。

[0067] 本发明还描述了用新型Toll样受体(TLR)三重激动剂佐剂平台配制抗原,该平台使用组合化学将三种不同的TLR激动剂连接在一起以形成一种佐剂复合物。可以评估TLR三重激动剂佐剂组的免疫调节活性,以确定它们是否引发独特的抗原特异性免疫应答(例如在体外和/或体内)。可以评估出最好的候选抗原,以帮助生产有效的疫苗。

[0068] 本发明还描述了疫苗设计和制备的策略。例如,本发明描述了一种使用优化的微球捕获系统的疫苗抗原(Ag)捕获和体内递送平台。可使用标签或其它化学交联剂将抗原连接到微球。例如,使用体外转录翻译(IVTT)系统或在体内(大肠杆菌)从含有抗原序列的质粒中表达带His标签的蛋白。链霉亲和素包被的微球可以与tris-NTA生物素接头结合,然后用于捕获在大肠杆菌中表达的或来自IVTT反应的蛋白质。在具有或不具有TLR激动剂佐剂的情况下,直接施用所得的Ag缀合微球,以监测抗体释放的动力学和同型。Ag以约200,000/微珠的密度进行包被。免疫原性研究显示稳健且持久的Ag特异性应答。这表明特定蛋白通过结合到微球上而从复杂混合物中分离,并且可通过高通量和可扩展的方式直接进行免疫原性测试。本发明不限于该方法,并且本发明不限于His标签。

[0069] 根据表3制备各疫苗制剂。各制剂通过(皮下)注射给予小鼠。第一针为第0天;强化针1为第14天。每组4只小鼠。表3和图6示出了测量的中点滴度。Ag仅代表抗原;AV代表Addavax;AV+TLR代表Addavax、CpG、MPLA、TLR2/6。图6示出了抗体滴度。用艰难梭菌TcdB的无毒区段免疫,在小鼠中诱导出高抗体水平。抗体水平的增加量大于3log。免疫应答对一种381个氨基酸的免疫原(全长毒素为2367个氨基酸)具有特异性。针对该381个氨基酸的免疫原诱导的抗体也与全长毒素反应。

[0070] 表3:中点滴度(血清稀释度)

疫苗制剂	TD1	TcdB-FL (全长)
可溶性TD1	6,888	43
可溶性TD1+明矾	23,468	141
可溶性TD1+AV	101,987	1,334
可溶性TD1+AV/MPLA/CpG/TLR2/6	237,546	13,323

1 μ M微珠	56,325	3,559
1 μ M微珠+AV	51,819	490
1 μ M微珠+AV/MPLA/CpG TLR2/6	262,557	25,095
0.2 μ M微珠	137,560	1,015
0.2 μ M微珠+AV	102,618	469
0.2 μ M微珠+AV/MPLA/CpG/TLR2/6	126,405	12,215

[0072] 本发明还描述了用于改善抗体或结合剂的抗毒素活性的方法和用于开发同时靶向多个目的表位(例如,本文毒素上的多个中和表位)的多结构域抗体或结合剂的方法。

[0073] 本发明描述了靶向中和表位使TcdB失活从而治疗CDI(例如,用药物、小分子、结合剂等)。本发明还描述了基于中和表位的疫苗开发。免疫原或疫苗可以使全毒素失活,例如通过抑制作为其毒性的先决条件的单个结构域的生物功能,或通过促进胞外活化导致其在攻击细胞之前灭活。

实施例

[0074] 以下是本发明的非限制性示例。应当理解的是,该示例并非意在以任何方式限制本发明。等效或替代方案均在本发明公开的范围内。

[0075] 方法

[0076] 使用了由艰难梭菌M68株制备的TcdB。TcdB全毒素及其GTD的表达如前文所述。将编码四种VHH(5D、E3、7F和B39)的基因、由VPI10463株制备的TcdB的GTD(第1-542位残基,称为GTD^{VPI10463})、截短的TcdB Delivery/RBD(第1072-1433位残基,TcdB¹⁰⁷²⁻¹⁴³³)和TD1(第1072-1452位残基,在C末端具有10 $\times$ His标签)克隆到经修饰的pET28a载体中,该载体在所有蛋白质的N末端均具有6 $\times$ His/SUMO(酿酒酵母*Saccharomyces cerevisiae* Smt3p)标签。将TcdB片段(第1-1805位残基,TcdB¹⁻¹⁸⁰⁵)克隆到经修饰的pET22b载体中,该载体具有引入在SUMO标签和TcdB¹⁻¹⁸⁰⁵之间的twin-Strep标签和C-末端的6 $\times$ His标签。所有突变体均通过两步PCR生成,并通过DNA测序进行验证。

[0077] 在大肠杆菌BL21-Star株(DE3)(Invitrogen)中表达5D、E3、7F、B39、GTD^{VPI10463}、TcdB¹⁻¹⁸⁰⁵、TcdB¹⁰⁷²⁻¹⁴³³和TD1。用含有卡那霉素或氨苄青霉素的LB培养基于37 $^{\circ}$ C培养细菌。当OD₆₀₀达到约0.8时,将温度降至16 $^{\circ}$ C。用1mM IPTG(异丙基-b-D-硫代半乳糖苷)诱导表达,并于16 $^{\circ}$ C持续培养过夜。离心收集细胞,并于-80 $^{\circ}$ C储存直至使用。

[0078] 使用Ni²⁺-NTA(亚硝基三乙酸,Qiagen)亲和树脂在含有50mM Tris、400mM NaCl和10mM咪唑,pH 8.5的缓冲液中纯化带His₆标签的TcdB、GTD和带His₆-SUMO标签的5D、E3、7F、B39、GTD^{VPI10463}、TcdB¹⁻¹⁸⁰⁵、TcdB¹⁰⁷²⁻¹⁴³³和TD1。蛋白质用高咪唑缓冲液(50mM Tris、400mM NaCl和300mM咪唑,pH 8.5)洗脱,然后于4 $^{\circ}$ C使用含有20mM Tris、1mM TCEP和40mM NaCl,pH 8.5的缓冲液进行透析。通过SUMO蛋白酶切割5D、E3、7F、B39、GTD^{VPI10463}、TcdB¹⁰⁷²⁻¹⁴³³和TD1的His₆-SUMO标签。这些蛋白以及TcdB全毒素和具有未切割His-标签的GTD,通过MonoQ离子交换层析法(GE Healthcare)在含有20mM Tris,pH8.5的缓冲液中进一步纯化,并采用NaCl梯度洗脱。通过SUMO蛋白酶切割后,TcdB¹⁻¹⁸⁰⁵使用链霉亲和素树脂进一步纯化。

[0079] 将纯化的TcdB全毒素与3种纯化的VHH以1:2:2:2的摩尔比在冰上混合2小时,从而组装TcdB-5D-E3-7F复合物。然后通过MonoQ离子交换层析法在20mM Tris,pH 8.5缓冲液中

纯化复合物,接着在含20mM Tris、1mM TCEP和40mM NaCl,pH 8.5的缓冲液中进行Superose 6分子排阻层析(SEC;GE Healthcare)。通过将纯化的GTD、GTD^{VP110463}和TcdB¹⁰⁷²⁻¹⁴³³分别与E3、7F和5D以1:2的摩尔比在冰上混合2小时,制备GTD-E3、GTD^{VP110463}-7F、TcdB¹⁰⁷²⁻¹⁴³³-5D复合物。随后,使用MonoQ离子交换柱(20mM Tris,pH 8.5)和Superdex-200Increase SEC(20mM Tris、1mM TCEP和40mM NaCl,pH 8.5)进一步纯化。将所有蛋白质复合物浓缩至约10mg/mL,并于-80℃储存直至使用。

[0080] 如前文所述,由SSRL beamline 4-2进行与小角度X射线散射结合的串联在线分子排阻层析(SEC-SAXS)实验。将纯化的TcdB全毒素交换到含有磷酸盐缓冲盐水PBS和5mM DTT,pH 7.4的缓冲液或含20mM乙酸钠、50mM NaCl和5mM DTT,pH 5.0的缓冲液中,然后浓缩到20mg/mL。使用Superdex-200 Increase PC 3.2/300柱(GE Healthcare)分别在pH5.0和7.4下收集SEC-SAXS数据。

[0081] 对于TcdB的DSSO交联,PBS缓冲液(pH 7.4)中的TcdB全毒素(50μL,10μM)与DSSO以1:100的摩尔比在室温下反应1小时。通过加入50倍过量的碳酸氢铵来淬灭交联反应10分钟,并使用FASP方案对所得产物进行酶解。简而言之,如前文所述,将交联的蛋白质转移到Milipore MicroconTM Ultracel PL-30(30kDa过滤器)中,进行还原/烷基化,并且依次使用Lys-C/胰蛋白酶进行酶解。将所得消化产物脱盐并通过肽SEC分级。收集含有交联肽的分级产物,用于后续的MSⁿ分析。在这项工作中,进行了三次生物学重复。

[0082] 使用与Orbitrap Fusion LumosTM质谱仪在线结合的Thermo ScientificTM Dionex UltiMate 3000系统进行LC MSⁿ分析。使用50cm×75μm AcclaimTM PepMapTM C18柱以82分钟内1%至25%ACN的梯度、300nL/min的流速来分离肽。利用两种不同类型的采集方法来最大限度地鉴定DSSO交联肽。

[0083] 对于TcdB的单分子FRET分析,VHH-7F和B39都含有隐藏的二硫键,使得天然半胱氨酸无法标记。通过诱变将半胱氨酸残基引入到7F(-1位置处)的N-末端中或B39(G42C)的外露表面环中。除了纯化期间在所有缓冲液中加入了5mM DTT,突变体VHH的表达和纯化都与野生型蛋白类似。纯化的7F用受体染料(Alexa-647马来酰亚胺)标记,而B39用供体染料(Alexa-555马来酰亚胺)(Thermo Fisher Scientific)标记。通过UV-Vis光谱法测定的标记效率>90%。在pH 6.8下,使用的EZ-Link NHS-PEG4-生物素(Thermo Fisher Scientific)对纯化的5D进行生物素化,以优先标记N-末端胺。用Superose 6 SEC进一步纯化与Alexa-647标记的7F、Alexa-555标记的B39和生物素标记的5D复合的TcdB全毒素,以除去过量的VHH。

[0084] 在加入链霉亲和素之前,用生物素化的牛血清白蛋白以及2%Biolipicure 203和0.2%Biolipicure 206的混合物(NOF America公司)依次钝化已清洁的石英载玻片。经过这种处理后,在比用于实现单分子之间的光学分辨率的100pM浓度高几个数量级的浓度下,预制的TcdB-3VHH复合物未显示出与载玻片的非特异性结合。

[0085] 在如此低的蛋白质浓度下,非共价结合的VHH会部分解离,且测量过程要求在每个pH条件下进行七次重复的表面处理,所以必须快速进行测量。使用基于棱镜的全内反射荧光显微镜对样品进行成像。使用637nm激光二极管(Coherent公司,Santa Clara,CA)激发样品用于检测Alexa-647,且使用532nm二极管泵浦固体激光器(Laser Quantum USA.Fremont,CA)激发样品用于检测Alexa-555。使用包含645nm分色镜的Optosplit比率计

图像分离器(Cairn Research Ltd,Faversham,UK)分离来自供体和受体的发射光,其中585/70带通滤波器用于供体通道,670/30带通滤波器用于受体通道(IDEX Health&Science.,Rochester,NY)。将重复图像以10Hz的帧速率转发到单个iXon DU-897EMCCD照相机(Andor Technologies,Belfast,UK)。

[0086] 在自写MATLAB脚本中处理数据,以对重复图像进行交叉关联,并提取强度高于基线的衍射极限斑点的时间轨迹。根据各复合物的荧光强度随时间变化的轨迹,仅选择那些含有单一供体和受体染料且在单一时间步长中与基线呈反相关光漂白关系的复合物。根据反相关光漂白事件的程度,可以对每分子进行 γ 标准化,使得能够报告绝对的FRET效率。将FRET效率编制成与高斯函数拟合的直方图。

[0087] 为了确保FRET变化不是光物理变化的结果,测量了游离染料、染料标记的VHH和与TcdB复合的单个染料标记的VHH的相对量子产率和荧光各向异性。所有测量均使用与smFRET相同的缓冲液,在pH 7(50mM Hepes,100mM NaCl,pH 7)和pH 5(50mM乙酸钠,100mM NaCl,pH 5)下、染料浓度为10nM时进行。

[0088] 使用2.0mm的激发狭缝和2.0mm的发射狭缝在ISS PC1光子计数光谱仪上记录整体荧光。分别在637nm和532nm激发Alexa-555和Alexa-647标记的样品。使用相同比色皿测量吸光度,确定用于荧光反应的样品浓度。以最大值周围20nm窗口的总和为发射强度。通过将强度标准化为pH 7下游离染料的发射量来计算相对量子产率。使用2.0mm的激发狭缝和2.0mm的发射狭缝收集各向异性度测量值。分别记录567nm和670nm处供体和受体的发射值。所有测量均进行三次,并报告平均值和标准误差。

[0089] 使用Zetasizer Nano S(Malvern Panalytical)进行动态光散射(DLS)实验。室温下,在0.2mg/mL浓度下测定200 μ L比色皿内的PBS缓冲液中的TcdB。使用Zetasizer 7.13版软件分析数据。

[0090] 对于钙黄绿素染料释放试验,按照制造商的方案,使用Avanti Mini挤出机通过挤出法制备脂质体。简而言之,将指定摩尔比的脂质(Avanti Polar Lipid)在氯仿中混合,然后在氮气中干燥,并真空放置过夜。将干燥后的脂质进行再水化,并进行五轮冻融循环。按照制造商的说明书,使用Avanti Mini挤出机穿过200nm孔膜挤出来制备单室脂质体。

[0091] 将含有55%的1,2-二油酰基-sn-甘油-3-磷酸胆碱(DOPC)、15%的1,2-二油酰基-sn-甘油-3-磷酸-L-丝氨酸(DOPS)和30%的胆固醇(10mg/mL)的干燥的脂质重悬于含150mM NaCl、20mM Hepes、1mM EDTA、50mM钙黄绿素,pH 7.0的缓冲液中。通过脱盐(Zeba)将游离的钙黄绿素染料与包载钙黄绿素的脂质体分离。在Spectramax M2e比色皿模块上测量荧光,经493nm激发,发射光为525nm。在试验中,脂质体在含150mM NaCl、20mM乙酸钠、1mM EDTA,pH 4.6的缓冲液中稀释到最终浓度0.3mM,并孵育至荧光信号稳定。加入单独的TcdB(0-25nM)或用5D或7F预孵育的TcdB(TcdB:VHH摩尔比为1:2),并记录荧光强度7分钟。加入0.1%Trion X-100终止反应。荧光变化百分比计算方式为 $((F - F_{\text{初始}}) / (F_{\text{最终}} - F_{\text{初始}}))$ 。根据曲线的线性部分的斜率,推导得出钙黄绿素染料释放的初始速率。实验各重复三次。

[0092] 如前文所述的方法对膜去极化进行了测量,但做了一些修改。简而言之,在含200mM NaCl、1mM KCl和10mM Hepes,pH 7.0的缓冲液中制备由55%DOPC、15%DOPS和30%胆固醇组成的脂质体。为了产生跨膜正电位(+135mV),将脂质体在含200mM KCl、1mM NaCl和10mM乙酸钠,pH 4.6的缓冲液中稀释至最终浓度0.1mM。使用12 μ M ANS监测膜电位。0秒时

加入缬氨霉素至最终浓度30nM。180秒时,加入单独的100nM TcdB全毒素或加入用0.02-1 μ M 5D或1 μ M 7F预孵育的TcdB,经380nm激发后测量490nm荧光强度7分钟。加入来自Bacillus anerinolyticus (Sigma-Aldrich)的2 μ M短杆菌肽终止反应。存在短杆菌肽时,相对于最大变化值的荧光变化值的计算方式为 $((F-F_{\text{初始}})/(F_{\text{最终}}-F_{\text{初始}}))$ 。实验各重复三次。

[0093] 在25 μ L的20mM Tris-HCl缓冲液(pH 8.0)中进行TcdB自动处理试验,该缓冲液含有0.4 μ M的TcdB全毒素或TcdB¹⁻¹⁸⁰⁵、指定浓度的InsP6,可含有或不含7F(2 μ M)。将反应混合物于37℃孵育1小时,然后在SDS样品缓冲液中煮沸5分钟以终止反应。通过4-20%SDS-PAGE检测样品,通过考马斯蓝染色法对TcdB片段显色。

[0094] 全长TcdB的晶体结构

[0095] 来自艰难梭菌M68株的全长TcdB全毒素在经验证的巨大芽孢杆菌(Bacillus megaterium)系统中表达,并纯化至高均一度。经广泛的结晶筛选和优化,并使用同步加速器测试大量晶体后,在由TcdB和三个中和VHH(5D、E3和7F)组成的异四聚体复合物晶体上以3.87Å的分辨率收集到最佳X射线衍射数据。TcdB-VHH复合物在pH 5.2结晶,该pH值是胞内体中生理相关的pH值(图1A、图1B和表2)。除了两个小区域(第944-949位残基和第1032-1047位残基)由于极高的结构灵活性无可见电子密度外,已经构建出TcdB全毒素的完整结构(图1B)。

[0096] 晶体结构表明TcdB由三个主要成分组成。GTD(第1-544位残基)和CPD(第545-841位残基)形成中心部分,该中心部分涉及广泛的结构域间相互作用。Delivery/RBD(第842-1834位残基)形成一个延伸的模块,与一侧的GTD和CPD相互作用,并指向远离GTD/CPD的方向。最重大的发现是细长的CROP结构域(第1835-2367位残基),该结构域从CPD和Delivery/RBD的连接处出现,并在相反的方向上伸展约130Å,从而呈钩状围绕着GTD弯曲(图1B)。胞内体的pH条件下TcdB的总体结构与中性pH条件下衍生自EM研究的TcdB和TcdA的结构模型不同,其中CROP平行于Delivery/RBD并与其相互作用。此外,观察到TcdB的疏水成孔区域(Delivery/RBD中的第957-1129位残基)的构象与接近中性pH条件下的TcdA片段的构象不同。这可能代表了胞内体的pH条件下罕见的TcdB的成孔中间态,其被中和抗体(5D)“冻结”。

[0097] CROP结构域的独特结构

[0098] TcdB的CROP由两种类型的重复序列组成,包括20-23个残基的二十个短重复序列(称为SR)以及30个残基的四个长重复序列(称为LR)(图2A)。每个SR由后接柔性环的 β -发夹组成,而每个LR具有三个 β -股,其与前述SR的 β -发夹一起形成扭曲式反平行 β -折叠。因为由SR组成的 β -螺线管(β -solenoid)的直的、棒状的区段被穿插的LR打断,所以引起CROP弯曲,导致约132-146°的扭结(图2B,图2C)。从结构上来看,CROP可以划分为四个等同的单元(称为CROP I-IV),每个单元由SR1-SR2-SR3-LR-SR4-SR5模块组成(图2C)。CROP I-IV的叠加产生了约0.9-2.6Å的Ca均方根差(r.m.s.d.)。

[0099] 有趣的是,在与所有其它SR相同的Delivery/RBD的C末端发现了未被识别的SR模块(第1815-1834位残基)。这种新的SR与上游长环和短 α 螺旋一起形成结构不同的模块(第1792-1834位残基),由于该模块能将Delivery/RBD连接到细长的CROP,因此在本文中称为“铰链”。此外,铰链直接与CPD(第742-765位残基,称为 β -瓣)中的三链 β -折叠和3-螺旋束(第766-841位残基,称为3-HB)相互作用,其中三链 β -折叠对于CPD活化来说至关重要,并且其中3-螺旋束位于由GTD、CPD、Delivery/RBD和CROP包围的缝隙中(图2D,图2E)。因为铰链

处于关键位置,所以它主要负责介导TcdB的所有四个结构域之间的结构连通。该铰链的功能作用得到了早期研究的支持,表明该区域的缺失大大降低了毒性。另外,铰链附近的高变序列可使由高毒力的艰难梭菌027核糖体分型株和其它低毒力株制备的TcdB变体在毒性和抗原性方面存在差异。

[0100] TcdB在中性pH和酸性pH条件下显示出不同的结构

[0101] 因为TcdB全毒素的结构衍生自酸性pH条件下生长的晶体,因此在pH 5.0和pH 7.4时,使用与SAXS结合的在线分子排阻层析(SEC-SAXS)进一步检查其溶液结构。曲线拟合分析发现,基于该晶体结构计算的散射图谱与pH 5.0时的实验散射图谱几乎相同,表明TcdB的溶液结构与pH 5.0时的晶体结构相似。然而,pH 7.4时的实验SAXS数据的散射轮廓与晶体结构的计算图谱之间在中间角(中间q)区域上的不一致表明TcdB在中性pH条件下采用了不同的构象(图3A)。Guinier和P(r)分析发现,pH 5.0和7.4时的 R_g 值相似,但pH 5.0时的 D_{max} (约233.0Å)则大于pH 7.4时的数据(约205.0Å)。pH 5.0时的TcdB的 D_{max} 与根据该晶体结构预测的值(约247Å)相当。而pH 7.4时的TcdB全毒素的 D_{max} 较小,与由GTD、CPD和Delivery/RBD(约203Å)组成的TcdB核心的预测值相当。因此,这表明pH 7.4时,细长的CROP可以朝TcdB核心摆动,以采用更紧凑的构象。

[0102] 为了更好地研究pH 7.4时的CROP的构象,采用了XL-MS策略,使用DSSO(二琥珀酰亚胺亚砷,一种含亚砷的MS可切割交联剂)测定了TcdB的结构域间相互作用。pH 7.4时,在TcdB中总共鉴定出87种交联,代表27种结构域间相互作用和60种结构域内相互作用。其中,在GTD和CPD之间、GTD和Delivery/RBD之间、以及CPD和Delivery/RBD之间分别鉴定出8对、4对和8对独特的交联肽(图3B)。将XL-MS数据映射到该晶体结构时,几乎所有这些交联满足30Å的距离截止值,表明与TcdB的晶体结构有良好的相关性。

[0103] 有趣的是,在CROP和Delivery/RBD之间鉴定出7对交联肽,对应于在该晶体结构中测量的范围在90Å至210Å之间的Ca-Ca距离。这表明,在中性pH条件下,TcdB的CROP可以比在该晶体结构中观察到的更接近Delivery/RBD。具体地,K1965和K1977残基周围的CROP的中心部分和K2234和K2249残基周围的CROP的C-末端尖端必须能够在Delivery/RBD的约30Å范围内移动(图3C)。这种新构象将会与来源于SAXS研究的pH 7.4时的TcdB的 D_{max} 一致,并且与在中性pH条件下TcdA的“闭合”构象相似。由于XL-MS除了稳定结构之外还能够捕获动态和瞬时接触,因此仍然不清楚在中性pH条件下TcdB花费在“闭合的”TcdA样构象上的时间。

[0104] CROP的pH依赖性结构灵活性

[0105] 接着,使用smFRET检测CROP的pH依赖性构象变化。smFRET是一种用于检测蛋白质结构和构象变化的成熟方法,可以鉴定出异质混合物或动态混合物中的单个物质。由于TcdB具有九个半胱氨酸残基且C699对于CPD功能来说至关重要,因此将三个VHH(7F、B39和5D)作为分子工具,用来标记和捕获TcdB,而不是对毒素进行化学标记。具体地,将受体染料(Alexa-647)与在标记TcdB全毒素核心的7F的-1位置处引入的半胱氨酸残基连接。将供体染料(Alexa-555)与B39连接,该B39与CROP IV(PDB代码:4NC2)特异性结合。鉴于TcdB全毒素的结构,两个染料之间的距离约为47Å。这两个染料标记的VHH之间的能量转移用于监测CROP的移动(图3D)。基于整体FRET研究,生物素标记5D对TcdB构象变化没有影响,用于在钝化的石英显微镜载玻片上对TcdB进行免疫下拉。将三个VHH与TcdB预组装,并通过分子排阻

层析纯化复合物。

[0106] 根据各异四聚体TcdB-VHH复合物的荧光强度随时间变化的轨迹,仅选择那些含有单一供体和受体染料且在单一时间步长中与基线呈反相关光漂白关系的复合物。利用反相关光漂白事件的程度,对每个分子进行 γ 标准化,使得能够报告绝对的FRET效率。将FRET效率编制成揭示pH 5.0和7.0时的单一FRET峰值的直方图(图3E)。单一的峰值的存在将与静态结构或比100ms的时间分箱更快、不能通过单一的FRET对来区分的动态平均一致。

[0107] 在pH 5.0 (0.532 ± 0.015) 和pH 7.0 (0.484 ± 0.007) 时的平均FRET效率中观察到了统计学显著差异,该差异支持TcdB表现出pH依赖性构象变化(图3E)。根据pH 5.0时的染料标记VHH之间的平均FRET效率进行简单计算,得到了估算距离 $49.9 \pm 0.05 \text{ \AA}$,与TcdB全毒素在酸性pH(约 $47 \text{ \AA}$)条件下的晶体结构一致。在pH 5.5和pH 5.25时,观察到类似的结果。在pH 7.0时,平均FRET效率表明标记位点之间的距离增加到 $51.5 \pm 0.05 \text{ \AA}$ 。单一的FRET对不足以相对于TcdB的其余部分定位CROP,而构象动力学中的任何变化都会影响FRET到距离的简单转化。在pH 7.0 (0.113 ± 0.002) 时,与pH 5.0 (0.141 ± 0.026) 相比,表观平均FRET略微增加,同时分布宽度在统计学上显著减少了25%,与构象动力学速率的增加一致。

[0108] 迄今为止,在TcdB中已经确定了两个限制性结构状态:由晶体结构、SAXS和smFRET研究所支持的酸性pH条件下的“开放式”构象,以及由SAXS和XL-MS研究揭示的中性pH条件下的“闭合式”构象(图3D)。这些数据共同表明,在中性pH条件下,CROP相对于TcdB的核心可对整体构象进行采样,并且这样的蛋白质动力学在smFRET中100ms的积分时间内无法分辨。由于CROP和TcdB核心之间缺乏稳定接触,并且连接Delivery/RBD和CROP的铰链中可能存在结构重排,应该允许这样的构象采样。

[0109] TcdB在胞内体pH条件下的成孔中间态

[0110] Delivery/RBD用于保护疏水性成孔区域(第957-1129位残基),预期在胞内体酸化后释放以形成将GTD和CPD递送至细胞质的孔隙。TcdB的成孔活性也有助于体外观察到的细胞坏死。酸性pH条件下的TcdB全毒素与中性pH条件下的TcdA片段之间的结构对比揭示了成孔区域(TcdB中的第1032-1134位残基和TcdA中的第1033-1135位残基)的同源C-末端部分存在巨大差异(图4A,图4B)。在TcdA中,该区域采用了混合 α/β 构型,其中疏水残基被保护在主要由Delivery/RBD中的 β -折叠形成的连续凹槽中(图4C,图4D)。然而,在TcdB的酸性构象中,可能是由于极高的灵活性,第1032-1047位残基无可见电子密度,表明这些残基在胞内体pH条件下展开并从毒素核心脱离。此外,TcdB中相当于TcdA中 $\alpha 2$ 的残基展开成环,而TcdB中相当于TcdA中 $\beta 3$ 和 $\alpha 3$ 的一部分的残基组装成新的占据与TcdA中原始 $\alpha 3$ 相同的区域的螺旋。由于这种转变,TcdB中的疏水残基(第1084-1094位残基,相当于TcdA中 $\alpha 3$ 的C末端部分)作为延伸环突出。有趣的是,构象变化没有扩散到TcdB与5D结合的区域,5D的构象与TcdA中观察到的构象保持相似。

[0111] 为了进一步分析酸性pH和5D对成孔区域中观察到的构象变化的影响,测定了pH 8.5时Delivery/RBD片段(即与5D复合的TcdB¹⁰⁷²⁻¹⁴³³)的晶体结构(表2)。发现pH 8.5时在TcdB¹⁰⁷²⁻¹⁴³³中观察到的成孔区域采用了TcdA样中性pH构象。因此,这一发现表明,在TcdB全毒素中观察到的成孔区域的新构象可能代表由胞内体pH诱导的中间态。

[0112] 此外,还发现5D与TcdB在pH 8.5和5.2时的结合模式几乎相同,涉及了5D的所有三个互补决定区(CDR)。5D的整体结合亲和力通过广泛的极性相互作用和疏水相互作用而进

一步增强,这些相互作用涉及成孔区域外的TcdB残基。因此,5D可以固定TcdB中 β 4- β 5的构象,防止 β 4- β 5- α 4模块中发生pH诱导的构象变化。先前的诱变研究表明,在TcdB中5D结合位点周围引入的突变有效地抑制了成孔和细胞毒性,而仅突变5D靶向的L1107 (L1107K) 即可导致毒性降低>1,000倍。这些发现表明,5D很可能抑制TcdB在胞内体pH条件下成孔所必需的构象变化。

[0113] 为了验证这一假设,使用了两次互补试验检测5D是如何影响TcdB的膜插入。通过监测TcdB渗透装载钙黄绿素的脂质体的能力,发现了pH 4.6时TcdB以蛋白质浓度依赖性方式增加了钙黄绿素的释放速率(图4E)。当TcdB与5D一起预孵育时,TcdB诱导的染料释放速率显著降低。与GTD结合的7F作为对照物对TcdB诱导的染料释放没有影响。进一步研究了5D或7F对TcdB消散脂质体中缬氨霉素诱导的膜电位的影响,发现是5D而非7F降低了TcdB的膜去极化能力(图4F)。

[0114] 综上所述,这些发现表明5D通过阻止成孔区域完成必要的pH诱导构象变化来中和TcdB。值得注意的是,5D识别的成孔区域在大梭菌葡糖基化毒素(LCGT)家族中高度保守,该家族包括TcdA和TcdB、诺维梭菌(*C. novyi*) α 毒素(Tcn α)、索氏梭菌(*C. sordellii*)致死性和出血性毒素(TcsL和TcsH)和产气荚膜梭菌(*C. perfringens*)毒素(TpeL)(图4C)。因此,成孔区域的这一部分代表了用于开发靶向TcdA、TcdB、其它LCGT的广谱疫苗和抗体的良好靶标、或其它合适靶标。

[0115] TcdB自动处理的调节

[0116] 进入细胞后,InsP6对CPD的激活是调节TcdA和TcdB病理的关键步骤。总的来说,除了 β -瓣外,TcdB全毒素中apo-CPD和与InsP6结合的CPD片段(PDB:3PEE)的结构非常相似(r.m.s.d.约1.1Å)(图5A,图5B)。将apo-CPD的结构与基于TcdB切割序列($G^{542}SL^{544}$)(PDB:3PA8)结合到InsP6或肽抑制剂的CPD片段的结构进行比较,发现 β -瓣部分占据了TcdB全毒素中CPD的P1底物袋,从而阻止底物结合。在CPD片段中,InsP6使 β -瓣旋转约90°(图5B),从而通过正确排列活性位点和底物袋来活化CPD。然而,在TcdB全毒素中 β -瓣的这种旋转是禁止的,否则其会与随后的3-HB发生空间冲突(图5D)。

[0117] 除了通过InsP6进行变构调节之外,一些研究表明,CROP还影响着TcdB的自动处理。使用TcdB全毒素和不含铰链和CROP(第1-1805残基)的截短型TcdB,对InsP6诱导的GTD切割效率进行了比较。发现InsP6诱导的GTD切割在TcdB的第¹⁻¹⁸⁰⁵位残基中的效率更高,表明CROP和铰链有助于抑制TcdB全毒素中的CPD功能。此外,在不存在CROP时,相比不含铰链(第1-1795位残基)的TcdB片段,含有铰链(第1-1832位残基)的TcdB片段显示出更弱的InsP6依赖性GTD切割。这些数据表明铰链参与了TcdB自动处理的调节。值得注意的是,在TcdB全毒素中,铰链与 β -瓣和3-HB相互作用,一起形成了TcdB中连接所有四个结构域的“心脏”(图5C,图5D)。由于 β -瓣和3-HB对于InsP6结合和CPD激活之间的偶联非常重要,因此与pH依赖性CROP移动相关的铰链的结构重排可有助于CPD功能的调节。

[0118] VHH 7F和E3揭示了在GTD上的两个不同的中和表位

[0119] 7F抑制GTD切割,但不直接与CPD相互作用。相反,7F与紧邻切割位点(L544)的GTD的C-末端结合。值得注意的是,7F的CDR3通过广泛的极性相互作用和疏水相互作用与易切键上游的 α 螺旋(第525-539位残基)以及相邻的 α 螺旋(第137-158位残基)结合。这种相互作用干扰易切键向CPD切割位点移动以及GTD相对于CPD的正确定位,从而抑制GTD切割。

[0120] E3通过特异性靶向GTD来抑制Rho葡糖基化并阻断TcdB的细胞病变作用。在使用GTD片段或TcdB全毒素的两个独立解析的晶体结构中,E3以类似方式与N-末端四螺旋束(第1-90位残基)结合。更具体地,E3通过广泛的极性相互作用和疏水相互作用识别GTD中的第2和第3螺旋(第21-64位残基)。由于GTD-Rho复合物的结构尚未有报道,所以仍不清楚E3是如何影响GTD-Rho的相互作用或催化作用的。在其它LCGT成员的葡糖基转移酶结构域中也发现了同源四螺旋束,其可能涉及葡糖基转移酶结构域的质膜结合,表明E3可干扰GTD的膜结合。因此,GTD-E3复合体的结构为进一步验证和利用这些机制作为对抗TcdB和潜在的其它LCGT成员的新策略奠定了基础。

[0121] 尽管已经示出和描述了本发明的优选实施例,但是对于本领域技术人员来说显而易见的是,可以对其进行不超出所附权利要求的范围的修改。因此,本发明的范围仅由以下权利要求限定。在一些实施例中,在本专利申请中呈现的附图按比例绘制,包括角度、尺寸比等。在一些实施例中,附图仅是代表性的,并且权利要求不受附图的尺寸的限制。在一些实施方案中,本文中使用短语“包含”描述的本发明包括可以描述为“基本上由...组成”或“由...组成”的实施方案,并且因此满足要求保护使用短语“基本上由...组成”或“由...组成”的本发明的一个或多个实施方案的书面描述要求。

[0122] 在以下权利要求中引用的附图标记仅仅是为了易于审查本专利申请,且为示例性的,并且不旨在以任何方式将权利要求的范围限制到在附图中具有对应附图标记的特定特征。

SEQUENCE LISTING

<110> The Regents of the University of California

Jin, Rongsheng

<120> 艰难梭菌的疫苗组合物

<130> P21119206WP

<150> 62/851,040

<151> 2019-05-21

<160> 22

<170> PatentIn version 3.5

<210> 1

<211> 2366

<212> PRT

<213> Clostridium difficile

<400> 1

```

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

```

Gln Lys Glu Glu Asn Pro Asp Leu Ile Ile Asp Asp Ile Val Lys Thr		
195	200	205
Tyr Leu Ser Asn Glu Tyr Ser Lys Asp Ile Asp Glu Leu Asn Ala Tyr		
210	215	220
Ile Glu Glu Ser Leu Asn Lys Val Thr Glu Asn Ser Gly Asn Asp Val		
225	230	235
Arg Asn Phe Glu Glu Phe Lys Thr Gly Glu Val Phe Asn Leu Tyr Glu		
	245	250
		255
Gln Glu Leu Val Glu Arg Trp Asn Leu Ala Gly Ala Ser Asp Ile Leu		
	260	265
		270
Arg Val Ile Leu Lys Asn Ile Gly Gly Val Tyr Leu Asp Val Asp Met		
	275	280
		285
Leu Pro Gly Ile His Pro Asp Leu Phe Lys Asp Ile Asn Lys Pro Asp		
	290	295
		300
Ser Val Lys Thr Ala Val Asp Trp Glu Glu Met Gln Leu Glu Ala Ile		
305	310	315
		320
Met Lys His Lys Glu Tyr Ile Pro Glu Tyr Thr Ser Lys His Phe Asp		
	325	330
		335
Thr Leu Asp Glu Glu Val Gln Ser Ser Phe Glu Ser Val Leu Ala Ser		
	340	345
		350
Lys Ser Asp Lys Ser Glu Ile Phe Leu Pro Leu Gly Asp Ile Glu Val		
	355	360
		365
Ser Pro Leu Glu Val Lys Ile Ala Phe Ala Lys Gly Ser Ile Ile Asn		
	370	375
		380
Gln Ala Leu Ile Ser Ala Lys Asp Ser Tyr Cys Ser Asp Leu Leu Ile		
385	390	395
		400
Lys Gln Ile Gln Asn Arg Tyr Lys Ile Leu Asn Asp Thr Leu Gly Pro		
	405	410
		415
Ile Ile Ser Gln Gly Asn Asp Phe Asn Thr Thr Met Asn Asn Phe Gly		
	420	425
		430
Glu Ser Leu Gly Ala Ile Ala Asn Glu Glu Asn Ile Ser Phe Ile Ala		
	435	440
		445
Lys Ile Gly Ser Tyr Leu Arg Val Gly Phe Tyr Pro Glu Ala Asn Thr		
	450	455
		460
Thr Ile Thr Leu Ser Gly Pro Thr Ile Tyr Ala Gly Ala Tyr Lys Asp		
465	470	475
		480
Leu Leu Thr Phe Lys Glu Met Ser Ile Asp Thr Ser Ile Leu Ser Ser		
	485	490
		495
Glu Leu Arg Asn Phe Glu Phe Pro Lys Val Asn Ile Ser Gln Ala Thr		

500	505	510
Glu Gln Glu Lys Asn Ser Leu Trp	Gln Phe Asn Glu Glu Arg Ala Lys	
515	520	525
Ile Gln Phe Glu Glu Tyr Lys Lys	Asn Tyr Phe Glu Gly Ala Leu Gly	
530	535	540
Glu Asp Asp Asn Leu Asp Phe Ser	Gln Asn Thr Val Thr Asp Lys Glu	
545	550	555
Tyr Leu Leu Glu Lys Ile Ser Ser	Ser Thr Lys Ser Ser Glu Arg Gly	
565	570	575
Tyr Val His Tyr Ile Val Gln Leu	Gln Gly Asp Lys Ile Ser Tyr Glu	
580	585	590
Ala Ala Cys Asn Leu Phe Ala Lys	Asn Pro Tyr Asp Ser Ile Leu Phe	
595	600	605
Gln Lys Asn Ile Glu Asp Ser Glu	Val Ala Tyr Tyr Tyr Asn Pro Thr	
610	615	620
Asp Ser Glu Ile Gln Glu Ile Asp	Lys Tyr Arg Ile Pro Asp Arg Ile	
625	630	635
Ser Asp Arg Pro Lys Ile Lys Leu	Thr Phe Ile Gly His Gly Lys Ala	
645	650	655
Glu Phe Asn Thr Asp Ile Phe Ala	Gly Leu Asp Val Asp Ser Leu Ser	
660	665	670
Ser Glu Ile Glu Thr Ala Ile Gly	Leu Ala Lys Glu Asp Ile Ser Pro	
675	680	685
Lys Ser Ile Glu Ile Asn Leu Leu	Gly Cys Asn Met Phe Ser Tyr Ser	
690	695	700
Val Asn Val Glu Glu Thr Tyr Pro	Gly Lys Leu Leu Leu Arg Val Lys	
705	710	715
Asp Lys Val Ser Glu Leu Met Pro	Ser Met Ser Gln Asp Ser Ile Ile	
725	730	735
Val Ser Ala Asn Gln Tyr Glu Val	Arg Ile Asn Ser Glu Gly Arg Arg	
740	745	750
Glu Leu Leu Asp His Ser Gly Glu	Trp Ile Asn Lys Glu Glu Ser Ile	
755	760	765
Ile Lys Asp Ile Ser Ser Lys Glu	Tyr Ile Ser Phe Asn Pro Lys Glu	
770	775	780
Asn Lys Ile Ile Val Lys Ser Lys	Asn Leu Pro Glu Leu Ser Thr Leu	
785	790	795
Leu Gln Glu Ile Arg Asn Asn Ser	Asn Ser Ser Asp Ile Glu Leu Glu	
805	810	815

Glu Lys Val Met Leu Ala Glu Cys Glu Ile Asn Val Ile Ser Asn Ile		
820	825	830
Glu Thr Gln Val Val Glu Glu Arg Ile Glu Glu Ala Lys Ser Leu Thr		
835	840	845
Ser Asp Ser Ile Asn Tyr Ile Lys Asn Glu Phe Lys Leu Ile Glu Ser		
850	855	860
Ile Ser Asp Ala Leu Cys Asp Leu Lys Gln Gln Asn Glu Leu Glu Asp		
865	870	875
Ser His Phe Ile Ser Phe Glu Asp Ile Ser Glu Thr Asp Glu Gly Phe		
	885	890
Ser Ile Arg Phe Ile Asn Lys Glu Thr Gly Glu Ser Ile Phe Val Glu		
	900	905
Thr Glu Lys Thr Ile Phe Ser Glu Tyr Ala Asn His Ile Thr Glu Glu		
	915	920
Ile Ser Lys Ile Lys Gly Thr Ile Phe Asp Thr Val Asn Gly Lys Leu		
	930	935
Val Lys Lys Val Asn Leu Asp Thr Thr His Glu Val Asn Thr Leu Asn		
945	950	955
Ala Ala Phe Phe Ile Gln Ser Leu Ile Glu Tyr Asn Ser Ser Lys Glu		
	965	970
Ser Leu Ser Asn Leu Ser Val Ala Met Lys Val Gln Val Tyr Ala Gln		
	980	985
Leu Phe Ser Thr Gly Leu Asn Thr Ile Thr Asp Ala Ala Lys Val Val		
	995	1000
Glu Leu Val Ser Thr Ala Leu Asp Glu Thr Ile Asp Leu Leu Pro		
	1010	1015
Thr Leu Ser Glu Gly Leu Pro Ile Ile Ala Thr Ile Ile Asp Gly		
	1025	1030
Val Ser Leu Gly Ala Ala Ile Lys Glu Leu Ser Glu Thr Ser Asp		
	1040	1045
Pro Leu Leu Arg Gln Glu Ile Glu Ala Lys Ile Gly Ile Met Ala		
	1055	1060
Val Asn Leu Thr Thr Ala Thr Thr Ala Ile Ile Thr Ser Ser Leu		
	1070	1075
Gly Ile Ala Ser Gly Phe Ser Ile Leu Leu Val Pro Leu Ala Gly		
	1085	1090
Ile Ser Ala Gly Ile Pro Ser Leu Val Asn Asn Glu Leu Val Leu		
	1100	1105
Arg Asp Lys Ala Thr Lys Val Val Asp Tyr Phe Lys His Val Ser		

1115	1120	1125
Leu Val Glu Thr Glu Gly	Val Phe Thr Leu Leu	Asp Asp Lys Val
1130	1135	1140
Met Met Pro Gln Asp Asp	Leu Val Ile Ser Glu	Ile Asp Phe Asn
1145	1150	1155
Asn Asn Ser Ile Val Leu	Gly Lys Cys Glu Ile	Trp Arg Met Glu
1160	1165	1170
Gly Gly Ser Gly His Thr	Val Thr Asp Asp Ile	Asp His Phe Phe
1175	1180	1185
Ser Ala Pro Ser Ile Thr	Tyr Arg Glu Pro His	Leu Ser Ile Tyr
1190	1195	1200
Asp Val Leu Glu Val Gln	Lys Glu Glu Leu Asp	Leu Ser Lys Asp
1205	1210	1215
Leu Met Val Leu Pro Asn	Ala Pro Asn Arg Val	Phe Ala Trp Glu
1220	1225	1230
Thr Gly Trp Thr Pro Gly	Leu Arg Ser Leu Glu	Asn Asp Gly Thr
1235	1240	1245
Lys Leu Leu Asp Arg Ile	Arg Asp Asn Tyr Glu	Gly Glu Phe Tyr
1250	1255	1260
Trp Arg Tyr Phe Ala Phe	Ile Ala Asp Ala Leu	Ile Thr Thr Leu
1265	1270	1275
Lys Pro Arg Tyr Glu Asp	Thr Asn Ile Arg Ile	Asn Leu Asp Ser
1280	1285	1290
Asn Thr Arg Ser Phe Ile	Val Pro Ile Ile Thr	Thr Glu Tyr Ile
1295	1300	1305
Arg Glu Lys Leu Ser Tyr	Ser Phe Tyr Gly Ser	Gly Gly Thr Tyr
1310	1315	1320
Ala Leu Ser Leu Ser Gln	Tyr Asn Met Gly Ile	Asn Ile Glu Leu
1325	1330	1335
Ser Glu Ser Asp Val Trp	Ile Ile Asp Val Asp	Asn Val Val Arg
1340	1345	1350
Asp Val Thr Ile Glu Ser	Asp Lys Ile Lys Lys	Gly Asp Leu Ile
1355	1360	1365
Glu Gly Ile Leu Ser Thr	Leu Ser Ile Glu Glu	Asn Lys Ile Ile
1370	1375	1380
Leu Asn Ser His Glu Ile	Asn Phe Ser Gly Glu	Val Asn Gly Ser
1385	1390	1395
Asn Gly Phe Val Ser Leu	Thr Phe Ser Ile Leu	Glu Gly Ile Asn
1400	1405	1410

Ala Ile Ile Glu Val Asp Leu Leu Ser Lys Ser Tyr Lys Leu Leu		
1415	1420	1425
Ile Ser Gly Glu Leu Lys Ile Leu Met Leu Asn Ser Asn His Ile		
1430	1435	1440
Gln Gln Lys Ile Asp Tyr Ile Gly Phe Asn Ser Glu Leu Gln Lys		
1445	1450	1455
Asn Ile Pro Tyr Ser Phe Val Asp Ser Glu Gly Lys Glu Asn Gly		
1460	1465	1470
Phe Ile Asn Gly Ser Thr Lys Glu Gly Leu Phe Val Ser Glu Leu		
1475	1480	1485
Pro Asp Val Val Leu Ile Ser Lys Val Tyr Met Asp Asp Ser Lys		
1490	1495	1500
Pro Ser Phe Gly Tyr Tyr Ser Asn Asn Leu Lys Asp Val Lys Val		
1505	1510	1515
Ile Thr Lys Asp Asn Val Asn Ile Leu Thr Gly Tyr Tyr Leu Lys		
1520	1525	1530
Asp Asp Ile Lys Ile Ser Leu Ser Leu Thr Leu Gln Asp Glu Lys		
1535	1540	1545
Thr Ile Lys Leu Asn Ser Val His Leu Asp Glu Ser Gly Val Ala		
1550	1555	1560
Glu Ile Leu Lys Phe Met Asn Arg Lys Gly Ser Thr Asn Thr Ser		
1565	1570	1575
Asp Ser Leu Met Ser Phe Leu Glu Ser Met Asn Ile Lys Ser Ile		
1580	1585	1590
Phe Val Asn Phe Leu Gln Ser Asn Ile Lys Phe Ile Leu Asp Ala		
1595	1600	1605
Asn Phe Ile Ile Ser Gly Thr Thr Ser Ile Gly Gln Phe Glu Phe		
1610	1615	1620
Ile Cys Asp Glu Asn Asn Asn Ile Gln Pro Tyr Phe Ile Lys Phe		
1625	1630	1635
Asn Thr Leu Glu Thr Asn Tyr Thr Leu Tyr Val Gly Asn Arg Gln		
1640	1645	1650
Asn Met Ile Val Glu Pro Asn Tyr Asp Leu Asp Asp Ser Gly Asp		
1655	1660	1665
Ile Ser Ser Thr Val Ile Asn Phe Ser Gln Lys Tyr Leu Tyr Gly		
1670	1675	1680
Ile Asp Ser Cys Val Asn Lys Val Val Ile Ser Pro Asn Ile Tyr		
1685	1690	1695
Thr Asp Glu Ile Asn Ile Thr Pro Val Tyr Glu Thr Asn Asn Thr		

1700	1705	1710
Tyr Pro Glu Val Ile Val	Leu Asp Ala Asn Tyr	Ile Asn Glu Lys
1715	1720	1725
Ile Asn Val Asn Ile Asn	Asp Leu Ser Ile Arg	Tyr Val Trp Ser
1730	1735	1740
Asn Asp Gly Asn Asp Phe	Ile Leu Met Ser Thr	Ser Glu Glu Asn
1745	1750	1755
Lys Val Ser Gln Val Lys	Ile Arg Phe Val Asn	Val Phe Lys Asp
1760	1765	1770
Lys Thr Leu Ala Asn Lys	Leu Ser Phe Asn Phe	Ser Asp Lys Gln
1775	1780	1785
Asp Val Pro Val Ser Glu	Ile Ile Leu Ser Phe	Thr Pro Ser Tyr
1790	1795	1800
Tyr Glu Asp Gly Leu Ile	Gly Tyr Asp Leu Gly	Leu Val Ser Leu
1805	1810	1815
Tyr Asn Glu Lys Phe Tyr	Ile Asn Asn Phe Gly	Met Met Val Ser
1820	1825	1830
Gly Leu Ile Tyr Ile Asn	Asp Ser Leu Tyr Tyr	Phe Lys Pro Pro
1835	1840	1845
Val Asn Asn Leu Ile Thr	Gly Phe Val Thr Val	Gly Asp Asp Lys
1850	1855	1860
Tyr Tyr Phe Asn Pro Ile	Asn Gly Gly Ala Ala	Ser Ile Gly Glu
1865	1870	1875
Thr Ile Ile Asp Asp Lys	Asn Tyr Tyr Phe Asn	Gln Ser Gly Val
1880	1885	1890
Leu Gln Thr Gly Val Phe	Ser Thr Glu Asp Gly	Phe Lys Tyr Phe
1895	1900	1905
Ala Pro Ala Asn Thr Leu	Asp Glu Asn Leu Glu	Gly Glu Ala Ile
1910	1915	1920
Asp Phe Thr Gly Lys Leu	Ile Ile Asp Glu Asn	Ile Tyr Tyr Phe
1925	1930	1935
Glu Asp Asn Tyr Arg Gly	Ala Val Glu Trp Lys	Glu Leu Asp Gly
1940	1945	1950
Glu Met His Tyr Phe Ser	Pro Glu Thr Gly Lys	Ala Phe Lys Gly
1955	1960	1965
Leu Asn Gln Ile Gly Asp	Asp Lys Tyr Tyr Phe	Asn Ser Asp Gly
1970	1975	1980
Val Met Gln Lys Gly Phe	Val Ser Ile Asn Asp	Asn Lys His Tyr
1985	1990	1995

Phe Asp Asp Ser Gly Val Met Lys Val Gly Tyr Thr Glu Ile Asp		
2000	2005	2010
Gly Lys His Phe Tyr Phe Ala Glu Asn Gly Glu Met Gln Ile Gly		
2015	2020	2025
Val Phe Asn Thr Glu Asp Gly Phe Lys Tyr Phe Ala His His Asn		
2030	2035	2040
Glu Asp Leu Gly Asn Glu Glu Gly Glu Glu Ile Ser Tyr Ser Gly		
2045	2050	2055
Ile Leu Asn Phe Asn Asn Lys Ile Tyr Tyr Phe Asp Asp Ser Phe		
2060	2065	2070
Thr Ala Val Val Gly Trp Lys Asp Leu Glu Asp Gly Ser Lys Tyr		
2075	2080	2085
Tyr Phe Asp Glu Asp Thr Ala Glu Ala Tyr Ile Gly Leu Ser Leu		
2090	2095	2100
Ile Asn Asp Gly Gln Tyr Tyr Phe Asn Asp Asp Gly Ile Met Gln		
2105	2110	2115
Val Gly Phe Val Thr Ile Asn Asp Lys Val Phe Tyr Phe Ser Asp		
2120	2125	2130
Ser Gly Ile Ile Glu Ser Gly Val Gln Asn Ile Asp Asp Asn Tyr		
2135	2140	2145
Phe Tyr Ile Asp Asp Asn Gly Ile Val Gln Ile Gly Val Phe Asp		
2150	2155	2160
Thr Ser Asp Gly Tyr Lys Tyr Phe Ala Pro Ala Asn Thr Val Asn		
2165	2170	2175
Asp Asn Ile Tyr Gly Gln Ala Val Glu Tyr Ser Gly Leu Val Arg		
2180	2185	2190
Val Gly Glu Asp Val Tyr Tyr Phe Gly Glu Thr Tyr Thr Ile Glu		
2195	2200	2205
Thr Gly Trp Ile Tyr Asp Met Glu Asn Glu Ser Asp Lys Tyr Tyr		
2210	2215	2220
Phe Asp Pro Glu Thr Lys Lys Ala Cys Lys Gly Ile Asn Leu Ile		
2225	2230	2235
Asp Asp Ile Lys Tyr Tyr Phe Asp Glu Lys Gly Ile Met Arg Thr		
2240	2245	2250
Gly Leu Ile Ser Phe Glu Asn Asn Asn Tyr Tyr Phe Asn Glu Asn		
2255	2260	2265
Gly Glu Met Gln Phe Gly Tyr Ile Asn Ile Glu Asp Lys Met Phe		
2270	2275	2280
Tyr Phe Gly Glu Asp Gly Val Met Gln Ile Gly Val Phe Asn Thr		

2285	2290	2295
Pro Asp Gly Phe Lys Tyr	Phe Ala His Gln Asn	Thr Leu Asp Glu
2300	2305	2310
Asn Phe Glu Gly Glu Ser	Ile Asn Tyr Thr Gly	Trp Leu Asp Leu
2315	2320	2325
Asp Glu Lys Arg Tyr Tyr	Phe Thr Asp Glu Tyr	Ile Ala Ala Thr
2330	2335	2340
Gly Ser Val Ile Ile Asp	Gly Glu Glu Tyr Tyr	Phe Asp Pro Asp
2345	2350	2355
Thr Ala Gln Leu Val Ile	Ser Glu	
2360	2365	
<210> 2		
<211> 543		
<212> PRT		
<213> Clostridium difficile		
<400> 2		
Met Ser Leu Val Asn Arg Lys Gln Leu Glu Lys Met Ala Asn Val Arg		
1	5	10
Phe Arg Val Gln Glu Asp Glu Tyr Val Ala Ile Leu Asp Ala Leu Glu		
20	25	30
Glu Tyr His Asn Met Ser Glu Asn Thr Val Val Glu Lys Tyr Leu Lys		
35	40	45
Leu Lys Asp Ile Asn Ser Leu Thr Asp Thr Tyr Ile Asp Thr Tyr Lys		
50	55	60
Lys Ser Gly Arg Asn Lys Ala Leu Lys Lys Phe Lys Glu Tyr Leu Val		
65	70	75
Ile Glu Ile Leu Glu Leu Lys Asn Ser Asn Leu Thr Pro Val Glu Lys		
85	90	95
Asn Leu His Phe Ile Trp Ile Gly Gly Gln Ile Asn Asp Thr Ala Ile		
100	105	110
Asn Tyr Ile Asn Gln Trp Lys Asp Val Asn Ser Asp Tyr Asn Val Asn		
115	120	125
Val Phe Tyr Asp Ser Asn Ala Phe Leu Ile Asn Thr Leu Lys Lys Thr		
130	135	140
Ile Ile Glu Ser Ala Ser Asn Asp Thr Leu Glu Ser Phe Arg Glu Asn		
145	150	155
Leu Asn Asp Pro Glu Phe Asn His Thr Ala Phe Phe Arg Lys Arg Met		
165	170	175
Gln Ile Ile Tyr Asp Lys Gln Gln Asn Phe Ile Asn Tyr Tyr Lys Ala		

			180				185				190					
Gln	Lys	Glu	Glu	Asn	Pro	Asp	Leu	Ile	Ile	Asp	Asp	Ile	Val	Lys	Thr	
			195				200				205					
Tyr	Leu	Ser	Asn	Glu	Tyr	Ser	Lys	Asp	Ile	Asp	Glu	Leu	Asn	Ala	Tyr	
			210				215				220					
Ile	Glu	Glu	Ser	Leu	Asn	Lys	Val	Thr	Glu	Asn	Ser	Gly	Asn	Asp	Val	
225							230				235				240	
Arg	Asn	Phe	Glu	Glu	Phe	Lys	Thr	Gly	Glu	Val	Phe	Asn	Leu	Tyr	Glu	
			245				250				255					
Gln	Glu	Leu	Val	Glu	Arg	Trp	Asn	Leu	Ala	Gly	Ala	Ser	Asp	Ile	Leu	
			260				265				270					
Arg	Val	Ala	Ile	Leu	Lys	Asn	Ile	Gly	Gly	Val	Tyr	Leu	Asp	Val	Asp	
			275				280				285					
Met	Leu	Pro	Gly	Ile	His	Pro	Asp	Leu	Phe	Lys	Asp	Ile	Asn	Lys	Pro	
290							295				300					
Asp	Ser	Val	Lys	Thr	Ala	Val	Asp	Trp	Glu	Glu	Met	Gln	Leu	Glu	Ala	
305							310				315				320	
Ile	Met	Lys	His	Lys	Glu	Tyr	Ile	Pro	Glu	Tyr	Thr	Ser	Lys	His	Phe	
			325				330				335					
Asp	Thr	Leu	Asp	Glu	Glu	Val	Gln	Ser	Ser	Phe	Glu	Ser	Val	Leu	Ala	
			340				345				350					
Ser	Lys	Ser	Asp	Lys	Ser	Glu	Ile	Phe	Leu	Pro	Leu	Gly	Asp	Ile	Glu	
			355				360				365					
Val	Ser	Pro	Leu	Glu	Val	Lys	Ile	Ala	Phe	Ala	Lys	Gly	Ser	Ile	Ile	
370							375				380					
Asn	Gln	Ala	Leu	Ile	Ser	Ala	Lys	Asp	Ser	Tyr	Cys	Ser	Asp	Leu	Leu	
385							390				395				400	
Ile	Lys	Gln	Ile	Gln	Asn	Arg	Tyr	Lys	Ile	Leu	Asn	Asp	Thr	Leu	Gly	
			405				410				415					
Pro	Ile	Ile	Ser	Gln	Gly	Asn	Asp	Phe	Asn	Thr	Thr	Met	Asn	Asn	Phe	
			420				425				430					
Gly	Glu	Ser	Leu	Gly	Ala	Ile	Ala	Asn	Glu	Glu	Asn	Ile	Ser	Phe	Ile	
			435				440				445					
Ala	Lys	Ile	Gly	Ser	Tyr	Leu	Arg	Val	Gly	Phe	Tyr	Pro	Glu	Ala	Asn	
450							455				460					
Thr	Thr	Ile	Thr	Leu	Ser	Gly	Pro	Thr	Ile	Tyr	Ala	Gly	Ala	Tyr	Lys	
465							470				475				480	
Asp	Leu	Leu	Thr	Phe	Lys	Glu	Met	Ser	Ile	Asp	Thr	Ser	Ile	Leu	Ser	
			485				490				495					

Ser	Glu	Leu	Arg	Asn	Phe	Glu	Phe	Pro	Lys	Val	Asn	Ile	Ser	Gln	Ala
500				505				510							
Thr	Glu	Gln	Glu	Lys	Asn	Ser	Leu	Trp	Gln	Phe	Asn	Glu	Glu	Arg	Ala
515				520				525							
Lys	Ile	Gln	Phe	Glu	Glu	Tyr	Lys	Lys	Asn	Tyr	Phe	Glu	Gly	Ala	
530				535				540							
<210> 3															
<211> 381															
<212> PRT															
<213> Clostridium difficile															
<400> 3															
Leu	Thr	Thr	Ala	Thr	Thr	Ala	Ile	Ile	Thr	Ser	Ser	Leu	Gly	Ile	Ala
1				5				10				15			
Ser	Gly	Phe	Ser	Ile	Leu	Leu	Val	Pro	Leu	Ala	Gly	Ile	Ser	Ala	Gly
20				25				30							
Ile	Pro	Ser	Leu	Val	Asn	Asn	Glu	Leu	Val	Leu	Arg	Asp	Lys	Ala	Thr
35				40				45							
Lys	Val	Val	Asp	Tyr	Phe	Lys	His	Val	Ser	Leu	Val	Glu	Thr	Glu	Gly
50				55				60							
Val	Phe	Thr	Leu	Leu	Asp	Asp	Lys	Val	Met	Met	Pro	Gln	Asp	Asp	Leu
65				70				75				80			
Val	Ile	Ser	Glu	Ile	Asp	Phe	Asn	Asn	Asn	Ser	Ile	Val	Leu	Gly	Lys
85				90				95							
Cys	Glu	Ile	Trp	Arg	Met	Glu	Gly	Gly	Ser	Gly	His	Thr	Val	Thr	Asp
100				105				110							
Asp	Ile	Asp	His	Phe	Phe	Ser	Ala	Pro	Ser	Ile	Thr	Tyr	Arg	Glu	Pro
115				120				125							
His	Leu	Ser	Ile	Tyr	Asp	Val	Leu	Glu	Val	Gln	Lys	Glu	Glu	Leu	Asp
130				135				140							
Leu	Ser	Lys	Asp	Leu	Met	Val	Leu	Pro	Asn	Ala	Pro	Asn	Arg	Val	Phe
145				150				155				160			
Ala	Trp	Glu	Thr	Gly	Trp	Thr	Pro	Gly	Leu	Arg	Ser	Leu	Glu	Asn	Asp
165				170				175							
Gly	Thr	Lys	Leu	Leu	Asp	Arg	Ile	Arg	Asp	Asn	Tyr	Glu	Gly	Glu	Phe
180				185				190							
Tyr	Trp	Arg	Tyr	Phe	Ala	Phe	Ile	Ala	Asp	Ala	Leu	Ile	Thr	Thr	Leu
195				200				205							
Lys	Pro	Arg	Tyr	Glu	Asp	Thr	Asn	Ile	Arg	Ile	Asn	Leu	Asp	Ser	Asn
210				215				220							

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

<210> 4
 <211> 410
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> TD1 Immunogen (aa 1072-1452) + peptide linker + 10x His Tag
 <220>
 <221> MISC_FEATURE
 <222> (383) .. (398)
 <223> Peptide linker
 <220>
 <221> MISC_FEATURE
 <222> (399) .. (408)
 <223> His tag
 <400> 4

Ser Leu Thr Thr Ala Thr Thr Ala Ile Ile Thr Ser Ser Leu Gly Ile			
1	5	10	15
Ala Ser Gly Phe Ser Ile Leu Leu Val Pro Leu Ala Gly Ile Ser Ala			
	20	25	30

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

340	345	350
Ser Lys Ser Tyr Lys Leu Leu Ile Ser Gly Glu Leu Lys Ile Leu Met		
355	360	365
Leu Asn Ser Asn His Ile Gln Gln Lys Ile Asp Tyr Ile Gly Glu Phe		
370	375	380
Ser Ser Gly His Ile Asp Asp Asp Asp Ser His Met Leu Glu His His		
385	390	395
His His His His His His His His Gly Met		400
405	410	
<210> 5		
<211> 421		
<212> PRT		
<213> Clostridium difficile		
<400> 5		
Thr Ser Asp Pro Leu Leu Arg Gln Glu Ile Glu Ala Lys Ile Gly Ile		
1	5	10
Met Ala Val Asn Leu Thr Thr Ala Thr Thr Ala Ile Ile Thr Ser Ser		
20	25	30
Leu Gly Ile Ala Ser Gly Phe Ser Ile Leu Leu Val Pro Leu Ala Gly		
35	40	45
Ile Ser Ala Gly Ile Pro Ser Leu Val Asn Asn Glu Leu Val Leu Arg		
50	55	60
Asp Lys Ala Thr Lys Val Val Asp Tyr Phe Lys His Val Ser Leu Val		
65	70	75
Glu Thr Glu Gly Val Phe Thr Leu Leu Asp Asp Lys Val Met Met Pro		
85	90	95
Gln Asp Asp Leu Val Ile Ser Glu Ile Asp Phe Asn Asn Asn Ser Ile		
100	105	110
Val Leu Gly Lys Cys Glu Ile Trp Arg Met Glu Gly Gly Ser Gly His		
115	120	125
Thr Val Thr Asp Asp Ile Asp His Phe Phe Ser Ala Pro Ser Ile Thr		
130	135	140
Tyr Arg Glu Pro His Leu Ser Ile Tyr Asp Val Leu Glu Val Gln Lys		
145	150	155
Glu Glu Leu Asp Leu Ser Lys Asp Leu Met Val Leu Pro Asn Ala Pro		
165	170	175
Asn Arg Val Phe Ala Trp Glu Thr Gly Trp Thr Pro Gly Leu Arg Ser		
180	185	190
Leu Glu Asn Asp Gly Thr Lys Leu Leu Asp Arg Ile Arg Asp Asn Tyr		

195	200	205
Glu Gly Glu Phe Tyr Trp Arg Tyr Phe Ala Phe Ile Ala Asp Ala Leu		
210	215	220
Ile Thr Thr Leu Lys Pro Arg Tyr Glu Asp Thr Asn Ile Arg Ile Asn		
225	230	235
Leu Asp Ser Asn Thr Arg Ser Phe Ile Val Pro Ile Ile Thr Thr Glu		
245	250	255
Tyr Ile Arg Glu Lys Leu Ser Tyr Ser Phe Tyr Gly Ser Gly Gly Thr		
260	265	270
Tyr Ala Leu Ser Leu Ser Gln Tyr Asn Met Gly Ile Asn Ile Glu Leu		
275	280	285
Ser Glu Ser Asp Val Trp Ile Ile Asp Val Asp Asn Val Val Arg Asp		
290	295	300
Val Thr Ile Glu Ser Asp Lys Ile Lys Lys Gly Asp Leu Ile Glu Gly		
305	310	315
Ile Leu Ser Thr Leu Ser Ile Glu Glu Asn Lys Ile Ile Leu Asn Ser		
325	330	335
His Glu Ile Asn Phe Ser Gly Glu Val Asn Gly Ser Asn Gly Phe Val		
340	345	350
Ser Leu Thr Phe Ser Ile Leu Glu Gly Ile Asn Ala Ile Ile Glu Val		
355	360	365
Asp Leu Leu Ser Lys Ser Tyr Lys Leu Leu Ile Ser Gly Glu Leu Lys		
370	375	380
Ile Leu Met Leu Asn Ser Asn His Ile Gln Gln Lys Ile Asp Tyr Ile		
385	390	395
Gly Phe Asn Ser Glu Leu Gln Lys Asn Ile Pro Tyr Ser Phe Val Asp		
405	410	415
Ser Glu Gly Lys Glu		
420		

<210> 6

<211> 481

<212> PRT

<213> Clostridium difficile

<400> 6

Leu Leu Pro Thr Leu Ser Glu Gly Leu Pro Ile Ile Ala Thr Ile Ile
1 5 10 15
Asp Gly Val Ser Leu Gly Ala Ala Ile Lys Glu Leu Ser Glu Thr Ser
20 25 30
Asp Pro Leu Leu Arg Gln Glu Ile Glu Ala Lys Ile Gly Ile Met Ala

35	40	45
Val Asn Leu Thr Thr Ala Thr Thr Ala Ile Ile Thr Ser Ser Leu Gly		
50	55	60
Ile Ala Ser Gly Phe Ser Ile Leu Leu Val Pro Leu Ala Gly Ile Ser		
65	70	75
Ala Gly Ile Pro Ser Leu Val Asn Asn Glu Leu Val Leu Arg Asp Lys		
85	90	95
Ala Thr Lys Val Val Asp Tyr Phe Lys His Val Ser Leu Val Glu Thr		
100	105	110
Glu Gly Val Phe Thr Leu Leu Asp Asp Lys Val Met Met Pro Gln Asp		
115	120	125
Asp Leu Val Ile Ser Glu Ile Asp Phe Asn Asn Asn Ser Ile Val Leu		
130	135	140
Gly Lys Cys Glu Ile Trp Arg Met Glu Gly Gly Ser Gly His Thr Val		
145	150	155
Thr Asp Asp Ile Asp His Phe Phe Ser Ala Pro Ser Ile Thr Tyr Arg		
165	170	175
Glu Pro His Leu Ser Ile Tyr Asp Val Leu Glu Val Gln Lys Glu Glu		
180	185	190
Leu Asp Leu Ser Lys Asp Leu Met Val Leu Pro Asn Ala Pro Asn Arg		
195	200	205
Val Phe Ala Trp Glu Thr Gly Trp Thr Pro Gly Leu Arg Ser Leu Glu		
210	215	220
Asn Asp Gly Thr Lys Leu Leu Asp Arg Ile Arg Asp Asn Tyr Glu Gly		
225	230	235
Glu Phe Tyr Trp Arg Tyr Phe Ala Phe Ile Ala Asp Ala Leu Ile Thr		
245	250	255
Thr Leu Lys Pro Arg Tyr Glu Asp Thr Asn Ile Arg Ile Asn Leu Asp		
260	265	270
Ser Asn Thr Arg Ser Phe Ile Val Pro Ile Ile Thr Thr Glu Tyr Ile		
275	280	285
Arg Glu Lys Leu Ser Tyr Ser Phe Tyr Gly Ser Gly Gly Thr Tyr Ala		
290	295	300
Leu Ser Leu Ser Gln Tyr Asn Met Gly Ile Asn Ile Glu Leu Ser Glu		
305	310	315
Ser Asp Val Trp Ile Ile Asp Val Asp Asn Val Val Arg Asp Val Thr		
325	330	335
Ile Glu Ser Asp Lys Ile Lys Lys Gly Asp Leu Ile Glu Gly Ile Leu		
340	345	350

Ser Thr Leu Ser Ile Glu Glu Asn Lys Ile Ile Leu Asn Ser His Glu
 355 360 365
 Ile Asn Phe Ser Gly Glu Val Asn Gly Ser Asn Gly Phe Val Ser Leu
 370 375 380
 Thr Phe Ser Ile Leu Glu Gly Ile Asn Ala Ile Ile Glu Val Asp Leu
 385 390 395 400
 Leu Ser Lys Ser Tyr Lys Leu Leu Ile Ser Gly Glu Leu Lys Ile Leu
 405 410 415
 Met Leu Asn Ser Asn His Ile Gln Gln Lys Ile Asp Tyr Ile Gly Phe
 420 425 430
 Asn Ser Glu Leu Gln Lys Asn Ile Pro Tyr Ser Phe Val Asp Ser Glu
 435 440 445
 Gly Lys Glu Asn Gly Phe Ile Asn Gly Ser Thr Lys Glu Gly Leu Phe
 450 455 460
 Val Ser Glu Leu Pro Asp Val Val Leu Ile Ser Lys Val Tyr Met Asp
 465 470 475 480
 Asp
 <210> 7
 <211> 533
 <212> PRT
 <213> Clostridium difficile
 <400> 7
 Met Ser Leu Val Asn Arg Lys Gln Leu Glu Lys Met Ala Asn Val Arg
 1 5 10 15
 Phe Arg Val Gln Glu Asp Glu Tyr Val Ala Ile Leu Asp Ala Leu Glu
 20 25 30
 Glu Tyr His Asn Met Ser Glu Asn Thr Val Val Glu Lys Tyr Leu Lys
 35 40 45
 Leu Lys Asp Ile Asn Ser Leu Thr Asp Thr Tyr Ile Asp Thr Tyr Lys
 50 55 60
 Lys Ser Gly Arg Asn Lys Ala Leu Lys Lys Phe Lys Glu Tyr Leu Val
 65 70 75 80
 Ile Glu Ile Leu Glu Leu Lys Asn Ser Asn Leu Thr Pro Val Glu Lys
 85 90 95
 Asn Leu His Phe Ile Trp Ile Gly Gly Gln Ile Asn Asp Thr Ala Ile
 100 105 110
 Asn Tyr Ile Asn Gln Trp Lys Asp Val Asn Ser Asp Tyr Asn Val Asn
 115 120 125
 Val Phe Tyr Asp Ser Asn Ala Phe Leu Ile Asn Thr Leu Lys Lys Thr

130	135	140
Ile Ile Glu Ser Ala Ser Asn Asp Thr Leu Glu Ser Phe Arg Glu Asn		
145	150	155
Leu Asn Asp Pro Glu Phe Asn His Thr Ala Phe Phe Arg Lys Arg Met		160
	165	170
Gln Ile Ile Tyr Asp Lys Gln Gln Asn Phe Ile Asn Tyr Tyr Lys Ala		175
	180	185
Gln Lys Glu Glu Asn Pro Asp Leu Ile Ile Asp Asp Ile Val Lys Thr		190
	195	200
Tyr Leu Ser Asn Glu Tyr Ser Lys Asp Ile Asp Glu Leu Asn Ala Tyr		205
	210	220
Ile Glu Glu Ser Leu Asn Lys Val Thr Glu Asn Ser Gly Asn Asp Val		
225	230	235
Arg Asn Phe Glu Glu Phe Lys Thr Gly Glu Val Phe Asn Leu Tyr Glu		240
	245	250
Gln Glu Leu Val Glu Arg Trp Asn Leu Ala Gly Ala Ser Asp Ile Leu		255
	260	265
Arg Val Ala Ile Leu Lys Asn Ile Gly Gly Val Tyr Leu Asp Val Asp		270
	275	280
Met Leu Pro Gly Ile His Pro Asp Leu Phe Lys Asp Ile Asn Lys Pro		285
	290	295
Asp Ser Val Lys Thr Ala Val Asp Trp Glu Glu Met Gln Leu Glu Ala		300
305	310	315
Ile Met Lys His Lys Glu Tyr Ile Pro Glu Tyr Thr Ser Lys His Phe		320
	325	330
Asp Thr Leu Asp Glu Glu Val Gln Ser Ser Phe Glu Ser Val Leu Ala		335
	340	345
Ser Lys Ser Asp Lys Ser Glu Ile Phe Leu Pro Leu Gly Asp Ile Glu		350
	355	360
Val Ser Pro Leu Glu Val Lys Ile Ala Phe Ala Lys Gly Ser Ile Ile		365
	370	375
Asn Gln Ala Leu Ile Ser Ala Lys Asp Ser Tyr Cys Ser Asp Leu Leu		380
385	390	395
Ile Lys Gln Ile Gln Asn Arg Tyr Lys Ile Leu Asn Asp Thr Leu Gly		400
	405	410
Pro Ile Ile Ser Gln Gly Asn Asp Phe Asn Thr Thr Met Asn Asn Phe		415
	420	425
Gly Glu Ser Leu Gly Ala Ile Ala Asn Glu Glu Asn Ile Ser Phe Ile		430
	435	440
		445

Ala Lys Ile Gly Ser Tyr Leu Arg Val Gly Phe Tyr Pro Glu Ala Asn		
450	455	460
Thr Thr Ile Thr Leu Ser Gly Pro Thr Ile Tyr Ala Gly Ala Tyr Lys		
465	470	475 480
Asp Leu Leu Thr Phe Lys Glu Met Ser Ile Asp Thr Ser Ile Leu Ser		
	485	490 495
Ser Glu Leu Arg Asn Phe Glu Phe Pro Lys Val Asn Ile Ser Gln Ala		
	500	505 510
Thr Glu Gln Glu Lys Asn Ser Leu Trp Gln Phe Asn Glu Glu Arg Ala		
	515	520 525
Lys Ile Gln Phe Glu		
530		
<210> 8		
<211> 593		
<212> PRT		
<213> Clostridium difficile		
<400> 8		
Met Ser Leu Val Asn Arg Lys Gln Leu Glu Lys Met Ala Asn Val Arg		
1	5	10 15
Phe Arg Val Gln Glu Asp Glu Tyr Val Ala Ile Leu Asp Ala Leu Glu		
	20	25 30
Glu Tyr His Asn Met Ser Glu Asn Thr Val Val Glu Lys Tyr Leu Lys		
	35	40 45
Leu Lys Asp Ile Asn Ser Leu Thr Asp Thr Tyr Ile Asp Thr Tyr Lys		
	50	55 60
Lys Ser Gly Arg Asn Lys Ala Leu Lys Lys Phe Lys Glu Tyr Leu Val		
65	70	75 80
Ile Glu Ile Leu Glu Leu Lys Asn Ser Asn Leu Thr Pro Val Glu Lys		
	85	90 95
Asn Leu His Phe Ile Trp Ile Gly Gly Gln Ile Asn Asp Thr Ala Ile		
	100	105 110
Asn Tyr Ile Asn Gln Trp Lys Asp Val Asn Ser Asp Tyr Asn Val Asn		
	115	120 125
Val Phe Tyr Asp Ser Asn Ala Phe Leu Ile Asn Thr Leu Lys Lys Thr		
	130	135 140
Ile Ile Glu Ser Ala Ser Asn Asp Thr Leu Glu Ser Phe Arg Glu Asn		
145	150	155 160
Leu Asn Asp Pro Glu Phe Asn His Thr Ala Phe Phe Arg Lys Arg Met		
	165	170 175

Gln Ile Ile Tyr Asp Lys Gln Gln Asn Phe Ile Asn Tyr Tyr Lys Ala		
180	185	190
Gln Lys Glu Glu Asn Pro Asp Leu Ile Ile Asp Asp Ile Val Lys Thr		
195	200	205
Tyr Leu Ser Asn Glu Tyr Ser Lys Asp Ile Asp Glu Leu Asn Ala Tyr		
210	215	220
Ile Glu Glu Ser Leu Asn Lys Val Thr Glu Asn Ser Gly Asn Asp Val		
225	230	235
Arg Asn Phe Glu Glu Phe Lys Thr Gly Glu Val Phe Asn Leu Tyr Glu		
245	250	255
Gln Glu Leu Val Glu Arg Trp Asn Leu Ala Gly Ala Ser Asp Ile Leu		
260	265	270
Arg Val Ala Ile Leu Lys Asn Ile Gly Gly Val Tyr Leu Asp Val Asp		
275	280	285
Met Leu Pro Gly Ile His Pro Asp Leu Phe Lys Asp Ile Asn Lys Pro		
290	295	300
Asp Ser Val Lys Thr Ala Val Asp Trp Glu Glu Met Gln Leu Glu Ala		
305	310	315
Ile Met Lys His Lys Glu Tyr Ile Pro Glu Tyr Thr Ser Lys His Phe		
325	330	335
Asp Thr Leu Asp Glu Glu Val Gln Ser Ser Phe Glu Ser Val Leu Ala		
340	345	350
Ser Lys Ser Asp Lys Ser Glu Ile Phe Leu Pro Leu Gly Asp Ile Glu		
355	360	365
Val Ser Pro Leu Glu Val Lys Ile Ala Phe Ala Lys Gly Ser Ile Ile		
370	375	380
Asn Gln Ala Leu Ile Ser Ala Lys Asp Ser Tyr Cys Ser Asp Leu Leu		
385	390	395
Ile Lys Gln Ile Gln Asn Arg Tyr Lys Ile Leu Asn Asp Thr Leu Gly		
405	410	415
Pro Ile Ile Ser Gln Gly Asn Asp Phe Asn Thr Thr Met Asn Asn Phe		
420	425	430
Gly Glu Ser Leu Gly Ala Ile Ala Asn Glu Glu Asn Ile Ser Phe Ile		
435	440	445
Ala Lys Ile Gly Ser Tyr Leu Arg Val Gly Phe Tyr Pro Glu Ala Asn		
450	455	460
Thr Thr Ile Thr Leu Ser Gly Pro Thr Ile Tyr Ala Gly Ala Tyr Lys		
465	470	475
Asp Leu Leu Thr Phe Lys Glu Met Ser Ile Asp Thr Ser Ile Leu Ser		

	485		490		495
Ser Glu Leu Arg Asn Phe Glu Phe Pro Lys Val Asn Ile Ser Gln Ala					
	500		505		510
Thr Glu Gln Glu Lys Asn Ser Leu Trp Gln Phe Asn Glu Glu Arg Ala					
	515		520		525
Lys Ile Gln Phe Glu Glu Tyr Lys Lys Asn Tyr Phe Glu Gly Ala Leu					
	530		535		540
Gly Glu Asp Asp Asn Leu Asp Phe Ser Gln Asn Thr Val Thr Asp Lys					
545		550		555	560
Glu Tyr Leu Leu Glu Lys Ile Ser Ser Ser Thr Lys Ser Ser Glu Arg					
	565		570		575
Gly Tyr Val His Tyr Ile Val Gln Leu Gln Gly Asp Lys Ile Ser Tyr					
	580		585		590
Glu					
<210>	9				
<211>	573				
<212>	PRT				
<213>	Clostridium difficile				
<400>	9				
Met Ser Leu Val Asn Arg Lys Gln Leu Glu Lys Met Ala Asn Val Arg					
1	5		10		15
Phe Arg Val Gln Glu Asp Glu Tyr Val Ala Ile Leu Asp Ala Leu Glu					
	20		25		30
Glu Tyr His Asn Met Ser Glu Asn Thr Val Val Glu Lys Tyr Leu Lys					
	35		40		45
Leu Lys Asp Ile Asn Ser Leu Thr Asp Thr Tyr Ile Asp Thr Tyr Lys					
	50		55		60
Lys Ser Gly Arg Asn Lys Ala Leu Lys Lys Phe Lys Glu Tyr Leu Val					
65		70		75	80
Ile Glu Ile Leu Glu Leu Lys Asn Ser Asn Leu Thr Pro Val Glu Lys					
	85		90		95
Asn Leu His Phe Ile Trp Ile Gly Gly Gln Ile Asn Asp Thr Ala Ile					
	100		105		110
Asn Tyr Ile Asn Gln Trp Lys Asp Val Asn Ser Asp Tyr Asn Val Asn					
	115		120		125
Val Phe Tyr Asp Ser Asn Ala Phe Leu Ile Asn Thr Leu Lys Lys Thr					
	130		135		140
Ile Ile Glu Ser Ala Ser Asn Asp Thr Leu Glu Ser Phe Arg Glu Asn					
145		150		155	160

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

465	470	475	480
Asp Leu Leu Thr Phe	Lys Glu Met Ser Ile	Asp Thr Ser Ile	Leu Ser
	485	490	495
Ser Glu Leu Arg Asn	Phe Glu Phe Pro Lys	Val Asn Ile Ser	Gln Ala
	500	505	510
Thr Glu Gln Glu Lys	Asn Ser Leu Trp Gln	Phe Asn Glu Glu	Arg Ala
	515	520	525
Lys Ile Gln Phe Glu	Glu Tyr Lys Lys Asn	Tyr Phe Glu Gly	Ala Leu
	530	535	540
Gly Glu Asp Asp Asn	Leu Asp Phe Ser Gln	Asn Thr Val Thr	Asp Lys
545	550	555	560
Glu Tyr Leu Leu Glu	Lys Ile Ser Ser Ser	Thr Lys Ser	
	565	570	
<210>	10		
<211>	254		
<212>	PRT		
<213>	Clostridium difficile		
<400>	10		
Pro Ser Leu Val Asn	Asn Glu Leu Val Leu	Arg Asp Lys Ala	Thr Lys
1	5	10	15
Val Val Asp Tyr Phe	Lys His Val Ser Leu	Val Glu Thr Glu	Gly Val
	20	25	30
Phe Thr Leu Leu Asp	Asp Lys Val Met Met	Pro Gln Asp Asp	Leu Val
	35	40	45
Ile Ser Glu Ile Asp	Phe Asn Asn Asn Ser	Ile Val Leu Gly	Lys Cys
	50	55	60
Glu Ile Trp Arg Met	Glu Gly Gly Ser Gly	His Thr Val Thr	Asp Asp
65	70	75	80
Ile Asp His Phe Phe	Ser Ala Pro Ser Ile	Thr Tyr Arg Glu	Pro His
	85	90	95
Leu Ser Ile Tyr Asp	Val Leu Glu Val Gln	Lys Glu Glu Leu	Asp Leu
	100	105	110
Ser Lys Asp Leu Met	Val Leu Pro Asn Ala	Pro Asn Arg Val	Phe Ala
	115	120	125
Trp Glu Thr Gly Trp	Thr Pro Gly Leu Arg	Ser Leu Glu Asn	Asp Gly
	130	135	140
Thr Lys Leu Leu Asp	Arg Ile Arg Asp Asn	Tyr Glu Gly Glu	Phe Tyr
145	150	155	160
Trp Arg Tyr Phe Ala	Phe Ile Ala Asp Ala	Leu Ile Thr Thr	Leu Lys

	165		170		175
Pro Arg Tyr Glu Asp Thr Asn Ile Arg Ile Asn Leu Asp Ser Asn Thr					
	180		185		190
Arg Ser Phe Ile Val Pro Ile Ile Thr Thr Glu Tyr Ile Arg Glu Lys					
	195		200		205
Leu Ser Tyr Ser Phe Tyr Gly Ser Gly Gly Thr Tyr Ala Leu Ser Leu					
	210		215		220
Ser Gln Tyr Asn Met Gly Ile Asn Ile Glu Leu Ser Glu Ser Asp Val					
225		230		235	240
Trp Ile Ile Asp Val Asp Asn Val Val Arg Asp Val Thr Ile					
	245		250		
<210> 11					
<211> 41					
<212> PRT					
<213> Clostridium difficile					
<400> 11					
Glu Tyr Val Ala Ile Leu Asp Ala Leu Glu Glu Tyr His Asn Met Ser					
1	5		10		15
Glu Asn Thr Val Val Glu Lys Tyr Leu Lys Leu Lys Asp Ile Asn Ser					
	20		25		30
Leu Thr Asp Thr Tyr Ile Asp Thr Tyr					
	35		40		
<210> 12					
<211> 392					
<212> PRT					
<213> Clostridium difficile					
<400> 12					
Glu Ser Ala Ser Asn Asp Thr Leu Glu Ser Phe Arg Glu Asn Leu Asn					
1	5		10		15
Asp Pro Glu Phe Asn His Thr Ala Phe Phe Arg Lys Arg Met Gln Ile					
	20		25		30
Ile Tyr Asp Lys Gln Gln Asn Phe Ile Asn Tyr Tyr Lys Ala Gln Lys					
	35		40		45
Glu Glu Asn Pro Asp Leu Ile Ile Asp Asp Ile Val Lys Thr Tyr Leu					
	50		55		60
Ser Asn Glu Tyr Ser Lys Asp Ile Asp Glu Leu Asn Ala Tyr Ile Glu					
65		70		75	80
Glu Ser Leu Asn Lys Val Thr Glu Asn Ser Gly Asn Asp Val Arg Asn					
	85		90		95

Phe	Glu	Glu	Phe	Lys	Thr	Gly	Glu	Val	Phe	Asn	Leu	Tyr	Glu	Gln	Glu
							100								
Leu	Val	Glu	Arg	Trp	Asn	Leu	Ala	Gly	Ala	Ser	Asp	Ile	Leu	Arg	Val
Ala	Ile	Leu	Lys	Asn	Ile	Gly	Gly	Val	Tyr	Leu	Asp	Val	Asp	Met	Leu
Pro	Gly	Ile	His	Pro	Asp	Leu	Phe	Lys	Asp	Ile	Asn	Lys	Pro	Asp	Ser
Val	Lys	Thr	Ala	Val	Asp	Trp	Glu	Glu	Met	Gln	Leu	Glu	Ala	Ile	Met
Lys	His	Lys	Glu	Tyr	Ile	Pro	Glu	Tyr	Thr	Ser	Lys	His	Phe	Asp	Thr
Leu	Asp	Glu	Glu	Val	Gln	Ser	Ser	Phe	Glu	Ser	Val	Leu	Ala	Ser	Lys
Ser	Asp	Lys	Ser	Glu	Ile	Phe	Leu	Pro	Leu	Gly	Asp	Ile	Glu	Val	Ser
Pro	Leu	Glu	Val	Lys	Ile	Ala	Phe	Ala	Lys	Gly	Ser	Ile	Ile	Asn	Gln
Ala	Leu	Ile	Ser	Ala	Lys	Asp	Ser	Tyr	Cys	Ser	Asp	Leu	Leu	Ile	Lys
Gln	Ile	Gln	Asn	Arg	Tyr	Lys	Ile	Leu	Asn	Asp	Thr	Leu	Gly	Pro	Ile
Ile	Ser	Gln	Gly	Asn	Asp	Phe	Asn	Thr	Thr	Met	Asn	Asn	Phe	Gly	Glu
Ser	Leu	Gly	Ala	Ile	Ala	Asn	Glu	Glu	Asn	Ile	Ser	Phe	Ile	Ala	Lys
Ile	Gly	Ser	Tyr	Leu	Arg	Val	Gly	Phe	Tyr	Pro	Glu	Ala	Asn	Thr	Thr
Ile	Thr	Leu	Ser	Gly	Pro	Thr	Ile	Tyr	Ala	Gly	Ala	Tyr	Lys	Asp	Leu
Leu	Thr	Phe	Lys	Glu	Met	Ser	Ile	Asp	Thr	Ser	Ile	Leu	Ser	Ser	Glu
Leu	Arg	Asn	Phe	Glu	Phe	Pro	Lys	Val	Asn	Ile	Ser	Gln	Ala	Thr	Glu
Gln	Glu	Lys	Asn	Ser	Leu	Trp	Gln	Phe	Asn	Glu	Glu	Arg	Ala	Lys	Ile
Gln	Phe	Glu	Glu	Tyr	Lys	Lys	Asn								

<210> 13

<211> 52
 <212> PRT
 <213> Clostridium difficile
 <400> 13
 Pro Val Ser Glu Ile Ile Leu Ser Phe Thr Pro Ser Tyr Tyr Glu Asp
 1 5 10 15
 Gly Leu Ile Gly Tyr Asp Leu Gly Leu Val Ser Leu Tyr Asn Glu Lys
 20 25 30
 Phe Tyr Ile Asn Asn Phe Gly Met Met Val Ser Gly Leu Ile Tyr Ile
 35 40 45
 Asn Asp Ser Leu
 50
 <210> 14
 <211> 176
 <212> PRT
 <213> Clostridium difficile
 <400> 14
 Gly Leu Asp Val Asp Ser Leu Ser Ser Glu Ile Glu Thr Ala Ile Gly
 1 5 10 15
 Leu Ala Lys Glu Asp Ile Ser Pro Lys Ser Ile Glu Ile Asn Leu Leu
 20 25 30
 Gly Cys Asn Met Phe Ser Tyr Ser Val Asn Val Glu Glu Thr Tyr Pro
 35 40 45
 Gly Lys Leu Leu Leu Arg Val Lys Asp Lys Val Ser Glu Leu Met Pro
 50 55 60
 Ser Met Ser Gln Asp Ser Ile Ile Val Ser Ala Asn Gln Tyr Glu Val
 65 70 75 80
 Arg Ile Asn Ser Glu Gly Arg Arg Glu Leu Leu Asp His Ser Gly Glu
 85 90 95
 Trp Ile Asn Lys Glu Glu Ser Ile Ile Lys Asp Ile Ser Ser Lys Glu
 100 105 110
 Tyr Ile Ser Phe Asn Pro Lys Glu Asn Lys Ile Ile Val Lys Ser Lys
 115 120 125
 Asn Leu Pro Glu Leu Ser Thr Leu Leu Gln Glu Ile Arg Asn Asn Ser
 130 135 140
 Asn Ser Ser Asp Ile Glu Leu Glu Glu Lys Val Met Leu Ala Glu Cys
 145 150 155 160
 Glu Ile Asn Val Ile Ser Asn Ile Glu Thr Gln Val Val Glu Glu Arg
 165 170 175

<210> 15
 <211> 100
 <212> PRT
 <213> Clostridium difficile
 <400> 15
 Gln Tyr Glu Val Arg Ile Asn Ser Glu Gly Arg Arg Glu Leu Leu Asp
 1 5 10 15
 His Ser Gly Glu Trp Ile Asn Lys Glu Glu Ser Ile Ile Lys Asp Ile
 20 25 30
 Ser Ser Lys Glu Tyr Ile Ser Phe Asn Pro Lys Glu Asn Lys Ile Ile
 35 40 45
 Val Lys Ser Lys Asn Leu Pro Glu Leu Ser Thr Leu Leu Gln Glu Ile
 50 55 60
 Arg Asn Asn Ser Asn Ser Ser Asp Ile Glu Leu Glu Glu Lys Val Met
 65 70 75 80
 Leu Ala Glu Cys Glu Ile Asn Val Ile Ser Asn Ile Glu Thr Gln Val
 85 90 95
 Val Glu Glu Arg
 100

<210> 16
 <211> 541
 <212> PRT
 <213> Clostridium difficile
 <400> 16
 Met Ser Leu Ile Ser Lys Glu Glu Leu Ile Lys Leu Ala Tyr Ser Ile
 1 5 10 15
 Arg Pro Arg Glu Asn Glu Tyr Lys Thr Ile Leu Thr Asn Leu Asp Glu
 20 25 30
 Tyr Asn Lys Leu Thr Thr Asn Asn Asn Glu Asn Lys Tyr Leu Gln Leu
 35 40 45
 Lys Lys Leu Asn Glu Ser Ile Asp Val Phe Met Asn Lys Tyr Lys Thr
 50 55 60
 Ser Ser Arg Asn Arg Ala Leu Ser Asn Leu Lys Lys Asp Ile Leu Lys
 65 70 75 80
 Glu Val Ile Leu Ile Lys Asn Ser Asn Thr Ser Pro Val Glu Lys Asn
 85 90 95
 Leu His Phe Val Trp Ile Gly Gly Glu Val Ser Asp Ile Ala Leu Glu
 100 105 110
 Tyr Ile Lys Gln Trp Ala Asp Ile Asn Ala Glu Tyr Asn Ile Lys Leu

115	120	125
Trp Tyr Asp Ser Glu Ala Phe	Leu Val Asn Thr	Leu Lys Lys Ala Ile
130	135	140
Val Glu Ser Ser Thr Thr Glu Ala	Leu Gln Leu Leu Glu Glu Glu Ile	
145	150	155
Gln Asn Pro Gln Phe Asp Asn Met	Lys Phe Tyr Lys Lys Arg Met Glu	
165	170	175
Phe Ile Tyr Asp Arg Gln Lys Arg	Phe Ile Asn Tyr Tyr Lys Ser Gln	
180	185	190
Ile Asn Lys Pro Thr Val Pro Thr	Ile Asp Asp Ile Ile Lys Ser His	
195	200	205
Leu Val Ser Glu Tyr Asn Arg Asp	Glu Thr Val Leu Glu Ser Tyr Arg	
210	215	220
Thr Asn Ser Leu Arg Lys Ile Asn	Ser Asn His Gly Ile Asp Ile Ile	
225	230	235
Ser Arg Pro Ser Ser Ile Gly Leu	Asp Arg Trp Glu Met Ile Lys Leu	
245	250	255
Glu Ala Ile Met Lys Tyr Lys Lys	Tyr Ile Asn Asn Tyr Thr Ser Glu	
260	265	270
Asn Phe Asp Lys Leu Asp Gln Gln	Leu Lys Asp Asn Phe Lys Leu Ile	
275	280	285
Ile Glu Ser Lys Ser Glu Lys Ser	Glu Ile Phe Ser Lys Leu Glu Asn	
290	295	300
Leu Asn Val Ser Asp Leu Glu Ile	Lys Ile Ala Phe Ala Leu Gly Ser	
305	310	315
Val Ile Asn Gln Ala Leu Ile Ser	Lys Gln Gly Ser Tyr Leu Thr Asn	
325	330	335
Leu Val Ile Glu Gln Val Lys Asn	Arg Tyr Gln Phe Leu Asn Gln His	
340	345	350
Leu Asn Pro Ala Ile Glu Ser Asp	Asn Asn Phe Thr Asp Thr Thr Lys	
355	360	365
Ile Phe His Asp Ser Leu Phe Asn	Ser Ala Thr Ala Glu Asn Ser Met	
370	375	380
Phe Leu Thr Lys Ile Ala Pro Tyr	Leu Gln Val Gly Phe Met Pro Glu	
385	390	395
Ala Arg Ser Thr Ile Ser Leu Ser	Gly Pro Gly Ala Tyr Ala Ser Ala	
405	410	415
Tyr Tyr Arg Ala Asn Ser Leu Phe	Thr Glu Gln Glu Leu Leu Asn Ile	
420	425	430

Tyr Ser Gln Glu Leu Leu Asn Arg Gly Asn Leu Ala Ala Ala Ser Asp			
435	440	445	
Ile Val Arg Leu Leu Ala Leu Lys Asn Phe Gly Gly Val Tyr Leu Asp			
450	455	460	
Val Asp Met Leu Pro Gly Ile His Ser Asp Leu Phe Lys Thr Asp Phe			
465	470	475	480
Ile Asn Leu Gln Glu Asn Thr Ile Glu Lys Thr Leu Lys Ala Ser Asp			
	485	490	495
Leu Ile Glu Phe Lys Phe Pro Glu Asn Asn Leu Ser Gln Leu Thr Glu			
	500	505	510
Gln Glu Ile Asn Ser Leu Trp Ser Phe Asp Gln Ala Ser Ala Lys Tyr			
	515	520	525
Gln Phe Glu Lys Tyr Val Arg Asp Tyr Thr Gly Gly Ser			
530	535	540	
<210> 17			
<211> 380			
<212> PRT			
<213> Clostridium difficile			
<400> 17			
Met Ser Leu Ser Ile Ala Ala Thr Val Ala Ser Ile Val Gly Ile Gly			
1	5	10	15
Ala Glu Val Thr Ile Phe Leu Leu Pro Ile Ala Gly Ile Ser Ala Gly			
	20	25	30
Ile Pro Ser Leu Val Asn Asn Glu Leu Ile Leu His Asp Lys Ala Thr			
	35	40	45
Ser Val Val Asn Tyr Phe Asn His Leu Ser Glu Ser Lys Lys Tyr Gly			
50	55	60	
Pro Leu Lys Thr Glu Asp Asp Lys Ile Leu Val Pro Ile Asp Asp Leu			
65	70	75	80
Val Ile Ser Glu Ile Asp Phe Asn Asn Asn Ser Ile Lys Leu Gly Thr			
	85	90	95
Cys Asn Ile Leu Ala Met Glu Gly Gly Ser Gly His Thr Val Thr Gly			
	100	105	110
Asn Ile Asp His Phe Phe Ser Ser Pro Ser Ile Ser Ser His Ile Pro			
	115	120	125
Ser Leu Ser Ile Tyr Ser Ala Ile Gly Ile Glu Thr Glu Asn Leu Asp			
130	135	140	
Phe Ser Lys Lys Ile Met Met Leu Pro Asn Ala Pro Ser Arg Val Phe			
145	150	155	160

Ser Ile Asp Val Phe Met Asn Lys Tyr
35 40

<210> 19
 <211> 390
 <212> PRT
 <213> Clostridium difficile
 <400> 19
 Ser Ser Thr Thr Glu Ala Leu Gln Leu Leu Glu Glu Glu Ile Gln Asn
 1 5 10 15
 Pro Gln Phe Asp Asn Met Lys Phe Tyr Lys Lys Arg Met Glu Phe Ile
 20 25 30
 Tyr Asp Arg Gln Lys Arg Phe Ile Asn Tyr Tyr Lys Ser Gln Ile Asn
 35 40 45
 Lys Pro Thr Val Pro Thr Ile Asp Asp Ile Ile Lys Ser His Leu Val
 50 55 60
 Ser Glu Tyr Asn Arg Asp Glu Thr Val Leu Glu Ser Tyr Arg Thr Asn
 65 70 75 80
 Ser Leu Arg Lys Ile Asn Ser Asn His Gly Ile Asp Ile Ile Ser Arg
 85 90 95
 Pro Ser Ser Ile Gly Leu Asp Arg Trp Glu Met Ile Lys Leu Glu Ala
 100 105 110
 Ile Met Lys Tyr Lys Lys Tyr Ile Asn Asn Tyr Thr Ser Glu Asn Phe
 115 120 125
 Asp Lys Leu Asp Gln Gln Leu Lys Asp Asn Phe Lys Leu Ile Ile Glu
 130 135 140
 Ser Lys Ser Glu Lys Ser Glu Ile Phe Ser Lys Leu Glu Asn Leu Asn
 145 150 155 160
 Val Ser Asp Leu Glu Ile Lys Ile Ala Phe Ala Leu Gly Ser Val Ile
 165 170 175
 Asn Gln Ala Leu Ile Ser Lys Gln Gly Ser Tyr Leu Thr Asn Leu Val
 180 185 190
 Ile Glu Gln Val Lys Asn Arg Tyr Gln Phe Leu Asn Gln His Leu Asn
 195 200 205
 Pro Ala Ile Glu Ser Asp Asn Asn Phe Thr Asp Thr Thr Lys Ile Phe
 210 215 220
 His Asp Ser Leu Phe Asn Ser Ala Thr Ala Glu Asn Ser Met Phe Leu
 225 230 235 240
 Thr Lys Ile Ala Pro Tyr Leu Gln Val Gly Phe Met Pro Glu Ala Arg
 245 250 255
 Ser Thr Ile Ser Leu Ser Gly Pro Gly Ala Tyr Ala Ser Ala Tyr Tyr
 260 265 270

Arg Ala Asn Ser Leu Phe Thr Glu Gln Glu Leu Leu Asn Ile Tyr Ser
 275 280 285
 Gln Glu Leu Leu Asn Arg Gly Asn Leu Ala Ala Ala Ser Asp Ile Val
 290 295 300
 Arg Leu Leu Ala Leu Lys Asn Phe Gly Gly Val Tyr Leu Asp Val Asp
 305 310 315 320
 Met Leu Pro Gly Ile His Ser Asp Leu Phe Lys Thr Asp Phe Ile Asn
 325 330 335
 Leu Gln Glu Asn Thr Ile Glu Lys Thr Leu Lys Ala Ser Asp Leu Ile
 340 345 350
 Glu Phe Lys Phe Pro Glu Asn Asn Leu Ser Gln Leu Thr Glu Gln Glu
 355 360 365
 Ile Asn Ser Leu Trp Ser Phe Asp Gln Ala Ser Ala Lys Tyr Gln Phe
 370 375 380
 Glu Lys Tyr Val Arg Asp
 385 390
 <210> 20
 <211> 52
 <212> PRT
 <213> Clostridium difficile
 <400> 20
 Ser Leu Gly Tyr Ile Met Ser Asn Phe Lys Ser Phe Asn Ser Glu Asn
 1 5 10 15
 Glu Leu Asp Arg Asp His Leu Gly Phe Lys Ile Ile Asp Asn Lys Thr
 20 25 30
 Tyr Tyr Tyr Asp Glu Asp Ser Lys Leu Val Lys Gly Leu Ile Asn Ile
 35 40 45
 Asn Asn Ser Leu
 50
 <210> 21
 <211> 279
 <212> PRT
 <213> Clostridium difficile
 <400> 21
 Asn Lys Ile Pro Ser Asn Asn Val Glu Glu Ala Gly Ser Lys Asn Tyr
 1 5 10 15
 Val His Tyr Ile Ile Gln Leu Gln Gly Asp Asp Ile Ser Tyr Glu Ala
 20 25 30
 Thr Cys Asn Leu Phe Ser Lys Asn Pro Lys Asn Ser Ile Ile Ile Gln

35	40	45
Arg Asn Met Asn Glu Ser Ala Lys Ser Tyr Phe Leu Ser Asp Asp Gly		
50	55	60
Glu Ser Ile Leu Glu Leu Asn Lys Tyr Arg Ile Pro Glu Arg Leu Lys		
65	70	75
Asn Lys Glu Lys Val Lys Val Thr Phe Ile Gly His Gly Lys Asp Glu		
85	90	95
Phe Asn Thr Ser Glu Phe Ala Arg Leu Ser Val Asp Ser Leu Ser Asn		
100	105	110
Glu Ile Ser Ser Phe Leu Asp Thr Ile Lys Leu Asp Ile Ser Pro Lys		
115	120	125
Asn Val Glu Val Asn Leu Leu Gly Cys Asn Met Phe Ser Tyr Asp Phe		
130	135	140
Asn Val Glu Glu Thr Tyr Pro Gly Lys Leu Leu Leu Ser Ile Met Asp		
145	150	155
Lys Ile Thr Ser Thr Leu Pro Asp Val Asn Lys Asn Ser Ile Thr Ile		
165	170	175
Gly Ala Asn Gln Tyr Glu Val Arg Ile Asn Ser Glu Gly Arg Lys Glu		
180	185	190
Leu Leu Ala His Ser Gly Lys Trp Ile Asn Lys Glu Glu Ala Ile Met		
195	200	205
Ser Asp Leu Ser Ser Lys Glu Tyr Ile Phe Phe Asp Ser Ile Asp Asn		
210	215	220
Lys Leu Lys Ala Lys Ser Lys Asn Ile Pro Gly Leu Ala Ser Ile Ser		
225	230	235
Glu Asp Ile Lys Thr Leu Leu Leu Asp Ala Ser Val Ser Pro Asp Thr		
245	250	255
Lys Phe Ile Leu Asn Asn Leu Lys Leu Asn Ile Glu Ser Ser Ile Gly		
260	265	270
Asp Tyr Ile Tyr Tyr Glu Lys		
275		
<210> 22		
<211> 100		
<212> PRT		
<213> Clostridium difficile		
<400> 22		
Gln Tyr Glu Val Arg Ile Asn Ser Glu Gly Arg Lys Glu Leu Leu Ala		
1	5	10
His Ser Gly Lys Trp Ile Asn Lys Glu Glu Ala Ile Met Ser Asp Leu		

20	25	30
Ser Ser Lys Glu Tyr Ile Phe Phe Asp Ser Ile Asp Asn Lys Leu Lys		
35	40	45
Ala Lys Ser Lys Asn Ile Pro Gly Leu Ala Ser Ile Ser Glu Asp Ile		
50	55	60
Lys Thr Leu Leu Leu Asp Ala Ser Val Ser Pro Asp Thr Lys Phe Ile		
65	70	75
Leu Asn Asn Leu Lys Leu Asn Ile Glu Ser Ser Ile Gly Asp Tyr Ile		
85	90	95
Tyr Tyr Glu Lys		
100		

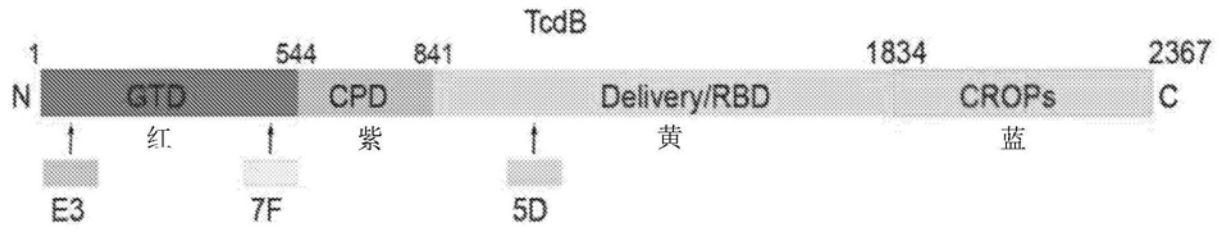


图1A

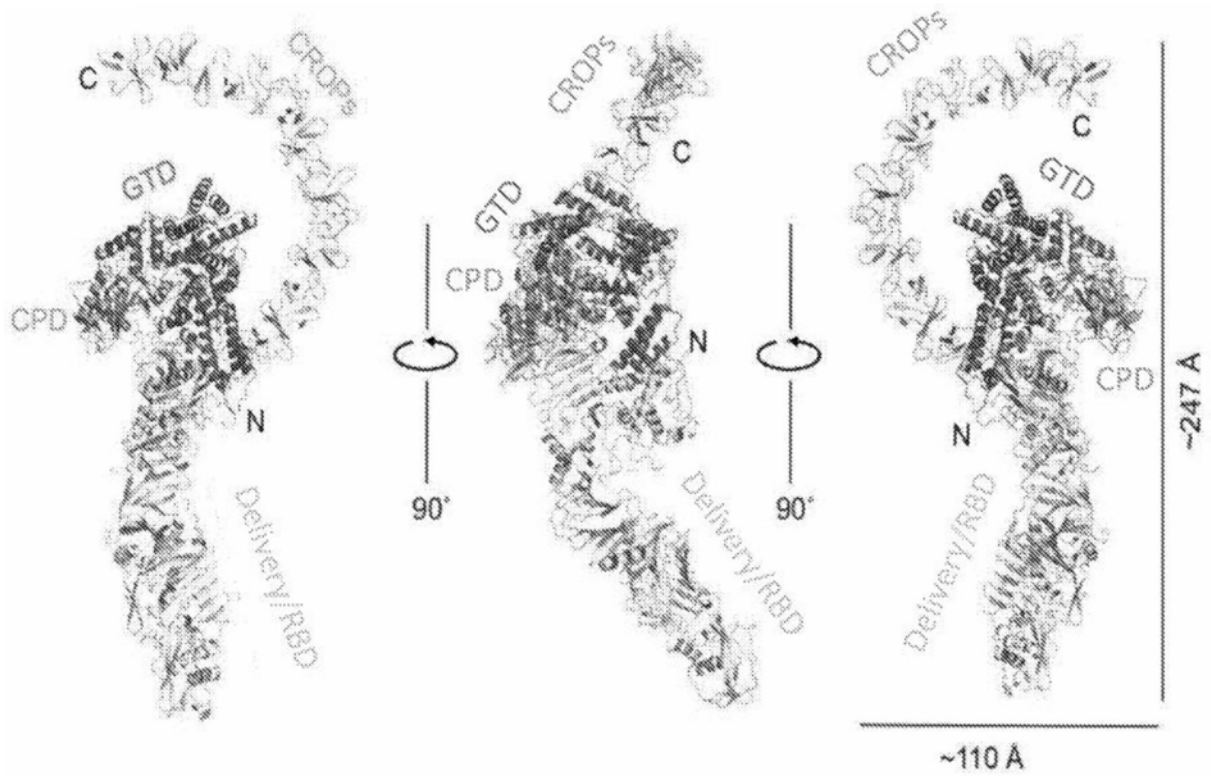


图1B

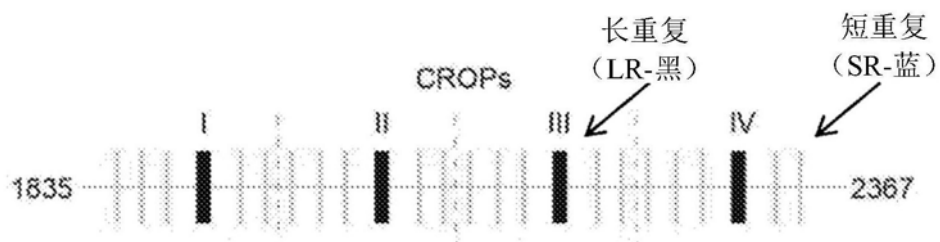


图2A

图 2B

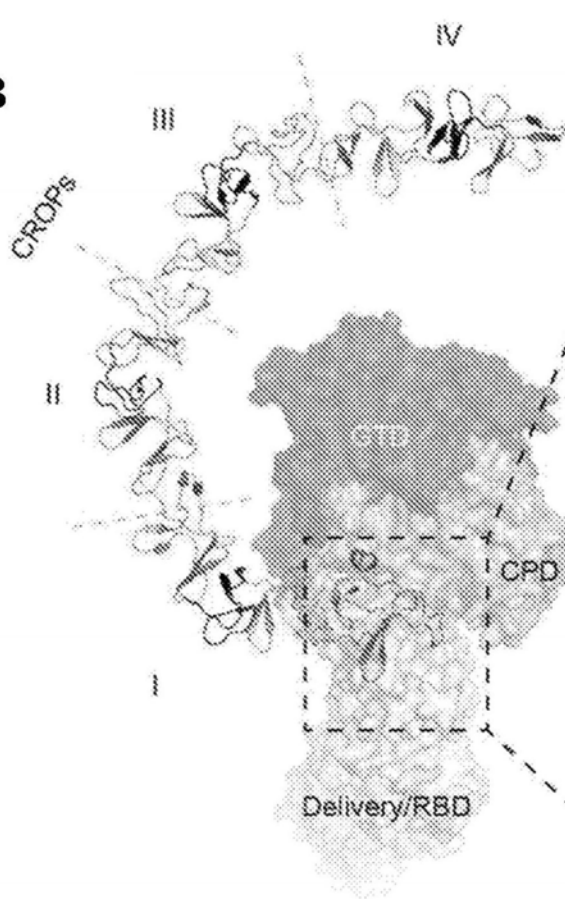


图 2D

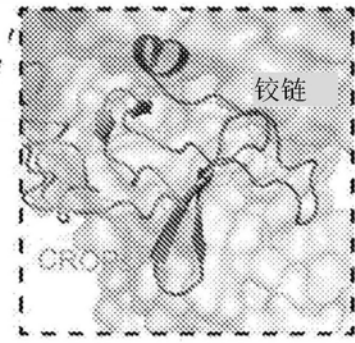


图 2E

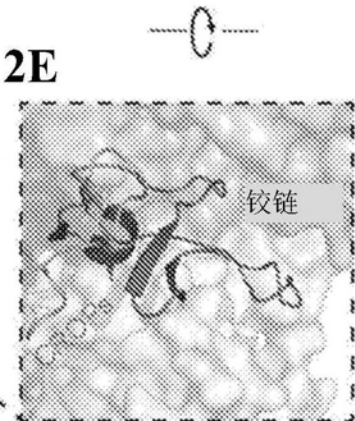
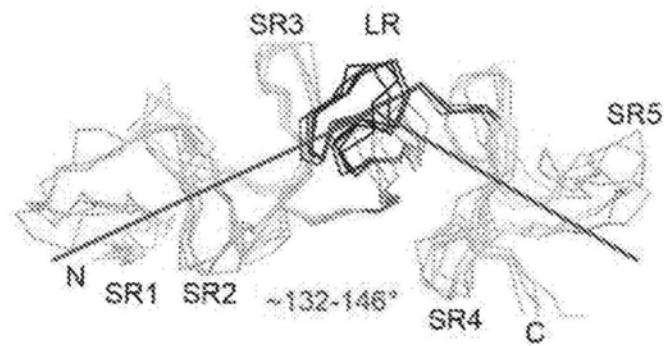


图 2C



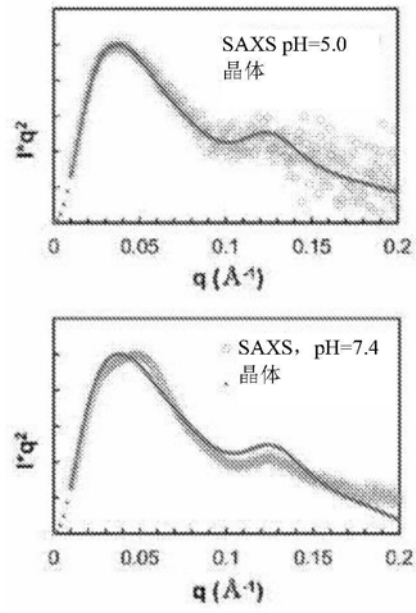


图3A

图 3B

XL-MS 测定的连接

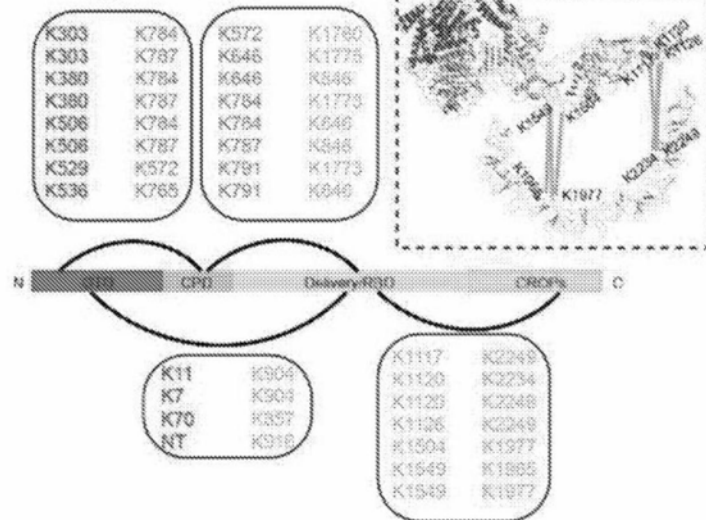


图 3C

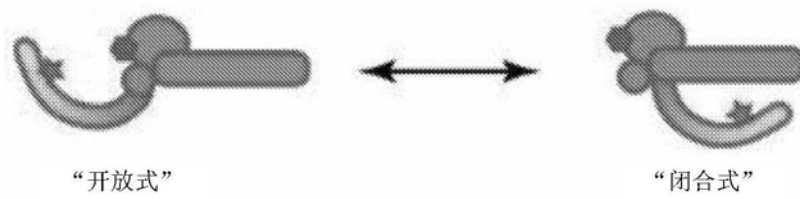


图3D

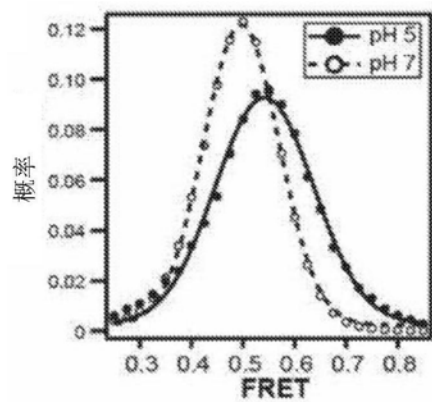


图3E

图 4A

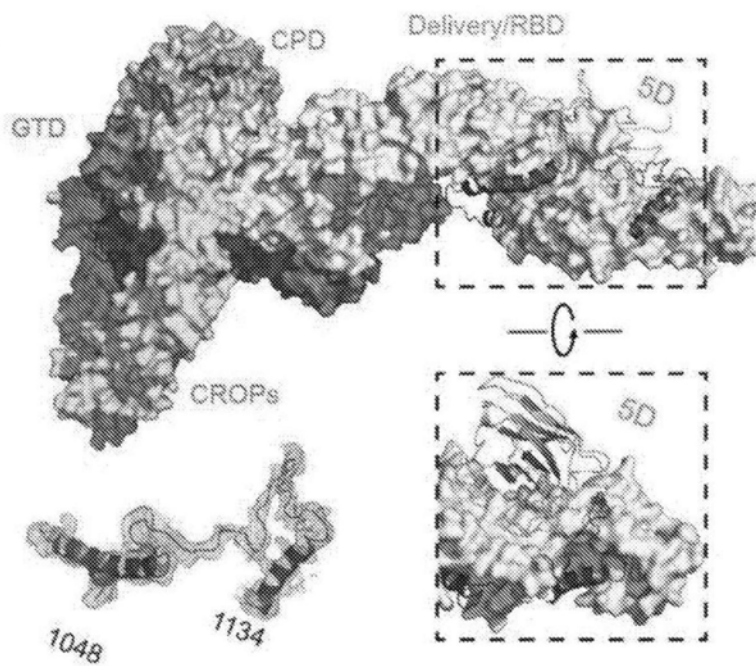


图 4B

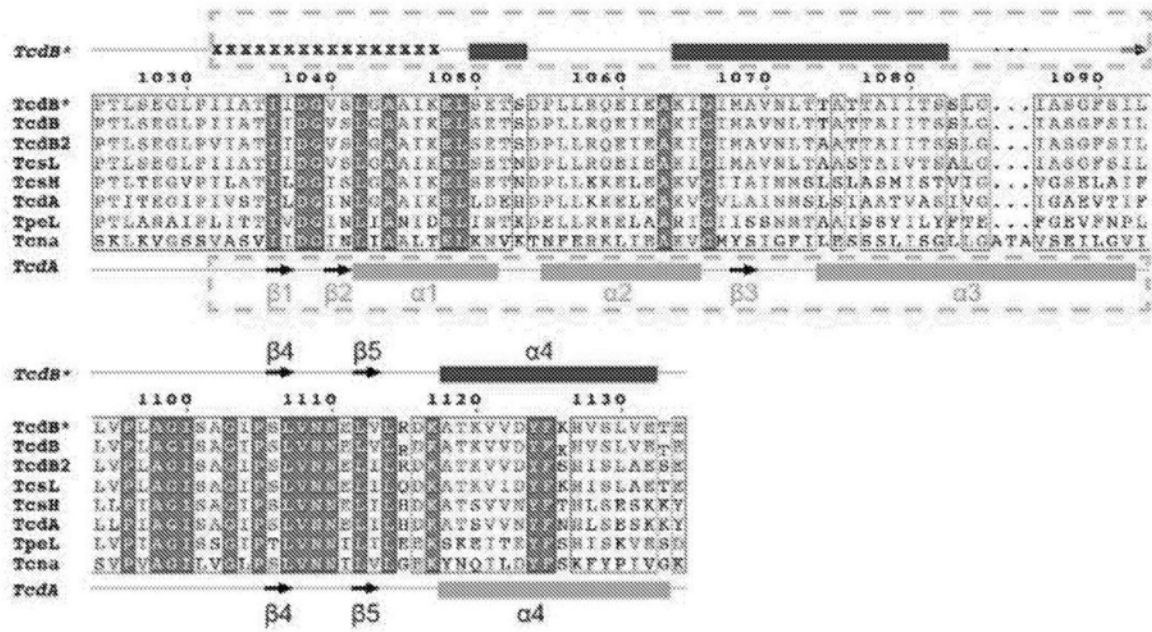


图4C

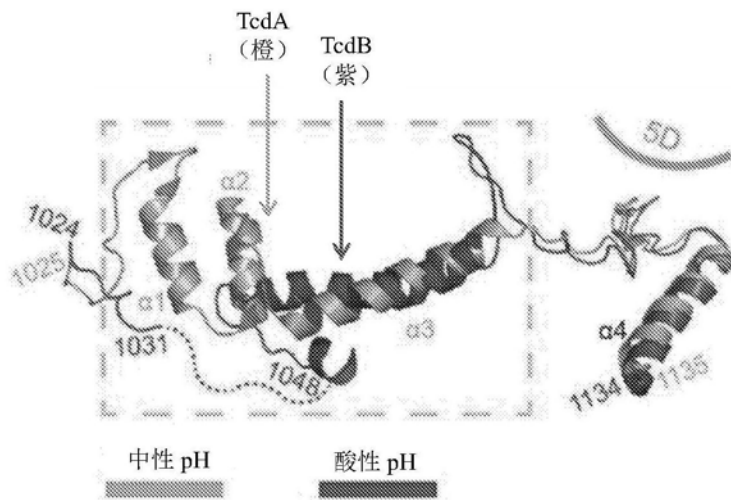


图4D

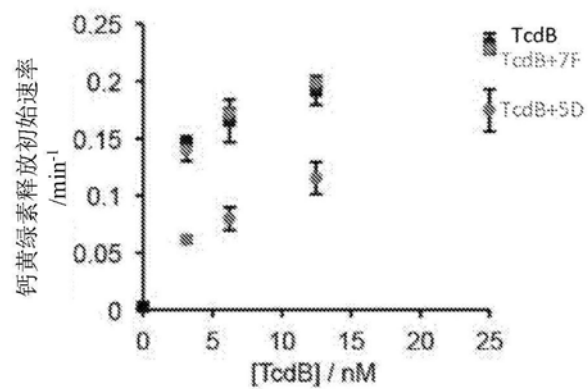


图4E

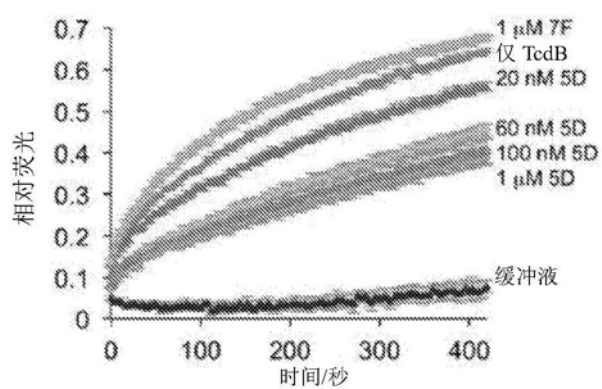


图4F

图 5A

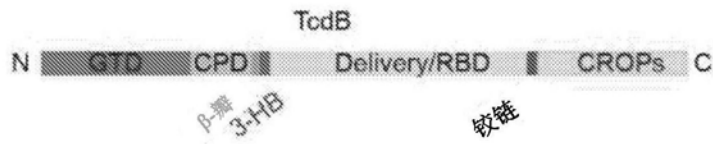


图 5B

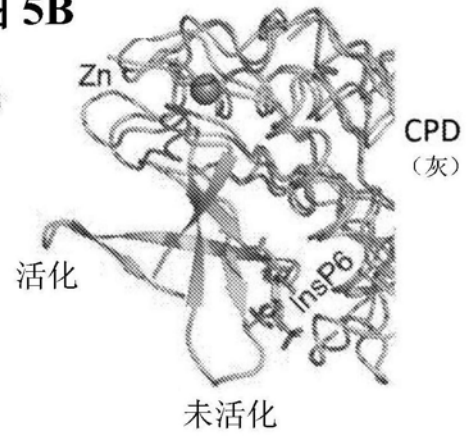


图 5C

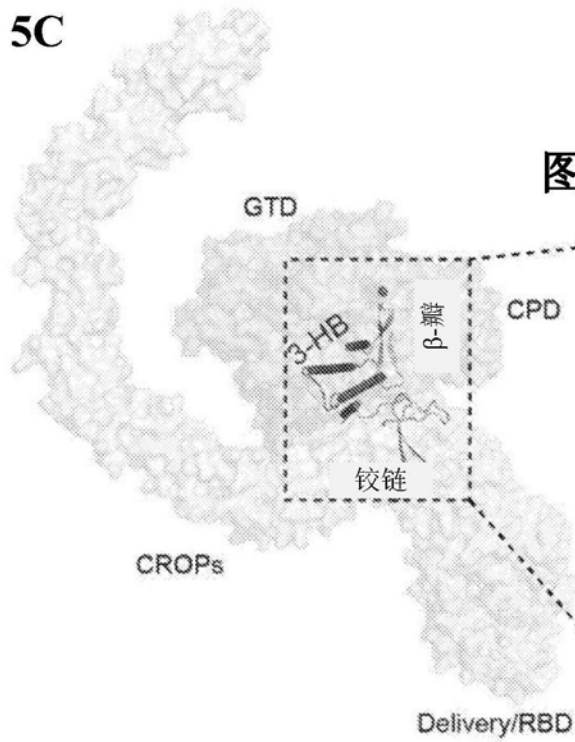
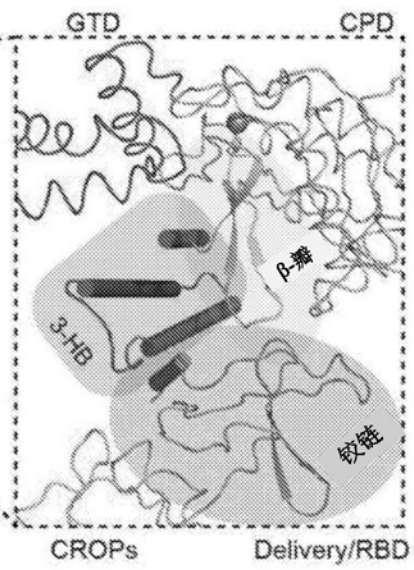


图 5D



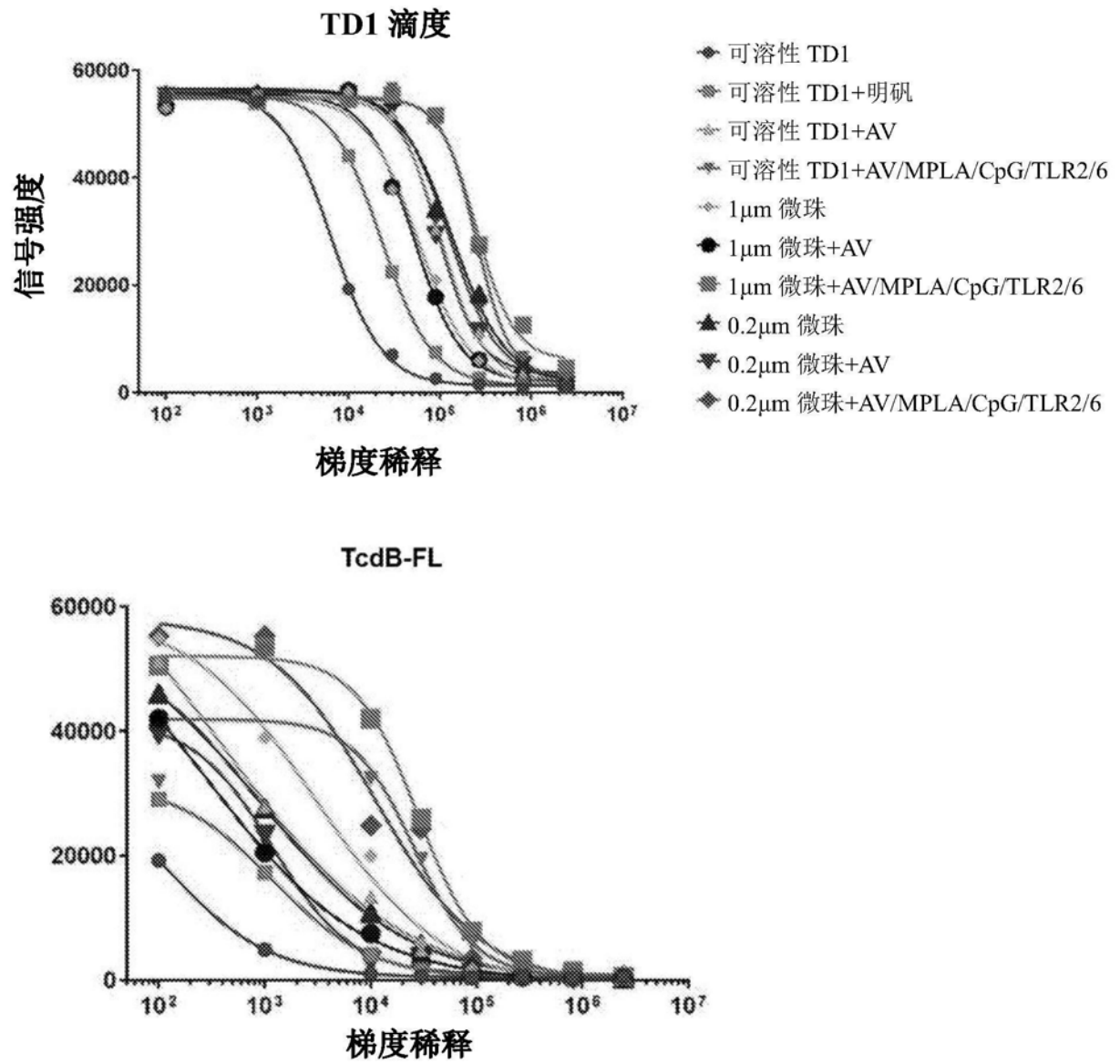


图6