(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 111417656 A
(43)申请公布日 2020.07.14

(21)申请号 201880055010.1

(22)申请日 2018.08.22

(30)优先权数据
201731029916 2017.08.23 IN
201831002945 2018.01.24 IN

(85)PCT国际申请进入国家阶段日
2020.02.24

(86)PCT国际申请的申请数据
PCT/EP2018/072699 2018.08.22

(87)PCT国际申请的公布数据
WO2019/038344 EN 2019.02.28

(71)申请人 西卡肿瘤解决方案有限公司
地址 爱尔兰科克

(72)发明人 N·杰姆桑加米塔

(74)专利代理机构 余姚德盛专利代理事务所（普通合伙）33239
代理人 周积德

(51)Int.Cl.
C07K 19/00(2006.01)

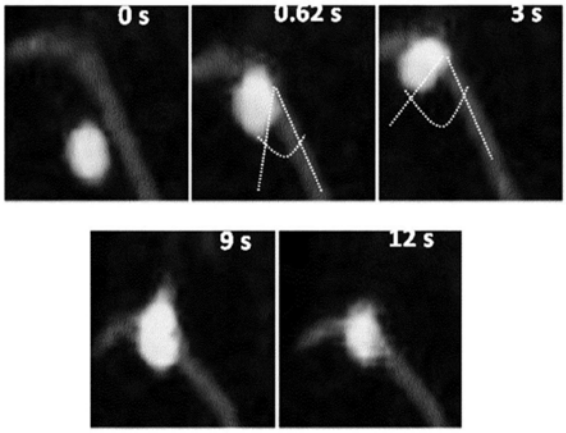
权利要求书4页 说明书42页
序列表19页 附图14页

(54)发明名称

细胞膜穿透缀合物

(57)摘要

细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中所述β螺旋蛋白的长度在5nm至25nm、合适地在10nm至15nm的范围内，并且宽度在1nm至5nm、合适地在1nm至3nm的范围内。还公开制备所述缀合物的方法及其用途。



CN 111417656 A

1.细胞穿透缀合物,包含与功能分子连接的重组β螺旋蛋白,其中所述β螺旋蛋白的长度在5nm至25nm的范围内,并且宽度在1nm至5nm的范围内。

2.权利要求1所述的缀合物,其中所述β螺旋蛋白的长度在10nm至15nm的范围内,并且宽度在1nm至3nm的范围内。

3.权利要求1或2所述的缀合物,其中所述β螺旋蛋白包含一个或多个选自下列的氨基酸阶梯结构:精氨酸阶梯、赖氨酸阶梯、天冬酰胺阶梯、天冬氨酸阶梯和谷氨酸阶梯。

4.权利要求3所述的缀合物,其中当存在时,所述精氨酸阶梯包含10至20个精氨酸残基;所述赖氨酸阶梯包含10至30个赖氨酸残基;所述天冬酰胺阶梯包含10至40个天冬酰胺残基;所述天冬氨酸阶梯包含10至40个天冬氨酸残基;并且所述谷氨酸阶梯包含10至40个谷氨酸残基。

5.权利要求1至4中任一项所述的缀合物,其中所述β螺旋蛋白的总电荷小于零。

6.权利要求5所述的缀合物,其中所述β螺旋蛋白的总电荷为-20至-60。

7.权利要求1至6中任一项所述的缀合物,其中所述β螺旋蛋白具有通过原子力显微镜测得的刚度参数K(β螺旋)为0.2至12N/m^2的β螺旋结构。

8.权利要求1至7中任一项所述的缀合物,其中所述β螺旋蛋白是五肽重复蛋白。

9.权利要求8所述的缀合物,其中所述β螺旋蛋白包含具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的串联重复的五肽。

10.权利要求1至7中任一项所述的缀合物,其中所述β螺旋蛋白由选自以下的序列表示:SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

11.权利要求1至10中任一项所述的缀合物,其中所述重组β螺旋蛋白通过选自聚乙二醇(PEG)、肽、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子与所述功能分子连接。

12.权利要求11所述的缀合物,其中当所述接头分子是肽时,该肽包含选自脂族氨基酸、芳香族氨基酸、及其组合的氨基酸。

13.权利要求1至12中任一项所述的缀合物,其中所述接头通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

14.权利要求1至13中任一项所述的缀合物,其中所述重组β螺旋蛋白通过酯键或酰胺键与功能分子连接。

15.权利要求1至14中任一项所述的缀合物,其中所述功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

16.权利要求1至15中任一项所述的缀合物,其中所述缀合物还包含信号序列,其中该信号序列将所述缀合物引导至特定细胞或细胞的一部分。

17.权利要求16所述的缀合物,其中所述信号序列选自SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16和SEQ ID NO:17。

18.权利要求1至17中任一项所述的缀合物,其中所述缀合物还包含磷脂酰胆碱分子。

19.权利要求16至18中任一项所述的缀合物,其中所述缀合物将所述功能分子转移至选自以下的位置:细胞器、核、内质网、线粒体、细胞膜、和P-钙黏着蛋白过度表达的乳腺癌

细胞。

20.权利要求19所述的缀合物，其中所述细胞器选自肌动蛋白丝、高尔基体、细胞膜、微管蛋白。

21.权利要求1至20中任一项所述的缀合物，其中所述缀合物用作细胞标记剂。

22.权利要求1至20中任一项所述的缀合物，其中所述缀合物用于将所述功能分子转移到细胞中。

23.一种将功能分子转移到细胞中的方法，该方法包括：

a)将所述功能分子连接到重组β螺旋蛋白以获得缀合物；

b)使所述缀合物与至少一种细胞接触；

其中步骤(b)中接触缀合物将所述功能分子转移到所述细胞中；并且其中所述β螺旋蛋白的长度在5nm至25nm的范围内，并且宽度在1nm至5nm的范围内。

24.权利要求23所述的方法，其中该方法在步骤(b)之后包括：

c)检测所述细胞内的所述缀合物的转移。

25.权利要求23或24所述的方法，其中所述β螺旋蛋白的长度在10nm至15nm的范围内，并且宽度在1nm至3nm的范围内。

26.权利要求23至25中任一项所述的方法，其中所述β螺旋蛋白包含一个或多个选自下列的氨基酸阶梯结构：精氨酸阶梯、赖氨酸阶梯、天冬酰胺阶梯、天冬氨酸阶梯和谷氨酸阶梯。

27.权利要求26所述的方法，其中当存在时，所述精氨酸阶梯包含10至20个精氨酸残基；所述赖氨酸阶梯包含10至30个赖氨酸残基；所述天冬酰胺阶梯包含10至40个天冬酰胺残基；所述天冬氨酸阶梯包含10至40个天冬氨酸残基；并且所述谷氨酸阶梯包含10至40个谷氨酸残基。

28.权利要求23至27中任一项所述的方法，其中所述β螺旋蛋白的总电荷小于零。

29.权利要求28所述的方法，其中所述β螺旋蛋白的总电荷为-20至-60。

30.权利要求23至29中任一项所述的方法，其中所述β螺旋蛋白具有通过原子力显微镜测得的刚度参数K(β螺旋)为0.2至12N/m^2的β螺旋结构。

31.权利要求23至30中任一项所述的方法，其中所述β螺旋蛋白是五肽重复蛋白。

32.权利要求31所述的方法，其中所述β螺旋蛋白包含具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的串联重复的五肽。

33.权利要求23至30中任一项所述的方法，其中所述β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

34.权利要求23至33中任一项所述的方法，其中所述功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

35.权利要求23至34中任一项所述的方法，其中所述功能分子通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

36.权利要求23至35中任一项所述的方法，其中所述细胞选自真核细胞、原核细胞及其组合。

37.权利要求23至36中任一项所述的方法,其中所述方法用于细胞标记。

38.权利要求1至20中任一项所述的细胞穿透缀合物在用于将功能分子传递到细胞中的用途,其中所述功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

39.权利要求1至20中任一项所述的细胞穿透缀合物在用于细胞穿透中的用途。

40.权利要求1至20中任一项所述的细胞穿透缀合物在用于细胞标记中的用途。

41.缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中所述至少一种重组β螺旋蛋白的长度在5nm至25nm的范围内,并且宽度在1nm至5nm的范围内。

42.权利要求41所述的缀合物,其中所述至少一种重组β螺旋蛋白的长度在10nm至15nm的范围内,并且宽度在1nm至3nm的范围内。

43.权利要求41或42所述的缀合物,其中所述β螺旋蛋白包含一个或多个选自下列的氨基酸阶梯结构:精氨酸阶梯、赖氨酸阶梯、天冬酰胺阶梯、天冬氨酸阶梯和谷氨酸阶梯。

44.权利要求43所述的缀合物,其中当存在时,所述精氨酸阶梯包含10至20个精氨酸残基;所述赖氨酸阶梯包含10至30个赖氨酸残基;所述天冬酰胺阶梯包含10至40个天冬酰胺残基;所述天冬氨酸阶梯包含10至40个天冬氨酸残基;并且所述谷氨酸阶梯包含10至40个谷氨酸残基。

45.权利要求41至44中任一项所述的缀合物,其中所述β螺旋蛋白的总电荷小于零。

46.权利要求45所述的缀合物,其中所述β螺旋蛋白的总电荷为-20至-60。

47.权利要求41至46中任一项所述的缀合物,其中所述β螺旋蛋白具有通过原子力显微镜测得的刚度参数K(β螺旋)为0.2至12N/m^2的β螺旋结构。

48.权利要求41至47中任一项所述的缀合物,其中所述β螺旋蛋白是五肽重复蛋白。

49.权利要求48所述的缀合物,其中所述β螺旋蛋白包含具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的串联重复的五肽。

50.权利要求41至47中任一项所述的缀合物,其中所述β螺旋蛋白由选自以下的序列表示:SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

51.权利要求41至50中任一项所述的缀合物,其中所述接头是选自聚乙二醇(PEG)、肽、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子。

52.权利要求51所述的缀合物,其中当所述接头分子是肽时,该肽包含选自脂族氨基酸、芳香族氨基酸、及其组合的氨基酸。

53.权利要求41至52中任一项所述的缀合物,其中所述接头通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

54.权利要求53所述的缀合物,其中所述接头通过酯键或酰胺键与所述重组β螺旋蛋白连接。

55.权利要求41至54中任一项所述的缀合物,其中所述核酸分子包含至少一种目的基因。

56.一种将核酸分子转移到细胞中的方法,该方法包括:

a)将所述核酸分子通过至少一种接头连接到至少一种重组β螺旋蛋白以获得缀合物;和

b)使所述缀合物与至少一种细胞接触;

其中接触所述缀合物将所述核酸分子转移到所述细胞中,并且其中所述β螺旋蛋白的长度在5nm至25nm的范围内,并且宽度在1nm至5nm的范围内。

57.权利要求56所述的方法,其中该方法在步骤(b)之后包括:

c)检测所述细胞内的所述缀合物的转移。

58.权利要求56或57所述的方法,其中所述重组β螺旋蛋白的长度在10nm至15nm的范围内,并且宽度在1nm至3nm的范围内。

59.权利要求56至58中任一项所述的方法,其中所述β螺旋蛋白包含一个或多个选自下列的氨基酸阶梯结构:精氨酸阶梯、赖氨酸阶梯、天冬酰胺阶梯、天冬氨酸阶梯和谷氨酸阶梯。

60.权利要求59所述的方法,其中当存在时,所述精氨酸阶梯包含10至20个精氨酸残基;所述赖氨酸阶梯包含10至30个赖氨酸残基;所述天冬酰胺阶梯包含10至40个天冬酰胺残基;所述天冬氨酸阶梯包含10至40个天冬氨酸残基;并且所述谷氨酸阶梯包含10至40个谷氨酸残基。

61.T权利要求56至60中任一项所述的方法,其中所述β螺旋蛋白的总电荷小于零。

62.权利要求61所述的方法,其中所述β螺旋蛋白的总电荷为-20至-60。

63.权利要求56至62中任一项所述的方法,其中所述β螺旋蛋白具有通过原子力显微镜测得的刚度参数K(β螺旋)为0.2至12N/m^2的β螺旋结构。

64.权利要求56至63中任一项所述的方法,其中所述β螺旋蛋白是五肽重复蛋白。

65.权利要求64所述的方法,其中所述β螺旋蛋白包含具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的串联重复的五肽。

66.权利要求56至63中任一项所述的方法,其中所述β螺旋蛋白由选自以下的序列表示:SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

67.权利要求56至66中任一项所述的方法,其中所述接头选自聚乙二醇(PEG)、肽、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合。

68.权利要求67所述的缀合物,其中当所述接头分子是肽时,该肽包含选自脂族氨基酸、芳香族氨基酸、及其组合的氨基酸。

69.权利要求56至68中任一项所述的方法,其中所述接头通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

70.权利要求69所述的缀合物,其中所述接头通过酯键或酰胺键与所述重组β螺旋蛋白连接。

71.权利要求56至70中任一项所述的方法,其中所述细胞选自原核细胞或真核细胞。

72.权利要求41至55中任一项所述的缀合物在作为转染剂中的用途。

73.权利要求41至55中任一项所述的缀合物在用于基因疗法中的用途。

细胞膜穿透缀合物

技术领域

[0001] 本公开广泛地涉及药物向细胞质的递送领域，并且特别地公开了包含与用于穿透细胞膜的功能分子连接的重组蛋白的缀合物，制备所述缀合物的方法及其用途。

背景技术

[0002] 细胞膜是将细胞的内部环境与其外部环境分隔开的半透膜。原核和真核细胞的膜虽然在某些性质和组成上有所不同，但都包含磷脂的半渗透性双层结构。细胞膜的半透性使它对能够穿透它的分子类型具有选择性。那些能够穿透细胞膜的分子有望用于细胞标记、细胞渗透、细胞递送、药物吸收、基因治疗以及涉及细胞膜渗透的许多其他应用。

[0003] 有某些肽可以穿透细胞膜并易位到胞质溶胶，这种肽称为细胞穿透肽(CPPs)。已经研究了CPP缀合物，其中CPP连接到一个或多个功能分子上，作为将各种生物活性分子跨膜运输的手段。例如，当使用CPP缀合物CPP-胰岛素处理时，Caco-2细胞中的胰岛素吸收急剧增加(6-8倍)(Liang et al.,Biochem.Biophys.Res.Commun.;2005;335(3):734-738)。对于包含Tat肽的缀合物，观察到相似的结果(同上)。另一项研究报道了使用短两亲性肽载体Pep-1，它可以穿透膜并在几种细胞系中递送各种肽和蛋白质(Morris et al.,Nature Biotechnol.,2001,1173-1176)。

[0004] CPP还被用于研究各种抗癌药物作为CPP缀合物的有效传递，由于CPP的特性，CPP缀合物比单独的药物更有效地穿透细胞膜。一项此类研究报道了将Tat蛋白与CK2抑制剂(P15)缀合使用来治疗实体瘤(Perea et al.,Cancer Res.2004,7127-7129)。

[0005] 也有几种分子转运蛋白可以跨细胞膜传递分子。由于胍基的数目和空间排列，已显示包含肽和非肽试剂的富含胍的分子转运蛋白(GR-MoTrs)可以穿透细胞膜。GR-MoTrs可以增强哺乳动物细胞内部各种货物的运输，包括小分子、金属、显像剂、铁颗粒和蛋白质(Wender et al.,Adv.Drug Deliv.Rev.2008,452-472;Wender et al.,Drug Discov.Today Technol.2012,e49-e55)。

[0006] US20130137644公开了由细胞穿透肽、核酸和亲水性聚合物组成的缀合物，所述缀合物可以以增加的效率穿透细胞膜。缀合物中使用的核酸被描述为优选为siRNA，其中聚乙二醇(PEG)为亲水性聚合物。

[0007] US20040176282公开了用于细胞递送核酸、多肽、荧光团、分子复合物的组合物的方法和用途。复合物的光活化分散刺激细胞渗透后生物活性分子在细胞内的释放。该系统有助于抑制分子的生物学功能，同时又是复合物的一部分，但是一旦进入细胞内部并在光激活后，它就可以分散并恢复其生物学活性。

[0008] 基因疗法是涉及功能分子穿过细胞膜屏障的一种有前途且广泛使用的技术的例子。基因治疗涉及将目的基因传递给细胞，以补偿基因的异常活性或提供有益的蛋白质。基因治疗已被证明可治疗慢性淋巴细胞性白血病、X连锁SCID、多发性骨髓瘤、血友病等疾病。许多威胁生命的疾病具有潜在的遗传起源，即该疾病是由于一种或多种相关基因显示的功能障碍或缺乏适当的功能所致。基因疗法已显示出治疗此类疾病的希望。然而，基因疗法在

跨细胞膜递送所需基因方面仍然面临挑战。迄今为止,已经使用了两种传递基因的方法-基于病毒和基于非病毒。

[0009] 用于基因治疗的基于病毒的方法利用减毒病毒作为载体,将所需的基因克隆并通过称为转导的过程转移到所需的细胞中。该方法具有将递送的基因适当整合到细胞基因组中的优点,但是具有其他缺点,其中之一是在不适当的基因组整合的情况下容易诱发癌症。基于非病毒的方法包括将分离的DNA注射到细胞中,以及将阳离子脂质用于质粒DNA的包裹(脂转染)。非病毒方法不需要将基因整合到基因组中,并且在将所需基因转移到组织中的其他细胞方面效率低下。因此,基因疗法虽然是治疗许多危及生命的疾病的有前途的和极好的技术,但它遭受基因传递到细胞中的问题。

[0010] 对于最常见的遗传性疾病,例如囊性纤维化或肌营养不良症,由于难以将遗传物质传递到细胞中,有效的基因治疗可能仍然是一个挑战。尚无一种简单的方法将基因传递到肺上皮或骨骼肌等组织的大量细胞中(Collins et al.,Proc.R.Soc.B.Vol.282.No.1821.The Royal Society,2015)。因此,需要一种有效的细胞递送机制,该机制可以大大增强基因治疗的益处,并扩大为许多威胁生命的疾病提供有希望的治疗方法的途径。

[0011] 另外,对于所有功能分子向细胞的递送而言,普遍的是,为细胞递送而开发的大多数机制都依赖于依赖内吞作用的机制来进入细胞内。易位效率是依赖于胞吞作用的途径中主要关注的领域,这是由于例如药物被困在内体中或在溶酶体中降解的可能性。因此,迫切需要设计新的穿透细胞膜的机制,该机制可以更加可靠和有效。

[0012] 发明概述

[0013] 参考以下描述和所附权利要求将更好地理解本主题的这些和其他特征、方面和优点。提供该概述是为了以简化形式介绍一些概念。本概述既不旨在标识所要求保护的主题的关键特征或必要特征,也不旨在用于限制所要求保护的主题的范围。

[0014] 在本公开的第一方面,提供细胞穿透缀合物,包含与功能分子连接的至少一种重组β螺旋蛋白分子,其中所述β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。

[0015] 在本公开的第二方面,提供一种将功能分子转移到细胞中的方法,该方法包括:(a)将所述功能分子连接到重组β螺旋蛋白以获得缀合物;和(b)使所述细胞穿透缀合物与至少一种细胞接触;其中使缀合物与至少一种细胞接触将核酸分子转移到所述细胞中,并且其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。

[0016] 在第三方面,本公开涉及本发明的第一方面的细胞穿透缀合物在用于将功能分子传递到细胞中的用途,其中所述功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0017] 在第四方面,本公开涉及本发明的第一方面的细胞穿透缀合物在用于细胞穿透中的用途。

[0018] 在第五方面,本公开涉及本发明的第一方面的细胞穿透缀合物在用于细胞标记中的用途。

[0019] 在本公开的第六方面,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适

地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。

[0020] 在本公开的第七方面，提供一种将核酸分子转移到细胞中的方法，该方法包括：(i)将所述核酸分子通过至少一种接头连接到至少一种重组β螺旋蛋白以获得缀合物；和(ii)使所述缀合物与至少一种细胞接触，其中使缀合物与至少一种细胞接触将核酸分子转移到所述细胞中，并且其中重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。

[0021] 在第八方面，本公开涉及本发明的第六方面的细胞穿透缀合物作为转染剂的用途。

[0022] 在第九方面，本公开涉及本发明的第六方面的细胞穿透缀合物用于基因疗法的用途。

[0023] 附图简述

[0024] 以下附图形成本说明书的一部分，并且被包括以进一步示出本公开的方面。通过参考附图并结合本文呈现的具体实施例的详细描述，可以更好地理解本公开。

[0025] 图1显示了根据本发明的分离的AlbG蛋白的CD(圆二色性)光谱。

[0026] 图2显示了根据本发明的含有纯化的质粒(含有AlbG和EfsQNR基因)的琼脂糖凝胶。

[0027] 图3显示了根据本发明的包含纯化的蛋白质(AlbG和EfsQNR)的聚丙烯酰胺凝胶。

[0028] 图4显示了EfsQNR和AlbG蛋白的MALDI-TOF质谱分析。

[0029] 图5示出了根据本公开的实施方案的通过紫外-可见分光光度法表征标记的蛋白质(AlbG-NHSC和EfsQNR-NHSC)的图示。

[0030] 图6示出了根据本公开的实施方案的通过标记的蛋白(缀合物)AlbG-NHSC、EfsQNR-NHSC和TtCuA-NHSC对Hela细胞的差异标记。

[0031] 图7示出了根据本发明的一个实施方案，通过用ATTO-520(市售绿色荧光染料)标记的EfsQNR标记Hela细胞。

[0032] 图8示出了根据本发明的一个实施方案，通过用ATTO-390(市售蓝色荧光染料)标记的EfsQNR标记Hela细胞。

[0033] 图9显示了根据本发明的一个实施方案，通过用ATTO-520(市售绿色荧光染料)标记的EfsQNR标记小胶质细胞(图A)；用ATTO-520标记的EfsQNR对小胶质细胞进行标记和差异标记(图B)。

[0034] 图10显示了根据本发明的一个实施方案，通过用ATTO-520(市售绿色荧光染料)标记的EfsQNR标记角质形成细胞细胞(图A)；用ATTO-520标记的EfsQNR对角质形成细胞进行标记和差异标记(图B)。

[0035] 图11示出了根据本公开的实施方案的通过用ATTO-520(市售绿色荧光染料)标记的EfsQNR标记SH-SY5Y细胞。

[0036] 图12示出了根据本发明的一个实施方案，通过用ATTO-520(市售绿色荧光染料)标记的EfsQNR标记小鼠ES细胞。

[0037] 图13示出了根据本公开的实施方式的通过用ATTO-520(市售绿色荧光染料)标记的EfsQNR标记大肠杆菌细胞。

[0038] 图14示出了根据本发明的一个实施方案，通过用ATTO-520(可商购的绿色荧光染

料)标记的EfsQNR标记酵母(克鲁维酵母)细胞。

[0039] 图15显示了与未处理的细胞(图A)相比,用缀合物EfsQNR-ATTO-647N标记的缀合物处理10分钟(图B)、1小时(图C)和3小时(图D)的HeLa细胞的标准FACS分选结果。

[0040] 图16显示了与未缀合的药物相比,根据本发明的一个实施方案,作为缀合物的一部分的药物的细胞摄取百分比。

[0041] 图17显示了与未缀合的药物相比,根据本公开的实施方案,在作为与蛋白质EfsQNR(标记的CYDD)的缀合物的一部分的化学疗法药物Cisplatin®治疗后,细胞(Hela和HepG2)的存活率百分比。

[0042] 图18显示了根据本公开的实施方案的携带mcherry基因的质粒载体。

[0043] 图19示出了根据本公开的实施方式的用于制备缀合物的方法。

[0044] 图20显示了根据本公开的实施方案的使用缀合物的转染的图示。

[0045] 图21显示了根据本发明用缀合物转染后的HeLa细胞的共聚焦显微镜图像;该缀合物包含编码RFP(红色荧光蛋白)的mcherry基因,该基因通过铜[II]菲咯啉与EfsQNR蛋白连接。

[0046] 图22显示了根据本发明的EfsQNR-ATTO 520缀合物对细胞膜的实时直接渗透。

[0047] 发明详述

[0048] 本领域技术人员将意识到,除了具体描述的内容以外,本公开内容还可以进行变化和修改。应该理解,本公开包括所有这样的变化和修改。本公开还单独地或共同地包括在本说明书中提及或指示的所有这样的步骤、特征、组合物和化合物,以及任何或更多这样的步骤或特征的任何和所有组合。

[0049] 定义

[0050] 为了方便起见,在进一步描述本公开之前,在此描述本说明书中使用的某些术语和实例。这些定义应根据本公开的其余部分来阅读,并且应被本领域技术人员理解。这里使用的术语具有本领域技术人员公认的含义,但是,为了方便和完整起见,下面列出了特定的术语及其含义。

[0051] 除非另有说明,否则本发明的实施采用化学、分子生物学、微生物学、重组DNA技术和化学方法的常规技术,这在本领域普通技术人员的能力范围内。下列文献中也对此类技术进行了说明,例如,M.R.Green,J.Sambrook,2012,Molecular Cloning:A Laboratory Manual,Fourth Edition,Books 1-3,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,NY;Ausubel,F.M.et al.(1995 and periodic supplements;Current Protocols in Molecular Biology,ch.9,13,and 16,John Wiley&Sons,New York,N.Y.);B.Roe,J.Crabtree,and A.Kahn,1996,DNA Isolation and Sequencing:Essential Techniques,John Wiley&Sons;J.M.Polak and James O'D.McGee,1990,In Situ Hybridisation:Principles and Practice,Oxford University Press;M.J.Gait(Editor),1984,Oligonucleotide Synthesis:A Practical Approach,IRL Press;and D.M.J.Lilley and J.E.Dahlberg,1992,Methods of Enzymology:DNA Structure Part A:Synthesis and Physical Analysis of DNA Methods in Enzymology,Academic Press。这些一般文本中的每一个均通过引用并入本文。

[0052] 冠词“一”、“一个”和“该”用于指代该冠词的语法对象中的一个或多个(即至少一

个)。

[0053] 如本文所使用的,术语“包括”是指必须包括任何所述要素,并且还可以可选地包括其他要素。“基本上由...组成”是指必须包括任何所列举的要素,排除了将对所列举要素的基本和新颖特征产生实质性影响的要素,并且可以可选地包括其他要素。“由...组成”是指除列出的元素以外的所有元素。这些术语中的每一个所定义的实施方案都在本发明的范围内。

[0054] 术语“包括”用于表示“包括但不限于”。“包括”和“包括但不限于”可互换使用。

[0055] 如本文所用,术语“细胞膜”是存在于原核和真核细胞中并且将细胞的内部和外部环境分开的生物膜。它充当半透屏障,可检查物质在细胞内外的运输,通常由磷脂双层形成。膜充当支撑物并有助于维持细胞的形状和结构。

[0056] 如本文所用,术语“β螺旋蛋白”是指形成β螺旋二级结构的蛋白。β螺旋蛋白由肽链的相邻β链之间通常平行的缔合形成。β螺旋蛋白可以是右旋β螺旋或左旋β螺旋,这取决于螺旋结构的卷曲方向。

[0057] 如本文所用,术语“五肽重复蛋白(PRP)”是指由串联重复的五肽组成的β螺旋蛋白。在实施方案中,串联重复的五肽具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$。PRP家族在原核和真核家族中有超过500个成员。

[0058] 如本文所用,术语“功能分子”是在细胞内具有效用的任何分子。适用于本发明的功能分子的实例包括染料、药物分子、蛋白质、酶、抗体和核酸。

[0059] 如本文所用,术语“细胞穿透肽”或“CPP”涉及促进细胞摄取/摄取各种功能分子的肽序列。细胞穿透肽通常直接通过细胞膜递送功能分子,从而避免了胞吞介导的细胞进入途径。

[0060] 如本文所用,术语“P-钙黏着蛋白”是指在正常组织中具有稳态功能的细胞间粘附分子。该分子的过度表达与在乳腺、卵巢、前列腺、子宫内膜、皮肤、胃、胰腺和结肠肿瘤中的显着的肿瘤促进作用有关。

[0061] 如本文所用,术语“NHS-香豆素”或“NHSC”是指广泛用于细胞生物学技术的荧光染料。它是分子量为358.35g/mol的7-(二乙氨基)香豆素-3-羧酸N-琥珀酰亚胺酯的通用名称。NHS-香豆素(NHSC)的激发波长为445nm,发射波长为482nm。在荧光显微镜下观察时,它发出绿色荧光,表明与该染料缀合的分子的位置和定量。

[0062] 如本文所用,术语“磷脂酰胆碱”定义了一类以胆碱为首基的磷脂分子。磷脂酰胆碱可用作促进与细胞膜选择性结合和附着的信号分子。

[0063] 如本文所用,术语“Hoechst 33342”是指荧光染料的溶液,该溶液用于细胞成像技术中的DNA和细胞核的固定和活细胞染色。Hoechst 33342是一种具有细胞渗透性的DNA染色剂,其激发波长为460nm,发射波长为490nm,它优先与DNA的腺嘌呤(A)-胸腺嘧啶(T)区结合。

[0064] 如本文所用,术语“ATTO 520”、“ATTO 390”和“ATTO 647N”分别是指由ATTO-Tec GmbH开发并且可从Sigma Aldrich商购的荧光染料。

[0065] 如本文所用,术语“钌金属络合物”是指已知具有抗癌活性的钌金属的配合物。八面体钌(III)和钌(II)配合物在许多实验肿瘤上均表现出抗肿瘤活性。钌金属络合物被认为是规避铂基化合物副作用的极佳替代品。已证明可用于本发明的钌金属络合物的非限制

性实例是三羰基二氯钌(II)(来自Sigma Aldrich)。

[0066] 如本文所用,术语“核酸”是核苷酸的单链或双链共价连接的序列,其中每个核苷酸上的3’和5’末端通过磷酸二酯键连接。所述多核苷酸可以由脱氧核糖核苷酸碱基或核糖核苷酸碱基组成。核酸可以包括DNA和RNA,并且通常是合成制备的,但是也可以从天然来源中分离出来。核酸可进一步包括修饰的DNA或RNA,例如已被甲基化或已进行化学修饰的DNA或RNA,例如经7-甲基鸟苷进行5’-加帽,3’-加工如切割和聚腺苷酸化,和剪接、或用荧光团或其他化合物标记。核酸还可以包括合成核酸(XNA),例如己糖醇核酸(HNA)、环己烯核酸(CeNA)、苏糖核酸(TNA)、甘油核酸(GNA)、锁核酸(LNA)和肽核酸(PNA)。因此,在本文中使用术语“DNA”和“RNA”时,应理解这些术语不限于仅包括天然存在的核苷酸。核酸的大小(在本文中也称为“多核苷酸”)通常表示为双链多核苷酸的碱基对数(bp),或者在单链多核苷酸的情况下以核苷酸数(nt)表示。一千个bp或nt等于一千碱基(kb)。长度少于100个核苷酸的多核苷酸通常称为“寡核苷酸”。

[0067] 如本文所用,术语“3”(“3质数”)和“5”(“5质数”)具有其在本领域中的通常含义,即区分多核苷酸的末端。多核苷酸具有5’和3’端,并且多核苷酸序列通常沿5’至3’方向书写。

[0068] 在本发明的上下文中,术语“氨基酸”以其最广泛的含义使用,并且意在包括天然存在的Lα-氨基酸或残基。天然氨基酸的常用一字母缩写和三字母缩写在本文中使用:A=Ala;C=Cys;D=Asp;E=Glu;F=Phe;G=Gly;H=His;I=Ile;K=Lys;L=Leu;M=Met;N=Asn;P=Pro;Q=Gln;R=Arg;S=Ser;T=Thr;V=Val;W=Trp;Y=Tyr(Lehninger,A.L.,(1975)Biochemistry,2d ed.,pp.71-92,Worth Publishers,New York)。通用术语“氨基酸”进一步包括D-氨基酸、逆反氨基酸以及化学修饰的氨基酸(例如氨基酸类似物),通常不掺入蛋白质(例如正亮氨酸)的天然存在的氨基酸,以及具有本领域已知的氨基酸特征的化学合成化合物,例如β-氨基酸。例如,氨基酸定义中包括苯丙氨酸或脯氨酸的类似物或模拟物,其允许肽化合物的构象限制与天然Phe或Pro相同。这样的类似物和模拟物在本文中被称为相应氨基酸的“功能等同物”。氨基酸的其他实例由Roberts和Vellaccio,The Peptides:Analysis,Synthesis,Biology,Gross and Meiehofer,eds.,Vol.5p.341,Academic Press,Inc.,N.Y.1983列出,其通过引用并入本文。

[0069] “多肽”是通过肽键连接的氨基酸残基的聚合物,无论是天然产生的还是体外合成产生的。长度小于约12个氨基酸残基的多肽通常被称为“肽”,而长度在约12至约30个氨基酸残基之间的多肽可以被称为“寡肽”。如本文所用,术语“多肽”表示天然存在的多肽、前体形式或前蛋白的产物。多肽还可经历成熟或翻译后修饰过程,其可包括但不限于:糖基化、蛋白水解切割、脂化、信号肽切割、前肽切割、磷酸化等。本文使用的术语“蛋白质”是指包含一个或多个多肽链的大分子。

[0070] 当将术语“分离的”应用于多核苷酸或蛋白质序列时,表示该序列已从其天然来源的生物体中移出,因此没有外来或不需要的编码或调控序列。分离的序列适合用于组装本发明的组合物和纳米结构。这样的分离的序列可以包括cDNA和RNA。

[0071] 根据本发明,与本文描述的核酸或蛋白质序列的同源性不仅仅限于100%序列同一性。与本文指定的那些核酸或蛋白质序列密切相关的,证明其功能和/或生化等效性的任何核酸或蛋白质序列均被认为在由权利要求书限定的本发明的范围内。

[0072] 在本发明的上下文中，术语“信号序列”是指用于不同细胞器或核酸结合结构域例如锌指结合蛋白的核定位序列或识别序列。

[0073] 如本文所用，术语“载体”是指用作改善药物的递送和有效性的机制的物质。

[0074] 如本文所用，术语“稀释剂”(也称为填充剂、稀释剂或稀释剂)是指稀释剂。

[0075] 如本文所用，术语“赋形剂”是指充当药物或其他活性物质的媒介物或介质的非活性物质。赋形剂包括着色剂、保湿剂、防腐剂、润肤剂及其组合。

[0076] 除非另有定义，否则本文使用的所有技术和科学术语具有与本公开所属领域的普通技术人员通常所理解的相同含义。尽管类似于或等同于本文描述的那些方法和材料的任何方法和材料都可以用于本公开的实践或测试中，但是现在描述优选的方法和材料。本文提及的所有出版物均通过引用并入本文。

[0077] 序列

[0078] SEQ ID NO:1描述AlbG蛋白的氨基酸序列。

[0079] MPAKTLESKDYCGESFVSEDRSGQSLESIRFEDCTFRQCNFTEAELNRCKFRECEFVDCNLSLISIPQTSFMEVRFVDCKMLGVNWTSAQWPSVKMEGALSFERCILNDSLFYGLYLAGVKMVECRIHDANFTEADCEDADFTQSDLKGSTFHNTKLTGASFIDAVNYHIDIFHNDIKRARFSLPEAASLLNSLDIELSD

[0080] SEQ ID NO:2描述EfsQNR蛋白的氨基酸序列。

[0081] GSHMKITYPLPPNLPEQLPLLTNCQLEDEAILENHLYQQIDLPNQEVRNLVFRDAVFDHLSLANGQFASFDCSNVRFEACDFSNVEWLSGSFHRVTFLRCNLTGTNFADSYLKDCLFEDCKADYASFRFANFNLVHFNQTRLVESEFFEVTWKKLLLEACDLTESNWLNTSLKGLDFSQNTFERLTFSPNYLSGLKVTPEQAIYLASALGLVIT

[0082] SEQ ID NO:3描述来自黄粉虫的抗冻蛋白的氨基酸序列。

[0083] QCTGGADCTSCTGACTGCGNCPNAVTCTNSQHCVKANTCTGSTDCNTAQTCTNSKDCFEANTCTDSTNCYKATACTNSSGCPGH

[0084] SEQ ID NO:4描述来自Rhagium inquisitor的抗冻蛋白的氨基酸序列。

[0085] GYSCRAVGVDGRAVTDIQGTCHAKATGAGAMASGTSEPGSTSTATATGRGATARSTSTGRGTATTTATGTASATSNAIGQGTATTTATGSAGGRATGSATTSSSASQPTQTQTITGPGFQTAKSFARNTATTTVTASHHHHHH

[0086] SEQ ID NO:5描述来自云杉芽杆虫(Choristoneura fumiferana)的抗冻蛋白的氨基酸序列。

[0087] DGSCTNTNSQLSANSKCEKSTLTNCYVDKSEVYGTTCTGSRFDGVTITTSTSTGSRISGPGCKISTCIITGGVPAPSAACKISGCTFSAN

[0088] SEQ ID NO:6描述QNRB1蛋白的氨基酸序列。

[0089] GSHMALALVGEKIDRNRFTGEKIENSTFFNCDFSGADLSGTEFIGCQFYDRESQKGCNFSRAMLKDAIFKSCDLSMADFRNSSALGIEIRHCRAQGADFRGASFMNMITTRTWFCSAYITNTNLSYANFSKVVLEKCELWENRWIGAQVLGATFSGSDLSGGEFSTFDWRAANFTHCDLTNSELGDLDIRGVDLQGVKLDNYQASLLMERLGIAVIG

[0090] SEQ ID NO:7描述UDP-N-乙酰氨基葡萄糖酰基转移酶蛋白的氨基酸序列。

[0091] MIDKSAFVHPTAIVEEGASIGANAHIGPFCIVGPHVEIGEGTVLKSHVVVNGHTKIGRDNEIYQFASIGEVNQDLKYAGEPTRVEIGDRNRIRESVTIHRGTVQGGGLTKVGSDNLLMINAHIAHDCTVGNRCILANNATLAGHVSVDDFAIIGGMTAVHQFCIIGAHVMVGGCSGVAQDVPPYVIAQGNHATPFGVNIEGLKRRGFSREAITAIRNAYKLIYRSGKTLDEVKPEIAELAETYPEVKAFTDFFARSTRGLIR

[0092] SEQ ID NO:8描述来自点状诺氏菌的NP275蛋白的氨基酸序列。

[0093] MGSSHHHHHHSSGLVPRGSHMDVEKLRQLYAAGERDFSIVDLRGAVLENINLSGAILHGAMLDEANLQQANLSRADLSGATLNGADLRGANLSKADLSDAILDNAILEGAILDEAVLNQANLKAANLEQAILSHANIREADLSEANLEAADLSGADLAIADLHQANLHQAALERANLTGANLEDANLEGTILEGGNNNLAT

[0094] SEQ ID NO:9描述果胶裂解酶C的氨基酸序列。

[0095] ATDTGGYAATAGGNVTGAVSKTATSMQDIVNIIDAARLDANGKKVKGGAYPLVITYTGNEDSLINAAAANICGQWSKDPRGVEIKEFTKGITIIGANGSSANFGIWIKKSSDVVVQNMRIGYLPGGAKDGDMIRVDDSPNVWVDHNELFAANHECDGTPDNDTTFESAVDIKGASNTVTVSYNYIHGVKKVGLDGSSSSDTGRNITYHHNYYNDVNARLPLQRGGLVHAYNNLYTNITGSGLNVRQNGQALIENNWFEKAINPVTSRYDGKNFGTWVLKGNNITKPADFSTYSITWTADTKPYVNADSWTSTGTFPTVAYNYSPVSAQCVKDKLPGYAGVGKNLATLTSTACK

[0096] SEQ ID NO:10描述来自卡尔迪纤维素破坏者贝西氏菌的果胶酸裂合酶的氨基酸序列。

[0097] VGTNTGGVLVITDTIIVKSGQTYDGKGIKIIAQGMGDGSQSENQKPIFKLEKGANLKNVIIGAPGCDGIHCYGDNVVENVVWEDVGEDALTVKSEGVVEVIGGSAKEAADKVFQLNAPCTFKVKNFTATNIGKLVRQNGNTTFKVVIYLEDVTLNNVKSCVAKSDSPVSELWYHNLNVNNCKTLFEFPSQSQIHQY

[0098] SEQ ID NO:11描述嗜热甲烷菌的碳酸酐酶的氨基酸序列。

[0099] QEITVDEFSNIRENPVTPWNPEPSAPVIDPTAYIDPQASVIGEVTIGANVMVSPMASIRSDEGMPIFVGDRSNVQDGVVLHALETINEEGEPIEDNIVEVDGKEYAVYIGNNVSLAHQSQVHGPAAVGDDTFIGMQAFVFKSKVGNNCVLEPRSAAIGVTIPDGRYIPAGMVVTSQAEADKLPEVTDDYAYSHTNEAVVYVNVHLAEGYKETS

[0100] SEQ ID NO:12描述来自黑曲霉的果胶裂解酶A蛋白的氨基酸序列。

[0101] VGVSGSAEGFAKGVTGGGSATPVYPDTIDELVSYLGDDEARVIVLTKTFDFTDSEGTTTGTGCAPWGTASACQVAIDQDDWCENYEPDAPSVSVEYYNAGTLGITVTSNKSLIGEGSSGAIKGKGLRIVSGAENIIIQNIAVTDINPKYVWGGDAITLDDCDLVWIDHVTTARIGRQHYVLGTSADNRVSLTNNYIDGVSDYSATCDGYHYWAIYLDGDADLVTMKGNYIYHTSGRSPKVQDNTLLHAVNNYWYDISGHAFEIGEGGYVLAEGNVFQNVDTVLETYEGEAFTVPSSTAGEVCSTYLGRDCVINGFGSSGTFSEDSTSFLSDFEGKNIASASAYTSVASRVVANAGQGNL

[0102] SEQ ID NO:13描述TtCuA蛋白的氨基酸序列。

[0103] AYTLATHTAGVIPAGKLERVDPTTVRQEGPWADPAQAVVQTGPNQYTVYVLAFAFGYQPNPIEVPQGAEIVFKITSPDVIHGFHVEGTNINVEVLPGEVSTVRYTFKRPGEYRIICNQYCGLGHQNMFGTIVVKE

[0104] SEQ ID NO:14描述用于靶向细胞核的信号序列。

[0105] PAAKRVKCD

[0106] SEQ ID NO:15描述靶向细胞内质网的信号序列。

[0107] YPYDVPDYAKDEL

[0108] SEQ ID NO:16描述靶向细胞线粒体的信号序列。

[0109] MLSLRQSIRFFKPATRTLCSSRYLL

[0110] SEQ ID NO:17描述靶向表达P-钙黏着蛋白过度表达的乳腺癌细胞的信号序列。

[0111] LSTAADMQGVVTDGMASGLDKDYLKPDD

[0112] SEQ ID NO:18描述五肽重复蛋白中的共有序列。

[0113] $(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$

[0114] SEQ ID NO:19描述AlbG基因的核酸序列。

[0115] ATGCCGGCGAAAACCCTGGAAAGCAAAGATTATTGCGGCGAAAGCTTTGTGAGCGAAGATCGCAGCGGC

CAGAGCCTGGAAAGCATTCGCTTTGAAGATTGCACCTTTCGCCAGTGCAACTTTACCGAAGCGGAACTGAACCGCTG
CAAATTTCGCGAATGCGAATTTGTGGATTGCAACCTGAGCCTGATTAGCATTCCGCAGACCAGCTTTATGGAAGTGC
GCTTTGTGGATTGCAAAATGCTGGGCGTGAACTGGACCAGCGCGCAGGCGGGCGCGCTGAGCTTTGAACGCTGCATT

[0116] CTGAACGATAGCCTGTTTTATGGCCTGTATCTGGCGGGCGTGAAAATGGTGGAATGCCGCATTCATGAT
GCGAACTTTACCGAAGCGGATTGCGAAGATGCGGATTTTACCCAGAGCGATCTGAAAGGCAGCACCTTTCATAACAC
CAAACTGACCGGCGCGAGCTTTATTGATGCGGTGAACTATCATATTGATATTTTTCATAACGATATTAAACGCGCGC
GCTTTAGCCTGCCGGAAGCGGCGAGCCTGCTGAACAGCCTGGATATTGAACTGAGCGAT

[0117] SEQ ID NO:20描述EfsQNR基因的核酸序列。

[0118] GGCAGCCATATGAAAATTACCTATCCGCTGCCGCCGAACCTGCCGGAACAGCTGCCGCTGCTGACCAAC
TGCCAGCTGGAAGATGAAGCGATTCTGGAAAACCATCTGTATCAGCAGATTGATCTGCCGAACCAGGAAGTGCGCAA
CCTGGTGTTTCGCGATGCGGTGTTTGATCATCTGAGCCTGGCGAACGGCCAGTTTGCGAGCTTTGATTGCAGCAACG
TGCGCTTTGAAGCGTGCGATTTTAGCAACGTGGAATGGCTGAGCGGCAGCTTTCATCGCGTGACCTTTCTGCGCTGC
AACCTGACCGGCACCAACTTTGCGGATAGCTATCTGAAAGATTGCCTGTTTGAAGATTGCAAAGCGGATTATGCGAG
CTTTCGCTTTGCGAACTTTAACCTGGTGCATTTTAACCAGACCCGCCTGGTGGAAAGCGAATTTTTTGAAGTGACCT
GGAAAAAACTGCTGCTGGAAGCGTGCGATCTGACCGAAAGCAACTGGCTGAACACCAGCCTGAAAGGCCTGGATTTT
AGCCAGAACACCTTTGAACGCCTGACCTTTAGCCCGAACTATCTGAGCGGCCTGAAAGTGACCCCGGAACAGGCGAT
TTATCTGGCGAGCGCGCTGGGCCTGGTGATTACC

[0119] SEQ ID NO:21描述TtCuA基因的核酸序列

[0120] GCGTATACCCTGGCGACCCATACCGCGGGCGTGATTCCGGCGGGCAAACTGGAACGCGTGGATCCGACC
ACCGTGCGCCAGGAAGGCCCGTGGGCGGATCCGGCGCAGGCGGTGGTGCAGACCGGCCCGAACCAGTATACCGTGTA
TGTGCTGGCGTTTGCGTTTGGCTATCAGCCGAACCCGATTGAAGTGCCGCAGGGCGCGGAAATTGTGTTTAAAATTA
CCAGCCCGGATGTGATTCATGGCTTTCATGTGGAAGGCACCAACATTAACGTGGAAGTGCTGCCGGGCGAAGTGAGC
ACCGTGCGCTATACCTTTAAACGCCCGGGCGAATATCGCATTATTTGCAACCAGTATTGCGGCCTGGGCCATCAGAA
CATGTTTGGCACCATTGTGGTGAAAGAA

[0121] SEQ ID NO:22描述靶向细胞中肌动蛋白的信号序列。

[0122] GDVQKKRWLFETKPLD

[0123] SEQ ID NO:23描述靶向细胞中微管蛋白的信号序列。

[0124] VQSKCGSKDNIKHVPGGG

[0125] SEQ ID NO.24:描述锌指蛋白的氨基酸序列

[0126] MERPYACPVESCDRRFSDSSNLTRHIRIHTGQKPFQCRICMRNFSRSDHLTTHIRTHTGEKPFACDIC
GRKFARSDERKR HTKIHLRQKD

[0127] SEQ ID NO.25:描述mcherry基因的核酸序列。

[0128] GTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGA
GGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCC
AAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCA
AGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCG
CGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTCCAGGACGGCGAGTTCATCTAC
AAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCT
CCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGG

CCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAAC ATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACT CCACCGGCGGCATGGACGAGCTGTACAAGTAGTAATCTAGAGGGCCCTATTCTATAGTGTCACC.

[0129] SEQ ID NO.26:描述用于AlbG基因的PCR扩增的引物的核酸序列。

[0130] ATCCCGCTCATATGCCGGCCAAGACCCTTG

[0131] SEQ ID NO.27:描述用于AlbG基因的PCR扩增的引物的核酸序列。

[0132] ATCCCGCTCTCGAGTCAATCGGACAGCTCGATATC

[0133] SEQ ID NO.28:描述用于PCR扩增EfsQNR基因的引物的核酸序列。

[0134] ATCCCGCTCATATGAAAATAACTTATCCCTTGCCA

[0135] SEQ ID NO.29:描述用于EfsQNR基因的PCR扩增的引物的核酸序列。

[0136] ATCCCGCTCTCGAGTTAGGTAATCACCAAACCAAGT

[0137] 本发明还提供了包含重组蛋白和用于穿透细胞膜的功能性分子的缀合物及其用途,其与SEQ ID NO:1至12具有基本相似的序列同一性或同源性。术语“基本相似的序列同一性”表示约50%、60%、70%、80%、90%、95%至约99%同一性的序列相似性水平。序列同一性百分比可以使用常规方法确定,例如在Henikoff and Henikoff Proc.Natl.Acad.Sci.USA 1992;89:10915和Altschul et al.Nucleic Acids Res.1997;25:3389-3402 for nucleic acids;and for proteins via comparison after alignment using systems such as BLAST®中描述的那些。

[0138] 具有功能分子的细胞穿透缀合物

[0139] 鉴定用于治疗或预防威胁生命的疾病的新药仍然是研究的高度活跃领域。大多数这样的药物倾向于在细胞内具有靶标,并且为了达到靶标,它们将必须穿过半透膜,这在许多情况下是不直接或有效的。因此,仍然需要开发新颖的机制来穿透细胞膜。另一方面,旨在揭示各种细胞机制的科学研究和研究也正在寻求发现渗透细胞膜的新方法,这些方法可以帮助标记细胞及其各种细胞器。同样,这对于将细胞内所需的材料用于科学实验非常有用。最近报道中描述的大多数细胞穿透分子都涉及细胞进入的内吞机制。

[0140] 为了避免通过内吞作用吸收分子的缺点,例如药物在不同类型的内体区室中的捕获和降解,这些内体区室最终与诸如降解的溶酶体的细胞降解区室融合,本文公开了一种穿透细胞膜的缀合物,其可以穿透细胞膜以进入细胞内部,并且还可以用于在细胞内部递送各种货物,包括染料、药物、蛋白质、酶、抗体和核酸。

[0141] 本公开的范围不受本文描述的具体实施方案的限制,这些具体实施方案仅旨在示例的目的。如本文所述,功能等效的产品、组合物和方法显然在本公开的范围内。

[0142] 本发明公开了一种能够穿透细胞膜的缀合物。缀合物包含与功能分子连接的重组β螺旋蛋白或重组β螺旋蛋白部分,其中蛋白质的最长尺寸(定义为其长度)在5nm-25nm范围内,宽度或直径(定义为基本垂直于其长度的蛋白质结构的尺寸)在1nm-5nm范围内。定义的蛋白质尺寸是在固态(通过X射线晶体学或原子力显微镜表征)或溶液状态(动态光散射测量)中测量的。

[0143] 在本发明的实施方案中,缀合物的蛋白质结构可以具有大于5nm、7.5nm、10nm、11nm、12nm、13nm或14nm的最长尺寸或长度。在实施方案中,缀合物的蛋白质结构可具有小于25nm、、20nm、17.5nm、15nm、14nm、13nm、12nm或11nm的最长尺寸或长度。合适地,长度在

5nm至25nm的范围内，更合适地在10nm至15nm的范围内，甚至更合适地在11nm至14nm或12nm至13nm的范围内。在实施方案中，缀合物的蛋白质结构的宽度或直径在基本上垂直于其长度的尺寸上可以为至少1nm、1.2nm、1.3nm、1.4nm、1.5nm、1.6nm、1.7nm、1.8nm、1.9nm或2.0nm。在实施方案中，缀合物的蛋白质结构可具有小于5.0mm、4.5nm、4.0nm、3.5nm、3.0nm、2.9nm、2.8nm、2.7nm、2.6nm、2.5nm、2.4nm、2.3nm、2.2nm、2.1nm或2.0nm的最长尺寸或长度。适当地，宽度在1nm至5nm的范围内，更合适地在1nm至3nm的范围内，甚至更合适地在1.5nm至2.5nm的范围内。

[0144] 本发明缀合物的蛋白质部分的物理尺寸的替代物或替代定义是其β-螺旋结构及其分子量的组合。基于其物理尺寸或分子量的蛋白质大小定义可以互换使用。在本发明的实施方案中，蛋白质部分的分子量可以为至少30kDa。适当地，蛋白质部分的分子量可以为至少35kDa、40kDa、45kDa或50kDa。在本发明的实施方案中，蛋白质部分的分子量可以为至多100kDa。适当地，蛋白质部分的分子量可以为至多90kDa、80kDa、70kDa、60kDa、55kDa或50kDa。合适地，本发明的缀合物的蛋白质部分的分子量范围为30-100kDa，更合适地为40-60kDa，甚至更合适地为48-55kDa。

[0145] 不希望被理论所束缚，可以设想，本发明的缀合物在细胞渗透中显示的有益特性是由于蛋白质部分的物理尺寸，如上文概述的尺寸或分子量所定义。蛋白质的另一个特征是刚性，这种刚性主要源于罕见的β螺旋二级结构。适当地，蛋白质可以比要穿透的膜更坚硬。根据细胞类型，典型的细胞膜具有0.005至0.02N/m^2的刚度(由Atomic Force Mircroscopy;Hayashi,“Tensile Properties and Local Stiffness of Cells”;Mechanics of Biological Tissue,pp 137-152测量)。合适地，本发明的蛋白质具有典型地为0.7至12N/m^2的β-螺旋的刚性或高刚性参数(K)(Keten et al.,Cell Mol.Bioeng.2009;2;66-74,其通过引用并入本文)。

[0146] 被认为对本发明的缀合物的细胞渗透效率有影响的另一个特征是蛋白质的β-螺旋结构中的氨基酸的电荷分布和排列。具体而言，β螺旋蛋白结构中带电荷的赖氨酸、精氨酸天冬酰胺、天冬氨酸和/或谷氨酸“阶梯”的存在促进了细胞膜的渗透，并促进了蛋白序列的总负电荷。

[0147] 关于β-螺旋蛋白质结构中氨基酸残基的术语“阶梯”定义为沿蛋白质表面长度交替排列的带正电(赖氨酸、精氨酸和/或天冬酰胺)和带负电的残基(天冬氨酸和/或谷氨酸)。

[0148] 不希望受理论的束缚，根据本发明的蛋白质的β-螺旋结构中带电的“阶梯”结构的存在可以促进蛋白质与细胞膜脂质分子的相互作用(例如与脂质分子的羟基形成氢键)和/或蛋白质首先附着在带负电的细胞膜上。

[0149] 由于带负电荷的残基，根据本发明的蛋白质的总负电荷可以促进排斥，从而导致蛋白质的角运动，使其直立在膜上并刺穿膜。

[0150] 在本发明的实施方案中，缀合物的蛋白质包含暴露于带正电和带负电的氨基酸残基的表面交替排列的阶梯。在实施方案中，该蛋白质包含精氨酸阶梯(10-30Arg残基)、赖氨酸阶梯(10-30Lys残基)、天冬酰胺阶梯(10-40Asn残基)、天冬氨酸(10-40Asp残基)和谷氨酸(10-40Glu残基)中的至少一个。

[0151] 在本发明的实施方案中，蛋白质的总形式电荷可能为零，也可能为非零。适当地，

总形式收费不为零，更适当地，总形式收费低于零(负)。在实施方案中，蛋白质的总形式电荷低于-10(即比其更负)。适当地，蛋白质的总形式电荷低于-20、-25、-30、-35、-40、-45或-50。更合适的是，总表格费用低于-20。在实施方案中，蛋白质的总形式电荷高于(即，负电荷小于)-80。合适地，蛋白质的总形式电荷大于-70、-65、-60、-55、-50、-45、-40、-35或-30。更合适的是，总形式电荷高于-60。在实施方案中，蛋白质的总形式电荷在-10至-80的范围内，更合适地在-20至-60的范围内。

[0152] 可以通过增加沿蛋白质表面带正电荷的残基(例如精氨酸)的数量来优化缀合物的总形式电荷来调节细胞穿透的效率和细胞穿透的机制，并且为了获得细胞器的特异性或特定癌细胞的特异性，可以将该序列在信号传导序列的N末端突变(SEQ ID NO:14、15、16或17和/或磷脂酰胆碱)。

[0153] 在一个实施方案中，适用于直接细胞渗透(即非内吞性)的本发明的缀合物的蛋白质可以具有一个或多个以下结构参数：

[0154] 刚度参数K(β螺旋)为0.2到12N/m^2的β螺旋结构；

[0155] 长度在5nm至25nm之间；

[0156] 直径在1nm至5nm之间；

[0157] 分子量介于25KDa至100KDa之间；

[0158] 沿长度方向交替排列的表面阶梯暴露在蛋白质表面上的正电荷和负电荷残基：精氨酸阶梯(10-30Arg残基)、赖氨酸阶梯(10-30Lys残基)、天冬酰胺阶梯(10-40Asn残基)、天冬氨酸(10-40Asp残基)和谷氨酸(10-40Glu残基)；

[0159] 总形式电荷(-20至-60)。

[0160] 在本发明的实施方案中，β-螺旋蛋白与功能分子之间的连接子可以通过β-螺旋蛋白与功能分子之间的直接连接形成，例如通过酰胺(或肽)或酯共价键，或通过金属配位；或者接头可以采取接头分子的形式。适当地，该连接是通过共价或非共价键或相互作用。当接头是接头分子时，接头分子可以采取可逆地连接β-螺旋蛋白和功能分子的任何合适的形式。在一些实施方案中，接头分子可以选自肽或蛋白质、PEG(聚乙二醇)接头、有机分子、金属缀合物、药物-金属缀合物、核酸结合结构域和核酸嵌入分子。

[0161] 在实施方案中，本发明公开具有重组β螺旋蛋白的缀合物，重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，其中β螺旋蛋白具有与功能分子连接的特定序列。合适地特定序列是五肽重复。在实施方案中，连接功能分子的重组β螺旋蛋白的共有序列是$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$。五肽重复蛋白的例子是SEQ ID No:1、2、6和8。

[0162] 在本公开的实施方案中，重组β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11和SEQ ID NO:12。

[0163] 在本发明的实施方案中，功能分子可以是任何有机分子(抗癌药、抗生素、NSAIDS、止痛药或任何其他药物分子、荧光染料、杀虫剂、杀虫剂等)、药物-金属络合物、金属、抗体、蛋白质、多糖、核酸、肽、核定位信号、量子点和纳米颗粒。

[0164] 缀合物可用于在缀合物穿透细胞膜的能力的促进下在细胞内转移功能分子。通过与缀合物或功能分子或两者相关或整合的适当定位信号，本发明的缀合物可用于将功能性

分子靶向细胞内部的特定部分，例如存在于细胞内部的细胞器。

[0165] 在实施方案中，本发明还公开了使用本发明的细胞穿透缀合物在细胞内转移功能分子的方法。所公开的方法可以进一步用于细胞标记、细胞渗透以及将任何功能分子靶向细胞器。

[0166] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。

[0167] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中β螺旋蛋白是五肽重复蛋白。

[0168] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中β螺旋蛋白包含具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的串联重复的五肽。

[0169] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

[0170] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中β螺旋蛋白是具有SEQ ID NO:1所示的序列的AlbG。

[0171] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中β螺旋蛋白是具有SEQ ID NO:2所示的序列的EfsQNR。

[0172] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:3所示的序列的抗冻蛋白。

[0173] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:4所示的序列的抗冻蛋白。

[0174] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:5所示的序列的抗冻蛋白。

[0175] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:6所示的序列的QNRB1。

[0176] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:7所示的序列的UDP

N-乙酰氨基葡萄糖酰基转移酶。

[0177] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:8所示的序列的NP275。

[0178] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:9所示的序列的果胶裂解酶C。

[0179] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:10所示的序列的果胶酸裂合酶。

[0180] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:11所示的序列的碳酸酐酶。

[0181] 在本公开的实施方案中，提供细胞穿透缀合物，包含与功能分子连接的重组β螺旋蛋白，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:12所示的序列的果胶裂解酶A。

[0182] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0183] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是染料。

[0184] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是药物。

[0185] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是金属。

[0186] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是药物金属络合物。

[0187] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是蛋白。

[0188] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是酶。

[0189] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是抗体。

[0190] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是核酸。

[0191] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是多糖。

[0192] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是核定位信号。

[0193] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子是纳米粒子。

[0194] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中所述功能分子通

过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

[0195] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子通过共价键连接重组β螺旋蛋白。

[0196] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中功能分子通过非共价键连接重组β螺旋蛋白。

[0197] 在本公开的实施方案中，提供本文所述包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中缀合物还包含信号序列。

[0198] 在本公开的实施方案中，提供本文所述包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中缀合物还包含磷脂分子。

[0199] 在本公开的实施方案中，提供本文所述包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中缀合物还包含选自SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16和SEQ ID NO:17的信号序列。

[0200] 在本公开的实施方案中，提供本文所述包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中缀合物还包含磷脂酰胆碱分子。

[0201] 在本公开的实施方案中，提供本文所述包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中缀合物还包含SEQ ID NO:14所示的信号序列。

[0202] 在本公开的实施方案中，提供本文所述包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中缀合物还包含SEQ ID NO:15所示的信号序列。

[0203] 在本公开的实施方案中，提供本文所述包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中缀合物还包含SEQ ID NO:16所示的信号序列。

[0204] 在本公开的实施方案中，提供本文所述包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中缀合物还包含SEQ ID NO:17所示的信号序列。

[0205] 在本公开的实施方案中，提供用于将功能分子转移到细胞核的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中缀合物还包含SEQ ID NO:14所示的信号序列。

[0206] 在本公开的实施方案中，提供本文所述用于将功能分子转移到细胞核的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中缀合物还包含SEQ ID NO:14所示的信号序列。

[0207] 在本公开的实施方案中，提供本文所述用于将功能分子转移到细胞核的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0208] 在本公开的实施方案中，提供用于将功能分子转移到细胞的内质网的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中缀合物还包含SEQ ID NO:15所示的信号序列。

[0209] 在本公开的实施方案中，提供本文所述用于将功能分子转移到细胞的内质网的包

含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中缀合物还包含SEQ ID NO:15所示的信号序列。

[0210] 在本公开的实施方案中，提供本文所述用于将功能分子转移到细胞的内质网的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0211] 在本公开的实施方案中，提供用于将功能分子转移到细胞的线粒体的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中缀合物还包含SEQ ID NO:16所示的信号序列。

[0212] 在本公开的实施方案中，提供本文所述用于将功能分子转移到细胞的线粒体的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中缀合物还包含SEQ ID NO:16所示的信号序列。

[0213] 在本公开的实施方案中，提供本文所述用于将功能分子转移到细胞的线粒体的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0214] 在本公开的实施方案中，提供用于将功能分子转移到P-钙黏着蛋白过度表达的乳腺癌细胞的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中缀合物还包含SEQ ID NO:17所示的信号序列。

[0215] 在本公开的实施方案中，提供本文所述用于将功能分子转移到P-钙黏着蛋白过度表达的乳腺癌细胞的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中缀合物还包含SEQ ID NO:17所示的信号序列。

[0216] 在本公开的实施方案中，提供本文所述用于将功能分子转移到P-钙黏着蛋白过度表达的乳腺癌细胞的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0217] 在本公开的实施方案中，提供用于将功能分子转移到细胞膜的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中所述缀合物还包含磷脂酰胆碱分子。

[0218] 在本公开的实施方案中，提供本文所述用于将功能分子转移到细胞膜的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID

NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中所述缀合物还包含磷脂酰胆碱分子。

[0219] 在本公开的实施方案中，提供本文所述用于将功能分子转移到细胞膜的包含连接功能分子的重组β螺旋蛋白的细胞穿透缀合物，其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0220] 在本公开的实施方案中，提供一种将功能分子转移到细胞中的方法，该方法包括：(a)将所述功能分子连接到重组β螺旋蛋白以获得缀合物；(b)使所述缀合物与至少一种细胞接触，其中接触缀合物将功能分子转移到细胞中，并且其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。在实施方案中，该方法还包括在步骤(c)、步骤(d)后检测所述细胞内的所述缀合物的转移。

[0221] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是五肽重复蛋白。

[0222] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白包含具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的串联重复的五肽。

[0223] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

[0224] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:1所示的序列的AlbG。

[0225] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:2所示的序列的EfsQNR。

[0226] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:3所示的序列的抗冻蛋白。

[0227] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:4所示的序列的抗冻蛋白。

[0228] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:5所示的序列的抗冻蛋白。

[0229] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:6所示的序列的QNRB1。

[0230] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:7所示的序列的UDP N乙酰氨基葡萄糖酰基转移酶。

[0231] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:8所示的序列的NP275。

[0232] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:9所示的序列的果胶裂解酶C。

[0233] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白是具有SEQ ID NO:10所示的序列的果胶酸裂合酶。

[0234] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺

旋蛋白是具有SEQ ID NO:11所示的序列的碳酸酐酶。

[0235] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中β螺旋蛋白具有SEQ ID NO:12所示的序列。

[0236] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0237] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是染料。

[0238] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是药物。

[0239] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是金属。

[0240] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是药物金属络合物。

[0241] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是蛋白。

[0242] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是酶。

[0243] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是抗体。

[0244] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是核酸。

[0245] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是多糖。

[0246] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是核定位信号。

[0247] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子是纳米粒子。

[0248] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中功能分子通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

[0249] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中细胞选自真核细胞、原核细胞及其组合。

[0250] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中细胞是原核细胞。

[0251] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中细胞是真核细胞。

[0252] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中β螺旋蛋白由SEQ ID NO:1所示的序列表示,并且功能分子是NHS-香豆素染料。

[0253] 在本公开的实施方案中,提供本文所述将功能分子转移到细胞中的方法,其中β螺

旋蛋白由SEQ ID NO:2所示的序列表示，并且功能分子是NHS-香豆素染料。

[0254] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中β螺旋蛋白由SEQ ID NO:2所示的序列表示，并且功能分子的钌金属络合物。

[0255] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中该方法用于细胞标记。

[0256] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞中的方法，其中该方法用于将功能分子传递到细胞中。

[0257] 在本公开的实施方案中，提供一种将功能分子转移到细胞器的方法，该方法包括：(a)将所述功能分子连接到重组β螺旋蛋白以获得缀合物；(b)将选自SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16的任何一个信号序列进一步结合到缀合物中；(c)使步骤(b)的缀合物接触至少一种细胞；其中接触步骤(b)的缀合物将功能分子转移到细胞内的细胞器，其中细胞器选自细胞核、内质网和线粒体，并且其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。在实施方案中，该方法还包括在步骤(c)、步骤(d)后检测所述细胞内的所述缀合物的转移。

[0258] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞器的方法，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

[0259] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞器的方法，其中信号序列如SEQ ID NO:14所示。

[0260] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞的内质网的方法，其中信号序列如SEQ ID NO:15所示。

[0261] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞的线粒体的方法，其中信号序列如SEQ ID NO:16所示。

[0262] 在本公开的实施方案中，提供本文所述将功能分子转移到细胞器的方法，其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0263] 在本公开的实施方案中，提供本文所述将功能分子转移到P-钙黏着蛋白过度表达的乳腺癌细胞的方法，该方法包括：(a)使功能分子连接重组β螺旋蛋白以获得缀合物；(b)还将SEQ ID NO:17所示的信号序列引入缀合物中；(c)使步骤(b)的缀合物接触至少一种P-钙黏着蛋白过度表达的乳腺癌细胞，其中接触步骤(b)的缀合物将功能分子转移到P-钙黏着蛋白过度表达的乳腺癌细胞中，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。在实施方案中，该方法还包括在步骤(c)、步骤(d)后检测所述细胞内的所述缀合物的转移。

[0264] 在本公开的实施方案中，提供本文所述将功能分子转移到P-钙黏着蛋白过度表达的乳腺癌细胞的方法，其中重组β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11和SEQ ID NO:12。

[0265] 在本公开的实施方案中，提供本文所述将功能分子转移到P-钙黏着蛋白过度表达

的乳腺癌细胞的方法，其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0266] 在本公开的实施方案中，提供本文所述的方法，其中该方法用于将功能分子定向递送到细胞中。

[0267] 在本公开的实施方案中，提供本文所述的方法，其中该方法用于标记细胞。

[0268] 在本公开的实施方案中，提供本文所述的方法，其中该方法用于在细胞中靶向递送药物。

[0269] 在本公开的实施方案中，提供本文所述的细胞穿透缀合物，其中重组β螺旋蛋白用于将功能分子传递到细胞中，并且其中功能分子选自染料、药物、金属、药物金属络合物、蛋白、酶、抗体、核酸、多糖、核定位信号、纳米粒子及其组合。

[0270] 在本公开的实施方案中，提供了一种将功能分子转移至细胞内存在的肌动蛋白的方法，该方法包括：(a)将所述功能分子连接到重组β螺旋蛋白以获得缀合物；(b)还将SEQ ID NO:22所示的信号序列引入缀合物中；(c)使步骤(b)的缀合物接触至少一种细胞；其中接触步骤(b)中的缀合物将所述功能分子转移到所述细胞中，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。在实施方案中，该方法还包括在步骤(c)、步骤(d)后检测所述细胞内的所述缀合物的转移。

[0271] 在本公开的实施方案中，提供了一种将功能分子转移至细胞内存在的微管蛋白的方法，该方法包括：(a)将所述功能分子连接到重组β螺旋蛋白以获得缀合物；(b)还将SEQ ID NO:23所示的信号序列引入缀合物中；(c)使步骤(b)的缀合物接触至少一种细胞；其中接触步骤(b)中的缀合物将所述功能分子转移到P-钙黏着蛋白过度表达的乳腺癌细胞中，其中β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。在实施方案中，该方法还包括在步骤(c)、步骤(d)后检测所述细胞内的所述缀合物的转移。

[0272] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中重组β螺旋蛋白用于细胞穿透。

[0273] 在本公开的实施方案中，提供本文描述的细胞穿透缀合物，其中重组β螺旋蛋白用于细胞标记。

[0274] 具有核酸的细胞穿透分子

[0275] 在实施方案中，本发明提供了解决基因治疗所面临的作为功能分子的核酸片段的递送问题的解决方案。本文件公开了一种缀合物，其包含与可以穿透细胞膜的核酸分子连接的重组β螺旋蛋白。在一个实施方案中，连接的重组β螺旋蛋白可以通过接头元件、分子、部分或部分与一个或多个核酸分子，合适地为单个核酸分子连接。

[0276] 显示包含重组β螺旋蛋白、接头和质粒的缀合物成功穿透细胞膜以建立形成质粒一部分的基因的表达。缀合物避免了内吞作用的途径，并通过直接穿透膜而穿过细胞膜。因此，缀合物克服了通过内吞作用进入所面临的问题，同时有效地穿透了细胞膜。

[0277] 在实施方案中，本发明公开了一种能够穿透细胞膜的缀合物。在一些实施方案中，缀合物包含重组β螺旋蛋白，其长度在5nm-25nm的范围内，合适地为10nm-15nm，并且宽度在1nm-5nm的范围内，合适地为1nm-3nm，通过接头连接至核酸分子。在一个实施方案中，重组β螺旋蛋白可以是具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的五肽重复蛋白。在本公开的

另外实施方案中，重组β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11和SEQ ID NO:12。

[0278] 根据本公开，在一个实施方案中，重组β螺旋蛋白通过接头元件、分子、部分或部分与核酸分子连接。在实施方案中，接头元件可以是通过共价或非共价键的直接连接。当接头是接头分子时，它可以选自蛋白质、金属缀合物、药物-金属缀合物、DNA结合结构域和核酸嵌入分子。在一个实施方案中，包含缀合物的一部分的核酸分子包含至少一个目的基因。在一个实施方案中，核酸分子与包含重组β螺旋蛋白和接头的复合物连接以形成本发明中公开的缀合物。在一个实施方案中，缀合物能够通过穿透细胞膜进入细胞，并且形成核酸分子的一部分的基因能够在细胞内部表达。

[0279] 本发明还公开了使用本发明的缀合物在细胞内转移核酸分子的方法。在一个实施方案中，所公开的方法可以进一步用于基因治疗技术，以促进用于在细胞中递送目的基因的非病毒方法，从而有效地补偿细胞内的缺陷基因或蛋白质。

[0280] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。

[0281] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

[0282] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是五肽重复蛋白。

[0283] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白包含具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的串联重复的五肽。

[0284] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中接头是选自蛋白、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子。

[0285] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、

SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中接头是选自蛋白、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子。

[0286] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中所述接头通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

[0287] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中接头是选自蛋白、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子，并且其中所述接头通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

[0288] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中所述核酸分子包含至少一种目的基因。

[0289] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中所述核酸分子包含至少一种目的基因。

[0290] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中接头是选自蛋白、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子，并且其中所述核酸分子包含至少一种目的基因。

[0291] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:1所示的序列的AlbG。

[0292] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少一种接头；和(c)至少一种核酸分子，其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内，并且其中β螺旋蛋白是具有SEQ ID NO:2所示的序列的EfsQNR。

[0293] 在本公开的实施方案中，提供缀合物，包含：(a)至少一种重组β螺旋蛋白；(b)至少

一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:3所示的序列的抗冻蛋白。

[0294] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:4所示的序列的抗冻蛋白。

[0295] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:5所示的序列的抗冻蛋白。

[0296] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:6所示的序列的QNRB1。

[0297] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:7所示的序列的UDP N乙酰氨基葡萄糖酰基转移酶。

[0298] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:8所示的序列的NP275.

[0299] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:9所示的序列的果胶裂解酶C。

[0300] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:10所示的序列的果胶酸裂合酶。

[0301] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:11所示的序列的碳酸酐酶。

[0302] 在本公开的实施方案中,提供缀合物,包含:(a)至少一种重组β螺旋蛋白;(b)至少一种接头;和(c)至少一种核酸分子,其中至少一种重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内,并且宽度在1nm至5nm、合适地在1nm-3nm的范围内,并且其中β螺旋蛋白是具有SEQ ID NO:12所示的序列的果胶裂解酶A。

[0303] 在本公开的实施方案中，提供一种将核酸分子转移到细胞中的方法，该方法包括：(i)使核酸分子连接包含下列的缀合物：(a)至少一种重组β螺旋蛋白；和(b)至少一种接头，以获得缀合物；和(ii)使缀合物接触至少一种细胞，其中接触所述缀合物将所述核酸分子转移到所述细胞中，并且其中重组β螺旋蛋白的长度在5nm至25nm、合适地在10nm-15nm的范围内，并且宽度在1nm至5nm、合适地在1nm-3nm的范围内。

[0304] 在本公开的实施方案中，提供本文所述将核酸分子转移到细胞中的方法，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合。

[0305] 在本公开的实施方案中，提供本文所述将核酸分子转移到细胞中的方法，其中β螺旋蛋白是五肽重复蛋白。

[0306] 在本公开的实施方案中，提供本文所述将核酸分子转移到细胞中的方法，其中β螺旋蛋白包含具有共有序列$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$的串联重复的五肽。

[0307] 在本公开的实施方案中，提供本文所述将核酸分子转移到细胞中的方法，其中接头是选自蛋白、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子。

[0308] 在本公开的实施方案中，提供本文所述将核酸分子转移到细胞中的方法，其中β螺旋蛋白由选自以下的序列表示：SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12及其组合，并且其中接头是选自蛋白、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子。

[0309] 在本公开的实施方案中，提供本文所述将核酸分子转移到细胞中的方法，其中接头通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

[0310] 在本公开的实施方案中，提供本文所述将核酸分子转移到细胞中的方法，其中接头是选自蛋白、金属缀合物、药物金属络合物、DNA结合结构域、核酸嵌入分子及其组合的接头分子，并且其中所述接头通过共价键、非共价键及其组合与所述重组β螺旋蛋白连接。

[0311] 在本公开的实施方案中，提供本文所述将核酸分子转移到细胞中的方法，其中细胞是原核细胞或真核细胞。

[0312] 在本公开的实施方案中，提供本文描述的缀合物，其中缀合物用作转染剂。

[0313] 在本公开的实施方案中，提供本文描述的缀合物，其中缀合物用于基因疗法。

[0314] 实施例-具有功能分子的细胞穿透分子

[0315] 现在将通过工作实例来说明本发明，所述工作实例旨在说明本发明的工作，而并非旨在限制性地暗示对本发明范围的任何限制。除非另有定义，否则本文中使用的所有技术和科学术语具有与本公开所属领域的普通技术人员通常理解的相同含义。尽管与本文描述的那些类似或等同的方法和材料可以用于所公开的过程和组合物的实践中，但是本文描述示例性的过程、装置和材料。应当理解，本公开不限于特定的方法和所描述的实验条件，因为这样的过程和条件可以变化。

[0316] 在以下实施例中，提供了进行质粒和蛋白质研究的方法和方案。还提供了用于获得缀合物、蛋白质标记的方案，以及使用蛋白质-药物缀合物和蛋白质-标记缀合物进行细

胞渗透研究的方法。结果部分具体描述本发明中公开的缀合物的细胞渗透能力的概念证明。根据公开的方法,已经使用本发明的缀合物进行了不同细胞系的体外标记。还已经进行了用药物-蛋白质缀合物的测定以研究缀合物通过细胞渗透增强药物吸收的能力。

[0317] 材料和方法

[0318] 染料Hoechst 33342购自Invitrogen。NHS-香豆素,ATTO 520-NHS,ATTO 390-NHS和ATTO 647N-NHS均购自Sigma。用于紫外-可见分光光度法和CD光谱的有机溶剂和试剂购自Sigma和Merck。用于研究质粒表达和蛋白质纯化的试剂购自Sigma Aldrich和Merck。MTT测定所需的试剂购自MP Biomedicals。钌金属络合物购自Sigma。

[0319] 实施例1

[0320] 质粒/蛋白质研究-含有SEQ ID NO:19中所示的AlbG基因的质粒是通过先前公开的方法获得的(Vetting et al.Acta Crystallogr.Sect.F Struct.Biol.Cryst.Commun.2011:67(3):296-302,其内容通过引用并入本文,并在下面提供具体细节)。

[0321] 使用X albilineans(ATCC 29184;Pieretti et al.,BMC Genomics,2009,10:616,1-15)染色体DNA作为模板,通过标准PCR技术扩增AlbG的开放阅读框。使用分别含有NdeI和XhoI限制位点的寡核苷酸AlbGF(5'-ATCCCGCTCATATGCCGGCCAAGACCCTTG-3')和AlbGR(5'-ATCCCGCTCTCGAGTCAATCGGACAGCTCGATATC-3')。将PCR片段克隆到pET-28a(+)中,并在大肠杆菌菌株BL21(DE3)中表达带有凝血酶可裂解的N末端His6标签的重组AlbG。对于摇瓶生长,向补充有卡那霉素(35μg/ml)的1升Luria肉汤培养基中接种10ml过夜培养物,并在37℃下孵育。使培养物生长至对数中期(A_{600}为～0.8),冷却至20℃,用0.5mM IPTG诱导,然后在20℃孵育过夜。所有纯化程序均在4℃下进行。通过3000g离心收集细胞,将其重新悬浮在缓冲液A[50mM Tris-HCl pH 7.8中,其中包含300mM NaCl、蛋白酶抑制剂、溶菌酶(5μg/ml)和DNase I(0.1μg/ml)]并搅拌20分钟。然后通过超声裂解细胞,并通过以10000g离心30分钟除去细胞碎片。将上清液加载到预先用缓冲液A平衡过的镍-亚硝酸三乙酸镍(Ni-NTA)柱上,并用十倍柱体积的相同缓冲液洗涤。结合的蛋白质用线性0至0.3M咪唑梯度洗脱,合并峰部分并浓缩。

[0322] SEQ ID NO:20所示的包含EfsQNR基因的质粒是通过先前公开的方法获得的。(Hegde et al.,Antimicrob.Agents Chemother.2011 Jan;55(1):110-117,其内容通过引用并入本文,并在下面提供详细信息)。

[0323] 通过使用粪肠球菌V583(ATCC 700802;ENTFA 226185)染色体DNA作为模板,通过标准PCR技术扩增EfsQNR的开放阅读框。使用分别含有NdeI和XhoI限制性位点的寡核苷酸(5′-ATCCCGCTCATATGAAAATAACTTATCCCTTGCCA-3′)和(5'-ATCCCGCTCTCGAGTTAGGTAATCACCAAACCAAGT-3')。将PCR片段克隆到pET-28a(+)中,并在大肠杆菌菌株BL21(DE3)中表达带有凝血酶可裂解的N末端His6标签的重组EfsQNR。为了摇瓶生长,将补充有卡那霉素(35μg/ml)的1升Luria肉汤培养基用10ml过夜培养物接种并在37℃下孵育。将培养物生长至对数中期(A_{600}～0.8),冷却至20℃,用0.5mM异丙基-β-d-硫代吡喃半乳糖吡喃糖苷(IPTG)诱导,并在20℃下进一步孵育过夜。

[0324] 通过1,200g离心收集细胞,并将其重悬于含有蛋白酶抑制剂、溶菌酶(5μg/ml)和DNase I(0.1μg/ml)的缓冲液A(50mM Tris-HCl[pH 7.8],300mM NaCl)中,并将混合物搅拌

20分钟。然后通过超声处理裂解细胞，并通过在10,000g下离心30分钟去除细胞碎片。将上清液加载到用缓冲液A预平衡的Ni-NTA柱上，并用10倍柱体积的相同缓冲液洗涤。用线性0至0.3M咪唑梯度洗脱结合的蛋白质，合并级分并浓缩。

[0325] 作为比较例，通过Biochemistry，2008，47，1309-1318中的方法获得了包含SEQ ID NO:13所示的TtCuA基因的质粒，该文献通过引用并入本文。

[0326] 实施例2

[0327] 毒性测定-进行了行业标准的MTT测定(来自Sigma Aldrich)来分析β螺旋蛋白-细胞毒性药物缀合物对哺乳动物细胞(例如HeLa细胞)的毒性。通过使用标准方案培养细胞。将一百万个细胞接种在共聚焦平板(1cm培养皿)中，生长6-8h。将细胞毒性药物和EfsQNR蛋白的混合物(以2:1的比例混合)添加到细胞中，并在室温下孵育15分钟至24小时至72小时。温育后，将细胞用PBS洗涤并用MTT处理，并进一步温育24小时。然后洗涤细胞，并分析其在570nm的吸光度值。具有活跃代谢功能的活细胞将MTT转化为紫色formazan产物，其最大吸收度接近570nm。死细胞无法将MTT转化为formazan，因此，通过分析570nm处的吸光度值，可以计算出给定蛋白质或任何其他分子的活细胞百分比。

[0328] 实施例3

[0329] 用荧光染料标记蛋白质-荧光染料被认为是研究缀合物细胞膜穿透能力的功能分子的一个例子。通过在黑暗条件下进行一系列反应，用染料标记重组蛋白。待标记的蛋白质收集在PBS缓冲液(1X和pH 7.3)中，染料的收集浓度是蛋白质的2-3倍。将蛋白质添加到0.1M碳酸钠缓冲液(pH 8.5)中，然后添加染料。将该染料非常缓慢地(每次3μl)添加到含有蛋白质的缓冲液中，该蛋白质保持在冰上并偶尔摇晃。将所得溶液用铝箔包裹，以避开任何光。将该溶液在室温下搅拌1小时，然后通过凝胶过滤进行纯化，即使其通过带有1X PBS缓冲液的脱盐柱或PD 10柱。使用紫外-可见光分光光度法表征所得的标记蛋白质，以确定染料与蛋白质的比例。

[0330] 实施例4

[0331] 细胞对标记蛋白的摄取-通过用不同的标记蛋白(例如AlbG(SEQ ID NO:1)、EfsQNR(SEQ ID NO:2)和作为对比例的TtCuA(SEQ ID NO:13)。TtCuA是一种来自嗜热栖热菌(Thermus thermophilus)的细胞色素氧化酶c蛋白。它由SEQ ID NO:13所示的氨基酸序列表示。将细胞以一百万的浓度接种在共聚焦平板(1cm培养皿)中，并使其生长6-8小时。随后，将标记的蛋白质以不同的浓度添加到细胞中，并在5％CO_2和37℃的标准条件下孵育。然后在3小时和24小时后用HBSS(汉克斯空白盐溶液)或PBS洗涤细胞两次，并在荧光显微镜和/或共聚焦激光扫描显微镜下观察。

[0332] 通过用不同的标记蛋白如AlbG(SEQ ID NO:1)、EfsQNR(SEQ ID NO:2)处理E-Coli和酵母(克鲁维酵母)细胞来检查标记蛋白的摄取。接种大肠杆菌或克鲁维酵母细胞并生长过夜，直到OD达到0.6。将细胞稀释至50倍，然后加入3微摩尔标记的蛋白(缀合物)。将细胞在标准生长条件(37℃摇动条件)下进一步生长24小时。然后将处理的细胞离心并在PBS中洗涤3-4次。然后将细胞分散在100微升PBS中，然后在荧光/孔镜下成像。

[0333] 实施例5

[0334] 药物的细胞摄取-使用本发明的细胞穿透缀合物，将HeLa细胞用于确定药物的摄取增加。将HeLa细胞以10,000的浓度接种在96孔板中，并使其生长6-8小时。将钌金属络合

物和EfsQNR以1:1、1:2和1:3的不同比例混合以形成结合物，然后将其加入细胞中。钌与EfsQNR的比例为1:2的缀合物产生增强的结果。如先前所述，通过MTT测定法监测24小时后的细胞死亡，并与HeLa细胞摄取钌-EfsQNR缀合物的百分比相关。

[0335] 实施例1至5的结果

[0336] 蛋白质和缀合物的表征

[0337] 结构特异性-AlbG蛋白的大小由公开的晶体结构pdb id:2xt2.pdb确定。通过在Pymol®中测量二聚体结构的端到端原子距离，确定长度为10.7nm，宽度(直径)为2.6nm。总形式电荷为-27。交替排列正电荷和负电荷残基:存在精氨酸阶梯(总共20个Arg残基)、赖氨酸(14个残基)和天冬氨酸(30个Asp残基)和谷氨酸(31个Glu残基)。天冬酰胺(20个Asn残基)沿着蛋白质表面阶梯状分布。

[0338] EfsQNR蛋白的大小由发布的晶体结构pdb id:2w7z.pdb确定。通过在Pymol®中测量二聚体结构的端到端原子距离，可以确定蛋白质的长度为10.9nm，宽度(直径)为2.8nm。总形式电荷为-41。交替排列的正电荷和负电荷残基:存在精氨酸阶梯(总共16个Arg残基)、赖氨酸(12个残基)和天冬氨酸(25个Asp残基)和谷氨酸(33个Glu残基)。天冬酰胺(33个Asn残基)沿着蛋白质表面阶梯状分布。

[0339] 为了比较，从公开的晶体结构pdb id:2CuA.pdb确定了TtCuA蛋白的大小。通过测量Pymol®结构的端到端原子距离，可以确定长度为3.3nm，宽度(直径)为1.9nm。该结构包含5个精氨酸残基、4个赖氨酸残基、9个谷氨酸残基和6个天冬酰胺残基，总电荷为-3。

[0340] 图1显示了用于研究的纯化的AlbG蛋白(SEQ ID NO:1)的CD光谱。将蛋白质在PBS中透析，并使用1-5微摩尔浓度的浓度在200-300nm范围内测量CD光谱。如有需要，可用PBS进一步稀释。该蛋白质在220nm处CD的负值表示β片层(β桶或β螺旋结构)的排列。

[0341] 图2显示了用于表达β螺旋蛋白AlbG和EfsQNR的纯化质粒，每份一式三份显示相同的结果(AlbG-1至AlbG-3和EfsQNR-1至EfsQNR-3)。

[0342] 图3显示了聚丙烯酰胺凝胶中AlbG和EfsQNR蛋白的纯化蛋白条带。在进行SDS-凝胶电泳后，注意到单体形式的两种蛋白质的分子量约为30kDa。AlbG和EfsQNR的二聚体形式的大约分子量约为48kDa。蛋白、AlbG和EfsQNR用于与功能分子形成缀合物，然后用于研究细胞穿透功效。使用HeLa细胞进一步测定蛋白质的任何细胞毒性，以确定研究过程(请参见下面的“细胞毒性研究”)。

[0343] 图4显示了EfsQNR蛋白和AlbG蛋白的MALDI TOF质谱分析结果。对于该实验，将蛋白质在水中透析，并使用芥子酸作为具有0.1％TFA和乙腈的基质制备样品。使用1-10微摩尔的蛋白质。结果表明，分子量为约26.5kDa的EfsQNR蛋白和分子量为约23kDa的AlbG蛋白。

[0344] 细胞毒性研究

[0345] 如先前所述，MTT分析用于研究AlbG和EfsQNR蛋白对HeLa细胞的毒性。对于毒性测定，一式三份使用安慰剂，对照、AlbG(30μM)和EfsQNR(30μM)的样品。表1描述用MTT处理后每个样品在570nm的吸光度值。通过观察该值，可以理解，与对照细胞相比，用30μM的AlbG和30μM的EfsQNR处理的细胞显示出95％的活力。这清楚地表明，处于测试浓度的测试蛋白质对HeLa细胞无毒，因此是用于进一步实验的安全浓度。

[0346] 此处提供的表1描述MTT分析中570nm处的吸光度值，用于评估蛋白质的毒性：

[0347] 表1：

[0348]

1	2	3	4	5	6	7	8	9	10	11	12
PLACEBO			对照			EfsQNR			AlbG		
0.049	0.05	0.046	1.341	1.348	1.413	0.901	1.111	0.971	1.002	1.03	0.932

[0349] 该实验的安慰剂是pH 8.0的20mM Tris缓冲液，而所考虑的对照只是HeLa细胞，未添加任何蛋白质或其他底物。

[0350] 标记研究

[0351] 图4显示了通过前面提到的步骤对标记的AlbG和EfsQNR缀合物进行的紫外-可见分光光度法(请参见上面的“用荧光染料标记蛋白质”)。分析紫外-可见光谱以计算两种缀合物的标记率。标记率表示每个蛋白质分子附着的染料数量。通过测量染料的吸光度值(在一个波长取决于用于标记的染料)与蛋白质在280nm的吸光度之比来计算。在该实验中用于标记的染料是NHS-香豆素。对于AlbG蛋白，染料:蛋白分子的比例为1.95。对于EfsQNR，发现染料:蛋白质分子之比为5.5。紫外可见光谱证实了NHS-香豆素染料标记了AlbG和EfsQNR蛋白。在这项研究中，标记的蛋白被进一步用于建立细胞渗透。

[0352] 细胞渗透研究

[0353] 图5显示了用NHS-C标记的AlbG、EfsQNR和TtCuA蛋白(缀合物)处理的HeLa细胞的荧光显微图像。将标记蛋白质(缀合物)的摄入量与未经处理的HeLa细胞进行比较，将其视为本研究的对照。赫斯特蓝(Hoechst Blue)是一种核标记染料，与用于标记蛋白质的NHS-C染料一起使用。蛋白质AlbG和EfsQNR是长度为5nm-25nm，合适地为10nm-15nm，宽度为1nm-5nm，合适地为1nm-3nm的β螺旋蛋白，它们代表了本发明的重组β螺旋蛋白。然而，TtCuA是一种长度为4nm左右的蛋白质，并且由形成β桶的β链组成，从而代表了本发明的重组β螺旋蛋白的比较例。考虑到这两种不同类别的蛋白质，与不同类型的β螺旋蛋白质相比，本发明的缀合物具有优异的细胞穿透能力。

[0354] 在图5中，图A代表仅用核标记染料(Hoechst Blue)处理的对照细胞，表示缺少较淡的染料，图B代表用Hoechst Blue染料和TtCuA-NHSC标记的蛋白质(缀合物)处理的细胞，面板C代表用Hoechst Blue染料和AlbG-NHSC标记的蛋白质(缀合物)处理的细胞，面板D代表用Hoechst Blue染料和EfsQNR-NHSC标记的蛋白质(缀合物)处理的细胞。在观察该图时，可以理解，如图B所示，在细胞内可以看到明显更少的荧光，这表明在这些细胞内较少存在TtCuA-NHSC缀合物，而在图C和图D所代表的细胞内可以看到更大的荧光，这表明在细胞内存在更大的AlbG-NHSC缀合物和EfsQNR-NHSC缀合物。因此，可以理解，与AlbG-NHSC缀合物和EfsQNR-NHSC缀合物相比，TtCuA-NHSC缀合物不能有效地穿透细胞膜以进入细胞内。

[0355] 在比较图C和D时，可以观察到，与AlbG-NHSC缀合物相比，EfsQNR-NHSC缀合物能够更有效地穿透膜。因此，三种缀合物的细胞穿透能力可以按照细胞摄取的递增顺序来概括，如TtCuA-NHSC<AlbG-NHSC<EfsQNR-NHSC。这进一步证明了包含重组β螺旋蛋白的缀合物的细胞穿透能力，所述重组β螺旋蛋白的长度和宽度在5nm-25nm的范围内，合适地为10nm-15nm，而宽度分别在1nm-5nm的范围内，合适地为1nm-3nm。与包含TtCuA的缀合物相比，独立地包含AlbG和EfsQNR的缀合物能够显示出增强的细胞渗透性。

[0356] 由于包含EfsQNR蛋白的缀合物显示出最高的穿透细胞膜的能力,因此对EfsQNR蛋白用不同的染料进行了进一步研究,以增强对哺乳动物细胞的标记。

[0357] 图22显示了通过缀合物与细胞膜的直接相互作用进行的实时细胞进入过程。为了在荧光显微镜下观察细胞,它们用与核结合的染料Hoechst 33342(蓝色,在图22中显示为黑色)和与质膜结合的染料Alexa 594-WGA(红色,在图22中显示为浅)处理过,表明EfsQNR-ATTO-520NHS(绿色,在图22中以浅色圆圈显示)通过避免任何内吞机制的直接机制有效地渗透到细胞内部。

[0358] 哺乳动物细胞的标记

[0359] 图7描述包含与染料缀合的EfsQNR的缀合物在HeLa细胞的标记中的能力。图6描述根据上述标记方法制备的用EfsQNR-ATTO-520NHS标记的细胞。染料Hoechst 33342(蓝色,在图7中显示为黑色)结合到细胞核,染料Alexa 594-WGA(红色,在图7中显示为浅色)与质膜结合。通过将EfsQNR蛋白与绿色荧光染料ATTO-520NHS(在图7中显示为最轻的染料)连接而形成缀合物。可以清楚地观察到,与显示未用缀合物处理的HeLa细胞的对照板A相比,EfsQNR-ATTO-520NHS缀合物穿透HeLa细胞以导致细胞的有效标记(图B)。

[0360] 图8描述用缀合物处理的HeLa细胞的相似结果,该缀合物是通过将EfsQNR蛋白与蓝色荧光染料ATTO-390NHS连接而形成的(在图8中更亮)

[0361] 图9描述用EfsQNR-ATTO-520NHS标记的小胶质细胞的相似结果,所述EfsQNR-ATTO-520NHS是根据上述标记方法制备的。图A显示了来自细胞内部的染料的显着荧光。B图显示了小胶质细胞,该小胶质细胞还用与细胞核结合的染料Hoechst 33342(蓝色,在图9中显示为黑色)和与质膜结合的染料Alexa 594-WGA(红色,在图9中显得更浅)处理,表明EfsQNR-ATTO-520NHS已有效渗透到细胞内部。

[0362] 图10描述用缀合物EfsQNR-ATTO-520NHS标记的角质形成细胞的相似结果,该缀合物是根据上述标记方法制备的。图A显示了来自细胞内部的染料的显着荧光。B图显示了另外用染料Hoechst 33342(蓝色,在图10中显示为黑色)处理的角质形成细胞,其与细胞核结合,表明EfsQNR-ATTO-520NHS已经有效地渗透进入细胞内,并存在于细胞核附近。

[0363] 图11描述用缀合物EfsQNR-ATTO-520NHS标记的SH-SY5Y细胞的相似结果,所述缀合物是根据上述标记方法制备的。

[0364] 图12描述用EfsQNR-ATTO-520NHS标记的小鼠ES细胞的相似结果,该细胞是根据上述标记方法制备的。图12显示了用与细胞核结合的染料Hoechst 33342(蓝色,在图12中显示为黑色)和与质膜结合的染料Alexa 594-WGA(红色,在图12中显得更浅)处理的小鼠ES细胞,表明EfsQNR-ATTO-520NHS已经有效地渗透进入细胞内。

[0365] 非哺乳动物细胞的标记

[0366] 图13和14描述包含与染料缀合的EfsQNR的缀合物在标记大肠杆菌和酵母(克鲁维酵母)细胞中的能力。图13和14分别描述用根据上述标记方法制备的EfsQNR-ATTO-520NHS标记的大肠杆菌和酵母(克鲁维酵母)细胞。

[0367] 用缀合物处理的HeLa细胞的FACS分选

[0368] 图15显示了用缀合物EfsQNR-ATTO-647N标记的缀合物处理的HeLa细胞的标准FACS分选结果。与对照(图A)相比,仅处理细胞10分钟(图B)会导致染料的显着吸收,而对照在图1h(图C)持续增加直至测量的3小时最大值(图D)。

[0369] 药物吸收研究

[0370] 图16描述HeLa细胞中钌金属络合物$Ru(CO)_3Cl$甘氨酸盐的细胞摄取。将钌金属络合物与EfsQNR蛋白缀合以形成由钌金属络合物-EfsQNR组成的缀合物。如上所述，用钌金属络合物-EfsQNR缀合物处理HeLa细胞(参见上文“细胞渗透研究”)，并通过如上所述进行MTT分析检查细胞的活力，从而检测复合物的摄取。通过观察该图(图14)，可以理解，仅用钌金属络合物处理时，细胞摄取仅为0.5％，而用钌金属络合物-EfsQNR蛋白缀合物处理时，该络合物的细胞摄取增加至24.4％。因此，可以确定的是，与单独的钌金属络合物相比，包含EfsQNR蛋白的缀合物促进了50倍的细胞摄取。通过使用本发明的缀合物将许多拯救生命的药物递送至靶标，从而减少药物的剂量并随后减少副作用，可以进一步利用这种增强的摄取。

[0371] 图17显示了哺乳动物细胞(HeLa和HepG2)的活力研究结果，该细胞用包含与化学药物Cisplatin®(顺铂或顺二氨二氯铂(II)(CDDP))相连的EfsQNR蛋白的缀合物培养，图17中标记为“CYDD”。使用标准方案培养哺乳动物细胞(HeLa:A组;或HepG2:B组)。将一百万个细胞接种在共聚焦平板(1cm培养皿)中，生长8小时。然后加入包含Cisplatin®和EfsQNR的缀合物(根据上述方法以2:1的摩尔比制备)，并将培养物在室温下保持15分钟。将细胞在标准生长条件下保持24小时和72小时。然后将细胞用PBS洗涤并加入MTT溶液。形成formazan染料后，将其溶于DMSO中，并测量550nm处的吸光度。从结果显而易见，当根据本发明将药物作为缀合物递送时，在较低的剂量下证明了在两种测试的细胞类型中具有相似的Cisplatin®毒性。

[0372] 注意，在所有测试的实验中，当EfsQNR蛋白替换为AlbG蛋白时，可获得相似的结果。

[0373] 优点

[0374] 总的来说，可以得出结论，包含特定长度和宽度的重组β螺旋蛋白以及功能分子的缀合物可以充当高效的细胞膜穿透缀合物。与内吞进入机制相比，缀合物直接穿透细胞膜的能力提供了优势。在本公开中提供了AlbG-NHSC和EfsQNR-NHSC缀合物的工作实例，其证明了它们穿透细胞膜的能力。EfsQNR-dye缀合物因其有效穿透细胞膜的能力而显示出增强各种哺乳动物和非哺乳动物细胞标记的能力。已显示，包含EfsQNR和/或AlbG的缀合物可增加药物在HeLa和HepG2细胞中的细胞摄取，包括钌金属络合物和Cisplatin®。可以进一步利用这种能力来提高癌细胞对不同抗癌药物的吸收，通过有效地选择性靶向癌细胞并同时减少药物需求，可以促进治疗。减少抗癌药的剂量也可以规避与这类药物的给药有关的副作用。

[0375] 实施例-具有核酸的细胞穿透分子

[0376] 在以下段落中，已经提供了用于通过穿透细胞膜而在细胞内递送具有目的基因的核酸的工作实例。由于使用缀合物作为本发明的一个实施方案，其进一步包含核酸分子，使得递送成为可能。实施例还描述在穿透细胞膜进入细胞内部后目的基因的表达。

[0377] 该实施例描述使用本发明的缀合物的一个实施方案将目的基因转染到细胞中，其中目的基因是编码RFP(红色荧光蛋白)的载体。含有mcherry基因的质粒与蛋白质连接形成缀合物，其中缀合物包含EfsQNR蛋白质和铜[II]菲咯啉。该缀合物用于转染HeLa细胞。

mcherry基因在HeLa细胞中的成功表达被证明是建立结合物穿透细胞膜并转染细胞能力的概念证明。

[0378] 材料与方法

[0379] 菲咯啉铜配合物是商业购买的。

[0380] 实施例6

[0381] 蛋白质研究-按照先前发表的方法(Hegde et al.Antimicrob.Agents and Chemother.2011:55(1):110-7,其通过引用并入本文),进行有关EfsQNR质粒表达,温育条件和纯化的研究,以获得EfsQNR蛋白。使用Bio-Rad试剂盒进行SDS-PAGE(聚丙烯酰胺凝胶电泳),并使用12%聚丙烯酰胺凝胶可视化蛋白质条带。

[0382] 实施例7

[0383] 缀合物的制备

[0384] 如本发明所公开的,制备了用于转染细胞的缀合物,该缀合物包含EfsQNR(SEQ ID NO:2)、铜[II]菲咯啉和具有目的基因的质粒。

[0385] 图18描述携带mcherry基因(SEQ ID NO:25)的质粒的载体图,该质粒在本实施例中用于制备缀合物。EfsQNR蛋白(SEQ ID NO:2)是具有五肽重复序列(如SEQ ID NO:18所示)的β螺旋蛋白,其长度在5nm-25nm的范围内,合适地,在10nm-15nm的范围内,并且宽度在1nm-5nm的范围内,合适地1nm-3nm的范围内。铜[II]菲咯啉是在本实施例中用作接头的核酸嵌入剂。

[0386] 将通过控制EfsQNR质粒的表达获得的EfsQNR蛋白与铜[II]菲咯啉复合物复合,以获得包含EfsQNR和铜[II]菲咯啉的复合物。根据本发明的一个实施方案,将获得的复合物进一步与具有mcherry基因的质粒反应以获得缀合物。获得的缀合物用于转染HeLa细胞。

[0387] 图19描述用于制备可用于转染的缀合物的两种方案。方案1建议使用铜[II]菲咯啉作为与蛋白质形成复合物的接头。所获得的复合物进一步与具有目的基因的质粒反应以获得用于转染的缀合物。

[0388] 方案2建议使用DNA结合蛋白作为与蛋白质形成复合物的接头。所获得的复合物进一步与具有目的基因的质粒反应以获得用于转染的缀合物。

[0389] 在本示例中,铜[II]菲咯啉用作接头以形成复合物,重组β螺旋蛋白-EfsQNR(SEQ ID NO:2)与包含mcherry基因的质粒一起使用。以下方案用于制备复合物和缀合物:

[0390] 将蛋白质EfsQNR与摩尔比为1:2的铜[II]菲咯啉混合,并在室温下温育30分钟以获得复合物。将含有mcherry基因的质粒与上一步获得的复合物混合,并在4℃下孵育30分钟,以获得缀合物。如此获得的缀合物用于转染HeLa细胞。

[0391] 实施例8

[0392] 使用缀合物转染HeLa细胞

[0393] 图20描述使用实施例9中获得的缀合物进行转染实验的方案。在本实施例中,使用以下条件进行转染实验。

[0394] 将实施例8中获得的包含mcherry质粒的缀合物添加到HeLa细胞中(接种后8小时)。在37℃和5%CO_2条件下孵育48小时,使细胞生长。孵育后,在共聚焦显微镜下观察细胞。

[0395] 图21描述转染实验后HeLa细胞的共聚焦显微镜图像。可以观察到RFP的表达为

HeLa细胞内部较亮的(红色)点,从而证明了本文公开的缀合物的细胞穿透力和转染能力。

[0396] 尽管本文提供的实施例显示了蛋白EfsQNR(SEQ ID NO:2)在制备用于转染的缀合物中的用途,但这些蛋白具有SEQ ID NO:1和SEQ ID NO:3至SEQ ID NO:12也可以有效地用于形成缀合物。同样,重组β螺旋蛋白具有如SEQ ID NO:18所示的五肽重复序列,并且长度在5nm-25nm的范围内,合适地为10nm-15nm,宽度在1nm-5nm的范围内,合适地为1nm-3nm,可用于制备本公开中所述的缀合物。类似地,本实施例描述使用铜[II]菲咯啉作为接头,但是其他分子(例如DNA结合蛋白、金属缀合物、药物金属缀合物和其他核酸嵌入分子)也可以用于形成缀合物以有效用作转染剂。一类DNA结合蛋白,即分子量小于12kDa的锌指蛋白也可以用于形成缀合物。SEQ ID NO:24描述一种可用作接头的锌指蛋白。在本公开中,携带mcherry基因的质粒(图18)已用于制备用于转染的缀合物,并且mcherry基因的表达已被证明是一种概念证明。考虑到基本上任何用于转染目的目的基因都可以与本文公开的复合物复合以形成用于基因治疗的缀合物。

[0397] 优点

[0398] 本公开内容提供了包含目的基因的缀合物,所述目的基因具有穿透细胞膜并表达目的基因的能力。因此,该缀合物可以用于转染并且具有用于基因治疗的巨大潜力。众所周知,将基因递送至细胞是基因治疗领域的主要挑战,所公开的缀合物在该领域开辟了新途径。所公开的缀合物易于制备并且可以与多种用于基因治疗的基因复合。

[0399] 尽管这里已经详细公开了本发明的特定实施例,但是这仅是示例性的并且仅出于说明的目的。前述实施例并非旨在限制本发明的范围。发明人认为可以在不脱离本发明的精神和范围的情况下对本发明进行各种替换、变更和修改。

[0400] 表2:序列表

[0401]

SEQ ID	序列
SEQ ID NO: 1-蛋白	MPAKTLESKDYCGESFVSEDRSGQSLESI RFEDCTFRQCNFTEAELNRCKFRECEFVD CNLSLISIPQTSFMEVRFVDCKMLGVNWT SAQWPSVKMEGALSFERCILNDSLFYGLY LAGVKMVECRIHDANFTEADCEDADFTQ SDLKGSTFHNTKLTGASFIDAVNYHIDIFH NDIKRARFSLPEAASLLNSLDIELSD
SEQ ID NO: 2-蛋白	GSHMKITYPLPPNLPEQLPLLTNCQLEDE AILENHLYQQIDLPNQEVRNLVFRDAVFD HLSLANGQFASFDCSNVRFEACDFSNVE WLSGSFHRVTFLRCNLTGTNFADSYLKD CLFEDCKADYASFRFANFNLVHFNQTRL VESEFFEVTWKKLLLEACDLTESNWLNTS LKGLDFSQNTFERLTFSPNYLSGLKVTPE QAIYLASALGLVIT
SEQ ID NO: 3-蛋白	QCTGGADCTSCTGACTGCGNCPNAVTCT NSQHCVKANTCTGSTDCNTAQTCTNSKD CFEANTCTDSTNCYKATACTNSSGCPGH
SEQ ID NO: 4-蛋白	GYSCRAVGVDGRAVTDIQGTCHAKATGA

[0402]

	GAMASGTSEPGSTSTATATGRGATARSTS TGRGTATTTATGTASATSNAIGQGTATTT ATGSAGGRATGSATTSSSASQPTQTQTIT GPGFQTAKSFARNTATTTVTASHHHHHH
SEQ ID NO: 5-蛋白	DGSCTNTNSQLSANSKCEKSTLTNCYVD KSEVYGTTCTGSRFDGVTITTSTSTGSRIS GPGCKISTCIITGGVPAPSAAC KISGCTFSAN
SEQ ID NO: 6-蛋白	GSHMALALVGEKIDRNRFTGEKIENSTFF NCDFSGADLSGTEFIGCQFYDRESQKGCN FSRAMLKDAIFKSCDLSMADFRNSSALGI EIRHCRAQGADFRGASFMNMITTRTWFCS AYITNTNLSYANFSKVVLEKCELWENRWI GAQVLGATFSGSDLSGGEFSTFDWRAAN FTHCDLTNSELGDLDIRGVDLQGVKLDN YQASLLMERLGIAVIG
SEQ ID NO: 7-蛋白	MIDKSAFVHPTAIVEEGASIGANAHIGPFC IVGPHVEIGEGTVLKSHVVVNGHTKIGRD NEIYQFASIGEVNQDLKYAGEPTRVEIGD RNRIRESVTIHRGTVQGGGLTKVGSDNLL MINAHIAHDCTVGNRCILANNATLAGHV SVDDFAIIGGMTAVH

[0403]

	QFCIIGAHVMVGGCSGVAQDVPPYVIAQ GNHATPFGVNIEGLKRRGFSREAITAIRNA YKLIYRSGKTLDEVKPEIAELA ETYPEVKAFTDFFARSTRGLIR
SEQ ID NO: 8-蛋白	MGSSHHHHHHSSGLVPRGSHMDVEKLRQ LYAAGERDFSIVDLRGAVLENINLSGAILH GAMLDEANLQQANLSRADLSGATLNGA DLRGANLSKADLSDAILDNAILEGAILDE AVLNQANLKAANLEQAILSHANIREADLS EANLEAADLSGADLAIA DLHQANLHQAALERANLTGANLEDANLE GTILEGGNNNLAT
SEQ ID NO: 9-蛋白	ATDTGGYAATAGGNVTGAVSKTATSMQ DIVNIIDAARLDANGKKVKGGAYPLVITY TGNEDSLINAAAANICGQWSKDPRGVEIK EFTKGITIIGANGSSANFGIWIKKSSDVVV QNMRIGYLPGGAKDGDMIRVDDSPNVW VDHNELFAANHECDGTPDNDTTFESAVDI KGASNTVTVSYNYIHGVKKVGLDGSSSS DTGRNITYHHNYYNDVNARLPLQRGGLV HAYNNLYTNITGSGLNVRQNGQALIENN WFEKAINPVTSRYDGKNFGTWVLKGNNI

[0404]

	TKPADFSTYSITWTADTKPYVNADSWTST GTFPTVAY NYSPVSAQCVKDKLPGYAGVGKNLATLT STACK
SEQ ID NO: 10-蛋白	VGTNTGGVLVITDTIIVKSGQTYDGKGIKI IAQGMGDGSQSENQKPIFKLEKGANLKN VIIGAPGCDGIHCYGDNVVENVVWEDVG EDALTVKSEGVVEVIGGSAKEAADKVFQ LNAPCTFKVKNFTATNIGKLVRQNGNTTF KVVIYLEDVTLNNVKSCVAKSDSPVSEL WYHNLNVNNCKTLFEFPSQSQIHQY
SEQ ID NO: 11-蛋白	QEITVDEFSNIRENPVTPWNPEPSAPVIDPT AYIDPQASVIGEVTIGANVMVSPMASIRS DEGMPIFVGDRSNVQDGVVLHALETINEE GEPIEDNIVEVDGKEYAVYIGNNVSLAHQ SQVHGPAAVGDDTFIGMQAFVFKSKVGN NCVLEPRSAAIGVTIPDGRYIPAGMVVTS QAEADKLPEVTDDYAYSHTNEAVVYVN VHLAEGYKETS
SEQ ID NO: 12-蛋白	VGVSGSAEGFAKGVTGGGSATPVYPDTID ELVSYLGDDEARVIVLTKTFDFTDSEGTT TGTGCAPWGTASACQVAIDQDDWCENY

[0405]

	EPDAPSVSVEYYNAGTLGITVTSNKSLIGE GSSGAIKGKGLRIVSGAENIIIQNIAVTDIN PKYVWGGDAITLDD CDLVWIDHVTTARIGRQHYVLGTSADNR VSLTNNYIDGVSDYSATCDGYHYWAIYL DGDADLVTMKGNYIYHTSGRSPKVQDNT LLHAVNNYWYDISGHAFEIGEGGYVLAE GNVFQNVDTVLETYEGEAFTVPSSTAGEV CSTYLGRDCVINGFGSSGT FSEDSTSFLSDFEGKNIASASAYTSVASRV VANAGQGNL
SEQ ID NO: 13-蛋白	AYTLATHTAGVIPAGKLERVDPTTVRQEG PWADPAQAVVQTGPNQYTVYVLAFAFG YQPNPIEVPQGAEIVFKITSPDVIHGFHVE GTNINVEVLPGEVSTVRYTFKRPGEYRIIC NQYCGLGHQNMFGTIVVKE
SEQ ID NO: 14-肽	PAAKRVKCD
SEQ ID NO: 15-肽	YPYDVPDYAKDEL
SEQ ID NO: 16-肽	MLSLRQSIRFFKPATRTLCSSRYLL
SEQ ID NO: 17-肽	LSTAADMQGVVTDGMASGLDKDYLKPD D
SEQ ID NO: 18- 共	$(STAV)_1(DN)_2(LF)_3(STR)_4(G)_5$

[0406]

有序列	
SEQ ID NO: 19-DNA	ATGCCGGCGAAAACCCTGGAAAGCAAA GATTATTGCGGCGAAAGCTTTGTGAGCG AAGATCGCAGCGGCCAGAGCCTGGAAA GCATTCGCTTTGAAGATTGCACCTTTCG CCAGTGCAACTTTACCGAAGCGGAACT GAACCGCTGCAAATTTCGCGAATGCGA ATTTGTGGATTGCAACCTGAGCCTGATT AGCATTCCGCAGACCAGCTTTATGGAA GTGCGCTTTGTGGATTGCAAAATGCTGG GCGTGAACTGGACCAGCGCGCAGGCGG GCGCGCTGAGCTTTGAACGCTGCATT CTGAACGATAGCCTGTTTTATGGCCTGT ATCTGGCGGGCGTGAAAATGGTGGAAT GCCGCATTCATGATGCGAACTTTACCGA AGCGGATTGCGAAGATGCGGATTTTAC CCAGAGCGATCTGAAAGGCAGCACCTT TCATAACACCAAACTGACCGGCGCGAG CTTTATTGATGCGGTGAACTATCATATT GATATTTTTCATAACGATATTAAACGCG CGCGCTTTAGCCTGCCGGAAGCGGCGA GCCTGCTGAACAGCCTGGATATTGAACT

[0407]

	GAGCGAT
SEQ ID NO: 20-DNA	GGCAGCCATATGAAAATTACCTATCCGC
	TGCCGCCGAACCTGCCGGAACAGCTGC
	CGCTGCTGACCAACTGCCAGCTGGAAG
	ATGAAGCGATTCTGGAAAACCATCTGT
	ATCAGCAGATTGATCTGCCGAACCAGG
	AAGTGCGCAACCTGGTGTTTCGCGATGC
	GGTGTTTGATCATCTGAGCCTGGCGAAC
	GGCCAGTTTGCGAGCTTTGATTGCAGCA
	ACGTGCGCTTTGAAGCGTGCGATTTTAG
	CAACGTGGAATGGCTGAGCGGCAGCTT
	TCATCGCGTGACCTTTCTGCGCTGCAAC
	CTGACCGGCACCAACTTTGCGGATAGCT
	ATCTGAAAGATTGCCTGTTTGAAGATTG
	CAAAGCGGATTATGCGAGCTTTCGCTTT
	GCGAACTTTAACCTGGTGCATTTTAACC
	AGACCCGCCTGGTGGAAAGCGAATTTTT
	TGAAGTGACCTGGAAAAAACTGCTGCT
	GGAAGCGTGCGATCTGACCGAAAGCAA
	CTGGCTGAACACCAGCCTGAAAGGCCT
	GGATTTTAGCCAGAACACCTTTGAACGC
	CTGACCTTTAGCCCGAACTATCTGAGCG

[0408]

	GCCTGAAAGTGACCCCGGAACAGGCGA TTTATCTGGCGAGCGCGCTGGGCCTGGT GATTACC
SEQ ID NO: 21-DNA	GCGTATACCCTGGCGACCCATACCGCG GGCGTGATTCCGGCGGGCAAACTGGAA CGCGTGGATCCGACCACCGTGCGCCAG GAAGGCCCGTGGGCGGATCCGGCGCAG GCGGTGGTGCAGACCGGCCCGAACCAG TATACCGTGTATGTGCTGGCGTTTGCGT TTGGCTATCAGCCGAACCCGATTGAAGT GCCGCAGGGCGCGGAAATTGTGTTTAA AATTACCAGCCCGGATGTGATTCATGGC TTTCATGTGGAAGGCACCAACATTAACG TGGAAGTGCTGCCGGGCGAAGTGAGCA CCGTGCGCTATACCTTTAAACGCCCGGG CGAATATCGCATTATTTGCAACCAGTAT TGCGGCCTGGGCCATCAGAACATGTTTG GCACCATTGTGGTGAAAGAA
SEQ ID NO: 22-肽	GDVQKKRWLFETKPLD
SEQ ID NO: 23-肽	VQSKCGSKDNIKHVPGGG
SEQ ID NO: 24-肽	MERPYACPVESCDRRFSDSSNLTRHIRIHT GQKPFQCRICMRNFSRSDHLTTHIRTHTGE

[0409]

	KPFACDICGRKFARSDERKR HTKIHLRQKD
SEQ ID NO: 25-肽	GTGAGCAAGGGCGAGGAGGATAACATG GCCATCATCAAGGAGTTCATGCGCTTCA AGGTGCACATGGAGGGCTCCGTGAACG GCCACGAGTTCGAGATCGAGGGCGAGG GCGAGGGCCGCCCCTACGAGGGCACCC AGACCGCCAAGCTGAAGGTGACCAAGG GTGGCCCCCTGCCCTTCGCCTGGGACAT CCTGTCCCCTCAGTTCATGTACGGCTCC AAGGCCTACGTGAAGCACCCCGCCGAC ATCCCCGACTACTTGAAGCTGTCCTTCC CCGAGGGCTTCAAGTGGGAGCGCGTGA TGAACTTCGAGGACGGCGGCGTGGTGA CCGTGACCCAGGACTCCTCCCTCCAGGA CGGCGAGTTCATCTACAAGGTGAAGCTG CGCGGCACCAACTTCCCCTCCGACGGCC CCGTAATGCAGAAGAAGACCATGGGCT GGGAGGCCTCCTCCGAGCGGATGTACCC CGAGGACGGCGCCCTGAAGGGCGAGAT CAAGCAGAGGCTGAAGCTGAAGGACGG CGGCCACTACGACGCTGAGGTCAAGAC

	CACCTACAAGGCCAAGAAGCCCGTGCA GCTGCCCGGCGCCTACAACGTCAACATC AAGTTGGACATCACCTCCCACAACGAG GACTACACCATCGTGGAACAGTACGAAC GCGCCGAGGGCCGCCACTCCACCGGCG GCATGGACGAGCTGTACAAGTAGTAATC TAGAGGGCCCTATTCTATAGTGTCACC
SEQ ID NO: 26-DNA	ATCCCGCTCATATGCCGGCCAAGACCCT TG
SEQ ID NO: 27-DNA	ATCCCGCTCTCGAGTCAATCGGACAGCT CGATATC
SEQ ID NO: 28-DNA	ATCCCGCTCATATGAAAATAACTTATCCC TTGCCA
SEQ ID NO: 29-DNA	ATCCCGCTCTCGAGTTAGGTAATCACCA AACCAAGT

[0410]

SEQUENCE LISTING
<110> 西卡肿瘤解决方案有限公司
<120> 细胞膜穿透缀合物
<130> P48161WO/DJC/JG
<150> IN201831002945
<151> 2018-01-24
<150> IN201731029916
<151> 2017-08-23
<160> 29
<170> BiSSAP 1.3.6
<210> 1
<211> 200
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> AlbG蛋白的氨基酸序列(Amino acid sequence of AlbG protein)
<400> 1

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

145 150 155 160
Ser Phe Ile Asp Ala Val Asn Tyr His Ile Asp Ile Phe His Asn Asp
165 170 175
Ile Lys Arg Ala Arg Phe Ser Leu Pro Glu Ala Ala Ser Leu Leu Asn
180 185 190
Ser Leu Asp Ile Glu Leu Ser Asp
195 200

<210> 2
<211> 214
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> EfsQNR蛋白的氨基酸序列(Amino acid sequence of EfsQNR protein)
<400> 2
Gly Ser His Met Lys Ile Thr Tyr Pro Leu Pro Pro Asn Leu Pro Glu
1 5 10 15
Gln Leu Pro Leu Leu Thr Asn Cys Gln Leu Glu Asp Glu Ala Ile Leu
20 25 30
Glu Asn His Leu Tyr Gln Gln Ile Asp Leu Pro Asn Gln Glu Val Arg
35 40 45
Asn Leu Val Phe Arg Asp Ala Val Phe Asp His Leu Ser Leu Ala Asn
50 55 60
Gly Gln Phe Ala Ser Phe Asp Cys Ser Asn Val Arg Phe Glu Ala Cys
65 70 75 80
Asp Phe Ser Asn Val Glu Trp Leu Ser Gly Ser Phe His Arg Val Thr
85 90 95
Phe Leu Arg Cys Asn Leu Thr Gly Thr Asn Phe Ala Asp Ser Tyr Leu
100 105 110
Lys Asp Cys Leu Phe Glu Asp Cys Lys Ala Asp Tyr Ala Ser Phe Arg
115 120 125
Phe Ala Asn Phe Asn Leu Val His Phe Asn Gln Thr Arg Leu Val Glu
130 135 140
Ser Glu Phe Phe Glu Val Thr Trp Lys Lys Leu Leu Leu Glu Ala Cys
145 150 155 160
Asp Leu Thr Glu Ser Asn Trp Leu Asn Thr Ser Leu Lys Gly Leu Asp
165 170 175
Phe Ser Gln Asn Thr Phe Glu Arg Leu Thr Phe Ser Pro Asn Tyr Leu

180 185 190

Ser Gly Leu Lys Val Thr Pro Glu Gln Ala Ile Tyr Leu Ala Ser Ala

195 200 205

Leu Gly Leu Val Ile Thr

210

<210> 3

<211> 84

<212> PRT

<213> 黄粉虫(Tenebrio molitor)

<220>

<221>

<222>

<223> 来自黄粉虫的抗冻蛋白的氨基酸序列(Amino acid sequence of anti-freeze protein from Tenebrio molitor)

<400> 3

Gln Cys Thr Gly Gly Ala Asp Cys Thr Ser Cys Thr Gly Ala Cys Thr

1 5 10 15

Gly Cys Gly Asn Cys Pro Asn Ala Val Thr Cys Thr Asn Ser Gln His

20 25 30

Cys Val Lys Ala Asn Thr Cys Thr Gly Ser Thr Asp Cys Asn Thr Ala

35 40 45

Gln Thr Cys Thr Asn Ser Lys Asp Cys Phe Glu Ala Asn Thr Cys Thr

50 55 60

Asp Ser Thr Asn Cys Tyr Lys Ala Thr Ala Cys Thr Asn Ser Ser Gly

65 70 75 80

Cys Pro Gly His

<210> 4

<211> 143

<212> PRT

<213> Rhagium inquisitor

<220>

<221>

<222>

<223> 来自Rhagium inquisitor的抗冻蛋白的氨基酸序列(Acid sequence of anti-freeze protein from Rhagium inquisitor)

<400> 4

Gly Tyr Ser Cys Arg Ala Val Gly Val Asp Gly Arg Ala Val Thr Asp

1 5 10 15

Ile Gln Gly Thr Cys His Ala Lys Ala Thr Gly Ala Gly Ala Met Ala

20 25 30

Ser Gly Thr Ser Glu Pro Gly Ser Thr Ser Thr Ala Thr Ala Thr Gly
35 40 45

Arg Gly Ala Thr Ala Arg Ser Thr Ser Thr Gly Arg Gly Thr Ala Thr
50 55 60

Thr Thr Ala Thr Gly Thr Ala Ser Ala Thr Ser Asn Ala Ile Gly Gln
65 70 75 80

Gly Thr Ala Thr Thr Thr Ala Thr Gly Ser Ala Gly Gly Arg Ala Thr
85 90 95

Gly Ser Ala Thr Thr Ser Ser Ser Ala Ser Gln Pro Thr Gln Thr Gln
100 105 110

Thr Ile Thr Gly Pro Gly Phe Gln Thr Ala Lys Ser Phe Ala Arg Asn
115 120 125

Thr Ala Thr Thr Thr Val Thr Ala Ser His His His His His His
130 135 140

<210> 5

<211> 90

<212> PRT

<213> 杉芽杆虫(Choristoneura fumiferana)

<220>

<221>

<222>

<223> 来自云杉芽杆虫的抗冻蛋白的氨基酸序列(Amino acid sequence of anti-freeze protein from Spruce Budworm)

<400> 5

Asp Gly Ser Cys Thr Asn Thr Asn Ser Gln Leu Ser Ala Asn Ser Lys
1 5 10 15

Cys Glu Lys Ser Thr Leu Thr Asn Cys Tyr Val Asp Lys Ser Glu Val
20 25 30

Tyr Gly Thr Thr Cys Thr Gly Ser Arg Phe Asp Gly Val Thr Ile Thr
35 40 45

Thr Ser Thr Ser Thr Gly Ser Arg Ile Ser Gly Pro Gly Cys Lys Ile
50 55 60

Ser Thr Cys Ile Ile Thr Gly Gly Val Pro Ala Pro Ser Ala Ala Cys
65 70 75 80

Lys Ile Ser Gly Cys Thr Phe Ser Ala Asn
85 90

<210> 6

<211> 217

<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> QNRB1蛋白的氨基酸序列(Amino acid sequence of QNRB1 protein)
<400> 6

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

<210> 7
<211> 262
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> UDP-N-乙酰氨基葡萄糖酰基转移酶蛋白的氨基酸序列(Amino acid sequence of UDP-N-acetylglucosamine acyltransferase protein)

<400> 7

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

260

<210> 8

<211> 201

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 来自点状诺氏菌的NP275蛋白的氨基酸序列(Amino acid sequence of NP275 protein from Nostoc punctiforme)

<400> 8

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

<210> 9

<211> 353

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 果胶裂解酶C的氨基酸序列(Amino acid sequence of pectate lyase C protein)

<400> 9

Ala Thr Asp Thr Gly Gly Tyr Ala Ala Thr Ala Gly Gly Asn Val Thr
1 5 10 15
Gly Ala Val Ser Lys Thr Ala Thr Ser Met Gln Asp Ile Val Asn Ile
20 25 30
Ile Asp Ala Ala Arg Leu Asp Ala Asn Gly Lys Lys Val Lys Gly Gly
35 40 45
Ala Tyr Pro Leu Val Ile Thr Tyr Thr Gly Asn Glu Asp Ser Leu Ile
50 55 60
Asn Ala Ala Ala Ala Asn Ile Cys Gly Gln Trp Ser Lys Asp Pro Arg
65 70 75 80
Gly Val Glu Ile Lys Glu Phe Thr Lys Gly Ile Thr Ile Ile Gly Ala
85 90 95
Asn Gly Ser Ser Ala Asn Phe Gly Ile Trp Ile Lys Lys Ser Ser Asp
100 105 110
Val Val Val Gln Asn Met Arg Ile Gly Tyr Leu Pro Gly Gly Ala Lys
115 120 125
Asp Gly Asp Met Ile Arg Val Asp Asp Ser Pro Asn Val Trp Val Asp
130 135 140
His Asn Glu Leu Phe Ala Ala Asn His Glu Cys Asp Gly Thr Pro Asp
145 150 155 160
Asn Asp Thr Thr Phe Glu Ser Ala Val Asp Ile Lys Gly Ala Ser Asn
165 170 175
Thr Val Thr Val Ser Tyr Asn Tyr Ile His Gly Val Lys Lys Val Gly
180 185 190
Leu Asp Gly Ser Ser Ser Ser Asp Thr Gly Arg Asn Ile Thr Tyr His
195 200 205
His Asn Tyr Tyr Asn Asp Val Asn Ala Arg Leu Pro Leu Gln Arg Gly
210 215 220
Gly Leu Val His Ala Tyr Asn Asn Leu Tyr Thr Asn Ile Thr Gly Ser
225 230 235 240
Gly Leu Asn Val Arg Gln Asn Gly Gln Ala Leu Ile Glu Asn Asn Trp

245 250 255
Phe Glu Lys Ala Ile Asn Pro Val Thr Ser Arg Tyr Asp Gly Lys Asn
260 265 270
Phe Gly Thr Trp Val Leu Lys Gly Asn Asn Ile Thr Lys Pro Ala Asp
275 280 285
Phe Ser Thr Tyr Ser Ile Thr Trp Thr Ala Asp Thr Lys Pro Tyr Val
290 295 300
Asn Ala Asp Ser Trp Thr Ser Thr Gly Thr Phe Pro Thr Val Ala Tyr
305 310 315 320
Asn Tyr Ser Pro Val Ser Ala Gln Cys Val Lys Asp Lys Leu Pro Gly
325 330 335
Tyr Ala Gly Val Gly Lys Asn Leu Ala Thr Leu Thr Ser Thr Ala Cys
340 345 350
Lys

<210> 10
<211> 196
<212> PRT
<213> 卡尔迪纤维素破坏者贝西氏菌(Caldicellulosiruptor bescii)
<220>
<221>
<222>
<223> 来自卡尔迪纤维素破坏者贝西氏菌的果胶酸裂合酶的氨基酸序列(Amino acid sequence of pectate lyase from Caldicellulosiruptor bescii)
<400> 10
Val Gly Thr Asn Thr Gly Gly Val Leu Val Ile Thr Asp Thr Ile Ile
1 5 10 15
Val Lys Ser Gly Gln Thr Tyr Asp Gly Lys Gly Ile Lys Ile Ile Ala
20 25 30
Gln Gly Met Gly Asp Gly Ser Gln Ser Glu Asn Gln Lys Pro Ile Phe
35 40 45
Lys Leu Glu Lys Gly Ala Asn Leu Lys Asn Val Ile Ile Gly Ala Pro
50 55 60
Gly Cys Asp Gly Ile His Cys Tyr Gly Asp Asn Val Val Glu Asn Val
65 70 75 80
Val Trp Glu Asp Val Gly Glu Asp Ala Leu Thr Val Lys Ser Glu Gly
85 90 95
Val Val Glu Val Ile Gly Gly Ser Ala Lys Glu Ala Ala Asp Lys Val
100 105 110
Phe Gln Leu Asn Ala Pro Cys Thr Phe Lys Val Lys Asn Phe Thr Ala

115 120 125

Thr Asn Ile Gly Lys Leu Val Arg Gln Asn Gly Asn Thr Thr Phe Lys
130 135 140

Val Val Ile Tyr Leu Glu Asp Val Thr Leu Asn Asn Val Lys Ser Cys
145 150 155 160

Val Ala Lys Ser Asp Ser Pro Val Ser Glu Leu Trp Tyr His Asn Leu
165 170 175

Asn Val Asn Asn Cys Lys Thr Leu Phe Glu Phe Pro Ser Gln Ser Gln
180 185 190

Ile His Gln Tyr
195

<210> 11

<211> 213

<212> PRT

<213> 嗜热甲烷菌(Methanosarcina thermophila)

<220>

<221>

<222>

<223> 嗜热甲烷菌的碳酸酐酶的氨基酸序列(Amino acid sequence of carbonic anhydrase from Methanosarcina thermophila)

<400> 11

Gln Glu Ile Thr Val Asp Glu Phe Ser Asn Ile Arg Glu Asn Pro Val
1 5 10 15

Thr Pro Trp Asn Pro Glu Pro Ser Ala Pro Val Ile Asp Pro Thr Ala
20 25 30

Tyr Ile Asp Pro Gln Ala Ser Val Ile Gly Glu Val Thr Ile Gly Ala
35 40 45

Asn Val Met Val Ser Pro Met Ala Ser Ile Arg Ser Asp Glu Gly Met
50 55 60

Pro Ile Phe Val Gly Asp Arg Ser Asn Val Gln Asp Gly Val Val Leu
65 70 75 80

His Ala Leu Glu Thr Ile Asn Glu Glu Gly Glu Pro Ile Glu Asp Asn
85 90 95

Ile Val Glu Val Asp Gly Lys Glu Tyr Ala Val Tyr Ile Gly Asn Asn
100 105 110

Val Ser Leu Ala His Gln Ser Gln Val His Gly Pro Ala Ala Val Gly
115 120 125

Asp Asp Thr Phe Ile Gly Met Gln Ala Phe Val Phe Lys Ser Lys Val
130 135 140

Gly Asn Asn Cys Val Leu Glu Pro Arg Ser Ala Ala Ile Gly Val Thr
145 150 155 160
Ile Pro Asp Gly Arg Tyr Ile Pro Ala Gly Met Val Val Thr Ser Gln
165 170 175
Ala Glu Ala Asp Lys Leu Pro Glu Val Thr Asp Asp Tyr Ala Tyr Ser
180 185 190
His Thr Asn Glu Ala Val Val Tyr Val Asn Val His Leu Ala Glu Gly
195 200 205
Tyr Lys Glu Thr Ser
210

<210> 12
<211> 359
<212> PRT
<213> 黑曲霉(Aspergillus niger)
<220>
<221>
<222>
<223> 来自黑曲霉的果胶裂解酶A蛋白的氨基酸序列(Amino acid sequence of Pectin lyase A protein from Aspergillus niger)

<400> 12
Val Gly Val Ser Gly Ser Ala Glu Gly Phe Ala Lys Gly Val Thr Gly
1 5 10 15
Gly Gly Ser Ala Thr Pro Val Tyr Pro Asp Thr Ile Asp Glu Leu Val
20 25 30
Ser Tyr Leu Gly Asp Asp Glu Ala Arg Val Ile Val Leu Thr Lys Thr
35 40 45
Phe Asp Phe Thr Asp Ser Glu Gly Thr Thr Thr Gly Thr Gly Cys Ala
50 55 60
Pro Trp Gly Thr Ala Ser Ala Cys Gln Val Ala Ile Asp Gln Asp Asp
65 70 75 80
Trp Cys Glu Asn Tyr Glu Pro Asp Ala Pro Ser Val Ser Val Glu Tyr
85 90 95
Tyr Asn Ala Gly Thr Leu Gly Ile Thr Val Thr Ser Asn Lys Ser Leu
100 105 110
Ile Gly Glu Gly Ser Ser Gly Ala Ile Lys Gly Lys Gly Leu Arg Ile
115 120 125
Val Ser Gly Ala Glu Asn Ile Ile Ile Gln Asn Ile Ala Val Thr Asp
130 135 140
Ile Asn Pro Lys Tyr Val Trp Gly Gly Asp Ala Ile Thr Leu Asp Asp

145 150 155 160
Cys Asp Leu Val Trp Ile Asp His Val Thr Thr Ala Arg Ile Gly Arg
165 170 175
Gln His Tyr Val Leu Gly Thr Ser Ala Asp Asn Arg Val Ser Leu Thr
180 185 190
Asn Asn Tyr Ile Asp Gly Val Ser Asp Tyr Ser Ala Thr Cys Asp Gly
195 200 205
Tyr His Tyr Trp Ala Ile Tyr Leu Asp Gly Asp Ala Asp Leu Val Thr
210 215 220
Met Lys Gly Asn Tyr Ile Tyr His Thr Ser Gly Arg Ser Pro Lys Val
225 230 235 240
Gln Asp Asn Thr Leu Leu His Ala Val Asn Asn Tyr Trp Tyr Asp Ile
245 250 255
Ser Gly His Ala Phe Glu Ile Gly Glu Gly Gly Tyr Val Leu Ala Glu
260 265 270
Gly Asn Val Phe Gln Asn Val Asp Thr Val Leu Glu Thr Tyr Glu Gly
275 280 285
Glu Ala Phe Thr Val Pro Ser Ser Thr Ala Gly Glu Val Cys Ser Thr
290 295 300
Tyr Leu Gly Arg Asp Cys Val Ile Asn Gly Phe Gly Ser Ser Gly Thr
305 310 315 320
Phe Ser Glu Asp Ser Thr Ser Phe Leu Ser Asp Phe Glu Gly Lys Asn
325 330 335
Ile Ala Ser Ala Ser Ala Tyr Thr Ser Val Ala Ser Arg Val Val Ala
340 345 350
Asn Ala Gly Gln Gly Asn Leu
355

<210> 13
<211> 135
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> TtCuA蛋白的氨基酸序列(Amino acid sequence of TtCuA protein)
<400> 13
Ala Tyr Thr Leu Ala Thr His Thr Ala Gly Val Ile Pro Ala Gly Lys
1 5 10 15
Leu Glu Arg Val Asp Pro Thr Thr Val Arg Gln Glu Gly Pro Trp Ala

20 25 30

Asp Pro Ala Gln Ala Val Val Gln Thr Gly Pro Asn Gln Tyr Thr Val
35 40 45

Tyr Val Leu Ala Phe Ala Phe Gly Tyr Gln Pro Asn Pro Ile Glu Val
50 55 60

Pro Gln Gly Ala Glu Ile Val Phe Lys Ile Thr Ser Pro Asp Val Ile
65 70 75 80

His Gly Phe His Val Glu Gly Thr Asn Ile Asn Val Glu Val Leu Pro
85 90 95

Gly Glu Val Ser Thr Val Arg Tyr Thr Phe Lys Arg Pro Gly Glu Tyr
100 105 110

Arg Ile Ile Cys Asn Gln Tyr Cys Gly Leu Gly His Gln Asn Met Phe
115 120 125

Gly Thr Ile Val Val Lys Glu
130 135

<210> 14

<211> 9

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 用于靶向细胞核的信号序列(Signal sequence for targeting the nucleus of the cell)

<400> 14

Pro Ala Ala Lys Arg Val Lys Cys Asp
1 5

<210> 15

<211> 13

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 靶向细胞内质网的信号序列(Signal sequence for targeting the endoplasmic reticulum of the cell)

<400> 15

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Lys Asp Glu Leu
1 5 10

<210> 16
<211> 25
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> 靶向细胞线粒体的信号序列(Signal sequence for targeting the mitochondria of the cell)
<400> 16
Met Leu Ser Leu Arg Gln Ser Ile Arg Phe Phe Lys Pro Ala Thr Arg
1 5 10 15
Thr Leu Cys Ser Ser Arg Tyr Leu Leu
20 25
<210> 17
<211> 28
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> 靶向表达P-钙黏着蛋白过度表达的乳腺癌细胞的信号序列(Signal sequence for targeting the P-cadherin-over expressing breast cancer cells)
<400> 17
Leu Ser Thr Ala Ala Asp Met Gln Gly Val Val Thr Asp Gly Met Ala
1 5 10 15
Ser Gly Leu Asp Lys Asp Tyr Leu Lys Pro Asp Asp
20 25
<210> 18
<211> 5
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> 五肽重复蛋白中的共有序列(Consensus sequence in a pentapeptide-repeat protein)
<220>
<221> UNSURE

<222> (1)..(1)
<223> Ser or Thr or Ala or Val
<220>
<221> UNSURE
<222> (2)..(2)
<223> Asp or Asn
<220>
<221> UNSURE
<222> (3)..(3)
<223> Leu or Phe
<220>
<221> UNSURE
<222> (4)..(4)
<223> Ser or Thr or Arg
<400> 18
Xaa Xaa Xaa Xaa Gly
1 5
<210> 19
<211> 582
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> AlbG基因的核酸序列(Nucleic acid sequence of AlbG gene)
<400> 19
atgccggcga aaaccctgga aagcaaagat tattgcggcg aaagctttgt gagcgaagat 60
cgcagcggcc agagcctgga aagcattcgc tttgaagatt gcacctttcg ccagtgcaac 120
tttaccgaag cggaactgaa ccgctgcaaa tttcgcgaat gcgaatttgt ggattgcaac 180
ctgagcctga ttagcattcc gcagaccagc tttatggaag tgcgctttgt ggattgcaaa 240
atgctgggcg tgaactggac cagcgcgcag gcgggcgcgc tgagctttga acgctgcatt 300
ctgaacgata gcctgtttta tggcctgtat ctggcgggcg tgaaaatggt ggaatgccgc 360
attcatgatg cgaactttac cgaagcggat tgcgaagatg cggattttac ccagagcgat 420
ctgaaaggca gcacctttca taacaccaaa ctgaccggcg cgagctttat tgatgcggtg 480
aactatcata ttgatatttt tcataacgat attaaacgcg cgcgctttag cctgccggaa 540
gcggcgagcc tgctgaacag cctggatatt gaactgagcg at 582
<210> 20
<211> 642
<212> DNA

<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> EfsQNR基因的核酸序列(nucleic acid sequence of EfsQNR gene)
<400> 20
ggcagccata tgaaaattac ctatccgctg ccgccgaacc tgccggaaca gctgccgctg 60
ctgaccaact gccagctgga agatgaagcg attctggaaa accatctgta tcagcagatt 120
gatctgccga accaggaagt gcgcaacctg gtgtttcgcg atgcggtgtt tgatcatctg 180
agcctggcga acggccagtt tgcgagcttt gattgcagca acgtgcgctt tgaagcgtgc 240
gattttagca acgtggaatg gctgagcggc agctttcatc gcgtgacctt tctgcgctgc 300
aacctgaccg gcaccaactt tgcggatagc tatctgaaag attgcctgtt tgaagattgc 360
aaagcggatt atgcgagctt tcgctttgcg aactttaacc tggtgcattt taaccagacc 420
cgcctggtgg aaagcgaatt ttttgaagtg acctggaaaa aactgctgct ggaagcgtgc 480
gatctgaccg aaagcaactg gctgaacacc agcctgaaag gcctggattt tagccagaac 540
acctttgaac gcctgacctt tagcccgaac tatctgagcg gcctgaaagt gaccccggaa 600
caggcgattt atctggcgag cgcgctgggc ctggtgatta cc 642
<210> 21
<211> 405
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<221>
<222>
<223> TtCuA基因的核酸序列(Nucleic acid sequence of TtCuA gene)
<400> 21
gcgtataccc tggcgaccca taccgcgggc gtgattccgg cgggcaaact ggaacgcgtg 60
gatccgacca ccgtgcgcca ggaaggcccg tgggcggatc cggcgcaggc ggtggtgcag 120
accggcccga accagtatac cgtgtatgtg ctggcgtttg cgtttggcta tcagccgaac 180
ccgattgaag tgccgcaggg cgcggaaatt gtgtttaaaa ttaccagccc ggatgtgatt 240
catggctttc atgtggaagg caccaacatt aacgtggaag tgctgccggg cgaagtgagc 300
accgtgcgct ataccttaa acgcccgggc gaatatcgca ttatttgcaa ccagtattgc 360
ggcctgggcc atcagaacat gtttggcacc attgtggtga aagaa 405
<210> 22
<211> 16
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<221>

<222>

<223> 靶向细胞中肌动蛋白的信号序列(Signal sequence for targeting actin in the cell)

<400> 22

Gly Asp Val Gln Lys Lys Arg Trp Leu Phe Glu Thr Lys Pro Leu Asp
1 5 10 15

<210> 23

<211> 18

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 靶向细胞中微管蛋白的信号序列(Signal sequence for targeting tubulin in the cell)

<400> 23

Val Gln Ser Lys Cys Gly Ser Lys Asp Asn Ile Lys His Val Pro Gly
1 5 10 15

Gly Gly

<210> 24

<211> 90

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 锌指蛋白的氨基酸序列(Amino acid sequence of a zinc finger protein)

<400> 24

Met Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe
1 5 10 15

Ser Asp Ser Ser Asn Leu Thr Arg His Ile Arg Ile His Thr Gly Gln
20 25 30

Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp
35 40 45

His Leu Thr Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala
50 55 60

Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg
65 70 75 80

His Thr Lys Ile His Leu Arg Gln Lys Asp

85 90

<210> 25

<211> 740

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> mcherry基因的核酸序列(Nucleic acid sequence of mcherry gene)

<400> 25

gtgagcaagg gcgaggagga taacatggcc atcatcaagg agttcatgcg cttcaaggtg 60

cacatggagg gctccgtgaa cggccacgag ttcgagatcg agggcgaggg cgagggccgc 120

ccctacgagg gcacccagac cgccaagctg aaggtgacca agggtggccc cctgcccttc 180

gcctgggaca tcctgtcccc tcagttcatg tacggctcca aggcctacgt gaagcacccc 240

gccgacatcc ccgactactt gaagctgtcc ttccccgagg gcttcaagtg ggagcgcgtg 300

atgaacttcg aggacggcgg cgtggtgacc gtgacccagg actcctccct ccaggacggc 360

gagttcatct acaaggtgaa gctgcgcggc accaacttcc cctccgacgg ccccgtaatg 420

cagaagaaga ccatgggctg ggaggcctcc tccgagcgga tgtaccccga ggacggcgcc 480

ctgaagggcg agatcaagca gaggctgaag ctgaaggacg gcggccacta cgacgctgag 540

gtcaagacca cctacaaggc caagaagccc gtgcagctgc ccggcgccta caacgtcaac 600

atcaagttgg acatcacctc ccacaacgag gactacacca tcgtggaaca gtacgaacgc 660

gccgagggcc gccactccac cggcggcatg gacgagctgt acaagtagta atctagaggg 720

ccctattcta tagtgtcacc 740

<210> 26

<211> 30

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 用于AlbG基因的PCR扩增的引物的核酸序列(Nucleic acid sequence of a PCR primer AlbG gene)

<400> 26

atcccgctca tatgccggcc aagacccttg 30

<210> 27

<211> 35

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 用于AlbG基因的PCR扩增的引物的核酸序列(Nucleic acid sequence of a PCR primer AlbG gene)

<400> 27

atcccgctct cgagtcaatc ggacagctcg atatc 35

<210> 28

<211> 35

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 用于PCR扩增EfsQNR基因的引物的核酸序列(Nucleic acid sequence of a PCR primer EfsQNR gene)

<400> 28

atcccgctca tatgaaaata acttatccct tgcca 35

<210> 29

<211> 36

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<221>

<222>

<223> 用于EfsQNR基因的PCR扩增的引物的核酸序列(Nucleic acid sequence of a PCR primer EfsQNR gene)

<400> 29

atcccgctct cgagttaggt aatcaccaaa ccaagt 36

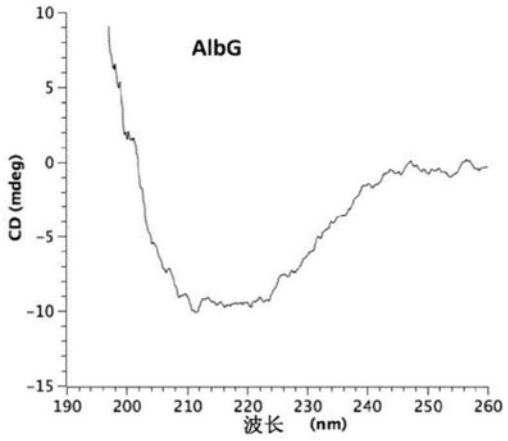

AlbG
CD (mdeg)
10
5
0
-5
-10
-15
190
200
210
220
230
240
250
260
波长 (nm)

图1

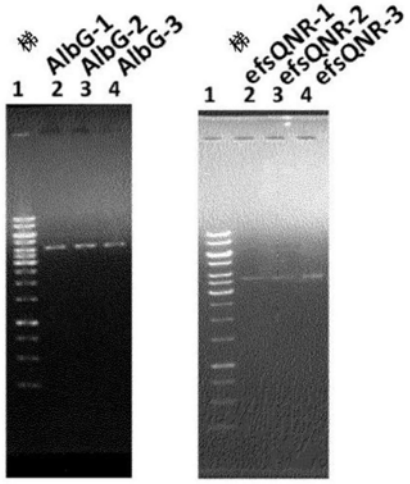

梯
AlbG-1
AlbG-2
AlbG-3
1 2 3 4
梯
efsQNR-1
efsQNR-2
efsQNR-3
1 2 3 4

图2

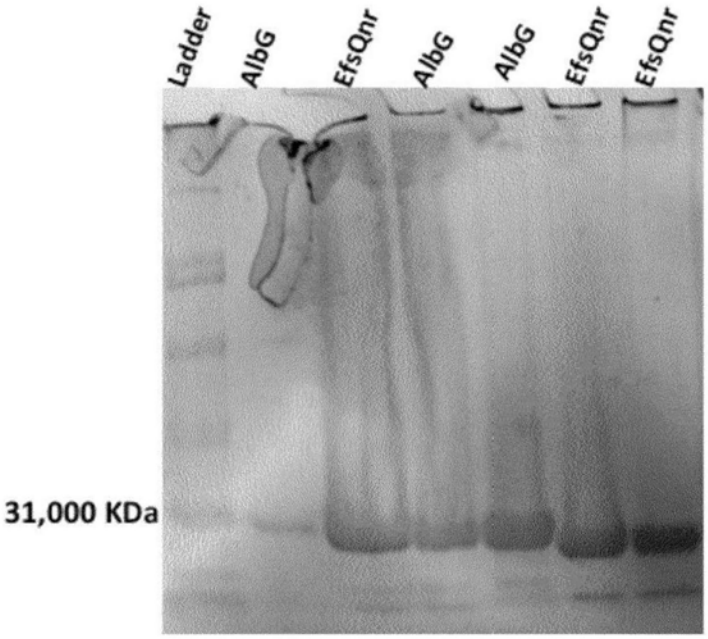


图3

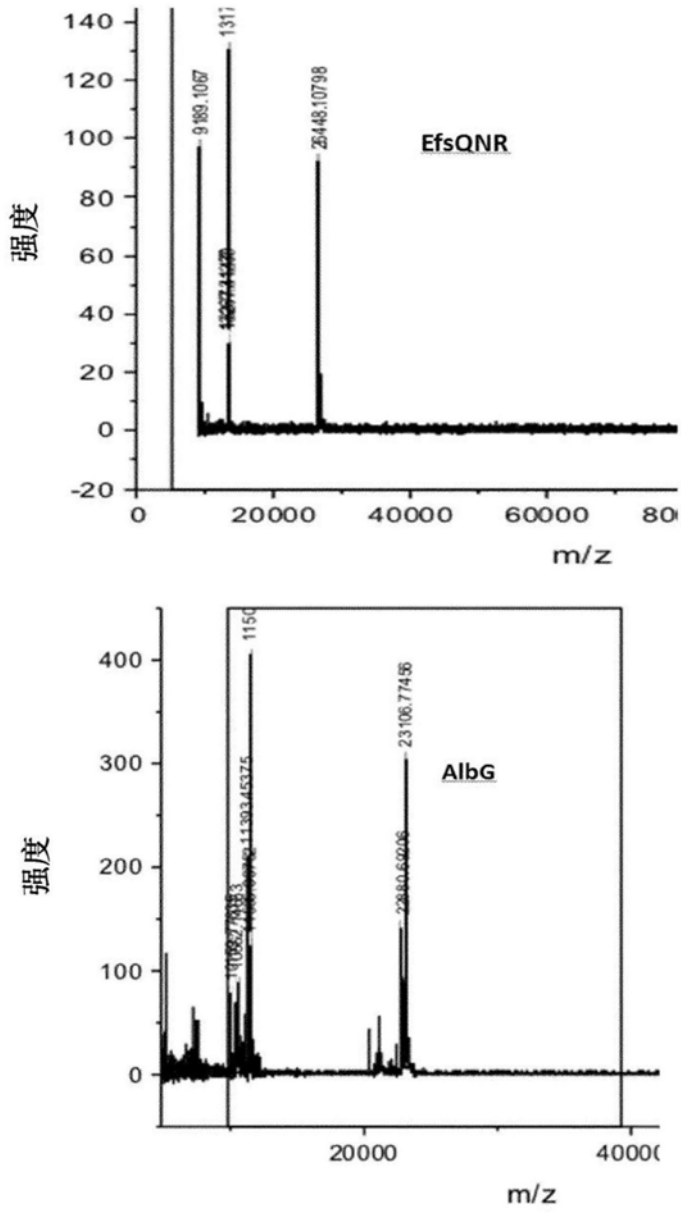


图4

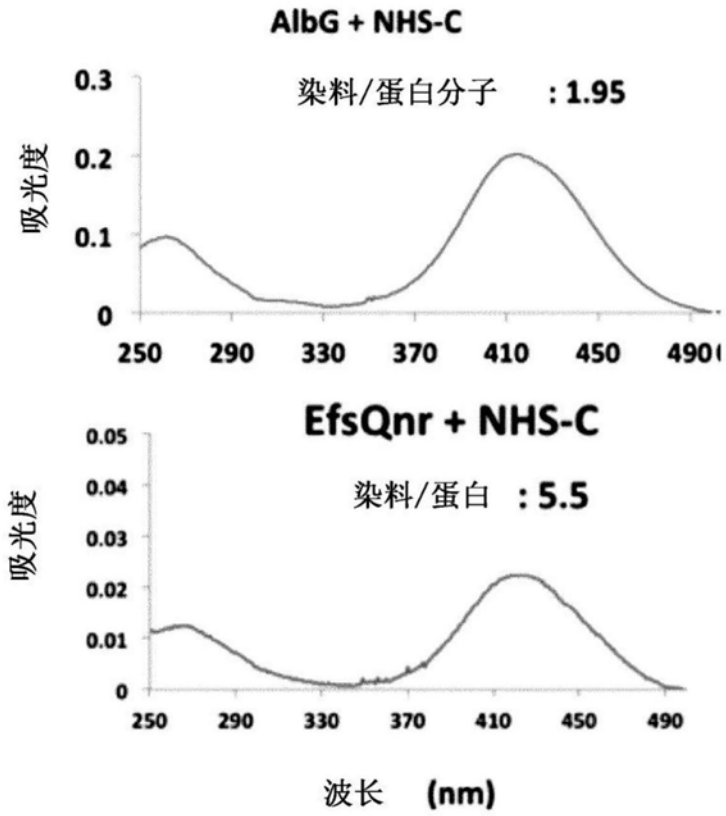

AlbG + NHS-C
染料/蛋白分子 : 1.95
吸光度
0.3
0.2
0.1
0
250 290 330 370 410 450 490
EfsQnr + NHS-C
染料/蛋白 : 5.5
吸光度
0.05
0.04
0.03
0.02
0.01
0
250 290 330 370 410 450 490
波长 (nm)

图5

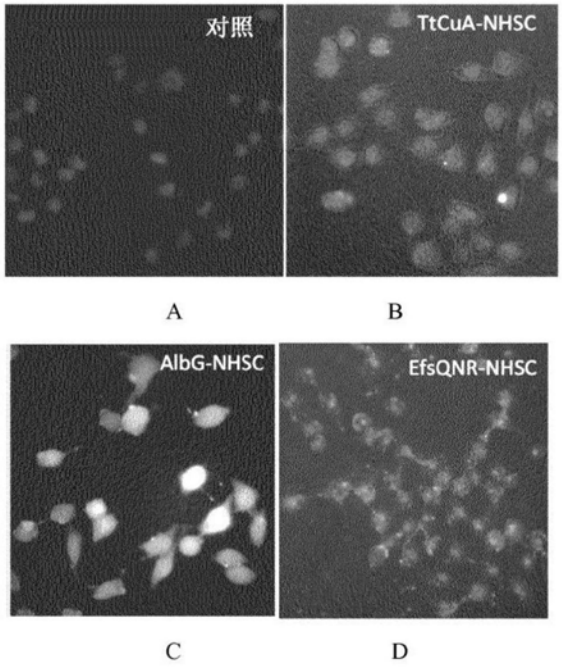

对照
TtCuA-NHSC
A
B
AlbG-NHSC
EfsQNR-NHSC
C
D

图6

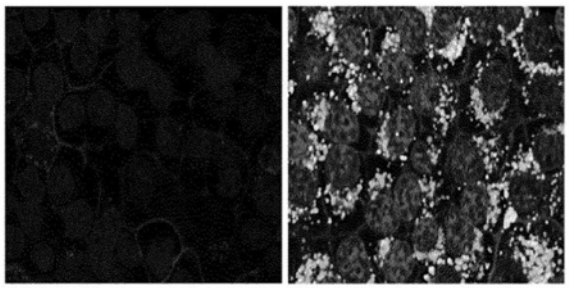


图7

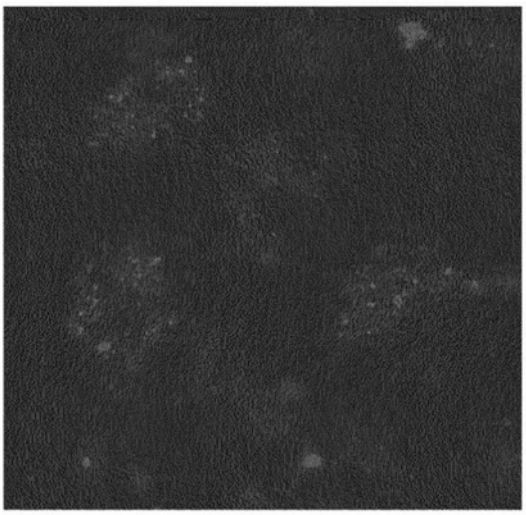


图8

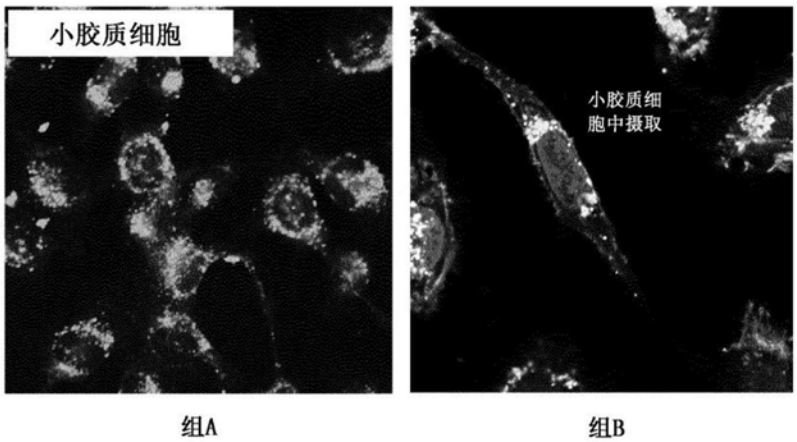


图9

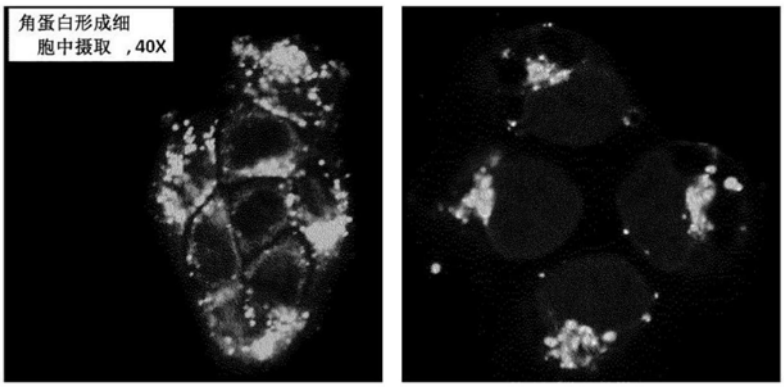


图10

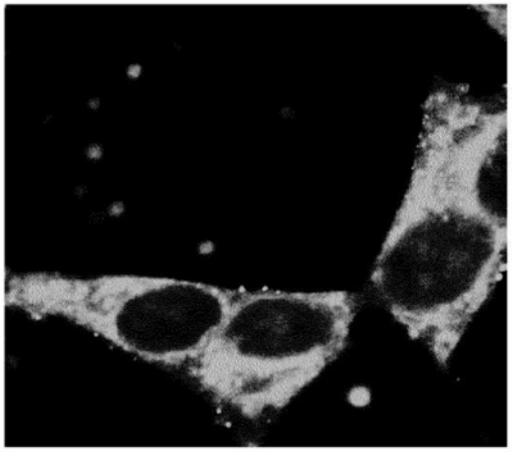


图11

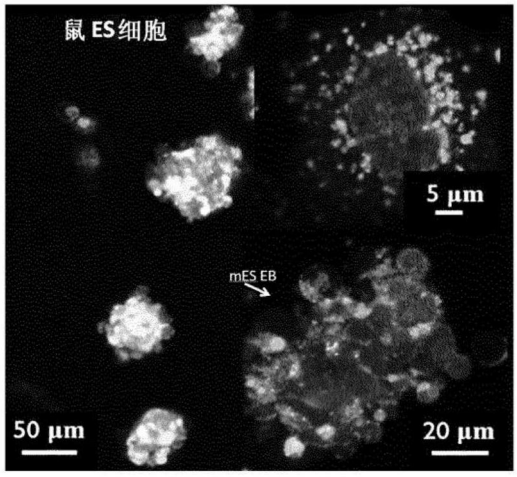


图12

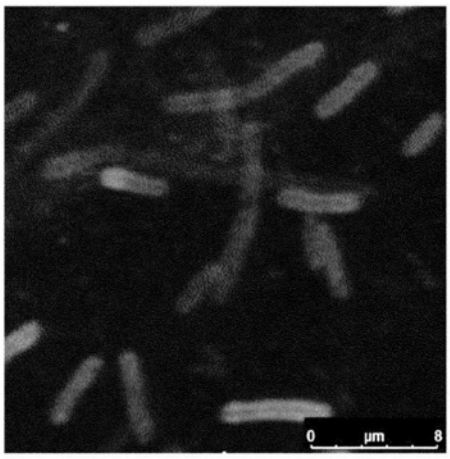


图13

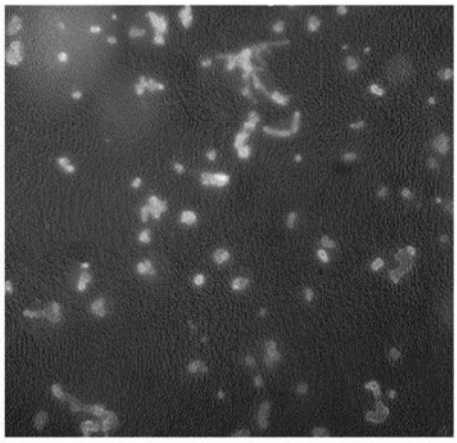


图14

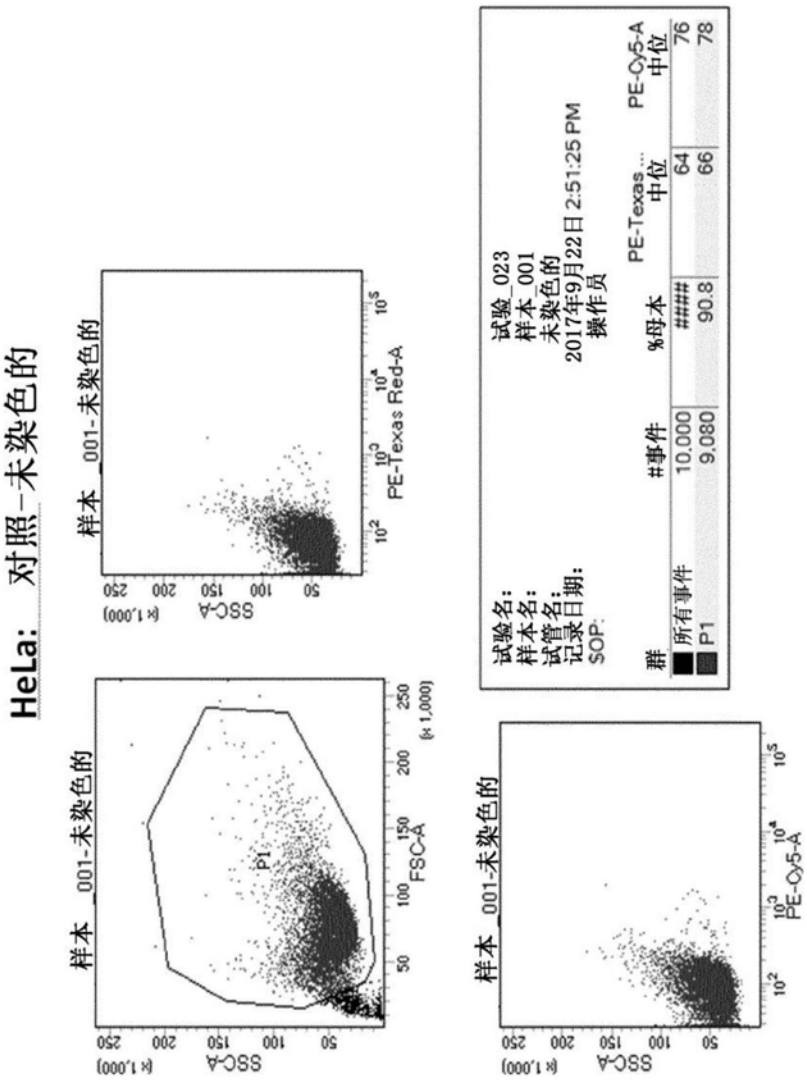

HeLa: 对照-未染色的
样本_001-未染色的
SSC-A
PE-Texas Red-A
FSC-A
P1
PE-Cy5-A
试验名: 试验_023
样本名: 样本_001
试管名: 未染色的
记录日期: 2017年9月22日 2:51:25 PM
SOP: 操作员
群
#事件
%母本
PE-Texas... 中位
PE-Cy5-A 中位
所有事件 10,000 #### 64 76
P1 9,080 90.8 66 78

图15:组A

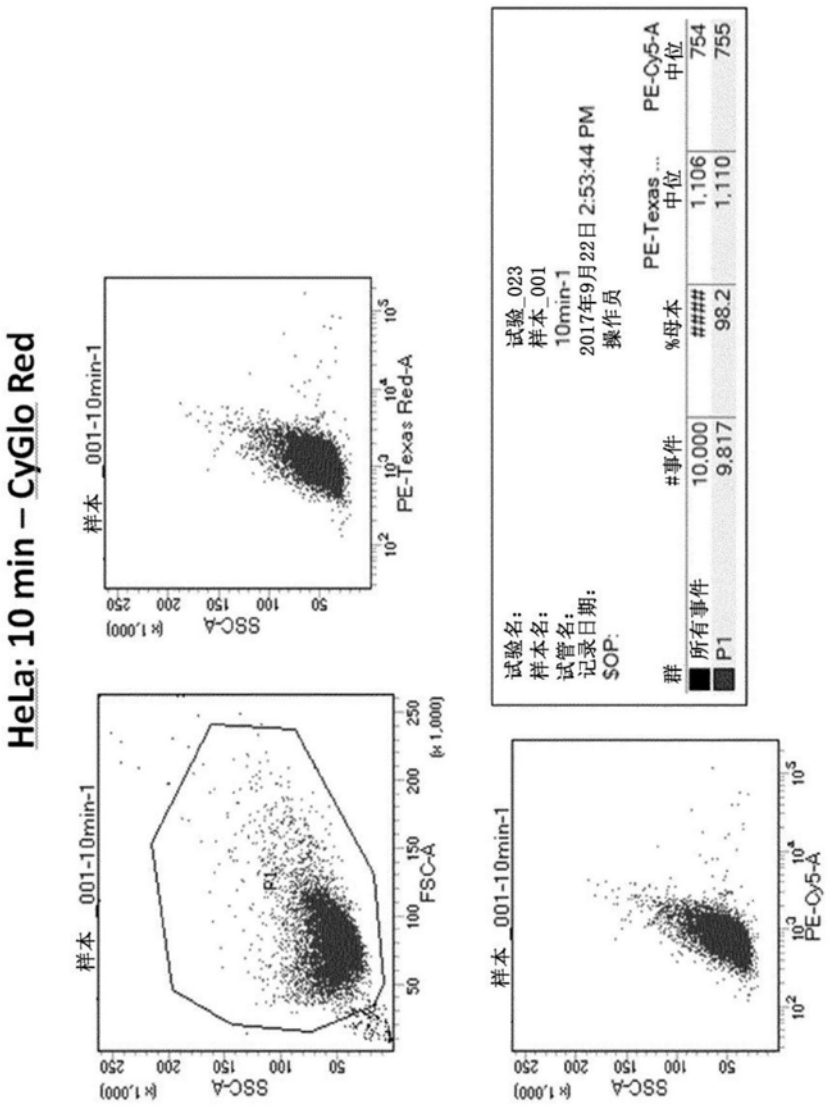

HeLa: 10 min – CyGlo Red
样本_001-10min-1
SSC-A
PE-Texas Red-A
FSC-A
P1
PE-Cy5-A
试验名: 试验_023
样本名: 样本_001
试管名: 10min-1
记录日期: 2017年9月22日 2:53:44 PM
SOP: 操作员
群 #事件 %母本 PE-Texas ... 中位 PE-Cy5-A 中位
所有事件 10,000 #### 1,106 754
P1 9,817 98.2 1,110 755

图15:组B

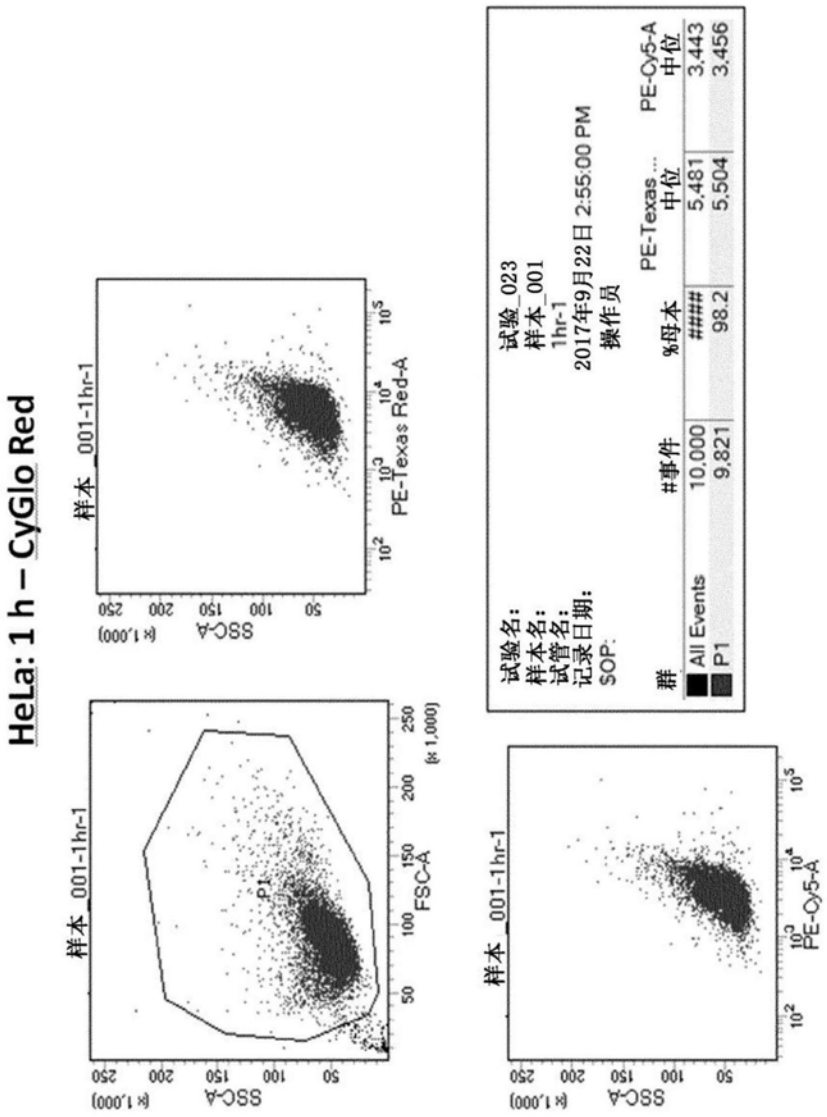

HeLa: 1 h – CyGlo Red
样本_001-1hr-1
SSC-A
(x 1,000)
PE-Texas Red-A
FSC-A
P1
PE-Cy5-A
试验名:
样本名:
试管名:
记录日期:
SOP:
试验_023
样本_001
1hr-1
2017年9月22日 2:55:00 PM
操作员
群
#事件
%母本
PE-Texas ... 中位
PE-Cy5-A 中位
All Events 10,000 #### 5,481 3,443
P1 9,821 98.2 5,504 3,456

图15:组C

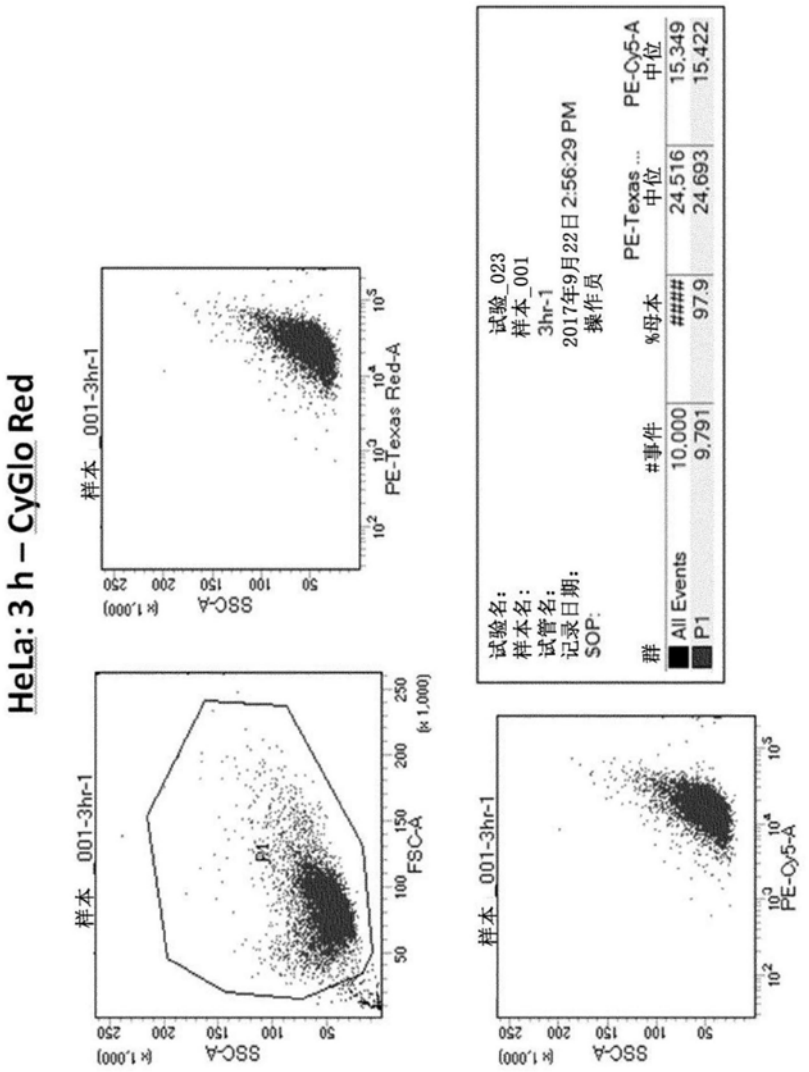

HeLa: 3 h – CyGlo Red
样本_001-3hr-1
SSC-A
PE-Texas Red-A
FSC-A
P1
PE-Cy5-A
试验名:
样本名:
试管名:
记录日期:
SOP:
试验_023
样本_001
3hr-1
2017年9月22日 2:56:29 PM
操作员
群
#事件
%母本
PE-Texas ... 中位
PE-Cy5-A 中位
All Events 10,000 ##### 24.516 15.349
P1 9,791 97.9 24.693 15.422

图15:组D

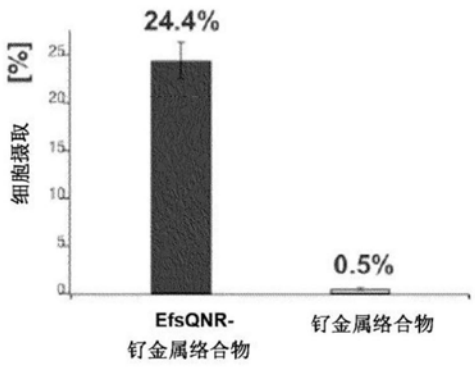

细胞摄取 [%]
24.4%
0.5%
EfsQNR-钌金属络合物
钌金属络合物

图16

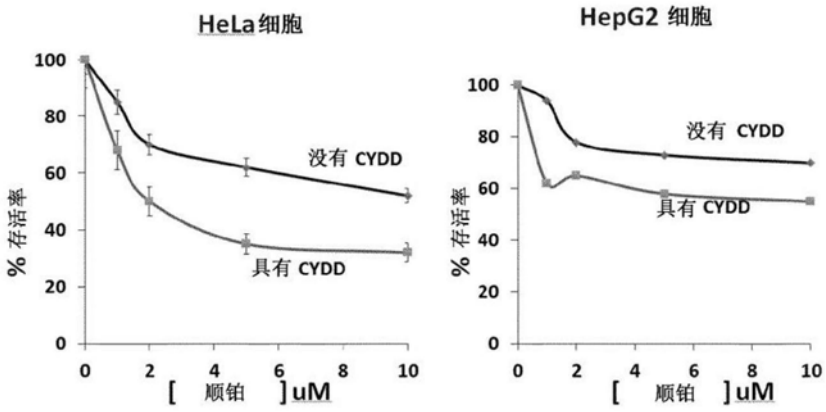

HeLa细胞
HepG2 细胞
% 存活率
没有 CYDD
具有 CYDD
[顺铂]uM
0
20
40
60
80
100
0
2
4
6
8
10

图17

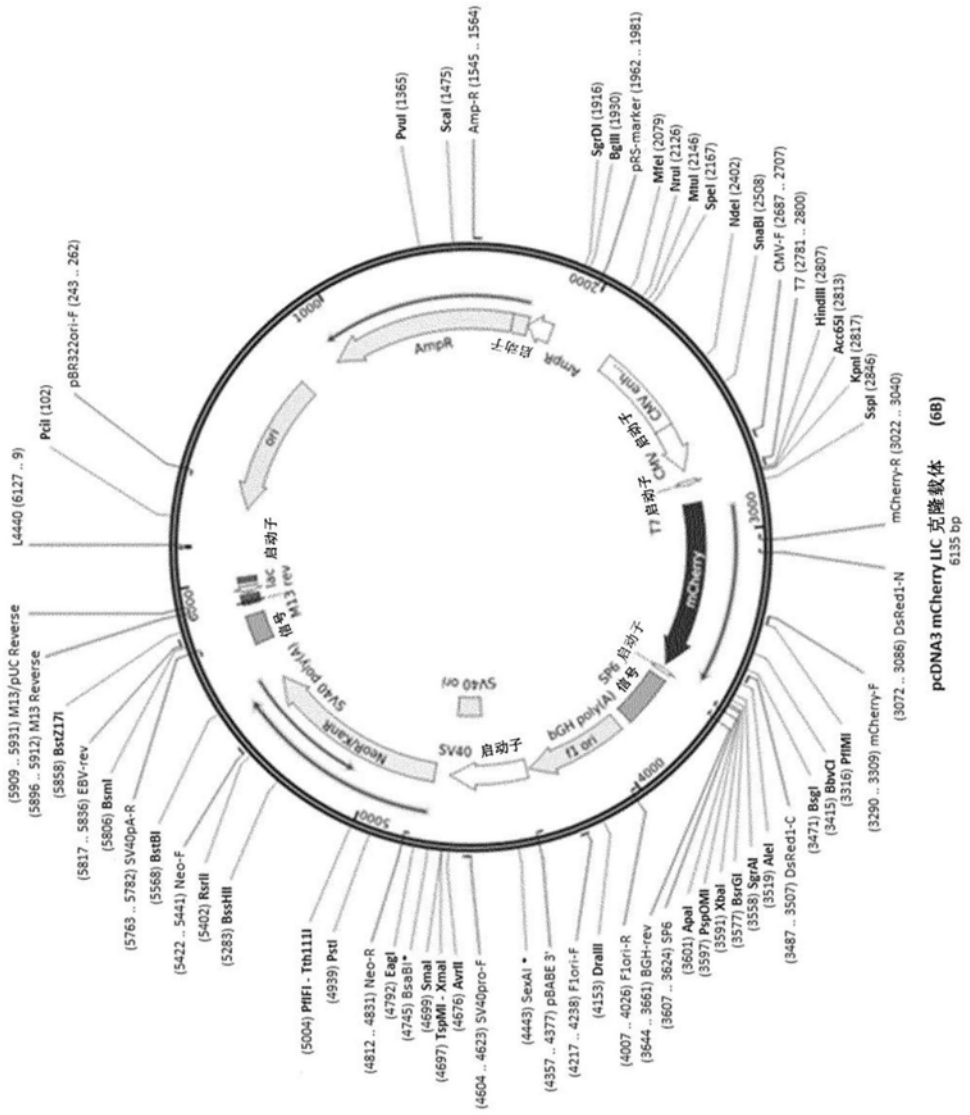

pcDNA3 mCherry LIC 克隆载体 (6B)
6135 bp
CMV 启动子
T7 启动子
mCherry
SP6 启动子
BGH poly(A) 信号
f1 ori
SV40 启动子
SV40 poly(A) 信号
Neo/Kan
SV40 ori
lac 启动子
M13 rev 信号
ori
AmpR 启动子
AmpR
CMV enh.
2000
3000
4000
5000
6000
1000

图18

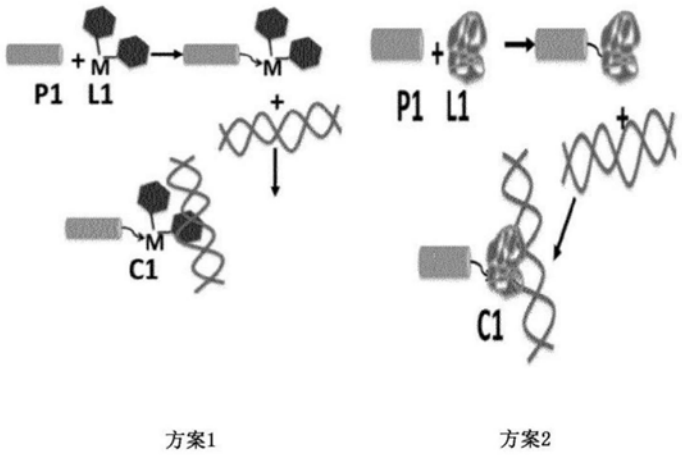


图19

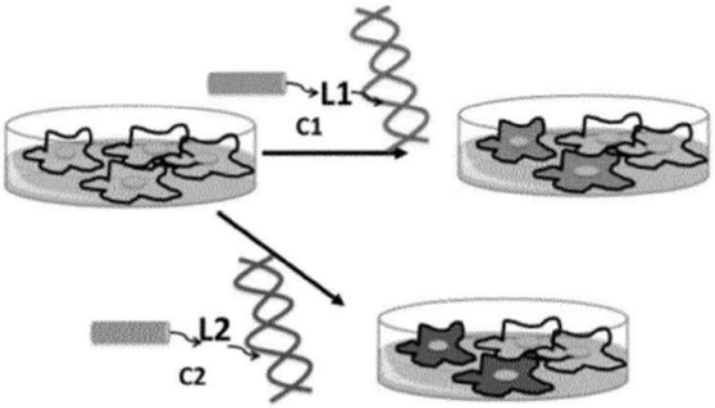


图20

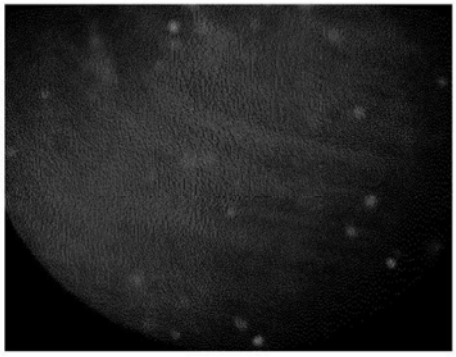


图21

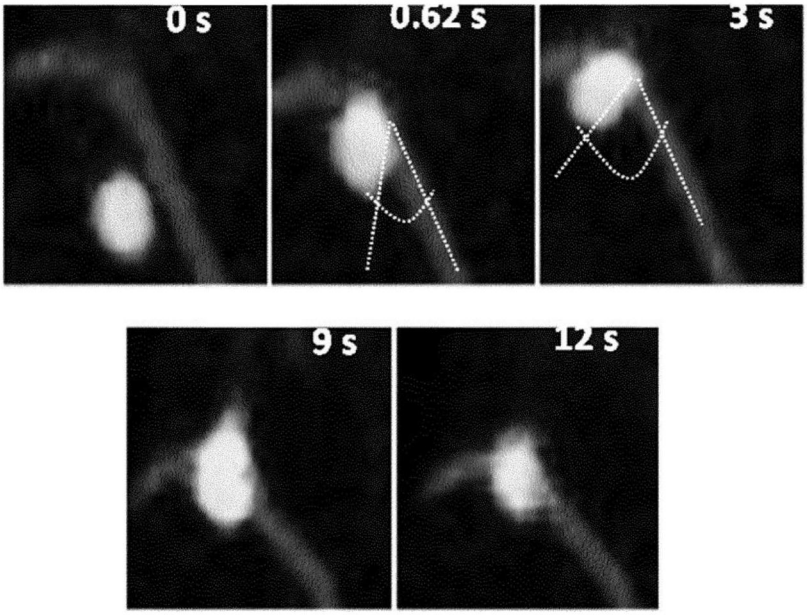


图22