



(12)发明专利申请

(10)申请公布号 CN 106279433 A

(43)申请公布日 2017.01.04

(21)申请号 201610670405.5

(51)Int.Cl.

(22)申请日 2012.03.02

*C07K 19/00*(2006.01)

(30)优先权数据

*C07K 7/08*(2006.01)

61/449,225 2011.03.04 US

*C07K 14/00*(2006.01)

PCT/US2011/041928 2011.06.25 US

*C12N 9/24*(2006.01)

13/168,969 2011.06.25 US

*C12N 15/62*(2006.01)

(62)分案原申请数据

*C12N 15/11*(2006.01)

201280020811.7 2012.03.02

*A61K 38/47*(2006.01)

(71)申请人 夏尔人类遗传治疗公司

*A61K 47/48*(2006.01)

地址 美国马萨诸塞州

*A61P 3/00*(2006.01)

*A61P 25/00*(2006.01)

(72)发明人 保罗·马蒂尼 米歇尔·孔奇诺

(74)专利代理机构 北京市君合律师事务所

11517

代理人 顾云峰 吴龙瑛

权利要求书4页 说明书12页

序列表20页 附图6页

(54)发明名称

用于多肽组成物的肽连接物及其使用方法

(57)摘要

GAPGGGGGAAAAAGCGGG (SEQ ID NO. 1)

本文公开了新的肽连接物和包含所述连接物的多肽组成物(例如嵌合多肽)以及所述多肽组成物的使用方法。所述组成物和方法对于向目标细胞、组织或器官靶向/递送目标多肽或蛋白(例如治疗性多肽)以便治疗各种疾病或障碍(例如溶酶体贮积症)来说,特别有用。

1. 一种多肽组成物,其包含:
  - a) 第一肽;
  - b) 第二肽;以及
  - c) 布置在所述第一肽与所述第二肽之间的连接物,所述连接物包含SEQ ID NO.1的氨基酸序列的一个或多个顺序重复序列。
2. 权利要求1的多肽组成物,其中所述连接物SEQ ID NO.1的3个顺序重复序列。
3. 权利要求1或2的多肽组成物,其中所述连接物还在SEQ ID NO.1的3'末端处包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP)。
4. 权利要求3的多肽组成物,其中所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。
5. 权利要求1至4任一项的多肽组成物,其中所述第一肽包含SEQ ID NO.4的氨基酸序列。
6. 权利要求1至5任一项的多肽组成物,其中所述第二肽包含受体结合结构域。
7. 权利要求1至6任一项的多肽组成物,其中所述第二肽包含SEQ ID NO.6的氨基酸序列。
8. 一种多肽组成物,其包含:
  - a) 包含SEQ ID NO.4的氨基酸序列的第一肽;
  - b) 包含SEQ ID NO.6的氨基酸序列的第二肽;以及
  - c) 布置在所述第一肽与所述第二肽之间的连接物,所述连接物包含SEQ ID NO.1的氨基酸序列的一个或多个顺序重复序列。
9. 权利要求8的多肽组成物,其中所述连接物包含SEQ ID NO.1的3个顺序重复序列。
10. 权利要求8或9的多肽组成物,其中所述连接物还在SEQ ID NO.1的3'末端处包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP)。
11. 权利要求8至10任一项的多肽组成物,其中所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。
12. 一种多肽组成物,其包含:
  - a) 包含SEQ ID NO.4的氨基酸序列的第一肽;
  - b) 包含SEQ ID NO.6的氨基酸序列的第二肽;以及
  - c) 布置在所述第一肽与所述第二肽之间的连接物,所述连接物包含SEQ ID NO.2的氨基酸序列。
13. 权利要求12的多肽组成物,其中所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。
14. 一种多肽连接物,其包含18个连续氨基酸残基,其中所述连接物在所述连接物的前5个氨基酸残基内包含1个脯氨酸残基,并包含选自一个或多个甘氨酸残基和一个或多个丙氨酸残基的17个氨基酸残基。
15. 权利要求14的多肽连接物,其中所述连接物的3'末端还包含3个连续氨基酸,所述3个连续氨基酸包含1个脯氨酸残基和选自甘氨酸和丙氨酸的2个氨基酸残基。
16. 一种多肽连接物,其包含21个连续氨基酸残基,其中所述连接物包含:
  - a) 所述连接物的前5个氨基酸内的第一脯氨酸残基;

- b)选自一个或多个甘氨酸残基和一个或多个丙氨酸残基的19个氨基酸残基;以及
- c)所述连接物的后5个氨基酸内的第二脯氨酸残基。

17. 权利要求14至16任一项的多肽连接物,其中所述一个或多个丙氨酸残基的任一个被一个或多个丝氨酸残基取代。

18. 权利要求17的多肽连接物,其中所述连接物包含的甘氨酸和丝氨酸残基数量为丙氨酸残基的两倍。

19. 一种多肽连接物,其由SEQ ID NO.7的氨基酸序列(GGGGGAAAAAGGGGG)构成。

20. 一种多肽连接物组成物,其包含权利要求19的多肽连接物的一个或多个顺序重复序列。

21. 权利要求20的多肽连接物组成物,其中所述多肽连接物组成物的5'末端还包含3个连续氨基酸,所述3个连续氨基酸包含1个脯氨酸残基和选自甘氨酸和丙氨酸的2个氨基酸残基。

22. 一种多肽连接物,其由SEQ ID NO.1的氨基酸序列(GAPGGGGAAAAAGGGGG)构成。

23. 一种多肽连接物组成物,其包含权利要求22的多肽连接物的一个或多个顺序重复序列。

24. 权利要求21或23的多肽连接物组成物,其中所述多肽连接物组成物的3'末端还包含3个连续氨基酸,所述3个连续氨基酸包含1个脯氨酸残基和选自甘氨酸和丙氨酸的2个氨基酸残基。

25. 权利要求21、23或24的多肽连接物组成物,其中所述3个连续氨基酸包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP)。

26. 一种多肽连接物,其由SEQ ID NO.2的氨基酸序列(GAPGGGGAAAAAGGGGGAPGGGGAAAAAGGGGGAPGGGGAAAAAGGGGGGAP)构成。

27. 一种多肽组成物,其包含:

a)第一肽;

b)第二肽;以及

c)布置在所述第一肽与所述第二肽之间的连接物,其中所述连接物包含权利要求14至26的多肽连接物或多肽连接物组成物中的任一种。

28. 一种多核苷酸,其编码包含下列氨基酸序列的多肽连接物:

a)SEQ ID NO.1;

b)SEQ ID NO.2;或

c)SEQ ID NO.1的一个或多个顺序重复序列。

29. 权利要求23的多核苷酸,其中所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代,并且所述多核苷酸编码所述一个或多个取代的丝氨酸残基。

30. 一种多核苷酸,其编码包含下列氨基酸序列的多肽组成物:

a)第一肽;

b)第二肽;和

c)布置在所述第一肽与所述第二肽之间的连接物,所述连接物包含SEQ ID NO.1的氨基酸序列的一个或多个顺序重复序列。

31. 权利要求30的多核苷酸,其中所述多肽组成物的所述连接物包含SEQ ID NO.1的3

个顺序重复序列。

32. 权利要求30或31的多核苷酸,其中所述多肽组成物的所述连接物的3'末端还包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP),并且所述多核苷酸编码所述GAP氨基酸序列。

33. 权利要求31的多核苷酸,其中所述多肽组成物的所述连接物包含SEQ ID NO.2。

34. 权利要求30至33的多核苷酸,其中所述多肽组成物的所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代,并且所述多核苷酸编码所述一个或多个取代的丝氨酸残基。

35. 权利要求30至34任一项的多核苷酸,其中所述多肽组成物的所述第一肽包含SEQ ID NO.4的氨基酸序列。

36. 权利要求30至35任一项的多核苷酸,其中所述多肽组成物的所述第二肽包含受体结合结构域。

37. 权利要求28至36任一项的多核苷酸,其中所述第二肽包含SEQ ID NO.6的氨基酸序列。

38. 一种表达载体,其包含权利要求28至37任一项的多核苷酸。

39. 一种重组细胞,其包含权利要求28至37任一项的多核苷酸。

40. 一种重组细胞,其包含权利要求38的表达载体。

41. 一种药物组合物,其包含权利要求1至7或27任一项的多肽组成物和可药用载体。

42. 一种药物组合物,其包含权利要求8至13任一项的多肽组成物。

43. 一种药物组合物,其包含权利要求28至34任一项的多核苷酸和可药用载体。

44. 一种药物组合物,其包含权利要求35至37任一项的多核苷酸。

45. 一种药物组合物,其包含权利要求39或40的重组细胞和可药用载体。

46. 一种生产多肽组成物的方法,所述方法包括将权利要求39或40的重组细胞在适合于所述多肽表达的条件下培养。

47. 一种向需要的对象递送治疗性多肽的方法,所述方法包括向所述对象给药权利要求1至13或权利要求27的多肽组成物中的任一种。

48. 一种向需要的对象递送治疗性多肽的方法,所述方法包括向所述对象给药包含下列组分的多肽组成物:

a) 包含SEQ ID NO.4的氨基酸序列的第一肽;

b) 包含SEQ ID NO.6的氨基酸序列的第二肽;以及

c) 布置在所述第一肽与所述第二肽之间的连接物,所述连接物包含SEQ ID NO.1的氨基酸序列的一个或多个顺序重复序列。

49. 权利要求48的方法,其中所述连接物还在SEQ ID NO.1的3'末端处包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP)。

50. 权利要求48的方法,其中所述连接物包含SEQ ID NO.2的氨基酸序列。

51. 权利要求48至50任一项的方法,其中所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。

52. 一种向需要的对象递送治疗性多肽的方法,所述方法包括向所述对象给药一种或多种权利要求38的表达载体。

53. 一种向需要的对象递送治疗性多肽的方法,所述方法包括向所述对象给药权利要

求39或40的一种或多种重组细胞。

54.一种用于治疗溶酶体贮积症的方法,所述方法包括向需要的对象给药有效量的权利要求41至45的药物组合物中的任一种。

55.权利要求54的方法,其中所述溶酶体贮积症是沙费利波综合征。

56.一种治疗沙费利波综合征的方法,所述方法包括向需要的患者给药有效量的权利要求42至45的药物组合物中的任一种。

57.一种治疗沙费利波综合征的方法,所述方法包括向需要的患者给药有效量的包含多肽组成物的药物组合物,所述多肽组成物包含:

a)包含SEQ ID NO.4的氨基酸序列的第一肽;

b)包含SEQ ID NO.6的氨基酸序列的第二肽;以及

c)布置在所述第一肽与所述第二肽之间的连接物,所述连接物包含SEQ ID NO.1的氨基酸序列的一个或多个顺序重复序列。

58.权利要求57的方法,其中所述连接物还在SEQ ID NO.1的3'末端处包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP)。

59.权利要求57的方法,其中所述连接物包含SEQ ID NO.2的氨基酸序列。

60.权利要求54至59任一项的方法,其中所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。

61.权利要求54至60任一项的方法,其中所述药物组合物经肠胃外给药。

62.权利要求54至60任一项的方法,其中所述药物组合物经口服给药。

## 用于多肽组成物的肽连接物及其使用方法

[0001] 相关申请

[0002] 本申请是申请号为201280020811.7、申请日为2012年3月2日、发明名称为“用于多肽组成物的肽连接物及其使用方法”的中国发明专利申请的分案申请。原申请为国际申请号为PCT/US2012/027561的国家阶段申请,该国际申请要求申请日为2011年3月4日,申请号为61/449,225的美国临时申请的优先权,并且是2011年6月25日提交的美国申请号13/168,969和2011年6月25日提交的国际申请号PCT/US2011/041928的部分连续案。所述申请的全部教示内容在此引为参考。

### 发明领域

[0003] 本发明涉及新的肽连接物和包含这样的连接物的多肽组成物(例如嵌合多肽)及其使用和制备方法。

[0004] 发明背景

[0005] 溶酶体贮积症是由特定溶酶体酶的缺乏或失活引起的一类超过40种罕见且遗传的代谢障碍。具体来说,溶酶体贮积症由催化脂类或被称为糖胺聚糖的复杂糖蛋白的逐步代谢的溶酶体酶的缺乏或失活引起。作为这种代谢缺陷的结果,代谢前体在患病对象的细胞、组织以及特别是细胞溶酶体中逐步积累。这种蛋白积累引起永久性、渐进性的细胞损伤,其影响患病对象的外表、身体能力、器官和系统功能以及在大多数病例中智力发育。尽管酶的缺乏影响每个组织,但患者的临床表现通常可能随着例如酶缺乏或缺损的程度而变。溶酶体贮积症还可能与一定程度的神经元细胞丧失相关,主要引起神经症状包括智力迟钝、严重运动缺损、身体失能、寿命缩短和/或上述的组合。

[0006] 溶酶体贮积症没有治愈方法,治疗通常是治标措施,提供给对象主要是为了提高他们的生活质量。对于患有溶酶体贮积症的对象来说,酶替代疗法(ERT)是一种有用的治疗选项。ERT一般包括向患者肠胃外给药天然或重组产生的蛋白和/或酶。批准的疗法被静脉内给药于患者,并且一般对于治疗由根本的酶缺乏引起的身体和外周症状有效。为了有效地治疗溶酶体贮积症,所给药的治疗剂(例如缺乏的溶酶体酶)必须在被输注到患者血流中后分配到患病细胞和组织中。

[0007] 为了实现必需的酶在患病细胞和组织中的分配,一般将酶靶向到特异性细胞表面受体,所述受体通过受体介导的内吞作用将酶运输到细胞中。例如,在高歇氏病中,将缺乏的酶、葡萄糖脑苷脂酶通过酶上暴露的甘露糖残基与靶细胞(网状内皮细胞)上丰富表达的甘露糖受体的结合,靶向到适合的细胞。在缺乏甘露糖受体的细胞中,已提出使用不依赖胰岛素样生长因子/阳离子的甘露糖-6-磷酸受体(IGF-II/CI-MPR)将缺乏的溶菌酶递送至细胞(Kornfeld, S., 1987 Biochem Soc Trans 18:367-374)。IGF-II/CI-MPR受体存在于许多哺乳动物细胞类型的表面上,因此提供了将含有受体配体(例如IGFII或甘露糖-6-磷酸)的蛋白靶向到广泛的各种细胞和组织、包括中枢神经系统的手段。然而,尽管已对如何将缺失的溶酶体酶靶向到适合的组织有一定了解,但对许多溶酶体贮积症(例如沙费利波(Sanfilippo)综合征、法伯氏(Farber's)病等)仍没有有效的疗法。因此,在本领域中对于

可用于将药剂导向必需组织以治疗疾病(例如溶酶体贮积症)的组合物、特别是可以肠胃外给药的组合物以及方法,仍存在着需求。

[0008] 发明简述

[0009] 对于促进有功能的治疗剂(例如蛋白、多肽)向目标组织的运输和递送的组合物和方法,存在着需求。这样的组合物和方法可用于治疗大量疾病或障碍,特别是治疗溶酶体贮积症如沙费利波病。

[0010] 本文中描述了包含肽连接物的新的组合物、包含通过所述肽连接物联结的多肽的多肽组成物和相关的多核苷酸、载体、细胞和药物组合物。所描述的连接物序列可操作地联结两个目标肽/多肽,使得通过所述连接物连接的多肽的表达和活性(例如受体结合和/或酶活性)持久且最优。包含所述肽连接物的多肽组成物促进目标多肽/蛋白向特定细胞和/或组织的定向递送。

[0011] 因此,本发明的实施方式提供了一种多肽组成物,其包含第一肽/多肽、第二肽/多肽以及放置在所述第一肽与第二肽之间的连接物,所述连接物包含一个或多个顺序或串联重复的SEQ ID NO.1的氨基酸序列(GAPGGGGAAAAAGGGGG)。在某些实施方式中,所述多肽组成物的连接物包含3个顺序或串联重复的SEQ ID NO.1序列,并且在某些实施方式中,所述连接物还在SEQ ID NO.1的3'末端处包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP)。在其他实施方式中,所述连接物的一个或多个丙氨酸残基可以被一个或多个丝氨酸残基取代。在某些实施方式中,所述多肽组成物的第一肽包含SEQ ID NO.4的氨基酸序列。在其他实施方式中,所述第二肽包含受体结合结构域,并且在其他实施方式中,所述第二肽包含SEQ ID NO.6的氨基酸序列。

[0012] 在某些实施方式中,多肽组成物包含:第一肽,其包含SEQ ID NO.4的氨基酸序列;第二肽,其包含SEQ ID NO.6的氨基酸序列;以及布置在所述第一肽与所述第二肽之间的连接物,其包含一个或多个顺序或串联重复的SEQ ID NO.1的氨基酸序列。在某些实施方式中,所述连接物包含SEQ ID NO.1的3个顺序重复序列。在其他实施方式中,所述连接物还可以在SEQ ID NO.1的3'末端处包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP)。在其他实施方式中,所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。

[0013] 在某些实施方式中,多肽组成物包含:第一肽,其包含SEQ ID NO.4的氨基酸序列;第二肽,其包含SEQ ID NO.6的氨基酸序列;以及布置在所述第一肽与第二肽之间的连接物,其包含SEQ ID NO.2的氨基酸序列。在某些实施方式中,所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。

[0014] 本发明还提供了本文中描述的肽/多肽连接物。例如,在某些实施方式中,多肽连接物包含18个连续氨基酸残基,其中所述连接物在所述连接物的前5个氨基酸残基中包含1个脯氨酸残基,并包含选自一个或多个甘氨酸残基和一个或多个丙氨酸残基的17个氨基酸残基。在其他实施方式中,所述多肽连接物的3'末端还包含3个连续氨基酸,其包含1个脯氨酸残基和选自甘氨酸和丙氨酸的2个氨基酸残基。在某些实施方式中,所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。

[0015] 在某些实施方式中,多肽连接物包含21个连续氨基酸残基,其中所述连接物包含在所述连接物的前5个氨基酸残基中的第一脯氨酸残基;选自一个或多个甘氨酸残基和一个或多个丙氨酸残基的19个氨基酸残基;以及在所述连接物的后5个氨基酸中的第二脯氨酸残基。

酸残基。

[0016] 在某些实施方式中,所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。在其他实施方式中,所述多肽连接物包含的甘氨酸和丝氨酸残基的数量为丙氨酸残基的两倍。

[0017] 在某些实施方式中,所述多肽连接物由 SEQ ID NO.7 的氨基酸序列 (GGGGGAAAAAGGGGG) 构成。

[0018] 在其他实施方式中,本发明提供了多肽连接物组成物,其包含 SEQ ID NO.7 的氨基酸序列的一个或多个顺序重复序列构成的多肽连接物。在其他实施方式中,所述多肽连接物组成物的 5' 末端还包含 3 个连续氨基酸,其包含 1 个脯氨酸残基和选自甘氨酸和丙氨酸的 2 个氨基酸残基。

[0019] 在其他实施方式中,多肽连接物由 SEQ ID NO.1 的氨基酸序列构成。在某些实施方式中,本发明提供了一种多肽连接物组成物,其包含、由 SEQ ID NO.1 的一个或多个顺序重复序列构成的多肽连接物。

[0020] 在其他实施方式中,上述多肽连接物组成物的 3' 末端还包含 3 个连续氨基酸,其包含 1 个脯氨酸残基和选自甘氨酸和丙氨酸的 2 个氨基酸残基。在某些实施方式中,所述 3 个连续氨基酸残基包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP)。

[0021] 在某些实施方式中,多肽连接物由 SEQ ID NO.2 的氨基酸序列 (GAPGGGGGAAAAAGGGGGGAPGGGGGAAAAAGGGGGGAPGGGGGAAAAAGGGGGGAP) 构成。

[0022] 在某些实施方式中,本发明提供了一种多肽组成物,其包含:第一肽;第二肽;以及布置在所述第一肽与所述第二肽之间的多肽连接物,其中所述连接物包含任何上述多肽连接物或多肽连接物组成物。

[0023] 在其他实施方式中,本发明还提供了编码本文中描述的多肽连接物和/或多肽组成物的多核苷酸。因此,某些实施方式提供了编码以下多肽的多核苷酸:包含 SEQ ID NO.1 的氨基酸序列的多肽;包含 SEQ ID NO.2 的氨基酸序列的多肽;或者包含 SEQ ID NO.1 的一个或多个顺序重复序列的氨基酸序列的多肽。在某些实施方式中,所述多肽的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代,并且所述多核苷酸编码所述一个或多个取代的丝氨酸残基。

[0024] 在其他实施方式中,多核苷酸编码多肽组成物,其包含下列氨基酸序列:第一肽;第二肽;以及布置在所述第一肽和所述第二肽之间的连接物,其包含 SEQ ID NO.1 的氨基酸序列的一个或多个顺序重复序列。在某些实施方式中,所述多肽组成物的连接物包含 3 个顺序或串联重复的 SEQ ID NO.1 序列。在其他实施方式中,所述多肽组成物的连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代,并且所述多核苷酸编码所述一个或多个取代的丝氨酸残基。在上述多核苷酸的其他实施方式中,所述多肽组成物的连接物的 3' 末端还包含氨基酸序列甘氨酸-丙氨酸-脯氨酸(GAP),并且所述多核苷酸编码所述 GAP 氨基酸序列。

[0025] 在某些实施方式中,多核苷酸编码所述多肽组成物的第一肽,其包含 SEQ ID NO.4 的氨基酸序列。在其他实施方式中,所述多核苷酸编码所述多肽组成物的第二肽,其包含受体结合结构域,并且在某些实施方式中,所述第二肽包含 SEQ ID NO.6 的氨基酸序列。

[0026] 本文中描述还描述了包含上述编码本文中描述的多肽组成物的多核苷酸的表达载体。本



文描述的其他实施方式涉及包含上述多核苷酸和表达载体的重组细胞。

[0027] 此外,本发明提供了药物组合物,其包含本文描述的多肽组成物和可药用载体。例如,在某些实施方式中,所述药物组合物包含多肽组成物,其包含:包含SEQ ID NO.4的氨基酸序列的第一肽,以及包含SEQ ID NO.6的氨基酸序列的第二肽。在这些实施方式的某些中,布置在所述第一肽与第二肽之间的连接物包含一个或多个顺序或串联重复的SEQ ID NO.1序列(例如1、2、3、4、5、6、7、8、9、10个或更多个顺序或串联重复的SEQ ID NO.1序列),并且在这些实施方式的其他一些中,所述连接物包含SEQ ID NO.2的氨基酸序列。

[0028] 在某些实施方式中,本发明还提供了包含本文描述的多核苷酸和可药用载体的药物组合物。本发明还提供了包含本文描述的重组细胞和可药用载体的药物组合物。

[0029] 在本文中描述还描述了生产多肽组成物的方法,所述方法包括将本文描述的重组细胞在适合于所述多肽表达的条件下进行培养。

[0030] 本文中描述还描述了向需要的对象递送治疗性多肽的方法。因此,本文公开的某些实施方式是向需要的对象递送治疗性多肽的方法,所述方法包括向所述对象给药本文中描述的任何上述多肽组成物。

[0031] 在其他实施方式中,向需要的对象递送治疗性多肽的方法包括向所述对象给药多肽组成物,其包含:包含SEQ ID NO.4的氨基酸序列的第一肽;包含SEQ ID NO.6的氨基酸序列的第二肽;以及布置在所述第一肽与所述第二肽之间的连接物,其包含SEQ ID NO.1的氨基酸序列的一个或多个顺序重复序列。在所述方法的其他实施方式中,所述连接物还在SEQ ID NO.1的3'末端处包含GAP。在所述方法的某些实施方式中,所述连接物包含SEQ ID NO.2的氨基酸序列。在所述方法的其他实施方式中,所述连接物的一个或多个丙氨酸残基被一个或多个丝氨酸残基取代。

[0032] 在某些实施方式中,向需要的对象递送治疗性多肽的方法包括向所述对象给药一种或多种本文中描述的上述表达载体。某些其他实施方式涉及向需要的对象递送治疗性多肽的方法,所述方法包括向所述对象给药一种或多种本文中描述的上述重组细胞。

[0033] 本文中描述还描述了治疗溶酶体贮积症的方法,在某些实施方式中,所述方法包括向需要的对象给药有效量的本文中描述的任一种上述药物组合物(包括例如本文中描述的多肽组成物、多核苷酸和/或宿主细胞)。在某些实施方式中,所述溶酶体贮积症是沙费利波综合征。

[0034] 本文中描述的其他实施方式涉及治疗沙费利波综合征的方法。在某些实施方式中,所述方法包括向需要的患者给药有效量的本文中描述的上述药物组合物。

[0035] 在其他实施方式中,治疗沙费利波综合征的方法包括向需要的患者给药有效量的药物组合物,其包含:包含SEQ ID NO.4的氨基酸序列的第一肽;包含SEQ ID NO.6的氨基酸序列的第二肽;以及布置在所述第一肽和所述第二肽之间的连接物,其包含SEQ ID NO.1的一个或多个顺序重复序列。在所述方法的某些实施方式中,所述连接物还在SEQ ID NO.1的3'末端处包含氨基酸序列GAP。在所述方法的其他实施方式中,所述连接物包含SEQ ID NO.2的氨基酸序列。

[0036] 在所述方法的某些实施方式中,所述药物组合物肠胃外给药。

[0037] 通过结合随附的实施例参考下面的发明详述,本发明的上面讨论的和许多其他特点和伴随的优点将被更好地理解。本文描述的各种实施方式是互补的,并且可以根据本文

中包含的教导以专业技术人员所理解的方式组合或一起使用。

[0038] 附图简述

[0039] 图1示出了肽连接物的氨基酸序列,其包含联结到甘氨酸-丙氨酸-甘氨酸(GAG)重复序列的甘氨酸-丙氨酸-脯氨酸(GAP)序列(SEQ ID NO.1)。

[0040] 图2示出了肽连接物的氨基酸序列,其包含在3'末端处联结到GAP序列的3个顺序或串联重复的SEQ ID NO.1序列(SEQ ID NO.2)。

[0041] 图3A示出了通过GAP连接物联结的 $\alpha$ -N-乙酰葡萄糖胺糖苷酶-胰岛素样生长因子(NaGlu-IGFII)构建物。图3B示出了通过SEQ ID NO.2的连接物联结的NaGlu-IGFII构建物(NaGlu-GAG<sub>3</sub>-IGFII)。

[0042] 图4A示出了人类NaGlu的核苷酸序列(SEQ ID NO.3)。图4B示出了人类NaGlu的氨基酸序列(SEQ ID NO.4)。

[0043] 图5A示出了人类IGFII的核苷酸序列(SEQ ID NO.5)。图5B示出了人类IGFII的氨基酸序列(SEQ ID NO.6)。

[0044] 图6示出了包含GAG重复序列的肽连接物的氨基酸序列(SEQ ID NO.7)。

[0045] 图7示出了在图3A中示出的NaGlu-IGFII构建物的氨基酸序列(SEQ ID NO.8)。

[0046] 图8示出了在图3B中示出的NaGlu-IGFII构建物的氨基酸序列(SEQ ID NO.9)。

[0047] 图9示出了两种不同NaGlu-IGFII蛋白构建物(NaGlu-GAG<sub>3</sub>-IGFII和NaGlu-IGFII)在HT1080细胞中的活性水平的比较,并证实了与野生型NaGlu相比,NaGlu-GAG<sub>3</sub>-IGFII具有非常高的活性水平,而NaGlu-IGFII具有非常低的活性水平。

[0048] 图10通过western印迹示出了NaGlu-GAG<sub>3</sub>-IGFII和NaGlu-IGFII多肽的表达,并显示NaGlu-IGFII蛋白经历降解,而野生型NaGlu和NaGlu-GAG<sub>3</sub>-IGFII不经历降解。

[0049] 图11示出了人类成纤维细胞(HF1156)对NaGlu-GAG<sub>3</sub>-IGFII多肽的摄入,并证实NaGlu-GAG<sub>3</sub>-IGFII容易被人类细胞摄入。

[0050] 图12示出了肽连接物的氨基酸序列(SEQ ID NO.10)。

[0051] 图13示出了肽连接物的氨基酸序列(SEQ ID NO.11)。

[0052] 图14示出了肽连接物的氨基酸序列(SEQ ID NO.12)。

[0053] 发明详述

[0054] 本文中描述的组合物提供了一种制造(例如设计、工程化改造)嵌合或融合多肽的手段。多肽组成物还可以提供一种便于将药剂(例如多肽/肽、蛋白质和/或酶)递送到目标细胞、组织或器官的手段。具体来说,所述组成物和方法可用于将药剂选择性递送到需要它的对象的适合组织,从而治疗疾病或障碍。这些治疗性组成物可以是多核苷酸或多肽,其包含通过连接物序列联结到靶向剂(例如细胞受体配体蛋白)的治疗剂(例如蛋白质/酶),所述连接物序列允许所述治疗剂和/或靶向剂的正确表达、折叠和活性。在某些情况下,治疗性组成物包含通过连接物序列连接到细胞表面受体配体蛋白的溶酶体蛋白或酶。治疗性多肽组成物可用于治疗诸如溶酶体贮积症的障碍。

[0055] 当在本文中使用时,短语“溶酶体贮积症”是指与一种或多种溶酶体酶的异常表达或缺乏相关的一类遗传疾病。这些酶的缺乏引起代谢产物在患病对象的溶酶体中的有害积累。代表性的溶酶体贮积症包括天冬氨酰葡萄糖胺尿症、胆固醇酯贮积症、胱氨酸贮积症、Danon病、法布里病、法伯氏病、岩藻糖苷贮积症、I/II型半乳糖唾液酸贮积症、1、2、3型高歇

氏病、球形细胞脑白质营养不良/克拉伯病、II型糖原贮积症/庞帕病、I/III型GM1神经节苷脂贮积症、I型GM2神经节苷脂贮积症/泰-萨病、II型GM2神经节苷脂贮积症/山德霍夫氏病、I/II型 $\alpha$ -甘露糖苷贮积症、 $\beta$ -甘露糖苷贮积症、异染色性脑白质营养不良(MLD)、I型粘脂贮积症/I/II型唾液酸贮积症、II/III型粘脂贮积症、III型粘脂贮积症假Hurler多营养障碍、黏多糖贮积症(例如I、II、IIIA、IIIB、IIIC、HID、IVA、IVB、VI、VII和IX型)、多硫酸酯酶缺乏症、神经元蜡样脂褐质贮积症(例如巴登氏-婴儿期、婴儿后期和成人型)、尼曼-皮克病(例如A、B、C1C2型)、I/II型Schindler病、唾液酸贮积症、沙费利波病和沃尔曼氏病(酸性脂酶缺乏症)。在本发明的一个方面,组合物包含治疗剂,所述治疗剂的缺乏与溶酶体贮积症相关。

[0056] 本文中描述了多肽组成物和编码所述多肽组成物的多核苷酸,其中所述多肽组成物包含通过本文公开的连接物序列相连的第一和第二肽/多肽。本发明人令人吃惊地发现,当连接两个蛋白质序列(例如第一多肽和第二多肽)的连接物包含一个或多个顺序或串联重复的SEQ ID NO.1(GAPGGGGGAAAAAGGGG)时,导致产生表达良好且活性(例如生物活性)高的多肽。当在本文中使用时,术语“多肽”或“肽”是指通常专一地由肽键连接的氨基酸残基的聚合物,其可以天然地(例如分离、基本上纯化或纯化)或合成地(例如通过化学合成)生产。通过非宿主DNA分子的表达而生产的多肽是“异源”肽或多肽。构成多肽的“氨基酸残基”可以是天然或非天然的氨基酸残基,其通过肽键和/或不同于肽键的键相连。氨基酸残基可以采取D-构型或L-构型。在某些情况下,在本文中指称的多肽是通过重组核酸的表达来生产的蛋白质、肽或其片段。在某些实施方式中,本文中描述的多肽组成物包含通过连接物序列(例如SEQ ID NO.2)相连的两个多肽,其中所述两个多肽之一是能够给药以治疗疾病或障碍的肽(例如治疗性肽),另一个多肽是可用于将治疗性肽递送到靶细胞、组织或器官的肽(例如靶向肽)。

[0057] 本文中描述的连接物或多肽连接物是指被设计用于连接(例如联结、联系)两个蛋白质序列的肽序列,其中所述连接物肽序列在天然情况下通常不会布置在所述两个蛋白质序列之间。在本发明的情形中,短语“联系”、“联结”或“连接”一般是指两个毗邻或相邻氨基酸序列之间的功能性连接,以产生通常在自然界中不存在的多肽。在某些实施方式中,连接可用于指称例如一个或多个治疗剂肽和一个或多个靶向剂(例如结合或受体配体肽)的氨基酸序列的共价连接。一般来说,相连的蛋白是彼此毗邻或相邻的,并且在相连时保留其相应的可操作性和功能。构成本文公开的嵌合多肽的肽利用居间的包含一个或多个氨基酸的肽连接物相连。这样的连接物可能提供所需的柔性,以允许所述嵌合多肽的所需表达、活性和/或构象定位。典型的氨基酸连接物一般被设计成柔性的,或在两个蛋白质组成部分之间插入诸如 $\alpha$ -螺旋的结构。可以将连接物融合到编码溶酶体酶的多肽的N-端或C-端,或者内部插入。连接物也被称为间隔物。

[0058] 连接物肽序列可以具有任何适合于连接一个或多个目标蛋白的长度,并且优选地被设计成具有足够的柔性,以允许它联结的肽中的一个或两个的正确折叠和/或功能和/或活性。因此,连接物肽可以具有不超过3个、不超过5个、不超过10个、不超过15个、不超过20个、不超过25个、不超过30个、不超过35个、不超过40个、不超过45个、不超过50个、不超过55个、不超过60个、不超过65个、不超过70个、不超过75个、不超过80个、不超过85个、不超过90个、不超过95个或不超过100个氨基酸的长度。在某些实施方式中,连接物肽可以具有至少3

个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少12个、至少15个、至少18个、至少20个、至少25个、至少30个、至少35个、至少40个、至少45个或至少50个氨基酸的长度。在某些实施方式中,连接物包含至少10个且不超过60个氨基酸、至少10个且不超过55个氨基酸、至少10个且不超过50个氨基酸、至少10个且不超过45个氨基酸、至少10个且不超过40个氨基酸、至少10个且不超过35个氨基酸、至少10个且不超过30个氨基酸、至少10个且不超过25个氨基酸、至少10个且不超过20个氨基酸或至少10个且不超过15个氨基酸。在某些实施方式中,连接物包含12至57个氨基酸,并且在特定实施方式中,包含57个氨基酸。在包含连接物的多肽组成物中,连接物肽序列(例如氨基酸序列)的5'末端(例如端)与一个蛋白序列(例如全长蛋白或蛋白结构域、片段或变体)的3'末端相邻并共价连接,并且此外,连接物氨基酸序列的3'末端与另一个蛋白序列的5'末端相邻并共价连接。以这种方式产生的多肽组成物通常被称为融合或嵌合蛋白/多肽,并且通常通过在适合的系统中表达(例如转录、翻译)编码所述多肽组成物的核酸序列来制造。制造融合和/或嵌合多肽的手段在本领域中是公知的(参见例如Sambrook等,《分子克隆实验指南》(Molecular Cloning:A Laboratory Manual),Cold Springs Harbor Laboratory,1992,New York,在此以其全文引为参考)。

[0059] 在某些实施方式中,连接物的氨基酸序列包含甘氨酸、丙氨酸和/或丝氨酸氨基酸残基。本发明人发现,在肽连接物中使用简单氨基酸(例如具有简单侧链(例如H、CH<sub>3</sub>或CH<sub>2</sub>OH)和/或非支链的氨基酸)是有利的,这是由于在这些氨基酸上分支侧链的缺少,在连接物中并因此在多肽组成物中提供了更高的柔性(例如二维或三维柔性)。此外,本发明人发现,交替的甘氨酸、丙氨酸和/或丝氨酸残基在连接物中提供了甚至更高的有序性和更高的柔性。氨基酸可以以与连接物的剩余功能(例如产生被表达的和/或有活性的多肽)相符的任何方式交替/重复。在任何连接物中,丙氨酸氨基酸残基可以被丝氨酸取代。因此,连接物中的氨基酸可以每1个(例如GAGA、GSGS)、每2个(例如GGAAGGAA、GGSSGGSS)、每3个、每4个、每5个、每6个、每7个、每8个、每9个或每10个或更多个氨基酸重复,或者氨基酸可以以上述的任何组合重复。在某些实施方式中,氨基酸每5个氨基酸重复,并且连接物由一个或多个甘氨酸-丙氨酸-甘氨酸重复序列构成。例如,肽连接物可以由GGGGGAAAAAGGGGG(SEQ ID NO.7)或GGGGGSSSSSGGGGG(SEQ ID NO:10)重复序列构成。

[0060] 此外,本发明人发现,将脯氨酸残基置于连接物的前5个(例如5'末端)和/或后5个(例如3'末端)氨基酸内,在连接物中提供了其他益处。例如,连接物或间隔物可以是GAP(SEQ ID NO.11)或GGGGGP(SEQ ID NO.12)。不受理论限制,据信与作为柔性氨基酸的甘氨酸、丙氨酸和丝氨酸不同,其环状侧链导致刚性的脯氨酸,可能在其他氨基酸残基都是柔性的连接物的末端附近产生扭结,从而使通过连接物相连的多肽保持适当分离。因此,连接物的氨基酸序列可以在第一、第二、第三、第四或第五个氨基酸残基中具有脯氨酸,或在连接物的倒数第一、倒数第二、倒数第三、倒数第四或倒数第五个氨基酸残基中具有脯氨酸。在某些实施方式中,肽连接物可以包含18个连续氨基酸残基,其中一个氨基酸残基是脯氨酸,其位于连接物的前5个氨基酸残基的任一个处,其余的17个氨基酸残基由甘氨酸和丙氨酸残基(例如一个或多个甘氨酸和一个或多个丙氨酸残基)构成。连接物的甘氨酸和丙氨酸氨基酸残基可以包含甘氨酸和丙氨酸残基的任何组合,包括上述甘氨酸-丙氨酸-甘氨酸重复序列。连接物还可以包含3个连续氨基酸,其包含第二脯氨酸、甘氨酸残基和/或丙氨酸残

基,以产生包含21个氨基酸的肽连接物。所述第二脯氨酸也可以是连接物氨基酸序列中的后5个氨基酸中的任一个。

[0061] 上述肽连接物的两侧可以具有一个或多个氨基酸序列,其由一个或多个所需的限制性内切酶位点编码。大量内切酶切割位点(例如EcoRI、BamHI、HindIII、AscI位点等)在本领域中是公知的,包含在连接物(和/或多肽)的核酸序列中的所述切割位点的选择,由专业技术人员最好地确定,并且所述位点一般根据待连接的相应核酸序列来选择。根据需要和/或要求,在连接物序列的每一端上的内切酶限制性位点可以是相同位点或不同的限制性位点。在某些实施方式中,连接物的甘氨酸-丙氨酸-甘氨酸氨基酸重复序列在连接物氨基酸序列的5'和/或3'末端处具有甘氨酸-丙氨酸-脯氨酸(GAP)氨基酸序列(例如SEQ ID NO.2)。在这种情形中,GAP序列由代表AscI限制性内切酶位点的核酸序列编码。在某些实施方式中,连接物的氨基酸序列包含SEQ ID NO.1(GAPGGGGGAAAAAGGGGG)。在其他实施方式中,连接物的氨基酸序列包含SEQ ID NO.1的一个或多个(例如1、2、3、4个或以上)顺序重复序列。本发明人发现,即使一个SEQ ID NO.1的重复序列也能改进连接物的功能,其中3个和4个SEQ ID NO.1的重复序列在允许被连接的多肽的表达和/或活性方面最为有效。在某些实施方式中,SEQ ID NO.1的一个或多个顺序重复序列在连接物的氨基酸序列的3'末端/端处进一步包含GAP序列。在某些实施方式中,连接物的氨基酸序列包含SEQ ID NO.2(GAPGGGGGAAAAAGGGGGGAPGGGGGAAAAAGGGGGGAPGGGGGAAAAAGGGGGGAP)。

[0062] 在某些实施方式中,适合的连接物或间隔物可能包含的序列与SEQ ID NO.2的序列具有至少50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、98%或99%的同一性。

[0063] 在本文描述的多肽组成物中,两个多肽(例如第一多肽和第二多肽)可以通过任何上述连接物多肽重组地联结,其中所述连接物被布置在两个多肽之间。例如,在某些实施方式中,多肽或组合物包含由包含SEQ ID NO.1或SEQ ID NO.2的连接物重组联结的第一和第二多肽。两个多肽可以是两种不同蛋白质或相同蛋白质的任何氨基酸序列,包括全长蛋白、蛋白片段或部分、功能性蛋白片段或部分、功能性蛋白结构域等。当在本文中使用时,“功能性片段”或“部分”意指比整个成熟或天然蛋白更小,但足以保留成熟或天然蛋白的一种或多种所需生物活性(例如对于待治疗障碍来说足以保留治疗性或改善性的生物活性)。因此,可以对氨基酸序列或多肽进行修饰,例如在多肽序列中插入、缺失和/或取代氨基酸,其方式使得所述修饰不显著干扰多肽编码功能性药剂的能力。

[0064] 在某些实施方式中,多肽组成物的一种蛋白是具有所需活性的肽,而另一种多肽将具有所需活性的多肽递送或靶向到特定细胞或组织。当在本文中使用时,短语“靶向配体或结合肽”是指用于将药剂(例如蛋白、多肽)定向和递送到特定位点以发挥所需活性的氨基酸序列。在特定实施方式中,一种多肽的所需活性是治疗性活性或预防性活性(例如治疗、替代、抑制、预防、增强、降低或改善)。例如,在某些实施方式中,本文描述的多肽组成物包含在溶酶体贮积症/障碍中缺乏的一种或多种酶和/或蛋白质。例如,本公开的组合物可以包含一种或多种治疗剂,其包含源自于下列一种或多种物质的氨基酸序列或由所述氨基酸序列构成:天冬氨酰葡糖胺酶,酸性脂酶,半胱氨酸转运蛋白,Lamp-2, $\alpha$ -半乳糖苷酶A,脂蛋白脂酶(LPL),神经酰胺酶, $\alpha$ -L-岩藻糖苷酶, $\beta$ -氨基己糖苷酶A, $\beta$ -葡糖醛酸糖苷酶,GM2神经节苷脂活化蛋白, $\alpha$ -D-甘露糖苷酶, $\beta$ -D-甘露糖苷酶,芳基硫酸酯酶A,皂角苷B,神经氨

酸苷酶,  $\alpha$ -N-乙酰葡萄糖胺糖苷酶, 磷酸转移酶, 磷酸转移酶, L-艾杜糖苷酸酶, 艾杜糖酸-2-硫酸酯酶, 艾杜硫酸酯酶, 乙酰肝素-N-硫酸酯酶, 肝素硫酸胺酶,  $\alpha$ -N-乙酰葡萄糖胺糖苷酶, N-乙酰转移酶, N-乙酰氨基葡萄糖6-硫酸酯酶, 半乳糖6-硫酸酯酶,  $\beta$ -半乳糖苷酶, N-乙酰氨基半乳糖4-硫酸酯酶, N-乙酰氨基葡萄糖6-硫酸酯酶, 透明质酸酶-氨基葡萄糖苷酶, 多硫酸酯酶, 棕榈酰基蛋白硫酸酯酶, 三肽基肽酶I, 酸性神经鞘磷脂酶,  $\alpha$ -半乳糖苷酶B, 唾液酸, 以及上述物质的功能性片段、亚基和组合。在某些实施方式中, 多肽组成物的蛋白质中的一种(例如治疗性蛋白)包含N-乙酰- $\alpha$ -氨基葡萄糖苷酶(NaGlu), 特别是人类NaGlu或NaGlu的功能性部分、片段、变体、突变体或衍生物。据信, 溶酶体酶NaGlu的丧失造成溶酶体贮积症沙费利波综合征。

[0065] 在某些实施方式中, 多肽组成物的多肽中的一种包含细胞表面受体配体, 并且在特定实施方式中, 所述多肽是IGFII, 其为IGFII/阳离子不依赖性甘露糖6-磷酸受体(IGFII/CI-MPR)的配体之一。IGFII/CI-MPR识别哺乳动物细胞中添加到新合成的溶酶体酶上的寡糖的甘露糖6-磷酸(Man6-P)组成部分。由于Man6-P与IGFII/CI-MPR的相互作用调节将新合成的酶带到溶酶体的正常细胞内运输, 因此认为IGFII/CI-MPR是可用于将溶酶体酶递送到细胞的受体机制。因此, 上述组合物依赖于受体-介导的转胞吞作用将连接的蛋白(例如治疗性蛋白)递送到目标细胞(例如内皮细胞、巨噬细胞或神经元细胞)。在生理上, 依赖于受体介导的运输将大分子(例如蛋白质、多肽)运输到细胞内, 并通常涉及配体识别和大分子(例如IGFI或IGFII组成部分)与靶细胞上的特定受体结合结构域(例如阳离子不依赖性甘露糖-6磷酸受体(CI-MPR)、IGFI受体、IGFII受体或IGFII/CI-MPR)的结合。在配体与受体上的结合结构域识别和结合之后, 受体-配体复合物经历细胞(例如内皮细胞或巨噬细胞)的胞吞, 并由此使得复合物被内化。然后配体可以被运输穿过细胞(例如内皮细胞、神经元细胞、神经胶质细胞、血管周细胞和/或脑膜细胞)的近腔(abluminal)膜, 并进入到适合的组织中(例如中枢神经系统的组织例如脑或脊髓组织)。在本文中描述的某些实施方式中, 多肽组成物的结合肽或靶向肽包含SEQ ID NO.4, 其为IGFII的8至67位氨基酸残基。

[0066] 本文中还想设想了将功能性蛋白标记物或标签包含在公开的组合物中, 以提供分离和/或检测翻译的多肽或蛋白质的其他手段。适合的标记物和标签在本领域中是公知的, 例如包括但不限于萤光素酶、绿色荧光蛋白、碱性磷酸酶、辣根过氧化物酶、myc标签、FLAG标签、eTag和多组氨酸标签。在优选实施方式中, 这样的标记物和标签能够提供可检测信号, 以便于例如在编码所述多肽组成物的氨基酸序列在所需细胞和组织(例如CNS组织)中分配后鉴定这样的标记物或标签。

[0067] 本文描述的包含通过连接物序列相连的两个多肽的多肽组成物, 可以通过合成方法生产(例如化学合成), 或由多核苷酸(例如核酸)序列编码和表达(例如转录和翻译)。当在本文中使用时, “多核苷酸”是指毗连的共价连接的核酸或核酸分子, 例如脱氧核糖核酸(DNA)或核糖核酸(RNA)、寡核苷酸、通过本领域已知的任何适合手段(例如连接或聚合酶链反应(PCR))产生的片段、和/或通过任何连接、剪切、内切核酸酶作用和外切核酸酶作用产生的片段。多核苷酸分子可以由天然存在的核苷酸(例如DNA和RNA)的单体构成。在某些实施方式中, 本文描述的多肽组成物由同源多核苷酸序列编码。多核苷酸使用本领域技术人员已知的重组DNA技术(参见例如Sambrook等, 1992)来产生。根据本发明, 一般来说将一个或多个靶向剂(例如配体/结合蛋白如IGFII或其生物活性片段)可操作地联结到编码治疗

剂(例如酶NaGlu或其生物活性片段)的核酸或氨基酸序列。在本文描述的包含通过连接物序列相连的至少两个基因的核苷酸序列的某些多核苷酸分子中,可以产生代表两种蛋白的杂合体的融合或嵌合多肽。本文描述的多肽组成物还可以包含适合的调控元件,其可以被克隆到表达载体中并在适合的宿主中表达。用于设计、表达和纯化融合蛋白的重组方法在本领域中是已知的(参见例如Sambrook等,1992)。

[0068] 本文中还想设想了含有上述多核苷酸的表达载体和包含所述多核苷酸或表达载体的重组细胞。“表达载体”是编码在宿主细胞中表达的基因的多核苷酸分子。通常,表达载体包含转录启动子、基因和转录终止子。基因表达通常置于启动子控制之下,并且这样的基因被称为“可操作地连接到”启动子。同样地,如果调控元件调节启动子的活性,调控元件和启动子也可以可操作地连接。用于载体表达的“重组”或“宿主”细胞是含有多核苷酸分子例如本文描述的多核苷酸的细胞。使用这样的表达载体和/或重组细胞可以在体外生产大量蛋白质,因此,本文中设想了生产本公开的多肽组成物的方法。这样的方法包括将包含多核苷酸(例如表达载体)的重组和/或宿主细胞,在适合于从多核苷酸表达多肽的条件下进行培养。具有蛋白质合成能力的任何细胞可用于此目的(例如动物、细菌、酵母或昆虫细胞)。如果需要特定蛋白质修饰,动物细胞以及特别是哺乳动物细胞可能是必需的。可用于表达多肽组成物的细胞包括但不限于HT1080、HF1156、中华仓鼠卵巢(CHO)细胞、CHO-K1细胞、HeLa细胞、Vero细胞、FAO(肝细胞)、人类3T3细胞、A20细胞、EL4细胞、HepG2细胞、J744A细胞、Jurkat细胞、P388D1细胞、RC-4B/c细胞、SK-N-SH细胞、Sp2/mIL-6细胞、SW480细胞、3T6Swiss细胞等。在各种细胞/系统中,适合于蛋白质的表达的条件依赖于所述细胞/系统,并且在本领域中是公知的(Sambrook等,1992)。

[0069] 还设想了可以给药到对象(例如患有疾病或障碍的对象)以实现所需治疗效果(例如分配到目标细胞和组织中)的药物组合物。本文中设想的药物组合物包括例如核酸或氨基酸序列,其编码与一种或多种靶向配体(例如IGFII的片段)可操作连接的一种或多种治疗剂(例如溶酶体酶NaGlu),或包含所述核酸或氨基酸序列的载体或细胞。这样的氨基酸或核酸可以单独给药,但是优选地与至少一种其他试剂或赋形剂(例如可药用赋形剂例如缓冲盐水、右旋糖和纯水)组合给药。

[0070] 适合的可药用赋形剂优选地使悬浮或溶解在其中的蛋白质、酶、核酸、氨基酸和/或多肽稳定,并便于将这样的蛋白质、酶、核酸、氨基酸和/或多肽加工成可以给药于对象的药物组合物。所述药物组合物可以通过任意途径给药,包括但不限于口服、静脉内、肌肉内、动脉内、髓内、鞘内、室内、透皮、皮下、腹膜内、鼻内、肠胃外、表面、舌下或直肠手段。

[0071] 适合于肠胃外给药的药物制剂可以配制在水性溶液,特别是生理相容缓冲液例如生理缓冲盐水中。水性注射悬液可以含有增加悬液粘度的物质,例如羧甲基纤维素钠、山梨糖醇或右旋糖酐。任选地,悬液还可以含有适合的稳定剂或增加化合物的溶解性的试剂,以允许制备高度浓缩的溶液。关于制剂和给药的技术的其他详细情况,可以在最新版的《Remington制药学》(Remington's Pharmaceutical Science, Maack Publishing Co., Easton, Pa.)中找到。在制备药物组合物之后,可以将它们置于适合的容器中并附有标签以指示适应病症的治疗(例如沙费利波综合征的治疗)。这样的标签可以包括但不限于计算向对象给药的药物组合物的有效量、适合的给药时间安排、可接受的给药途径和预期副作用的说明。

[0072] 还设想了通过将本文中公开的多肽组成物和/或表达载体和/或重组细胞给药于需要它们的对象,将治疗性多肽递送到需要其的对象的方法。还设想了通过将有效量的本公开多肽组成物给药于需要它们的对象,来治疗溶酶体贮积症(例如沙费利波综合征)的方法。当在本文中使用时,术语“对象”是指任何哺乳动物(例如人类、小鼠、大鼠、狗、猫、猪、猴、马),特别是人类。在某些实施方式中,对象是成年人、青少年或婴儿。本发明还设想了在子宫内给药组成物和/或执行治疗方法。本文公开的组成物和方法可以使用上述给药途径来给药,并且在某些实施方式中,多肽组成物肠胃外给药。

[0073] 当在本文中使用时,短语“有效量”是指获得所需临床效果所需的治疗剂和/或多肽的量,并主要根据药物组合中包含的治疗剂的总量来确定。一般来说,治疗剂的有效量足以获得对对象有意义的益处(例如治疗、调节、治愈、预防和/或改善潜伏的疾病或病症)。例如,本文描述的药物组合物的有效量可以是足以实现所需治疗和/或预防效果的量,例如足以调节溶酶体酶受体或它们的活性,从而治疗这样的溶酶体贮积症或其症状的量。一般来说,给药到需要的对象的治疗剂(例如重组溶酶体酶)的量取决于对象的特征。这样的特征包括对象的病症、总体健康、年龄、性别和体重。本领域的普通技术人员能够根据这些以及其他相关因素容易地确定适合的剂量。此外,可以任选地使用客观和主观测定法两者来鉴定最适剂量范围。

## 实施例

[0074] 为了最大化溶酶体酶N-乙酰- $\alpha$ -氨基葡萄糖苷酶(Naglu)的摄入,将编码成熟IGFII的8-67位残基的表达盒同框融合到全长人类Naglu开放阅读框的C-端。该构建物的设计类似于由LeBowitz等(PNAS101(9):3083-3088,2004)描述的葡萄糖醛酸苷酶-IGFII融合蛋白,所述作者观察到当融合到IGFII时,葡萄糖醛酸苷酶通过IGFII阳离子不依赖性甘露糖-6-磷酸受体的细胞摄入极大地增加,所述IGFII以高亲和性结合于所述受体上的独特位点。为了确定这样的融合蛋白的好处,产生了两种表达质粒(图3A),一种表达全长野生型人类Naglu,另一种表达带有由3个残基甘氨酸-丙氨酸-脯氨酸(GAP)构成的连接物区的Naglu-IGFII融合蛋白(LeBowitz,2004)。

[0075] 为野生型Naglu和Naglu-IGFII两者产生了稳定的HT1080哺乳动物细胞系。在将稳定细胞系以 $1 \times 10^6$ 个细胞/ml的密度接种并在33°C培养24小时后,通过western印迹监测在调制培养基中的蛋白表达,并通过测量产荧光底物4MU-N-乙酰- $\alpha$ -D-氨基葡萄糖的切割来测定活性。确定了Naglu-IGFII的根据细胞数量归一化的活性,与对无标签的Naglu所观察到的相比低10倍(参见图9)。通过western印迹明显发现,当将样品载量对细胞数量归一化时,Naglu-IGFII的表达明显低于无标签的Naglu(图10)。此外,在Naglu-IGFII表达期间还发生显著量的降解,正如在western印迹的Naglu-IGFII道中,由跑到上方条带之下的较低分子量条带所证实的(箭头,图10)。在包含3次独立的稳定转染的所有所发展出的Naglu-IGFII稳定克隆中,观察到这样的表达和分解水平。事实上,当分析各个Naglu-IGFII克隆的表达水平和活性时,在通过western印迹观察到的降解量与样品中Naglu-IGFII的活性水平之间存在非常明显的关联性。因此,在调制培养基中表现出更高Naglu-IGFII水平的克隆,通常在western印迹上具有相应的更高的降解水平。怀疑由于布置IGFII蛋白,Naglu-IGFII相对无活性,并且怀疑在从Naglu-IGFII克隆获取的样品中观察到的活性水平主要由IGFII



蛋白的剪切造成,引起细胞调制的培养基样品中有活性的无标签Naglu的水平增加。

[0076] 假设增加Naglu与IGFII之间的连接物长度,通过允许产生构象上更稳定的折叠蛋白和/或防止IGFII与Naglu活性位点之间的任何潜在干扰,可以提高Naglu-IGFII融合蛋白的表达和活性。为了增加连接物长度,产生了互补的寡核苷酸,其由甘氨酸-丙氨酸-甘氨酸重复序列构成,并在两侧带有由AscI限制性内切酶位点编码的原始GAP连接物。然后将所述寡核苷酸连接到AscI位点中。在该连接中,将3个GAG寡核苷酸合并到连接物中,产生长为57个氨基酸的连接物(图3B和图8)。从该构建物产生的稳定克隆得到有活性(图9)并且以与Naglu相近的水平表达(图10)的Naglu-GAG<sub>3</sub>-IGFII蛋白。这证实了较长的连接物允许Naglu-IGFII融合蛋白的正确折叠和酶活性。此外,试验了重组Naglu-GAG<sub>3</sub>-IGFII和Naglu在人类成纤维细胞(HF1156)中的摄入。Naglu-GAG<sub>3</sub>-IGFII容易地被细胞摄入,并且所述摄入以剂量依赖性方式被IGFII但不被甘露糖6-磷酸(M6P)抑制。正如预期的,缺少M6P的重组Naglu在细胞中未显示出摄入(图11)。

[0077] 因此,本发明人令人吃惊地发现了产生高表达且有活性的Naglu-IGFII多肽的肽连接物。根据这一工作,相信较长的连接物(例如长于3个氨基酸)导致蛋白的正确折叠/活性并防止其降解。因此,这种以及类似的肽连接物可普遍用于连接其他蛋白并产生异源多肽。

[0078] 尽管已根据某些实施方式对本发明的某些组合和方法进行了具体描述,但上面的实例仅用于示例本发明的化合物而不打算对其进行限制。

[0079] 当在说明书和权利要求书中使用时,无具体数量的指称应该被理解为包括复数指称物,除非清楚地指明不是如此。在组的一个或多个成员之间包括“或”的权利要求项或描述,被认为满足一个、一个以上或所有的组的成员存在于、使用于给定产物或方法中,或以其他方式与给定产物或方法相关,除非指明不是如此或从上下文明显看出不是如此。本发明包括其中所述组中仅有一个确定成员存在于、使用于给定产物或方法中或以其他方式与给定产物或方法相关的实施方式。本发明还包括其中一个以上或全部组的成员存在于、使用于给定产物或方法中或以其他方式与给定产物或方法相关的实施方式。此外,应该理解,本发明涵盖将来自于一个或多个所列权利要求项的一个或多个限制、要素、条款、描述性项等,引入到从属于同一基本权利要求项的另一个权利要求项(或者如果适当的话任何其他权利要求项)中的所有变化、组合和排列方式,除非另有指明或本领域普通技术人员可以明显看出发生矛盾或不协调。当要素作为名单出现时(例如采取马库什组或类似格式),应该理解所述要素的每个亚组也被公开,并且可以从所述组中移除任何要素。应该理解,一般来说,当本发明或本发明的方面被称为包含特定要素、特点等时,本发明或本发明的方面的某些实施方式由这样的要素、特点等构成,或基本上由其构成。出于简化的目的,那些实施方式不是每种情况都用如此多的词句在本文中具体陈述。还应该理解,本发明的任何实施方式或方面都可以从权利要求书中明确排除,不论在说明书中是否叙述了所述具体的排除。在本文中为描述本发明的背景和提供关于其实践的其他详情而参考的出版物、网址和其他参考材料,在此引为参考。

## 序列表

<110> 夏尔人类遗传治疗公司  
 <120> 用于多肽组合物多肽组成物的肽连接物及其使用方法  
 <130> SHIR-018-W01  
 <140> PCT/US12/027561  
 <141> 2012-03-02  
 <150> 61/449,225  
 <151> 2011-03-04  
 <160> 12  
 <170> PatentIn version 3.5

<210> 1  
 <211> 18  
 <212> PRT  
 [0001] <213> Artificial

<220>  
 <223> Peptide Linker

<400> 1

Gly Ala Pro Gly Gly Gly Gly Gly Ala Ala Ala Ala Ala Gly Gly Gly  
 1                    5                    10                    15

Gly Gly

<210> 2  
 <211> 57  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Peptide Linker

<400> 2

Gly Ala Pro Gly Gly Gly Gly Gly Ala Ala Ala Ala Ala Gly Gly Gly  
1 5 10 15

Gly Gly Gly Ala Pro Gly Gly Gly Gly Gly Ala Ala Ala Ala Ala Gly  
20 25 30

Gly Gly Gly Gly Gly Ala Pro Gly Gly Gly Gly Gly Ala Ala Ala Ala  
35 40 45

Ala Gly Gly Gly Gly Gly Gly Ala Pro  
50 55

<210> 3

<211> 2228

<212> DNA

<213> Artificial

[0002]

<220>

<223> NaGlu-IGFII Fusion Construct

<400> 3

atggaggcgg tggcggtggc cgcggcgggtg ggggtccttc tcctggccgg ggccgggggc 60

gcggcaggcg acgaggcccg ggaggcggcg gccgtgcggg cgctcgtggc ccgctgctg 120

ggcccaggcc ccgcggccga ctctccgtg tcggtggagc gcctctggc tgccaagccg 180

ggcttgaca cctacagcct gggcggcggc ggcgcggcgc gcgtgcgggt gcgcggctcc 240

acgggcgtgg cggccgcgcg gggctgcaec cgctacctgc gcgacttctg tggetgccac 300

gtggcctggt ccggctctca gctgcgctg ccgcggccac tgccagccgt gccgggggag 360

ctgaccgagg ccacgcccaa caggtaccgc tattaccaga atgtgtgeac gcaaagctac 420

tccttcgtgt ggtgggactg ggcccgtgg gagegagaga tagactggat ggcgctgaat 480

ggcatcaacc tggcactggc ctggagcggc caggaggcca tctggcagcg ggtgtacctg 540

	gecttgggce tgaccagge agagatcaat gatttcttta ctggctctgc ctctctggee	600
	tggggcgaa tgggcaacct gcaacactgg gatggcccc tgcecccctc ctggcacate	660
	aagcagcttt acctgcagca cgggtctctg gaccagatgc gctccttegg catgacceca	720
	gtgctgcctg caticgcggg gcatgttccc gaggctgtca ccagggtgtt ccctcaggtc	780
	aatgtcaaga agatgggcag ttggggccac tttaaactgtt cctactcctg ctctctcctt	840
	ctggetccgg aagaacccat attcccctc ategggagct ctctctgaga gactgatca	900
	aagagtttgg cacagaccac atctatgggg ccgacaactt caatgagatg cagccacctt	960
	cctcagagcc ctectacctt gcegcageca ccaactgcegt ctatgagccc atgactgcag	1020
	tggatactga ggctgtgtgg ctgctccaag gctggctctt ccagcaccag ccgcagttct	1080
	ggggcccgc ccagatcagg gctgtgctgg gactgtgccc ccgtggccc ctctctgttc	1140
	tggacctgtt tgctgagage cagcctgtgt ataaccgcac tgccctcttc caggccagc	1200
[0003]	cttcatctg gtgcatgctg cacaactttg ggggaaacca tggctttttt ggagccctag	1260
	aggctgtgaa cggaggecca gaagctgccc gctcttccc caactccacc atggtaggea	1320
	cggcatggc ccccagggc atcagccaga acgaagtgtt ctattccctc atggetgagc	1380
	tgggetggcg aaaggacca gtgccagatt tgccagcctg ggtgaccagc ttgcccgc	1440
	ggcggtatgg ggtctccac ccggacgcag gggcagcgtg gaggetactg ctccggagtg	1500
	tgfacaactg ctccggggag gctgcaggg gccacaatcg tagcccctg gtcaggcggc	1560
	cgctccatac gatgaatacc agcatctggt acaaccgac tgatgtgttt gaggcctgge	1620
	ggetgtget cacatctget cctcctctg ccaccagccc cgcctccgc taagaactgc	1680
	tggacctcac teggcaggca gtgcaggagc tggctagctt gtaactatgag gaggcaagaa	1740
	gcgcctacct gagcaaggag ctggcctccc tgttgagggc tggaggcgtc ctggcctatg	1800
	agctgtgccc ggcactggac gagggtctgg ctagtgcag ccgcttcttg ctgggcagct	1860

ggctagagca ggccccgagca gcggcagtea gtgaggccga ggccgattc tacgagcaga 1920  
 acagccgcta ccagctgacc ttgtgggggc cagaaggcaa catctctggac tatgccaaca 1980  
 agcagctggc ggggttggtg gccaaactact acaccctctg ctggcggett ttctggagg 2040  
 cgctggttga cagtgtggcc cagggcatec cttccaaca gcaccagttt gacaaaaatg 2100  
 tcttccaact ggagcaggee ttctttctca gcaagcagag gtaeccceag cageccgcgag 2160  
 gagacactgt ggacctggee aagaagatct tctcaataa ttacccegc tgggtggccg 2220  
 gctcttgg 2228

<210> 4  
 <211> 743  
 <212> PRT  
 <213> Artificial

<220>

[0004] <223> NaGlu-IgFII Fusion Construct

<400> 4

Met Glu Ala Val Ala Val Ala Ala Ala Val Gly Val Leu Leu Leu Ala  
 1 5 10 15

Gly Ala Gly Gly Ala Ala Gly Asp Glu Ala Arg Glu Ala Ala Ala Val  
 20 25 30

Arg Ala Leu Val Ala Arg Leu Leu Gly Pro Gly Pro Ala Ala Asp Phe  
 35 40 45

Ser Val Ser Val Glu Arg Ala Leu Ala Ala Lys Pro Gly Leu Asp Thr  
 50 55 60

Tyr Ser Leu Gly Gly Gly Gly Ala Ala Arg Val Arg Val Arg Gly Ser  
 65 70 75 80

Thr Gly Val Ala Ala Ala Ala Gly Leu His Arg Tyr Leu Arg Asp Phe  
 85 90 95

Cys Gly Cys His Val Ala Trp Ser Gly Ser Gln Leu Arg Leu Pro Arg  
 100 105 110

Pro Leu Pro Ala Val Pro Gly Glu Leu Thr Glu Ala Thr Pro Asn Arg  
 115 120 125

Tyr Arg Tyr Tyr Gln Asn Val Cys Thr Gln Ser Tyr Ser Phe Val Trp  
 130 135 140

Trp Asp Trp Ala Arg Trp Glu Arg Glu Ile Asp Trp Met Ala Leu Asn  
 145 150 155 160

[0005] Gly Ile Asn Leu Ala Leu Ala Trp Ser Gly Gln Glu Ala Ile Trp Gln  
 165 170 175

Arg Val Tyr Leu Ala Leu Gly Leu Thr Gln Ala Glu Ile Asn Glu Phe  
 180 185 190

Phe Thr Gly Pro Ala Phe Leu Ala Trp Gly Arg Met Gly Asn Leu His  
 195 200 205

Thr Trp Asp Gly Pro Leu Pro Pro Ser Trp His Ile Lys Gln Leu Tyr  
 210 215 220

Leu Gln His Arg Val Leu Asp Gln Met Arg Ser Phe Gly Met Thr Pro  
 225 230 235 240

Val Leu Pro Ala Phe Ala Gly His Val Pro Glu Ala Val Thr Arg Val  
 245 250 255

Phe Pro Gln Val Asn Val Thr Lys Met Gly Ser Trp Gly His Phe Asn  
 260 265 270

Cys Ser Tyr Ser Cys Ser Phe Leu Leu Ala Pro Glu Asp Pro Ile Phe  
 275 280 285

Pro Ile Ile Gly Ser Leu Phe Leu Arg Glu Leu Ile Lys Glu Phe Gly  
 290 295 300

Thr Asp His Ile Tyr Gly Ala Asp Thr Phe Asn Glu Met Gln Pro Pro  
 305 310 315 320

Ser Ser Glu Pro Ser Tyr Leu Ala Ala Ala Thr Thr Ala Val Tyr Glu  
 325 330 335

[0006] Ala Met Thr Ala Val Asp Thr Glu Ala Val Trp Leu Leu Gln Gly Trp  
 340 345 350

Leu Phe Gln His Gln Pro Gln Phe Trp Gly Pro Ala Gln Ile Arg Ala  
 355 360 365

Val Leu Gly Ala Val Pro Arg Gly Arg Leu Leu Val Leu Asp Leu Phe  
 370 375 380

Ala Glu Ser Gln Pro Val Tyr Thr Arg Thr Ala Ser Phe Gln Gly Gln  
 385 390 395 400

Pro Phe Ile Trp Cys Met Leu His Asn Phe Gly Gly Asn His Gly Leu  
 405 410 415

Phe Gly Ala Leu Glu Ala Val Asn Gly Gly Pro Glu Ala Ala Arg Leu  
 420 425 430

Phe Pro Asn Ser Thr Met Val Gly Thr Gly Met Ala Pro Glu Gly Ile  
 435 440 445

Ser Gln Asn Glu Val Val Tyr Ser Leu Met Ala Glu Leu Gly Trp Arg  
 450 455 460

Lys Asp Pro Val Pro Asp Leu Ala Ala Trp Val Thr Ser Phe Ala Ala  
 465 470 475 480

Arg Arg Tyr Gly Val Ser His Pro Asp Ala Gly Ala Ala Trp Arg Leu  
 485 490 495

Leu Leu Arg Ser Val Tyr Asn Cys Ser Gly Glu Ala Cys Arg Gly His  
 500 505 510

[0007] Asn Arg Ser Pro Leu Val Arg Arg Pro Ser Leu Gln Met Asn Thr Ser  
 515 520 525

Ile Trp Tyr Asn Arg Ser Asp Val Phe Glu Ala Trp Arg Leu Leu Leu  
 530 535 540

Thr Ser Ala Pro Ser Leu Ala Thr Ser Pro Ala Phe Arg Tyr Asp Leu  
 545 550 555 560

Leu Asp Leu Thr Arg Gln Ala Val Gln Glu Leu Val Ser Leu Tyr Tyr  
 565 570 575

Glu Glu Ala Arg Ser Ala Tyr Leu Ser Lys Glu Leu Ala Ser Leu Leu  
 580 585 590

Arg Ala Gly Gly Val Leu Ala Tyr Glu Leu Leu Pro Ala Leu Asp Glu  
 595 600 605



Val Leu Ala Ser Asp Ser Arg Phe Leu Leu Gly Ser Trp Leu Glu Gln  
610 615 620

Ala Arg Ala Ala Ala Val Ser Glu Ala Glu Ala Asp Phe Tyr Glu Gln  
625 630 635 640

Asn Ser Arg Tyr Gln Leu Thr Leu Trp Gly Pro Glu Gly Asn Ile Leu  
645 650 655

Asp Tyr Ala Asn Lys Gln Leu Ala Gly Leu Val Ala Asn Tyr Tyr Thr  
660 665 670

Pro Arg Trp Arg Leu Phe Leu Glu Ala Leu Val Asp Ser Val Ala Gln  
675 680 685

[0008] Gly Ile Pro Phe Gln Gln His Gln Phe Asp Lys Asn Val Phe Gln Leu  
690 695 700

Glu Gln Ala Phe Val Leu Ser Lys Gln Arg Tyr Pro Ser Gln Pro Arg  
705 710 715 720

Gly Asp Thr Val Asp Leu Ala Lys Lys Ile Phe Leu Lys Tyr Tyr Pro  
725 730 735

Arg Trp Val Ala Gly Ser Trp  
740

<210> 5

<211> 183

<212> DNA

<213> Homo sapiens

<400> 5

ctttggggcg gggagctggt ggacaccctc cagttcgtct gtggggaccg cggtttctac 60  
 ttcagcaggc ccgcaagecg tgtgagccgt cgcagccgtg gcatacgttga ggagtgcctg 120  
 ttcgcagct gtgacctggc cctcctggag acgtaactgtg ctacccccgc caagtccgag 180  
 tga 183

<210> 6  
 <211> 60  
 <212> PRT  
 <213> Homo sapiens

<400> 6

Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val Cys Gly Asp  
 1 5 10 15

Arg Gly Phe Tyr Phe Ser Arg Pro Ala Ser Arg Val Ser Arg Arg Ser  
 20 25 30

[0009]

Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala Leu  
 35 40 45

Leu Glu Thr Tyr Cys Ala Thr Pro Ala Lys Ser Glu  
 50 55 60

<210> 7  
 <211> 15  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Peptide Linker

<400> 7

Gly Gly Gly Gly Gly Ala Ala Ala Ala Ala Gly Gly Gly Gly Gly  
 1 5 10 15

<210> 8

<211> 806

<212> PRT

<213> Artificial

<220>

<223> NaGlu-IgFII Fusion Construct

<400> 8

Met Glu Ala Val Ala Val Ala Ala Ala Val Gly Val Leu Leu Leu Ala  
1 5 10 15

Gly Ala Gly Gly Ala Ala Gly Asp Glu Ala Arg Glu Ala Ala Ala Val  
20 25 30

Arg Ala Leu Val Ala Arg Leu Leu Gly Pro Gly Pro Ala Ala Asp Phe  
35 40 45

[0010]

Ser Val Ser Val Glu Arg Ala Leu Ala Ala Lys Pro Gly Leu Asp Thr  
50 55 60

Tyr Ser Leu Gly Gly Gly Ala Ala Arg Val Arg Val Arg Gly Ser  
65 70 75 80

Thr Gly Val Ala Ala Ala Ala Gly Leu His Arg Tyr Leu Arg Asp Phe  
85 90 95

Cys Gly Cys His Val Ala Trp Ser Gly Ser Gln Leu Arg Leu Pro Arg  
100 105 110

Pro Leu Pro Ala Val Pro Gly Glu Leu Thr Glu Ala Thr Pro Asn Arg  
115 120 125

Tyr Arg Tyr Tyr Gln Asn Val Cys Thr Gln Ser Tyr Ser Phe Val Trp  
 130 135 140

Trp Asp Trp Ala Arg Trp Glu Arg Glu Ile Asp Trp Met Ala Leu Asn  
 145 150 155 160

Gly Ile Asn Leu Ala Leu Ala Trp Ser Gly Gln Glu Ala Ile Trp Gln  
 165 170 175

Arg Val Tyr Leu Ala Leu Gly Leu Thr Gln Ala Glu Ile Asn Glu Phe  
 180 185 190

Phe Thr Gly Pro Ala Phe Leu Ala Trp Gly Arg Met Gly Asn Leu His  
 195 200 205

[0011] Thr Trp Asp Gly Pro Leu Pro Pro Ser Trp His Ile Lys Gln Leu Tyr  
 210 215 220

Leu Gln His Arg Val Leu Asp Gln Met Arg Ser Phe Gly Met Thr Pro  
 225 230 235 240

Val Leu Pro Ala Phe Ala Gly His Val Pro Glu Ala Val Thr Arg Val  
 245 250 255

Phe Pro Gln Val Asn Val Thr Lys Met Gly Ser Trp Gly His Phe Asn  
 260 265 270

Cys Ser Tyr Ser Cys Ser Phe Leu Leu Ala Pro Glu Asp Pro Ile Phe  
 275 280 285

Pro Ile Ile Gly Ser Leu Phe Leu Arg Glu Leu Ile Lys Glu Phe Gly  
 290 295 300

Thr Asp His Ile Tyr Gly Ala Asp Thr Phe Asn Glu Met Gln Pro Pro  
 305 310 315 320

Ser Ser Glu Pro Ser Tyr Leu Ala Ala Ala Thr Thr Ala Val Tyr Glu  
 325 330 335

Ala Met Thr Ala Val Asp Thr Glu Ala Val Trp Leu Leu Gln Gly Trp  
 340 345 350

Leu Phe Gln His Gln Pro Gln Phe Trp Gly Pro Ala Gln Ile Arg Ala  
 355 360 365

Val Leu Gly Ala Val Pro Arg Gly Arg Leu Leu Val Leu Asp Leu Phe  
 370 375 380

[0012] Ala Glu Ser Gln Pro Val Tyr Thr Arg Thr Ala Ser Phe Gln Gly Gln  
 385 390 395 400

Pro Phe Ile Trp Cys Met Leu His Asn Phe Gly Gly Asn His Gly Leu  
 405 410 415

Phe Gly Ala Leu Glu Ala Val Asn Gly Gly Pro Glu Ala Ala Arg Leu  
 420 425 430

Phe Pro Asn Ser Thr Met Val Gly Thr Gly Met Ala Pro Glu Gly Ile  
 435 440 445

Ser Gln Asn Glu Val Val Tyr Ser Leu Met Ala Glu Leu Gly Trp Arg  
 450 455 460

Lys Asp Pro Val Pro Asp Leu Ala Ala Trp Val Thr Ser Phe Ala Ala  
 465 470 475 480

Arg Arg Tyr Gly Val Ser His Pro Asp Ala Gly Ala Ala Trp Arg Leu  
 485 490 495

Leu Leu Arg Ser Val Tyr Asn Cys Ser Gly Glu Ala Cys Arg Gly His  
 500 505 510

Asn Arg Ser Pro Leu Val Arg Arg Pro Ser Leu Gln Met Asn Thr Ser  
 515 520 525

Ile Trp Tyr Asn Arg Ser Asp Val Phe Glu Ala Trp Arg Leu Leu Leu  
 530 535 540

Thr Ser Ala Pro Ser Leu Ala Thr Ser Pro Ala Phe Arg Tyr Asp Leu  
 545 550 555 560

[0013] Leu Asp Leu Thr Arg Gln Ala Val Gln Glu Leu Val Ser Leu Tyr Tyr  
 565 570 575

Glu Glu Ala Arg Ser Ala Tyr Leu Ser Lys Glu Leu Ala Ser Leu Leu  
 580 585 590

Arg Ala Gly Gly Val Leu Ala Tyr Glu Leu Leu Pro Ala Leu Asp Glu  
 595 600 605

Val Leu Ala Ser Asp Ser Arg Phe Leu Leu Gly Ser Trp Leu Glu Gln  
 610 615 620

Ala Arg Ala Ala Ala Val Ser Glu Ala Glu Ala Asp Phe Tyr Glu Gln  
 625 630 635 640

Asn Ser Arg Tyr Gln Leu Thr Leu Trp Gly Pro Glu Gly Asn Ile Leu  
 645 650 655

Asp Tyr Ala Asn Lys Gln Leu Ala Gly Leu Val Ala Asn Tyr Tyr Thr  
 660 665 670

Pro Arg Trp Arg Leu Phe Leu Glu Ala Leu Val Asp Ser Val Ala Gln  
 675 680 685

Gly Ile Pro Phe Gln Gln His Gln Phe Asp Lys Asn Val Phe Gln Leu  
 690 695 700

Glu Gln Ala Phe Val Leu Ser Lys Gln Arg Tyr Pro Ser Gln Pro Arg  
 705 710 715 720

Gly Asp Thr Val Asp Leu Ala Lys Lys Ile Phe Leu Lys Tyr Tyr Pro  
 725 730 735

Arg Trp Val Ala Gly Ser Trp Gly Ala Pro Leu Cys Gly Gly Glu Leu  
 740 745 750

[0014]

Val Asp Thr Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser  
 755 760 765

Arg Pro Ala Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu  
 770 775 780

Cys Cys Phe Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala  
 785 790 795 800

Thr Pro Ala Lys Ser Glu  
 805

<210> 9

<211> 860

<212> PRT

<213> Artificial

&lt;220&gt;

&lt;223&gt; NaGlu-IgFII Fusion Construct

&lt;400&gt; 9

Met Glu Ala Val Ala Val Ala Ala Ala Val Gly Val Leu Leu Leu Ala  
1 5 10 15

Gly Ala Gly Gly Ala Ala Gly Asp Glu Ala Arg Glu Ala Ala Ala Val  
20 25 30

Arg Ala Leu Val Ala Arg Leu Leu Gly Pro Gly Pro Ala Ala Asp Phe  
35 40 45

Ser Val Ser Val Glu Arg Ala Leu Ala Ala Lys Pro Gly Leu Asp Thr  
50 55 60

[0015] Tyr Ser Leu Gly Gly Gly Gly Ala Ala Arg Val Arg Val Arg Gly Ser  
65 70 75 80

Thr Gly Val Ala Ala Ala Ala Gly Leu His Arg Tyr Leu Arg Asp Phe  
85 90 95

Cys Gly Cys His Val Ala Trp Ser Gly Ser Gln Leu Arg Leu Pro Arg  
100 105 110

Pro Leu Pro Ala Val Pro Gly Glu Leu Thr Glu Ala Thr Pro Asn Arg  
115 120 125

Tyr Arg Tyr Tyr Gln Asn Val Cys Thr Gln Ser Tyr Ser Phe Val Trp  
130 135 140

Trp Asp Trp Ala Arg Trp Glu Arg Glu Ile Asp Trp Met Ala Leu Asn  
145 150 155 160



	Gly	Ile	Asn	Leu	Ala	Leu	Ala	Trp	Ser	Gly	Gln	Glu	Ala	Ile	Trp	Gln	
						165				170					175		
	Arg	Val	Tyr	Leu	Ala	Leu	Gly	Leu	Thr	Gln	Ala	Glu	Ile	Asn	Glu	Phe	
						180				185					190		
	Phe	Thr	Gly	Pro	Ala	Phe	Leu	Ala	Trp	Gly	Arg	Met	Gly	Asn	Leu	His	
						195				200					205		
	Thr	Trp	Asp	Gly	Pro	Leu	Pro	Pro	Ser	Trp	His	Ile	Lys	Gln	Leu	Tyr	
						210				215					220		
	Leu	Gln	His	Arg	Val	Leu	Asp	Gln	Met	Arg	Ser	Phe	Gly	Met	Thr	Pro	
						225				235					240		
[0016]	Val	Leu	Pro	Ala	Phe	Ala	Gly	His	Val	Pro	Glu	Ala	Val	Thr	Arg	Val	
						245				250					255		
	Phe	Pro	Gln	Val	Asn	Val	Thr	Lys	Met	Gly	Ser	Trp	Gly	His	Phe	Asn	
						260				265					270		
	Cys	Ser	Tyr	Ser	Cys	Ser	Phe	Leu	Leu	Ala	Pro	Glu	Asp	Pro	Ile	Phe	
						275				280					285		
	Pro	Ile	Ile	Gly	Ser	Leu	Phe	Leu	Arg	Glu	Leu	Ile	Lys	Glu	Phe	Gly	
						290				295					300		
	Thr	Asp	His	Ile	Tyr	Gly	Ala	Asp	Thr	Phe	Asn	Glu	Met	Gln	Pro	Pro	
						305				310					315		320
	Ser	Ser	Glu	Pro	Ser	Tyr	Leu	Ala	Ala	Ala	Thr	Thr	Ala	Val	Tyr	Glu	
						325				330					335		

Ala Met Thr Ala Val Asp Thr Glu Ala Val Trp Leu Leu Gln Gly Trp  
 340 345 350

Leu Phe Gln His Gln Pro Gln Phe Trp Gly Pro Ala Gln Ile Arg Ala  
 355 360 365

Val Leu Gly Ala Val Pro Arg Gly Arg Leu Leu Val Leu Asp Leu Phe  
 370 375 380

Ala Glu Ser Gln Pro Val Tyr Thr Arg Thr Ala Ser Phe Gln Gly Gln  
 385 390 395 400

Pro Phe Ile Trp Cys Met Leu His Asn Phe Gly Gly Asn His Gly Leu  
 405 410 415

[0017] Phe Gly Ala Leu Glu Ala Val Asn Gly Gly Pro Glu Ala Ala Arg Leu  
 420 425 430

Phe Pro Asn Ser Thr Met Val Gly Thr Gly Met Ala Pro Glu Gly Ile  
 435 440 445

Ser Gln Asn Glu Val Val Tyr Ser Leu Met Ala Glu Leu Gly Trp Arg  
 450 455 460

Lys Asp Pro Val Pro Asp Leu Ala Ala Trp Val Thr Ser Phe Ala Ala  
 465 470 475 480

Arg Arg Tyr Gly Val Ser His Pro Asp Ala Gly Ala Ala Trp Arg Leu  
 485 490 495

Leu Leu Arg Ser Val Tyr Asn Cys Ser Gly Glu Ala Cys Arg Gly His  
 500 505 510

Asn Arg Ser Pro Leu Val Arg Arg Pro Ser Leu Gln Met Asn Thr Ser  
 515 520 525

Ile Trp Tyr Asn Arg Ser Asp Val Phe Glu Ala Trp Arg Leu Leu Leu  
 530 535 540

Thr Ser Ala Pro Ser Leu Ala Thr Ser Pro Ala Phe Arg Tyr Asp Leu  
 545 550 555 560

Leu Asp Leu Thr Arg Gln Ala Val Gln Glu Leu Val Ser Leu Tyr Tyr  
 565 570 575

Glu Glu Ala Arg Ser Ala Tyr Leu Ser Lys Glu Leu Ala Ser Leu Leu  
 580 585 590

[0018] Arg Ala Gly Gly Val Leu Ala Tyr Glu Leu Leu Pro Ala Leu Asp Glu  
 595 600 605

Val Leu Ala Ser Asp Ser Arg Phe Leu Leu Gly Ser Trp Leu Glu Gln  
 610 615 620

Ala Arg Ala Ala Ala Val Ser Glu Ala Glu Ala Asp Phe Tyr Glu Gln  
 625 630 635 640

Asn Ser Arg Tyr Gln Leu Thr Leu Trp Gly Pro Glu Gly Asn Ile Leu  
 645 650 655

Asp Tyr Ala Asn Lys Gln Leu Ala Gly Leu Val Ala Asn Tyr Tyr Thr  
 660 665 670

Pro Arg Trp Arg Leu Phe Leu Glu Ala Leu Val Asp Ser Val Ala Gln  
 675 680 685

Gly Ile Pro Phe Gln Gln His Gln Phe Asp Lys Asn Val Phe Gln Leu  
 690 695 700

Glu Gln Ala Phe Val Leu Ser Lys Gln Arg Tyr Pro Ser Gln Pro Arg  
 705 710 715 720

Gly Asp Thr Val Asp Leu Ala Lys Lys Ile Phe Leu Lys Tyr Tyr Pro  
 725 730 735

Arg Trp Val Ala Gly Ser Trp Gly Ala Pro Gly Gly Gly Gly Gly Ala  
 740 745 750

Ala Ala Ala Ala Gly Gly Gly Gly Gly Gly Ala Pro Gly Gly Gly Gly  
 755 760 765

[0019] Gly Ala Ala Ala Ala Ala Gly Gly Gly Gly Gly Gly Ala Pro Gly Gly  
 770 775 780

Gly Gly Gly Ala Ala Ala Ala Ala Gly Gly Gly Gly Gly Gly Ala Pro  
 785 790 795 800

Leu Cys Gly Gly Glu Leu Val Asp Thr Leu Gln Phe Val Cys Gly Asp  
 805 810 815

Arg Gly Phe Tyr Phe Ser Arg Pro Ala Ser Arg Val Ser Arg Arg Ser  
 820 825 830

Arg Gly Ile Val Glu Glu Cys Cys Phe Arg Ser Cys Asp Leu Ala Leu  
 835 840 845

Leu Glu Thr Tyr Cys Ala Thr Pro Ala Lys Ser Glu  
 850 855 860

<210> 10  
<211> 15  
<212> PRT  
<213> Artificial

<220>  
<223> Peptide Linker

<400> 10

Gly Gly Gly Gly Gly Ser Ser Ser Ser Ser Gly Gly Gly Gly Gly  
1                      5                      10                      15

<210> 11  
<211> 3  
<212> PRT  
<213> Artificial

[0020] <220>  
<223> Peptide Linker

<400> 11

Gly Ala Pro  
1

<210> 12  
<211> 6  
<212> PRT  
<213> Artificial

<220>  
<223> Peptide Linker

<400> 12

Gly Gly Gly Gly Gly Pro  
1                      5

GAPGGGGGAAAAAGGGGG (SEQ ID NO. 1)

图1

GAPGGGGGAAAAAGGGGG  
GAPGGGGGAAAAAGGGGG  
GAPGGGGGAAAAAGGGGG  
GAP (SEQ ID NO. 2)

图2

pXD671-NAGLU

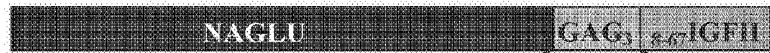


pXD671-NAGLU-IGFII



图3A

pXD671-NAGLU-GAG<sub>3</sub>-IGFII



连接物:  
...G-A-P-G-G-G-G-G-A-A-A-A-A-G-G-G-G-G-  
G-A-P-G-G-G-G-G-A-A-A-A-A-G-G-G-G-G-  
G-A-P-G-G-G-G-G-A-A-A-A-A-G-G-G-G-G-  
G-A-P...  
(SEQ ID NO. 2)

图3B

ATGGAGGCGGTGGCGGTGGCCGCGGCGGTGGGGTCCCTTCTCCTGGCCG  
GGGCCGGGGGCGCGGCAGGCGACGAGGCCCGGGAGGCGGCGGCGGTGCG  
GGCGCTCGTGGCCCGGCTGCTGGGGCCAGGCCCGCGGCGGACTTCTCC  
GTGTCCGTGGAGCGCGCTCTGGCTGCCAAGCCGGGCTTGGACACCTACA  
GCCTGGGCGGCGGCGGCGCGGCGCGCTGCGGGTGC GCGGCTCCACGGG  
CGTGGCGGCCGCCGCGGGGCTGCACCGCTACCTGCGCGACTTCTGTGGC  
TGCCACGTGGCCTGGTCCGGCTCTCAGCTGCGCCTGCCGCGGCCACTGC  
CAGCCGTGCCGGGGGAGCTGACCGAGGCCACGCCAACAGGTACCGCTA  
TTACCAGAATGTGTGCACGCAAAGCTACTCCTTCGTGTGGTGGGACTGG  
GCCCCTGGGAGCGAGAGATAGACTGGATGGCGCTGAATGGCATCAACC  
TGGCACTGGCCTGGAGCGGCCAGGAGGCCATCTGGCAGCGGGTGTACCT  
GGCCTTGGGCCTGACCCAGGCAGAGATCAATGAGTTCTTTACTGGTCCT  
GCCTTCCTGGCCTGGGGGCGAATGGGCAACCTGCACACCTGGGATGGCC  
CCCTGCCCCCCTCCTGGCACATCAAGCAGCTTTACCTGCAGCACCGGGT  
CCTGGACCAGATGCGCTCCTTCGGCATGACCCAGTGCTGCCTGCATTC  
GCGGGGCATGTTCCCGAGGCTGTCAACAGGGTGTTCCTCAGGTCAATG  
TCACGAAGATGGGCAGTTGGGGCCACTTTAACTGTTCCCTACTCCTGCTC  
CTTCCTTCTGGCTCCGGAAGACCCCATATTTCCCATCATCGGGAGCTCT  
TCCTGCGAGAGCTGATCAAAGAGTTTGGCACAGACCACATCTATGGGGC  
CGACACTTTCAATGAGATGCAGCCACCTTCCTCAGAGCCCTCCTACCTT  
GCCGCAGCCACCACTGCCGTCTATGAGGCCATGACTGCAGTGGATACTG  
AGGCTGTGTGGCTGCTCCAAGGCTGGCTCTTCCAGCACCAGCCGCAGTT  
CTGGGGGCCCCGCCAGATCAGGGCTGTGCTGGGAGCTGTGCCCCGTGGC  
CGCCTCCTGGTTCTGGACCTGTTTGTGAGAGCCAGCCTGTGTATAACC  
GCACTGCCTCCTTCCAGGGCCAGCCCTTCATCTGGTGCATGCTGCACAA  
CTTTGGGGGAAACCATGGTCTTTTTTGGAGCCCTAGAGGCTGTGAACGGA  
GGCCCAGAAGCTGCCCGCCTCTTCCCCAACTCCACCATGGTAGGCACGG  
GCATGGCCCCCGAGGGGCATCAGCCAGAACGAAGTGGTCTATTCCCTCAT  
GGCTGAGCTGGGCTGGCGAAAGGACCCAGTGCCAGATTTGGCAGCCTGG  
GTGACCAGCTTTGCCGCCCGGCGGTATGGGGTCTCCACCCGGACGCAG  
GGGCAGCGTGGAGGCTACTGCTCCGGAGTGTGTACAACCTGCTCCGGGGA  
GGCCTGCAGGGGCCACAATCGTAGCCCGCTGGTCAGGCGGCCGTCCCTA  
CAGATGAATACCAGCATCTGGTACAACCGATCTGATGTGTTTGAGGCCT  
GGCGGCTGCTGCTCACATCTGCTCCCTCCCTGGCCACCAGCCCCGCCTT  
CCGCTACGACCTGCTGGACCTCACTCGGCAGGCAGTGCAGGAGCTGGTC  
AGCTTGTACTATGAGGAGGCAAGAAGCGCCTACCTGAGCAAGGAGCTGG  
CCTCCCTGTTGAGGGCTGGAGGCGTCCTGGCCTATGAGCTGCTGCCGGC  
ACTGGACGAGGTGCTGGCTAGTGACAGCCGCTTCTTGCTGGGCAGCTGG  
CTAGAGCAGGCCCGAGCAGCGGCAGTCAGTGAGGCCGAGGCCGATTTCT  
ACGAGCAGAACAGCCGCTACCAGCTGACCTTGTGGGGGCCAGAAGGCAA

图4A

CATCCTGGACTATGCCAACAAGCAGCTGGCGGGGTGGTGGCCAACTAC  
TACACCCCTCGCTGGCGGCTTTTCTGGAGGCGCTGGTTGACAGTGTGG  
CCCAGGGCATCCCTTTCCAACAGCACCAGTTTGACAAAAATGTCTTCCA  
ACTGGAGCAGGCCTTCGTTCTCAGCAAGCAGAGGTACCCAGCCAGCCG  
CGAGGAGACACTGTGGACCTGGCCAAGAAGATCTTCTCAAATATTACC  
CCCGCTGGGTGGCCGGCTCTTGG (SEQ ID NO. 3)

MEAVAVAAVGVLLLAGAGGAAGDEAREAAVRLVARLLGPGFAADFS  
VSVERALAAKPGLDTYSLGGGAARVRVRGSTGVAAAAGLHRYLRDFCG  
CHVAWSGSQRLRPLPAVPGELTEATPNRYRYQNVCTQSYSFVWWDW  
ARWEREIDWMALNGINLALAWSGQEAIWQRVYLALGLTQAEINEFFTGP  
AFLAWGRMGNLHTWDGPLPPSWHIKQLYLQHRVLDQMRSFGMTPVLPFA  
AGHVPEAVTRVFPQVNVTKMGSWGHFNCSYSCSFLAPEDPIFPIIGSL  
FLRELIKEFGTDHIYGADTFNEMQPPSSEPSYLAAATTAVYEAMTAVDT  
EAVWLLQGWLFOHQPFQFWGPAQIRAVLGAVPRGRLLVLDLFAESQPVYT  
RTASFQGGQPFIWCM LNHNFGGNHGLEGALEAVNGGPEAARLFPNSTMVGT  
GMAPEGISQNEVVYSLMAELGWRKDPVPLAAWVTSFAARRYGVSHFDA  
GAAWRLLLRSVYNCSGEACRGNRSPLVRRPSLQMNNTSIWYNRSDVFEA  
WRLLLTSAPSLATSPAFLYDLDDLTRQAVQELVSLYEEARSAYLSKEL  
ASLLRAGGVLAYELLPALDEVLASDSRFLGSWLEQARAAVSEAEADF  
YEQNSRYQLTLWGPEGNILDYANKQLAGLVANYTTPRWRLFLEALVDSV  
AQGIPFQQHQFDKNVFQLEQAFVLSKQRYPSQPRGDTVDLAKKIFLKYY  
PRWVAGSW (SEQ ID NO. 4)

图4B

CTTTGCGGCGGGGAGCTGGTGGACACCCTCCAGTTCGTCTGTGGGGACC  
GCGGCTTCTACTTCAGCAGGCCCGCAAGCCGTGTGAGCCGTCGCAGCCG  
TGGCATCGTTGAGGAGTGCTGTTTCCGCAGCTGTGACCTGGCCCTCCTG  
GAGACGTACTGTGCTACCCCGCCAAGTCCGAGTGA (SEQ ID NO.  
5)

图5A

LCGGELVDTLQFVCGDRGFYFSRPASRVSRRSRGIVEECCFRSCDLALL  
ETYCATEPAKSE (SEQ ID NO. 6)

图5B



GGGGGAAAAAGGGGG (SEQ ID NO. 7)

图6

MEAVAVAAAVGVLLLAGAGGAAGDEAREAAAVRALVARLLGPGPAADFS  
VSVERALAAKPLDITYSLGGGGAARVRVRGSTGVAAAAGLHRYLRDFCG  
CHVAWSGSQRLRPLPAVPGELTEATPNRYRYQNVCTQSYSFVWWDW  
ARWEREIDWMALNGINLALAWSGOEAIWQRVYLALGLTQAEINEFFTGP  
AFLAWGRMGNLHTWDGPLPPSWHIKQLYLQHRVLDQMRSEFGMTPVLPFAF  
AGHVPEAVTRVFPQVNVTKMGSWG HFNC SYSCS FLLAPEDPI FPIIGSL  
FLRELIKEFGTDHIYGADTFNEMQPPSSEPSYLA AATTAVYEAMTAVDT  
EAVWLLQGWL FQHQPQFWGPAQIRAVLGAVPRGRLLVLDLFAESQPVYT  
RTASFQGPFIWCMLHNFGGNHGLFGALEAVNGGPEAARLFPNSTMVGT  
GMAPEGISQNEVVYSLMAELGWRKDPV PDLAAWVTSFAARRYGVSH PDA  
GAAWRLLLR SVYNC SGEACRGNRSPLVRRPSLQMNTSIWYNRSDVFEA  
WRLLLT SAPSLATSPAFRYD LLDLTRQAVQELVSLYEEARSAYLSKEL  
ASLLRAGGV LAYELL PALDEV LASDSR FLLGSWLEQARAAAVSEAEADF  
YEQNSRYQLTLWGPEGNI LDYANKQLAGLVANYYT PRWRLFLEALVDSV  
AQGI PFQHQFQFDKNVFQLEQAFVLSKQRYPSQPRGDTV DLAKKI FLKYY  
PRWVAGSWGAPLCGGELVDTLQFVCGDRGFYFSRPASRVSRRSRGIVEE  
CCFRSCDLALLETYCATPAKSE (SEQ ID NO. 8)

图7

MEAVAVAAAVGVLLLAGAGGAAGDEAREAAAVRALVARLLGPGPAADFS  
 VSVERALAAKPGLDYSLGGGGAARVRVRGSTGVAAAAGLHRYLRDFCG  
 CHVAWSGSQRLRPLPAVPGELTEATPNRYRYQQNVCTQSYSFVWWDW  
 ARWEREIDWMALNGINLALAWSGQEAIWQRVYLALGLTQAEINEFFTGP  
 AFLAWGRMGNLHTWDGPLPPSWHIKQLYLQHRVLDQMRSEFGMTPVLPFAF  
 AGHVPEAVTRVFPQVNVTKMGSWGHFNCSYSCSFLAPEDPIFPIIGSL  
 FLRELIKEFGTDHIYGADTFNEMQPPSSEPSYLA AATTAVYEAMTAVDT  
 EAVWLLQGWL FQHQPQFWGPAQIRAVLGAVPRGRLLVLDLFAESQPVYT  
 RTASFQGGQPF IWCM LHNFGGNHGLEGALEAVNGGPEAARLFPNSTMVGT  
 GMAPEGISQNEVVYSLMAELGWRKDPV PDLAAWVTSFAARRYGVSH PDA  
 GAAWRLLLRSVYNCSGEACRGHNRSP LVRRPSLQMNTSIWYNRSDVFEA  
 WRLLLTSA PSLATSPAFRYDLLDLTRQAVQELVSLYEEARSAYLSKEL  
 ASLLRAGGV LAYELLPALDEV LASDSRFL LGSWLEQARAAA VSEAEADF  
 YEQNSRYQLTLWGPEGNI LDYANKQLAGLVANYYT PRWRLFLEALVDSV  
 AQGIPFQQHQFDKNVFQLEQAFVLSKQRYPSQPRGDTVDLAKKIFLKY Y  
 PRWVAGSWGAPGGGGGAAAAAGGGGGGAPGGGGGAAAAAGGGGGGAPGG  
 GGGAAAAAGGGGGGAPLCGGELVDTLQFVCGDRGFYFSRPASRVSRRSR  
 GIVEECCFRSCDLALLE TYCATPAKSE (SEQ ID NO. 9)

图8

在33℃生长的稳定细胞系的上清液中归一化的Naglu活性

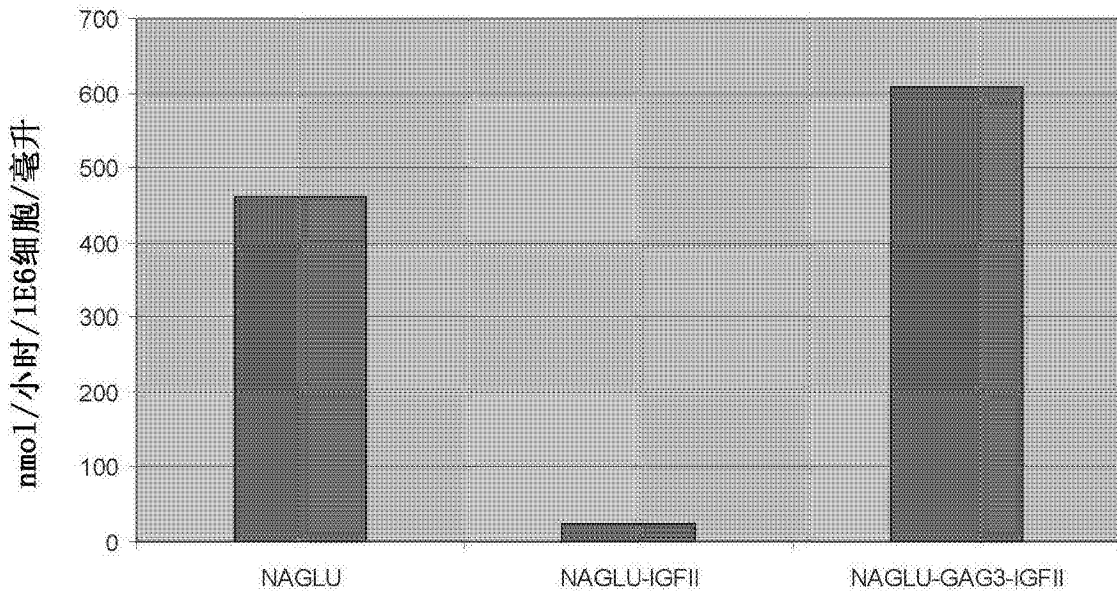


图9

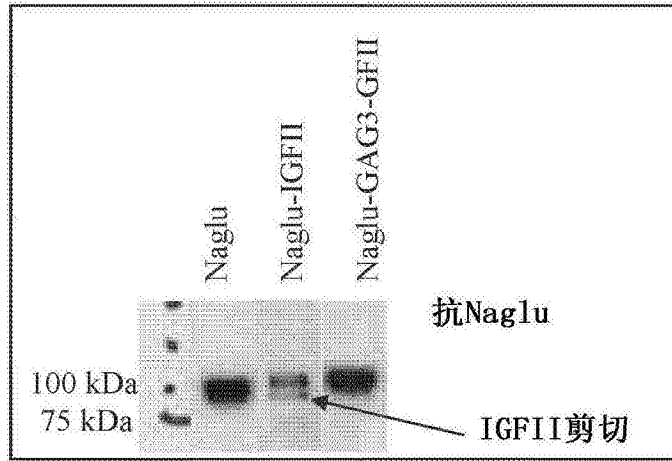


图10

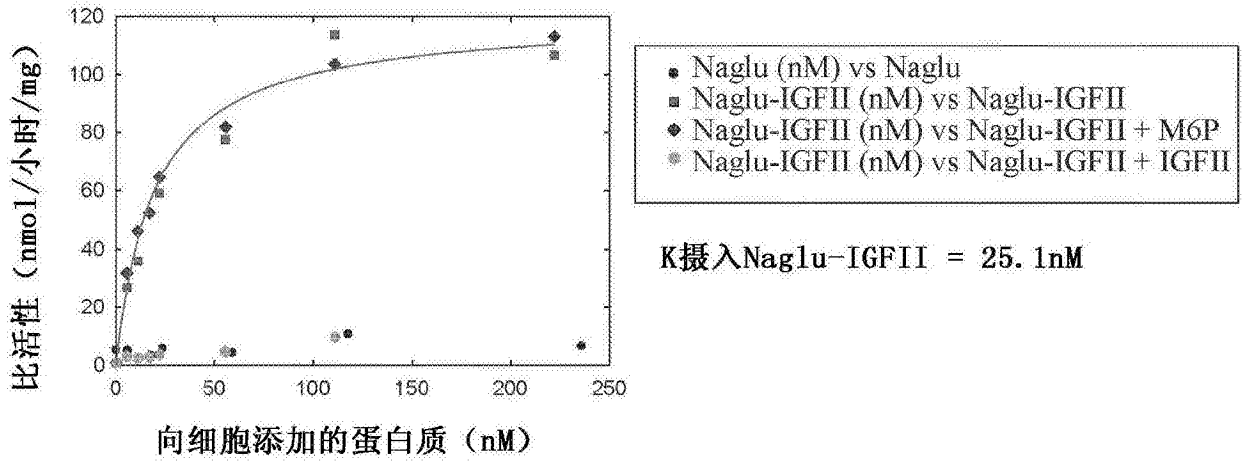


图11

GGGGGSSSSSSGGGGG (SEQ ID NO. 10) GAP (SEQ ID NO. 11)

图12

图13

GGGGGP (SEQ ID NO. 12)

图14