



(12) 发明专利申请

(10) 申请公布号 CN 118369352 A

(43) 申请公布日 2024. 07. 19

(21) 申请号 202280080968.2
(22) 申请日 2022.08.10
(30) 优先权数据
20210100770 2021.11.04 GR
(85) PCT国际申请进入国家阶段日
2024.06.06
(86) PCT国际申请的申请数据
PCT/CA2022/051225 2022.08.10
(87) PCT国际申请的公布数据
W02023/077210 EN 2023.05.11
(71) 申请人 病童医院
地址 加拿大安大略
(72) 发明人 R·梅尔尼克 G·拜尔哈茨
S·吉尔 S·苏吉曼-马兰戈斯

(74) 专利代理机构 北京律盟知识产权代理有限公司 11287
专利代理师 刘媛媛

(51) Int.Cl.
C07K 19/00 (2006.01)
A61K 47/64 (2006.01)
A61P 35/00 (2006.01)
C07K 14/35 (2006.01)
C07K 14/36 (2006.01)
C07K 14/195 (2006.01)
C07K 16/00 (2006.01)
C07K 16/28 (2006.01)
C12N 15/62 (2006.01)

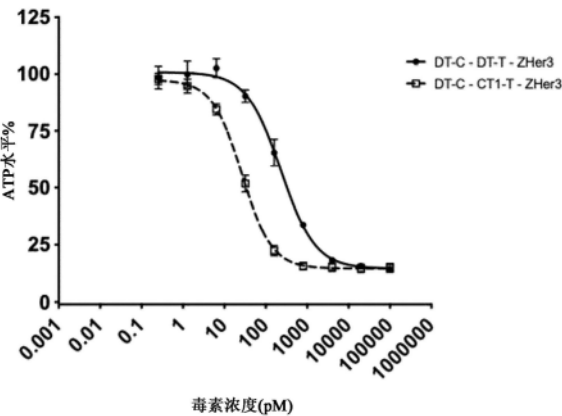
权利要求书9页 说明书37页
序列表（电子公布） 附图16页

(54) 发明名称

用于细胞递送的新颖细菌易位结构域和包含其的重组多肽

(57) 摘要

一般来说,本公开提供用于细胞递送的新颖细菌易位结构域。一个实例是来自SEQ ID NO:2的龟奥斯特维克氏菌蛋白质的易位结构域,所述易位结构域具有SEQ ID NO:3的氨基酸序列。还描述了来自其它细菌菌株和物种的易位酶结构域(例如SEQ ID NO:4至17和36至48)。描述了包含这些易位结构域的重组多肽。所述重组多肽旨在用于递送负荷分子,包括可以用于治疗疾病的治疗性多肽。



1.一种通式(I)的重组多肽,

A-B-C(i)

其中:

A是负荷分子;

B是易位多肽,其包含:

a) 来自以下的易位结构域:

i. SEQ ID NO:2的龟奥斯特维克氏菌蛋白质,

ii. GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌蛋白质,

iii. GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质,

iv. GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质,

v. GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质,

vi. GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌蛋白质,

vii. GenBank登录号JZ58907.1的松土壤链霉菌蛋白质,

viii. GenBank登录号WP_079110321.1的链霉菌属MBT76蛋白质,

ix. GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质,

x. GenBank登录号WP_095582082.1的白网链霉菌蛋白质,

xi. GenBank登录号WP_133259917.1的松生弯曲嗜酸菌蛋白质,

xii. GenBank登录号WP_073156187.1的嗜脲清野氏菌蛋白质,

xiii. GenBank登录号WP_189053160.1的吐鲁番长丝菌蛋白质,

xiv. GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质,

xv. GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质;

xvi. GenBank登录号EIZ2913133.1的产气克雷伯氏菌蛋白质,

xvii. GenBank登录号MCH0551590.1的链霉菌属MUM 178J蛋白质,

xviii. GenBank登录号MBB4677777.1的嗜冷克洛斯氏菌蛋白质,

xix. GenBank登录号WP_143261759.1的异库茨涅尔氏菌属NRRL B-24872蛋白质,

xx. GenBank登录号WP_156051914.1的白色异库茨涅尔氏菌蛋白质,

xxi. GenBank登录号WP_199893204.1的链霉菌属AV19蛋白质,

xxii. GenBank登录号WP_147264604.1的链霉菌属NRBC_110611蛋白质,

xxiii. GenBank登录号WP_209513619.1的注射器链霉菌蛋白质,

xxiv. GenBank登录号RJQ69589.1的假诺卡氏菌科细菌YIM PH 21723蛋白质,

xxv. GenBank登录号WP_143218892.1的曼谷放线动孢菌蛋白质,

xxvi. GenBank登录号MBF6055834.1的优洛菌素链霉菌蛋白质,

xxvii. GenBank登录号WP_169790908.1的灭病菌素链霉菌蛋白质,或

xxviii. GenBank登录号WP_157868472.1的卡廷加链霉菌蛋白质,或

b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域;并且

C是靶向部分。

2.一种通式(I)的重组多肽,

A-B-C(i)

其中:

A是负荷分子；
B是易位多肽,其包含：
a) 来自以下的易位结构域：
i. SEQ ID NO:2的龟奥斯特维克氏菌蛋白质，
ii. GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌蛋白质，
iii. GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质，
iv. GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质，
v. GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质，
vi. GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌蛋白质，
vii. GenBank登录号JZ58907.1的松土壤链霉菌蛋白质，
viii. GenBank登录号WP_079110321.1的链霉菌属MBT76蛋白质，
ix. GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质，
x. GenBank登录号WP_133259917.1的松生弯曲嗜酸菌蛋白质，
xi. GenBank登录号WP_189053160.1的吐鲁番长丝菌蛋白质，
xii. GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质，或
xiii. GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质
b) 与a) 中定义的所述易位结构域具有至少80%同一性的易位结构域；
c) 包含根据SEQ ID No:36至48中任一者的氨基酸序列的易位结构域；或
d) 与c) 中定义的所述易位结构域具有至少80%同一性的易位结构域,并且C是靶向部分。

3. 一种通式(I)的重组多肽，

A-B-C(i)

其中：

A是负荷分子；

B是易位多肽,其包含：

a) 来自以下的易位结构域：

i. SEQ ID NO:2的龟奥斯特维克氏菌蛋白质，

ii. GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质，

iii. GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质，

iv. GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质，或

v. GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质；或

b) 与a) 中定义的所述易位结构域具有至少80%同一性的易位结构域；并且

C是靶向部分。

4. 一种通式(I)的重组多肽，

A-B-C(i)

其中：

A是负荷分子；

B是易位多肽,其包含：

a) 来自以下的易位结构域：

- i. SEQ ID NO:2的龟奥斯特维克氏菌蛋白质,
- ii. GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质;或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域;并且
- C是靶向部分。

5. 根据权利要求1至4中任一权利要求所述的重组多肽,
其中在a)中:

来自所述龟奥斯特维克氏菌蛋白质的所述易位结构域具有SEQ ID NO:3的氨基酸序列,

来自GenBank登录号PSJ28985.1的所述无淀粉酶链孢囊菌蛋白质的所述易位结构域具有SEQ ID NO:4的氨基酸序列,

来自GenBank登录号SDT83331.1的所述链霉菌属TLI 053蛋白质的所述易位结构域具有SEQ ID NO:5的氨基酸序列,

来自GenBank登录号WP_160159328.1的所述链霉菌属SLBN-118蛋白质的所述易位结构域具有SEQ ID NO:6的氨基酸序列,

来自GenBank登录号WP_168096531.1的所述链霉菌属AA8蛋白质的所述易位结构域具有SEQ ID NO:7的氨基酸序列,

来自GenBank登录号WP_078659863.1的所述玫瑰轮丝链霉菌蛋白质的所述易位结构域具有SEQ ID NO:8的氨基酸序列,

来自GenBank登录号JZ58907.1的所述松土壤链霉菌蛋白质的所述易位结构域具有SEQ ID NO:9的氨基酸序列,

来自GenBank登录号WP_079110321.1的所述链霉菌属MBT76蛋白质的所述易位结构域具有SEQ ID NO:10的氨基酸序列,

来自GenBank登录号WP_120757473.1的所述克伦基链霉菌蛋白质的所述易位结构域具有SEQ ID NO:11的氨基酸序列,

来自GenBank登录号WP_095582082.1的所述白网链霉菌蛋白质的所述易位结构域具有SEQ ID NO:12的氨基酸序列,

来自GenBank登录号WP_133259917.1的所述松生弯曲嗜酸菌蛋白质的所述易位结构域具有SEQ ID NO:13的氨基酸序列,

来自GenBank登录号WP_073156187.1的所述嗜脲清野氏菌蛋白质的所述易位结构域具有SEQ ID NO:14的氨基酸序列,

来自GenBank登录号WP_189053160.1的所述吐鲁番长丝菌蛋白质的所述易位结构域具有SEQ ID NO:15的氨基酸序列,

来自GenBank登录号WP_219106995.1的所述奥斯特维克氏菌属TVS 96-490-7B蛋白质的所述易位结构域具有SEQ ID NO:16的氨基酸序列,并且

来自GenBank登录号WP_162873017.1的所述龟奥斯特维克氏菌LK16-18蛋白质的所述易位结构域具有SEQ ID NO:17的氨基酸序列。

6. 根据权利要求1至4中任一权利要求所述的重组多肽,其中所述易位多肽包含:

a) 来自所述龟奥斯特维克氏菌蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:3的氨基酸序列,或

- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
7. 根据权利要求1或2所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号PSJ28985.1的所述无淀粉酶链孢囊菌蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:4的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
8. 根据权利要求1至3中任一权利要求所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号SDT83331.1的所述链霉菌属TLI 053蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:5的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
9. 根据权利要求1或2所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号WP_160159328.1的所述链霉菌属SLBN-118蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:6的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
10. 根据权利要求1或2所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号WP_168096531.1的所述链霉菌属AA8蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:7的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
11. 根据权利要求1或2所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号WP_078659863.1的所述玫瑰轮丝链霉菌蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:8的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
12. 根据权利要求1或2所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号JZ58907.1的所述松土壤链霉菌蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:9的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
13. 根据权利要求1或2所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号WP_079110321.1的所述链霉菌属MBT76蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:10的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
14. 根据权利要求1至3中任一权利要求所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号WP_120757473.1的所述克伦基链霉菌蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:11的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
15. 根据权利要求1所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号WP_095582082.1的所述白网链霉菌蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:12的氨基酸序列,或
- b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
16. 根据权利要求1或2所述的重组多肽,其中所述易位多肽包含:
- a) 来自GenBank登录号WP_133259917.1的所述松生弯曲嗜酸菌蛋白质的所述易位结构

域,其中所述易位结构域具有SEQ ID NO:13的氨基酸序列,或

b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

17. 根据权利要求1所述的重组多肽,其中所述易位多肽包含:

a) 来自GenBank登录号WP_073156187.1的所述嗜脉清野氏菌蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:14的氨基酸序列,或

b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

18. 根据权利要求1或2所述的重组多肽,其中所述易位多肽包含:

a) 来自GenBank登录号WP_189053160.1的所述吐鲁番长丝菌蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:15的氨基酸序列,或

b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

19. 根据权利要求1至3中任一权利要求所述的重组多肽,其中所述易位多肽包含:

a) 来自GenBank登录号WP_219106995.1的所述奥斯特维克氏菌属TVS 96-490-7B蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:16的氨基酸序列,或

b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

20. 根据权利要求1至4中任一权利要求所述的重组多肽,其中所述易位多肽包含:

a) 来自GenBank登录号WP_116115734.1的所述龟奥斯特维克氏菌LK16-18蛋白质的所述易位结构域,其中所述易位结构域具有SEQ ID NO:17的氨基酸序列,或

b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

21. 根据权利要求1至20中任一权利要求所述的重组多肽,其中所述易位结构域如b)中所定义并且在其全长上与a)中定义的所述易位结构域具有至少85%同一性。

22. 根据权利要求1至20中任一权利要求所述的重组多肽,其中所述易位结构域如b)中所定义并且在其全长上与a)中定义的所述易位结构域具有至少90%同一性。

23. 根据权利要求1至20中任一权利要求所述的重组多肽,其中所述易位结构域如b)中所定义并且在其全长上与a)中定义的所述易位结构域具有至少95%同一性。

24. 根据权利要求1至20中任一权利要求所述的重组多肽,其中所述易位结构域如b)中所定义并且在其全长上与a)中定义的所述易位结构域具有至少98%同一性。

25. 根据权利要求1至20中任一权利要求所述的重组多肽,其中所述易位结构域如b)中所定义并且在其全长上与a)中定义的所述易位结构域具有至少99%同一性。

26. 根据权利要求1至20中任一权利要求所述的重组多肽,其中所述易位多肽如a)中所定义。

27. 根据权利要求1至26中任一权利要求所述的重组多肽,其中A和B通过氨基酸接头分隔开。

28. 根据权利要求27或28所述的重组多肽,其中所述氨基酸接头包含 $(G_4S)_2$ 。

29. 根据权利要求27或28所述的重组多肽,其中所述氨基酸接头是可裂解的,优选地其中所述氨基酸接头包含蛋白酶识别位点,优选地其中所述蛋白酶识别位点被半胱氨酸残基包围,更优选地其中所述蛋白酶识别位点是弗林蛋白酶识别位点。

30. 根据权利要求29所述的重组多肽,其中所述氨基酸接头包含SEQ ID NO:32,其包含所述弗林蛋白酶识别位点。

31. 根据权利要求30所述的重组多肽,其中所述 $(G_4S)_2$ 相对于所述SEQ ID NO:32位于N

端。

32. 根据权利要求1至27中任一权利要求所述的重组多肽,其中B和C通过氨基酸接头分隔开。

33. 根据权利要求27或28所述的重组多肽,其中所述氨基酸接头包含 $(G_4S)_2$ 。

34. 根据权利要求32或33所述的重组多肽,其中所述氨基酸接头包含SEQ ID NO:33。

35. 根据权利要求34所述的重组多肽,其中所述 $(G_4S)_2$ 相对于所述SEQ ID NO:33位于N端。

36. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:22的氨基酸1至821具有至少80%同一性、优选与SEQ ID NO:22的氨基酸1至821具有90%同一性、更优选与SEQ ID NO:22的氨基酸1至821具有95%同一性、甚至更优选与SEQ ID NO:22的氨基酸1至821具有100%同一性的氨基酸序列。

37. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:22的序列。

38. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:23的氨基酸1至822具有至少80%同一性、优选与SEQ ID NO:23的氨基酸1至822具有90%同一性、更优选与SEQ ID NO:23的氨基酸1至822具有95%同一性、甚至更优选与SEQ ID NO:23的氨基酸1至822具有100%同一性的氨基酸序列。

39. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:23的序列。

40. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:24的氨基酸1至822具有至少80%同一性、优选与SEQ ID NO:24的氨基酸1至822具有90%同一性、更优选与SEQ ID NO:24的氨基酸1至822具有95%同一性、甚至更优选与SEQ ID NO:24的氨基酸1至822具有100%同一性的氨基酸序列。

41. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:24的序列。

42. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:25的氨基酸1至821具有至少80%同一性、优选与SEQ ID NO:25的氨基酸1至821具有90%同一性、更优选与SEQ ID NO:25的氨基酸1至821具有95%同一性、甚至更优选与SEQ ID NO:25的氨基酸1至821具有100%同一性的氨基酸序列。

43. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:25的序列。

44. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:26的氨基酸1至806具有至少80%同一性、优选与SEQ ID NO:26的氨基酸1至806具有90%同一性、更优选与SEQ ID NO:26的氨基酸1至806具有95%同一性、甚至更优选与SEQ ID NO:26的氨基酸1至806具有100%同一性的氨基酸序列。

45. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:26的序列。

46. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:27的氨基酸1至811具有至少80%同一性、优选与SEQ ID NO:27的氨基酸1至811具有90%同一性、更优选与SEQ ID NO:27的氨基酸1至811具有95%同一性、甚至更优选与SEQ ID NO:27的氨基酸1至811具有100%同一性的氨基酸序列。

47. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:27的序列。

48. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:28的氨基酸1至822具有至少80%同一性、优选与SEQ ID NO:28的氨基酸1至822具有90%同一性、

更优选与SEQ ID NO:28的氨基酸1至822具有95%同一性、甚至更优选与SEQ ID NO:28的氨基酸1至822具有100%同一性的氨基酸序列。

49. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:28的序列。

50. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:29的氨基酸1至796具有至少80%同一性、优选与SEQ ID NO:29的氨基酸1至796具有90%同一性、更优选与SEQ ID NO:29的氨基酸1至796具有95%同一性、甚至更优选与SEQ ID NO:29的氨基酸1至796具有100%同一性的氨基酸序列。

51. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:29的序列。

52. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含与SEQ ID NO:30的氨基酸1至822具有至少80%同一性、优选与SEQ ID NO:30的氨基酸1至822具有90%同一性、更优选与SEQ ID NO:30的氨基酸1至822具有95%同一性、甚至更优选与SEQ ID NO:30的氨基酸1至822具有100%同一性的氨基酸序列。

53. 根据权利要求1至4中任一权利要求所述的重组多肽,其包含SEQ ID NO:30的序列。

54. 根据权利要求1至35中任一权利要求所述的重组多肽,其中所述靶向部分包含靶向多肽或适体。

55. 根据权利要求48所述的重组多肽,其中所述靶向多肽包含抗体、抗体的结合片段、亲和体、肽、亲和素、DARPin或受体配体。

56. 根据权利要求55所述的重组多肽,其中所述靶向多肽包含针对Her3的亲和体。

57. 根据权利要求56所述的重组多肽,其中所述针对Her3的亲和体包含SEQ ID NO:19的氨基酸序列。

58. 根据权利要求59所述的重组多肽,其中所述靶向多肽包含 $\alpha v \beta 6$ 整合素的受体配体。

59. 根据权利要求52所述的重组多肽,其中所述 $\alpha v \beta 6$ 整合素的受体配体包含SEQ ID NO:20的氨基酸序列。

60. 根据权利要求54所述的重组多肽,其中所述靶向部分包含至少两种靶向多肽、至少两种适体或靶向多肽与适体的组合。

61. 根据权利要求60所述的重组多肽,其中所述至少两种靶向多肽选自以下的群组:抗体、抗体的结合片段、亲和体、肽、亲和素、DARPin、受体配体及其组合。

62. 根据权利要求60或61所述的重组多肽,其中所述至少两种靶向多肽包含针对Her3的亲和体。

63. 根据权利要求62所述的重组多肽,其中所述针对Her3的亲和体包含SEQ ID NO:19的氨基酸序列。

64. 根据权利要求60至63中任一权利要求所述的重组多肽,其中所述至少两种靶向多肽包含 $\alpha v \beta 6$ 整合素的受体配体。

65. 根据权利要求64所述的重组多肽,其中所述 $\alpha v \beta 6$ 整合素的受体配体包含SEQ ID NO:20的氨基酸序列。

66. 根据权利要求60至65中任一权利要求所述的重组多肽,其中所述至少两种靶向多肽、所述至少两种适体或所述组合通过氨基酸接头分隔开。

67. 根据权利要求66所述的重组多肽,其中所述氨基酸接头包含 $(G_4S)_2$ 。

68. 根据权利要求1至35和54至67中任一权利要求所述的重组多肽,其中所述靶向部分

结合到细胞表面蛋白质。

69. 根据权利要求68所述的重组多肽,其中所述细胞表面蛋白质是谱系特异性或组织特异性的。

70. 根据权利要求67所述的重组多肽,其中所述细胞表面蛋白质是普遍表达的。

71. 根据权利要求67所述的重组多肽,其中所述细胞表面蛋白质表达于疾病细胞中。

72. 根据权利要求67所述的重组多肽,其中所述细胞表面蛋白质对于疾病细胞是特异性的并且不表达于相应健康细胞中。

73. 根据权利要求67所述的重组多肽,其中与相应健康细胞相比,所述细胞表面蛋白质在疾病细胞中具有升高的表达。

74. 根据权利要求71至73中任一权利要求所述的重组多肽,其中所述疾病细胞是癌细胞。

75. 根据权利要求1至35和54至74中任一权利要求所述的重组多肽,其中所述负荷分子包含治疗性多肽。

76. 根据权利要求75所述的重组多肽,其中所述治疗性多肽包含细胞毒性多肽,优选多肽毒素或其功能片段。

77. 根据权利要求76所述的重组多肽,其中所述细胞毒性多肽包含来自白喉毒素的催化结构域。

78. 根据权利要求76所述的重组多肽,其中所述细胞毒性多肽包含来自龟毒素的催化结构域。

79. 根据权利要求78所述的重组多肽,其中所述催化结构域来自:

SEQ ID NO:2中所阐述的龟毒素(CT1),

SEQ ID NO:21中所阐述的龟毒素(CT2),或

SEQ ID NO:35中所阐述的龟毒素(CT3)。

80. 根据权利要求79所述的重组多肽,

其中来自SEQ ID NO:2中所阐述的龟毒素(CT1)的所述催化结构域具有根据SEQ ID NO:2的氨基酸位置1至186的氨基酸序列;

其中来自SEQ ID NO:21中所阐述的龟毒素(CT2)的所述催化结构域具有根据SEQ ID NO:21的氨基酸位置1至186的氨基酸序列;或

其中来自SEQ ID NO:35中所阐述的龟毒素(CT3)的所述催化结构域具有根据SEQ ID NO:35的氨基酸位置1至191的氨基酸序列。

81. 根据权利要求75所述的重组多肽,其中所述治疗性多肽包含疾病状态中缺乏的蛋白质或其功能片段。

82. 根据权利要求75所述的重组多肽,其中所述治疗性多肽包含来自创伤弧菌的Ras/Rap1特异性内肽酶(RRSP)(SEQ ID NO:18)。

83. 根据权利要求1至35和54至74中任一权利要求所述的重组多肽,其中所述负荷分子包含用于“点击”化学的N端半胱氨酸残基。

84. 根据权利要求1至35、54至74和83中任一权利要求所述的重组多肽,其中所述负荷分子包含核酸分子。

85. 一种核酸,其编码根据权利要求1至81中任一权利要求所述的重组多肽。

86. 一种组合物,其包含根据权利要求1至82中任一权利要求所述的重组多肽以及可接受的赋形剂、稀释剂或载剂。

87. 一种药物组合物,其包含根据权利要求1至82中任一权利要求所述的重组多肽以及药学上可接受的赋形剂、稀释剂或载剂。

88. 一种将负荷分子递送到细胞的方法,其包含使所述细胞与根据权利要求1至82中任一权利要求所述的重组多肽接触。

89. 一种根据权利要求1至82中任一权利要求所述的重组多肽的用途,其用于将所述负荷分子递送到细胞。

90. 根据权利要求1至82中任一权利要求所述的重组多肽,其用于将所述负荷分子递送到细胞。

91. 一种治疗受试者的癌症的方法,其包含向所述受试者投与根据权利要求1至82中任一权利要求所述的重组多肽。

92. 一种根据权利要求1至82中任一权利要求所述的重组多肽的用途,其用于治疗受试者的癌症。

93. 根据权利要求1至82中任一权利要求所述的重组多肽,其用于治疗受试者的癌症。

用于细胞递送的新颖细菌易位结构域和包含其的重组多肽

[0001] 相关申请的交叉引用

[0002] 本申请要求2021年11月4日提交的标题为“用于细胞递送的新颖细菌易位结构域和包含其的重组多肽 (NOVEL BACTERIAL TRANSLOCATION DOMAINS AND RECOMBINANT POLYPEPTIDES COMPRISING THEM FOR USE IN CELLULAR DELIVERY)”的希腊专利申请第20210100770号的优先权。

技术领域

[0003] 本公开大体上涉及一种递送平台。更特定地,本公开涉及一种用于将运载物(cargo)分子递送到细胞的基于细菌毒素的平台。

背景技术

[0004] 治疗性分子通常难以递送到细胞。其通常不容易穿透生物膜。免疫毒素是一类由细菌毒素(如白喉毒素(DT))构成的生物治疗药物,所述细菌毒素已通过重新靶向其受体结合结构域(RBD)以靶向癌症受体并且通过递送靶向细胞内肿瘤蛋白的酶运载物而改换用途,例如用于癌症靶向疗法。然而,针对白喉的全球疫苗接种计划已经引起针对DT和基于DT的治疗剂的群体水平免疫。

[0005] 仍需要提供将治疗剂(包括基于蛋白质的治疗剂)递送到细胞的可能性的递送平台。

发明内容

[0006] 本公开的目的是消除或减轻先前方法的至少一个缺点。

[0007] 在第一方面中,本公开提供一种通式(I)的重组多肽:

[0008] A-B-C(i)

[0009] 其中:

[0010] A是运载物分子;

[0011] B是易位多肽,其包含:

[0012] a) 来自以下的易位结构域:

[0013] SEQ ID NO:2的龟奥斯特维克氏菌(*Austwickia chelonae*)蛋白质,

[0014] GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌(*Streptosporangium nondiastaticum*)蛋白质,

[0015] GenBank登录号SDT83331.1的链霉菌属(*Streptomyces* sp.)TLI 053蛋白质,

[0016] GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质,

[0017] GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质,

[0018] GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌(*Streptomyces roseoverticillatus*)蛋白质,

[0019] GenBank登录号JZ58907.1的松土壤链霉菌(*Streptomyces piniterrae*)蛋白质,

- [0020] GenBank登录号WP_079110321.1的链霉菌属MBT76蛋白质,
- [0021] GenBank登录号WP_120757473.1的克伦基链霉菌 (*Streptomyces klenkii*) 蛋白质,
- [0022] GenBank登录号WP_095582082.1的白网链霉菌 (*Streptomyces albireticuli*) 蛋白质,
- [0023] GenBank登录号WP_133259917.1的松生弯曲嗜酸菌 (*Streptacidiphilus pinicola*) 蛋白质,
- [0024] GenBank登录号WP_073156187.1的嗜脲清野氏菌 (*Seinonella peptonophila*) 蛋白质,
- [0025] GenBank登录号WP_189053160.1的吐鲁番长丝菌 (*Longimycelium tulufanense*) 蛋白质,
- [0026] GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质,
- [0027] GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质;
- [0028] GenBank登录号EIZ2913133.1的产气克雷伯氏菌 (*Klebsiella aerogenes*) 蛋白质,
- [0029] GenBank登录号MCH0551590.1的链霉菌属MUM 178J蛋白质,
- [0030] GenBank登录号MBB4677777.1的嗜冷克洛斯氏菌 (*Crossiella cryophila*) 蛋白质,
- [0031] GenBank登录号WP_143261759.1的异库茨涅尔氏菌属 (*Allokutzneria* sp.) NRRL B-24872蛋白质,
- [0032] GenBank登录号WP_156051914.1的白色异库茨涅尔氏菌 (*Allokutzneria albata*) 蛋白质,
- [0033] GenBank登录号WP_199893204.1的链霉菌属AV19蛋白质,
- [0034] GenBank登录号WP_147264604.1的链霉菌属NRBC_110611蛋白质,
- [0035] GenBank登录号WP_209513619.1的注射器链霉菌 (*Streptomyces syringium*) 蛋白质,
- [0036] GenBank登录号RJQ69589.1的假诺卡氏菌科细菌 (*Pseudonocardiaceae bacterium*) YIM PH 21723蛋白质,
- [0037] GenBank登录号WP_143218892.1的曼谷放线动孢菌 (*Actinokineospora bangkokensis*) 蛋白质,
- [0038] GenBank登录号MBF6055834.1的优洛菌素链霉菌 (*Streptomyces eurocidicus*) 蛋白质,
- [0039] GenBank登录号WP_169790908.1的灭病菌素链霉菌 (*Streptomyces pathocidini*) 蛋白质,或
- [0040] GenBank登录号WP_157868472.1的卡廷加链霉菌 (*Streptomyces caatingaensis*) 蛋白质,或
- [0041] b) 与a) 中定义的所述易位结构域具有至少80%同一性的易位结构域;并且
- [0042] C是靶向部分。
- [0043] 在一个方面中,提供编码如本文所定义的重组多肽的核酸。

- [0044] 在一个方面中,提供包含如本文所定义的核酸的载体。
- [0045] 在另一方面中,本公开提供一种组合物,其包含如本文所定义的重组多肽以及可接受的赋形剂、稀释剂或载剂。
- [0046] 在一个方面实施例中,提供一种药物组合物,其包含如本文所定义的重组多肽以及药学上可接受的赋形剂、稀释剂或载剂。
- [0047] 在一个方面中,提供一种将运载物分子递送到细胞的方法,其包含使所述细胞与如本文所定义的重组多肽接触。
- [0048] 在一个方面中,提供如本文所定义的重组多肽用于将运载物分子递送到细胞的用途。
- [0049] 在一个方面中,提供如本文所定义的重组多肽用于制备用于将运载物分子递送到细胞的药剂的用途。
- [0050] 在一个方面中,提供如本文所定义的重组多肽,其用于将运载物分子递送到细胞。
- [0051] 在一个方面中,提供一种治疗受试者的癌症的方法,其包含向所述受试者投与如本文所定义的重组多肽。
- [0052] 在一个方面中,提供如本文所定义的重组多肽用于治疗受试者的癌症的用途。
- [0053] 在一个方面中,提供如本文所定义的重组多肽用于制备用于治疗受试者的癌症的药剂的用途。
- [0054] 在一个方面中,提供如本文所定义的重组多肽,其用于治疗受试者的癌症。
- [0055] 所属领域的技术人员在结合附图查阅了特定实施例的以下描述后,将容易理解本公开的其它方面和特征。

附图说明

- [0056] 现在将参考附图仅通过举例的方式描述本公开的实施例。
- [0057] 图1A描绘白喉毒素的细胞进入机制的示意图。
- [0058] 图1B描绘DT和CT1的结构域组织,展示催化结构域(C)和桥接弗林蛋白酶(furin)识别位点(F),之后是易位酶(T)和受体结合结构域(R)。
- [0059] 图1C描绘白喉毒素和CT1的晶体结构。
- [0060] 图2A展示评估CT1催化结构域功能和释放的实验的结果。DT和含有CT1的C结构域的DT嵌合体展示对HEK293T细胞(b)的类似毒性。
- [0061] 图2B展示DT和含有CT1的C结构域的DT嵌合体显示对DPH4^{-/-}细胞的细胞活力没有影响。
- [0062] 图2C展示当DT和CT1与细胞裂解物一起培育时,在弗林蛋白酶识别位点处存在裂解。
- [0063] 图3A展示DT和含有来自CT1的弗林蛋白酶位点的DT嵌合体对非洲绿猴肾细胞(Vero cell)是等效的,证明两种毒素的有效运载物释放。
- [0064] 图3B展示通过CT1-T介导的DT-C递送,非洲绿猴肾细胞中的蛋白质合成减少。
- [0065] 图3C展示CT1-T介导的非天然运载物RRSP的递送。RRSP对RAS突变细胞,如CFPAC-1有毒。
- [0066] 图3D展示通过RRSP裂解细胞内RAS证实了递送。

[0067] 图3E展示通过测量HPAF II细胞上的细胞活力,CT1-T介导的DT-C递送。出乎意料的是,当与DT-R之外的靶向结构域复合时,CT1-T是比DT-T更有效的易位酶。

[0068] 图3F展示通过测量HPAF II细胞上的细胞活力,CT1-T介导的RRSP递送。

[0069] 图4A展示评估CT1人血清结合和中和的实验的结果。为了定量人血清中预先存在的抗DT或抗CT1抗体的水平,将DT或CT1固定在Nunc MaxiSorp™板上,并且与各种稀释度的人血清一起培育。随后将孔与缀合到HRP的抗人IgG抗体一起培育,所述抗体使用TMB试剂显色。在630nm下读取吸光度。

[0070] 图4B展示评估CT1人血清结合和中和的进一步实验的结果。为了确定中和抗体对DT或CT1使细胞中毒的能力的影响,将各种DT/CT1毒素嵌合体(指定)与人血清(或PBS)一起培育,随后添加到非洲绿猴肾细胞中,并且测量蛋白质合成水平。计算EC50值并且绘制与PBS对照的差异倍数。人血清对CT1使细胞中毒的能力没有影响,证明其未被人血清中和。

[0071] 图5是DT CT1之间保守的关键残基的示意图。

[0072] 图6A和6B描绘易位酶的功能表征的结果。

[0073] 图7A展示人血清中的抗体识别DT,但未展示与CT1的结合。

[0074] 图7B展示人血清中的抗体识别来自DT(而非CT1)的易位酶。

[0075] 图8展示抗DT抗体不中和基于CT1的免疫毒素。

[0076] 图9展示说明各种易位酶的关系的系统发生树。

具体实施方式

[0077] 一般来说,本公开提供用于细胞递送的新颖细菌易位结构域。描述了包含这些易位结构域的重组多肽。重组多肽旨在用于递送运载物分子,包括治疗性多肽。

[0078] 重组多肽

[0079] 在一个方面中,提供一种通式(I)的重组多肽:

[0080] A-B-C(i)

[0081] 其中:

[0082] A是运载物分子;

[0083] B是易位多肽,其包含:

[0084] a) 来自以下的易位结构域:

[0085] SEQ ID NO:2的龟奥斯特维克氏菌蛋白质,

[0086] GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌蛋白质,

[0087] GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质,

[0088] GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质,

[0089] GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质,

[0090] GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌蛋白质,

[0091] GenBank登录号JZ58907.1的松土壤链霉菌蛋白质,

[0092] GenBank登录号WP_0791103f21.1的链霉菌属MBT76蛋白质,

[0093] GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质,

[0094] GenBank登录号WP_095582082.1的白网链霉菌蛋白质,

[0095] GenBank登录号WP_133259917.1的松生弯曲嗜酸菌蛋白质,

- [0096] GenBank登录号WP_073156187.1的嗜脉清野氏菌蛋白质,
- [0097] GenBank登录号WP_189053160.1的吐鲁番长丝菌蛋白质,
- [0098] GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质,
- [0099] GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质;
- [0100] GenBank登录号EIZ2913133.1的产气克雷伯氏菌蛋白质,
- [0101] GenBank登录号MCH0551590.1的链霉菌属MUM 178J蛋白质,
- [0102] GenBank登录号MBB4677777.1的嗜冷克洛斯氏菌蛋白质,
- [0103] GenBank登录号WP_143261759.1的异库茨涅尔氏菌属NRRL B-24872蛋白质,
- [0104] GenBank登录号WP_156051914.1的白色异库茨涅尔氏菌蛋白质,
- [0105] GenBank登录号WP_199893204.1的链霉菌属AV19蛋白质,
- [0106] GenBank登录号WP_147264604.1的链霉菌属NRBC_110611蛋白质,
- [0107] GenBank登录号WP_209513619.1的注射器链霉菌蛋白质,
- [0108] GenBank登录号RJQ69589.1的假诺卡氏菌科细菌YIM PH 21723蛋白质,
- [0109] GenBank登录号WP_143218892.1的曼谷放线动孢菌蛋白质,
- [0110] GenBank登录号MBF6055834.1的优洛菌素链霉菌蛋白质,
- [0111] GenBank登录号WP_169790908.1的灭病菌素链霉菌蛋白质,或
- [0112] GenBank登录号WP_157868472.1的卡廷加链霉菌蛋白质,或
- [0113] b) 与a) 中定义的所述易位结构域具有至少80%同一性的易位结构域;并且
- [0114] C是靶向部分。
- [0115] 在一个实施例中,a), 易位结构域来自:
- [0116] -SEQ ID NO:2的龟奥斯特维克氏菌蛋白质,
- [0117] -GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌蛋白质,
- [0118] -GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质,
- [0119] -GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质,
- [0120] -GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质,
- [0121] -GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌蛋白质,
- [0122] -GenBank登录号JZ58907.1的松土壤链霉菌蛋白质,
- [0123] -GenBank登录号WP_079110321.1的链霉菌属MBT76蛋白质,
- [0124] -GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质,
- [0125] -GenBank登录号WP_133259917.1的松生弯曲嗜酸菌蛋白质,
- [0126] -GenBank登录号WP_189053160.1的吐鲁番长丝菌蛋白质,
- [0127] -GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质,或
- [0128] -GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质。
- [0129] 在一个实施例中,在a) 中,易位结构域来自:
- [0130] -SEQ ID NO:2的龟奥斯特维克氏菌蛋白质,
- [0131] -GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质,
- [0132] -GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质,
- [0133] -GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质,或
- [0134] -GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质。

- [0135] 在一个实施例中,在a)中,易位结构域来自:
- [0136] -SEQ ID NO:2的龟奥斯特维克氏菌蛋白质,
- [0137] -GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质。
- [0138] 在一个实施例中,在a)中:
- [0139] i.来自本文中称为CT1的龟奥斯特维克氏菌蛋白质的易位结构域具有SEQ ID NO:3的氨基酸序列,
- [0140] ii.来自GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌蛋白质的易位结构域具有SEQ ID NO:4的氨基酸序列,
- [0141] iii.来自GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质的易位结构域具有SEQ ID NO:5的氨基酸序列,
- [0142] iv.来自GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质的易位结构域具有SEQ ID NO:6的氨基酸序列,
- [0143] v.来自GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质的易位结构域具有SEQ ID NO:7的氨基酸序列,
- [0144] vi.来自GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌蛋白质的易位结构域具有SEQ ID NO:8的氨基酸序列,
- [0145] vii.来自GenBank登录号JZ58907.1的松土壤链霉菌蛋白质的易位结构域具有SEQ ID NO:9的氨基酸序列,
- [0146] viii.来自GenBank登录号WP_079110321.1的链霉菌属MBT76蛋白质的易位结构域具有SEQ ID NO:10的氨基酸序列,
- [0147] ix.来自GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质的易位结构域具有SEQ ID NO:11的氨基酸序列,
- [0148] x.来自GenBank登录号WP_095582082.1的白网链霉菌蛋白质的易位结构域具有SEQ ID NO:12的氨基酸序列,
- [0149] xi.来自GenBank登录号WP_133259917.1的松生弯曲嗜酸菌蛋白质的易位结构域具有SEQ ID NO:13的氨基酸序列,
- [0150] xii.来自GenBank登录号WP_073156187.1的嗜脉清野氏菌蛋白质的易位结构域具有SEQ ID NO:14的氨基酸序列,
- [0151] xiii.来自GenBank登录号WP_189053160.1的吐鲁番长丝菌蛋白质的易位结构域具有SEQ ID NO:15的氨基酸序列,
- [0152] xiv.来自GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质的易位结构域具有SEQ ID NO:16的氨基酸序列,
- [0153] xv.来自GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质的易位结构域具有SEQ ID NO:17的氨基酸序列,并且
- [0154] xvi.所述易位结构域具有SEQ ID No:36至48中任一者的氨基酸序列。
- [0155] 在一个实施例中,在a)中:
- [0156] i.来自本文中称为CT1的龟奥斯特维克氏菌蛋白质的易位结构域具有SEQ ID NO:3的氨基酸序列,
- [0157] ii.来自GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌蛋白质的易位结构域具

有SEQ ID NO:4的氨基酸序列,

[0158] iii.来自GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质的易位结构域具有SEQ ID NO:5的氨基酸序列,

[0159] iv.来自GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质的易位结构域具有SEQ ID NO:6的氨基酸序列,

[0160] v.来自GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质的易位结构域具有SEQ ID NO:7的氨基酸序列,

[0161] vi.来自GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌蛋白质的易位结构域具有SEQ ID NO:8的氨基酸序列,

[0162] vii.来自GenBank登录号JZ58907.1的松土壤链霉菌蛋白质的易位结构域具有SEQ ID NO:9的氨基酸序列,

[0163] viii.来自GenBank登录号WP_079110321.1的链霉菌属MBT76蛋白质的易位结构域具有SEQ ID NO:10的氨基酸序列,

[0164] ix.来自GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质的易位结构域具有SEQ ID NO:11的氨基酸序列,

[0165] xi.来自GenBank登录号WP_133259917.1的松生弯曲嗜酸菌蛋白质的易位结构域具有SEQ ID NO:13的氨基酸序列,

[0166] xiii.来自GenBank登录号WP_189053160.1的吐鲁番长丝菌蛋白质的易位结构域具有SEQ ID NO:15的氨基酸序列,

[0167] xiv.来自GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质的易位结构域具有SEQ ID NO:16的氨基酸序列,并且

[0168] xv.来自GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质的易位结构域具有SEQ ID NO:17的氨基酸序列。

[0169] 如本文所提及,“易位多肽”意指包含易位结构域的多肽。

[0170] 如本文所提及的“易位结构域”是一种多肽序列,其用于促进存在有所述序列的蛋白质跨细胞膜转运,从而促进细胞进入。这种活性可以例如使用本文所描述的分析来评估。

[0171] 在一个实施例中,易位多肽包含:

[0172] a) 来自龟奥斯特维克氏菌蛋白质CT1的易位结构域,其中所述易位结构域具有SEQ ID NO:3的氨基酸序列(本文中称为“CT1-T”),或

[0173] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0174] 在一个实施例中,易位多肽包含:

[0175] a) 来自GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:4的氨基酸序列,或

[0176] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0177] 在一个实施例中,易位多肽包含:

[0178] a) 来自GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:5的氨基酸序列,或

[0179] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0180] 在一个实施例中,易位多肽包含:

- [0181] a) 来自GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:6的氨基酸序列,或
- [0182] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0183] 在一个实施例中,易位多肽包含:
- [0184] a) 来自GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:7的氨基酸序列,或
- [0185] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0186] 在一个实施例中,易位多肽包含:
- [0187] a) 来自GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:8的氨基酸序列,或
- [0188] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0189] 在一个实施例中,易位多肽包含:
- [0190] a) 来自GenBank登录号JZ58907.1的松土壤链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:9的氨基酸序列,或
- [0191] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0192] 在一个实施例中,易位多肽包含:
- [0193] a) 来自GenBank登录号WP_079110321.1的链霉菌属MBT76蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:10的氨基酸序列,或
- [0194] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0195] 在一个实施例中,易位多肽包含:
- [0196] a) 来自GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:11的氨基酸序列,或
- [0197] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0198] 在一个实施例中,易位多肽包含:
- [0199] a) 来自GenBank登录号WP_095582082.1的白网链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:12的氨基酸序列,或
- [0200] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0201] 在一个实施例中,易位多肽包含:
- [0202] a) 来自GenBank登录号WP_133259917.1的松生弯曲嗜酸菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:13的氨基酸序列,或
- [0203] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0204] 在一个实施例中,易位多肽包含:
- [0205] a) 来自GenBank登录号WP_073156187.1的嗜脲清野氏菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:14的氨基酸序列,或
- [0206] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0207] 易位多肽包含:
- [0208] a) 来自GenBank登录号WP_189053160.1的吐鲁番长丝菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:15的氨基酸序列,或
- [0209] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0210] 在一个实施例中,易位多肽包含:

[0211] a) 来自GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:16的氨基酸序列,或

[0212] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0213] 在一个实施例中,易位多肽包含:

[0214] a) 来自GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:17的氨基酸序列,或

[0215] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0216] 在一个实施例中,易位多肽包含:

[0217] a) 来自GenBank登录号EIZ2913133.1的产气克雷伯氏菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:36的氨基酸序列,或

[0218] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0219] 在一个实施例中,易位多肽包含:

[0220] a) 来自GenBank登录号MCH0551590.1的链霉菌属MUM 178J蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:37的氨基酸序列,或

[0221] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0222] 在一个实施例中,易位多肽包含:

[0223] a) 来自GenBank登录号MBB4677777.1的嗜冷克洛斯氏菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:38的氨基酸序列,或

[0224] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0225] 在一个实施例中,易位多肽包含:

[0226] a) 来自GenBank登录号WP_143261759.1的异库茨涅尔氏菌属NRRL B-24872蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:39的氨基酸序列,或

[0227] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0228] 在一个实施例中,易位多肽包含:

[0229] a) 来自GenBank登录号WP_156051914.1的白色异库茨涅尔氏菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:40的氨基酸序列,或

[0230] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0231] 在一个实施例中,易位多肽包含:

[0232] a) 来自GenBank登录号WP_199893204.1的链霉菌属AV19蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:41的氨基酸序列,或

[0233] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0234] 在一个实施例中,易位多肽包含:

[0235] a) 来自GenBank登录号WP_147264604.1的链霉菌属NRBC_110611蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:42的氨基酸序列,或

[0236] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。

[0237] 在一个实施例中,易位多肽包含:

[0238] a) 来自GenBank登录号WP_209513619.1的注射器链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:43的氨基酸序列,或

- [0239] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0240] 在一个实施例中,易位多肽包含:
- [0241] a) 来自GenBank登录号RJQ69589.1的假诺卡氏菌科细菌YIM PH 21723蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:44的氨基酸序列,或
- [0242] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0243] 在一个实施例中,易位多肽包含:
- [0244] a) 来自GenBank登录号WP_143218892.1的曼谷放线动孢菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:45的氨基酸序列,或
- [0245] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0246] 在一个实施例中,易位多肽包含:
- [0247] a) 来自GenBank登录号MBF6055834.1的优洛菌素链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:46的氨基酸序列,或
- [0248] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0249] 在一个实施例中,易位多肽包含:
- [0250] a) 来自GenBank登录号WP_169790908.1的灭病菌素链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:47的氨基酸序列,或
- [0251] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0252] 在一个实施例中,易位多肽包含:
- [0253] a) 来自GenBank登录号WP_157868472.1的卡廷加链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:48的氨基酸序列,或
- [0254] b) 与a)中定义的所述易位结构域具有至少80%同一性的易位结构域。
- [0255] GenBank登录号WP_162873017.1之龟奥斯特维克氏菌LK16-18蛋白质在本文中可以被称为“CT2”。GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质在本文中可以被称为“CT3”。
- [0256] 在一个实施例中,易位结构域如以上实施例中任一者中的b)中所定义,并且在其全长上与a)中定义的易位结构域具有至少85%同一性。在一个实施例中,易位结构域如以上实施例中任一者中的b)中所定义,并且在其全长上与a)中定义的易位结构域具有至少90%同一性。在一个实施例中,易位结构域如以上实施例中任一者中的b)中所定义,并且在其全长上与a)中定义的易位结构域具有至少95%同一性。在一个实施例中,易位结构域如以上实施例中任一者中的b)中所定义,并且在其全长上与a)中定义的易位结构域具有至少98%同一性。在一个实施例中,易位结构域如以上实施例中任一者中的b)中所定义,并且在其全长上与a)中定义的易位结构域具有至少99%同一性。
- [0257] 在一个实施例中,易位多肽如以上实施例中任一者中的a)中所定义。
- [0258] 在一个实施例中,易位多肽包含:
- [0259] a) 来自龟奥斯特维克氏菌蛋白质CT1的易位结构域,其中所述易位结构域具有SEQ ID NO:3的氨基酸序列。
- [0260] 在一个实施例中,易位多肽包含:
- [0261] a) 来自GenBank登录号PSJ28985.1的无淀粉酶链孢囊菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:4的氨基酸序列。

[0262] 在一个实施例中,易位多肽包含:

[0263] a) 来自GenBank登录号SDT83331.1的链霉菌属TLI 053蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:5的氨基酸序列。

[0264] 在一个实施例中,易位多肽包含:

[0265] a) 来自GenBank登录号WP_160159328.1的链霉菌属SLBN-118蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:6的氨基酸序列。

[0266] 在一个实施例中,易位多肽包含:

[0267] a) 来自GenBank登录号WP_168096531.1的链霉菌属AA8蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:7的氨基酸序列。

[0268] 在一个实施例中,易位多肽包含:

[0269] a) 来自GenBank登录号WP_078659863.1的玫瑰轮丝链霉菌蛋白质的易位结构域。

[0270] 在一个实施例中,易位多肽包含:

[0271] a) 来自GenBank登录号JZ58907.1的松土壤链霉菌蛋白质的易位结构域。

[0272] 在一个实施例中,易位多肽包含:

[0273] a) 来自GenBank登录号WP_079110321.1的链霉菌属MBT76蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:10的氨基酸序列。

[0274] 在一个实施例中,易位多肽包含:

[0275] a) 来自GenBank登录号WP_120757473.1的克伦基链霉菌蛋白质的易位结构域。

[0276] 在一个实施例中,易位多肽包含:

[0277] a) 来自GenBank登录号WP_095582082.1的白网链霉菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:12的氨基酸序列。

[0278] 在一个实施例中,易位多肽包含:

[0279] a) 来自GenBank登录号WP_133259917.1的松生弯曲嗜酸菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:13的氨基酸序列。

[0280] 在一个实施例中,易位多肽包含:

[0281] a) 来自GenBank登录号WP_073156187.1的嗜脲清野氏菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:14的氨基酸序列。

[0282] 易位多肽包含:

[0283] a) 来自GenBank登录号WP_189053160.1的吐鲁番长丝菌蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:15的氨基酸序列。

[0284] 在一个实施例中,易位多肽包含:

[0285] a) 来自GenBank登录号WP_219106995.1的奥斯特维克氏菌属TVS 96-490-7B蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:16的氨基酸序列。

[0286] 在一个实施例中,易位多肽包含:

[0287] a) 来自GenBank登录号WP_162873017.1的龟奥斯特维克氏菌LK16-18蛋白质的易位结构域,其中所述易位结构域具有SEQ ID NO:17的氨基酸序列。

[0288] 在一些实施例中,将具有SEQ ID NO:12(来自GenBank登录号WP_095582082.1的白网链霉菌蛋白质)和14(来自GenBank登录号WP_073156187.1的嗜脲清野氏菌蛋白质)的氨基酸序列的易位结构域从上述实施例中排除。

[0289] 在一些实施例中,易位多肽可以包含全长蛋白质的功能性截短体,其包含上述相应易位结构域中的任一者,其中易位结构域的功能得以维持。

[0290] 在一个实施例中,A和B通过接头分隔开。在一个实施例中,A和B通过氨基酸接头分隔开。在一个实施例中,氨基酸接头包含 $(G_4S)_2$ 。在一个实施例中,接头是可裂解的。在一个实施例中,接头包含蛋白酶识别位点。在一个实施例中,蛋白酶识别位点是弗林蛋白酶识别位点。在一个实施例中,蛋白酶识别位点被半胱氨酸残基包围以允许二硫键形成以形成分子内环。在一个实施例中,氨基酸接头包含被半胱氨酸残基包围的弗林蛋白酶识别位点。在一个示例性实施例中,氨基酸接头包含SEQ ID NO:32,其包含弗林蛋白酶识别位点和包围半胱氨酸残基。在一个实施例中,所述 $(G_4S)_2$ 相对于所述SEQ ID NO:32位于N端。在一个实施例中,接头是自裂解的。在一个实施例中,接头是自清除的。在一个实施例中,接头包含自动处理结构域。

[0291] 在一个实施例中,B和C通过接头分隔开。在一个实施例中,B和C通过氨基酸接头分隔开。在一个实施例中,氨基酸接头包含 $(G_4S)_2$ 。在一个实施例中,氨基酸接头包含 $(G_4S)_2$ 。在一个实施例中,氨基酸接头包含SEQ ID NO:33。在一个实施例中,其中所述 $(G_4S)_2$ 相对于所述SEQ ID NO:33位于N端。

[0292] 在一个实施例中,重组多肽包含与SEQ ID NO:22的氨基酸1至821具有至少80%同一性、优选与SEQ ID NO:22的氨基酸1至821具有90%同一性、更优选与SEQ ID NO:22的氨基酸1至821具有95%同一性、甚至更优选与SEQ ID NO:22的氨基酸1至821具有100%同一性的氨基酸序列。

[0293] 在一个实施例中,重组多肽包含SEQ ID NO:22的序列。

[0294] 在一个实施例中,重组多肽包含与SEQ ID NO:23的氨基酸1至822具有至少80%同一性、优选与SEQ ID NO:23的氨基酸1至822具有90%同一性、更优选与SEQ ID NO:23的氨基酸1至822具有95%同一性、甚至更优选与SEQ ID NO:23的氨基酸1至822具有100%同一性的氨基酸序列。

[0295] 在一个实施例中,重组多肽包含SEQ ID NO:23的序列。

[0296] 在一个实施例中,重组多肽包含与SEQ ID NO:24的氨基酸1至822具有至少80%同一性、优选与SEQ ID NO:24的氨基酸1至822具有90%同一性、更优选与SEQ ID NO:24的氨基酸1至822具有95%同一性、甚至优选与SEQ ID NO:24的氨基酸1至822具有100%同一性的氨基酸序列。

[0297] 在一个实施例中,重组多肽包含SEQ ID NO:24的序列。

[0298] 在一个实施例中,重组多肽包含与SEQ ID NO:25的氨基酸1至821具有至少80%同一性、优选与SEQ ID NO:25的氨基酸1至821具有90%同一性、更优选与SEQ ID NO:25的氨基酸1至821具有95%同一性、甚至更优选与SEQ ID NO:25的氨基酸1至821具有100%同一性的氨基酸序列。

[0299] 在一个实施例中,重组多肽包含SEQ ID NO:25的序列。

[0300] 在一个实施例中,重组多肽包含与SEQ ID NO:26的氨基酸1至806具有至少80%同一性、优选与SEQ ID NO:26的氨基酸1至806具有90%同一性、更优选与SEQ ID NO:26的氨基酸1至806具有95%同一性、甚至更优选与SEQ ID NO:26的氨基酸1至806具有100%同一性的氨基酸序列。

[0301] 在一个实施例中,重组多肽包含SEQ ID NO:26的序列。

[0302] 在一个实施例中,重组多肽包含与SEQ ID NO:27的氨基酸1至811具有至少80%同一性、优选与SEQ ID NO:27的氨基酸1至811具有90%同一性、更优选与SEQ ID NO:27的氨基酸1至811具有95%同一性、甚至更优选与SEQ ID NO:27的氨基酸1至811具有100%同一性的氨基酸序列。

[0303] 在一个实施例中,重组多肽包含SEQ ID NO:27的序列。

[0304] 在一个实施例中,重组多肽包含与SEQ ID NO:28的氨基酸1至822具有至少80%同一性、优选与SEQ ID NO:28的氨基酸1至822具有90%同一性、更优选与SEQ ID NO:28的氨基酸1至822具有95%同一性、甚至更优选与SEQ ID NO:28的氨基酸1至822具有100%同一性的氨基酸序列。

[0305] 在一个实施例中,重组多肽包含SEQ ID NO:28的序列。

[0306] 在一个实施例中,重组多肽包含与SEQ ID NO:29的氨基酸1至796具有至少80%同一性、优选与SEQ ID NO:29的氨基酸1至796具有90%同一性、更优选与SEQ ID NO:29的氨基酸1至796具有95%同一性、甚至更优选与SEQ ID NO:29的氨基酸1至796具有100%同一性的氨基酸序列。

[0307] 在一个实施例中,重组多肽包含SEQ ID NO:29的序列。

[0308] 在一个实施例中,重组多肽包含与SEQ ID NO:30的氨基酸1至822具有至少80%同一性、优选与SEQ ID NO:30的氨基酸1至822具有90%同一性、更优选与SEQ ID NO:30的氨基酸1至822具有95%同一性、甚至更优选与SEQ ID NO:30的氨基酸1至822具有100%同一性的氨基酸序列。

[0309] 在一个实施例中,重组多肽包含SEQ ID NO:30的序列。

[0310] 在一个实施例中,靶向部分包含靶向多肽或适体。

[0311] 在一个实施例中,靶向多肽包含抗体、抗体的结合片段、亲和体、亲和素、DARPin或受体配体。

[0312] 在一个实施例中,靶向多肽包含针对Her3的亲和体。在一个实施例中,针对Her3的亲和体包含SEQ ID NO:19的氨基酸序列。

[0313] 在一个实施例中,靶向多肽包含 $\alpha v\beta 6$ 整合素的受体配体。在一个实施例中, $\alpha v\beta 6$ 整合素的受体配体包含SEQ ID NO:20的氨基酸序列。

[0314] 在一个实施例中,靶向部分包含至少两种靶向多肽、至少两种适体或靶向多肽与适体的组合。在一个实施例中,至少两种靶向多肽选自以下的群组:抗体、抗体的结合片段、亲和体、肽、亲和素、DARPin、受体配体及其组合。在一个实施例中,至少两种靶向多肽包含针对Her3的亲和体。在一个实施例中,针对Her3的亲和体包含SEQ ID NO:19的氨基酸序列。在一个实施例中,至少两种靶向多肽包含 $\alpha v\beta 6$ 整合素的受体配体。在一个实施例中, $\alpha v\beta 6$ 整合素的受体配体包含SEQ ID NO:20的氨基酸序列。在一个实施例中,至少两种靶向多肽包含针对Her3的亲和体和 $\alpha v\beta 6$ 整合素的受体配体两者,优选地其中前者包含SEQ ID NO:19的氨基酸序列并且后者包含SEQ ID NO:20的氨基酸序列。在一个实施例中,至少两种靶向多肽、至少两种适体或组合通过氨基酸接头分隔开。在一个实施例中,氨基酸接头包含 $(G_4S)_2$ 。

[0315] 在一个实施例中,靶向部分结合到细胞表面蛋白质。

[0316] 在一个实施例中,细胞表面蛋白质是谱系特异性或组织特异性的。

- [0317] 在一个实施例中,细胞表面蛋白质是普遍表达的。
- [0318] 在一个实施例中,细胞表面蛋白质在疾病细胞中表达。
- [0319] 在一个实施例中,细胞表面蛋白质对于疾病细胞是特异性的并且不在相应健康细胞中表达。
- [0320] 在一个实施例中,与相应健康细胞相比,细胞表面蛋白质在疾病细胞中具有升高的表达。
- [0321] 在一个实施例中,疾病细胞是癌细胞。
- [0322] 在一个实施例中,运载物分子包含运载物多肽。
- [0323] 运载物多肽可以包含需要细胞递送的任何多肽。运载物多肽可以包含酶或其具有大体上相同活性的活性片段。“大体上相同活性”意思指酶的核心功能在片段中大体上未改变。
- [0324] 运载物多肽可以具有小于10kDa、大于10kDa、大于20kDa、大于30kDa、大于50kDa、大于100kDa或大于150kDa的分子量。
- [0325] 运载物多肽包含基因组修饰蛋白质。基因组修饰蛋白质包含锌指核酸酶(ZFN)、转录激活因子样效应物核酸酶(TALEN)或CRISPR(成簇规律间隔短回文重复序列)蛋白质。CRISPR蛋白质可以是Cas9。运载物多肽可以包含基因组修饰蛋白质与核酸(如指导核酸)的复合物。举例来说,Cas9可以与核酸(如指导RNA),如crRNA、trRNA和/或sgRNA复合。
- [0326] 在一个实施例中,运载物分子包含治疗性多肽。
- [0327] “治疗性多肽”意思指细胞递送可以用于治疗目的的任何蛋白质。举例来说,众所周知,许多人类疾病或病症由蛋白质缺乏引起或以蛋白质缺乏为特征。治疗性蛋白质涵盖递送可以改善或校正这种缺乏的蛋白质。治疗性蛋白质可以用于替代疾病或病症中缺乏的蛋白质。治疗性蛋白质可以是疾病或病症中缺乏的蛋白质。然而,治疗性蛋白质不一定需要与疾病或病症中缺乏的蛋白质相同。举例来说,治疗性蛋白质可以是缺乏蛋白质的活性片段或修饰形式。治疗性蛋白质还可以部分或完全地在功能上补偿疾病或病症背后的蛋白质缺乏。治疗性蛋白质还可以改善或校正细胞缺乏特定蛋白质的下游或次级效应。
- [0328] 在一个实施例中,治疗性多肽包含细胞毒性多肽,优选多肽毒素或其功能片段。在一个实施例中,细胞毒性多肽包含来自白喉毒素的催化结构域。在一个实施例中,细胞毒性多肽包含来自龟毒素,如来自本文所描述的CT1(SEQ ID NO:2)、CT2(SEQ ID NO:21)或CT3(SEQ ID NO:35)的催化结构域。
- [0329] 在实施例中,催化结构域来自CT1。在一个实施例中,来自SEQ ID NO:2中所阐述的龟毒素(CT1)的催化结构域具有根据SEQ ID NO:2的氨基酸位置1至186的氨基酸序列。
- [0330] 在实施例中,催化结构域来自CT2。在一个实施例中,来自SEQ ID NO:21中所阐述的龟毒素(CT2)的催化结构域具有根据SEQ ID NO:21的氨基酸位置1至186的氨基酸序列。
- [0331] 在实施例中,催化结构域来自CT3。在一个实施例中,来自SEQ ID NO:35中所阐述的龟毒素(CT3)的催化结构域具有根据SEQ ID NO:35的氨基酸位置1至191的氨基酸序列。
- [0332] 在一个实施例中,治疗性多肽包含疾病状态中缺乏的蛋白质或其功能片段。
- [0333] 在一个实施例中,治疗性多肽包含来自创伤弧菌(*Vibrio vulnificus*)的Ras/Rap1特异性内肽酶(RRSP),例如SEQ ID NO:18中所阐述。在一些实施例中,治疗性多肽可以与RRSP具有至少80%同一性。在一些实施例中,治疗性多肽可以与RRSP具有至少90%同一

性。在一些实施例中,治疗性多肽可以与RRSP具有至少95%同一性。在一些实施例中,治疗性多肽可以与RRSP具有至少98%同一性。这些序列变体可以保留与全长RRSP大体上相同的活性。

[0334] 在一个实施例中,运载物分子包含用于“点击”化学生物缀合的N端半胱氨酸残基。

[0335] 在一个实施例中,运载物分子包含核酸分子。

[0336] 可以在比对的全长上计算本文所描述的序列鉴别百分比。

[0337] 在一些实施例中,本文所提及的氨基酸序列可以涵盖与参考序列(如下表1中所阐述的那些)相比的序列差异。这些可以是变体、突变、插入或缺失。在一些应用中,确保蛋白质的主要功能大体上没有改变或消除可能很重要,但这可以容易地测试,例如使用本文所描述的分析。本文所描述的氨基酸序列可以包含与参考序列具有75%或更高、80%或更高、85%或更高、90%或更高、95%或更高或99%或更高同一性的序列。氨基酸序列可以涵盖保守氨基酸取代。所属领域中已知的保守氨基酸取代如下,其中保守可取代候选氨基酸展示在括号中:Ala (Gly、Ser);Arg (Gly、Gln);Asn (Gln;His);Asp (Glu);Cys (Ser);Gln (Asn、Lys);Glu (Asp);Gly (Ala、Pro);His (Asn;Gln);Ile (Leu;Val);Leu (Ile;Val);Lys (Arg;Gln);Met (Leu、Ile);Phe (Met、Leu、Tyr);Ser (Thr;Gly);Thr (Ser;Val);Trp (Tyr);Tyr (Trp;Phe);Val (Ile;Leu)。一些所谓的“功能性”变体、突变、插入或缺失涵盖功能与参考序列的功能大体上相同的序列,例如其来源于参考序列。这可以使用与本文所描述的分析类似的分析容易地测试。

[0338] 核酸和载体

[0339] 在一个方面中,提供编码如本文所定义的重组多肽的核酸。在一个实施例中,核酸是DNA或RNA。RNA可以是mRNA。

[0340] 技术人员将容易了解,例如归因于遗传密码的简并性,存在许多编码本文所描述的重组多肽的方式,所有这些方式都涵盖在某些实施例中。如果蛋白质功能保持大体上完整,那么也可以允许缺失、插入和取代。举例来说,可以涵盖核酸可以与野生型或参考序列具有75%或更高、80%或更高、85%或更高、90%或更高、95%或更高或99%或更高的同一性。上文提到的核酸还可以视其预期表达于的生物体或表达系统而定进行密码子优化。

[0341] 在一个方面中,提供包含如本文所定义的核酸的载体。

[0342] 在一个实施例中,提供包含如本文所定义的核酸或如本文所定义的载体的宿主细胞。宿主细胞可以被转化或转染。

[0343] 组合物

[0344] 在一个方面中,提供一种组合物,其包含如本文所定义的重组多肽以及可接受的赋形剂、稀释剂或载剂。

[0345] 在方面实施例中,提供一种药物组合物,其包含如本文所定义的重组多肽以及药学上可接受的赋形剂、稀释剂或载剂。

[0346] 药学上可接受的载剂包括溶剂、稀释剂、液体媒剂、分散或悬浮助剂、表面活性剂、等渗剂、增稠或乳化剂、防腐剂、固体粘合剂或润滑剂。可以选择载剂以延长适合于所选投药途径的持续释放的停留时间。示例性载剂包括糖(如葡萄糖和蔗糖)、淀粉(如玉米淀粉和马铃薯淀粉)、纤维(如纤维素及其衍生物、羧甲基纤维素钠、乙基纤维素、乙酸纤维素)、粉末状黄芪胶、麦芽、明胶、滑石、可可脂、栓剂蜡、油(如花生油、棉籽油、红花油、芝麻油、橄榄

油、玉米油和大豆油)；二醇(如丙二醇)、酯(如油酸乙酯和月桂酸乙酯)、琼脂、缓冲剂(如氢氧化镁和氢氧化铝)、海藻酸、无热原水、等渗盐水、林格氏溶液(Ringer's solution)、乙醇、磷酸盐缓冲溶液、无毒相容润滑剂(如月桂基硫酸钠和硬脂酸镁)、着色剂、脱模剂、包衣剂、甜味剂、调味剂、芳香剂、防腐剂以及抗氧化剂。

[0347] 组合物可以通过任何可接受的途径投与受试者,如局部(如通过粉末、软膏或滴剂)、口服、直肠、粘膜、舌下、肠胃外、脑池内、阴道内、腹膜内、经颊、经眼或鼻内。

[0348] 用于口服投药的液体剂型可以包括乳液、微乳液、溶液、悬浮液、糖浆和酏剂。液体剂型可以含有惰性稀释剂(如水或其它溶剂)、增溶剂和乳化剂(如乙醇、异丙醇、碳酸乙酯、乙酸乙酯、苯甲醇、苯甲酸苯甲酯、丙二醇、1,3-丁二醇、二甲基甲酰胺)、油(如棉籽油、花生油、玉米油、胚芽油、橄榄油、蓖麻油和芝麻油)、甘油、四氢糠醇、聚乙二醇和脱水山梨糖醇的脂肪酸酯以及其混合物。除惰性稀释剂之外,口服组合物还可以包括佐剂,如湿润剂、乳化剂和悬浮剂、甜味剂、调味剂和芳香剂。

[0349] 用于局部或经皮投与本发明药物组合物的剂型包括软膏、糊剂、乳膏、洗剂、凝胶、粉末、溶液、喷雾剂、吸入剂或贴剂。在无菌条件下将活性剂与药学上可接受的载剂和任何所需的防腐剂或缓冲剂按要求混合。

[0350] 可以根据已知技术使用适合的分散剂或湿润剂和悬浮剂来调配可注射制剂,如无菌可注射水性或油性悬浮液。无菌可注射制剂也可以是在无毒肠胃外可接受的稀释剂或溶剂中的无菌可注射溶液、悬浮液或乳液,例如呈在1,3-丁二醇中的溶液形式。可以采用的可接受的媒剂和溶剂中有水、林格氏溶液、U.S.P.和等渗氯化钠溶液。另外,无菌不挥发性油常规地用作溶剂或悬浮介质。出于此目的,可以采用任何温和不挥发性油,包括合成单甘油酯或二甘油酯。另外,在制备可注射剂时使用脂肪酸(如油酸)。可注射调配物可以在添加孢子之前例如通过经细菌截留过滤器过滤或通过并入灭菌剂来灭菌,所述灭菌剂呈可以在使用前溶解或分散于无菌水或其它无菌可注射介质中的无菌固体组合物形式。

[0351] 通常需要减缓来自皮下或肌肉内注射的药剂的吸收。肠胃外投与的活性剂的延迟吸收可以通过将药剂溶解或悬浮在油媒剂中来实现。通过形成药剂于生物可降解聚合物(如聚丙交酯-聚乙交酯)中的微胶囊基质来制造可注射积存形式。取决于活性剂与聚合物的比率和所采用的特定聚合物的性质,可以控制活性剂的释放速率。其它生物可降解聚合物的实例包括聚(原酸酯)和聚(酸酐)。还通过将药剂包覆于与身体组织相容的脂质体或微乳液中来制备积存可注射调配物。

[0352] 用于直肠或阴道投药的组合物优选是栓剂,所述栓剂可以通过将本发明的活性剂与适合的无刺激性赋形剂或载剂(如可可脂、聚乙二醇或栓剂蜡)混合而制备,所述赋形剂或载剂在环境温度下是固体,但在体温下是液体并且因此在直肠或阴道腔中熔融并释放活性剂。

[0353] 用于口服、粘膜或舌下投药的固体剂型包括胶囊、片剂、丸剂、粉末和颗粒。在此类固体剂型中,活性剂与以下各者混合:至少一种药学上可接受的惰性赋形剂或载剂,如柠檬酸钠或磷酸二钙;填充剂或增量剂,如淀粉、蔗糖、葡萄糖、甘露糖醇和硅酸;粘合剂,例如羧甲基纤维素、海藻酸盐、明胶、聚乙烯吡咯烷酮、蔗糖和阿拉伯胶;保湿剂,如甘油;崩解剂,如琼脂-琼脂、碳酸钙、马铃薯或木薯淀粉、海藻酸、某些硅酸盐和碳酸钠;溶解阻滞剂,如石蜡;吸收促进剂,如季铵化合物;湿润剂,例如鲸蜡醇和单硬脂酸甘油酯;吸收剂,如高岭土

和膨润土；以及润滑剂，如滑石、硬脂酸钙、硬脂酸镁、固体聚乙二醇、月桂基硫酸钠及其混合物。

[0354] 也可以使用类似类型的固体组合物作为使用如乳糖以及高分子量聚乙二醇等赋形剂的软和硬填充明胶胶囊中的填充剂。片剂、胶囊、丸剂以及颗粒的固体剂型可以用包衣和外壳(如肠溶衣、释放控制包衣和药物调配领域中熟知的其他包衣)来制备。在此类固体剂型中,活性剂可以与至少一种惰性稀释剂(如蔗糖或淀粉)混合。正常实践时,此类剂型还可以包含除惰性稀释剂以外的额外物质,如压片润滑剂和其它压片助剂,如硬脂酸镁和微晶纤维素。在胶囊、片剂和丸剂的情况下,剂型还可以包含缓冲剂。其可以任选地含有乳浊剂,并且还可以具有使其任选地在肠道某一部分中以延迟方式仅或优先释放活性剂的组成。可以使用的包埋组合物的实例包括聚合物物质和蜡。

[0355] 治疗有效量可以基于个体或基于确定的必要量来确定。鉴于待治疗的受试者选择个体受试者的剂量。可以调整剂量和投药以提供足够水平的活性剂或维持所需效果。可以考虑的因素包括疾病状态的严重程度、过去与传染原的接触、潜在的未来接触;受试者的年龄、体重和性别、饮食、投药时间和频率、药物组合、反应敏感性以及对疗法的耐受性/反应。持续释放组合物可以比速效组合物更不频繁地投与。

[0356] 方法和用途

[0357] 在一个方面中,提供一种将运载物分子递送到细胞的方法,其包含使所述细胞与如本文所定义的重组多肽接触。

[0358] 在一个方面中,提供如本文所定义的重组多肽用于将运载物分子递送到细胞的用途。

[0359] 在一个方面中,提供如本文所定义的重组多肽用于制备用于将运载物分子递送到细胞的药剂的用途。

[0360] 在一个方面中,提供如本文所定义的重组多肽,其用于将运载物分子递送到细胞。

[0361] 在一个方面中,提供一种治疗受试者的癌症的方法,其包含向所述受试者投与如本文所定义的重组多肽。

[0362] 在一个方面中,提供如本文所定义的重组多肽用于治疗受试者的癌症的用途。

[0363] 在一个方面中,提供如本文所定义的重组多肽用于制备用于治疗受试者的癌症的药剂的用途。

[0364] 在一个方面中,提供如本文所定义的重组多肽,其用于治疗受试者的癌症。

[0365] 在一个方面中,提供一种减轻细胞中的酶或蛋白质缺乏的方法,其包含使细胞如本文所描述的重组多肽接触。

[0366] 在一个方面中,提供如本文所描述的重组多肽用于减轻细胞中的酶或蛋白质缺乏的用途。

[0367] 在一个方面中,提供如本文所定义的重组多肽用于制备用于减轻细胞中的酶或蛋白质缺乏的药剂的用途。

[0368] 在一个方面中,提供如本文所描述的重组多肽,其用于减轻细胞中的酶或蛋白质缺乏。

[0369] 如本文所用,“减轻”意思指运载物分子校正或至少部分改善蛋白质或酶缺乏、缺乏蛋白质或酶的功能的一个方面或其下游或次级细胞效应或后果中的一或多者。

[0370] 在所述方法和用途的实施例中,可以释放运载物。

[0371] 实例

[0372] 实例1

[0373] 引言

[0374] 工程化的嵌合毒素已引起针对具有挑战性的疾病(如癌症)的新颖治疗剂的出现。免疫毒素是一类由细菌毒素(如白喉毒素(DT))构成的生物治疗药物,所述细菌毒素已通过重新靶向其受体结合结构域(RBD)以靶向癌症受体并且通过递送靶向细胞内肿瘤蛋白的酶运载物而改换用途用于癌症靶向疗法。然而,针对白喉的全球疫苗接种计划已经引起针对DT和基于DT的治疗剂的群体水平免疫。为了规避预先存在的针对DT的中和抗体的问题,研究了具有很低序列同一性的DT的远缘同源物是否可以保留DT的功能但避免DT特异性中和抗体。此处,已在结构和功能上表征了与DT仅具有38%同一性的来自龟奥斯特维克氏菌的推定基因序列。其被命名为龟毒素1(CT1)。其x射线晶体学结构已解析到2.50 Å,并且发现其结构与DT高度类似。使用多种生物化学分析,进行逐结构域分析以研究此DT样蛋白质充当毒素和用于治疗性蛋白质递送的新平台的能力。已证明,虽然这种新颖蛋白质的每个结构域可以作为毒素发挥其相应功能,但CT1的易位酶(CT1-T)可以被工程化以靶向非天然受体并且将非天然运载物递送到细胞中。重要的是,CT1不会被人血清中发现的预先存在的抗DT抗体识别,并且在将运载物递送到细胞中方面出乎意料地优于DT。龟毒素为毒素生物学提供了新颖见解,并且代表用于治疗性蛋白质递送的改进平台。

[0375] 材料和方法

[0376] 来自龟奥斯特维克氏菌的DT样蛋白质的结晶。

[0377] 归因于被鉴别为对DT功能性重要的关键残基的保守性(图5),来自物种龟奥斯特维克氏菌的最接近的DT样蛋白质(本文中称为CT1,表示“龟毒素1”)(SEQ ID NO:2)被选择作为替代免疫毒素支架的候选物。

[0378] SEQ ID NO:2来源于代表毒素的两个片段的两个ORF(参见GenBank登录号WP_143115263.1和WP_040322835.1)的组合。当相比于龟奥斯特维克氏菌的基因组序列(参见GenBank登录号NZ_BAGZ01000024.1)时,似乎基因组序列中的1个碱基对(bp)移码已使得全长毒素分离成2个ORF。通过删除1bp(NZ_BAGZ01000024.1C41398)来恢复阅读框,并且结果是全长毒素,其随后称为“龟毒素1(CT1)”(SEQ ID NO:2)。尚不清楚1bp插入是序列错误还是反映了龟奥斯特维克氏菌中的真正突变。在任何情况下,去除1bp插入以产生用于本文所描述的实验的蛋白质和易位结构域。

[0379] 为了确定CT1的结构,从整合DNA技术(Integrated DNA Technologies)订购大肠杆菌(E.coli)密码子优化的gBlock基因片段,并且通过吉布森组装(Gibson Assembly)克隆到Champion™pET SUMO大肠杆菌表达系统中。

[0380] 将50mL NiCo21(DE3)大肠杆菌细胞(新英格兰生物实验室(New England Biolabs))的起子培养物接种到1L LB培养基中,并且用0.1mM IPTG在18°C下诱导18小时。将细胞以5000rpm离心并且再悬浮于裂解缓冲液(1%蛋白酶抑制剂混合物、1mg/mL溶菌酶、0.01% Pierce™通用核酸酶抑制剂、20mM咪唑、500mM NaCl、20mM Tris-HCl pH7.5)中。在15000psi下通过Emulsiflex C3(奥维斯丁(Avestin))三次来裂解细胞。将全细胞裂解物以18000×g离心,并且将上清液通过0.45μm过滤器过滤并通过HisTrap FF粗柱(思拓凡

(Cytiva))。将蛋白质用50-75mM咪唑洗脱,缓冲液交换到150mM NaCl、20mM Tris-HCl pH 7.5中,并且与SUMO蛋白酶一起在4°C下培育过夜,以裂解6×His-SUMO亲和标签。使蛋白质流过HisTrap FF粗柱,并且收集流出物(蛋白质)并通过离心浓缩到8mg/mL。

[0381] 使用悬滴蒸气扩散来生长晶体。获得CT1晶体的条件含有2μL母液(0.2M氯化钙、0.1M Tris-HCl pH 8.5、25% (w/v) PEG4000)和1μL 8mg/mL蛋白质。将液滴经130μL 2M (NH₄)₂PO₄脱水45分钟,随后在液氮中冷冻。在先进光子源(Advanced Photon Source)的23-ID-D光束线上收集数据。

[0382] 使用Phenix软件包中的Phaser,通过使用具有无序环已去除的个别DT结构域(C结构域残基13-167、R结构域残基391-531、T结构域残基205-378)的多组分搜索模型来确定初始相位。使用phenix.refine和autobuild的迭代循环来细化结构。

[0383] 蛋白质合成分析

[0384] 将Vero-nLucP细胞(非洲绿猴肾细胞的纳米荧光素酶报告株系)以5000个细胞/孔涂铺在96孔白色透明底板(康宁(Corning))中。第二天,添加蛋白质毒素并且培育24小时,之后在SpectraMax M5e读板仪(美谷分子仪器(Molecular Devices))上使用NanoGlo®荧光素酶分析试剂盒(普洛麦格(Promega))读取细胞的发光信号。相对于未处理的细胞(100%纳米荧光素酶信号)校正数据。

[0385] 脂质体染料释放分析

[0386] 如先前所描述制备单层脂质体(DOPC, 0.8% DGS-NTA, 阿凡提极性脂质(Avanti Polar Lipids))。简单来说,将1,2-二油酰基-sn-甘油-3-磷酸胆碱(DOPC)(阿凡提极性脂质)与0.8%1,2-二油酰基-sn-甘油-3-[N-(5-氨基-1-羧基戊基)亚氨基二乙酸]琥珀酰基(镍盐)(DGS-NTA[Ni])(阿凡提极性脂质)合并,在真空干燥器中用N₂干燥1小时。将脂质再悬浮于20mM Tris pH 8、35mM 8-羟基芘-1,3,6-三磺酸(HPTS)、50mM对二甲苯-双吡啶鎓溴化物(DPX)(赛默飞世尔(Thermo Fischer))中,并且在干冰和42°C水浴中进行10次冻融循环,并且使用200μm过滤器进行10次挤出。随后使用Hi Prep 16/60Sephacryl S-300HR柱(通用电气医疗集团(GE Healthcare))和150mM NaCl、20mM Tris pH 8缓冲液通过凝胶过滤来纯化脂质体。蛋白质与脂质体以1:10,000的比率添加,其中最终脂质体浓度为约400μM,处于pH范围为4.0至7.5的150mM柠檬酸盐磷酸盐缓冲液中,pH间隔为0.5。在96孔不透明板(康宁)中进行分析,并且以20分钟的间隔监测荧光,每30秒获取一次读数(激发403nm,发射510nm)。通过添加0.3% Triton X-100,将数据归一化为总HPTS荧光的%。

[0387] 血清抗体结合ELISA

[0388] 用1% BSA封闭之后,将Nunc MaxiSorp™板(赛默飞世尔科技(Thermo Fisher Scientific))用2000ng蛋白质固定,并且随后与各种稀释度的人血清(冷冻合并人血清,赛达兰(Cedarlane))一起培育1小时。将孔用PBST(0.01%吐温)洗涤,并且随后与缀合到HRP的抗人IgG抗体(艾博抗(Abcam), ab102420)一起培育,所述抗体使用TMB试剂(赛默飞世尔科技)显色。在630nm下读取吸光度,并且将蛋白质孔相对于对照孔(无蛋白质,+人血清)进行校正。

[0389] 血清毒性分析

[0390] 将蛋白质毒素与人血清(冷冻合并人血清,赛达兰)或小鼠血清(无菌冷冻小鼠血清,赛达兰)以1:1比率在室温下一起培育30分钟。随后将样品添加到前一天在96孔白色透

明底板(康宁)中涂铺到5000个细胞/孔的Vero-nLucP细胞中。将细胞培育24小时,然后裂解细胞并且评估发光信号。将值相对于仅血清处理的细胞校正,所述细胞代表100%纳米荧光素酶信号和100%蛋白质合成。

[0391] 结果

[0392] 来自龟奥斯特维克氏菌的DT样蛋白质的结构表征

[0393] 为了找到对预先存在抗DT抗体不敏感但具有功能活性的替代DT样免疫毒素平台,选择了棒状杆菌属(*Corynebacterium* genus)之外进化上最接近的DT样蛋白质。从序列上看,来自龟奥斯特维克氏菌的DT样蛋白质(CT1)与DT具有38%同一性。催化活性残基(E148)和对底物(NAD)结合和配位重要的残基(H21、Y54、Y65)都是保守的,弗林蛋白酶识别位点也是如此,并且C186与C201之间形成的二硫键(在DT中)也存在于CT1序列中(表1)。此外,发现参与易位结构域的pH依赖性展开和孔形成的关键组氨酸残基和带电残基在来自CT1的序列的预测易位结构域中保守(表1)。为了测试CT1的推定结构域是否具有功能,将蛋白质克隆、表达并纯化以在结构和功能上进行研究(图1B)。

[0394] 使用悬滴蒸气扩散获得CT1的x射线晶体学结构。蛋白质成功结晶并且衍射到2.50 Å。不可能使用全长DT结构(pdb 1MDT)作为分子置换的搜索模型。然而,使用具有1MDT的部分搜索模型,解析了结构(图1C)。CT1与DT的RMSD为2.4 Å,并且其保留了与DT相同的三结构域、Y形架构。

[0395] CT1的C结构域具有功能并且具有与DT相同的细胞内靶标

[0396] CT1-C(CT1的催化结构域)与DT_C(DT催化结构域)的结构比对展示DT_C功能所需的关键残基之间的良好结构保守性。为了测试CT1-C是否具有功能,产生了用CT1-C置换DT_C的嵌合体(称为CT1-C-DT_T-DT_R)。在具有和不具有DPH4基因敲除(DPH4^{-/-})的HEK293T细胞上测试嵌合蛋白。这些细胞在白喉酰胺合成路径中有缺陷,并且产生没有白喉酰胺修饰的真核延伸因子2(eEF-2),并且因此对DT具有完全抗性(图2A和2B)。野生型HEK293T细胞对野生型DT和CT1-C-DT_T-DT_R均敏感,而白喉酰胺基因敲除细胞则不敏感。这表明CT1的催化结构域具有与DT相同的细胞内靶标。

[0397] CT1的C结构域有效地释放

[0398] DT中毒机制的一个重要部分是在进入胞质溶胶后从分子的其余部分释放C结构域。弗林蛋白酶识别位点在CT1中保守(CT1中的RAKR)。为了确认弗林蛋白酶位点被识别并且C结构域从分子的其余部分释放,将DT和CT1与哺乳动物细胞裂解物一起在37°C下培育过夜。DT和CT1均在C结构域与T结构域之间裂解(图2C)。当在非洲绿猴肾细胞上测试DTC-CT1-F-DTT-DTR时,其与DT具有类似的毒性(图3A),表明其被弗林蛋白酶有效裂解。

[0399] CT1含有功能性易位酶

[0400] 内体酸化引起DT_T重折叠并插入内体膜中,以及随后C结构域易位到胞质溶胶中。此过程被认为是由DT_T中的九个带电残基引发和驱动的,九个带电残基中的六个在CT1-T中保守。为了测试CT1-T是否形成DT样孔,纯化分离的T结构域并且在体外测试其从脂质体释放染料的能力(数据未展示)。CT1-T展示染料释放的pH依赖性增加,其中染料释放(解释为孔形成)在pH 5.5下开始(类似于DT_T)。

[0401] 进一步研究易位酶是否将运载物递送到细胞中。为此目的,产生嵌合DT,其中DT_T被交换为CT1-T(DT_C-CT1-T-DT_R),并且测量其对蛋白质合成的影响。CT1-T成功地将DT_C递送

到细胞中,如随着嵌合毒素的量增加观察到的信号减少所示(图3B)。此外,当将DT_C交换为靶向Ras肿瘤蛋白的非天然运载物(Ras/Rap1特异性蛋白酶;RRSP)时,其被证明成功地递送这种结构上不同的运载物(如CFPAC-1细胞上的细胞活力和评估细胞内Ras水平的蛋白质印迹所示;图3C和3D)。综合起来,这些实验证明CT1-T是能够将各种运载物递送到细胞的胞质组分中的功能性易位酶。

[0402] CT1的易位结构域可以适应不同的受体结合结构域

[0403] 在证明了CT1-T可以耐受N端上的操纵(可以使DT_C和RRSP易位)之后,接下来评估CT1-T是否可以耐受C端上的此类操纵。为此目的,产生嵌合体,其中DT_C-CT1-T-DT_R和RRSP-CT1-T-DT_R的受体结合结构域(DT_R)用靶向Her3(人表皮受体3)的亲受体(ZHer3:08699,下文中称为ZHer3)交换,并且在HPAF II细胞上测试这些构建体(图3E和3F)。发现不仅CT1-T可以被重新靶向到表达Her3的细胞,而且出人意料地,含CT1-T的构建体在运载物递送方面似乎比DT_T更有效。

[0404] CT1不被人血清识别或中和

[0405] 为了定量人血清中预先存在的抗DT或抗CT1抗体的水平,使用ELISA分析,其中DT或CT1固定在高结合板上并且与不同量的合并的人血清一起培育,之后使用缀合到HRP的抗IgG抗体来确定抗体结合水平(图4A)。DT在人血清存在下展示信号的剂量依赖性增加,但CT1却没有,表明CT1未被人血清中的抗体识别。

[0406] 为了确认DT被人血清中的抗体中和,将DT与合并的人血清一起培育30分钟,并且随后将其添加到非洲绿猴肾细胞中。观察到DT毒性的6个对数的变化,表明被人血清中和(图4B)。相比之下,CT1未被人血清中和。这表明CT1不被抗DT抗体中和。通过规避预先存在的中和抗DT抗体的问题,这为基于CT1的递送载体提供了优于基于DT的载体的优点。

[0407] 除了CT1的易位结构域(SEQ ID NO:3)之外,SEQ ID NO:12和14的易位结构域也被确定为功能性的。

[0408] 基于获得的结果,来自本文所描述的其他细菌菌株和物种并且如SEQ ID NO:4至11、13、15至17和36至48(及其相关序列)中所阐述的相关多肽序列也预期是在如本文所描述的多肽构建体内具有活性的功能性易位结构域。图9提供展示易位酶结构域之间的关系的发生树。

[0409] 实例2

[0410] 为了评估新颖易位酶相对于DT的功能,将每个易位酶序列“T”在细胞内RAS裂解酶RRSP(Ras/Rap1特异性肽酶,SEQ ID NO:18;即运载物)与称为ZHer3-A20的双受体结合结构域(由针对Her3的亲受体SEQ ID NO:19和称为A20FMDV2(A20)的针对 $\alpha v \beta 6$ 整合素的肽SEQ ID NO:20组成)之间克隆,产生构建体RRSP-T-ZHer3-A20(其中“T”指示易位酶)。将一系列蛋白质浓度的每种鉴别的构建体与人胰腺癌(HPAF-II)细胞一起培育72小时(图6A),和/或与上皮样细胞(H358细胞)一起培育72小时(图6B)。由于RRSP在递送至细胞中之后诱导凋亡,因此通过测量每种构建体在72小时培育之后杀死人胰腺癌细胞的能力来定量易位。

[0411] 图6A和6B展示易位酶的功能表征的结果。使用RRSP-T-ZHer3-A20模板,定量每种易位酶的功能。引起细胞活力降低最大毒性的50%的给定构建体的浓度(EC₅₀)表示在条形图中。更低值表示更有效的易位酶。来自松生弯曲嗜酸菌的易位酶未表达,并且因此无法评估。来自松土壤链霉菌和吐鲁番长丝菌的易位酶在分析中被确定为无功能,因为直到最高

测试剂量 (100nM) 都没有观察到毒性。剩余的五种易位酶在分析中具有功能。来自龟奥斯特维克氏菌的蛋白质的易位酶 (SEQ ID NO:2) 和来自龟奥斯特维克氏菌 LK16-18 的蛋白质的易位酶 (SEQ ID NO:17) 在细胞上最具活性并且比 DT 的易位酶更好。观察到的活性水平表明 RRSP 被有效释放。

[0412] 易位酶结构域白网链霉菌和嗜脉清野氏菌也被证明具有功能 (参见苏吉曼-马兰戈斯 (Sugiman-Marangos) 等人 2022, 其以全文引用的方式并入)。

[0413] 白喉类毒素疫苗是用于预防白喉疾病的全球疫苗接种计划的一部分。人血清中的抗 DT 抗体通过结合到 DT 并中和其功能来阻止 DT 的作用。不幸的是, 这些抗体也结合并中和基于 DT 的治疗剂。进行 ELISA 以评估人血清中的抗 DT 抗体识别全长毒素 DT 和 CT1 (SEQ ID NO:2) 以及易位酶 DT-T 和 CT1-T (SEQ ID NO:3) 的程度。如图 7A 中的 ELISA 数据所示, 人血清中的抗体识别 DT, 但未展示与 CT1 的结合。类似地, 发现针对来自 DT 的易位酶的高效价, 但未发现针对来自 CT1 的易位酶的高效价 (图 7B)。

[0414] 图 7A 和 7B 一起展示人血清中预先存在的抗 DT 抗体不结合或中和基于 CT1 的免疫毒素。将相应的蛋白质固定在板上并且与不同量的儿科人血清样品一起培育, 之后使用缀合到 HRP 的抗 IgG 抗体来确定抗体结合水平。DT 和 DT-T 均展示信号随人血清的剂量依赖性增加, 而 CT1 和 CT1-T 则没有。这表明全长 CT1 和仅其易位酶不含由人血清中的抗体识别的表位。

[0415] 接下来, 为了评估人血清中预先存在的抗 DT 抗体中和基于 DT 的治疗剂以及基于新颖易位结构域的相应治疗剂的功能的程度, 将基于 DT 的免疫毒素和来自龟奥斯特维克氏菌 LK16-18 的基于 CT 的免疫毒素与人血清一起培育。克隆并纯化基于 DT 和 CT 的免疫毒素, 其中相应毒素的 C 结构域和 T 结构域以重组方式连接到 ZHer3-A20 (DT₁₋₃₈₉-ZHer3-A20 或 CT2₁₋₃₉₁-ZHer3-A20)。如图 8 所示, 在不存在人血清 (PBS) 的情况下, DT、DT₁₋₃₈₉-ZHer3-A20 和 CT2₁₋₃₉₁-ZHer3-A20 都对细胞具有毒性。在人血清 (人) 存在下, 发现 DT 和 DT₁₋₃₈₉-ZHer3-A20 的毒性曲线向右偏移, 这表明人血清抑制/中和其功能。相反, 人血清对 CT2₁₋₃₉₁-ZHer3-A20 的毒性曲线没有影响, 这表明人血清不识别或抑制/中和基于 CT 的治疗剂。

[0416] 图 8 展示抗 DT 抗体不中和基于 CT 的免疫毒素。将免疫毒素与 PBS 对照或人血清一起培育 30 分钟, 并且随后添加到细胞中。在 72 小时评估细胞活力, 并且发现 DT₁₋₃₈₉-ZHer3-A20 在人血清培育后毒性降低约 3 个对数, 而 CT2₁₋₃₉₁-ZHer3-A20 没有变化。

[0417] 实例 3

[0418] 示例性注释的构建体序列

[0419] RRSP-T (白喉棒状杆菌 (C.diphtheria)) -ZHer3-A20 (SEQ ID NO:22)

[0420] GDKTKVVVDLAQIFTVQELKERAKVFAPKIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALTQLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMVVRWANEDPYLAPKHGKGETPSDLGFDKHYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYS GTGYDEL TNKLASATGVDEQLAVLLDDRKGLLIGEVHGSVDVNGLRVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELAMLKTKHLDVTLFENARANGMRIVALDANSSARPVQGT EHG LMYRAGAANNIAVEVLQNL PDGEKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGSCAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEHGPIKNKMSSEPNKTVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVF

AGANYAAWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGEL
VDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTKYQKA
AFIGKLQDDPSQSSELLSEAKKLNSQAPKGGGSGGGGSAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFE
K

[0421] 在上文中:

[0422] 残基1-510=RRSP

[0423] 残基511-724=来自白喉棒状杆菌的T结构域

[0424] 残基725-734=G4S2接头

[0425] 残基735-792=ZHer3:08699亲和体

[0426] 残基793-802=G4S2接头

[0427] 残基803-822=A20FMDV2肽

[0428] 残基823-840=凝血酶裂解位点和链霉素标签II

[0429] RRSP-T(龟奥斯特维克氏菌)-ZHer3-A20(SEQ ID NO:23)

[0430] GDKTKVVVDLAQIFTVQELKERAKVFAPKIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAE
HPTSGRNQALTQLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMVVRWANEDPYLA
PKHGYKGETPSDLGFDKHYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAF
YFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYSGTGYDELTKLASA
TGVDEQLAVLLDDRKGLLIGEVHGSVDNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELSA
MLKTKHLDVTLFENARANGMRIVALDANSSARPVQGTGHEGLMYRAGAANNIAVEVLQNLDPGEKFVAIYGKAHLQS
HKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGSCAGNRVRRSVG
SSLSCLSKINWKNVREKADALTKKVHADKEFMDKLSTHHQRGEAPSVEKTTALHNALLEHESFSALKGARASGKVGA
AATAAWGVAVAQAFTDPKADALTKAATLSVVPGLGQALGIADGIKHENTEEIVVQSISLAGLLAAQAIPVVGAV
DFGLLVYQLVETIVDLATHLSSAAANPAYSPGHKTQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTKYQKAA
FIGKLQDDPSQSSELLSEAKKLNSQAPKGGGSGGGGSAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK

[0431] 在上文中:

[0432] 残基1-510=RRSP

[0433] 残基511-520=G4S2接头

[0434] 残基521-536=具有弗林蛋白酶识别位点的白喉棒状杆菌序列

[0435] 残基537-711=来自龟奥斯特维克氏菌的T结构域

[0436] 残基712-723=来自白喉棒状杆菌的接头序列

[0437] 残基724-733=G4S2接头

[0438] 残基734-791=ZHer3:08699亲和体

[0439] 残基792-801=G4S2接头

[0440] 残基802-821=A20FMDV2肽

[0441] 残基822-839=凝血酶裂解位点和链霉素标签II

[0442] RRSP-T(龟奥斯特维克氏菌LK16-18)-ZHer3-A20(SEQ ID NO:24)

[0443] GDKTKVVVDLAQIFTVQELKERAKVFAPKIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAE
HPTSGRNQALTQLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMVVRWANEDPYLA
PKHGYKGETPSDLGFDKHYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAF

YFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYSGTGYDELTNKLASA
TGVDEQLAVLLDDRKGLLIGEVHGSDVNGLRVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELSA
MLKTKHLDTLTFENARANGMRIVALDANSSARPVNQGTGEHGLMYRAGAANNIAVEVLQNLDPGEKFVAIYGKAHLQS
HKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGSCAGNRVRRSVG
SSLSCLSKVKEQVREKSKKIIDNVKDNPEFMKKLSAHHHERGSAPTTEKITALHNELLDHESFSALKGARSSAGTAA
TAASAAWGLAVAQAFTNPKADDLTKATAVLSAVPGLGQALGIADGIKHHNTEEIVVQSSISLTALIAAQAIPVVGEL
VDFGLLAYQLVESIIDLTRQLSVITANPAYSPGHKTQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTQYQKA
AFIGKLQDDPSQSSELLSEAKKLNDQAPKGGGSGGGGSAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFE
K

[0444] 在上文中:

[0445] 残基1-510=RRSP

[0446] 残基511-520=G4S2接头

[0447] 残基521-536=具有弗林蛋白酶识别位点的白喉棒状杆菌序列

[0448] 残基537-712=来自龟奥斯特维克氏菌LK16-18的T结构域

[0449] 残基713-724=来自白喉棒状杆菌的接头序列

[0450] 残基725-734=G4S2接头

[0451] 残基735-792=ZHer3:08699亲和体

[0452] 残基793-802=G4S2接头

[0453] 残基803-822=A20FMDV2肽

[0454] 残基823-840=凝血酶裂解位点和链霉素标签II

[0455] RRSP-T (奥斯特维克氏菌属TVS96-490-7B) -ZHer3-A20 (SEQ ID NO:25)

[0456] GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAE
HPTSGRNQALTQLKEQVTSALFIGMKQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMVVRWANEDPYLA
PKHGYKGETPSDLGFDKHYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAF
YFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYSGTGYDELTNKLASA
TGVDEQLAVLLDDRKGLLIGEVHGSDVNGLRVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELSA
MLKTKHLDTLTFENARANGMRIVALDANSSARPVNQGTGEHGLMYRAGAANNIAVEVLQNLDPGEKFVAIYGKAHLQS
HKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGSCAGNRVRRSVG
SSLSCLSNIKWVRVQKSEELIKKLKDDENVKAVQERKEGVKPTSDDLQNLHKALIDHESFKELKSVHSNGVKAMD
AANAALWGANVARVFSKSDGLEKATAALAAVPGLGQVMGVADGVTHNTEEVVVQSVALAGFIAAQAIPVVGIV
DIGVLAYQFVEGVIDLSRQMMTSNARPAYSPGHKTQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTQYQKAA
FIGKLQDDPSQSSELLSEAKKLNDQAPKGGGSGGGGSAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK

[0457] 在上文中:

[0458] 残基1-510=RRSP

[0459] 残基511-520=G4S2接头

[0460] 残基521-536=具有弗林蛋白酶识别位点的白喉棒状杆菌序列

[0461] 残基537-711=来自奥斯特维克氏菌属TVS 96-490-7B的T结构域

[0462] 残基712-723=来自白喉棒状杆菌的接头序列

[0463] 残基724-733=G4S2接头

[0464] 残基734-791=ZHer3:08699亲和体

[0465] 残基792-801=G4S2接头

[0466] 残基802-821=A20FMDV2肽

[0467] 残基822-839=凝血酶裂解位点和链霉素标签II

[0468] RRSP-T (克伦基链霉菌) -ZHer3-A20 (SEQ ID NO:26)

[0469] GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAE
HPTSGRNQALTQLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMMVWRANEDPYLA
PKHGYKGETPSDLGFDKAYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAF
YFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYSGTGYDELTNKLASA
TGVDEQLAVLLDDRKGLLIGEVHGSVDNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELSA
MLKTKHLDVTLFENARANGMRIVALDANSSARPNNVQGTEHGLMYRAGAANNIAVEVLQNLDPGEKFVAIYGKAHLQS
HKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGSCAGNRVRRSVG
SSLSCLPVDWDKVEEKAKATAKKVAQDAEHVEKLPKRSPKGPTWGEAHSTAELSHAKVREVSGAHVAASAAGVGTWV
YGMAKTFSDKDATTLDKVAVTGAVVPGLGQALGIADGIQHGDP EAI AVNAVALAALAAQVVPVVG EVVDAVLLTEQ
LVEVLVDVFRATADPAYSPGHKTQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLT KYQKA AFIGKLQDDPSQ
SSELLSEAKKLND SQAPKGGGSGGGGSAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK

[0470] 在上文中:

[0471] 残基1-510=RRSP

[0472] 残基511-520=G4S2接头

[0473] 残基521-536=具有弗林蛋白酶识别位点的白喉棒状杆菌序列

[0474] 残基537-700=来自克伦基链霉菌的T结构域

[0475] 残基701-712=来自白喉棒状杆菌的接头序列

[0476] 残基713-722=G4S2接头

[0477] 残基723-780=ZHer3:08699亲和体

[0478] 残基781-790=G4S2接头

[0479] 残基791-810=A20FMDV2肽

[0480] 残基811-828=凝血酶裂解位点和链霉素标签II

[0481] RRSP-T (链霉菌属TLI053) -ZHer3-A20 (SEQ ID NO:27)

[0482] GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAE
HPTSGRNQALTQLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMMVWRANEDPYLA
PKHGYKGETPSDLGFDKAYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAF
YFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYSGTGYDELTNKLASA
TGVDEQLAVLLDDRKGLLIGEVHGSVDNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELSA
MLKTKHLDVTLFENARANGMRIVALDANSSARPNNVQGTEHGLMYRAGAANNIAVEVLQNLDPGEKFVAIYGKAHLQS
HKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGSCAGNRVRRSVG
SSLSCIPAASWEAIEHRKEMASAVARDTEYTKSLPNRHPKGPWSEAHHTSTSTHARVSAKSGAHVAVGAFVGSW
IYGMSETFANKNVTTLDKAAATVAIVPGIGHALGIAAALEHHDIEGVVNAISIAALAAQVVPVVG EIVDAALLAE
QLVEVLVHVFRASSTTEPAYSPGHKTQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLT KYQKA AFIGKLQDDPS
QSSELLSEAKKLND SQAPKGGGSGGGGSAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK

[0483] 在上文中:

[0484] 残基1-510=RRSP

[0485] 残基511-520=G4S2接头

[0486] 残基521-536=具有弗林蛋白酶识别位点的白喉棒状杆菌序列

[0487] 残基537-701=来自链霉菌属TLI053的T结构域

[0488] 残基702-713=来自白喉棒状杆菌的接头序列

[0489] 残基714-723=G4S2接头

[0490] 残基724-781=ZHer3:08699亲和体

[0491] 残基782-791=G4S2接头

[0492] 残基792-811=A20FMDV2肽

[0493] 残基812-829=凝血酶裂解位点和链霉素标签II

[0494] RRSP-T(吐鲁番长丝菌)-ZHer3-A20(SEQ ID NO:28)

[0495] GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAE
HPTSGRNQALTQLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMVVRWANEDPYLA
PKHGYKGETPSDLGFD AKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAF
YFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYS GTGYDEL TNKLASA
TGVDEQLAVLLDDRKGLLIGEVHGS DVNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELSA
MLKTKHLDVTLFENARANGMRIVALDANSSARP NVQGTEHGLMYRAGAANNIAVEVLQNL PDGEKFVAIYGKAHLQS
HKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGSCAGNRVRRSVG
SSLSCERKTLDKLDRKKIETRAKQVISKLAQDGLRRELPRRTTGHAEHEVRGVLAKSRGALHEIHGAATTVAMPL
ATVAWVEDMARVFREKNATTLDKAATVSEIVPVAGQVLGMADGIAHRDAETVAVNAVVLAAIAVSQAVPVVGELVDL
GLTAYAVVDVVVRLFGPAREVPITQESYWPAYSPGHKTQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTKYQ
KAAFIGKLQDDPSQSSELLSEAKKLNDSQAPKGGGSGGGGSSNAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQ
FEK

[0496] 在上文中:

[0497] 残基1-510=RRSP

[0498] 残基511-520=G4S2接头

[0499] 残基521-536=具有弗林蛋白酶识别位点的白喉棒状杆菌序列

[0500] 残基537-714=来自吐鲁番长丝菌的T结构域

[0501] 残基715-726=来自白喉棒状杆菌的接头序列

[0502] 残基727-736=G4S2接头

[0503] 残基737-794=ZHer3:08699亲和体

[0504] 残基795-804=G4S2接头

[0505] 残基805-824=A20FMDV2肽

[0506] 残基825-842=凝血酶裂解位点和链霉素标签II

[0507] RRSP-T(松土壤链霉菌)-ZHer3-A20(SEQ ID NO:29)

[0508] GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAE
HPTSGRNQALTQLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMVVRWANEDPYLA
PKHGYKGETPSDLGFD AKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAF

YFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYSGTGYDELTNKLASA
TGVDEQLAVLLDDRKGLLIGEVHGSVDNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELSA
MLKTKHLDTLTFENARANGMRIVALDANSSARPVNQGTGEHGLMYRAGAANNIAVEVLQNLDPGEKFVAIYGKAHLQS
HKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGGSGGGGSCAGNRVRRSVG
SSLSCELAKDKARDILSEAGNEVSLPQRSDGLSKQEITATAEATRSLGTGVHGAVSAAMVADWAHDVARTFADPK
ATKLDKAAAVTAIAPVIGQAVNIADGIQHHDKKTIVVNSLVLAAVVAAQAVPAVGEVVDAAIVADVFVEKLVGWFTP
TAKPGPEHVAPAYSPGHKTQPFGGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTQYQKAFIGKLQDDPSQSSELL
SEAKKLNSQAPKGGGGSGGGGSAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK

[0509] 在上文中:

[0510] 残基1-510=RRSP

[0511] 残基511-520=G4S2接头

[0512] 残基521-536=具有弗林蛋白酶识别位点的白喉棒状杆菌序列

[0513] 残基537-695=来自松土壤链霉菌的T结构域

[0514] 残基696-707=来自白喉棒状杆菌的接头序列

[0515] 残基708-717=G4S2接头

[0516] 残基718-775=ZHer3:08699亲和体

[0517] 残基776-786=G4S2接头

[0518] 残基786-805=A20FMDV2肽

[0519] 残基806-823=凝血酶裂解位点和链霉素标签II

[0520] RRSP-T (松生弯曲嗜酸菌) -ZHer3-A20 (SEQ ID NO:30)

[0521] GDKTKVVVDLAQIFTVQELKERAKVFAPKIGASYQGILDQLDLVHQAKGRDQIAASFELNKKINDYIAE
HPTSGRNQALTQLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVAIEEANGKHVGLTDMVVRWANEDPYLA
PKHGYKGETPSDLGFDKHYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKTVHLGYSYQELQDLTGAESVQMAF
YFLKEAAKKADPISGDSAEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDAWDRRYSGTGYDELTNKLASA
TGVDEQLAVLLDDRKGLLIGEVHGSVDNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQPLIDRYLATGVMSSSELSA
MLKTKHLDTLTFENARANGMRIVALDANSSARPVNQGTGEHGLMYRAGAANNIAVEVLQNLDPGEKFVAIYGKAHLQS
HKGIEGFVPGITHRLDLPALKVSDSNQFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGGSGGGGSCAGNRVRRSVG
SSLCKVNWKDAYDRSMNIAQDIDGSAEFRATAPARPAAGKAMPEADVTRLVSTSEDFLEKSTKTSGVKKALKAINN
QKMISWGSLLSNALANSHTWSDKNATNLDKAYAVVGGVPVLGEVIGIASGIDKQDAESIAVNTLSLVGIVAATVCP
LGATVEFVMIGYTAIKLMLSWFTVETIPAYSPGHKTQPFGGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTQYQKA
AFIGKLQDDPSQSSELLSEAKKLNSQAPKGGGGSGGGGSAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFE
K

[0522] 在上文中:

[0523] 残基1-510=RRSP

[0524] 残基511-520=G4S2接头

[0525] 残基521-536=具有弗林蛋白酶识别位点的白喉棒状杆菌序列

[0526] 残基537-712=来自松生弯曲嗜酸菌的T结构域

[0527] 残基713-724=来自白喉棒状杆菌的接头序列

[0528] 残基725-734=G4S2接头

- [0529] 残基735-792=ZHer3:08699亲和体
- [0530] 残基793-802=G4S2接头
- [0531] 残基803-822=A20FMDV2肽
- [0532] 残基823-840=凝血酶裂解位点和链霉素标签II
- [0533] DT-Her3-A20 (SEQ ID NO:31)
- [0534] GADDVVDSSKSFVMENFSSYHGTPKGYVDSIQKGIQKPKSGTQGNYYDDWKGFYSTDN KYDAAGYSVDN
ENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPFAEGSSS
VEYINNWEQAKALSVELEINFETRGRGQDAMYEMAQACAGNRVRRSVGSSLSCINLDWDVIRDKTKTKIESLKEH
GPIKNKMSESPNTVSEEKAKQYLEEFHQTALEHPSELKTVTGTPNFAGANYAAWAVNVAQVIDSETADNLEKT
TAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGELVDIGFAAYNFVESIINLFQVVHNSYNR
PAYSPGHKTQPFGGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTKYQKAFIGKLQDDPSQSSELLSEAKKLND SQ
APKGGGGSGGGGSSNAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK
- [0535] 在上文中:
- [0536] 残基1-389=来自白喉棒状杆菌的序列
- [0537] 残基390-399=G4S2接头
- [0538] 残基400-457=Her3:08699亲和体
- [0539] 残基458-467=G4S2接头
- [0540] 残基468-487=A20FMDV2肽
- [0541] 残基487-505=凝血酶裂解位点和链霉素标签II
- [0542] CT2-ZHer3-A20 (SEQ ID NO:32)
- [0543] YANDAVIADQSKTVDSFTSYHGAKPESFESVLAGIKKPESGSQGNHDPEWKGFTT DNKHAAGYTVSD
ESVMTGKAGGVVKVTYPGKTRVLAVKPLSATELKYTLGLAADKPLIDQLNNKDFINKFGE GASRVVLQMPFADGTSD
VEYIHNWEDATQLQVATEVRFDNLGKRGQDEMNRYMNLANCPSVSAVRVKNRPAKLCLSKVKWEQVREKSKKI IDNV
KDNPEFMKKLSAHHERGSAPTTEKITALHNELLDHESFSALKGARSSAGTAATAASAAWGLAVAQAFTNPKADDLT
KATAVLSAVPGLGQALGIADGIKHHNTEEIVVQSISLTALIAAQAIPVVGELVDFGLLAYQLVESIIDLTRQLSVIT
ANPPTVTHSSKLAGGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTKYQKAFIGKLQDDPSQSSELLSEAKKLND
SQAPKGGGGSGGGGSSNAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK
- [0544] 在上文中:
- [0545] 残基1-391=来自龟奥斯特维克氏菌LK16-18的序列
- [0546] 残基392-401=G4S2接头
- [0547] 残基402-459=Her3:08699亲和体
- [0548] 残基460-469=G4S2接头
- [0549] 残基470-489=A20FMDV2肽
- [0550] 残基491-507=凝血酶裂解位点和链霉素标签II。
- [0551] 表1:序列表

[0552]

全长蛋白质			
SEQ ID NO:	物种	序列	GenBank 登录号
1	白喉棒状杆菌	GADDVVDSSKSFVMENFSSYHGTPGYVDSIQKGIQKPKS GTQGNVDDDKGFYSTDNKYDAAGYSVDNENPLSGKAGGV VKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGT EEFIKRFQDGASRVVLSLPFAEGSSSVEYINNWEQAKALS VELEINFETRGRGQDAMYEMAQACAGNRVRRSVGSSLS CINLDWDVIRDKTCTKIESLKEHGPIKNKMSESPNKTVSE EKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAA WAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGA VHHNTEEIVAQSIALLSSLMVAQAIPLVGELVDIGFAAYNF VESIINLFQVHNSYNRPAYSPGHKTQPFLHDGYAVSWNT VEDSIIRTGFQGESGHDIKITAENTPLPIAGVLLPTIPGK LDVNKSKTHISVNGRKIRMRCRAIDGDVTFPCRPSVYVG NGVHANLHVAFHRSSEKIHSNEISSDSIGVLGYQKTVDH TKVNSKLSLFFFEIKS	来自 WP_072564851.1
2	龟奥斯特维克氏菌 (CT1)	YANDVVLKDQSTTVDSFTSYHGAKPESFNAVLTGIKKPEK GSQGNNDPDWKGFYTDDNKHAAAGYTVSDESVLGKAGGV VRVTYPGKTRILAVKSLSAELKGLGLDSAKPLIDQLND KSFLEKYGDGANRVVLKMPFADGTEDSEFIHNWKDAEQLS VETEVRFNDLGRGQDAMNSYMNMANCPSTSARAKRSPG KICLSKINWKNVREKADALTKKVHADKEFMDKLSTHHQRG EAPSVEKTTALHNALLEHESFSALKGARASGKVGAAASTA AWGVAVAQAFDTPKADALTKTAATLSVVPGLGQALGIADG IKHENTEEIVQSIISLAGLLAAQAIPVVGAEVDFGLLVYQ LVETIVDLATHLSSAAANPPEATDSVRPAVSLGLRAGWK TEEDAKLHIGSPYGMKFQRIVLSAEEGKEIPFVRAAVAVD SKFLKINGPRSFVVQNGIKTPMACFETEGNLAFCRPSRPI FLSSSSPATLHLSYVTNEHENGTIKNPTVDILGQRIVENK VITANKVSLVYKVDSSNTL	N/A
21	龟奥斯特维克氏菌 <i>LK16-18</i> (CT2)	YANDAVIADQSKTVDSFTSYHGAKPESFESVLAGIKKPES GSQGNHDPWEKGFYTDDNKHAAAGYTVSDESVMTGKAGGV VKVTYPGKTRVLAVKPLSATELKTYLGLAADKPLIDQLNN KDFINKFGEASRVVLQMPFADGTSDVEYIHNWEDATQLQ VATEVRFNDLGRGQDEMNRYMNLANCPSVSAVRVKRNP KLCLSKVKWEQVREKSKKIIDNVKDNPEFMKKLSAHERG SAPTTEKITALHNELLDHESFSALKGARSSAGTAATAASA AAWGLAVAQAFNPKADDLTKATAVLSAVPGLGQALGIAD GIKHHNTEEIVQSIISLTALIAAQAIPVVGELVDFGLLAY QLVESIIDLTRQLSVITANPPEVTHSSKLAKSNGLLAGW KTDQDGVLSLGAHPGKMTQLISMSAEKGQEIIPFTGAMIAV KKDLLTVNTPRVFVVQNGIRIPLRCSAAGETLTFCRPNHP	WP_162873017.1 Y64-L603

		VWLSEQHQATLHLSYVTQENESEKIDNPTVDILGQKIVDN KVSTANKVSLVYEVDRSNKL	
35	奥斯特维克氏菌属 <i>TVS 96-490-7B</i> (CT3)	GAFASPISPAIDKEKSKVVENFTSYHGTHQYIENVLKG NKPETGSQGNHDPDWKGFYTTDSKYAAAGYTPDEENPM KAGGVVKVSYTGKTKILALKDMPAGKIKGHLGLDPNKPL EQIKDQEFINKYGEADRVVLQMPFAEGSSEVEYINNWDN AKDLKVEPEVRFDQHGKRGQDAMYEMALASCESAGKAKR SVSKTCLSNIKWERVRQKSEELIKKLKDDENVKAVQERK EGVKPTSDDLQNLHKALIDHESFKELKSVHSNGVKAMDAA NAALWGANVARVFSDSKSDGLEKATAALAAPVGLGQVMGV ADGVTHHNTTEVVVQSVALAGFIAAQAIIPVVGIVDIGVL AYQFVEGVIDLSRQMMTSNARGIAEKRVVDIKDANGLSG GWLTPQDSKIHLPPIHYGKTQRLAVSANGDHNLPITGFKT SINPKFFNINESGSFIVQNGLKVPMNCSEEASKTEAVCIP VHPIWLTDKNNAIHVSQYQTKSPEGSRIDNPRLEVWGQIM HTEEGPKTNTINLGYSVDQEMPCYSSIKFENGGDYVGRLA VNSKDGKVTSHKMSGMRDKDKVDLSAFNSGEVLTTEISPT GGGAPTGPSPKIKYCSDDGIEASIKSWGTDKDAFLGFN	WP_219106995.1 G27-N663
易位结构域			
SEQ ID NO:	物种	易位结构域序列	源序列或 GenBank 登录号 (带坐标的全长序 列)
3	龟奥斯特维克氏菌 (CT1-T)	LSKINWKNVREKADALTKKVHADKEFMDKLSTHHQRGEAP SVEKTTALHNNALLEHESFSALKGARASGKVGAAASTAAWG VAVAQAFTDPKADALTAKAATLSVVPGLGQALGIADGIKH ENTEEIVVQSSISLAGLLAAQAIIPVVGAEVDFGLLVYQLVE TIVDLATHLSSAAAN	SEQ ID NO: 2 L204-N378
4	无淀粉酶链孢囊菌	LPVDWDKVEEKAKATAKKVAQDAEHVERLPKRSPKGPTWG EAHSTAELTHAKVREVSGAHVAASAAGVGTWVYGMAKTFS DKNATTLDKVAVTGAVVPGLGQALGIADGIQHGDPFAIAV NAVALAALAAQVVPVVGEEVDAVLLTEQLVEVLVDVFRT ATAD	PSJ28985.1 L231-D394
5	链霉菌属 <i>TLI 053</i>	IPAASWEAIEHRSEKEMASAVARDTEYTKSLPNRHPKGPTW SEAHTTSTSTHARVSAKSGAHVAVGAFAVGSWIYGMSETF ANKNVTTLDKAAATVAIVPGIGHALGIAAALEHHDIEGVV VNAISIAALAAQVVPVVGEEIVDAALLAEQLVEVLVHVFR ASTTE	SDT83331.1 I198-E362
6	链霉菌属 <i>SLBN-118</i>	LSEMDGEKVTKGQSVADTLKGHGNLQPLVKRTRAEPRTQ QYLPPTTRQQIDDAISAAQRVHPGVDATMKHLGSAQALAGF ALDVSQGEPLRAAADVTAMLPRI GPGLGNALGLADSI VHD DPEGIAVNAIGLGA FIVAQAI PVIGELADLGF AAYALIET LKSMFAGPTLG	WP_160159328.1 L631-G801
7	链霉菌属 <i>A48</i>	LPVDWDKVEEKAKATAKKVAQDAEHVERLPKRNPKGPTWG EAHSTAELTHAKVRAVSGAHVAATAAGVGTWVYGMAKTFS DKDATTLDKVAVTGAVVPGLGQALGIADGIQHGDPFAVAV NAVALAALAAQVVPVVGEEVDAALLTEQLVEVLVDVFRT ATAD	WP_168096531.1 L241-D404
8	玫瑰轮丝链霉菌	LPVDWDKVEEKAKATAKKVAQDAEHVERLPKRNPKGPTWG EAHSTAELTHAKVRAVSGAHVAATAAGVGTWVYGMAKTFS DKDATTLDKVAVTGAVVPGLGQALGIADGIQHGDPFAVAV NAVALAALAAQVVPVVGEEVDAALLTEQLVEVLVDVFRT ATAD	WP_078659863.1 L231-D394
9	松土壤链霉菌	ELAKDKARDILSEAGNEVSLPQRDSGLSKQEITATAEAT RSKLGTGVHGA VSAAMVADWAHDVARTFADPKATKLDKAA AVTAIAPVIGQAVNIADGIQHHDKKTIVVNSLVLA AVVAA	JZ58907.1 E225-A383

[0553]

[0554]

		QAVPAVGEVVDAAIVADVFVEKLVGWFTPTAKPGPEHVA	
10	链霉菌属 <i>MBT76</i>	LPVDWDKVEEKAKATAKKVAQDAEHAERLPKRNPKGPTWG EAHSTAELTHAKVREVSGTHVAATAAGVGTWVYGMAKTFS DKDATTLDKVAVTGAVVPGLGQALGIADGIQHGDP EAVAV NAVALAALAAAQVVPVVG EVVDAVLLTEQLVEVLVDVFRA ATAD	WP_079110321.1 L181-D344
11	克伦基链霉菌	LPVDWDKVEEKAKATAKKVAQDAEHVEKLPKRSPKGPTWG EAHSTAELSHAKVREVSGAHVAASAAGVGTWVYGMAKTFS DKDATTLDKVAVTGAVVPGLGQALGIADGIQHGDP EAIIV NAVALAALAAAQVVPVVG EVVDAVLLTEQLVEVLVDVFRT ATAD	WP_120757473.1 L227-D390
12	白网链霉菌	GINVAAYKADIEKLIKDVYEDPNFSDLKNRTGGPQKDKDT LKGYYERLKPKEVETLRPLKAGVSSAVGAAGAI SWAIGVAD AFTSENVSSFDKAAAVTAIVPGLGECVGIANAI DKRDEP LIINTISMAALMASAAVPVLAPIGVALDAGLAAAQGVATV LEYLEIGQ	WP_095582082.1 G232-Q399
13	松生弯曲嗜酸菌	KVNWKDAYDRSMNIAQDIDGSAEFRATAPARPAAGKAMPE ADVTRLVSTSEDFLEKSTKTSGVKKALKAINNQKMI SWGS LLSNALANSHTWSDKNATNLDKAYAVVGGVPVLGEVIGIA SGIDKQDAESI AVNTLSLVGIVAATVCPPLGATVEFVMIG YTAIKLMLS WFTVETI	WP_133259917.1 K246-I421
14	嗜脉清野氏菌	LDWEKIKTESQRIVKQIIEEHPELQSHSKNAVTDKEKLQK IYNDYAPKIDKLSSLKEGVS RATTALNIASWAAGLAETFS NKNADGLDKAAAVTAIIPGLGQAVGIANGIEKH DGEAIAI NSIALSALVVAQAIPIVGEIADVVGAGLILAGGLAQLIQS VSPDT	WP_073156187.1 L252-T416
15	吐鲁番长丝菌	ERKTLDKLDRKKIETRAKQVISKLAQDGDRLRRELPRRTT GHAHEEVRGVLAKSRGALHEIHGAATTVAMPLATVAVVED MARVFREKNATTLDKAATVSEIVPVAGQVLGMADGIAHRD AETVAVNAVVLAAI AVSQAVPVVGELVDLGLTAYAVVDV VRLFGPAREVPITQESY	WP_189053160.1 E249-W426
16	奥斯特维克氏菌属 <i>TVS 96-490-7B</i> (CT3-T)	LSNIKWERVRQKSEELIKKLKDDENVKAVQERKEGVKPT SDDLQNLHKLALIDHESFKELKSVHSGVKAMDAANAALWG ANVARVFSDSKSDGLEKATAALAAVPGLGQVMGVADGVTH HNTEEVVQSV ALAGFIAAQAI PVVGEIVDIGVLAYQFVE GVIDLSRQMMTSNAR	WP_219106995.1 L233-R407
17	龟奥斯特维克氏菌 <i>LK16-18</i> (CT2-T)	LSKVKEQVREKSKKIIDNVKDNPEFMKKLSAHHERSAP TTEKITALHNELLDHESFSALKGARSSAGTAATAASAAW GLAVAQAFTNPKADDLTKATAVLSAVPGLGQALGIADGIK HHNTEEIVVQSI SLTALIAAQAI PVVGEIVDFGLLAYQLV ESIIDLTRQLSVITAN	WP_116115734.1 L267-N442
36	产气克雷伯氏菌	GLNWGSIREKAVDKAKSVTSDKKFINSLPERGKNSPANKN DINTIYESSANHSTLKGGEVIHKASLLLWVANTADVLTS NSSSTLDKATAIIGIVPGIGDAFSLADSI EHDDVEGIITS SIALAAFTVAQAVPVVGELVDLALLTEAAISSLVNVISNY IQSTQR	EIZ2913133.1 G211-R376
37	链霉菌属 <i>MUM 178J</i>	ADLDWEKVRRAKETARQVAQDSEHVSGLPERRAEGLTKS EAHAVAERTHAKVSEVSGTRVAATAFGVGTWAYGMAQTFT DQNATTLDKVAVTTAIVPGVGNALGIADGIEHHDPEAVAV NAVALAALVAAQAVPVVGELVDTVLLTEQIVETFFVNI FQR ATAA	MCH0551590.1 A229-A392
38	嗜冷克洛斯氏菌	ETADSLGELFHQGMTIAGEAVAELDLEEALKSQKVAKGDA RAVVAGKVAAVGKQVGKSSALAKFKEVLKPSKGVIALWAV SALYTLAQEDVSTVDKAEAVLAIVPILGPLFSFGNSYRKE DVEGMVISTVALAATLIAFLCPPVGAGIGVGLAVYQVVRA FC	MBB4677777.1 E56-C217

[0555]

39	异库茨涅尔氏菌属 <i>NRRL B-24872</i>	SLVKGKTDQAQVSAVKQLFENSKTVATSLVADSDMADAIAIK DAKKAQKVGVPVDPKVKVTDPLVEKALKSKSFANTKAG KYLKNAQPSKAGTALWAATTVATFLDGSSSASDKAVAAVG IIPVVGQLASLIHTTIKDPSPDVEGIVTGTIGLAAAVITIA C	WP_143261759.1 S60-C220
40	白色异库茨涅尔氏 菌	KAIGATRESARKFVANILSKSEKATTTLAASNDYKSSMKG F SESFAKLKRAAAETKKPIPTADIKKLLTPHLDKVSTVKDF K KYFTNLKPSQGSAAATWLVNLAATVANEESSKADVAVAAAG I IPVVGQVAAIANSVSQGDVEGTVTSAIALAANIAIAVGAG	WP_156051914.1 K78-G240
41	链霉菌属 <i>AV19</i>	CDIGVRDHAIEVARKHQERLRAHGDDVEYVTTLDEFKPNK SLLSPLRETNGVLNDAAAAELGAKDFEAFSTRVTEALGRD LRAVHGLSTDITLLAKVARQAGRAAELGAKALPYVGIAAT GYAIAEDVKAGDYANLAFDSVAEGLQVAMVAQPELVPLLE PVLFAEQLAQIVYDEIAGWFKRQEQLDHRKQWDSAAEDL VEHRDGQWRDHLVGQALQKLF	WP_199893204.1 C260-P481
42	链霉菌属 <i>NRBC_110611</i>	CGWGDDGVAHTQERLRAQGGDVEYVTSLEEFKPNKALL SPLRETNKVLNDAAASAEELGAKDFETFSARVGAAFSRDLRA VQGLTTDTGLLAGVGRQVGRAGDLGVKVLPHYVGIATGYA VAEDAKAGDYANLAFDSVAEGLQVAMVAQPEVAPLLEPAL LAEQLAQLVYNEVAGWIAEQKQLEHDRAQWDGAVKDLVAH RDTQWRDRLVGQAVQELFP	WP_147264604.1 C231-P449
43	注射器链霉菌	CGIAAADDAKVARTHQERLRAQGADVEYVTSLEEFKPNK ALMSPLRETSKVLNGAAAAELGAKDFEAFSTQVTEALGRD LKAVHGLTTDTGLLAEVGRVGRAGELGVKILPYAGIAAT GYAVAEDAKAGDYGNLAFDSVAEGLQVAMVAQPELTPLLE PALLAEQLAQLVYDEVAGWIAEQKQLAHDRGQWDDAVKDL VAHRDTQWRNRLEGQALQELFP	WP_209513619.1 C232-P453
44	假诺卡氏菌科细菌 <i>YIM PH 21723</i>	RIGKAALAKATAEAQADAQKILKDPSEFKPPHPAPHQKLT AELTQVANDLAKRVSAHPGARDLTGLGKGLAKALKGVIS GVTAGIQAAEIAESAAGSNSGDLKVCALGPPIVGDIA GIVNAAAEAGDAVEIVKNVLTGLSMVSVAFPPPLAPVADAV AFVLEVAVFLGKYLWGLLF	RJQ69589.1 R276-F454
45	曼谷放线动孢菌	SLSQDQLDQAVADAQATAHQVIDDPRSFELPELAKGQVLS EQELAAALANGVAQDVAAANSNGGERAIVPAYVTRALSKAIP AIAEHGHDAFLKVSAREGVLTGNGALAKVTSVVGVPVPL GSLAGIVNEAVGISADDPAAIAQKALEIAKNVAGVALTV MAIAFPFPFAPVAALATAAWEVGAALGKYLGLL	WP_143218892.1 S322-L515
46	优洛菌素链霉菌	CETADGGDTAEEGKQPGEGRQATESDSTEGRNPPQGEAGPA EAGKPVADDEWVARVQRERLARHGSDEYVTSVAELEKNP KAVLSPLRETNTTLNKARDKPLTAAEYEFASKAAEAMSK DLKGVQAASDSTVLAKVGAKLARGGSIIVKVGPRVPLKRR LLRSSAASSR	MBF6055834.1 C280-R449
47	灭病菌素链霉菌	GTGRQDGGFSSEGREGAVTKEDLKGDPPQKVVELNRSRIAR HGADVYVTSLDALAANPKAVLIPLRNSSQTLDAALGKSH LTVDEFDRFSAELTQSFGRDLQAVENLSTSEKLLSSIGKG ARAGGEGLLKVVPFAGVAASSAALAEDLKSGENGDA	WP_169790908.1 G270-A425
48	卡廷加链霉菌	SADIARTHRERLRAHGDDVEYVTSLEEFKPNKALLSPLR ETGKVLNDAGNAGLSAGDFETFSARVGEAFERDLRAAHEL TTDTGLLARVGAEAGRAAELGAKALPFVGIATGYAVAED VKSGDYANLAFDSVAEGLQVAMTAQPEFA	WP_157868472.1 S241-A389
构建体和模块			
SEQ ID NO:	描述	序列	来源

18	创伤弧菌(RRSP)	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGMKQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVWRANEDPYLAPKHGYKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDA WDRRYSGTGYDELTKLASATGVDEQLAVLLDDRKGLLIG EVHGSDVNGLRVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSELSAMLKTKHLDVTLFENARANGMRIV ALDANSSARPVQGTGHEGLMYRAGAANNIAVEVLQNLDPG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVVYDDVANKPKITFKGSL	WP_011081430.1 G3580-L4089
19	Her3 亲和体 (ZHer3:08699)	AEAKYAKEKYNAYYEIWQLPNLTQYQKAAFIGKLQDDPSQ SSELLSEAKKLND SQAPK	纳扎里(Nazari)等 人(2019)
20	A20 肽 (A20FMDV2)	NAVPNLRGDLQVLAQKVART	迪卡拉(DiCara) 等人(2007)
22	RRSP - DT-T - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGMKQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVWRANEDPYLAPKHGYKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDA WDRRYSGTGYDELTKLASATGVDEQLAVLLDDRKGLLIG EVHGSDVNGLRVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSELSAMLKTKHLDVTLFENARANGMRIV ALDANSSARPVQGTGHEGLMYRAGAANNIAVEVLQNLDPG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGS CAGNRVRRSVGSSLSCLNLDWDVIRDKTKTKIESLKEHGP IKNKMSESPNKTVSEEKAKQYLEEFHQTALEHPELSELKT VTGTNPVFAGANYAAWAVNVAQVIDSETADNLEKTTAALS ILPGIGSVMGIA DGAVHHNTEEIVAQSIASLSLMVAQAIP LVGELVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHK TQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTQYQ KAAFIGKLQDDPSQSSELLSEAKKLND SQAPKGGGSGGG GSNAVPNLRGDLQVLAQKVARTQALVPRGSAWSHPQFEK	n/a
23	RRSP - CT1-T - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGMKQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVWRANEDPYLAPKHGYKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDA WDRRYSGTGYDELTKLASATGVDEQLAVLLDDRKGLLIG EVHGSDVNGLRVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSELSAMLKTKHLDVTLFENARANGMRIV ALDANSSARPVQGTGHEGLMYRAGAANNIAVEVLQNLDPG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGS CAGNRVRRSVGSSLSCLSKINWKNVREKADALTKKVHADK EFMDKLSTHHQRGEAPSVEKTTALHNALLEHESFSALKGA RASGKVGAAASTAAWGVAVQAFTDPKADALTKAATLSV VPGLGQALGIADGIKHENTEEIVVQSIISLAGLLAAQAIPV VGEAVDFGLLVYQLVETIVDLATHLSAAANPAYSPGHKT QPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTQYQ AAFIGKLQDDPSQSSELLSEAKKLND SQAPKGGGSGGGG	n/a

[0556]

[0557]

		SNAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK	
24	RRSP - CT2-T - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGMKQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVVRWANEDPYLAPKHGYKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSDRMDQIEGIYRSSHETDIDA WDRRYSGTGYDELTKLASATGVDEQLAVLLDDRKGLLIG EVHGSDVNGLRFRVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSSSELSAMLKTKHLDVTLFENARANGMRIV ALDANSSARPVQGEHGLMYRAGAANNIAVEVLQNLDPG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVYDDVANKPKITFKGSLGGGSGGGGS CAGNRVRRSVGSSLSCLSKVWEQVREKSKIIDNVKDNF EFMKKLSAHERGSAPTTEKITALHNELLDHESFSALKGA RSSAGTAATAASAAWGLAVAQFTNPKADDLTAKATAVLS AVPGLGQALGIADGIKHNTEEIVVQSI SLTALIAAQAIP VVGELVDFGLLAYQLVESIIDLTRQLSVITANPAYSPGHK TQPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTQYQ KAAFIGKLQDDPSQSSELLSEAKKLND SQAPKGGGSGGGG G SNAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK	n/a
25	RRSP - CT3-T - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGMKQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVVRWANEDPYLAPKHGYKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSDRMDQIEGIYRSSHETDIDA WDRRYSGTGYDELTKLASATGVDEQLAVLLDDRKGLLIG EVHGSDVNGLRFRVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSSSELSAMLKTKHLDVTLFENARANGMRIV ALDANSSARPVQGEHGLMYRAGAANNIAVEVLQNLDPG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVYDDVANKPKITFKGSLGGGSGGGGS CAGNRVRRSVGSSLSCLSNIKWERVRQKSEELIKKLKDE NVKKAVQERKEGVKPTSDDLQNLHKALIDHESFKELKSVH SNGVKAMDAANAALWGANVARVFSDSKSDGLEKATAALAA VPGLGQVMGVADGVTHNTEEIVVQSV ALAGFIAAQAIPV VGEIVDIGVLAYQFVEGVIDLSRQMMTSNARPAYSPGHKT QPFGGGSGGGGSAEAKYAKEKYNAYYEIWQLPNLTQYQ AAFIGKLQDDPSQSSELLSEAKKLND SQAPKGGGSGGGG G SNAVPNLRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK	n/a
26	RRSP - T (克伦基 链霉菌) - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILD QLDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQA LTQLKEQVTSALFIGMKQVAQAGIDAIAQTRPELAARIF MVAIEEANGKHVGLTDMVVRWANEDPYLAPKHGYKGETP SDLGFDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLD ESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKAD PISGDSAEMILLKKFADQSYLSQLDSDRMDQIEGIYRSS HETDIDAWDRRYSGTGYDELTKLASATGVDEQLAVLLD DRKGLLIGEVHGSDVNGLRFRVNEQMDALKKQGVTVIGLE HLRSDLAQPLIDRYLATGVMSSSELSAMLKTKHLDVTLFE NARANGMRIVALDANSSARPVQGEHGLMYRAGAANNI AVEVLQNLDPGGEKFVAIYGKAHLQSHKGIEGFVPGITHR LDLPALKVSDSNQFTVEQDDVSLRVYDDVANKPKITFK GSLGGGSGGGGSCAGNRVRRSVGSSLSCLPVDWDKVEE KAKATAKKVAQDAEHVEKLPKRSPKGPTWGEAHSTAELS HAKVREVSGAHVAASAAGVGTWVYGMAKTFSDKDATTLD	n/a

[0558]

		KVAVTGAVVPGLGQALGIADGIQHGDP EAI AVNAVALAA LAAQVVPVVG EVVDAVLLTEQLVEVLVDVFR TATADPA YSPGHKTQPF GGGSGGGSSAEAKYAKEKYNAYYEIWQL PNLT KYQKA AFIGKLQDDPSQSSELLSEAKKLND SQAPK GGGSGGGGSSNAV PNLRGDLQVLAQKVARTRQALVPRGS AWSHPQFEK	
27	RRSP - T (链霉菌 属 TL1053) - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVVRWANEDPYLAPKHGKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDA WDRRYS GTGYDEL TNKLASATGVDEQLAVLLDDRKGLLIG EVHGS DVNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSELSAMLKTKHLDVTLFENARANGMRIV ALDANSARP NVQGT EHG LMYRAGAANNIAVEVLQNL PDG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGS CAGNRVRRSVGSSLS CIP AASWEAIEHRSKEMASAVARDT EYTKSLPNRHPKGPTWSEAH TTSTSTHARVSAKSGAHVAV GAFAVGSWIYGMSETFANKNVTTLDKAAATVAIVPGIGHA LGIAAALEHHDIEGVVN AISIAALAAQVVPVVG EIVDA ALLAEQLVEVLVHVFRAS TTEPAYSPGHKTQPF GGGSGG GGSAEAKYAKEKYNAYYEIWQLPNLT KYQKA AFIGKLQDD PSQSSELLSEAKKLND SQAPKGGGSGGGGSSNAV PNLRGD LQVLAQKVARTRQALVPRGS AWSHPQFEK	n/a
28	RRSP - T (吐鲁番 长丝菌) - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVVRWANEDPYLAPKHGKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDA WDRRYS GTGYDEL TNKLASATGVDEQLAVLLDDRKGLLIG EVHGS DVNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSELSAMLKTKHLDVTLFENARANGMRIV ALDANSARP NVQGT EHG LMYRAGAANNIAVEVLQNL PDG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGS CAGNRVRRSVGSSLS CERKTLDKLDRKKIETRAKQVISKL AQDGD LRRELPRRTT TGHAEHEVRGVLAKSRGALHEIHGA ATTVAMPLATVAVVEDMARVFREKNATTLDKAATVSEIVP VAGQVLGMADGIAHRDAETVAVNAVVLAAIAVSQAVPVVG ELVDLGLTAYAVVDVVRLFGPAREVPITQESYWPAYSPG HKTQPF GGGSGGGSSAEAKYAKEKYNAYYEIWQLPNLT K YQKA AFIGKLQDDPSQSSELLSEAKKLND SQAPKGGGSG GGGSSNAV PNLRGDLQVLAQKVARTRQALVPRGS AWSHPQF EK	n/a
29	RRSP - T (松土壤 链霉菌) - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGKMQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVVRWANEDPYLAPKHGKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSRMDQIEGIYRSSHETDIDA WDRRYS GTGYDEL TNKLASATGVDEQLAVLLDDRKGLLIG EVHGS DVNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSELSAMLKTKHLDVTLFENARANGMRIV	n/a

[0559]

		ALDANSSARPNNVQGTGTEHGLMYRAGAANNIAVEVLQNLDPG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGS CAGNRVRRSVGSSLSCELAOKKARDILSEAGNEVSLPQRD SDGLSKQEITATAEATRSKLTGVHGAVSAMVADWAHDV ARTFADPKATKLDKAAAVTAIAPVIGQAVNIADGIQHHDK KTIVVNSLVLAADVAAQAVPAVGEVVDAAIADFVVEKLV GWFTPTAKPGPEHVAPAYSPGHKTQPFGGGSGGGGSAEA KYAKEKYNAYYEWQLPNLTKYQKAFIGKLQDDPSQSSE LLSEAKKLNDSPAPKGGGSGGGGSSNAVPLNRGDLQVLAQ KVARTRQALVPRGSAWSHPQFEK	
30	RRSP - T (松生弯 曲嗜酸菌) - ZHer3-A20 构建体	GDKTKVVVDLAQIFTVQELKERAKVFAKPIGASYQGILDQ LDLVHQAKGRDQIAASFELNKKINDYIAEHPTSGRNQALT QLKEQVTSALFIGMKQVAQAGIDAIAQTRPELAARIFMVA IEEANGKHVGLTDMVVRWANEDPYLAPKHGKGETPSDLG FDAKYHVDLGEHYADFKQWLETSQSNGLLSKATLDESTKT VHLGYSYQELQDLTGAESVQMAFYFLKEAAKKADPISGDS AEMILLKKFADQSYLSQLDSDRMDQIEGIYRSSHETDIDA WDRRYSGTGYDELTKLASATGVDEQLAVLLDDRKGLLIG EVHGSVDNGLRFVNEQMDALKKQGVTVIGLEHLRSDLAQP LIDRYLATGVMSSSELSAMLKTKHLDVTLFENARANGMRIV ALDANSSARPNNVQGTGTEHGLMYRAGAANNIAVEVLQNLDPG EKFVAIYGKAHLQSHKGIEGFVPGITHRLDLPALKVSDSN QFTVEQDDVSLRVVYDDVANKPKITFKGSLGGGSGGGGS CAGNRVRRSVGSSLSCKVNWKDAYDRSMNIAQDIDGSAEF RATAPARPAAGKAMPEADVTRLVSTSEDFLEKSTKTSQV KALKAINNQKMSWGSLLSNALANSHTWSDKNATNLDAKAY AVVGGVPVLGEVIGIASGIDKQDAESIAVNTLSLVGIVAA TVCPPLGATVEFVMIGYTAIKMLSWFTVETIPAYSPGHK TQPFGGGSGGGGSAEAKYAKEKYNAYYEWQLPNLTKYQ KAFIGKLQDDPSQSSELLSEAKKLNDSPAPKGGGSGGGG GSSNAVPLNRGDLQVLAQKVARTRQALVPRGSAWSHPQFEK	n/a
31	DT ₁₋₃₈₉ -ZHer3-A20 构建体	GADDVVDSSKSFVMENFSSYHGKPGYVDSIQKGIQKPKS GTQGNYYDDWKGFYSTDNKYDAGYSVDNENPLSGKAGGV VKVTPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGT EEFIKRFGDGASRVVLSLPPFAEGSSSVEYINNWEQAKALS VELEINFETRGRGQDAMEYMAQACAGNRVRRSVGSSLS CINLDWDVIRDKTKTKIESLKEHGPIKNKMSESPNKTVE EKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAA WAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGA VHHNTEEIVAQSIALLSLMVAQAIPLVGELVDIGFAAYNF VESIINLFQVVHNSYNRPAYSPGHKTQPFGGGSGGGGSA EAKYAKEKYNAYYEWQLPNLTKYQKAFIGKLQDDPSQS SELLSEAKKLNDSPAPKGGGSGGGGSSNAVPLNRGDLQVL AQKVARTRQALVPRGSAWSHPQFEK	n/a
32	CT ₂₁₋₃₉₁ -ZHer3-A2 0 构建体	YANDAVIADQSKTVDSFTSYHGAKPESFESVLGAIKKPES GSQGNHDPWEKGFYTNDKHAAGYTVSDESVMTGKAGGV VKVTPGKTRVLAVKPLSATELTKTYLGLAADKPLIDQLNN KDFINKFGEASRVVLQMPFADGTSDEYIHNWEDATQLQ VATEVRFDNLGKRGQDEMNRMYMLANCPVSASVRVKNPA KLCLSKVKEQVREKSKIIDNVKDNPEFMKKLSAHTHERG SAPTTEKITALHNELLDHESFALKGARSSAGTAATAASA AAWGLAVAQAFNPKADDLTKATAVLSAVPGLGQALGIAD GIKHHNTEEIVVQSI SLTALIAAQAI PVVGELVDFGLLAY QLVESIIDLTRQLSVITANPPTEVTHSSKLAGGGSGGGG SAEAKYAKEKYNAYYEWQLPNLTKYQKAFIGKLQDDPS QSSELLSEAKKLNDSPAPKGGGSGGGGSSNAVPLNRGDLQ VLAQKVARTRQALVPRGSAWSHPQFEK	n/a
33	来自白喉棒状杆菌	CAGNRVRRSVGSSSLSC	n/a

[0560]		的接头, 其含有弗林蛋白酶识别位点(具有包围半胱氨酸)	
	34	来自白喉棒状杆菌的 T 结构域后的接头	PAYSPGHKTQPF n/a

[0561] 参考文献

[0562] K.E. 奥勒尔 (Orrell, K.E.), M.J. 曼斯菲尔德 (Mansfield, M.J.), A.C. 多克塞 (Doxey, A.C.) 和 R.A. 梅尔尼克 (Melnik, R.A.) 艰难梭菌毒素 B 膜易位机构是进化上保守的蛋白质递送设备 (The *C. difficile* toxin B membrane translocation machinery is an evolutionarily conserved protein delivery apparatus). 自然通讯 (Nature Communications) 11, 1-11 (2020)。

[0563] M. 帕克 (Park, M.) 等人. 通过工程化的白喉毒素细胞内递送人嘌呤核苷磷酸化酶挽救靶细胞中的功能 (Intracellular Delivery of Human Purine Nucleoside Phosphorylase by Engineered Diphtheria Toxin Rescues Function in Target Cells). 分子药剂学 (Molecular Pharmaceutics) 15, 5217-5226 (2018)。

[0564] M. 纳扎里 (Nazari, M.); S. 扎马尼·库哈卢 (Zamani Koukhaloo, S.); S. 穆萨维 (Mousavi, S.); A. 米奈-德黑拉尼 (Minai-Tehrani, A.); R. 埃玛姆扎德 (Emamzadeh, R.); R. 切拉吉 (Cheraghi, R.) 开发 ZHER3 亲和体靶向纳米载体用于将基因递送到 HER3 过表达的乳腺癌细胞 (Development of a ZHER3-Affibody-Targeted Nano-Vector for Gene Delivery to HER3-Overexpressed Breast Cancer Cells). 大分子生物学 (Macromol. Biosci.) 2019, 19 (11)。

[0565] D. 迪卡拉 (DiCara, D.); C. 拉皮萨达 (Rapisarda, C.); J.L. 萨克利夫 (Sutcliffe, J.L.); S.M. 维奥莱特 (Violette, S.M.); P.H. 韦瑞伯 (Weinreb, P.H.); I.R. 哈特 (Hart, I.R.); M.J. 霍华德 (Howard, M.J.); J.F. 马歇尔 (Marshall, J.F.) $\text{A}\alpha\text{v}\beta 6$ 整合素配体中的 Arg-Gly-Asp 螺旋基序的结构-功能分析 (Structure-Function Analysis of Arg-Gly-Asp Helix Motifs in $\text{A}\alpha\text{v}\beta 6$ Integrin Ligands). 生物化学杂志 (J. Biol. Chem.) 2007, 282 (13), 9657-9665。

[0566] S.N. 苏吉曼-马兰戈斯 (Sugiman-Marangos, S.N.), S.K. 吉尔 (Gill, S.K.), M.J. 曼斯菲尔德 (Mansfield, M.J.) 等人. 远缘白喉毒素同源物的结构揭示进化上保守的毒素支架的功能决定因素 (Structures of distant diphtheria toxin homologs reveal functional determinants of an evolutionarily conserved toxin scaffold). 通讯生物学 (Commun Biol) 5, 375 (2022)。

[0567] 在先前描述中, 出于解释的目的, 阐述众多细节以便提供对实施例的透彻理解。然而, 所属领域的技术人员将明白, 不需要这些具体细节。在其它情况下, 以框图形式示出众所周知的电气结构和电路, 以免混淆理解。举例来说, 不提供关于将本文所述的实施例实施为软件例程、硬件电路、固件还是其组合的具体细节。

[0568] 上述实施例既定仅为实例。所属领域的技术人员可以对特定实施例实现更改、修改和变化。权利要求书的范围不应受本文中所阐述的具体实施例限制, 而是应以与整个说明书一致的方式解释。

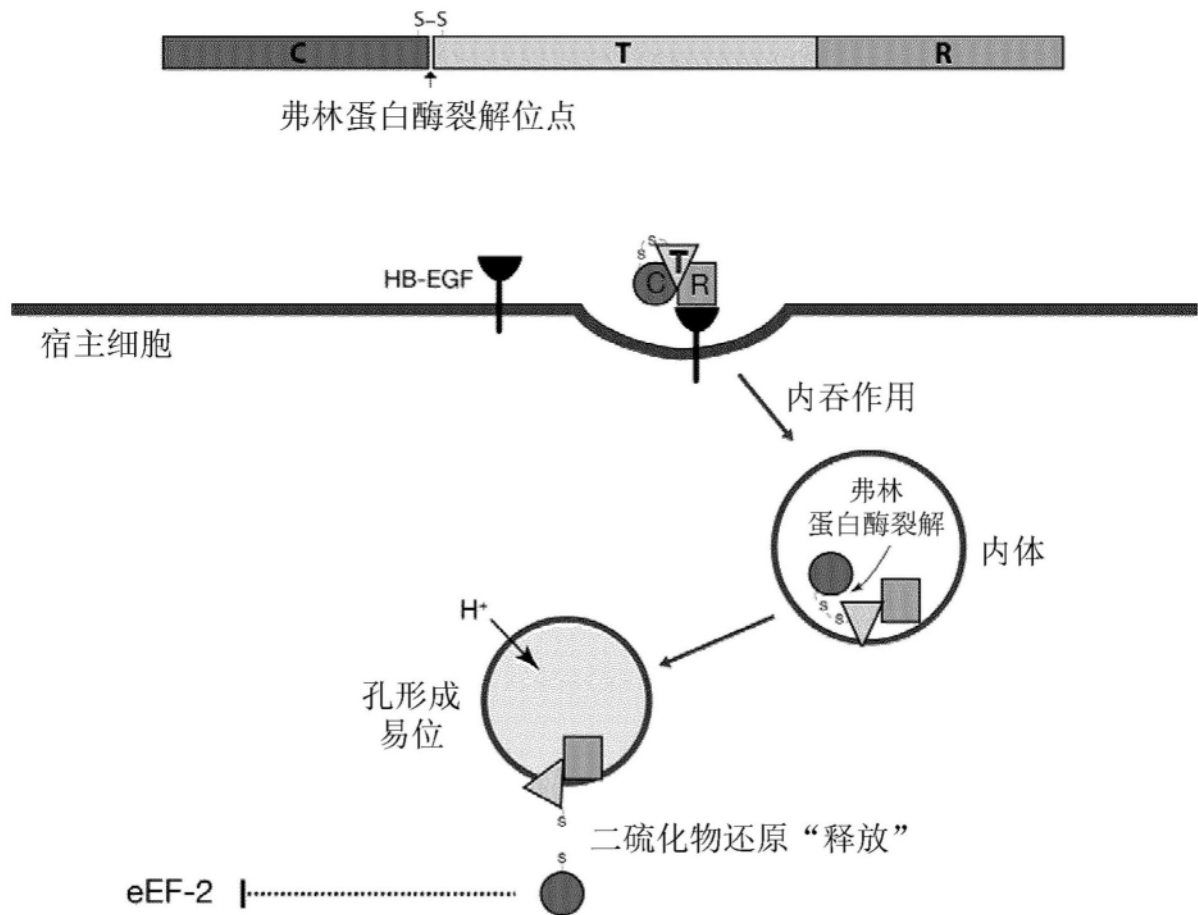


图1A

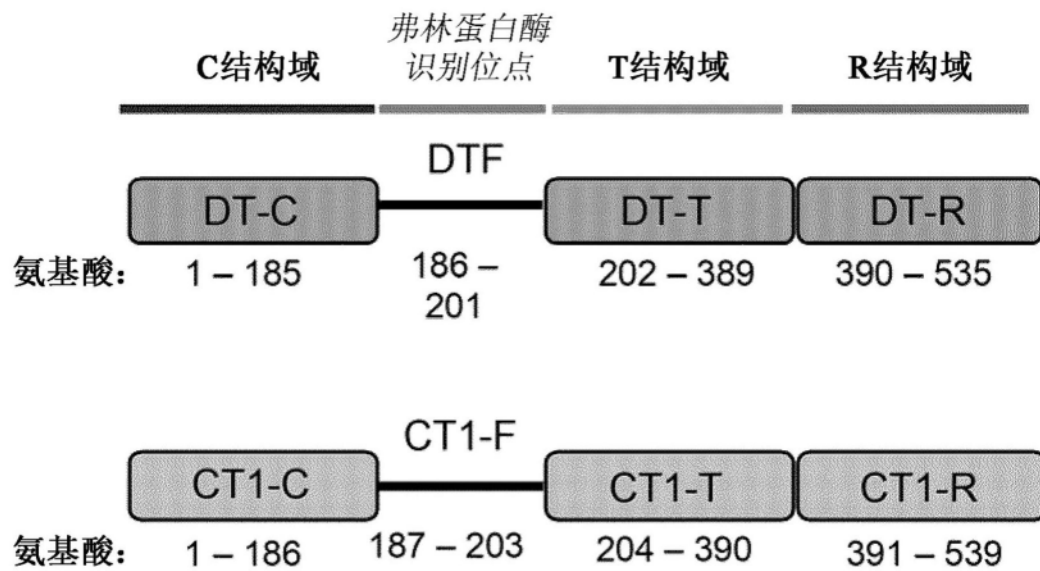


图1B



白喉毒素(DT)
(PDB ID: 1MDT)



龟毒素1 (CT1)
DT样(龟奥斯特维克氏菌)
分辨率: 2.50Å
与DT的RMSD: 2.4Å

图1C

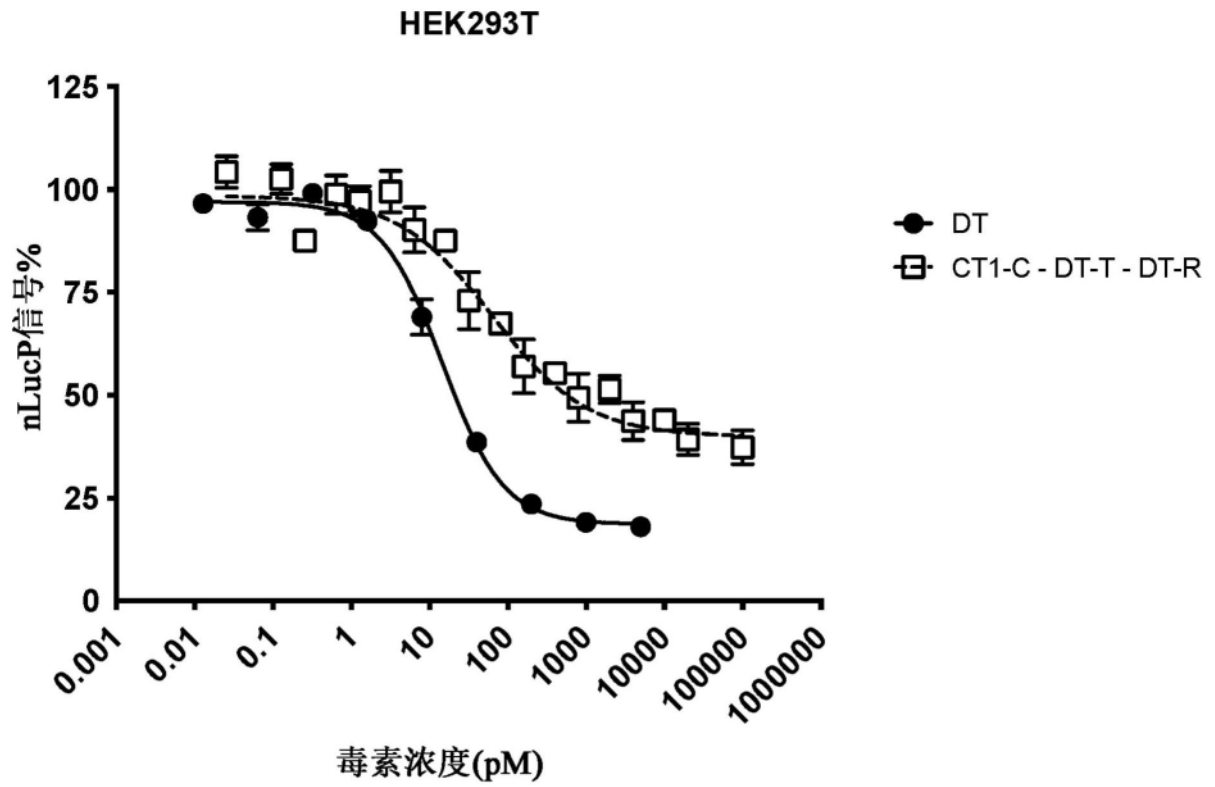


图2A

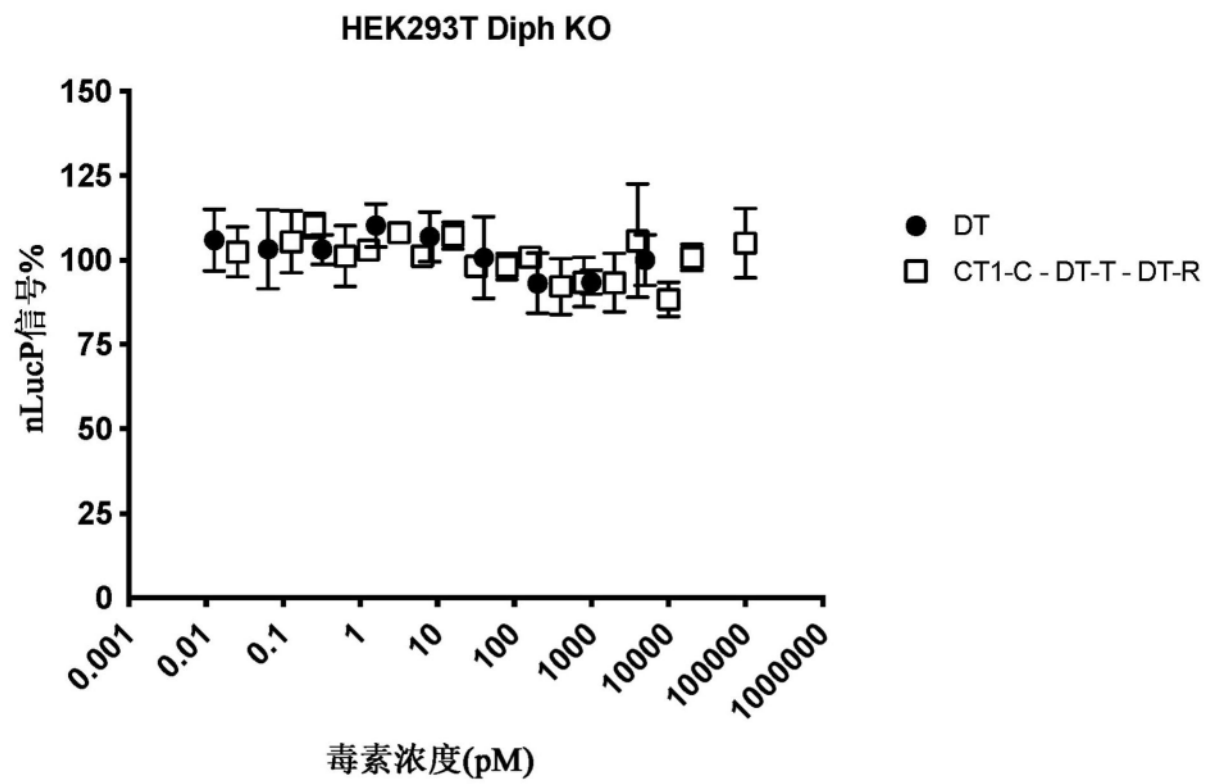


图2B

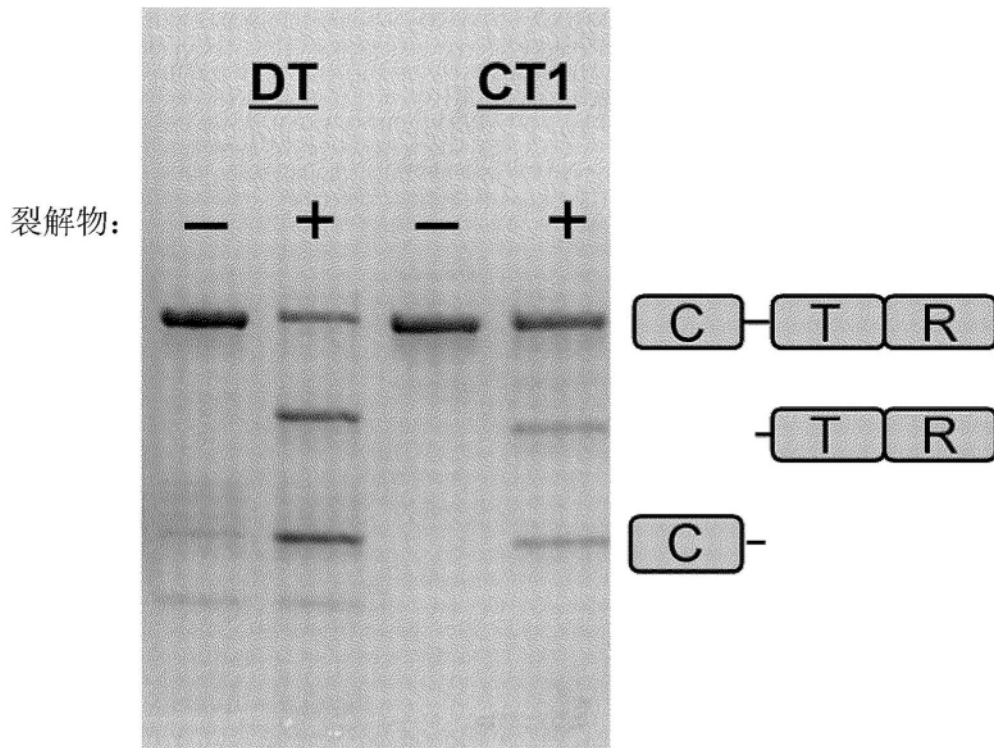


图2C

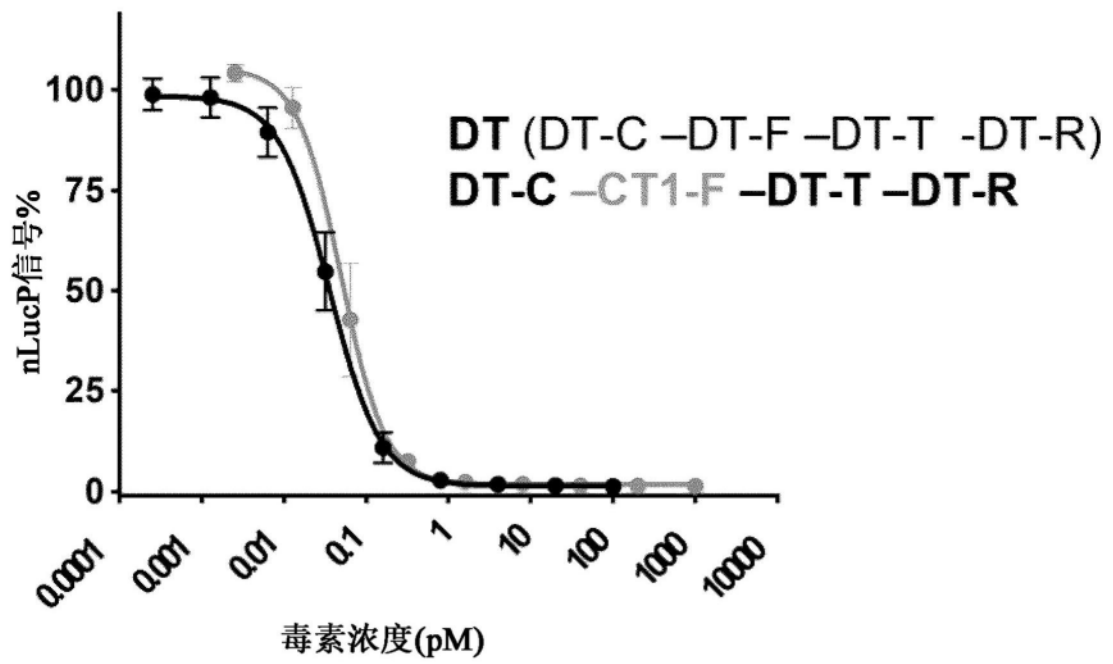


图3A

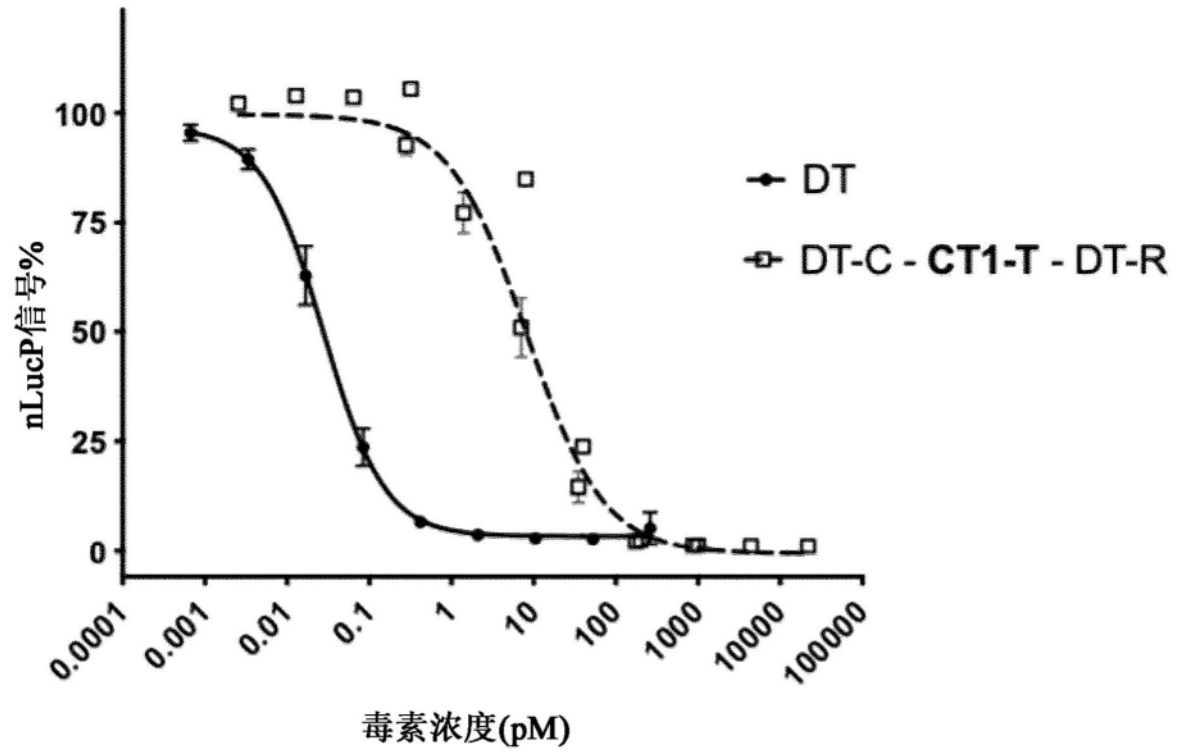


图3B

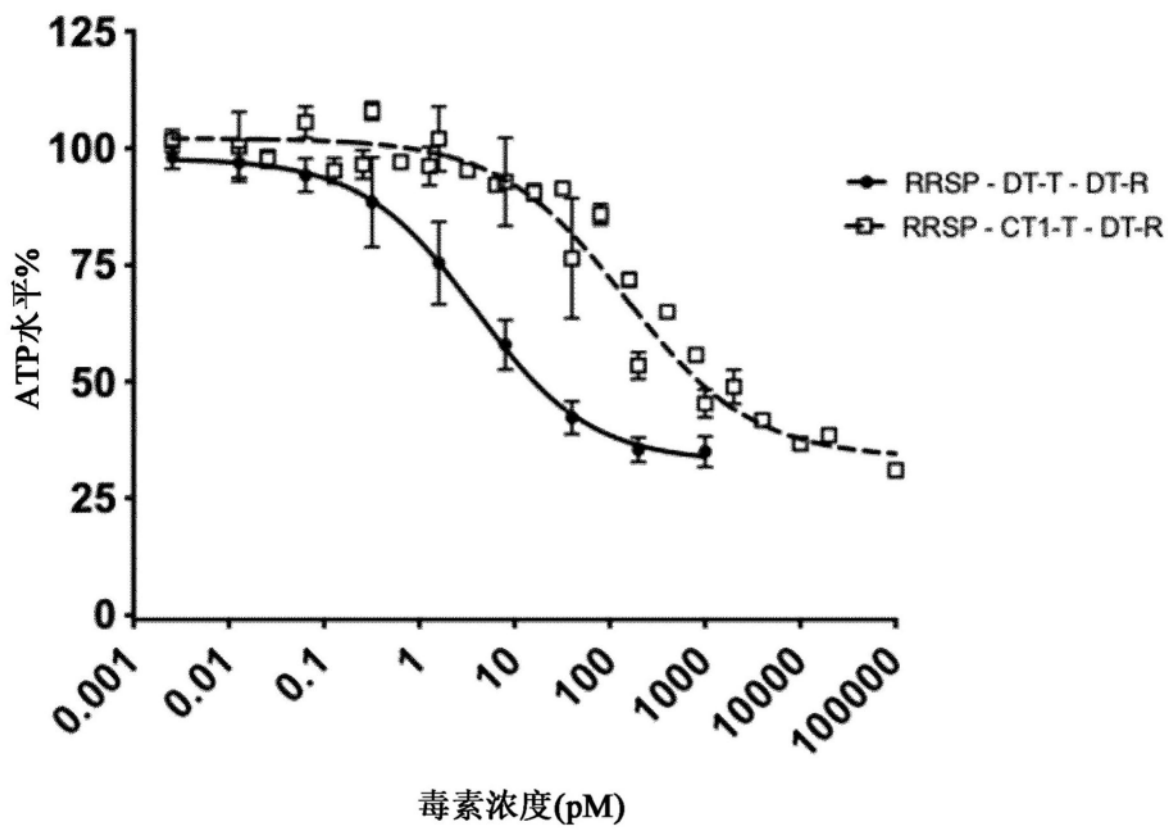


图3C

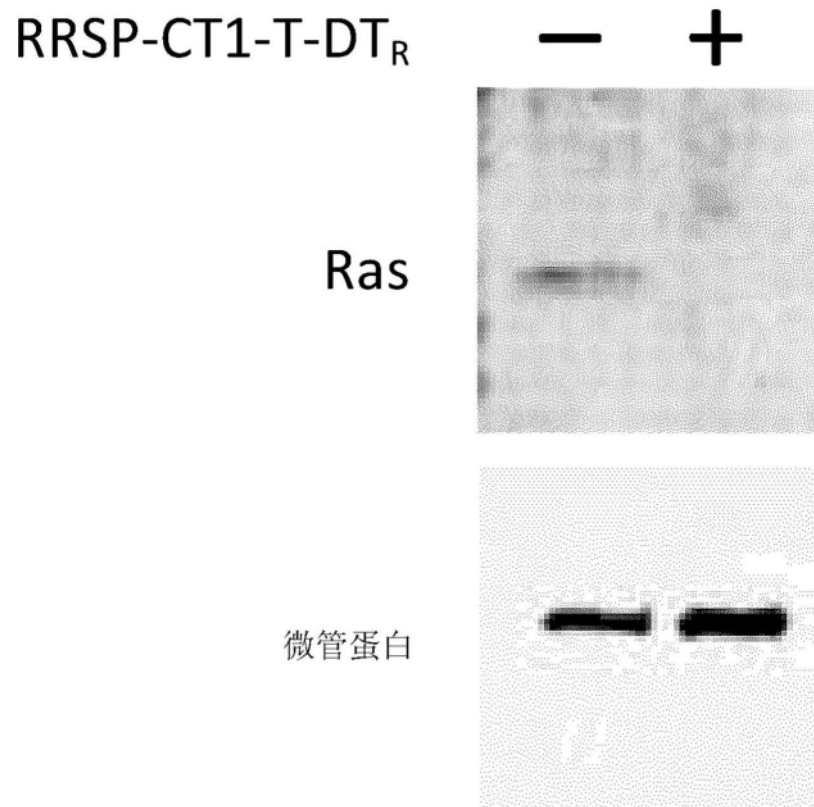


图3D

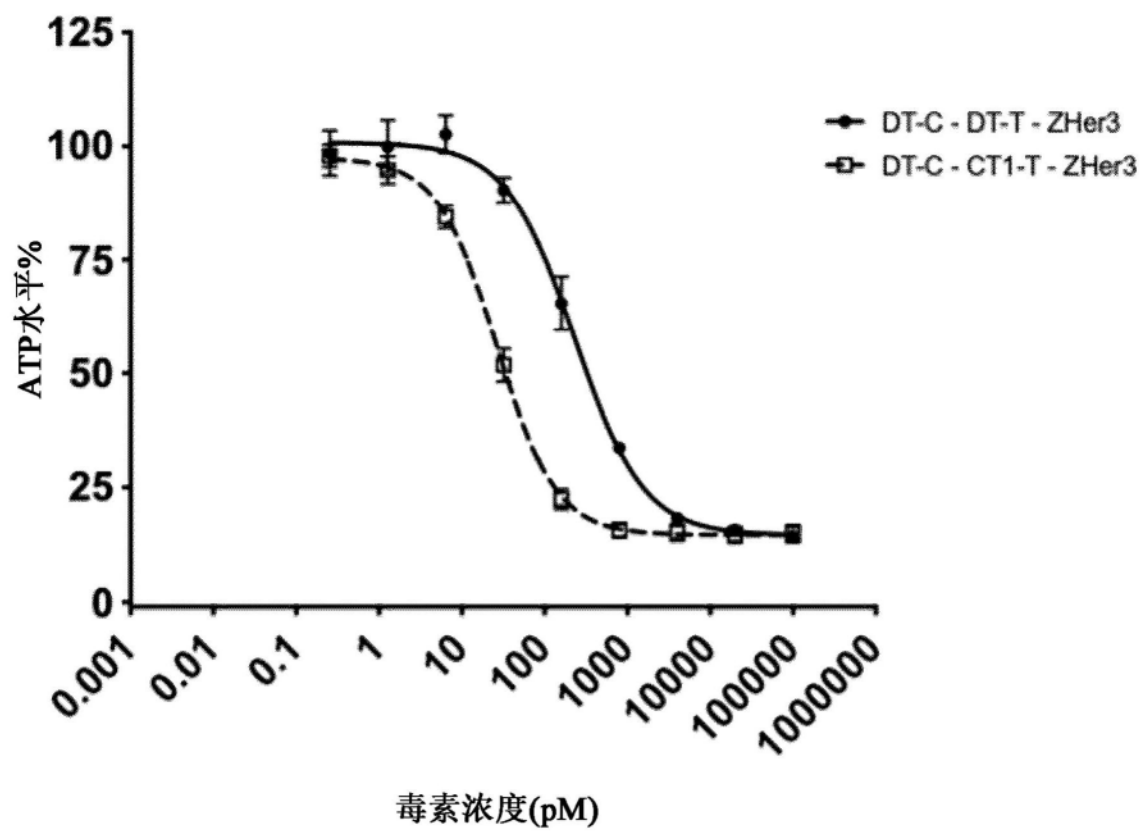


图3E

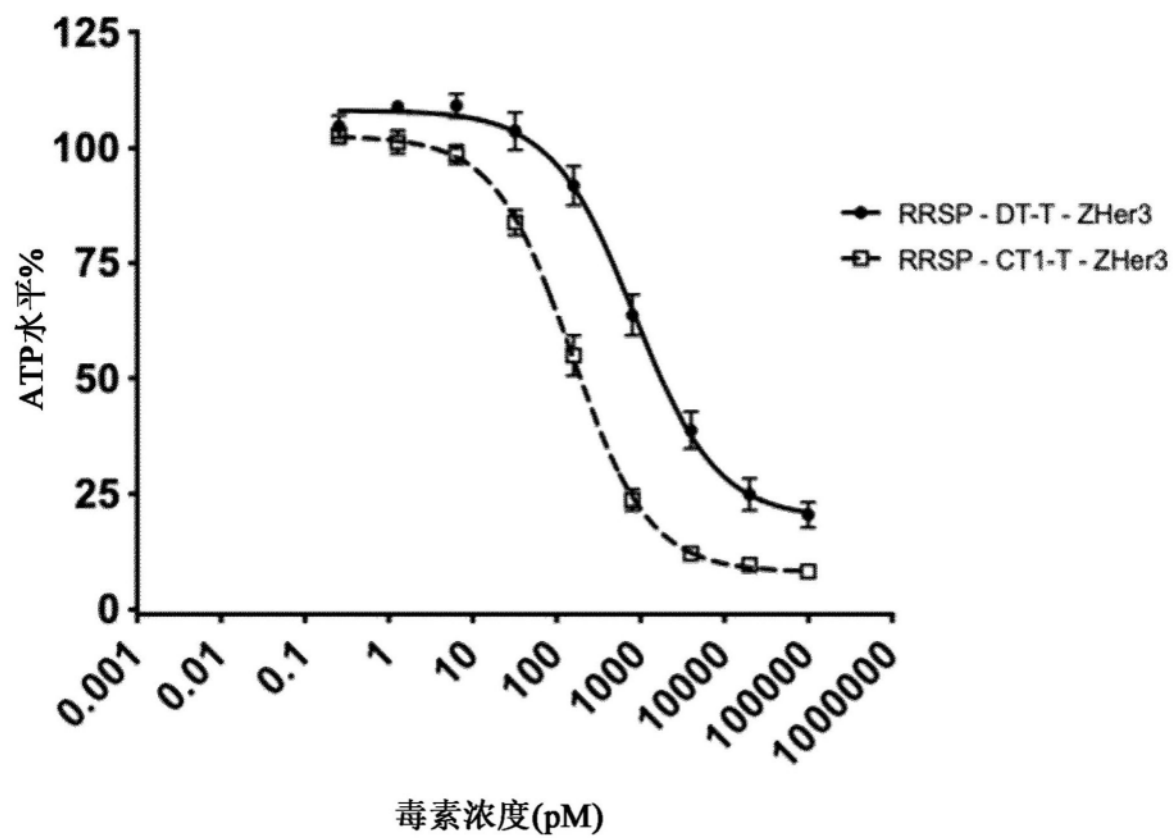


图3F

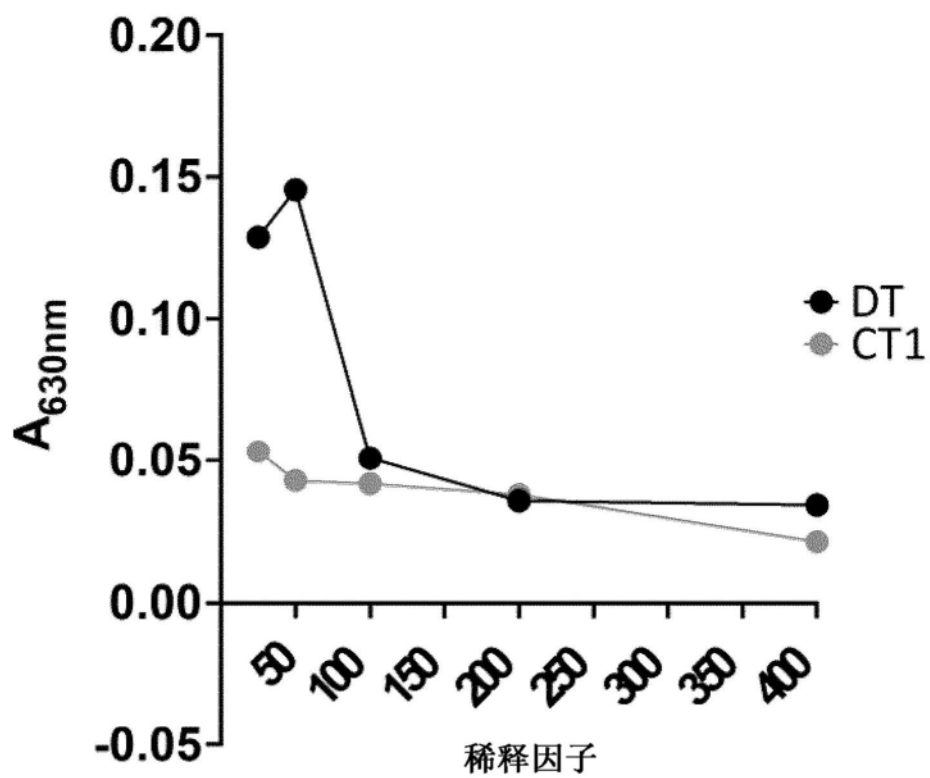


图4A

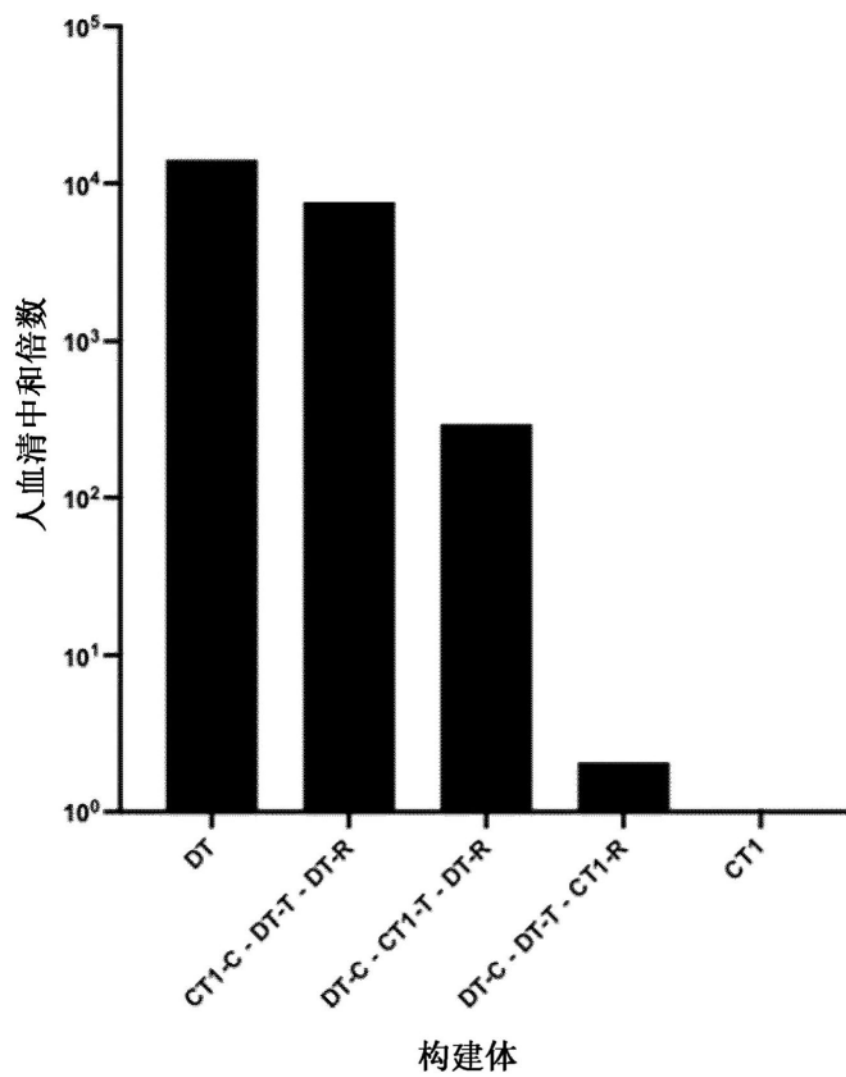
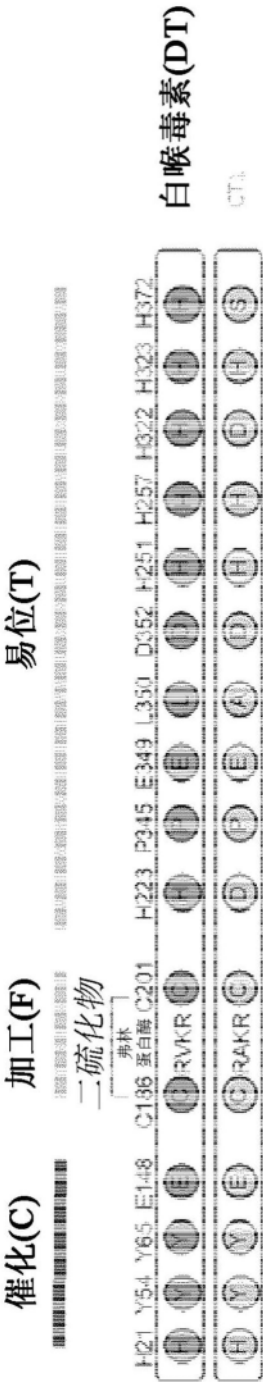


图4B



白喉毒素(DT)
CTA

图5

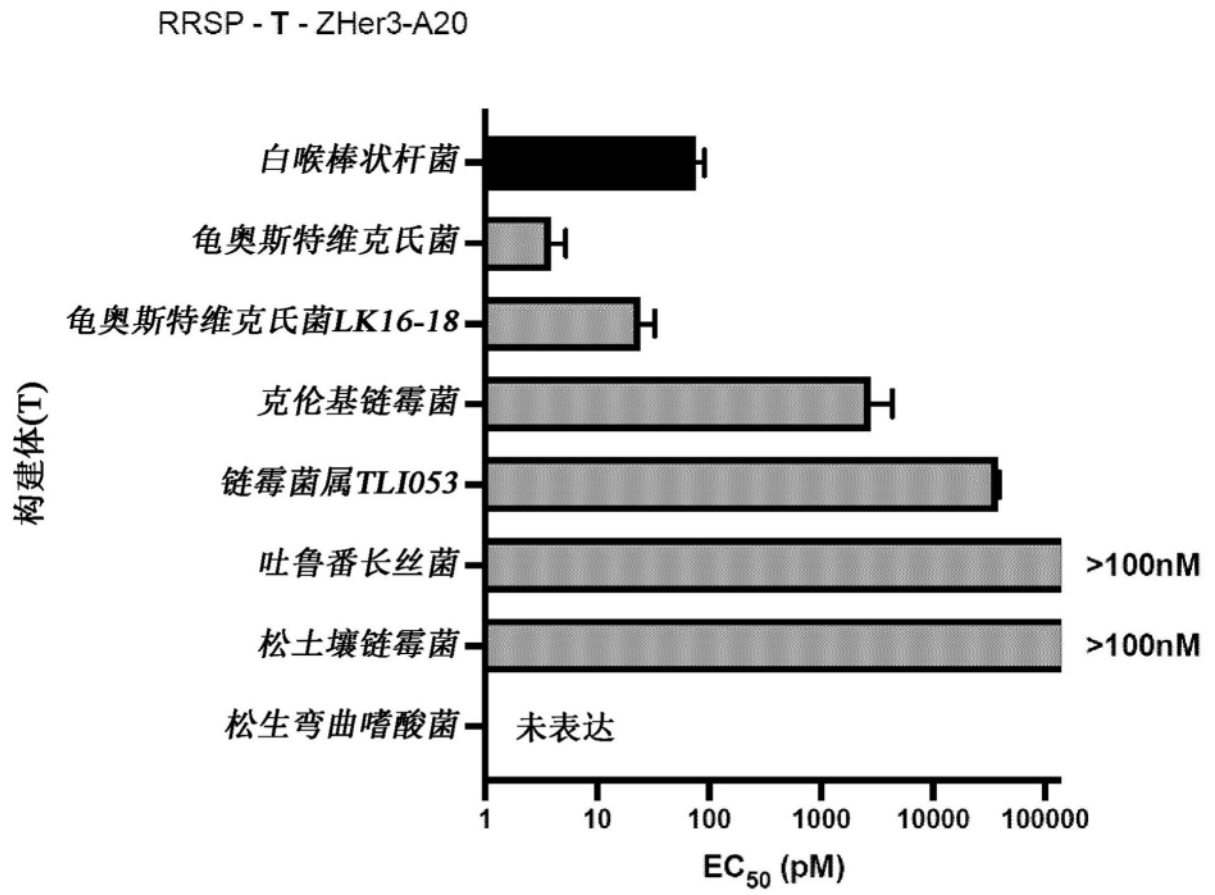


图6A

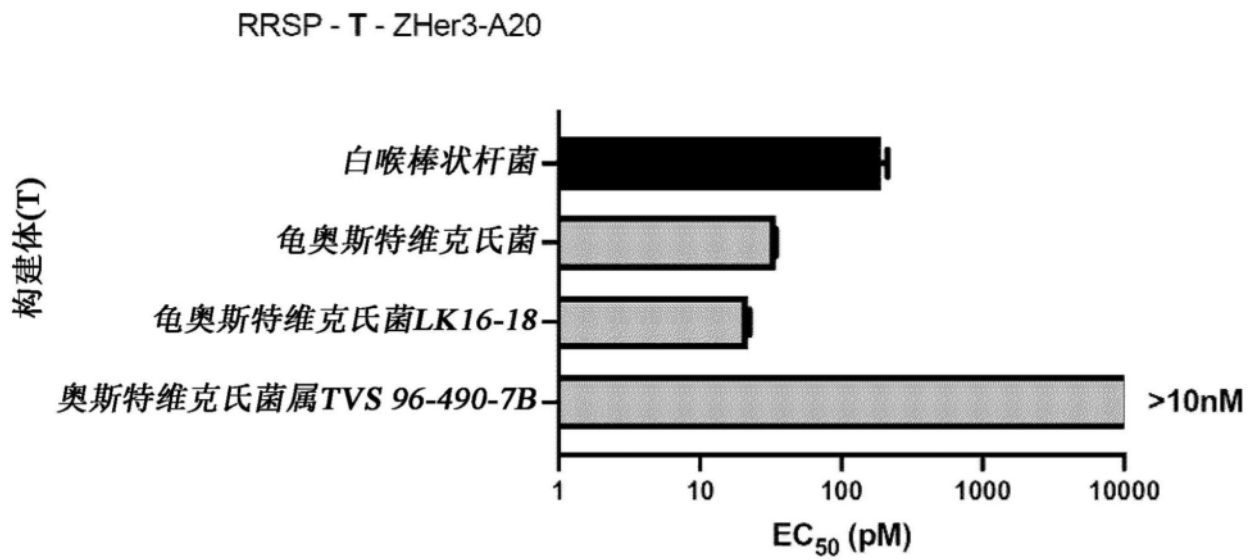


图6B

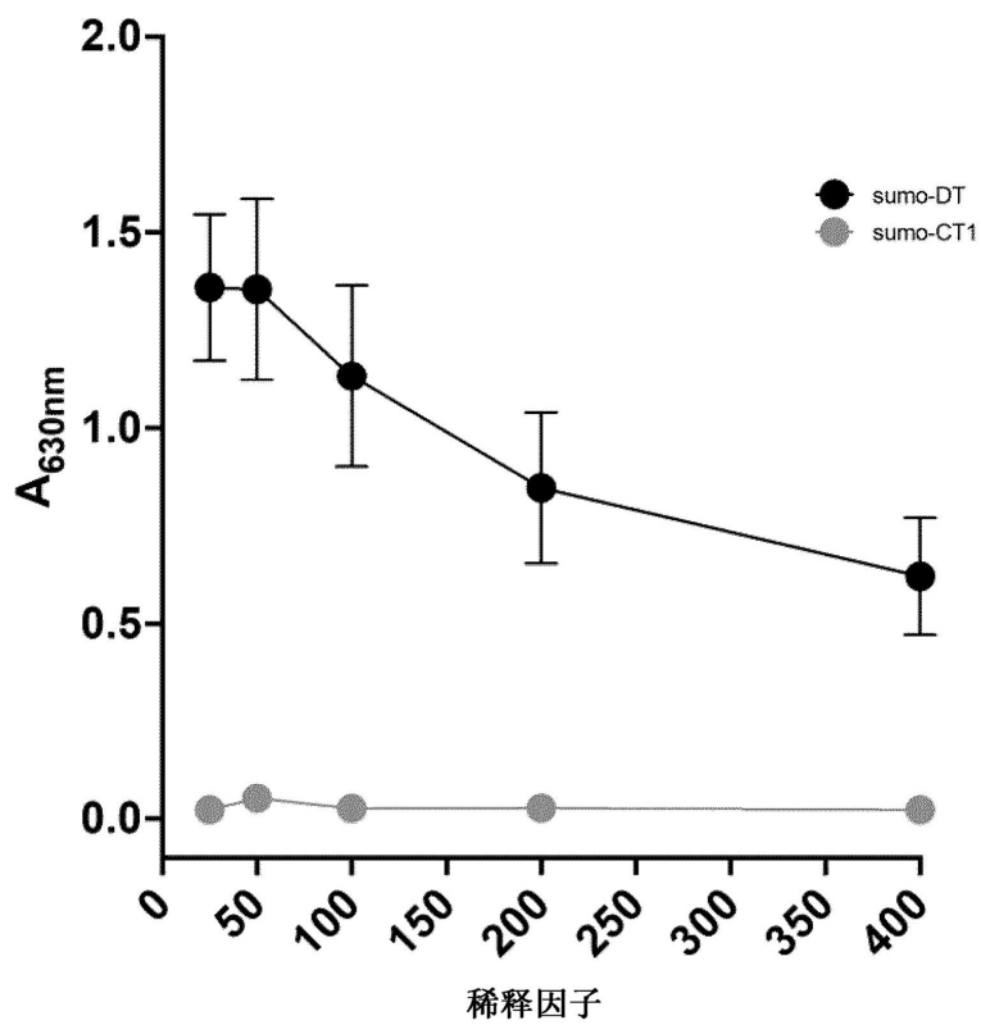


图7A

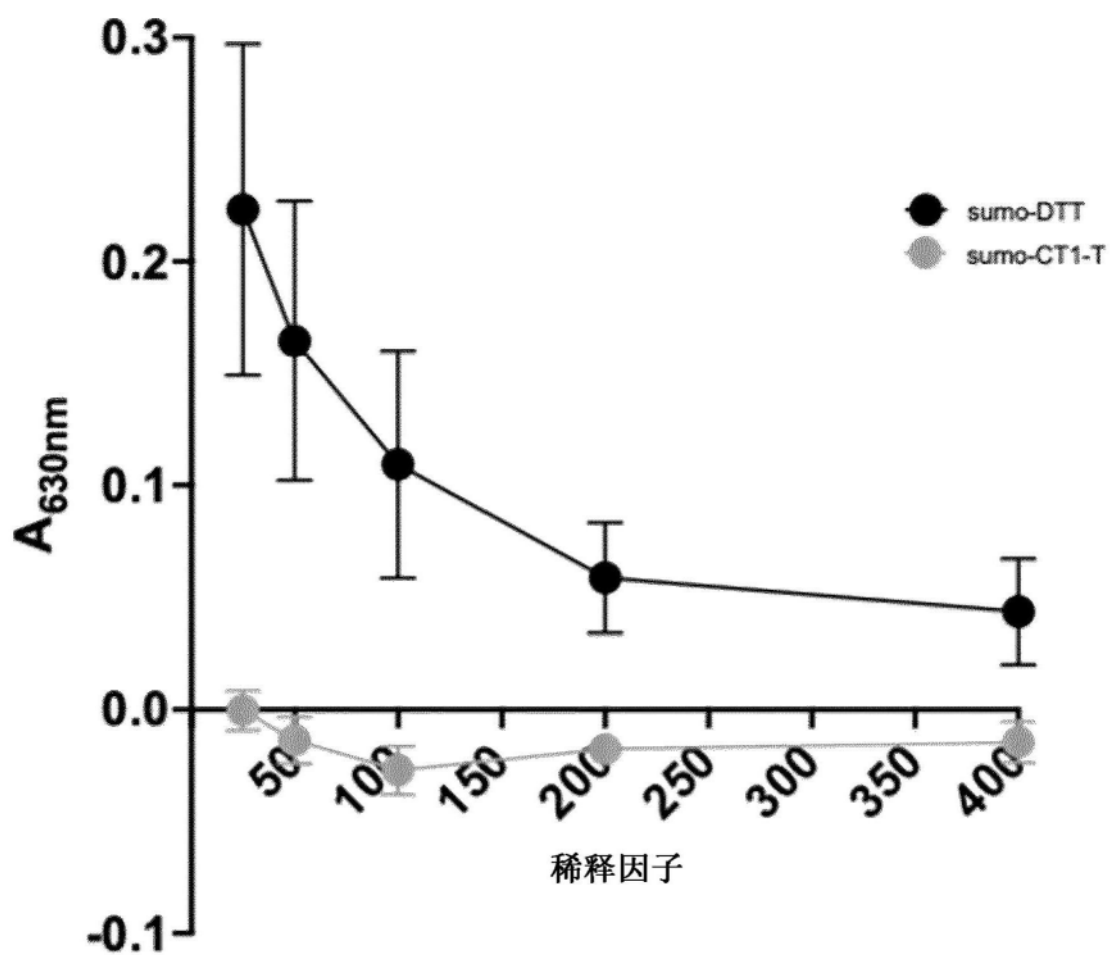


图7B

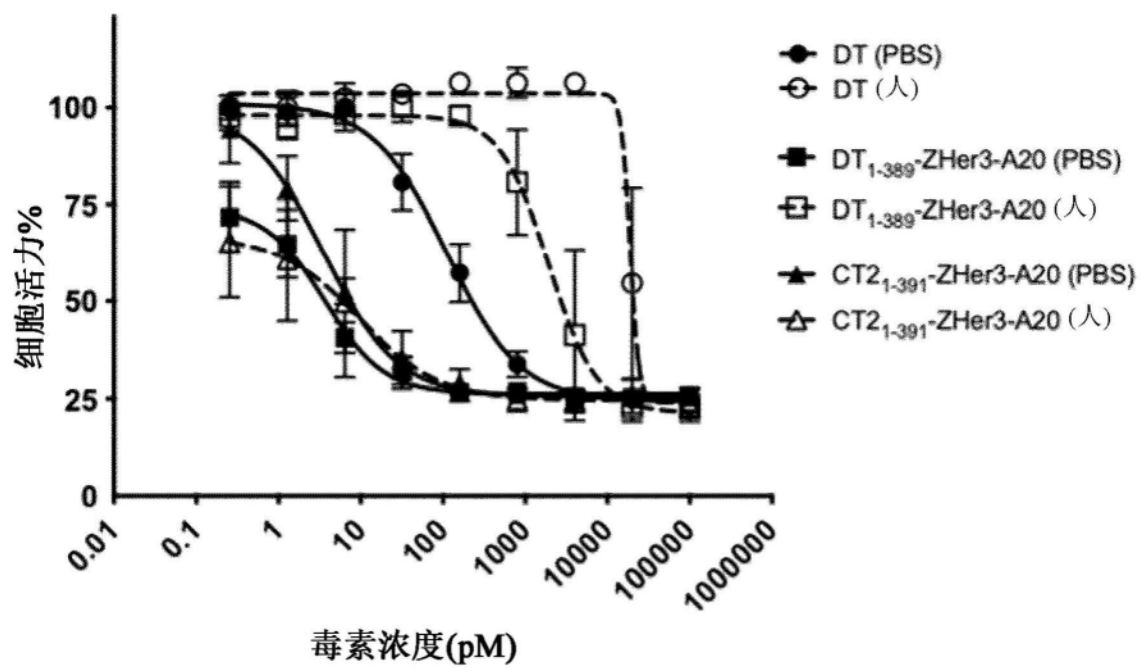


图8

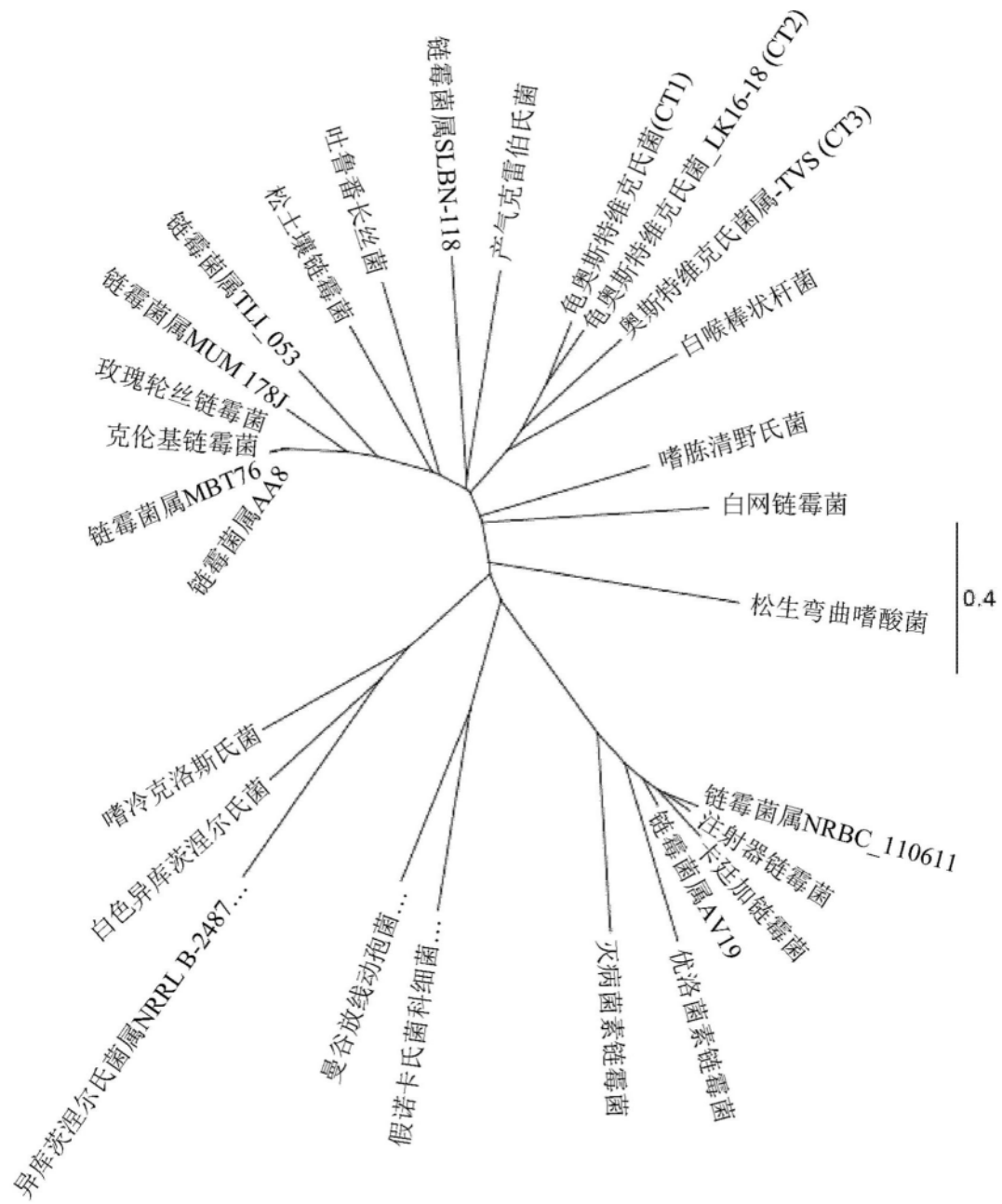


图9