(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 113454216 A
(43)申请公布日 2021.09.28

(21)申请号 201980078727.2

(22)申请日 2019.11.27

(30)优先权数据
62/772,907 2018.11.29 US
62/778,361 2018.12.12 US
62/780,442 2018.12.17 US

(85)PCT国际申请进入国家阶段日
2021.05.28

(86)PCT国际申请的申请数据
PCT/US2019/063798 2019.11.27

(87)PCT国际申请的公布数据
WO2020/113135 EN 2020.06.04

(71)申请人 旗舰先锋创新V股份有限公司
地址 美国马萨诸塞州

(72)发明人 D.A.贝里 C.罗兹 J.D.法雷利

(74)专利代理机构 北京市柳沈律师事务所 11105
代理人 张文辉

(51)Int.Cl.
C12N 9/78(2006.01)

权利要求书5页 说明书44页 附图5页

(54)发明名称

调节RNA的方法

(57)摘要

本披露总体上涉及用于调节RNA的方法和组合物,例如使用包含Pumilio同源结构域的多肽。

CN 113454216 A

1.一种多肽，该多肽包含：(a)RNA结合结构域，该RNA结合结构域包含多个(例如，2-50、10-30、或16-21个)RNA碱基结合基序，每个基序与RNA碱基结合，并且这些基序在该RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合，该RNA结合结构域连接至(b)异源RNA编辑结构域。

2.一种多肽，该多肽包含：(a)RNA结合结构域，该RNA结合结构域包含多个(例如，2-50、10-30、或16-21个)RNA碱基结合基序，每个基序与RNA碱基结合，并且这些基序在该RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合，该RNA结合结构域连接至(b)异源RNA编辑结构域，其中该多肽不包含核酸酶或其功能片段。

3.一种多肽，该多肽包含：(a)RNA结合结构域，该RNA结合结构域包含多个(例如，2-50、10-30、或16-21个)RNA碱基结合基序，每个基序与RNA碱基结合，并且这些基序在该RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合，该RNA结合结构域连接至(b)异源RNA编辑结构域，该异源RNA编辑结构域包含脱氨酶的催化结构域或其功能片段或变体。

4.一种多肽，该多肽包含：(a)RNA结合结构域，该RNA结合结构域包含多个(例如，2-50、10-30、或16-21个)RNA碱基结合基序，每个基序与RNA碱基结合，并且这些基序在该RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合，该RNA结合结构域连接至(b)异源RNA效应子，该异源RNA效应子包含剪接因子。

5.如任一前述权利要求所述的多肽，其中该多个RNA碱基结合基序包含至少3个(例如，至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、14-24个、15-23个、16-22个、16-21个、2-20个、2-15个、2-10个、2-8个、3-20个、3-15个、3-10个、3-8个、4-8个、多达25个、多达30个)PUM RNA结合基序。

6.如任一前述权利要求所述的多肽，其中该RNA结合结构域结合2-50个核苷酸(例如，14-30、15-26、16-21、2-40、2-30、2-25、2-20、5-50、5-40、5-30、5-25、5-20、5-15、2-18、2-15、2-12、2-10、2-9、2-8、3-20、3-15、3-10、3-9、3-8、4-12、4-10、4-9、4-8、5-10、5-9、5-8个核苷酸)的RNA序列。

7.如任一前述权利要求所述的多肽，其中该RNA结合结构域是90-500个氨基酸残基，例如，90-450个氨基酸残基、90-400个氨基酸残基、90-350个氨基酸残基、90-300个氨基酸残基、120-400个氨基酸残基。

8.如任一前述权利要求所述的多肽，其中该RNA结合结构域与野生型PUM-HD，例如野生型人PUM1-HD的相应的氨基酸序列具有至少80％同一性(例如，至少85％同一性、至少87％同一性、至少90％同一性、至少92％同一性、至少95％同一性、至少97％同一性、至少98％同一性、或99％同一性)并小于100％同一性。

9.如任一前述权利要求所述的多肽，其中该RNA结合结构域结合包含疾病相关突变的RNA序列。

10.如任一前述权利要求所述的多肽，其中该RNA结合结构域结合包含疾病相关突变的RNA序列并且该RNA编辑结构域编辑(例如，纠正)该疾病相关突变。

11.如任一前述权利要求所述的多肽，其中该RNA编辑结构域包含多肽，该多肽包含RNA脱氨酶(例如，腺苷脱氨酶或胞苷脱氨酶)的催化结构域或其功能片段或变体。

12.如任一前述权利要求所述的多肽，其中该RNA编辑结构域包含以下项的催化结构

域:作用于RNA的腺苷脱氨酶(ADAR)(例如,人ADAR 1、人ADAR2、人ADAR3、或人ADAR4);作用于tRNA的腺苷脱氨酶(ADAT);作用于RNA的胞嘧啶脱氨酶(CDAR);或其功能片段或变体。

13.如权利要求11或12所述的多肽,其中该脱氨酶的该催化结构域与表B中所示的序列至少80%相同(例如,至少85%、87%、90%、92%、95%、98%、99%、100%相同)。

14.如任一前述权利要求所述的多肽,其中该RNA编辑结构域修饰该靶RNA序列或包含该靶序列的RNA的至少1、2、3、4、5、6、7、8、9、或10(例如,1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3、1-2、2-10、2-9、2-8、2-7、2-6、2-5、2-4、2-3、3-10、3-9、3-8、3-7、3-6、3-5、3-4、4-10、4-9、4-8、4-7、4-6、4-5、5-10、5-9、5-8、5-7、5-6、6-10、6-9、6-8、6-7、7-10、7-9、7-8、8-10、8-9、或9-10)个核苷酸。

15.如任一前述权利要求所述的多肽,其中该RNA编辑结构域修饰该靶RNA序列或包含该靶序列的RNA的单个核苷酸。

16.如任一前述权利要求所述的多肽,其中该RNA编辑结构域将碱基改变为另一碱基,例如,将胞嘧啶改变为尿嘧啶;将腺苷改变为肌苷;或将鸟苷改变为腺苷。

17.如任一前述权利要求所述的多肽,其中该RNA编辑结构域修饰该靶RNA序列的氨基酸编码序列。

18.如权利要求17所述的多肽,其中对该靶RNA序列的该氨基酸编码序列的修饰改变了由该靶RNA序列编码的产物多肽的氨基酸序列。

19.如任一前述权利要求所述的多肽,其中该RNA编辑结构域修饰该靶RNA序列的至少1、2、3、4、5、6、7、8、9、或10(例如,1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3、1-2、2-10、2-9、2-8、2-7、2-6、2-5、2-4、2-3、3-10、3-9、3-8、3-7、3-6、3-5、3-4、4-10、4-9、4-8、4-7、4-6、4-5、5-10、5-9、5-8、5-7、5-6、6-10、6-9、6-8、6-7、7-10、7-9、7-8、8-10、8-9、或9-10)个核苷酸,并且任选地不超过该靶RNA序列的1、2、3、4、5、6、7、8、9、或10个核苷酸。

20.如任一前述权利要求所述的多肽,其中该RNA结合结构域结合RNA的二级结构。

21.如任一前述权利要求所述的多肽,其中该RNA结合结构域结合前体mRNA,例如,前体mRNA的内含子-外显子连接。

22.如任一前述权利要求所述的多肽,其中该多肽抑制(例如,抑制以下的形成)、去稳定、和/或消除该靶RNA序列或包含该靶RNA序列的RNA的二级结构。

23.如任一前述权利要求所述的多肽,其中该多肽改变该靶RNA序列或包含该靶RNA序列的RNA的剪接。

24.如权利要求23所述的多肽,其中该多肽抑制,例如消除该靶RNA序列或包含该靶RNA序列的RNA在剪接位点(例如,靶剪接位点)处的剪接,并且任选地不抑制该靶RNA序列或包含该靶RNA序列的RNA在一个或多个其他剪接位点(例如,一个或多个非靶剪接位点)处的剪接。

25.如任一前述权利要求所述的多肽,其中该多肽减少基因的表达,例如,编码该靶RNA序列的基因的表达。

26.如任一前述权利要求所述的多肽,其中该多肽降低由该靶RNA序列编码的产物多肽的水平。

27.如任一前述权利要求所述的多肽,其中该多肽消除在该靶RNA序列或包含该靶RNA序列的RNA中的终止密码子,例如,提前终止密码子。

28.如任一前述权利要求所述的多肽，其中该多肽在该靶RNA序列或包含该靶RNA序列的RNA中创建终止密码子，例如，提前终止密码子。

29.如任一前述权利要求所述的多肽，其中该RNA结合结构域的该多个RNA碱基结合基序中的至少2(例如，3、4、5、6、7、8、9或更多)个通过接头，例如氨基酸接头连接。

30.如任一前述权利要求所述的多肽，其中该RNA结合结构域和该RNA编辑结构域通过接头，例如氨基酸接头连接。

31.如任一前述权利要求所述的多肽，其中该多肽进一步包含剪接因子。

32.一种组合物，该组合物包含如任一前述权利要求所述的多肽、和反义寡核苷酸，该反义寡核苷酸包含与靶RNA序列互补的序列。

33.一种核酸，该核酸编码如任一前述权利要求所述的多肽。

34.如权利要求33所述的核酸，其中该核酸是RNA，例如mRNA。

35.一种组合物，该组合物包含如权利要求33或34所述的核酸、和反义寡核苷酸，该反义寡核苷酸包含与靶RNA序列互补的序列。

36.一种组合物，该组合物包含如权利要求33或34所述的核酸、和编码反义寡核苷酸的核酸，该反义寡核苷酸包含与靶RNA序列互补的序列。

37.一种表达载体(例如，质粒载体、病毒载体)，该表达载体包含如权利要求33或34所述的核酸。

38.一种宿主细胞(例如，细菌宿主细胞、哺乳动物宿主细胞)，该宿主细胞包含外源的如任一前述权利要求所述的多肽、如权利要求33或34所述的核酸、如权利要求35或36所述的组合物、或如权利要求37所述的载体。

39.一种GMP级药物组合物，该GMP级药物组合物包含如前述任一权利要求所述的多肽、核酸、载体、组合物、或宿主细胞以及药学上可接受的赋形剂。

40.如任一前述权利要求所述的多肽、核酸、载体、组合物、药物组合物、或宿主细胞，其包封或配制在药物载剂(例如，囊泡、脂质体、LNP)中。

41.一种修饰靶RNA(例如，改变靶RNA的序列)的方法，该方法包括使细胞、组织或受试者与如任一前述权利要求所述的多肽、核酸、载体、组合物、宿主细胞、或GMP级药物组合物以如下量和持续时间接触，所述量和持续时间足以使该多肽的RNA结合结构域与该细胞、组织或受试者中的靶RNA结合，并且使该多肽的RNA编辑结构域编辑该靶RNA。

42.如权利要求41所述的方法，其中该靶RNA是具有二级和/或三级结构的前体mRNA或mRNA。

43.如权利要求41或42所述的方法，其中该靶RNA是前体mRNA，例如，前体mRNA的内含子-外显子连接。

44.如任一前述权利要求所述的方法，其中该多肽改变该靶RNA的核苷酸序列。

45.如权利要求44所述的方法，其中改变包括修饰该靶RNA序列或包含该靶序列的RNA的至少1、2、3、4、5、6、7、8、9、或10(例如，1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3、1-2、2-10、2-9、2-8、2-7、2-6、2-5、2-4、2-3、3-10、3-9、3-8、3-7、3-6、3-5、3-4、4-10、4-9、4-8、4-7、4-6、4-5、5-10、5-9、5-8、5-7、5-6、6-10、6-9、6-8、6-7、7-10、7-9、7-8、8-10、8-9、或9-10)个核苷酸。

46.如权利要求44所述的方法，其中改变包含修饰该靶RNA序列或包含该靶序列的RNA

的单个核苷酸。

47.如权利要求44-46中任一项所述的方法，其中改变包括将碱基改变为另一碱基，例如，将胞嘧啶改变为尿嘧啶；将腺苷改变为肌苷；或将鸟苷改变为腺苷。

48.如权利要求44-47中任一项所述的方法，其中改变包括修饰该靶RNA序列的氨基酸编码序列。

49.如权利要求48所述的方法，其中对该靶RNA序列的该氨基酸编码序列的修饰改变了由该靶RNA序列编码的产物多肽的氨基酸序列。

50.如任一前述权利要求所述的方法，其中该靶RNA包含细胞、组织或受试者中的前体mRNA或mRNA，并且该多肽改变(例如，增加或减少)该前体mRNA或mRNA的二级或三级结构。

51.如任一前述权利要求所述的方法，其中该靶RNA包含细胞、组织或受试者中的前体mRNA或mRNA，并且该多肽改变该前体mRNA或mRNA的剪接。

52.如权利要求51所述的多肽，其中该多肽抑制，例如消除该前体mRNA或mRNA在剪接位点(例如，靶剪接位点)处的剪接，并且任选地不抑制该前体mRNA或mRNA在一个或多个其他剪接位点(例如，一个或多个非靶剪接位点)处的剪接。

53.如任一前述权利要求所述的药物组合物、多肽、核酸、载体、组合物、宿主细胞、或方法，其中该靶RNA包含EB病毒(EBV)mRNA，例如，EBV核抗原1(EBNA1)mRNA。

54.如任一前述权利要求所述的药物组合物、多肽、核酸、载体、组合物、宿主细胞、或方法，其中该靶RNA包含脊髓性肌神经元2(SMN2)mRNA。

55.如任一前述权利要求所述的药物组合物、多肽、核酸、载体、组合物、宿主细胞、或方法，其中该靶RNA包含GluA2 mRNA。

56.如任一前述权利要求所述的药物组合物、多肽、核酸、载体、组合物、宿主细胞、或方法，其中该多肽包含选自SEQ ID NO:13-21的氨基酸序列或与其具有至少80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、或99％同一性的氨基酸序列或相对于其具有不超过1、2、3、4、5、6、7、8、9、或10个碱基改变(例如，取代、缺失、或插入)的氨基酸序列。

57.如任一前述权利要求所述的药物组合物、多肽、核酸、载体、组合物、宿主细胞、或方法，其中该RNA结合结构域与靶RNA序列结合，该靶RNA序列包含选自SEQ ID NO:22-25或相对于其具有不超过1、2、3、4、5、6、7、8、9、或10个碱基改变的RNA序列。

58.一种治疗受试者，例如人受试者的疾病或障碍的方法，该方法包括向该受试者施用有效量的如任一前述权利要求所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞，从而治疗该疾病或障碍，

其中该疾病或障碍选自梅尔戈林综合征、塞克尔综合征4、Joubert综合征5、莱伯氏先天性黑蒙10；腓骨肌萎缩疾病，2型；腓骨肌萎缩疾病，2型；乌谢尔综合征，2C型；脊髓小脑性共济失调28；脊髓小脑性共济失调28；脊髓小脑性共济失调28；长QT综合征2；格伦-拉松综合征；遗传性果糖尿症；遗传性果糖尿症；神经母细胞瘤；神经母细胞瘤；卡尔曼综合征1；卡尔曼综合征1；卡尔曼综合征1；异染性脑白质营养不良、雷特综合征、10型肌萎缩性侧索硬化症、李-佛美尼综合征、囊性纤维化病、赫尔综合征、α-1-抗胰蛋白酶(A1AT)缺乏、帕金森氏症、阿耳茨海默病、白化病、肌萎缩性侧索硬化症、哮喘、b-地中海贫血、Cadasil综合征、腓骨肌萎缩疾病、慢性阻塞性肺病(COPD)、远端脊髓性肌萎缩(DSMA)、杜氏/贝克肌营养不良、营养不良性大疱性表皮松解、大疱性表皮松懈症、法布瑞氏症、因子V Leiden相关障碍、

家族性腺瘤、息肉病、半乳糖血症、戈谢病、葡萄糖-6-磷酸脱氢酶、血友病、遗传性血色素沉着症、亨特氏综合征、亨廷顿病、炎性肠病(IBD)、遗传性多凝集反应综合征、莱伯氏先天性黑蒙、莱施-奈恩综合征、林奇综合征、马凡综合征、黏多糖贮积症、肌肉萎缩症、I型和II型强直性肌营养不良、神经纤维瘤病、A型、B型和C型尼曼-皮克病、NY-eso1相关癌症、黑斑息肉综合征、苯酮尿症、庞培氏病、原发性睫状疾病、凝血酶原突变相关障碍例如凝血酶原G20210A突变、肺动脉高压、色素性视网膜炎、山德霍夫氏病、严重联合免疫缺陷综合征(SCID)、镰状细胞性贫血、脊髓性肌萎缩、斯特格氏病、泰-萨克斯病、乌谢尔综合征、X连锁免疫缺陷、斯特奇-韦伯综合征、以及癌症。

59.一种治疗感染或疑似感染EB病毒(EBV)的受试者(例如,人受试者)的方法,该方法包括向该受试者施用有效量的如任一前述权利要求所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞,从而治疗该感染或疑似感染EB病毒(EBV)的受试者。

60.如权利要求59所述的方法,其中该受试者患有单核细胞增多症或癌症(例如,伯基特淋巴瘤、霍奇金氏病和鼻咽癌)。

61.一种治疗患有脊髓性肌肉萎缩症(SMA)的受试者(例如,人受试者)的方法,该方法包括向该受试者施用有效量的如任一前述权利要求所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞,从而治疗该患有SMA的受试者。

62.一种治疗患有肌萎缩性侧索硬化(ALS)的受试者(例如,人受试者)的方法,该方法包括向该受试者施用有效量的如任一前述权利要求所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞,从而治疗该患有ALS的受试者。

调节RNA的方法

[0001] 相关申请

[0002] 本申请要求2018年11月29日提交的美国序列号62/772,907,2018年12月12日提交的美国序列号62/778,361和2018年12月17日提交的美国序列号62/780,442的优先权,其内容各自通过引用以其全文并入文中。

发明内容

[0003] 本文描述了用于改变RNA结构和功能以调节生物过程的组合物和方法。

[0004] 一级核苷酸序列决定了RNA的二级和三级结构。核苷酸的碱基配对形成了RNA配体(如蛋白质)结合所必需的茎、环和组合。这样,RNA的一级序列的编辑以及由此的二级和/或三级结构的编辑可以改变其配体结合特性,并提供调节下游过程而不改变配体(例如,RNA结合多肽)的功能的方式。本文描述了调节RNA一级、二级和三级结构和功能和/或剪接以影响由RNA-配体相互作用和/或RNA编码产物的表达实现的过程的组合物和相关方法。

[0005] 因此,在一方面,本披露涉及多肽,该多肽包含:(a)RNA结合结构域,该RNA结合结构域包含多个(例如,2-50、10-30、或16-21个)RNA碱基结合基序,每个基序与RNA碱基结合,并且这些基序在该RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合,该RNA结合结构域连接至(b)异源RNA编辑结构域。

[0006] 在另一个方面,本披露涉及多肽,该多肽包含:(a)RNA结合结构域,该RNA结合结构域包含多个(例如,2-50、10-30、或16-21个)RNA碱基结合基序,每个基序与RNA碱基结合,并且这些基序在该RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合,该RNA结合结构域连接至(b)异源RNA编辑结构域,其中该多肽不包含核酸酶或其功能片段。

[0007] 在另一个方面,本披露涉及多肽,该多肽包含:(a)RNA结合结构域,该RNA结合结构域包含多个(例如,2-50、10-30、或16-21个)RNA碱基结合基序,每个基序与RNA碱基结合,并且这些基序在该RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合,该RNA结合结构域连接至(b)异源RNA编辑结构域,该异源RNA编辑结构域包含脱氨酶的催化结构域或其功能片段或变体。

[0008] 在另一个方面,本披露涉及多肽,该多肽包含:(a)RNA结合结构域,该RNA结合结构域包含多个(例如,2-50、10-30、或16-21个)RNA碱基结合基序,每个基序与RNA碱基结合,并且这些基序在该RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合,该RNA结合结构域连接至(b)异源RNA效应子,该异源RNA效应子包含剪接因子。

[0009] 在一些实施例中,多个RNA碱基结合基序包含至少3个(例如,至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、14-24个、15-23个、16-22个、16-21个、2-20个、2-15个、2-10个、2-8个、3-20个、3-15个、3-10个、3-8个、4-8个、多达25个、多达30个)PUM RNA结合基序。

[0010] 在一些实施例中,该RNA结合结构域结合2-50个核苷酸(例如,14-30、15-26、16-21、2-40、2-30、2-25、2-20、5-50、5-40、5-30、5-25、5-20、5-15、2-18、2-15、2-12、2-10、2-9、2-8、3-20、3-15、3-10、3-9、3-8、4-12、4-10、4-9、4-8、5-10、5-9、5-8个核苷酸)的RNA序列。

[0011] 在一些实施例中，该RNA结合结构域是90-500个氨基酸残基，例如，90-450个氨基酸残基、90-400个氨基酸残基、90-350个氨基酸残基、90-300个氨基酸残基、120-400个氨基酸残基。

[0012] 在一些实施例中，该RNA结合结构域与野生型PUM-HD(例如，野生型人PUM1-HD)的相应氨基酸序列具有至少80％同一性(例如，至少85％同一性、至少87％同一性、至少90％同一性、至少92％同一性、至少95％同一性、至少97％同一性、至少98％同一性、或99％同一性)并小于100％同一性。

[0013] 在一些实施例中，该RNA结合结构域结合包含疾病相关突变的RNA序列。

[0014] 在一些实施例中，该RNA结合结构域结合包含疾病相关突变的RNA序列并且该RNA编辑结构域编辑(例如，纠正)疾病相关突变。

[0015] 在一些实施例中，该RNA编辑结构域包含多肽，该多肽包含RNA脱氨酶(例如，腺苷脱氨酶或胞苷脱氨酶)的催化结构域或其功能片段或变体。

[0016] 在一些实施例中，该RNA编辑结构域包含以下项的催化结构域：作用于RNA的腺苷脱氨酶(ADAR)(例如，人ADAR 1、人ADAR2、人ADAR3、或人ADAR4)；作用于tRNA的腺苷脱氨酶(ADAT)；作用于RNA的胞嘧啶脱氨酶(CDAR)；或其功能片段或变体。

[0017] 在一些实施例中，脱氨酶的催化结构域与表B中所示的序列至少80％相同(例如，至少85％、87％、90％、92％、95％、98％、99％、100％相同)。

[0018] 在一些实施例中，该RNA编辑结构域修饰靶RNA序列或包含该靶序列的RNA的至少1、2、3、4、5、6、7、8、9、或10(例如，1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3、1-2、2-10、2-9、2-8、2-7、2-6、2-5、2-4、2-3、3-10、3-9、3-8、3-7、3-6、3-5、3-4、4-10、4-9、4-8、4-7、4-6、4-5、5-10、5-9、5-8、5-7、5-6、6-10、6-9、6-8、6-7、7-10、7-9、7-8、8-10、8-9、或9-10)个核苷酸。

[0019] 在一些实施例中，该RNA编辑结构域修饰靶RNA序列或包含该靶序列的RNA的单个核苷酸。

[0020] 在一些实施例中，该RNA编辑结构域将碱基改变为另一碱基，例如，将胞嘧啶改变为尿嘧啶；将腺苷改变为肌苷；或将鸟苷改变为腺苷。

[0021] 在一些实施例中，该RNA编辑结构域修饰靶RNA序列的氨基酸编码序列。

[0022] 在一些实施例中，对靶RNA序列的氨基酸编码序列的修饰改变了由靶RNA序列编码的产物多肽的氨基酸序列。

[0023] 在一些实施例中，该RNA编辑结构域修饰靶RNA序列的至少1、2、3、4、5、6、7、8、9、或10(例如，1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3、1-2、2-10、2-9、2-8、2-7、2-6、2-5、2-4、2-3、3-10、3-9、3-8、3-7、3-6、3-5、3-4、4-10、4-9、4-8、4-7、4-6、4-5、5-10、5-9、5-8、5-7、5-6、6-10、6-9、6-8、6-7、7-10、7-9、7-8、8-10、8-9、或9-10)个核苷酸，并且任选地不超过靶RNA序列的1、2、3、4、5、6、7、8、9、或10个核苷酸。

[0024] 在一些实施例中，该RNA结合结构域结合RNA的二级结构。

[0025] 在一些实施例中，该RNA结合结构域结合前体mRNA，例如，前体mRNA的内含子-外显子连接。

[0026] 在一些实施例中，该多肽抑制(例如，抑制以下的形成)、去稳定、和/或消除靶RNA序列或包含该靶RNA序列的RNA的二级结构。

[0027] 在一些实施例中，该多肽改变靶RNA序列的剪接或包含该靶RNA序列的RNA的剪接。

[0028] 在一些实施例中，该多肽抑制(例如消除)靶RNA序列或包含该靶RNA序列的RNA在剪接位点(例如，靶剪接位点)处的剪接，并且任选地不抑制靶RNA序列或包含该靶RNA序列的RNA在一个或多个其他剪接位点(例如，一个或多个非靶剪接位点)处的剪接。

[0029] 在一些实施例中，该多肽减少基因的表达，例如，编码靶RNA序列的基因。

[0030] 在一些实施例中，该多肽降低由靶RNA序列编码的产物多肽的水平。

[0031] 在一些实施例中，该多肽消除在靶RNA序列或包含该靶RNA序列的RNA中的终止密码子，例如，提前终止密码子。

[0032] 在一些实施例中，该多肽在靶RNA序列或包含该靶RNA序列的RNA中创建终止密码子，例如，提前终止密码子。

[0033] 在一些实施例中，RNA结合结构域的多个RNA碱基结合基序中的至少2(例如，3、4、5、6、7、8、9或更多)个通过接头(例如，氨基酸接头)连接。

[0034] 在一些实施例中，该RNA结合结构域和RNA编辑结构域通过接头连接，例如氨基酸接头。

[0035] 在一些实施例中，该多肽进一步包含剪接因子。

[0036] 在另一个方面，本披露涉及组合物，该组合物包含本文所述的多肽和包含与靶RNA序列互补的序列的反义寡核苷酸。

[0037] 在另一个方面，本披露涉及编码本文所述的多肽的核酸。

[0038] 在一些实施例中，该核酸是RNA，例如mRNA。

[0039] 在另一个方面，本披露涉及组合物，该组合物包含本文所述的核酸和包含与靶RNA序列互补的序列的反义寡核苷酸。

[0040] 在另一个方面，本披露涉及组合物，该组合物包含本文所述的核酸和编码以下反义寡核苷酸的核酸，该反义寡核苷酸包含与靶RNA序列互补的序列。

[0041] 在另一个方面，本披露涉及表达载体(例如，质粒载体、病毒载体)，该表达载体包含本文所述的核酸。

[0042] 在另一个方面，本披露涉及宿主细胞(例如，细菌宿主细胞、哺乳动物宿主细胞)，该宿主细胞包含本文所述的多肽、核酸、组合物、或载体。

[0043] 在另一个方面，本披露涉及GMP级药物组合物，该药物组合物包含本文所述的多肽、核酸、载体、组合物、或宿主细胞，以及药学上可接受的赋形剂。

[0044] 在一些实施例中，本文所述的多肽、核酸、载体、组合物、药物组合物、或宿主细胞包封或配制在药物载剂(例如，囊泡、脂质体、LNP)中。

[0045] 在另一个方面，本披露涉及修饰靶RNA(例如，改变其序列)的方法，该方法包括使细胞、组织或受试者与本文所述的多肽、核酸、载体、组合物、宿主细胞、或GMP级药物组合物接触，所使用的量和持续的时间足以使多肽的RNA结合结构域与细胞、组织或受试者中的靶RNA结合，并且使多肽的RNA编辑结构域编辑靶RNA。

[0046] 在一些实施例中，该靶RNA是具有二级和/或三级结构的前体mRNA或mRNA。

[0047] 在一些实施例中，该靶RNA是前体mRNA，例如，前体mRNA的内含子-外显子连接。

[0048] 在一些实施例中，该多肽改变靶RNA的核苷酸序列。

[0049] 在一些实施例中，改变包括修饰靶RNA序列或包含该靶序列的RNA的至少1、2、3、4、5、6、7、8、9、或10(例如，1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3、1-2、2-10、2-9、2-8、2-7、2-

6、2-5、2-4、2-3、3-10、3-9、3-8、3-7、3-6、3-5、3-4、4-10、4-9、4-8、4-7、4-6、4-5、5-10、5-9、5-8、5-7、5-6、6-10、6-9、6-8、6-7、7-10、7-9、7-8、8-10、8-9、或9-10)个核苷酸。

[0050] 在一些实施例中,改变包含修饰靶RNA序列或包含该靶序列的RNA的单个核苷酸。

[0051] 在一些实施例中,改变包括将碱基改变为另一碱基,例如,将胞嘧啶改变为尿嘧啶;将腺苷改变为肌苷;或将鸟苷改变为腺苷。

[0052] 在一些实施例中,改变包括修饰靶RNA序列的氨基酸编码序列。

[0053] 在一些实施例中,对靶RNA序列的氨基酸编码序列的修饰改变了由靶RNA序列编码的产物多肽的氨基酸序列。

[0054] 在一些实施例中,该靶RNA包括细胞、组织或受试者中的前体mRNA或mRNA,并且该多肽改变(例如,增加或减少)前体mRNA或mRNA的二级或三级结构。

[0055] 在一些实施例中,该靶RNA包括细胞、组织或受试者中的前体mRNA或mRNA,并且该多肽改变前体mRNA或mRNA的剪接。

[0056] 在一些实施例中,该多肽抑制(例如消除)前体mRNA或mRNA在剪接位点(例如,靶剪接位点)处的剪接,并且任选地不抑制前体mRNA或mRNA在一个或多个其他剪接位点(例如,一个或多个非靶剪接位点)处的剪接。

[0057] 在一些实施例中,该靶RNA包含EB病毒(EBV)mRNA,例如,EBV核抗原1(EBNA1)mRNA。

[0058] 在一些实施例中,该靶RNA包含脊髓性肌神经元2(SMN2)mRNA。

[0059] 在一些实施例中,该靶RNA包含GluA2 mRNA。

[0060] 在一些实施例中,该多肽包含选自SEQ ID NO:13-21的氨基酸序列或与其具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%同一性的氨基酸序列或相对于其具有不超过1、2、3、4、5、6、7、8、9、或10个碱基改变(例如,取代、缺失、或插入)的氨基酸序列。

[0061] 在一些实施例中,该RNA结合结构域与以下靶RNA序列结合,该靶RNA序列包含选自SEQ ID NO:22-25或相对于其具有不超过1、2、3、4、5、6、7、8、9、或10个碱基改变的RNA序列。

[0062] 在另一个方面,本披露涉及治疗受试者(例如人受试者)的疾病或障碍的方法,该方法包括向受试者施用有效量的本文所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞,从而治疗该疾病或障碍,其中该疾病或障碍选自梅尔戈林综合征、塞克尔综合征4、Joubert综合征5、莱伯氏先天性黑蒙10;腓骨肌萎缩疾病,2型;腓骨肌萎缩疾病,2型;乌谢尔综合征,2C型;脊髓小脑性共济失调28;脊髓小脑性共济失调28;脊髓小脑性共济失调28;长QT综合征2;格伦-拉松综合征(Sjogren-Larsson syndrome);遗传性果糖尿症;遗传性果糖尿症;神经母细胞瘤;神经母细胞瘤;卡尔曼综合征1;卡尔曼综合征1;卡尔曼综合征1;异染性脑白质营养不良、雷特综合征、10型肌萎缩性侧索硬化症、李-佛美尼综合征(Li-Fraumeni syndrome)、囊性纤维化病、赫尔综合征、α-1-抗胰蛋白酶(A1AT)缺乏、帕金森氏症、阿耳茨海默病、白化病、肌萎缩性侧索硬化症、哮喘、b-地中海贫血、Cadasil综合征、腓骨肌萎缩疾病、慢性阻塞性肺病(COPD)、远端脊髓性肌萎缩(DSMA)、杜氏/贝克肌营养不良、营养不良性大疱性表皮松解、大疱性表皮松懈症、法布瑞氏症、因子V Leiden相关障碍、家族性腺瘤、息肉病、半乳糖血症、戈谢病、葡萄糖-6-磷酸脱氢酶、血友病、遗传性血色素沉着症、亨特氏综合征、亨廷顿病、炎性肠病(I BD)、遗传性多凝集反应综合征、莱伯氏先天性黑蒙、莱施-奈恩综合征、林奇综合征、马凡综合征、黏多糖贮积症、肌肉萎缩症、I型和II型强

直性肌营养不良、神经纤维瘤病、A型、B型和C型尼曼-皮克病、NY-eso1相关癌症、黑斑息肉综合征、苯酮尿症、庞培氏病、原发性睫状疾病、凝血酶原突变相关障碍(例如凝血酶原G20210A突变)、肺动脉高压、色素性视网膜炎、山德霍夫氏病、严重联合免疫缺陷综合征(SCID)、镰状细胞性贫血、脊髓性肌萎缩、斯特格氏病、泰-萨克斯病、乌谢尔综合征、X连锁免疫缺陷、斯特奇-韦伯综合征、以及癌症。

[0063] 在另一个方面,本披露涉及治疗感染或疑似感染EB病毒(EBV)的受试者(例如,人受试者)的方法,该方法包括向受试者施用有效量的本文所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞,从而治疗该感染或疑似感染EB病毒(EBV)的受试者。

[0064] 在一些实施例中,该受试者患有单核细胞增多症或癌症(例如,伯基特淋巴瘤、霍奇金氏病和鼻咽癌)。

[0065] 在另一个方面,本披露涉及治疗患有脊髓性肌肉萎缩症(SMA)的受试者(例如,人受试者)的方法,该方法包括向受试者施用有效量的本文所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞,从而治疗该患有SMA的受试者。

[0066] 在另一个方面,本披露涉及治疗患有肌萎缩性侧索硬化(ALS)的受试者(例如,人受试者)的方法,该方法包括向受试者施用有效量的本文所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞,从而治疗该患有ALS的受试者。

附图说明

[0067] 图1A和图1B示出了示例性RNA编辑器组成的示意图:GluA2.RBD-hADARDD(1A)和GluA2 mRNA序列的预期结果编辑示意图(1B)。

[0068] 图2示出了在脊髓性肌肉萎缩症患者中人SMN2剪接的示意图。

[0069] 图3示出了示例性的RNA编辑器组成:SMN2.RBD-hADARDD和SMN2mRNA序列预期产生的校正性编辑示意图。

[0070] 图4示出了以下示意图,其显示了EBNA1序列的编辑以增强病毒mRNA的二级结构,从而在宿主中诱导免疫应答,其中该二级结构如MFOLD所预测。

具体实施方式

[0071] 本发明描述了RNA编辑组合物和相关方法。本文所述的组合物(例如,药物组合物)包含多肽,该多肽包含RNA结合结构域,该RNA结合结构域包含多个(例如,2-50、2-30、15-30、16-21、5-20、5-15、5-10个)RNA碱基结合基序,每个基序与RNA碱基结合,并且这些基序在RNA结合结构域中有序排列以与靶RNA序列中连续顺序的RNA碱基结合,该RNA结合结构域连接至异源RNA编辑结构域,例如,脱氨酶,例如腺苷脱氨酶或胞苷脱氨酶。本文所述的组合物和方法可以用于修饰RNA序列,例如改变以下各项中的一种或多种:RNA的二级和/或三级结构;剪接;编码多肽的氨基酸序列;或编码多肽的表达水平,或增加或消除终止密码子(例如,提前终止密码子)。在实施例中,该RNA结合结构域结合RNA并且该RNA编辑结构域编辑RNA以减少或增加RNA的二级和/或三级结构,和/或改变RNA的剪接。在一些实施例中,该组合物减少双链RNA结构的量,例如,以减少对RNA的免疫应答。在一些实施例中,该组合物增加双链RNA结构的量,例如,以增加对RNA的免疫应答。在一些实施例中,该组合物纠正引起病理性剪接产物的疾病相关突变。

[0072] 定义

[0073] 如本文所用，术语“结构域”是指有助于生物分子的特定功能的生物分子的结构。结构域可以包含生物分子的连续区域(例如，连续序列)或不同的非连续区域(例如，非连续序列)。蛋白质结构域的实例包括但不限于RNA结合结构域、效应子结构域、RNA编辑结构域。

[0074] 如本文所用，术语“外源的”，当相对于生物分子(例如核酸序列或多肽)使用时，是指通过人为干预将生物分子引入宿主基因组、细胞或生物中。例如，使用重组DNA技术或其他方法添加到现有基因组、细胞、组织或受试者中的核酸对于现有核酸序列、细胞、组织或受试者而言是外源的。

[0075] 如本文所用，当用于参考第二元件来描述第一元件时，术语“异源的”意思是第一元件和第二元件自然界中不以如所描述的布置存在。例如，异源多肽、核酸分子、构建体或序列是指(a)多肽、核酸分子或多肽或核酸分子序列的一部分，其对于表达其的细胞而言不是天然的，(b)相对于其天然状态已发生改变或突变的多肽或核酸分子或多肽或核酸分子的一部分，或(c)具有与在类似条件下的天然表达水平相比改变的表达的多肽或核酸分子。例如，异源调节序列(例如启动子，增强子)可以用于调节基因或核酸分子的表达，其方式不同于基因或核酸分子通常在自然界中表达的方式。在另一个实例中，多肽或核酸序列的异源结构域(例如，多肽的RNA结合结构域或编码多肽的RNA结合结构域的核酸)可以相对于其他结构域布置，或者可以是不同的序列或相对于多肽的其他结构域或部分或其编码核酸来自不同来源。在某些实施例中，异源核酸分子可以存在于天然宿主细胞基因组中，但是可以具有改变的表达水平或具有不同的序列或两者。在其他实施例中，异源核酸分子对于宿主细胞或宿主基因组可能不是内源的，而是可能是已经通过转化(例如，转染，电穿孔)引入宿主细胞中，其中添加的分子可以整合到宿主基因组中或可以作为染色体外遗传物质短暂(例如mRNA)存在或半稳定存在超过一代(例如游离病毒载体、质粒或其他自我复制载体)。

[0076] 如本文所用，当应用于核酸序列时，术语“突变的”、“突变”及同源词是指与参考核酸序列(例如天然、野生型或非病理性核酸序列)相比，核酸序列中的核苷酸可以被插入、缺失或改变(例如，点突变)。

[0077] 如本文所用，“核酸分子”是指RNA和DNA分子两者，包括但不限于cDNA、基因组DNA、mRNA、tRNA，并且还包括合成的核酸分子，例如化学合成或重组产生的核酸分子，例如本文所述的核酸序列。核酸分子可以是双链或单链、其组合、环状或线性的。如果是单链，则核酸分子可以是有义链或反义链。如本领域技术人员将容易理解的，核酸序列可以进行化学或生物化学修饰或者可以含有非天然或衍生的核苷酸碱基。此类修饰包括例如标记，甲基化，用类似物取代一个或多个天然核苷酸，核苷酸间修饰，例如不带电荷的连接(例如，甲基膦酸酯、磷酸三酯、氨基磷酸酯、氨基甲酸酯等)、带电荷的连接(例如，硫代磷酸酯、二硫代磷酸酯等)，侧链部分(例如，多肽)，嵌入剂(例如，吖啶，补骨脂素等)，螯合剂，烷基化剂和经修饰的连接(例如，α异头核酸等等)。还包括合成的分子，它们模拟多核苷酸通过氢键和其他化学相互作用与指定序列结合的能力。这样的分子是本领域已知的，并且包括例如其中肽键替代分子主链中的磷酸键的那些。其他修饰可以包括，例如，其中核糖环包含桥接部分或其他结构(例如在“锁定”核酸中发现的修饰)的类似物。

[0078] 如本文所用，多肽的“RNA结合结构域”是多肽的特异性结合靶RNA序列的结构域。RNA结合结构域可以包含多个RNA碱基结合基序，每个基序都能够与RNA碱基特异性结合，并

且这些基序在RNA结合结构域中有序排列以与靶RNA序列中的连续顺序的RNA碱基结合。如本文所用,“PUM RNA结合基序”是与Pumilio同源结构域(PUM-HD)的RNA碱基结合重复序列同源的或由其衍生的基序。在实施例中,PUM RNA结合基序与PUM-HD的RNA碱基结合重复序列至少80%(例如,85%、87%、90%、92%、95%、97%、98%、99%或100%)相同,并且对特定RNA碱基具有结合特异性。在一些实施例中,该PUM RNA结合基序具有模块单元。在一些实施例中,该模块单元与RNA碱基腺嘌呤结合,其中模块单元氨基酸1是半胱氨酸,模块单元氨基酸2是酪氨酸,并且模块单元氨基酸5是谷氨酰胺。在一些实施例中,该模块单元与RNA碱基尿嘧啶结合,其中模块单元氨基酸1是天冬酰胺,模块单元氨基酸2是酪氨酸,并且模块单元氨基酸5是谷氨酰胺。在一些实施例中,该模块单元结合RNA碱基鸟嘌呤,其中模块单元氨基酸1是丝氨酸,模块单元氨基酸2是酪氨酸,并且模块单元氨基酸5是谷氨酸。在一些实施例中,该模块单元结合RNA碱基胞嘧啶,其中模块单元氨基酸1是丝氨酸,模块单元氨基酸2是酪氨酸,并且模块单元氨基酸5是精氨酸。在一些实施例中,该模块单元结合胞嘧啶,其中模块单元氨基酸1是丝氨酸,模块单元氨基酸2是酪氨酸,并且模块单元氨基酸5是精氨酸。例如,在Adamala等人2016.PNAS[美国科学院院报]113(19):E2579-E2588和在US 2016/0238593中找到了设计和制备此类模块单元以及RNA结合基序和结构域的方法。

[0079] 如本文所用,“RNA效应子”是作用于RNA以调节其结构和/或功能(例如编辑靶RNA的核苷酸序列)的部分。RNA效应子的实例是编辑靶RNA序列的一个或多个碱基的酶的催化结构域(“RNA编辑”结构域),例如脱氨酶(例如,将胞嘧啶编辑成尿嘧啶的胞苷脱氨酶、将腺苷编辑成肌苷的腺苷脱氨酶)的催化结构域,或APOBEC3A的催化结构域,据报道,APOBEC3A具有将G转化为A的能力(例如,如在Ahmadreza等人2015.PloS one[公共科学图书馆综合]10.3:e0120089)。此类酶包括作用于RNA的腺苷脱氨酶(ADAR)(例如,人ADAR 1、人ADAR2、人ADAR3、或人ADAR4);作用于tRNA的腺苷脱氨酶(ADAT)、作用于RNA的胞嘧啶脱氨酶(CDAR)、APOBEC、APOBEC3A A3A、TadA或CDA。

[0080] 如本文所用,术语“宿主”细胞是指已将蛋白质和/或遗传物质引入其中的细胞和/或其基因组。该术语不仅旨在指特定的受试者细胞和/或基因组,而且还指这种细胞的后代和/或这种细胞的后代的基因组。因为由于突变或环境影响,某些修饰可能在后代中发生,所以这样的后代实际上可能与亲本细胞不同,但仍包括在本文所用的术语“宿主细胞”的范围内。宿主基因组或宿主细胞可以是在培养物中生长的分离的细胞或细胞系,或者是从这种细胞或细胞系分离的基因组材料,或者可以是组成活组织或生物体的宿主细胞或宿主基因组。

[0081] 如本文所用,本文所述的组合物的术语“有效”或“足够”的量和/或时间是指如下的量和/或时间:当施用于细胞、组织或受试者(包括哺乳动物(例如人)时足以产生所希望的结果,包括在细胞水平、组织水平或临床结果上的效果,因此,“有效”或“足够”或其同义词取决于它所应用的背景。例如,在调节RNA结构的情况下,与不施用组合物(例如,多肽、核酸、载体等)获得的应答相比,该组合物的量足以实现对RNA结构的改变。与该量相对应的本文所述的给定组合物的量将根据多种因素而变化,例如给定的药剂、药物配制品、施用途径、疾病或障碍的类型、细胞、组织或受试者的身份(例如年龄、性别、体重)或接受治疗的宿主等,但仍然可以由本领域技术人员常规确定。同样,如本文所用,本披露的组合物的“治疗有效量”是与对照相比,在受试者中产生有益或所希望的结果的量。如本文所定义,本披露

的组合物的治疗有效量可以由普通技术人员通过本领域已知的常规方法容易地确定。剂量方案可以调整以提供最佳的治疗应答。

[0082] 如本文所用，术语“增加”和“减少”是指相对于参考，调节分别产生指标的功能、表达或活性更大或更小的量。例如，在施用本文所述的组合物之后，相对于施用前的量，如本文所述的RNA功能和/或结构(例如，表达或调节活性)可以在细胞、组织或受试者中增加或减少至少5％、10％、15％、20％、25％、30％、35％、40％、45％、50％、55％、60％、65％、70％、75％、80％、85％、90％、95％或98％或更多。通常，在受试者的背景下，在施用已达到所述作用的时间(例如，小时、天)，例如至少1周、1个月、3个月、或6个月时，或在开始治疗方案后，测量施用后的指标。

[0083] 如本文所用，“药物组合物”或“药物制剂”是对减轻、治疗或预防疾病具有药理学活性或其他直接效果的组合物或配制品，和/或其最终剂型或配制品，且指示供人使用。药物组合物典型地为GMP级，即符合用于人类的组合物的美国管理(FDA)规范。例如，典型地对GMP级组合物的内毒素进行测试，并符合具有少于规定量的内毒素的释放标准。

[0084] 如本文所用，“治疗(Treatment和treating)”是指对受试者进行的旨在改善、减轻、稳定(即不恶化)、预防或治愈疾病、病理状况或障碍的医疗管理。该术语包括活性治疗(旨在改善疾病、病理状况或障碍的治疗)、病因治疗(针对相关疾病、病理状况或障碍的原因的治疗)、姑息治疗(旨在缓解症状的治疗)、预防性治疗(旨在最大程度减少或部分或完全抑制相关疾病、病理状况或障碍发展的治疗)；以及支持治疗(用于补充另一治疗的治疗)。治疗还包括削减疾病或病症的程度；预防疾病或病症的传播；延迟或减缓疾病或病症的进展；改善或减轻疾病或病症；以及缓解(不管是部分缓解还是完全缓解)，不管是可检测的还是不可检测的。改善或减轻疾病或病症意指与缺乏治疗下的程度或时间进程相比，该疾病、障碍或病症的程度和/或不良临床表现减少和/或进展的时间进程被减缓或加长。“治疗”还可以意指与如果不接受治疗的预期生存期相比，延长的生存期。需要治疗的那些包括已经患有病症或障碍的那些，以及易于患有病症或障碍的那些，或其中要预防病症或障碍的那些。

[0085] RNA结合结构域

[0086] 本文所述的多肽RNA结合结构域特异性结合靶RNA序列。RNA结合结构域可以包含多个RNA碱基结合基序，每个基序都能够与RNA碱基特异性结合，并且这些基序在RNA结合结构域中有序排列例如以与靶RNA序列中的连续顺序的RNA碱基结合。RNA结合基序可以基于与Pumilio同源结构域(PUM-HD)的RNA碱基结合重复序列同源的或由其衍生的序列(“PUM RNA结合基序”)。在实施例中，PUM RNA结合基序与PUM-HD的RNA碱基结合基序至少80％(例如，85％、87％、90％、92％、95％、97％、98％、99％或100％)相同，并且对特定RNA碱基具有结合特异性。在PUM RNA结合基序中，基于结合核酸的沃森-克里克边缘的、受氢键或范德华相互作用控制的拓扑等效蛋白质表面上的保守位置，对靶RNA碱基的特异性进行工程化。这些拓扑结构用以下方式靶向RNA：用第1位和第5位的谷氨酸和丝氨酸以识别鸟嘌呤；用谷氨酰胺和半胱氨酸/丝氨酸以识别腺嘌呤；并且用谷氨酰胺和天冬酰胺以识别尿嘧啶。例如，在Lu等人2009.Curr Opin Struct Biol.[结构生物学最新观点]19(1):110-115;Adamala等人2016.PNAS[美国科学院院报]113(19):E2579-E2588和US 2016/0238593中找到了设计和制备此类模块单元以及RNA结合基序和结构域的方法。

[0087] 在实施例中，该RNA结合结构域与野生型PUM-HD(例如，野生型人PUM1-HD)的相应氨基酸序列具有至少80％同一性(例如，至少85％同一性、至少87％同一性、至少90％同一性、至少92％同一性、至少95％同一性、至少97％同一性、至少98％同一性、或99％同一性)并小于100％同一性。在一个实例中，表A中显示了靶向突变的HsPUM1-HD RNA结合基序和相关的识别核苷酸(来自Wang等人2002.Cell[细胞]110(4):501-12)。

[0088] 表A

[0089]

重复序列	重复序列的 AA 序列	靶残基	核苷酸
R1	HIMEFSQDQHGSRFIQLKLERATP AERQLVFNEILQ	Ser863, Arg864, Gln867	A
R3	HVLSLALQMYGCRVIQKALEFIPS DQQNEMVRELDG	Cys935, Arg936, Gln939	A
R5	QVFALSTHPYGCRVIQRILEHCLP DQTLPILEELHQ	Cys1007, Arg1008, Gln1011	A
R7	NVLVLSQHKFASNVVEKCVTHAS RTERAVLIDEVCTMNDGPHS	Ser1079, Ans1080, Glu1083	G
R2	AAYQLMVDVFGNYVIQKFFEFGS LEQKLALAERIRG	Asn899, Tyr900, Gln903	U
R4	HVLKCVKDQNGNHVVQKCIECV QPQSLQFIIDAFKG	Asn971, His972, Gln975	U
R6	HTEQLVQDQYGNYVIQHVLEHG RPEDKSKIVAEIRG	Asn1043, Tyr1044, Gln1047	U
R8	ALYTMMKDQYANYVVQKMIDV AEPGQRKIVMHKIRP HIATLRKYTYGKHILAKLEKYYM KNGVDLG	Tyr1123, Asn 1122, Gln1126	U

[0090] 例如，在重复序列7中结合尿嘧啶(U)而不是鸟嘌呤(G)，需要进行以下氨基酸残基改变:E1083Q、S1079N和N1080Y，如以下所述，例如Cheong和Tanaka.2006.PNAS[美国科学院院报]第103卷,37:13635-9。

[0091] 设计工程化的RNA结合结构域以结合靶RNA序列。典型地，RNA结合结构域结合2-50个RNA核苷酸(例如，2-50个核苷酸(例如，2-50、2-40、2-30、2-25、2-24、2-23、2-22、2-21、2-20、2-19、2-18、2-17、2-16、2-15、2-14、2-13、2-12、2-11、2-10、2-9、2-8、5-50、5-40、5-30、5-25、5-24、5-23、5-22、5-21、5-20、5-19、5-18、5-17、5-16、5-15、5-14、5-13、5-12、5-11、5-10、5-9、5-8、10-50、10-40、10-30、10-25、10-24、10-23、10-22、10-21、10-20、10-19、10-18、10-17、10-16、10-15、10-14、10-13、10-12、10-11、15-50、15-40、15-30、15-25、15-24、15-23、15-22、15-21、15-20、15-19、15-18、15-17、15-16、16-50、16-40、16-30、16-25、16-24、16-23、16-22、16-21、16-20、16-19、16-18、16-17、17-50、17-40、17-30、17-25、17-24、17-23、17-22、17-21、17-20、17-19、17-18、18-50、18-40、18-30、18-25、18-24、18-23、18-22、18-21、18-20、18-19、19-50、19-40、19-30、19-25、19-24、19-23、19-22、19-21、19-20、20-50、20-40、20-30、20-25、20-24、20-23、20-22、20-21、21-50、21-40、21-30、21-25、21-24、21-23、21-22、25-50、25-40、25-30、30-50、30-40,或40-50、3-20、3-15、3-10、3-9、3-8、4-12、4-10、4-9、4-8、5-10、5-9、5-8个核苷酸)的靶序列。在一些实施例中，该RNA结合结构域结合16-21个RNA核苷酸的靶序列。在一些实施例中，该RNA结合结构域结合至少16个RNA核苷酸(并且任选地不超过30、29、28、27、26、25、24、23、22、或21个RNA核苷酸)。多个RNA碱基结合基序可能包括至少3个(例如，至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、

2-20个、2-15个、2-10个、2-8个、3-20个、3-15个、3-10个、3-8个、4-8个、多达25个、多达30个)PUM RNA结合基序。在一些实施例中,该RNA结合结构域包含2-50、2-40、2-30、2-25、2-24、2-23、2-22、2-21、2-20、2-19、2-18、2-17、2-16、2-15、2-14、2-13、2-12、2-11、2-10、2-9、2-8、5-50、5-40、5-30、5-25、5-24、5-23、5-22、5-21、5-20、5-19、5-18、5-17、5-16、5-15、5-14、5-13、5-12、5-11、5-10、5-9、5-8、10-50、10-40、10-30、10-25、10-24、10-23、10-22、10-21、10-20、10-19、10-18、10-17、10-16、10-15、10-14、10-13、10-12、10-11、15-50、15-40、15-30、15-25、15-24、15-23、15-22、15-21、15-20、15-19、15-18、15-17、15-16、16-50、16-40、16-30、16-25、16-24、16-23、16-22、16-21、16-20、16-19、16-18、16-17、17-50、17-40、17-30、17-25、17-24、17-23、17-22、17-21、17-20、17-19、17-18、18-50、18-40、18-30、18-25、18-24、18-23、18-22、18-21、18-20、18-19、19-50、19-40、19-30、19-25、19-24、19-23、19-22、19-21、19-20、20-50、20-40、20-30、20-25、20-24、20-23、20-22、20-21、21-50、21-40、21-30、21-25、21-24、21-23、21-22、25-50、25-40、25-30、30-50、30-40、或40-50个PUM RNA结合基序,例如,PUM RNA结合基序的数目对应于所结合的RNA核苷酸的数目(例如,靶RNA序列的长度)。

[0092] 在一些实施例中,RNA结合结构域结合由GluA2(例如,人GluA2)基因编码的mRNA中的靶RNA序列。在一些实施例中,该RNA结合结构域与靶RNA序列结合,该靶RNA序列包含对应于人GluA2基因的1537-1552的核苷酸,或参考序列NM_000826中核苷酸1537-1552的50个碱基内的核酸序列。

[0093] 在一些实施例中,RNA结合结构域结合由SMN2(例如,人SMN2)基因编码的mRNA中的靶RNA序列。在一些实施例中,该RNA结合结构域与靶RNA序列结合,该靶RNA序列包含对应于人SMN2基因的31,995-32,010的核苷酸,或参考序列NM-022876中核苷酸31,995-32,010的50个碱基内的核酸序列。

[0094] 本文所述的RNA结合结构域可以是90-500个氨基酸残基,例如,90-450个氨基酸残基、90-400个氨基酸残基、90-350个氨基酸残基、90-300个氨基酸残基、120-400个氨基酸残基。RNA结合结构域可以结合RNA序列,例如mRNA序列,例如折叠成二级或三级结构的mRNA序列,例如双链RNA序列。RNA结合结构域可以结合RNA序列,例如mRNA序列,例如包含疾病相关突变(例如点突变)的mRNA序列。

[0095] 在一些实施例中,本文所述的PUM RNA结合基序与胞嘧啶结合。更特别地,可以将PUM RNA结合基序工程化以结合胞嘧啶,例如通过US 10233218 B2的方法,将其通过引用特此并入。在一些实施例中,RNA结合结构域包含一个或多个与胞嘧啶结合的PUM RNA结合基序。例如,结合胞嘧啶的PUM RNA结合基序可以包含具有式$X_1X_2X_3X_4X_5X_6X_7X_8X_9X_{10}X_{11}$的序列,其中:

[0096] X_1是谷氨酰胺(Q),X_2是组氨酸(H);X_3是甘氨酸(G);X_4选自包括以下的组:甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)和半胱氨酸(C);X_5是精氨酸(R);X_6是苯丙氨酸(F);X_7是异亮氨酸(I);X_8是精氨酸(R);X_9是亮氨酸(L);X_{10}是赖氨酸(K);并且X_{11}是亮氨酸(L);或

[0097] X_1是缬氨酸(V);X_2是苯丙氨酸(F);X_3是甘氨酸(G);X_4选自包括以下的组:甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)和半胱氨酸(C);X_5是酪氨酸(Y);X_6是缬氨酸(V);X_7是异亮氨酸(I);X_8是精氨酸(R);X_9是赖氨酸(K);X_{10}是苯丙氨酸(F);并且X_{11}是苯丙氨酸

(F)；或

[0098] X_1是甲硫氨酸(M)；X_2是酪氨酸(Y)；X_3是甘氨酸(G)；X_4选自包括以下的组：甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)和半胱氨酸(C)；X_5是精氨酸(R)；X_6是缬氨酸(V)；X_7是异亮氨酸(I)；X_8是精氨酸(R)；X_9是赖氨酸(K)；X_{10}是丙氨酸(A)；并且X_{11}是亮氨酸(L)；或

[0099] X_1是谷氨酰胺(Q)；X_2是天冬酰胺(N)；X_3是甘氨酸(G)；X_4选自包括以下的组：甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)和半胱氨酸(C)；X_5是组氨酸(H)；X_6是缬氨酸(V)；X_7是缬氨酸(V)；X_8是精氨酸(R)；X_9是赖氨酸(K)；X_{10}是半胱氨酸(C)；并且X_{11}是异亮氨酸(I)；或

[0100] X_1是脯氨酸(P)；X_2是酪氨酸(Y)；X_3是甘氨酸(G)；X_4选自包括以下的组：甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)和半胱氨酸(C)；X_5是精氨酸(R)；X_6是缬氨酸(V)；X_7是异亮氨酸(I)；X_8是精氨酸；(R)；X_9是精氨酸(R)；X_{10}是异亮氨酸(I)；并且X_{11}是亮氨酸(L)；或

[0101] X_1是谷氨酰胺(Q)；X_2是酪氨酸(Y)；X_3是甘氨酸(G)；X_4选自包括以下的组：甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)和半胱氨酸(C)；X_5是酪氨酸(Y)；X_6是缬氨酸(V)；X_7是异亮氨酸(I)；X_8是精氨酸；(R)；X_9是组氨酸(H)；X_{10}是缬氨酸(V)；并且X_{11}是亮氨酸(L)；或

[0102] X_1是赖氨酸(K)；X_2是苯丙氨酸(F)；X_3是丙氨酸(A)；X_4选自包括以下的组：甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)和半胱氨酸(C)；X_5是天冬酰胺(N)；X_6是缬氨酸(V)；X_7是缬氨酸(V)；X_8是精氨酸；(R)；X_9是赖氨酸(K)；X_{10}是半胱氨酸(C)；并且X_{11}是缬氨酸(V)；或

[0103] X_1是谷氨酰胺(Q)；X_2是酪氨酸(Y)；X_3是丙氨酸(A)；X_4选自包括以下的组：甘氨酸(G)、丙氨酸(A)、丝氨酸(S)、苏氨酸(T)和半胱氨酸(C)；X_5是酪氨酸(Y)；X_6是缬氨酸(V)；X_7是缬氨酸(V)；X_8是精氨酸；(R)；X_9是赖氨酸(K)；X_{10}是甲硫氨酸(M)；并且X_{11}是异亮氨酸(I)。

[0104] 例如，结合胞嘧啶的RNA结合基序可以包含氨基酸序列QYGGYVIRHVL(SEQ ID NO:100)。在一些实施例中，包含结合胞嘧啶的RNA结合基序的RNA结合结构域可以包含以下氨基酸序列：

[0105] GRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQHGSRFIQLKLERATPAERQLVFNEILQAAYQLMVDVFGNYVIQKFFEFGSLEQKLALAERIRGHVLSLALQMYGCRVIQKALEFIPSDQQVINEMVRELDGHVLKCVKDQNGNHVVQKCIECVQPQSLQFIIDAFKGQVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQHTEQLVQDQYGGYVIRHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPHSALYTMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYMKNGVDLG(SEQ ID NO:101)

[0106] 示例性RNA结合结构域

[0107] 示例性RNA结合结构域(例如，包含多个RNA结合基序(例如，多个PUM RNA结合基序或与PUM-HD同源的或由其衍生的序列))包括SEQ ID NO:13或15-21的RNA结合结构域或如由SEQ ID NO:4或6-12所编码的RNA结合结构域。在一些实施例中，RNA结合结构域包含与SEQ ID NO:13或15-21的RNA结合结构域具有至少80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、或99％同一性(或相对于其包含不超过1、2、3、4、5、6、7、8、9、或10个碱基改变)的氨基酸序列，或由与SEQ ID NO:4或6-12的RNA结合结构域编码序列具有至少80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％、或99％同一性的核酸序列编码。在一些实施例中，RNA结合结构域包含来自第一示例性RNA结合结构域的一个或多个

RNA结合基序以及来自第二示例性RNA结合结构域的一个或多个RNA结合基序。

[0108] 具有(G4S)3接头的双PUF设计(野生型PUF靶序列)

[0109] mRNA序列：

augggcaggagcaggcuuuuggaagauuuucgaaacaaccgCuaccccaauuuacaacugcgggagauugcugga
cauauaauggaauuuucccaagaccagcauggguccagauucauucagcugaaacuggagcgugccacaccagcug
agcgccagcuugucuucaaugaaauccuccaggcugccuaccaacucaugguggauguguuugguaauuacgucau
ucagaaguucuuugaauuuggcagucuugaacagaagcuggcuuuggcagaacggauucgaggccacguccuguc
auuggcacuacagauguauggcugccguguuauccagaaagcucuugaguuuauuccuucagaccagcagaaugag
augguucgggaacuagauggccaugucuugaagugugugaaagaucagaauggcaaucacgugguucagaaaugc
auugaaaugguacagccccagucuuugcaauuuaucaucgaugcguuuaagggacagguauuugccuuauccacac
auccuuauggcugccgagugauucagagaauccuggagcacugucucccugaccagacacucccuauuuuagagga
gcuucaccagcacacagagcagcuuguacaggaucaauauggaaauuauguaauccaacauguacuggagcacggu
cguccugaggauaaaagcaaaauuguagcagaaauccgaggcaauguacuuguauugagucagcacaaauuugcaa
gcaauguuguggagaaguguguuacucacgccucacguacggagcgcgcugugcucaucgaugaggugugcacca
ugaacgacgguccccacagugccuuauacaccaugaugaaggaccaguaugccaacuacgugguccagaagaugau
ugacguggcggagccaggccagcggaagaucgucaugcauaagauccggccccacaucgcaacucuucguaaguac
accuauggcaagcacauucuggccaagcuggagaaguacuacaugaagaacgguguugacuuagggGGAGGU
GGCGGAUCGGGAGGUGGCGGAUCGGGAGGUGGCGGAUCGggcaggagcaggcuuuu
ggaagauuuucgaaacaaccgCuaccccaauuuacaacugcgggagauugcuggacauauaauggaauuuucccaa
gaccagcauggguccagauucauucagcugaaacuggagcgugccacaccagcugagcgccagcuugucuucaaug
aaauccuccaggcugccuaccaacucaugguggauguguuugguaauuacgucauucagaaguucuuugaauuug
gcagucuugaacagaagcuggcuuuggcagaacggauucgaggccacguccugucauuggcacuacagauguaug
gcugccguguuauccagaaagcucuugaguuuauuccuucagaccagcagaaugagaugguucgggaacuagaug
gccaugucuugaagugugugaaagaucagaauggcaaucacgugguucagaaaugcauugaaauguguacagcccca
gucuuugcaauuuaucaucgaugcguuuaagggacagguauuugccuuauccacacauccuuauggcugccgagu
gauucagagaauccuggagcacugucucccugaccagacacucccuauuuuagaggagcuucaccagcacacagagc
agcuuguacaggaucaauauggaaauuauguaauccaacauguacuggagcacggucguccugaggauaaaagcaa
aauuguagcagaaauccgaggcaauguacuuguauugagucagcacaaauuugcaagcaauguuguggagaaagug
uguuacucacgccucacguacggagcgcgcugugcucaucgaugaggugugcaccaugaacgacgguccccacagu
gccuuauacaccaugaugaaggaccaguaugccaacuacgugguccagaagaugauugacguggcggagccaggcc
agcggaagaucgucaugcauaagauccggccccacaucgcaacucuucguaaguacaccuauggcaagcacauucug
gccaagcuggagaaguacuacaugaagaacgguguugacuuaggguga
(SEQ ID NO: 4)

[0110]

[0111] 蛋白质序列：

MGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQHGSRFIQLKLERATPAERQLV
FNEILQAAYQLMVDVFGNYVIQKFFEFGSLEQKLALAERIRGHVLSLALQMYGC
RVIQKALEFIPSDQQNEMVRELDGHVLKCVKDQNGNHVVQKCIECVQPQSLQFII
DAFKGQVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQHTEQLVQDQYGNYV
IQHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFASNVVEKCVTHASRTERAVLIDE
VCTMNDGPHSALYTMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRPHIATLR
KYTYGKHILAKLEKYYMKNGVDLGGGGGSGGGGSGGGGSGRSRLLEDFRNNR
YPNLQLREIAGHIMEFSQDQHGSRFIQLKLERATPAERQLVFNEILQAAYQLMVD
VFGNYVIQKFFEFGSLEQKLALAERIRGHVLSLALQMYGCRVIQKALEFIPSDQQ
NEMVRELDGHVLKCVKDQNGNHVVQKCIECVQPQSLQFIIDAFKGQVFALSTHP
YGCRVIQRILEHCLPDQTLPILEELHQHTEQLVQDQYGNYVIQHVLEHGRPEDKS
KIVAEIRGNVLVLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPHSALY
TMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILAKLEK
YYMKNGVDLG* (SEQ ID NO: 13)

[0112]

[0113] 具有E488Q突变的ADAR2DD(氨基酸299-701)的结构域

[0114] mRNA序列：

[0115] AUGCUCCACCUCGACCAAACACCCAGCAGACAGCCUAUCCCUUCCGAAGGA
CUGcagcugcauuuaccgcagguuuuagcugacgcugucucacgccugguccuggguaaguuuggugaucugac
cgacaacuucuccucccccucacgcucgcagaaaagugcuggcuggagucgucaugacaacaggcacagauguuaaa
gaugccaaggugauaaguguuucuacaggaggcaaauguauuaauggugaauacaugagugaucguggccuugca
uuaaaugacugccaugcagaaauaauaucucggagauccuugcucagauuucuuuauacacaacuugagcuuuacu
uaaauaacaaagaugaucaaaaaagauccaucuuucagaaaucagagcgaggggguuuaggcugaaggagaaugu
ccaguuucaucuguacaucagcaccucucccuguggagaugccagaaucuucucaccacaugagccaauccuggaa
gaaccagcagauagacacccaaaucguaaagcaagaggacagcuacggaccaaaauagagucuggucaggggacgau
uccagugcgcuccaaugcgagcauccaaacgugggacggggugcugcaaggggagcggcugcucaccauguccugc
agugacaagauugcacgcuggaacgugguggggcauccagggaucacugcucagcauuuucguggagcccauuuac
uucucgagcaucauccugggcagccuuuaccacggggaccaccuuuccagggccauguaccagcggaucuccaaca
uagaggaccugccaccucucuacacccucaacaagccuuugcucaguggcaucagcaaugcagaagcacggcagcca
gggaaggccccaacuucagugucaacuggacgguaggcgacuccgcuauugaggucaucaacgccacgacugggga
aggaugagcugggccgcgcgucccgccuguguaagcacgcguuguacugucgcuggaugcgugugcacggcaagg
uucccucccacuuacuacgcuccaagauuaccaagcccaacguguaccaugaguccaagcuggcggcaaaggaguac
caggccgccaaggcgcgucuguucacagccuucaucaaggcggggcuggggccuggguggagaagcccaccgagc
aggaccaguucucacucacgCCUUGA
(SEQ ID NO: 5)

[0116] 蛋白质序列：

[0117] MLHLDQTPSRQPIPSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVL
AGVVMTTGTDVKDAKVISVSTGGKCINGEYMSDRGLALNDCHAEIISRRSLLRFL
YTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPI
LEEPADRHPNRKARGQLRTKIESGQGTIPVRSNASIQTWDGVLQGERLLTMSCSD
KIARWNVVGIQGSLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTL
NKPLLSGISNAEARQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHA
LYCRWMRVHGKVPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGL
GAWVEKPTEQDQFSLTP* (SEQ ID NO: 14)

[0118] 与ADAR2DD融合的双PUF设计的融合多肽(野生型PUF靶序列)

[0119] mRNA序列：

[0120] AUGGACUAUAAGGACCACGACGGAGACUACAAGGAUCAUGAUAUUGAUUA
CAAAGACGAUGACGAUAAGAUGGCCCCAAAGAAGAAGCGGAAGGUCGGUA
UCCACGGAGUCCCAGCAGCCCUCCACCUCGACCAAACACCCAGCAGACAGC
CUAUCCCUUCCGAAGGACUGcagcugcauuuaccgcagguuuuagcugacgcugucucacgccugg
uccuggguaaguuuggugaucugaccgacaacuucuccucccccucacgcucgcagaaaagugcuggcuggagucgu
caugacaacaggcacagauguuaaagaugccaaggugauaaguguuucuacaggaggcaaauguauuaauggugaa
uacaugagugaucguggccuugcauuaaaugacugccaugcagaaauaauaucucggagauccuugcucagauuuc
uuuauacacaacuugagcuuuacuuaaaauaacaaagaugaucaaaaaagauccaucuuucagaaaucagagcgaggg
ggguuuaggcugaaggagaauguccaguuucaucuguacaucagcaccucucccuguggagaugccagaaucuuc
ucaccacaugagccaauccuggaagaaccagcagauagacacccaaaucguaaagcaagaggacagcuacggaccaaa
auagagucuggucaggggacgauuccagugcgcuccaaugcgagcauccaaacgugggacggggugcugcaaggg
gagcggcugcucaccauguccugcagugacaagauugcacgcuggaacgugguggggcauccagggaucacugcuca

gcauuuucguggagcccauuuacuucucgagcaucauccugggcagccuuuaccacggggaccaccuuuccagggc
cauguaccagcggaucuccaacauagaggaccugccaccucucuacacccucaacaagccuuugcucaguggcauca
gcaaugcagaagcacggcagccagggaaggccccc aacuucagugucaacuggacgguaggcgacuccgcuauuga
ggucaucaacgccacgacugggaaggaugagcugggccgcgcgucccgccuguguaagcacgcguuguacugucgc
uggaugcgugugcacggcaagguucccucccacuuacuacgcuccaagauuaccaagcccaacguguaccaugagu
ccaagcuggcggcaaaggaguaccaggccgccaaggcgcgucuguucacagccuucaucaaggcggggcuggggc
cuggguggagaagcccaccgagcaggaccaguucucacucacgCCUGGAGGUGGCGGAUCGGGAG
GUGGCGGAUCGGGAGGUGGCGGAUCGggcaggagcaggcuuuuggaagauuuucgaaacaac
cgCuaccccaauuuacaacugcgggagauugcuggacauauaauggaauuuucccaagaccagcauggguccagau
ucauucagcugaaacuggagcgugccacaccagcugagcgccagcuugucuucaaugaaauccuccaggcugccua
ccaacucaugguggauguguuugguaauuacgucauucagaaguucuuugaauuuggcagucuugaacagaagcu
ggcuuuggcagaacggauucgaggccacguccugucauuggcacuacagauguauggcugccguguuauccagaa
agcucuugaguuuauuccuucagaccagcagaaugagaugguucgggaacuagauggccaugucuugaaguguu
gaaagaucagaauggcaaucacgugguucagaaaugcauugaauguguacagccccagucuuugcaauuuaucauc
gaugcguuuaagggacagguauuugccuuauccacacauccuuauggcugccgagugauucagagaauccuggag
cacugucucccugaccagacacucccuauuuuagaggagcuucaccagcacacagagcagcuuguacaggaucaaua
uggaaauuauguaauccaacauguacuggagcacggucguccugaggauaaaagcaaaauuguagcagaaauccga
ggcaauguacuuguauugagucagcacaaauuugcaagcaauguuguggagaaguguguuacucacgccucacgu
acggagcgcgcugugcucaucgaugaggugugcaccaugaacgacgguccccacagugccuuauacaccaugauga
[0121] aggaccaguaugccaacuacgugguccagaagaugauugacguggcggagccaggccagcggaagaucgucaugca
uaagauccggccccacaucgcaacucuucguaaguacaccuauggcaagcacauucuggccaagcuggagaaguacu
acaugaagaacgguguugacuuaggGGAGGUGGCGGAUCGGGAGGUGGCGGAUCGGGA
GGUGGCGGAUCGggcaggagcaggcuuuuggaagauuuucgaaacaaccgCuaccccaauuuacaacugc
gggagauugcuggacauauaauggaauuuucccaagaccagcauggguccagauucauucagcugaaacuggagcg
ugccacaccagcugagcgccagcuugucuucaaugaaauccuccaggcugccuaccaacucaugguggauguguuu
gguaauuacgucauucagaaguucuuugaauuuggcagucuugaacagaagcuggcuuuggcagaacggauucga
ggccacguccugucauuggcacuacagauguauggcugccguguuauccagaaagcucuugaguuuauuccuuca
gaccagcagaaugagaugguucgggaacuagauggccaugucuugaaguguguugaaagaucagaauggcaaucacg
ugguucagaaaugcauugaauguguacagccccagucuuugcaauuuaucaucgaugcguuuaagggacagguau
uugccuuauccacacauccuuauggcugccgagugauucagagaauccuggagcacugucucccugaccagacacu
cccuauuuuagaggagcuucaccagcacacagagcagcuuguacaggaucaauauggaaauuauguaauccaacau
guacuggagcacggucguccugaggauaaaagcaaaauuguagcagaaauccgaggcaauguacuuguauugaguc
agcacaaauuugcaagcaauguuguggagaaguguguuacucacgccucacguacggagcgcgcugugcucaucga
ugaggugugcaccaugaacgacgguccccacagugccuuauacaccaugaugaaggaccaguaugccaacuacgug
guccagaagaugauugacguggcggagccaggccagcggaagaucgucaugcauaagauccggccccacaucgcaa
cucuucguaaguacaccuauggcaagcacauucuggccaagcuggagaaguacuacaugaagaacgguguugacuu
agggAGCGGCGGCAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGGCCAAG
AAGAAGAAGuga
(SEQ ID NO: 6)

[0122] 蛋白质序列：

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAALHLDQTPSRQPI
PSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDV
[0123] KDAKVISVSTGGKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNK
DDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPNRK
ARGQLRTKIESGQGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQG

[0124] SLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEA RQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGK VPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQ FSLTPGGGGSGGGGSGGGGSGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQH GSRFIQLKLERATPAERQLVFNEILQAAYQLMVDVFGNYVIQKFFEFGSLEQKLA LAERIRGHVLSLALQMYGCRVIQKALEFIPSDQQNEMVRELDGHVLKCVKDQNG NHVVQKCIECVQPQSLQFIIDAFKGQVFALSTHPYGCRVIQRILEHCLPDQTLPILE ELHQHTEQLVQDQYGNYVIQHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFASNV VEKCVTHASRTERAVLIDEVCTMNDGPHSALYTMMKDQYANYVVQKMIDVAE PGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYMKNGVDLGGGGGSGGGGS GGGGSGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQHGSRFIQLKLERATPAE RQLVFNEILQAAYQLMVDVFGNYVIQKFFEFGSLEQKLALAERIRGHVLSLALQ MYGCRVIQKALEFIPSDQQNEMVRELDGHVLKCVKDQNGNHVVQKCIECVQPQ SLQFIIDAFKGQVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQHTEQLVQDQY GNYVIQHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFASNVVEKCVTHASRTERA VLIDEVCTMNDGPHSALYTMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRPHI ATLRKYTYGKHILAKLEKYYMKNGVDLGSGGKRPAATKKAGQAKKKK* (SEQ ID NO: 15)

靶向人GluA2(参考序列NM_000826)核苷酸序列(aucaugaucaagaagc(SEQ ID NO:22))的核苷酸1537-1552的具有(G4S)3接头的双PUF设计

[0125] mRNA序列：

[0126] augGGCCGCAGCCGCCUUUUGGAAGAUUUUCGAAACAACCGGUACCCCAAU UUACAACUGCGGGAGAUUGCCGGACAUAUAAUGGAAUUUUCCCAAGACCA GCAUGGGUCCAGAUUCAUUCGCCUGAAACUGGAGCGUGCCACACCAGCUG AGCGCCAGCUUGUCUUUAAUGAAAUCCUCCAGGCUGCCUACCAACUCAUGG UGGAUGUGUUUGGUAGUUACGUCAUUGAGAAGUUCUUUGAAUUUGGCAGU CUUGAACAGAAGCUGGCUUUGGCAGAACGGAUUCGAGGUCACGUCCUGUC AUUGGCACUACAGAUGUAUGGCUGCCGUGUUAUCCAGAAAGCUCUUGAGU UUAUUCCUUCAGACCAGCAGAAUGAGAUGGUUCGGGAACUAGAUGGCCAU GUCUUGAAGUGUGUGAAAGAUCAGAAUGGCUGUCACGUGGUUCAGAAAUG CAUUGAAUGUGUACAGCCCCAGUCUUUGCAAUUUAUCAUCGAUGCGUUUA AGGGCCAGGUAUUUGCCUUAUCCACACAUCCUUAUGGCUCCCGAGUGAUU GAGAGAAUCCUGGAGCACUGUCUCCCUGACCAGACACUCCCUAUUUUAGA GGAGCUUCACCAGCACACAGAGCAGCUUGUACAGGAUCAAUAUGGAUGUU AUGUAAUCCAACAUGUACUGGAGCACGGUCGUCCUGAGGAUAAAAGCAAA AUUGUAGCAGAAAUCCGAGGCAAUGUACUUGUAUUGAGUCAGCACAAAUU UGCAUGCAAUGUUGUGCAGAAGUGUGUUACUCACGCCUCACGUACGGAGC GCGCUGUGCUCAUCGAUGAGGUGUGCACCAUGAACGACGGUCCCCACAGU GCCUUAUACACCAUGAUGAAGGACCAGUAUGCCAGCUACGUGGUCCGCAA GAUGAUUGACGUGGCGGAGCCAGGCCAGCGGAAGAUCGUCAUGCAUAAGA UCCGACCCCACAUCGCAACUCUUCGUAAGUACACCUAUGGCAAGCACAUUC UGGCCAAGCUGGAGAAGUACUACAUGAAGAACGGUGUUGACUUAGGGGGA

GGUGGCGGAUCGGGAGGUGGCGGAUCGGGAGGUGGCGGAUCGGGCCGCAG
CCGCCUUUUGGAAGAUUUUCGAAACAACCGGUACCCCAAUUUACAACUGC
GGGAGAUUGCCGGACAUAUAAUGGAAUUUUCCCAAGACCAGCAUGGGAAC
AGAUUCAUUCAGCUGAAACUGGAGCGUGCCACACCAGCUGAGCGCCAGCU
UGUCUUUAAUGAAAUCCUCCAGGCUGCCUACCAACUCAUGGUGGAUGUGU
UUGGUUGUUACGUCAUUCAGAAGUUCUUUGAAUUUGGCAGUCUUGAACAG
AAGCUGGCUUUGGCAGAACGGAUUCGAGGUCACGUCCUGUCAUUGGCACU
ACAGAUGUAUGGCUCCCGUGUUAUCGAGAAAGCUCUUGAGUUUAUUCCUU
CAGACCAGCAGAAUGAGAUGGUUCGGGAACUAGAUGGCCAUGUCUUGAAG
UGUGUGAAAGAUCAGAAUGGCAAUCACGUGGUUCAGAAAUGCAUUGAAUG
UGUACAGCCCCAGUCUUUGCAAUUUAUCAUCGAUGCGUUUAAGGGACAGG
[0127] UAUUUGCCUUAUCCACACAUCCUUAUGGCUGCCGAGUGAUUCAGAGAAUC
CUGGAGCACUGUCUCCCCUGACCAGACACUCCCUAUUUUAGAGGAGCUUCAC
CAGCACACAGAGCAGCUUGUACAGGAUCAAUAUGGAAGUUAUGUAAUCCG
CCAUGUACUGGAGCACGGUCGUCCUGAGGAUAAAAGCAAAAUUGUAGCAG
AAAUCCGAGGCAAUGUACUUGUAUUGAGUCAGCACAAAUUUGCAAACAAU
GUUGUGCAGAAGUGUGUUACUCACGCCUCACGUACGGAGCGCGCUGUGCU
CAUCGAUGAGGUGUGCACCAUGAACGACGGUCCCCACAGUGCCUUAUACAC
CAUGAUGAAGGACCAGUAUGCCUGCUACGUGGUCCAGAAGAUGAUUGACG
UGGCGGAGCCAGGCCAGCGGAAGAUCGUCAUGCAUAAGAUCCGACCCCACA
UCGCAACUCUUCGUAAGUACACCUAUGGCAAGCACAUUCUGGCCAAGCUG
GAGAAGUACUACAUGAAGAACGGUGUUGACUUAGGGuga
(SEQ ID NO: 7)

[0128] 蛋白质序列：

MGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQHGSRFIRLKLERATPAERQLV
FNEILQAAYQLMVDVFGSYVIEKFFEFGSLEQKLALAERIRGHVLSLALQMYGCR
VIQKALEFIPSDQQNEMVRELDGHVLKCVKDQNGCHVVQKCIECVQPQSLQFIID
AFKGQVFALSTHPYGSRVIERILEHCLPDQTLPILEELHQHTEQLVQDQYGCYVIQ
HVLEHGRPEDKSKIVAEIRGNVLVLSQHKFACNVVQKCVTHASRTERAVLIDEV
CTMNDGPHSALYTMMKDQYASYVVRKMIDVAEPGQRKIVMHKIRPHIATLRKY
[0129] TYGKHILAKLEKYYMKNGVDLGGGGGSGGGGSGGGGSGRSRLLEDFRNNRYPN
LQLREIAGHIMEFSQDQHGNRFIQLKLERATPAERQLVFNEILQAAYQLMVDVFG
CYVIQKFFEFGSLEQKLALAERIRGHVLSLALQMYGSRVIEKALEFIPSDQQNEM
VRELDGHVLKCVKDQNGNHVVQKCIECVQPQSLQFIIDAFKGQVFALSTHPYGC
RVIQRILEHCLPDQTLPILEELHQHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVA
EIRGNVLVLSQHKFANNVVQKCVTHASRTERAVLIDEVCTMNDGPHSALYTMM
KDQYACYVVQKMIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYM
KNGVDLG* (SEQ ID NO: 16)

[0130] 与ADAR2DD融合的双PUF设计的融合多肽(PUF靶向人GluA2(参考序列NM_000826)核苷酸序列[aucaugaucaagaagc](SEQ ID NO:22)的核苷酸1537-1552)

[0131] mRNA序列：

[0132]

AUGGACUAUAAGGACCACGACGGAGACUACAAGGAUCAUGAUAUUGAUUA
CAAAGACGAUGACGAUAAGAUGGCCCCAAAGAAGAAGCGGAAGGUCGGUA
UCCACGGAGUCCCAGCAGCCCUCCACCUCGACCAAACACCCAGCAGACAGC
CUAUCCCUUCCGAAGGACUGcagcugcauuuaccgcagguuuuagcugacgcugucucacgccugg
uccuggguaaguuuggugaucugaccgacaacuucuccucccccucacgcucgcagaaaagugcuggcuggagucgu
caugacaacaggcacagauguuaaagaugccaaggugauaaaguguuucuacaggaggcaaauguauuaauggugaa
uacaugagugaucguggccuugcauuaaaugacugccaugcagaaauaauaucucggagauccuugcucagauuuc
uuuauacacaacuugagcuuuacuuaaaauaacaaagaugaucaaaaaagauccaucuuucagaaaucagagcgaggg
ggguuuaggcugaaggagaauguccaguuucaucuguacaucagcaccucucccuguggagaugccagaaucuuc
ucaccacaugagccaauccuggaagaaccagcagauagacacccaaaucguaaagcaagaggacagcuacggaccaaa
auagagucuggucaggggacgauuccagugcgcuccaaugcgagcauccaaacgugggacggggugcugcaaggg
gagcggcugcucaccauguccugcagugacaagauugcacgcuggaacgugguggggcauccagggaucacugcuca
gcauuuucguggagcccauuuacuucucgagcaucauccugggcagccuuuaccacggggaccaccuuuccagggc
cauguaccagcggaucuccaacauagaggaccugccaccucucuacacccucaacaagccuuugcucaguggcauca
gcaaugcagaagcacggcagccagggaaggcccccaacuucagugucaacuggacgguaggcgacuccgcuauuga
ggucaucaacgccacgacugggaaggaugagcugggccgcgcgucccgccuguguaagcacgcguuguacugucgc
uggaugcgugugcacggcaagguucccucccacuuacuacgcuccaagauuaccaagcccaacguguaccaugagu
ccaagcuggcggcaaaggaguaccaggccgccaaggcgcgucuguucacagccuucaucaaggcggggcugggggc
cuggguggagaagcccaccgagcaggaccaguucucacucacgCCUGGAGGUGGCGGAUCGGGAG
GUGGCGGAUCGGGAGGUGGCGGAUCGGGCCGCAGCCGCCUUUUGGAAGAU
UUUCGAAACAACCGGUACCCCAAUUUACAACUGCGGGAGAUUGCCGGACA
UAUAAUGGAAUUUUCCCAAGACCAGCAUGGGUCCAGAUUCAUUCGCCUGA
AACUGGAGCGUGCCACACCAGCUGAGCGCCAGCUUGUCUUUAAUGAAAUC
CUCCAGGCUGCCUACCAACUCAUGGUGGAUGUGUUUGGUAGUUACGUCAU
UGAGAAGUUCUUUGAAUUUGGCAGUCUUGAACAGAAGCUGGCUUUGGCAG
AACGGAUUCGAGGUCACGUCCUGUCAUUGGCACUACAGAUGUAUGGCUGC
CGUGUUAUCCAGAAAGCUCUUGAGUUUAUUCCUUCAGACCAGCAGAAUGA
GAUGGUUCGGGAACUAGAUGGCCAUGUCUUGAAGUGUGUGAAAGAUCAGA
AUGGCUGUCACGUGGUUCAGAAAUGCAUUGAAUGUGUACAGCCCCAGUCU
UUGCAAUUUAUCAUCGAUGCGUUUAAGGGCCAGGUAUUUGCCUUAUCCAC
ACAUCCUUAUGGCUCCCGAGUGAUUGAGAGAAUCCUGGAGCACUGUCUCC
CUGACCAGACACUCCCUAUUUUAGAGGAGCUUCACCAGCACACAGAGCAGC
UUGUACAGGAUCAAUAUGGAUGUUAUGUAAUCCAACAUGUACUGGAGCAC
GGUCGUCCUGAGGAUAAAAGCAAAAUUGUAGCAGAAAUCCGAGGCAAUGU
ACUUGUAUUGAGUCAGCACAAAUUUGCAUGCAAUGUUGUGCAGAAGUGUG
UUACUCACGCCUCACGUACGGAGCGCGCUGUGCUCAUCGAUGAGGUGUGC
ACCAUGAACGACGGUCCCCACAGUGCCUUAUACACCAUGAUGAAGGACCAG
UAUGCCAGCUACGUGGUCCGCAAGAUGAUUGACGUGGCGGAGCCAGGCCA
GCGGAAGAUCGUCAUGCAUAAGAUCCGACCCCACAUCGCAACUCUUCGUAA
GUACACCUAUGGCAAGCACAUUCUGGCCAAGCUGGAGAAGUACUACAUGA
AGAACGGUGUUGACUUAGGGGGAGGUGGCGGAUCGGGAGGUGGCGGAUCG
GGAGGUGGCGGAUCGGGCCGCAGCCGCCUUUUGGAAGAUUUUCGAAACAA
CCGGUACCCCAAUUUACAACUGCGGGAGAUUGCCGGACAUAUAAUGGAAU
UUUCCCAAGACCAGCAUGGGAACAGAUUCAUUCAGCUGAAACUGGAGCGU
GCCACACCAGCUGAGCGCCAGCUUGUCUUUAAUGAAAUCCUCCAGGCUGCC
UACCAACUCAUGGUGGAUGUGUUUGGUUGUUACGUCAUUCAGAAGUUCUU

[0133] UGAAUUUGGCAGUCUUGAACAGAAGCUGGCUUUGGCAGAACGGAUUCGAG
GUCACGUCCUGUCAUUGGCACUACAGAUGUAUGGCUCCCGUGUUAUCGAG
AAAGCUCUUGAGUUUAUUCCUUCAGACCAGCAGAAUGAGAUGGUUCGGGA
ACUAGAUGGCCAUGUCUUGAAGUGUGUGAAAGAUCAGAAUGGCAAUCACG
UGGUUCAGAAAUGCAUUGAAUGUGUACAGCCCCAGUCUUUGCAAUUUAUC
AUCGAUGCGUUUAAGGGACAGGUAUUUGCCUUAUCCACACAUCCUUAUGG
CUGCCGAGUGAUUCAGAGAAUCCUGGAGCACUGUCUCCCUGACCAGACACU
CCCUAUUUUAGAGGAGCUUCACCAGCACACAGAGCAGCUUGUACAGGAUC
AAUAUGGAAGUUAUGUAAUCCGCCAUGUACUGGAGCACGGUCGUCCUGAG
GAUAAAAGCAAAAUUGUAGCAGAAAUCCGAGGCAAUGUACUUGUAUUGAG
UCAGCACAAAUUUGCAAACAAUGUUGUGCAGAAGUGUGUUACUCACGCCU
CACGUACGGAGCGCGCUGUGCUCAUCGAUGAGGUGUGCACCAUGAACGAC
GGUCCCCACAGUGCCUUAUACACCAUGAUGAAGGACCAGUAUGCCUGCUAC
GUGGUCCAGAAGAUGAUUGACGUGGCGGAGCCAGGCCAGCGGAAGAUCGU
CAUGCAUAAGAUCCGACCCCACAUCGCAACUCUUCGUAAGUACACCUAUGG
CAAGCACAUUCUGGCCAAGCUGGAGAAGUACUACAUGAAGAACGGUGUUG
ACUUAGGGAGCGGCGGCAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGG
CCAAGAAGAAGAAGuga
(SEQ ID NO: 8)

蛋白质序列：

[0134] MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAALHLDQTPSRQPI
PSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDV
KDAKVISVSTGGKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNK
DDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPNRK
ARGQLRTKIESGQGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQG
SLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEA
RQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGK
VPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQ
FSLTPGGGGSGGGGSGGGGSGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQH
GSRFIRLKLERATPAERQLVFNEILQAAYQLMVDVFGSYVIEKFFEFGSLEQKLAL
AERIRGHVLSLALQMYGCRVIQKALEFIPSDQQNEMVRELDGHVLKCVKDQNGC
HVVQKCIECVQPQSLQFIIDAFKGQVFALSTHPYGSRVIERILEHCLPDQTLPILEE
LHQHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFACNVV
QKCVTHASRTERAVLIDEVCTMNDGPHSALYTMMKDQYASYVVRKMIDVAEP
GQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYMKNGVDLGGGGGSGGGGSG
GGGSGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQHGNRFIQLKLERATPAER
QLVFNEILQAAYQLMVDVFGCYVIQKFFEFGSLEQKLALAERIRGHVLSLALQM
YGSRVIEKALEFIPSDQQNEMVRELDGHVLKCVKDQNGNHVVQKCIECVQPQSL
QFIIDAFKGQVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQHTEQLVQDQYG
SYVIRHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFANNVVQKCVTHASRTERAV
LIDEVCTMNDGPHSALYTMMKDQYACYVVQKMIDVAEPGQRKIVMHKIRPHIA
TLRKYTYGKHILAKLEKYYMKNGVDLGSGGKRPAATKKAGQAKKKK* (SEQ ID
NO: 17)

[0135] 靶向人SMN2(参考序列NM-022876)核苷酸序列(ACAGGGUUUUAGACAA(SEQ ID NO: 24))的核苷酸31,995-32,010的具有(G4S)3接头的双PUF设计

[0136] mRNA序列：

[0137]

augGGCCGCAGCCGCCUUUUGGAAGAUUUUCGAAACAACCGGUACCCCAAU
UUACAACUGCGGGAGAUUGCCGGACAUAUAAUGGAAUUUUCCCAAGACCA
GCAUGGGUCCAGAUUCAUUCAGCUGAAACUGGAGCGUGCCACACCAGCUG
AGCGCCAGCUUGUCUUUAAUGAAAUCCUCCAGGCUGCCUACCAACUCAUGG
UGGAUGUGUUUGGUUGUUACGUCAUUCAGAAGUUCUUUGAAUUUGGCAGU
CUUGAACAGAAGCUGGCUUUGGCAGAACGGAUUCGAGGUCACGUCCUGUC
AUUGGCACUACAGAUGUAUGGCUCCCGUGUUAUCCGCAAAGCUCUUGAGU
UUAUUCCUUCAGACCAGCAGAAUGAGAUGGUUCGGGAACUAGAUGGCCAU
GUCUUGAAGUGUGUGAAAGAUCAGAAUGGCUGUCACGUGGUUCAGAAAUG
CAUUGAAUGUGUACAGCCCCAGUCUUUGCAAUUUAUCAUCGAUGCGUUUA
AGGGCCAGGUAUUUGCCUUAUCCACACAUCCUUAUGGCUCCCGAGUGAUU
GAGAGAAUCCUGGAGCACUGUCUCCCUGACCAGACACUCCCUAUUUUAGA
GGAGCUUCACCAGCACACAGAGCAGCUUGUACAGGAUCAAUAUGGAUGUU
AUGUAAUCCAACAUGUACUGGAGCACGGUCGUCCUGAGGAUAAAAGCAAA
AUUGUAGCAGAAAUCCGAGGCAAUGUACUUGUAUUGAGUCAGCACAAAUU
UGCAAACAAUGUUGUGCAGAAGUGUGUUACUCACGCCUCACGUACGGAGC
GCGCUGUGCUCAUCGAUGAGGUGUGCACCAUGAACGACGGUCCCCACAGU
GCCUUAUACACCAUGAUGAAGGACCAGUAUGCCAACUACGUGGUCCAGAA
GAUGAUUGACGUGGCGGAGCCAGGCCAGCGGAAGAUCGUCAUGCAUAAGA
UCCGACCCCACAUCGCAACUCUUCGUAAGUACACCUAUGGCAAGCACAUUC
UGGCCAAGCUGGAGAAGUACUACAUGAAGAACGGUGUUGACUUAGGGGGA
GGUGGCGGAUCGGGAGGUGGCGGAUCGGGAGGUGGCGGAUCGGGCCGCAG
CCGCCUUUUGGAAGAUUUUCGAAACAACCGGUACCCCAAUUUACAACUGC
GGGAGAUUGCCGGACAUAUAAUGGAAUUUUCCCAAGACCAGCAUGGGAAC
AGAUUCAUUCAGCUGAAACUGGAGCGUGCCACACCAGCUGAGCGCCAGCU
UGUCUUUAAUGAAAUCCUCCAGGCUGCCUACCAACUCAUGGUGGAUGUGU
UUGGUAAUUACGUCAUUCAGAAGUUCUUUGAAUUUGGCAGUCUUGAACAG
AAGCUGGCUUUGGCAGAACGGAUUCGAGGUCACGUCCUGUCAUUGGCACU
ACAGAUGUAUGGCUCCCGUGUUAUCGAGAAAGCUCUUGAGUUUAUUCCUU
CAGACCAGCAGAAUGAGAUGGUUCGGGAACUAGAUGGCCAUGUCUUGAAG
UGUGUGAAAGAUCAGAAUGGCAGUCACGUGGUUGAGAAAUGCAUUGAAUG
UGUACAGCCCCAGUCUUUGCAAUUUAUCAUCGAUGCGUUUAAGGGACAGG
UAUUUGCCUUAUCCACACAUCCUUAUGGCUCCCGAGUGAUUGAGAGAAUC
CUGGAGCACUGUCUCCCUGACCAGACACUCCCUAUUUUAGAGGAGCUUCAC
CAGCACACAGAGCAGCUUGUACAGGAUCAAUAUGGAUGUUAUGUAAUCCA
ACAUGUACUGGAGCACGGUCGUCCUGAGGAUAAAAGCAAAAUUGUAGCAG
AAAUCCGAGGCAAUGUACUUGUAUUGAGUCAGCACAAAUUUGCAAGCUAU
GUUGUGCGCAAGUGUGUUACUCACGCCUCACGUACGGAGCGCGCUGUGCU
CAUCGAUGAGGUGUGCACCAUGAACGACGGUCCCCACAGUGCCUUAUACAC
CAUGAUGAAGGACCAGUAUGCCUGCUACGUGGUCCAGAAGAUGAUUGACG
UGGCGGAGCCAGGCCAGCGGAAGAUCGUCAUGCAUAAGAUCCGACCCCACA

UCGCAACUCUUCGUAAGUACACCUAUGGCAAGCACAUUCUGGCCAAGCUG
GAGAAGUACUACAUGAAGAACGGUGUUGACUUAGGGuga
(SEQ ID NO: 9)

[0138] 蛋白质序列：

[0139] MGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQHGSRFIQLKLERATPAERQLV FNEILQAAYQLMVDVFGCYVIQKFFEFGSLEQKLALAERIRGHVLSLALQMYGS RVIRKALEFIPSDQQNEMVRELDGHVLKCVKDQNGCHVVQKCIECVQPQSLQFII DAFKGQVFALSTHPYGSRVIERILEHCLPDQTLPILEELHQHTEQLVQDQYGCYVI QHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFANNVVQKCVTHASRTERAVLIDE VCTMNDGPHSALYTMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRPHIATLR KYTYGKHILAKLEKYYMKNGVDLGGGGGSGGGGSGGGGSGRSRLLEDFRNNR YPNLQLREIAGHIMEFSQDQHGNRFIQLKLERATPAERQLVFNEILQAAYQLMVD VFGNYVIQKFFEFGSLEQKLALAERIRGHVLSLALQMYGSRVIEKALEFIPSDQQN EMVRELDGHVLKCVKDQNGSHVVEKCIECVQPQSLQFIIDAFKGQVFALSTHPY GSRVIERILEHCLPDQTLPILEELHQHTEQLVQDQYGCYVIQHVLEHGRPEDKSKI VAEIRGNVLVLSQHKFASYVVRKCVTHASRTERAVLIDEVCTMNDGPHSALYT MMKDQYACYVVQKMIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKY YMKNGVDLG* (SEQ ID NO: 18)

与ADAR2DD融合的双PUF设计的融合多肽(PUF靶向人SMN2(参考序列NM-022876)核苷酸序列[ACAGGGUUUUAGACAA](SEQ ID NO:25)的核苷酸31,995-32,010)

[0140] mRNA序列：

[0141] AUGGACUAUAAGGACCACGACGGAGACUACAAGGAUCAUGAUAUUGAUUA CAAAGACGAUGACGAUAAGAUGGCCCCAAAGAAGAAGCGGAAGGUCGGUA UCCACGGAGUCCCAGCAGCCCUCCACCUCGACCAAACACCCAGCAGACAGC CUAUCCCUUCCGAAGGACUGcagcugcauuuaccgcagguuuuagcugacgcugucucacgccugg uccuggguaaguuuggugaucugaccgacaacuucuccucccccucacgcucgcagaaaagugcuggcuggagucgu caugacaacaggcacagauguuaaagaugccaaggugauaaguguuucuacaggaggcaaauguauuaauggugaa uacaugagugaucguggccuugcauuaaaaugacugccaugcagaaauaauaucucggagauccuugcucagauuuc uuuauacacaacuugagcuuuacuuaaaauaacaaagaugaucaaaaaagauccaucuuucagaaaucagagcgaggg ggguuuaggcugaaggagaauguccaguuucaucuguacaucagcaccucucccuguggagaugccagaaucuuc ucaccacaugagccaauccuggaagaaccagcagauagacacccaaaucguaaagcaagaggacagcuacggaccaaa auagagucuggucaggggacgauuccagugcgcuccaaugcgagcauccaaacgugggacggggugcugcaaggg gagcggcugcucaccauguccugcagugacaagauugcacgcuggaacgugguggcauccagggaucacugcuca gcauuuucguggagcccauuuacuucucgagcaucauccugggcagccuuuaccacggggaccaccuuuccagggc cauguaccagcggaucuccaacauagaggaccugccaccucucuacacccucaacaagccuuugcucaguggcauca gcaaugcagaagcacggcagccagggaaggcccccaacuucagugucaacuggacgguaggcgacuccgcuauuga ggucaucaacgccacgacugggaaggaugagcugggccgcgcgucccgccuguguaagcacgcguuguacugucgc uggaugcgugugcacggcaaggguccccucccacuuacuacgcuccaagauuaccaagcccaacguguaccaugagu ccaagcuggcggcaaaggaguaccaggccgccaaggcgcgucuguucacagccuucaucaaggcggggcuggggc cuggguggagaagcccaccgagcaggaccaguucucacucacgCCUGGAGGUGGCGGAUCGGGAG GUGGCGGAUCGGGAGGUGGCGGAUCGGGCCGCAGCCGCCUUUUGGAAGAU UUUCGAAACAACCGGUACCCCAAUUUACAACUGCGGGAGAUUGCCGGACA UAUAAUGGAAUUUUCCCAAGACCAGCAUGGGUCCAGAUUCAUUCAGCUGA AACUGGAGCGUGCCACACCAGCUGAGCGCCAGCUUGUCUUUAAUGAAAUC CUCCAGGCUGCCUACCAACUCAUGGUGGAUGUGUUUGGUUGUUACGUCAU UCAGAAGUUCUUUGAAUUUGGCAGUCUUGAACAGAAGCUGGCUUUGGCAG AACGGAUUCGAGGUCACGUCCUGUCAUUGGCACUACAGAUGUAUGGCUCC CGUGUUAUCCGCAAAGCUCUUGAGUUUAUUCCUUCAGACCAGCAGAAUGA

[0142]

GAUGGUUCGGGAACUAGAUGGCCAUGUCUUGAAGUGUGUGAAAGAUCAGA
AUGGCUGUCACGUGGUUCAGAAAUGCAUUGAAUGUGUACAGCCCCAGUCU
UUGCAAUUUAUCAUCGAUGCGUUUAAGGGCCAGGUAUUUGCCUUAUCCAC
ACAUCCUUAUGGCUCCCGAGUGAUUGAGAGAAUCCUGGAGCACUGUCUCC
CUGACCAGACACUCCCUAUUUUAGAGGAGCUUCACCAGCACACAGAGCAGC
UUGUACAGGAUCAAUAUGGAUGUUAUGUAAUCCAACAUGUACUGGAGCAC
GGUCGUCCUGAGGAUAAAAGCAAAAUUGUAGCAGAAAUCCGAGGCAAUGU
ACUUGUAUUGAGUCAGCACAAAUUUGCAAACAAUGUUGUGCAGAAGUGUG
UUACUCACGCCUCACGUACGGAGCGCGCUGUGCUCAUCGAUGAGGUGUGC
ACCAUGAACGACGGUCCCCACAGUGCCUUAUACACCAUGAUGAAGGACCAG
UAUGCCAACUACGUGGUCCAGAAGAUGAUUGACGUGGCGGAGCCAGGCCA
GCGGAAGAUCGUCAUGCAUAAGAUCCGACCCCACAUCGCAACUCUUCGUAA
GUACACCUAUGGCAAGCACAUUCUGGCCAAGCUGGAGAAGUACUACAUGA
AGAACGGUGUUGACUUAGGGGGAGGUGGCGGAUCGGGAGGUGGCGGAUCG
GGAGGUGGCGGAUCGGGCCGCAGCCGCCUUUUGGAAGAUUUUCGAAACAA
CCGGUACCCCAAUUUACAACUGCGGGAGAUUGCCGGACAUAUAAUGGAAU
UUUCCCAAGACCAGCAUGGGAACAGAUUCAUUCAGCUGAAACUGGAGCGU
GCCACACCAGCUGAGCGCCAGCUUGUCUUUAAUGAAAUCCUCCAGGCUGCC
UACCAACUCAUGGUGGAUGUGUUUGGUAAUUACGUCAUUCAGAAGUUCUU
UGAAUUUGGCAGUCUUGAACAGAAGCUGGCUUUGGCAGAACGGAUUCGAG
GUCACGUCCUGUCAUUGGCACUACAGAUGUAUGGCUCCCGUGUUAUCGAG
AAAGCUCUUGAGUUUAUUCCUUCAGACCAGCAGAAUGAGAUGGUUCGGGA
ACUAGAUGGCCAUGUCUUGAAGUGUGUGAAAGAUCAGAAUGGCAGUCACG
UGGUUGAGAAAUGCAUUGAAUGUGUACAGCCCCAGUCUUUGCAAUUUAUC
AUCGAUGCGUUUAAGGGACAGGUAUUUGCCUUAUCCACACAUCCUUAUGG
CUCCCGAGUGAUUGAGAGAAUCCUGGAGCACUGUCUCCCUGACCAGACACU
CCCUAUUUUAGAGGAGCUUCACCAGCACACAGAGCAGCUUGUACAGGAUC
AAUAUGGAUGUUAUGUAAUCCAACAUGUACUGGAGCACGGUCGUCCUGAG
GAUAAAAGCAAAAUUGUAGCAGAAAUCCGAGGCAAUGUACUUGUAUUGAG
UCAGCACAAAUUUGCAAGCUAUGUUGUGCGCAAGUGUGUUACUCACGCCU
CACGUACGGAGCGCGCUGUGCUCAUCGAUGAGGUGUGCACCAUGAACGAC
GGUCCCCACAGUGCCUUAUACACCAUGAUGAAGGACCAGUAUGCCUGCUAC
GUGGUCCAGAAGAUGAUUGACGUGGCGGAGCCAGGCCAGCGGAAGAUCGU
CAUGCAUAAGAUCCGACCCCACAUCGCAACUCUUCGUAAGUACACCUAUGG
CAAGCACAUUCUGGCCAAGCUGGAGAAGUACUACAUGAAGAACGGUGUUG
ACUUAGGGAGCGGCGGCAAGCGGCCCGCCGCCACCAAGAAGGCCGGCCAGG
CCAAGAAGAAGAAGuga
(SEQ ID NO: 10)

[0143] 蛋白质序列：

[0144]

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAALHLDQTPSRQPI
PSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGVVMTTGTDV
KDAKVISVSTGGKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYLNNK
DDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPNRK
ARGQLRTKIESGQGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQG
SLLSIFVEPIYFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEA
RQPGKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGK

[0145] VPSHLLRSKITKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQ FSLTPGGGGSGGGGSGGGGSGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQH GSRFIQLKLERATPAERQLVFNEILQAAYQLMVDVFGCYVIQKFFEFGSLEQKLA LAERIRGHVLSLALQMYGSRVIRKALEFIPSDQQNEMVRELDGHVLKCVKDQNG CHVVQKCIECVQPQSLQFIIDAFKGQVFALSTHPYGSRVIERILEHCLPDQTLPILE ELHQHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFANNV VQKCVTHASRTERAVLIDEVCTMNDGPHSALYTMMKDQYANYVVQKMIDVAE PGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYMKNGVDLGGGGGSGGGGS GGGGSGRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQHGNRFIQLKLERATPAE RQLVFNEILQAAYQLMVDVFGNYVIQKFFEFGSLEQKLALAERIRGHVLSLALQ MYGSRVIEKALEFIPSDQQNEMVRELDGHVLKCVKDQNGSHVVEKCIECVQPQS LQFIIDAFKGQVFALSTHPYGSRVIERILEHCLPDQTLPILEELHQHTEQLVQDQYG CYVIQHVLEHGRPEDKSKIVAEIRGNVLVLSQHKFASYVVRKCVTHASRTERAV LIDEVCTMNDGPHSALYTMMKDQYACYVVQKMIDVAEPGQRKIVMHKIRPHIA TLRKYTYGKHILAKLEKYYMKNGVDLGSGGKRPAATKKAGQAKKKK* (SEQ ID NO: 19)

[0146] 剪接调节子hTRA2-β1

[0147] mRNA序列(PUF插入位点(例如,缺失位点)位于带下划线的粗体区域[BbsI盒]):

[0148] AUGGACUACAAGGACCACGAUGGAGAUUAUAAAGACCACGACAUCGACUA UAAGGACGACGACGACAAGAUGAGCGACAGCGGCGAGCAGAACUACGGCG AGAGAGAGUCCAGAAGCGCCAGCAGAUCCGGCUCCGCUCACGGAAGCGGA AAGAGCGCUAGACAUACCCCCGCCAGAAGCAGAUCCAAGGAGGAUUCUAG AAGGAGCAGAAGCAAGAGCAGAUCUAGAAGCGAAUCUAGAUCCAGAUCUA GAAGAAGCUCUAGAAGGCACUACACAAGGUCUAGAAGCAGAUCUAGAAGC CAUAGAAGAAGCAGAUCCAGAAGCUACUCUAGAGACUACAGAAGGAGACA CAGCCACUCCCACAGCCCUAUGUCCACAAGAAGAAGGCACGUGGGCAAUAG GGCCAACCCCGACCCUAACCCCAAGAAGAAGAGGAAGGUGGGC**UCCGGCG UCUUCggcGAAGACGGCAGC**GGCCCUAAGAAGAAGAGGAAGGUGGGCAGC AGCAGCAUCACCAAGAGACCCCACACCCCUACCCCCGGCAUCUACAUGGGC AGACCCACCUACGGCUCCUCUAGAAGGAGAGACUACUACGACAGAGGCUAC GAUAGAGGCUACGACGAUAGAGAUUAUUACUCUAGAUCCUACAGAGGCGG CGGAGGAGGCGGAGGCGGAUGGAGAGCUGCCCAAGACAGAGACCAGAUCU AUAGAAGAAGGAGCCCCAGCCCCUACUAUAGCAGAGGCGGCUACAGAUCU AGAUCUAGAUCUAGAAGCUAUAGCCCCAGAAGAUACGGCGGCAGCUACCC UUACGACGUGCCCGACUACGCCUGA (SEQ ID NO: 11)

[0149] 蛋白质序列(PUF插入于带下划线的粗体区域[BbsI盒]):

[0150] MDYKDHDGDYKDHDIDYKDDDDKMSDSGEQNYGERESRSASRSGSAHGSGKS ARHTPARSRSKEDSRRSRSKSRSRSESRSRSRRSSRRHYTRSRSRSRSHRRSRSRS YSRDYRRRHSHSHSPMSTRRRHVGNRANPDPNPKKKRKVG**SGVFGEDGS**GPKK KRKVGSSSITKRPHTPTPGIYMGRPTYGSSRRRDYYDRGYDRGYDDRDYYSRSY RGGGGGGGGWRAAQDRDQIYRRRSPSPYYSRGGYRSRSRSRSYSPRRYGGSYPY DVPDYA* (SEQ ID NO: 20)

[0151] 具有双PUF设计的hTRA2-β1的融合多肽

[0152] mRNA序列:

AUGGACUACAAGGACCACGAUGGAGAUUAUAAAGACCACGACAUCGACUA
UAAGGACGACGACGACAAGAUGAGCGACAGCGGCGAGCAGAACUACGGCG
AGAGAGAGUCCAGAAGCGCCAGCAGAUCCGGCUCCGCUCACGGAAGCGGA
AAGAGCGCUAGACAUACCCCCGCCAGAAGCAGAUCCAAGGAGGAUUCUAG
AAGGAGCAGAAGCAAGAGCAGAUCUAGAAGCGAAUCUAGAUCCAGAUCUA
GAAGAAGCUCUAGAAGGCACUACACAAGGUCUAGAAGCAGAUCUAGAAGC
CAUAGAAGAAGCAGAUCCAGAAGCUACUCUAGAGACUACAGAAGGAGACA
CAGCCACUCCCACAGCCCUAUGUCCACAAGAAGAAGGCACGUGGGCAAUAG
GGCCAACCCCGACCCUAACCCCAAGAAGAAGAGGAAGGUGGGCGGAGGUG
GCGGAUCGggcaggagcaggcuuuuggaagauuuucgaaacaaccgCuaccccaauuuacaacugcgggaga
uugcuggacauauaauggaauuuucccaagaccagcauggguccagauucauucagcugaaacuggagcgugccac
accagcugagcgccagcuugucuucaaugaaauccuccaggcugccuaccaacucaugguggauguguuugguaau
uacgucauucagaaguucuuugaauuuggcagucuugaacagaagcuggcuuuggcagaacggauucgaggccac
guccugucauuggcacuacagauguauggcugccguguuauccagaaagcucuugaguuuauuccuucagaccag
cagaaugagaugguucgggaacuagauggccaugucuugaagugugugaaagaucagaauggcaaucacgugguu
cagaaaugcauugaaugugguacagccccagucuuugcaauuuaucaucgaugcguuuaagggacagguauuugcc
uuauccacacauccuuauggcugccgagugauucagagaauccuggagcacugucucccugaccagacacucccua
uuuuagaggagcuucaccagcacacagagcagcuuguacaggaucaauauggaaauuauguaauccaacauguacu
ggagcacggucguccugaggauaaaagcaaaauuguagcagaaauccgaggcaauguacuuguauugagucagcac
aaauuugcaagcaauguuguggagaagugugguuacucacgccucacguacggagcgcgcugugcucaucgaugag
gugugcaccaugaacgacggucccacagugccuuauacaccaugaugaaggaccaguaugccaacuacguggucc
agaagaugauugacguggcggagccaggccagcggaagaucgucaugcauaagauccggccccacaucgcaacucu
ucguaaguacaccuauggcaagcacauucuggccaagcuggagaaguacuacaugaagaacgguguugacuuaggg
[0153] GGAGGUGGCGGAUCGGGAGGUGGCGGAUCGGGAGGUGGCGGAUCGggcaggag
caggcuuuuggaagauuuucgaaacaaccgCuaccccaauuuacaacugcgggagauugcuggacauauaauggaa
uuuuccccaagaccagcauggguccagauucauucagcugaaacuggagcgugccacaccagcugagcgccagcuug
ucuucaaugaaauccuccaggcugccuaccaacucaugguggauguguuugguaauuacgucauucagaaguucu
uugaauuuggcagucuugaacagaagcuggcuuuggcagaacggauucgaggccacguccugucauuggcacuac
agauguauggcugccguguuauccagaaagcucuugaguuuauuccuucagaccagcagaaugagaugguucggg
aacuagauggccaugucuugaagugugugaaagaucagaauggcaaucacgugguucagaaaugcauugaaugug
uacagccccagucuuugcaauuuaucaucgaugcguuuaagggacagguauuugccuuauccacacauccuuaugg
cugccgagugauucagagaauccuggagcacugucucccugaccagacacucccuauuuuagaggagcuucaccag
cacacagagcagcuuguacaggaucaauauggaaauuauguaauccaacauguacuggagcacggucguccugagg
auaaaagcaaaauuguagcagaaauccgaggcaauguacuuguauugagucagcacaaauuugcaagcaauguugu
ggagaaguguguuacucacgccucacguacggagcgcgcugugcucaucgaugaggugugcaccaugaacgacgg
uccccacagugccuuauacaccaugaugaaggaccaguaugccaacuacguggguccagaagaugauugacguggcg
gagccaggccagcggaagaucgucaugcauaagauccggccccacaucgcaacucuucguaaguacaccuauggcaa
gcacauucuggccaagcuggagaaguacuacaugaagaacgguguugacuuaggGAGCGGCGGCGGCCC
UAAGAAGAAGAGGAAGGUGGGCAGCAGCAGCAUCACCAAGAGACCCCACA
CCCCUACCCCCGGCAUCUACAUGGGCAGACCCACCUACGGCUCCUCUAGAA
GGAGAGACUACUACGACAGAGGCUACGAUAGAGGCUACGACGAUAGAGAU
UAUUACUCUAGAUCCUACAGAGGCGGCGGAGGAGGCGGAGGCGGAUGGAG
AGCUGCCCAAGACAGAGACCAGAUCUAUAGAAGAAGGAGCCCCAGCCCCUA
CUAUAGCAGAGGCGGCUACAGAUCUAGAUCUAGAUCUAGAAGCUAUAGCC
CCAGAAGAUACGGCGGCAGCUACCCUUACGACGUGCCCGACUACGCCUGA
(SEQ ID NO: 12)

[0154] 蛋白质序列：

[0155] MDYKDHDGDYKDHDIDYKDDDDKMSDSGEQNYGERESRSASRSGSAHGSGKS ARHTPARSRSKEDSRRSRSKSRSRSESRSRSRRSSRRHYTRSRSRSRSHRRSRSRS YSRDYRRRHSHSHSPMSTRRRHVGNRANPDPNPKKKRKVGGGGGSGRSRLLED FRNNRYPNLQLREIAGHIMEFSQDQHGSRFIQLKLERATPAERQLVFNEILQAAY QLMVDVFGNYVIQKFFEFGSLEQKLALAERIRGHVLSLALQMYGCRVIQKALEFI PSDQQNEMVRELDGHVLKCVKDQNGNHVVQKCIECVQPQSLQFIIDAFKGQVFA LSTHPYGCRVIQRILEHCLPDQTLPILEELHQHTEQLVQDQYGNYVIQHVLEHGRP EDKSKIVAEIRGNVLVLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPH SALYTMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILA KLEKYYMKNGVDLGGGGGSGGGGSGGGGSGRSRLLEDFRNNRYPNLQLREIAG HIMEFSQDQHGSRFIQLKLERATPAERQLVFNEILQAAYQLMVDVFGNYVIQKFF EFGSLEQKLALAERIRGHVLSLALQMYGCRVIQKALEFIPSDQQNEMVRELDGH VLKCVKDQNGNHVVQKCIECVQPQSLQFIIDAFKGQVFALSTHPYGCRVIQRILE HCLPDQTLPILEELHQHTEQLVQDQYGNYVIQHVLEHGRPEDKSKIVAEIRGNVL VLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPHSALYTMMKDQYAN YVVQKMIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYMKNGVDL GSGGGPKKKRKVGSSSITKRPHTPTPGIYMGRPTYGSSRRRDYYDRGYDRGYDD RDYYSRSYRGGGGGGGGWRAAQDRDQIYRRRSPSPYYSRGGYRSRSRSRSYSPR RYGGSYPYDVPDYA* (SEQ ID NO: 21)

RNA效应子

[0156] 在一些实施例中，本文所述的多肽包含RNA效应子。

[0157] 在一些实施例中，该RNA效应子不包含核酸酶(例如，内切核酸酶和/或外切核酸酶)活性。在一些实施例中，该RNA效应子不包含核酸酶(例如，内切核酸酶和/或外切核酸酶)。在一些实施例中，该RNA效应子不包含核酸酶或其功能片段。在一些实施例中，RNA效应子不破坏磷酸二酯键。

[0158] RNA效应子的另一个实例是剪接调节子，例如剪接因子。剪接调节子可以包括募集细胞剪接机器的一种或多种组分的试剂。剪接调节子还可以包含或抑制或阻断细胞剪接机器的一种或多种组分(例如，与靶RNA序列或包含该靶RNA序列的RNA)结合的试剂。在一些实施例中，剪接因子包含天然存在的细胞剪接机器的组分或其功能片段或变体。在一些实施例中，剪接因子包含细胞剪接机器的重组和/或合成组分或其功能片段或变体。在一些实施例中，剪接调节子(例如，剪接因子)包含Sam68、hnRNP G、SRSF1、hnRNP A1/A2、TDP-43、SRp-30c、PSF、或hnRNP M。

[0159] 在一些实施例中，该RNA效应子包含RNA编辑结构域，例如如下所述。

[0160] RNA编辑结构域

[0161] 本文所述的某些多肽包括RNA编辑结构域。

[0162] 在一些实施例中，RNA编辑结构域在RNA中产生取代。

[0163] 在一些实施例中，RNA编辑结构域在RNA中产生插入或缺失。在一些实施例中，该RNA编辑结构域在RNA中产生少于5、4、3、2或1个核苷酸的插入。在一些实施例中，该RNA编辑结构域在RNA中产生少于5、4、3、2或1个核苷酸的缺失。在一些实施例中，该RNA编辑结构域：(a)破坏磷酸二酯键，从而产生RNA的第一部分和RNA的第二部分，(b)任选地从该第一部分或第二部分添加或去除核苷酸，以及(c)将该第一部分与该第二部分重新连接。在一些实施例中，这种RNA编辑在RNA中产生一个或多个核苷酸的插入、缺失或取代。产生插入和缺失的RNA编辑描述于例如Benne“RNA editing in trypanosomes[锥虫中的RNA编辑]”European Journal of Biochemistry[欧洲生物化学杂志]221:1(1994)，第9-23页，将其通过引用以

其全文并入本文。

[0164] 在一些实施例中，该RNA编辑结构域包含编辑靶RNA序列的一个或多个碱基的酶的、其功能片段或变体(例如，胞苷或腺苷脱氨酶功能片段或变体)的催化结构域。RNA编辑结构域可以是包含RNA脱氨酶(例如，腺苷脱氨酶、胞苷脱氨酶)的催化结构域的多肽序列。例如，RNA编辑结构域是以下项的催化结构域：作用于RNA的腺苷脱氨酶(ADAR)(例如，人ADAR 1、人ADAR2、人ADAR3、或人ADAR4)；作用于tRNA的腺苷脱氨酶(ADAT)、作用于RNA胞嘧啶脱氨酶(CDAR)。在实施例中，脱氨酶的催化结构域包含与表B中鉴定的具有GenBank登录号的序列至少80％相同的(例如，至少85％、87％、90％、92％、95％、98％、99％、100％相同的)序列。在实施例中，脱氨酶的催化结构域包含与表B中所示的催化核心结构域序列至少80％相同的(例如，至少85％、87％、90％、92％、95％、98％、99％、100％相同的)序列。

[0165] 表B.胞苷和腺苷脱氨酶的催化核心结构域(Maas和Rich,BioEssays[生物学分析]22:790-802(2000)John Wiley&Sons,Inc.[约翰·威利父子公司])

[0166]

基因	物种	GenBank登录号	用clustal W1.8进行的序列比对(催化核心结构域)
APOBEC1	小鼠	U22262	SNHVEVNFLEKFTTERY-FRPTWFLSWSPCGECSR
APOBEC1	人	L26234	TNHVEVNFIKKFTSERD-FHPTWFLSWSPCWECSQ
APOBEC1	兔	OCU10695	TNHVEVNFLEKLTSEGR-LGPTWFLSWSPCWECSM
APOBEC2	人	AF161698	AAHAEEAFFNTILPAFD---PTWYVSSSPCAACAD
AID	小鼠	AF132979	GCHVELLFLRYISD-WD-LDPTWFTSWSPCYDCAR
ADAR1	人	HSU10439	DCHAEIISRRGFIRFLY-SELHLYISTAPCGDGAL
ADAR1	非洲爪蟾(X.laevis)	XLU88065	DCHAEVVSRRGFIRFLY-SQLHLYISTAPCGDGAL
ADAR2	人	HSU76420	DCHAEIISRRSLLRFLY-TQLHLYISTSPCGDARI
ADAR2	大鼠	RNU43534	DCHAEIISRRSLLRFLY-AQLHLYISTSPCGDARI

[0167]

RED2	大鼠	RNU74586	DCHAEIVARRAFLHFLY-TQLHLYVSTSPCGDARL
ADAR	秀丽隐杆线虫	AF051275	DCHAEILARRGLLRFLY-SEVHLFINTAPCGVARI
ADAR	黑腹果蝇	DMBN35H14a	DSHAEIVSRRCLLKYLY-AQFHLYINTAPCGDARI
dCMP/CMP	人	L12136	VCHAELNAIMN-KNSTDVKGCSMYVALFPCNECAK
dCMP/CMP	酵母	YSCDCD1	CLHAEENALLEAGRDRVGQNATLYCDTCPCLTCSV
CDA	人	L27943	GICAERTAIQKAVSEGY-KDFMQDDFISPCGACRQ
CDA	大肠杆菌	ECCDD	TVHAEQSAISHAWLSGE-KALAITVNYTPCGHCRQ
hypCDA	秀丽隐杆线虫	P30648	VVHAEMNAIIN-KRCTTLHDCTVYVTLFPCNKCAQ
hypCDA	大肠杆菌	P30134	TAHAEIMALRQGGLVMQ-NYRTLYVTLEPCVMCAG
hypCDA	流感嗜血杆菌	P44931	TAHAEIIALRNGAKNIQ-NYRTLYVTLEPCTMCAG
ADAT1	人	AF125188	DSHAEVIARRSFQRYLL-HQLVFFSSHTPCGDASI
ADAT1	酵母	SCE7297	DCHAEILALRGANTVLL-NRIALYISRLPCGDASM
ADAT1	黑腹果蝇	AF192530	DSHAEVLARRGFLRFLY-QELHFLSTQTPCGDACI
ADAT2	人	AL031320.6a	TRHAEMVAIDQVLDWCRQSGTVLYVTVEPCIMCAA
ADAT2	酵母	SCE242667	VAHAEFMGIDQIKAMLG-SRGTLYVTVEPCIMCAS
ADAT3	人	AC012615.1a	LLHAVMVCVDLVARGQGRGGYDLYVTREPCAMCAM
ADAT3	酵母	SCE242668	IDHSVMVGIRAVGERLR-EGVDVYLTHEPCSMCSM

[0168] 在一些实施例中，RNA编辑结构域包含靶向单链RNA(ssRNA)的脱氨酶。在一些实施例中，RNA编辑结构域包含靶向双链RNA(dsRNA)的脱氨酶。不希望受理论束缚，尽管mRNA典型地是单链RNA，但是mRNA可以包含形成dsRNA的二级结构元件，该dsRNA可以由靶向dsRNA的脱氨酶编辑。在一些实施例中，本文所述的组合物可以进一步包含与靶RNA序列具有互补性并且能够与靶RNA序列杂交的核酸(例如，反义寡核苷酸)。不希望受理论束缚，由与靶RNA

序列具有互补性的核酸(例如反义寡核苷酸)形成的dsRNA和靶RNA序列可以使靶RNA序列被靶向dsRNA的脱氨酶靶向,例如,在不存在形成dsRNA的mRNA二级结构的情况下。在一些实施例中,与靶RNA序列具有互补性的核酸包含DNA。在一些实施例中,与靶RNA序列具有互补性的核酸包含RNA。在一些实施例中,与靶RNA序列具有互补性的核酸包含一个或多个经修饰的或合成的核苷酸。

[0169] 与靶RNA序列具有互补性的示例性核酸(例如,反义寡核苷酸)例如GluA2 mRNA或SMN2 mRNA包括但不限于SEQ ID NO:26-29。

[0170] GLUA2的70nt靶向序列

[0171] mRNA序列(要修饰的残基呈粗体/下划线):

[0172] 5'-ggcuauggcaucgcaacaccuaaaggauccucauua**A**ggugggguggaauaguauaacaauaugcuaaaug-3'(SEQ ID NO: 26)

[0173] SMN2的66nt靶序列(PUF靶向呈粗体/下划线):

[0174] 5'-UUUUUUAACUUCCUUUAUUUUCCUU**ACAGGGUUUUAGACAA**AAUCAAAAAGAAGGAAGGUGCUCAC-3'(SEQ ID NO: 27)

[0175] GLUA2的反义寡核苷酸靶向序列

[0176] 5'-ACTATTCCACCCACC**G**TAATGAGGATCCTT-3'(SEQ ID NO: 28)

[0177] SMN2的反义寡核苷酸靶向序列

[0178] 5'-TCACTTTCATAATGCTGG-3'(SEQ ID NO:29)

[0179] 示例性RNA编辑结构域包括但不限于SEQ ID NO:14-21的RNA编辑结构域或如由SEQ ID NO:5-12所编码的RNA编辑结构域。在一些实施例中,RNA编辑结构域包含与SEQ ID NO:14-21的RNA编辑结构域具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%同一性(或相对于其包含不超过1、2、3、4、5、6、7、8、9、或10个改变)的氨基酸序列,或由与SEQ ID NO:5-12的RNA编辑结构域编码序列具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、或99%同一性的核酸序列编码。

[0180] 接头

[0181] 在一些实施例中,本文所述的多肽可以包括一个或多个接头。在一些实施例中,RNA结合结构域中的RNA碱基结合基序通过接头连接。在一些实施例中,RNBA结合结构域和RNA编辑结构域之间具有接头。接头可以是化学键,例如一个或多个共价键或非共价键。在一些实施例中,连接是共价的。在一些实施例中,连接是非共价的。在一些实施例中,接头是肽接头。这样的接头可以在2-30个氨基酸之间,或更长。在一些实施例中,使用接头,例如以提供二级和三级结构的分子柔性,或允许分开的结构域或基序起作用(例如,与靶标结合),同时使位阻最小化。接头可以包括本文所述的柔性的、刚性的和/或可切割的接头。在一些实施例中,接头包括至少一个甘氨酸、丙氨酸、和丝氨酸氨基酸以提供柔性。在一些实施例中,接头是疏水接头,例如包括带负电荷的磺酸盐基团、聚乙二醇(PEG)基团或焦磷酸二酯基团。在一些实施例中,接头是可切割的以选择性地从另一个释放部分(例如结构域),但是足够稳定以防止过早切割。

[0182] 常用的柔性接头具有的序列主要由Gly和Ser残基("GS"接头)段组成。柔性接头可以有用于连接需要一定程度的移动或相互作用的结构域,并且可以包括小的、非极性的(例

如Gly)或极性的(例如Ser或Thr)氨基酸。Ser或Thr的掺入还可以通过与水分子形成氢键来维持接头在水溶液中的稳定性,且因此减少了接头与蛋白部分之间的不利相互作用。

[0183] 刚性接头有用于保持各结构域之间的固定距离并维持它们的独立功能。当结构域的空间分离对于保持融合中一种或多种组分的稳定性或生物活性至关重要时,刚性接头也可以是有用的。刚性接头可以具有α螺旋结构或富含脯氨酸的序列(Pro-rich序列)、(XP)n,其中X表示任何氨基酸,优选Ala、Lys或Glu。

[0184] 可切割接头可以在体内释放游离的功能性结构域。在一些实施例中,接头可以在特异性条件下(例如在还原剂或蛋白酶的存在下)切割。体内可切割接头可利用二硫键的可逆性质。一个实例包括两个Cys残基之间的凝血酶敏感性序列(例如,PRS)。CPRSC的体外凝血酶处理导致凝血酶敏感性序列的切割,而可逆的二硫键保持完整。此类接头是已知的并且描述于,例如,Chen等人,2013.Fusion Protein Linkers:Property,Design and Functionality[融合蛋白接头:性质、设计和功能]Adv Drug Deliv Rev.[先进药物输送评论]65(10):1357-1369。融合蛋白中接头的体内切割也可以通过蛋白酶进行,该蛋白酶在某些条件下在体内、在特定细胞或组织中、或在受限的某些细胞区室内表达。许多蛋白酶的特异性在受限的区室中提供了对接头的缓慢裂切割。

[0185] 连接分子的实例包括疏水接头,如带负电荷的磺酸根基团;脂质,如聚(--CH2--)烃链,如聚乙二醇(PEG)基团,其不饱和变体,其羟基化变体,其酰胺化或其他含N的变体,非碳接头;碳水化合物接头;磷酸二酯接头,或能够共价连接破坏剂的两个或多个组分(例如两个多肽)的其他分子。也可以包括非共价接头(如多肽与之连接的疏水性脂质小球)例如通过多肽的疏水区域或多肽的疏水延伸,如富含亮氨酸、异亮氨酸、缬氨酸、或者也可能是丙氨酸、苯丙氨酸、或甚至酪氨酸、甲硫氨酸、甘氨酸的一系列残基或其他疏水残基。破坏剂的组分可以使用基于电荷的化学连接,使得破坏剂的带正电荷的组分与另一组分或核酸的带负电荷的部分连接。

[0186] 制备组合物的方法

[0187] 制备重组蛋白质或多肽(例如,本文所述的多肽)的方法在本领域中是常规的。通常,参见Smales和James(编辑),Therapeutic Proteins:Methods and Protocols[治疗性蛋白:方法和方案](Methods in Molecular Biology[分子生物学方法]),Humana Press[胡玛纳出版社](2005);以及Crommelin,Sindelar和Meibohm(编辑),Pharmaceutical Biotechnology:Fundamentals and Applications[药物生物技术:基础与应用],Springer[斯普林格出版社](2013)。

[0188] 本披露组合物的蛋白质或多肽可以通过使用标准固相技术来生化合成。这类方法包括排阻固相合成、部分固相合成方法、片段缩合、经典溶液合成。这些方法可以当肽相对较短(即,10kDa)和/或当它不能通过重组技术来产生(例如,未由核酸序列来编码)并且因此涉及不同化学时使用。固相合成程序是本领域众所周知的,并且进一步由以下描述:John Morrow Stewart和Janis Dillaha Young,Solid Phase Peptide Syntheses[固相肽合成],第2版,Pierce Chemical Company[皮尔斯化学公司],1984;和Coin,I.等人,Nature Protocols[自然实验手册],2:3247-3256,2007。

[0189] 对于较长多肽,可以使用重组的方法。制备重组治疗性多肽的方法在本领域是常规的。通常,参见Smales和James(编辑),Therapeutic Proteins:Methods and Protocols

[治疗性蛋白:方法和方案](Methods in Molecular Biology[分子生物学方法]),Humana Press[胡玛纳出版社](2005);以及Crommelin,Sindelar和Meibohm(编辑),Pharmaceutical Biotechnology:Fundamentals and Applications[药物生物技术:基础与应用],Springer[斯普林格出版社](2013)。产生治疗性药物蛋白质或多肽的示例性方法涉及在哺乳动物细胞中表达,尽管也可以使用昆虫细胞、酵母、细菌、或其他细胞,在适当的启动子控制下,产生重组蛋白。哺乳动物表达载体可以包含非转录元件,如复制起点、合适的启动子、以及其他5'或3'侧翼非转录序列、以及5'或3'非翻译序列,如必要的核糖体结合位点、聚腺苷酸化位点、剪接供体和接受位点、以及终止序列。源自SV40病毒基因组的DNA序列,例如SV40起点、早期启动子、剪接和聚腺苷酸化位点可以用于提供异源DNA序列表达所需的其他遗传元件。在以下文献中描述了用于与细菌、真菌、酵母、和哺乳动物细胞宿主一起使用的适当的克隆和表达载体:Green&Sambrook,Molecular Cloning:A Laboratory Manual[分子克隆-实验室手册](第四版),Cold Spring Harbor Laboratory Press[冷泉港实验室出版社](2012)。

[0190] 在需要大量多肽的情况下,其可以使用例如由以下文献描述的技术来产生:Brian Bray,Nature Reviews Drug Discovery[自然综述:药物发现],2:587-593,2003;以及Weissbach&Weissbach,1988,Methods for Plant Molecular Biology,Academic Press[植物分子生物学方法],Academic Press[学术出版社],纽约,第VIII节,第421-463页。不同哺乳动物细胞培养系统可以用于表达和制造重组蛋白。哺乳动物表达系统的实例包括但不限于CHO细胞、COS细胞、HeLA和BHK细胞系。在以下文献中描述了用于生产蛋白治疗剂的宿主细胞培养的过程:Zhou和Kantardjieff(编辑),Mammalian Cell Cultures for Biologics Manufacturing[用于生物制品制造的哺乳动物细胞培养](Advances in Biochemical Engineering/Biotechnology[生物化学工程/生物科技的进展]),Springer[斯普林格出版社](2014)。本文所述的组合物可包括载体,例如编码重组蛋白的病毒载体,例如慢病毒载体。在一些实施例中,载体,例如病毒载体,可以包含编码重组蛋白的核酸。病毒和噬菌体表达载体是通过传统的遗传技术产生的。为了将基因转移至分裂和非分裂细胞,病毒表达载体可以包括慢病毒或腺病毒(AAV)。为了将基因转移至中枢神经系统(CNS),可以使用AAV载体或M13噬菌体载体。

[0191] 在以下文献中描述了蛋白治疗剂的纯化:Protein Biotechnology: solation, Characterization,and Stabilization[蛋白生物技术:分离、表征、和稳定化],Humana Press[胡玛纳出版社](2013);以及Cutler,Protein Purification Protocols[蛋白纯化方案](Methods in Molecular Biology[分子生物学方法]),Humana Press[胡玛纳出版社](2010)。

[0192] 可以将本文所述的核酸或编码本文所述的蛋白质的核酸掺入载体中。包括源自逆转录病毒如慢病毒的那些载体是实现长期基因转移的合适工具,因为它们允许转基因的长期稳定整合及其在子细胞中的繁殖。载体的实例包括表达载体、复制载体、探针产生载体、和测序载体。可以按病毒载体的形式,将表达载体提供给细胞。病毒载体技术是本领域众所周知的,并且描述于多种病理学和分子生物学手册。可用作载体的病毒包括但不限于逆转录病毒、腺病毒、腺相关病毒、疱疹病毒和慢病毒。通常,合适的载体含有在至少一种生物体中有功能的复制起点、启动子序列、方便的限制性内切酶位点、和一种或多种选择性标记。

[0193] 天然的或合成的核酸的表达典型地通过以下实现:将编码目的基因的核酸可操作地连接至启动子,并且将构建体并入表达载体。载体可以适用于在真核生物中复制和整合。典型的克隆载体含有转录和翻译终止子、起始序列、和启动子,可用于希望的核酸序列的表达。

[0194] 另外的启动子元件,例如增强序列,可以调节转录起始的频率。典型地,这些序列位于转录起始位点上游30-110bp的区域中,尽管最近已显示多种启动子也含有转录起始位点下游的功能元件。启动子元件之间的间隔通常是灵活的,使得当反转元件或使元件相对彼此移动时能够保留启动子功能。在胸苷激酶(tk)启动子中,在活性开始下降之前,启动子元件之间的间隔可以增加至50bp。取决于启动子,似乎单个元件可以共同地或独立地发挥功能,以激活转录。

[0195] 合适的启动子的一个实例是即时早期巨细胞病毒(CMV)启动子序列。这一启动子序列是能够驱动与其可操作地连接的任何多核苷酸序列的高水平表达的强组成型启动子序列。合适的启动子的一些实施例是延伸生长因子-1α(EF-1α)。然而,还可以使用其他组成型启动子序列,包括但不限于猿猴病毒40(SV40)早期启动子、小鼠乳腺瘤病毒(MMTV)、人类免疫缺陷病毒(HIV)、长末端重复序列(LTR)启动子、MoMuLV启动子、禽白血病病毒启动子、EB病毒即时早期启动子、劳斯肉瘤病毒启动子、连同人类基因启动子(诸如但不限于肌动蛋白启动子、肌球蛋白启动子、血红蛋白启动子、以及肌酸激酶启动子)。

[0196] 本披露不应解释为限于使用任何特定启动子或启动子种类(例如组成型启动子)。例如,在一些实施例中,诱导型启动子被认为是本披露的一部分。在一些实施例中,诱导型启动子的使用提供了能够开启与其可操作地连接的多核苷酸序列的表达的分子开关(当需要这样的表达时)。在一些实施例中,诱导型启动子的使用提供了能够关闭表达的分子开关(当不需要表达时)。诱导型启动子的实例包括但不限于金属硫蛋白启动子、糖皮质激素启动子、孕酮启动子、和四环素启动子。

[0197] 在一些实施例中,待引入的表达载体还可以含有选择性标记基因或报告基因或二者,从而便于从寻求通过病毒载体转染或感染的细胞的群体中,鉴定和选择表达性细胞。在一些方面,选择性标记物可以在一段单独的DNA上进行,并且用于共转染程序。选择性标记物和报告基因二者侧翼可以是适当的表达控制序列,以使得能够在宿主细胞中表达。有用的选择性标记物可以包括例如抗生素抗性基因,例如neo等。

[0198] 在一些实施例中,可以将报告基因用于鉴定潜在转染的细胞和/或用于评估表达控制序列的功能。通常,报告基因是不存在于(报告基因的)受者来源中或不由受者来源表达,并且编码其表达由一些易于检测特性(例如酶活性或可视荧光)所证明的多肽的基因。在将DNA引入受体细胞后,在适当的时间测定报告基因的表达。合适的报告基因可包括编码荧光素酶、β-半乳糖苷酶、氯霉素乙酰转移酶、分泌的碱性磷酸酶、或绿色荧光蛋白基因的基因(例如,Ui-Tei等人,2000FEBS Letters[欧洲生化学会联合会快报]479:79-82)。合适的表达系统是熟知的,可以使用已知技术制备或商业获得。通常,具有显示报告基因最高表达水平的最小5'侧翼区的构建体被鉴定为启动子。此类启动子区可以连接至报告基因,并且用于评估药剂调节启动子驱动的转录的能力。

[0199] 应用

[0200] RNA编辑组合物(例如,本文所述的多肽、核酸、载体和宿主细胞)可以例如通过纠

正细胞、组织或受试者中RNA中的功能丧失突变(例如,一个或多个点突变)来满足治疗需求。例如,RNA编辑组合物(例如,本文所述的多肽、核酸、载体和宿主细胞)可以用于治疗与基因中突变(例如,一个或多个点突变)相关的疾病。

[0201] 本文所述的组合物(例如,本文所述的多肽、核酸、载体和宿主细胞)可以用于治疗疾病或病症。在一些实施例中,该疾病选自梅尔戈林综合征、塞克尔综合征4、Joubert综合征5、莱伯氏先天性黑蒙10;腓骨肌萎缩疾病,2型;腓骨肌萎缩疾病,2型;乌谢尔综合征,2C型;脊髓小脑性共济失调28;脊髓小脑性共济失调28;脊髓小脑性共济失调28;长QT综合征2;格伦-拉松综合征(Sjogren-Larsson syndrome);遗传性果糖尿症;遗传性果糖尿症;神经母细胞瘤;神经母细胞瘤;卡尔曼综合征1;卡尔曼综合征1;卡尔曼综合征1;异染性脑白质营养不良、雷特综合征、10型肌萎缩性侧索硬化症、李-佛美尼综合征(Li-Fraumeni syndrome)、囊性纤维化病、赫尔综合征、α-1-抗胰蛋白酶(A1AT)缺乏、帕金森氏症、阿耳茨海默病、白化病、肌萎缩性侧索硬化症、哮喘、b-地中海贫血、Cadasil综合征、腓骨肌萎缩疾病、慢性阻塞性肺病(COPD)、远端脊髓性肌萎缩(DSMA)、杜氏/贝克肌营养不良、营养不良性大疱性表皮松解、大疱性表皮松懈症、法布瑞氏症、因子V Leiden相关障碍、家族性腺瘤、息肉病、半乳糖血症、戈谢病、葡萄糖-6-磷酸脱氢酶、血友病、遗传性血色素沉着症、亨特氏综合征、亨廷顿病、炎性肠病(I BD)、遗传性多凝集反应综合征、莱伯氏先天性黑蒙、莱施-奈恩综合征、林奇综合征、马凡综合征、黏多糖贮积症、肌肉萎缩症、I型和II型强直性肌营养不良、神经纤维瘤病、A型、B型和C型尼曼-皮克病、NY-eso1相关癌症、黑斑息肉综合征、苯酮尿症、庞培氏病、原发性睫状疾病、凝血酶原突变相关障碍(例如凝血酶原G20210A突变)、肺动脉高压、色素性视网膜炎、山德霍夫氏病、严重联合免疫缺陷综合征(SCID)、镰状细胞性贫血、脊髓性肌萎缩、斯特格氏病、泰-萨克斯病、乌谢尔综合征、X连锁免疫缺陷、斯特奇-韦伯综合征、以及癌症。在一些实施例中,本披露涉及本文所述的组合物(例如,本文所述的多肽、核酸、载体、或宿主细胞)在制造用于治疗或预防疾病或障碍(例如遗传障碍)的药物中的用途,该疾病或障碍选自本文列出的疾病或障碍。

[0202] 配制、施用和递送

[0203] 在各种实施例中,本披露提供了本文所述的多肽、核酸、载体和宿主细胞的药物组合物,与药学上可接受的赋形剂一起配制。药学上可接受的赋形剂包括可用于制备通常安全,无毒和所需的药物组合物的赋形剂,并且包括兽医用途以及人类药物用途可接受的赋形剂。此类赋形剂可以是水性的或非水性的。适当的赋形剂可有助于例如组合物的稳定性、溶解性、缓冲作用。以下文献中描述了蛋白治疗剂的配制品:Meyer(编辑),Therapeutic Protein Drug Products:Practical Approaches to formulation inthe Laboratory, Manufacturing,and the Clinic[治疗性蛋白药物产品:实验室、制造和临床中配制品的实践方法],Woodhead Publishing Series[伍德海德出版系列](2012)。

[0204] 根据本披露的药物组合物可以按治疗有效量递送。精确的治疗有效量是对受试者具有所希望的治疗作用的组合物(例如本文所述的多肽、核酸、载体和宿主细胞)的量。该量将根据多种因素而变化,包括但不限于治疗性化合物的特征(包括活性、药代动力学、药效学和生物利用度)、受试者的生理状况(包括年龄、性别、疾病类型和阶段、一般身体条件、对给定剂量的应答性和药物类型)、配制品中一种或多种药学上可接受的载剂的性质、和/或施用途径。对受试者的施用方式可包括全身、肠胃外、肠内或局部。

[0205] 在一些实施例中,可以使用载体将本文所述的多肽或核酸组合物递送至细胞、组织或受试者。载体可以是例如质粒或病毒。在一些实施例中,递送是体内、体外、离体或原位的。在一些实施例中,病毒是腺相关病毒(AAV)、慢病毒、腺病毒。在一些实施例中,将本文所述的多肽或核酸组合物与病毒样颗粒或病毒体一起递送至细胞。在一些实施例中,递送使用一种以上的病毒、病毒样颗粒或病毒体。

[0206] 脂质体配制品

[0207] 适用作媒剂或载剂以递送本文所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞的示例性配制品包括微乳剂、单层、胶束、双层、囊泡或脂质颗粒。这些配制品为本文所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞提供了生物相容的和可生物降解的递送系统。

[0208] 脂质体提供了脂质颗粒的实例,这些脂质颗粒由排列在一个或多个球形双层中的两亲脂质构成。脂质体是单层的或多层的囊泡,其具有一个形成自亲脂材料的膜以及一个水性内部。水性部分包含要递送的本文所述的多肽、药物组合物、核酸、载体、组合物、或宿主细胞。阳离子脂质体具有能够融合至细胞壁的优势。非阳离子脂质体尽管不能够有效地与细胞壁融合,但是可以被体内巨噬细胞摄取。

[0209] 脂质体具有几个优点,包括直径小;生物相容性和可生物降解性;结合多种内容物的能力,例如水和脂溶性药物。脂质体可以保护包封在其内部隔室中的药物免于代谢和降解(Rosoff,于Pharmaceutical Dosage Forms[药物剂型],Lieberman,Rieger和Banker(编辑),1988,Marcel Dekker,Inc.[马塞尔德克尔公司],纽约,纽约州,第1卷,第245页)。在制备脂质体配制品方面的重要的考虑是脂质表面电荷、囊泡尺寸以及这些脂质体的水性体积。

[0210] 脂质体有两个广泛的类别。阳离子脂质体是带正电荷的脂质体,其与带负电荷的DNA分子相互作用而形成一种稳定的复合体。该带正电荷的DNA/脂质体复合物结合至带负电荷的细胞表面并且内化在内涵体中。由于核内体内部的酸性pH,脂质体破裂,释放它们的内容物至细胞胞质中(Wang等人,Biochem.Biophys.Res.Commun.[生物化学与生物物理学研究通讯],1987,147,980-985)。

[0211] pH敏感的或带负电荷的脂质体捕获DNA而不是与其形成复合体。由于该DNA与该脂类是带类似的电荷,所以形成排斥而不是复合。虽然如此,一些DNA被捕获在这些脂质体的水性内部中。pH敏感的脂质体已经用于将编码胸苷激酶基因的DNA递送至培养中的细胞单层。在靶标细胞内检测到外源基因的表达(Zhou等人,Journal of Controlled Release[控释杂志],1992,19,269-274)。

[0212] 脂质体组合物的一个主要类型包括磷脂,除了天然衍生的磷脂酰胆碱。例如中性脂质体组合物可以形成自二肉豆蔻酰磷脂酰胆碱(DMPC)或二棕榈酰磷脂酰胆碱(DPPC)。阴离子脂质体组合物通常可以形成自二肉豆蔻酰磷脂酰甘油,而阴离子融合脂质体主要形成自二油酰基磷脂酰乙醇胺(DOPE)。另一类型的脂质体组合物形成自磷脂酰胆碱(PC),例如像大豆PC,蛋PC。另一个类型形成自磷脂和/或磷脂酰胆碱和/或胆固醇的混合物。

[0213] 适用于将药物递送至皮肤的示例性非离子脂质体系统包括包含非离子表面活性剂和胆固醇的系统。包括Novasome™ I(二月桂酸甘油酯/胆固醇/聚氧乙烯-10-硬脂酰醚)以及Novasome™ II(二硬脂酸甘油酯/胆固醇/聚氧乙烯-10-硬脂酰醚)的非离子型脂质体

配制品用于将环孢霉素A递送入小鼠皮肤的真皮。结果表明这类非离子型脂质体系统可有效促进环孢菌素A沉积至皮肤的不同层中(Hu等人,S.T.P.Pharma.Sci.[S.T.P.制药科学],1994,4,6,466)。

[0214] 脂质体可以在空间上稳定以包括一种或多种专门的脂质,当掺入脂质体中时相对于缺少此类专门脂质的脂质体导致循环寿命增加。立体化学稳定的脂质体的实例是以下那些:在其中该脂质体的形成囊泡的脂质部分中的一部分(A)包括一种或多种糖脂,例如单唾液酸神经节苷酯G_{M1},或者(B)是用一种或多种亲水聚合物衍生,例如聚乙二醇(PEG)部分(Allen等人,FEBS Letters[FEBS通讯],1987,223,42;Wu等人,Cancer Research[癌症研究],1993,53,3765)。也可以使用长循环脂质体,例如隐形脂质体。此类脂质体总体描述于美国专利号5,013,556。本文所披露的化合物可以通过控释手段和/或例如以下文献中描述的那些递送装置来施用:美国专利号3,845,770;3,916,899;3,536,809;3,598,123;和4,008,719。

[0215] 包括一种或多种糖脂的不同脂质体在本领域内是已知的。Papahadjopoulos等人(Ann.N.Y.Acad.Sci.[纽约科学院年报],1987,507,64)报道了单唾液酸神经节苷酯G_{M1}、硫酸半乳糖脑苷脂以及磷脂酰肌醇改进脂质体的血液半衰期的能力。这些发现由Gabizon等人(Proc.Natl.Acad.Sci.U.S.A.[美国科学院院报],1988,85,6949)阐述。均为Allen等人的美国专利号4,837,028和WO 88/04924披露了包含(1)鞘磷脂和(2)神经节苷脂G_{M1}或硫酸半乳糖脑苷脂的脂质体。美国专利号5,543,152(Webb等人)披露了包含鞘磷脂的脂质体。包含1,2-sn-二肉豆蔻磷脂酰胆碱的脂质体在WO97/13499(利姆(Lim)等人)中披露。

[0216] 包含脂质的脂质体可以用一种或多种亲水性聚合物衍生,并且其制备方法是本领域已知的。Sunamoto等人(Bull.Chem.Soc.Jpn.,1980,53,2778)描述包含含有PEG部分的非离子去垢剂$2C_{1215G}$的脂质体。Illum等人(FEBS Lett.[FEBS通讯],1984,167,79)报道了用聚合物二醇对聚苯乙烯粒子的亲水包覆导致显著增长的血液半寿期。通过接合聚二醇(例如,PEG)的羧基所修饰的合成磷脂由Sears描述(美国专利号4,426,330和4,534,899)。Klibanov等人(FEBS Lett.,1990,268,235)描述了展示以下的实验:包含用PEG或PEG硬脂酸酯衍生的磷脂酰乙醇胺(PE)的脂质体具有显著增加的血液循环半寿期。Blume等人(Biochimica et Biophysica Acta[生物化学与生物物理学报],1990,1029,91)将这类观察结果扩展至其他PEG衍生的磷脂,例如,由二硬脂酰磷脂酰乙醇胺(DSPE)和PEG的组合组成的DSPE-PEG。授予Fisher的欧洲专利号EP 0 445 131 B1和WO 90/04384中描述了在它们的外表面上具有共价结合的PEG部分的脂质体。含有1-20摩尔%用PEG衍生的PE的脂质体组合物及其使用方法由Woodle等人(美国专利号5,013,556和5,356,633)和Martin等人(美国专利号5,213,804和欧洲专利号EP 0 496 813 B1)描述。包含许多其他脂质缀合物的脂质体在WO 91/05545和美国专利号5,225,212(均授予Martin等人)中和在WO 94/20073(Zalipsky等人)中披露。包含PEG修饰的神经酰胺脂类的脂质体描述于WO 96/10391(Choi等人)中。美国专利号5,540,935(Miyazaki等人)以及美国专利号5,556,948(Tagawa等人)描述了包含PEG的脂质体,它们可以进一步用官能部分在其表面衍生。

[0217] 一些包括核酸的脂质体在本领域内是已知的。授予Thierry等人的WO 96/40062公开了在脂质体中包封高分子量核酸的方法。授予Tagawa等人的美国专利号5,264,221披露了蛋白质结合的脂质体。授予Rahman等人的美国专利号5,665,710描述在脂质体中包封寡

脱氧核苷酸的某些方法。

[0218] 表面活性剂在配制品(例如乳剂(包括微乳剂)以及脂质体)中有广泛应用。对许多不同类型的表面活性剂(天然的与合成的)的特性进行分类与评级的最普通的方法是通过亲水/亲油平衡值(HLB)的使用。亲水基团(也称为“头”)的性质为对配制品中使用的不同表面活性剂进行分类提供了最有用的方法。已经综述了表面活性物质在药品、配制品和在乳剂中的用途(Rieger,于Pharmaceutical Dosage Forms[药物剂型],Marcel Dekker,Inc.[马塞尔德克尔公司],纽约,纽约州,1988,285页)。

[0219] 递送媒剂的其他实例包括纳米结构化的脂质载剂(NLC),其是经修饰的固体脂质纳米颗粒(SLN),其保留了SLN的特性,提高了药物的稳定性和载药量,并防止了药物泄漏。聚合物纳米颗粒(PNP)是药物递送的重要组成部分。这些纳米颗粒可以有效地将药物递送引导至特定靶并改善药物稳定性和受控的药物释放。也可以使用脂质体纳米颗粒(PLN),组合脂质体和聚合物。这些纳米颗粒具有PNP和脂质体的互补优势。PLN由核-壳结构构成;聚合物核提供了稳定的结构,磷脂壳提供了良好的生物相容性。对于综述,参见例如Li等人2017,Nanomaterials[纳米材料]7,122;doi:10.3390/nano7060122。

[0220] 在一些实施例中,可以将本文所述的核酸、载体或组合物包封在脂质配制品中,例如形成核酸-脂质颗粒。核酸脂质颗粒典型地包含阳离子脂质、非阳离子脂质以及阻止颗粒聚集的脂质(例如,PEG-脂质缀合物)。这些颗粒对于合成应用是有用的,因为它们展示出在静脉内(i.v.)注射之后延长的循环寿命并且在远端位点积累(例如在与施用位点物理分开的位点)。颗粒包括包封的缩合剂-核酸复合物,如在PCT公开号WO 00/03683中所述的。这些颗粒典型地具有约50nm至约150nm,更典型地约60nm至约130nm,更典型地约70nm至约110nm,最典型地约70nm至约90nm的平均直径,并且基本上是无毒的。另外,当出现在本发明的核酸-脂质颗粒中时,这些核酸在水溶液中抵抗核酸酶的降解。核酸-脂质颗粒及其制备方法披露于例如美国专利号5,976,567;5,981,501;6,534,484;6,586,410;6,815,432;以及PCT公开号WO 96/40964。

[0221] 在一个实施例中,脂质与药物的比率(质量/质量比率)(例如脂质与dsRNA的比率)将处于从约1:1至约50:1、从约1:1至约25:1、从约3:1至约15:1、从约4:1至约10:1、从约5:1至约9:1、或约6:1至约9:1的范围内。

[0222] 阳离子脂质可以是例如,N,N-二油烯基-N,N-二甲基氯化铵(DODAC)、N,N-二硬脂基-N,N-二甲基溴化铵(DDAB)、N-(I-(2,3-二油酰基氧基)丙基)-N,N,N-三甲基氯化铵(DOTAP)、N-(I-(2,3-二油烯基氧基)丙基)-N,N,N-三甲基氯化铵(DOTMA)、N,N-二甲基-2,3-二油烯基氧基)丙胺(DODMA)、1,2-二亚油基氧基-N,N-二甲基氨基丙烷(DLinDMA)、1,2-二亚麻基氧基-N,N-二甲基氨基丙烷(DLenDMA)、1,2-二亚油基氨基甲酰基氧基-3-二甲基氨基丙烷(DLin-C-DAP)、1,2-二亚油基氧基-3-(二甲基氨基)乙酰氧基丙烷(DLin-DAC)、1,2-二亚油基氧基-3-吗啉代丙烷(DLin-MA)、1,2-二亚油基酰基-3-二甲基氨基丙烷(DLinDAP)、1,2-二亚油基硫代-3-二甲基氨基丙烷(DLin-S-DMA)、1-亚油基酰基-2-亚油基氧基-3-二甲基氨基丙烷(DLin-2-DMAP)、1,2-二亚油基氧基-3-三甲基氨基丙烷氯化盐(DLin-TMA.Cl)、1,2-二亚油基酰基-3-三甲基氨基丙烷氯化盐(DLin-TAP.Cl)、1,2-二亚油基氧基-3-(N-甲基哌嗪)丙烷(DLin-MPZ)、或3-(N,N-二亚油基氨基)-1,2-丙二醇(DLinAP)、3-(N,N-二油烯基氨基)-1,2-丙二醇(DOAP)、1,2-二亚油基氧代-3-(2-N,N-二甲

基氨基)乙氧基丙烷(DLin-EG-DMA)、1,2-二亚麻基氧基-N,N-二甲基氨基丙烷(DLinDMA)、2,2-二亚油基-4-二甲基氨基甲基-[1,3]-二氧戊环(DLin-K-DMA)或其类似物、(3aR,5s,6aS)-N,N-二甲基-2,2-二((9Z,12Z)-十八-9,12-二烯)四氢-3aH-环戊[d][1,3]二氧杂环戊烯-5-胺、(6Z,9Z,28Z,31Z)-三十七碳-6,9,28,31-四-19-基4-(二甲基氨基)丁酸酯(MC3)、1,1'-(2-(4-(2-((2-(双(2-羟基月桂基)氨基)乙基)(2-羟基月桂基)氨基)乙基)哌嗪-1-基)乙基脲二基)二十二烷-2-醇(Tech G1)、或其混合物。阳离子脂类可以占颗粒中存在的总脂质的从约20mol%至约50mol%或约40mol%。

[0223] 在一个实施例中,脂质颗粒包含40%2,2-二亚油基-4-二甲基氨基乙基-[1,3]-二氧戊环:10%DSPC:40%胆固醇:10%PEG-C-DOMG(摩尔百分数),其粒度为63.0±20nm和0.027siRNA/脂质比率。

[0224] 非阳离子脂类可以是一种阴离子脂类或一种中性脂类,包括但不限于:二硬脂酰磷酸卵磷酯(DSPC)、二油酰基磷酸卵磷酯(DOPC)、二棕榈酰磷酸卵磷酯(DPPC)、二油酰基磷脂酰甘油(DOPG)、二棕榈酰磷脂酰甘油(DPPG)、二油酰基-磷脂酰乙醇胺(DOPE)、棕榈酰油酰基磷酸卵磷酯(POPC)、棕榈酰油酰基磷脂酰乙醇胺(POPE)、二油酰基-磷脂酰乙醇胺4-(N-马来酰亚胺甲基)-环己烷-1-羧酸酯(DOPE-mal)、二棕榈酰磷脂酰基乙醇胺(DPPE)、二肉豆蔻酰磷酸乙醇胺(DMPE)、二硬脂酰-磷脂酰-乙醇胺(DSPE)、16-O-一甲基PE、16-O-二甲基PE、18-1-反式PE、1-硬脂酰-2-油酰基-磷脂酰乙醇胺(SOPE)、胆固醇,或其混合物。非阳离子脂类(如果包括胆固醇的话)可以占颗粒中存在的总脂质的从约5mol%至约90mol%、约10mol%、或约58mol%。

[0225] 抑制颗粒聚集的缀合脂类可以是例如一种聚乙二醇(PEG)-脂类,其包括但不限于PEG-PEG-(DAG)、PEG-二烷氧基丙基(DAA)、PEG-磷脂、PEG-神经酰胺(Cer),或其混合物。PEG-DAA缀合物可以是,例如PEG-二月桂基氧丙基(Ci_2)、PEG-二肉豆蔻基氧丙基(Ci_4)、PEG-二棕榈基氧丙基(Ci_6)或PEG-二硬脂基氧丙基$(C]_8$)。防止粒子聚集的缀合脂质可以占颗粒中存在的总脂质的从0mol%至约20mol%或约2mol%。

[0226] 在一些实施例中,核酸-脂质颗粒进一步包括胆固醇,该胆固醇例如占颗粒中存在的总脂质的约10mol%到约60mol%或约48mol%。

[0227] 在一个实施例中,该配制品是MC3,其包含描述于以下文献中的配制品,例如2010年6月10日提交的国际申请号PCT/US 10/28224,其通过引用特此并入。含有配制品的MC3的合成和结构描述于例如,WO 2013/155204的第114-119页,通过引用并入。在一些实施例中,MC3配制品包含DLin-M-C3-DMA(即,(6Z,9Z,28Z,31Z)-三十七碳-6,9,28,31-四-19-基4-(二甲基氨基)丁酸酯)的制剂。

[0228] 在一些实施例中,本文所述的多肽、核酸、载体或宿主细胞组合物可以配制在脂质体或其他类似囊泡中。脂质体是球形囊泡结构,所述球形囊泡结构由围绕内部水性隔室的单层或多层的脂质双层和相对不可渗透的外部亲脂性磷脂双层构成。脂质体可以是阴离子的、中性的或阳离子的。脂质体具有生物相容性,无毒,可以递送亲水性和亲脂性药物分子,保护其货物免受血浆酶的降解,并将其负载运输穿过生物膜和血脑屏障(BBB)(关于综述,参见,例如,Spuch和Navarro,Journal of Drug Delivery[药物递送杂志],第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。

[0229] 囊泡可由几种不同类型的脂质制成;然而,磷脂最常用于产生脂质体作为药物载

剂。囊泡可包括但不限于DOTMA、DOTAP、DOTIM、DDAB，单独使用或与胆固醇一起产生DOTMA和胆固醇、DOTAP和胆固醇、DOTIM和胆固醇以及DDAB和胆固醇。制备多层囊泡脂质的方法是本领域已知的(参见例如美国专利号6,693,086，其关于多层囊泡脂质制备的教导通过引用并入本文)。尽管当脂质膜与水溶液混合时，囊泡的形成是自发的，但也可以通过使用均质器、超声仪或挤压装置以振荡的形式施加力来加快囊泡的形成(关于综述，参见例如，Spuch和Navarro,Journal of Drug Delivery[药物递送杂志],第2011卷,文章ID 469679,第12页,2011.doi:10.1155/2011/469679)。可通过挤出通过具有减小尺寸的过滤器来制备挤出的脂质，如Templeton等人,Nature Biotech[自然生物技术],15:647-652,1997中所述，其关于挤出脂质制备的教导通过引用并入文中。

[0230] 脂质纳米颗粒(LNP)是为本文所述的药物组合物提供生物相容性和可生物降解的递送系统的载剂的另一个实例。纳米结构化的脂质载剂(NLC)是经修饰的固体脂质纳米颗粒(SLN)，其保留了SLN的特性，提高了药物的稳定性和载药量，并防止了药物泄漏。聚合物纳米颗粒(PNP)是药物递送的重要组成部分。这些纳米颗粒可以有效地将药物递送引导至特定靶并改善药物稳定性和受控的药物释放。也可以使用脂质聚合物纳米颗粒(PLN)，其是一种组合了脂质体和聚合物的新型载剂。这些纳米颗粒具有PNP和脂质体的互补优势。PLN由核-壳结构构成;聚合物核提供了稳定的结构，磷脂壳提供了良好的生物相容性。这样，这两种组分提高了药物包封有效率，促进了表面修饰，并防止了水溶性药物的泄漏。对于综述，参见例如Li等人2017,Nanomaterials[纳米材料]7,122;doi:10.3390/nano7060122。

[0231] 外来体也可用作本文所述的组合物和系统的药物递送媒剂。对于综述，参见Ha等人2016年7月.Acta Pharmaceutica Sinica B.[药学学报]第6卷,第4期,第287-296页;https://doi.org/10.1016/j.apsb.2016.02.001。

[0232] 所有的公开物、专利申请、专利、以及在此引用的其他出版物和参考文献(例如，序列数据库参考数字)通过引用以其全文并入本文。例如，本文例如在本文的任何表或实例中提及的所有GenBank、Unigene和Entrez序列都通过引用并入本文。除非另有说明，否则本文指定的序列登录号(包括本文的任何表中的序列登录号)是指截至2018年11月29日的当前数据库条目。当一个基因或蛋白质引用多个序列登录号时，所有序列变体都包括在内。

[0233] 实例

[0234] 通过以下实例进一步说明本发明。提供这些实例仅出于说明目的，而不应以任何方式解释为限制本发明的范围或内容。

[0235] 实例1:融合构建体的设计与表达

[0236] 该实例描述了包含与RNA编辑结构域融合的RNA结合结构域的融合蛋白的设计和生产。

[0237] RNA结合结构域:如上和Cheong和Tanaka.2006.PNAS[美国科学院院报]第103卷,37:13635-9所述，将靶RNA的8个核苷酸的序列转化为拓扑蛋白识别码。使用例如编码PUM1的pTYB3质粒的定点诱变用Quick Change II XL定点诱变试剂盒(Stratagene公司，拉荷亚，加利福尼亚州)，将该密码整合到PUM1(SEQ ID NO:1)的RNA结合结构域中，所述RNA结合结构域是GenBank:AAG31807.1的氨基酸序列的Gly 828至Gly 1176。

[0238] RNA编辑结构域:设计构建体，该构建体包含具有E488Q突变的人ADAR2的催化活性结构域(hADAR2DD)(SEQ ID NO:2的aa 276-701)，以增强脱氨酶活性，如描述于Kuttan和

Bass.2012.PNAS[美国科学院院报]2012和Phelps,Kelly J等人Nucleic Acids Research[核酸研究]2015。

[0239] 从上述质粒合成上述RNA结合结构域和RNA编辑结构域的ORF的相应序列,并使用Sinnamon等人2017PNAS[美国科学院院报]114.44(2017):E9395-E9402中所述的引物通过聚合酶链反应(PCR)进行扩增,然后按照制造商的说明书,使用Gibson Assembly ®方案(新英格兰生物实验室公司)将其克隆到氨苄青霉素抗性的pcDNA-CMV载体主链中,其中将RNA结合结构域在hADAR2DD的C末端框内融合。用DNA测序确认构建体。

[0240] 融合蛋白可以在大肠杆菌菌株BL21(DE3)细胞中表达。质粒在大肠杆菌细胞中表达,在Lennox LB培养基(西格玛公司(Sigma),美国)中于37℃生长过夜。通过在6000g下离心30min收获细胞,然后将其重悬于裂解缓冲液中,进行超声处理并纯化,如Wang,X.等人2002中所述。通过以20,000rpm旋转30min来清除裂解物,然后将其加载到10ml Ni-NTA琼脂糖柱(凯杰公司,美国)上。将洗脱液用Sephedex75凝胶过滤柱纯化,然后在10mM Tris(pH 7.4)、150mM NaCl和2mM二硫苏糖醇(DTT)中浓缩至约5.5mg/ml。将等分试样在液氮中速冻并在-80℃下储存,如Wang,X.等人2002和Dawson,T.R.2003。

[0241] 通过SDS/PAGE和考马斯亮蓝染色来确认蛋白纯化。将融合蛋白的峰级分在100μg/ml牛血清白蛋白(BSA)中连续稀释,并通过SDS-PAGE与已知浓度的BSA(作为标准)比较进行解析。

[0242] 实例2:编辑效率测定

[0243] 该实例描述了评估如本文所述制备的融合蛋白的编辑效率的测定。

[0244] 将Panoply TM人ADAR敲低HEK293细胞(创造生物基因公司(Creative Biogene),Shirley,纽约)以(3×105/孔)的密度接种到聚d-赖氨酸包衣的24孔板中,维持在补充有10%胎牛血清、1%青霉素-链霉素溶液、1mM的丙酮酸钠和2mM的谷氨酰胺的杜尔贝科改良伊格尔培养基(DMEM)中在37℃、5%CO2下持续24小时。根据制造商的方案,用融合蛋白的构建体与lipofectamine 2000一起转染细胞,并在转染后维持72h。按照制造商的说明书,使用TRIZOL(英杰公司(Invitrogen))从细胞中分离总RNA,然后对1ug总RNA进行DNA酶I处理,然后进行逆转录,从而验证RNA编辑效率。用iScript cDNA合成试剂盒(伯乐公司(BioRad),赫拉克勒斯,加利福尼亚州)与随机选择的RT引物合成cDNA,然后使其经受PCR扩增。直接对产物进行测序,以比较经主题融合蛋白转染的ADAR缺陷细胞与其时间匹配对照的(A)至肌苷(I)取代,如描述于Wettengel等人2016.Nucleic Acids Research[核酸研究]45,5:2797-2808。

[0245] 实例3:示例性ORF点突变的RNA编辑

[0246] 该实例证实了本发明的融合多肽编辑ORF mRNA的能力。

[0247] 在此实例中,RNA编辑用于改变与神经元障碍中离子通量失调相关的转运蛋白的序列和功能。在神经元中,由自然存在的ADAR复合物编辑近99%的α-氨基-3-羟基-5-甲基-4-异噁唑丙酸(AMPA)复合物GluA2亚基。GluA2的此编辑将极性谷氨酰胺(Q)的密码子转换为带电荷精氨酸(R)的密码子。这种转换导致的运动神经元(其中GluA2已被编辑)中的Ca2+渗透性丧失。肌萎缩性侧索硬化(ALS)患者表现出这种GluA2编辑的丧失,并导致运动神经元中钙相关的兴奋毒性。本发明的多肽可用于编辑人GluA2mRNA的靶密码子,以在所得蛋白质中产生纠正性氨基酸取代Q607R。GluA2的氨基酸607的密码子包含人GluA2核苷酸序列

(NCBI参考序列NM_001083620.1)的核苷酸1555-1557。因此,本发明的效应子多肽包括人ADAR的催化结构域,其将相关密码子CAG(谷氨酰胺)编辑为CIG,其被读作CGG(精氨酸),并连接至RNA结合(靶向)结构域,该RNA结合(靶向)结构域将特异性地结合人GluA2核苷酸序列的核苷酸1555-1557上游的序列(参见图1)。特别地,如下构建效应子多肽:

[0248] RNA结合结构域:改变PUM1-HD的序列以结合要编辑的靶密码子的上游8个核苷酸的序列(人GluA2核苷酸序列的氨基酸1545-1552:caagaagc),以如下创建GluA2.RBD:

[0249] 重复序列1:突变S863C和Q867R

[0250] 野生型:HIMEFSQDQHGSRFIQLKLERATPAERQLVFNEILQ

[0251] 用于GluA2识别的突变体:

[0252] HIMEFSQDQHGCRFIRLKLERATPAERQLVFNEILQ

[0253] 重复序列2:突变N899S和Y867R

[0254] 野生型:AAYQLMVDVFGNYVIQKFFEFGSLEQKLALAERIRG

[0255] 用于GluA2识别的突变体:

[0256] AAYQLMVDVFGSRVIQKFFEFGSLEQKLALAERIRG

[0257] 重复序列3:突变C935S、R936N和Q939E

[0258] 野生型:HVLSLALQMYGCRVIQKALEFIPSDQQNEMVRELDG

[0259] 用于GluA2识别的突变体:

[0260] AAYQLMVDVFGSNVIEKFFEFGSLEQKLALAERIRG

[0261] 重复序列4:突变N971S和H972R

[0262] 野生型:HVLKCVKDQNGNHVVQKCIECVQPQSLQFIIDAFKG

[0263] 用于GluA2识别的突变体:

[0264] HVLKCVKDQNGSRVVQKCIECVQPQSLQFIIDAFKG

[0265] 重复序列5:无突变

[0266] 野生型:QVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQ

[0267] 用于GluA2识别的突变体:

[0268] QVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQ

[0269] 重复序列6:N1043S、Y1044N和Q1047E

[0270] 野生型:HTEQLVQDQYGNYVIQHVLEHGRPEDKSKIVAEIRG

[0271] 用于GluA2识别的突变体:

[0272] HTEQLVQDQYGSNVIEHVLEHGRPEDKSKIVAEIRG

[0273] 重复序列7:无突变

[0274] 野生型:

[0275] NVLVLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPHS

[0276] 用于GluA2识别的突变体:

[0277] NVLVLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPHS

[0278] 重复序列8:N1122C和Q1126R

[0279] 野生型:ALYTMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRP

[0280] HIATLRKYTYGKHILAKLEKYYMKNGVDLG

[0281] 用于GluA2识别的突变体:

[0282] ALYTMMKDQYACYVVRKMIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYMKNGVDLG

[0283] RNA编辑结构域:RNA编辑结构域包含人ADAR2DD的催化结构域,并如实例1中进行设计和制备。如Adamala等人2016.PNAS[美国科学院院报]113.19:E2579-E2588所述进行融合构建体GluA2.RBD-ADAR2DD的PCR连接和扩增。用于该实例的方法的可替代示例性构建体包括本文所公开的RNA编辑结构域、RNA效应子和/或多肽(和/或编码其的核酸),例如,SEQ ID NO:16或17(和/或SEQ ID NO:7或8)。可替代的示例性靶RNA序列包括对应于人GluA2(参考序列NM_000826)的核苷酸1537-1552的mRNA序列或人GluA2的核苷酸1537-1552的50个核苷酸内的序列。

[0284] 测定:上述GluA2.RBD-ADAR2DD构建体用于以下测试模型:

[0285] 小鼠神经母细胞瘤(N2A)细胞,将细胞以每孔1×10^3个细胞的密度接种在24孔板中并维持在补充有10%胎牛血清、1%青霉素-链霉素溶液、1mM丙酮酸钠和2mM谷氨酰胺的伊格尔最低必需培养基(EMEM)中在37℃、5%CO_2下过夜。24h后,用ADAR siRNA慢病毒(abmcat:iV037759a)质粒转染细胞。通过针对ADAR表达水平的蛋白质印迹来确认ADAR敲低。

[0286] 24小时后,用Lipofectamine 2000(赛默飞世尔科技公司)和125ng上述GluA2.RBD-ADAR2DD质粒转染低Opti-MEM血清培养基(赛默飞世尔科技公司)中ADAR缺陷N2A细胞。72h后,通过按照制造商网站上的方案,用TRIZOL(英杰公司)从细胞中分离总RNA,然后对1ug总RNA进行DNA酶I处理,然后进行逆转录,从而验证RNA编辑。用iScript cDNA合成试剂盒(伯乐公司,赫拉克勒斯,加利福尼亚州)与GluA2 RT引物(fwd:CCATCGAAAGTGCTGAGGAT和rev:AGGGCTCTGCACTCCTCATA)合成cDNA,并使其经受PCR扩增。直接对产物进行测序,以比较经GluA2.RBD-ADAR2DD转染的ADAR缺陷细胞与其时间匹配的没有ADAR活性的对照的(A)至肌苷(I)取代率,如描述于Wettengel,Jacqueline等人。

[0287] 实例4:编辑前体mRNA以生成选择性剪接产物

[0288] 在脊髓性肌萎缩症(SMA)(婴儿死亡的主要遗传原因)中,由于SMN1基因的突变或缺失而缺乏SMN蛋白。(Hua等人2007.PLoS biology[公共科学图书馆·生物学]第5卷,4:e73.)人具有SMN2基因(智人运动神经元2的存活,着丝粒(SMN2),第5号染色体NCBI参考序列:NG_008728.1上的RefSeqGene),与能够产生SMN蛋白的SMN1几乎相同;然而,外显子7的第6位(转录本中C6U转变)处关键胞嘧啶(C)向胸苷(T)突变以及内含子7的第100位(A100G)处腺苷(A)向鸟苷(G)转变降低了剪接位点的识别,导致前体mRNA剪接事件中外显子7的跳过。(Singh等人2012.PLoS One[公共科学图书馆综合]7.11:e49595)。由于跳过外显子,后续的SMN蛋白不稳定且部分起功能,从而导致SMA表型。该实例描述了本文所述的示例性组合物的设计和制备,该组合物可减少SMN2的前体mRNA中外显子7的“剪接出”,从而挽救SMN的产生并消除疾病表型。

[0289] 质粒构建

[0290] RNA结合结构域:对SMN2前体mRNA进行修饰,以用SMN2 RNA结合-hADARDD融合构建体驱动包含外显子7。Hua等人2007.PLoS biology[公共科学图书馆·生物学]第5卷,4:e73中描述了SMN2的靶序列,其增强包含外显子7。为了靶向SMN2的外显子7,我们进行了PUM1-HD的定点诱变,以靶向8个核苷酸的序列:UUAGACAA(人SMN2 NCBI参考序列NG_008728.1的位置27003-27010)

[0291] 对PUM 1进行的突变如下：

[0292] 重复序列1:突变S863N和R864Y

[0293] 野生型:HIMEFSQDQHGSRFIQLKLERATPAERQLVFNEILQ

[0294] 用于SMN2识别的突变体：

[0295] HIMEFSQDQHGNYFIQLKLERATPAERQLVFNEILQ

[0296] 重复序列2:无突变

[0297] 野生型:AAYQLMVDVFGNYVIQKFFEFGSLEQKLALAERIRG

[0298] 用于SMN2识别的突变体：

[0299] AAYQLMVDVFGNYVIQKFFEFGSLEQKLALAERIRG

[0300] 重复序列3:无突变

[0301] 野生型:HVLSLALQMYGCRVIQKALEFIPSDQQNEMVRELDG

[0302] 用于SMN2识别的突变体：

[0303] AAYQLMVDVFGCRVIQKFFEFGSLEQKLALAERIRG

[0304] 重复序列4:N971S、H972N和Q975E

[0305] 野生型:HVLKCVKDQNGNHVVQKCIECVQPQSLQFIIDAFKG

[0306] 用于SMN2识别的突变体：

[0307] HVLKCVKDQNGSNVVEKCIECVQPQSLQFIIDAFKG

[0308] 重复序列5:无突变

[0309] 野生型:QVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQ

[0310] 用于SMN2识别的突变体：

[0311] QVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQ

[0312] 重复序列6:N1043C和Q1047R

[0313] 野生型:HTEQLVQDQYGNYVIQHVLEHGRPEDKSKIVAEIRG

[0314] 用于SMN2识别的突变体：

[0315] HTEQLVQDQYGCYVIRHVLEHGRPEDKSKIVAEIRG

[0316] 重复序列7:S1079C、N1080R和E1083Q

[0317] 野生型：

[0318] NVLVLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPHS

[0319] 用于SMN2识别的突变体：

[0320] NVLVLSQHKFACRVVQKCVTHASRTERAVLIDEVCTMNDGPHS

[0321] 重复序列8:N1122C和Y1123R

[0322] 野生型:ALYTMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYMKNGVDLG

[0323] 用于SMN2识别的突变体：

[0324] ALYTMMKDQYACRVVQKMIDVAEPGQRKIVMHKIRP

[0325] HIATLRKYTYGKHILAKLEKYYMKNGVDLG

[0326] RNA编辑结构域:RNA编辑结构域包含人ADAR2DD的催化结构域,并如实例1中进行设计和制备。

[0327] 如Adamala等人2016.PNAS[美国科学院院报]113.19:E2579-E2588所述进行融合

构建体SMN2.RBD-ADAR2DD的PCR连接和扩增。用于该实例的方法的可替代示例性构建体包括本文所公开的RNA编辑结构域、RNA效应子和/或多肽(和/或编码其的核酸),例如,SEQ ID NO:18或19(和/或SEQ ID NO:9或10)。可替代的示例性靶RNA序列包括对应于人SMN2(参考序列NM-022876)的核苷酸31,995-32,010的mRNA序列或人SMN2的核苷酸31,995-32,010的50个核苷酸内的序列。

[0328] 细胞培养和转染

[0329] 将人SMA I型成纤维细胞(柯瑞尔细胞库(Coriell Repositories))在转染前24小时进行铺板,并维持在37℃、5%CO2下在补充有10%非灭活FBS的DMEM中。在约50%汇合时,将细胞用0.5μg SMN2.RBD-hADARDD质粒瞬时转染。4h后,用新鲜培养基替换培养基。转染48h后提取总RNA。

[0330] 针对外显子包含的RT-PCR

[0331] 在测试细胞上进行用于检测SMN2的外显子7剪接的RT-PCR分析,如先前在Cho等人2014.Biochimica et Biophysica Acta[生物化学与生物物理学报](BBA)-Gene Regulatory Mechanisms[(BBA)-基因调节机制]1839.6:517-525中所述。对照细胞用表达hADARDD(没有RNA结合融合)的质粒转染。

[0332] 通过RiboEx试剂(Geneall公司)和乙醇沉淀从对照和测试哺乳动物细胞中提取总RNA。以20μl的总体积进行逆转录,包含1μg RNA、0.5μg寡dT、dNTP混合物(每种dNTP 0.5mM)、6mM MgCl2、4μl 5X ImProm-II Tm反应缓冲液和1μl ImProm-II TM逆转录酶(普洛麦格公司(Promega))。进行SMN+外显子7、SMN-外显子7和GAPDH对照的RT-PCR扩增,并在2%琼脂糖凝胶上用溴化乙锭溶液(0.5μ/ml)分析PCR产物(使用例如Cho等人的外显子6和外显子8PCR引物扩增的)。测试细胞产生更大的SMN2 mRNA,其包括外显子7。将PCR产物用DdeI(NEB)消化,并加载到5%天然聚丙烯酰胺凝胶上进行检测。

[0333] 实例5:编辑EBNA1的序列以诱导对Eb病毒的抗病毒应答

[0334] EB病毒(EBV)引起单核细胞增多症,并与许多人类癌症相关,包括伯基特淋巴瘤、霍奇金氏和鼻咽癌(Tellam等人2008.PNAS[美国科学院院报]105.27:9319-9324)。最初裂解性感染后,已证明EBV可以逃避免疫监视并持续潜伏性感染。在潜伏性感染期间,为了限制抗病毒的细胞毒性T应答,EBV编码的核抗原1(EBNA1)维持编码的蛋白质序列,但偏向于mRNA中使用的密码子,使得后续的二级结构不包含抗病毒应答和下游抗原呈递所必需的双链茎特征。EBNA中的甘氨酸-丙氨酸重复结构域(GAr)负责翻译效率和增强的免疫识别。在该结构域中,GAr结构域(UniprotKB P03211的位置87-352)内99%的甘氨酸残基和100%的丙氨酸残基由嘌呤密码子(GGG、GGA和GCA)构成,其含量明显高于人类平均水平甘氨酸和丙氨酸嘌呤密码子(分别为49.3%和33.3%)(Tellam)。该实例描述了本文描述的示例性组合物的设计和制备,以编辑EBNA1的序列以增强病毒mRNA的二级结构,从而诱导抗病毒应答。

[0335] 质粒构建

[0336] RNA结合结构域:在Tellam等人2008.PNAS[美国科学院院报]105.27:9319-9324(表1)中引用了EBV E1-GAr的序列。在该实例中,为了靶向EBV E1-Gar二级结构,我们在8个核苷酸的序列:GCGGGAGG处(其发现于Tellam等人表1中发现的天然EBNA1 Gar的105-mer核苷酸序列的20-27位

[0337] (5'TAAaggagcaggagcagga**gcgggagg**ggcaggagcaggaggggcaggagcaggaggaggggc

aggagcaggaggaggggcaggaggggcaggaggggcaggaAT-3'))进行了PUM1-HD的定点诱变。

[0338] 为了靶向此序列,对hPUM 1进行的突变如下:

[0339] 重复序列1:R864N和Q867E

[0340] 野生型:HIMEFSQDQHGSRFIQLKLERATPAERQLVFNEILQ

[0341] 突变体:HIMEFSQDQHGSNFIELKLERATPAERQLVFNEILQ

[0342] 重复序列2:N899C和Q903R

[0343] 野生型:AAYQLMVDVFGNYVIQKFFEFGSLEQKLALAERIRG

[0344] 突变体:AAYQLMVDVFGCYVIRKFFEFGSLEQKLALAERIRG

[0345] 重复序列3:C935S、R936N和Q939E

[0346] 野生型:HVLSLALQMYGCRVIQKALEFIPSDQQNEMVRELDG

[0347] 突变体:AAYQLMVDVFGSNVIEKFFEFGSLEQKLALAERIRG

[0348] 重复序列4:N971S、H972N和Q975E

[0349] 野生型:HVLKCVKDQNGNHVVQKCIECVQPQSLQFIIDAFKG

[0350] 突变体:HVLKCVKDQNGSNVVEKCIECVQPQSLQFIIDAFKG

[0351] 重复序列5:C1007S、R1008N和Q1011E

[0352] 野生型:QVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQ

[0353] 突变体:QVFALSTHPYGSNVIERILEHCLPDQTLPILEELHQ

[0354] 重复序列6:N1043S和Y1044R

[0355] 野生型:HTEQLVQDQYGNYVIQHVLEHGRPEDKSKIVAEIRG

[0356] 突变体:HTEQLVQDQYGSRVIQHVLEHGRPEDKSKIVAEIRG

[0357] 重复序列7:无突变

[0358] 野生型:

[0359] NVLVLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPHS

[0360] 重复序列8:N1122S、Y1123N和Q1126E

[0361] 野生型:ALYTMMKDQYANYVVQKMIDVAEPGQRKIVMHKIRP

[0362] HIATLRKYTYGKHILAKLEKYYMKNGVDLG

[0363] 突变体:ALYTMMKDQYASNVVEKMIDVAEPGQRKIVMHKIRP

[0364] HIATLRKYTYGKHILAKLEKYYMKNGVDLG

[0365] RNA编辑结构域:RNA编辑结构域包含人ADAR2DD的催化结构域,并如实例1中进行设计和制备。

[0366] 如Adamala等人2016.PNAS[美国科学院院报]113.19:E2579-E2588所述进行融合构建体EBVE1-GAr.RBD-ADAR2DD的PCR连接和扩增。

[0367] EBNA1表达构建体

[0368] 为了生成EBNA1表达构建体,将全长EBV编码的EBNA1(E1)和EBNA1 GAr序列(EBNA1-GA)的102-nt增量克隆到表达载体pcDNA3(英杰公司)中。然后,如Tellam中所述,将表达载体用编码GFP的序列(pEGFP-N1;克罗泰克公司(Clontech))进行框内亚克隆。

[0369] 细胞培养和转染

[0370] 将DG75(ATCC)或HEK293细胞维持在补充有2mM L-谷氨酰胺、100个单位/ml青霉素和100μg/ml链霉素加10%FCS的RPMI培养基1640中,如先前在Tellam,J.等人PNAS[美国科

学院院报]2008中所述。

[0371] EBNA1构建体的转染

[0372] 通过使用BioRad Gene Pulser(960μF,250V,0.4-cm间隙电极,300-μl测定体积,25℃)用10μg表达构建体转染细胞。用EBV转染后2小时,用0.5μg EBVE1-GAr.RBD-ADAR2DD基因质粒瞬时转染细胞,然后4h后,用新鲜培养基替换培养基。最终转染后24小时,收获细胞并使其经受SDS/PAGE并用抗GFP(1:2,000)或肌动蛋白mAb(1:1,000)进行免疫印迹,如在Tellam,J.等人PNAS[美国科学院院报]2008中所述。

[0373] 翻译测定

[0374] 用XbaI将EBNA1/pcDNA3表达构建体进行线性化和通过使用补充有50μCi[α-32P]UTP(安玛西亚生物科学(Amersham Biosciences))的核糖核酸探针体外转录系统(普洛麦格公司(Promega))经T7 RNA聚合酶转录1μg模板。对于翻译测定,通过使用补充有250μCi 35[S]甲硫氨酸(安玛西亚生物科学)的偶联的转录/翻译网状细胞裂解物系统(普洛麦格公司),用T7RNA聚合酶对EBNA1/pcDNA3载体进行体外转录和翻译。使裂解物经受SDS/PAGE和放射自显影,如在Tellam,J.等人PNAS[美国科学院院报]2008中所述。通过测序确认编辑。

[0375] SHAPE分析

[0376] 用MFOLD预测经编辑的序列的最低能量确认,如在Zuker,M.等人Nucleic Acid Res[核酸研究]2003中所述。

[0377] qRT-PCR

[0378] 通过使用MMLV SuperScript III逆转录酶(英杰公司)和锚定的寡(T)18引物组合随机六聚体进行每个样品的EBNA1的cDNA合成和1μg RNA分离。使用基于Sybr Green的荧光检测系统和ABI Prism 7900序列检测系统(应用生物系统公司(Applied Biosystems))的qRT-PCR来测量mRNA丰度。将核糖体蛋白P0(RPLP0;GenBank登录号NM_053275)用作所有样品的参考基因,如在Tellam,J.等人PNAS[美国科学院院报]2008中所述。

[0379] 每个qRT-PCR包含2.5ml的2×Sybr Green预混液(应用生物系统公司)、0.25ml的每种引物(每种的最终浓度均为500nM)、1.0ml的水和1.0ml的储备cDNA模板的1/10稀释液。循环条件应为:在95℃下持续15s并在60℃下持续1min进行40个循环。每次运行完成时,进行解离熔解曲线分析。

[0380] 测量蛋白质合成

[0381] 用EBNA1-GFP表达构建体以及EBVE1-GAr.RBD-ADAR2DD转染HEK293细胞。转染后二十四小时,将细胞在含有20μCi/ml 3[H]甲硫氨酸(安玛西亚生物科学)的生长培养基中于37℃进行标记持续12-14h。将细胞在PBS中洗涤,并在不含甲硫氨酸的生长培养基中在37℃下孵育30-min,然后用100μCi 35[S]甲硫氨酸进行30-min脉冲。脉冲后,将细胞在含1%Triton X-100和蛋白酶抑制剂的Tris缓冲盐水中裂解,并用蛋白A琼脂糖预先清除,并将裂解物用抗GFP或针对β-微管蛋白的mAb(西格玛公司)进行免疫沉淀。将免疫沉淀的样品添加到10ml闪烁液Ultima Gold(珀金埃尔默生命与分析科学公司(PerkinElmer Life and Analytical Sciences))中,并在Packard液体闪烁分析仪Tri-carb 2100TR上计数。

[0382] 实例6:序列

[0383] SEQ ID NO:1

[0384] PUM1的RNA结合结构域;GenBank的氨基酸序列的Gly 828至Gly 1176:AAG31807.1

[0385] GRSRLLEDFRNNRYPNLQLREIAGHIMEFSQDQHGSRFIQLKLERATPAERQLVFNEILQAAYQLMVDV FGNYVIQKFFEFGSLEQKLALAERIRGHVLSLALQMYGCRVIQKALEFIPSDQQNEMVRELDGHVLKCVKDQNGNHV VQKCIECVQPQSLQFIIDAFKGQVFALSTHPYGCRVIQRILEHCLPDQTLPILEELHQHTEQLVQDQYGNYVIQHVL EHGRPEDKSKIVAEIRGNVLVLSQHKFASNVVEKCVTHASRTERAVLIDEVCTMNDGPHSALYTMMKDQYANYVVQK MIDVAEPGQRKIVMHKIRPHIATLRKYTYGKHILAKLEKYYMKNGVDLG

[0386] SEQ ID NO:2

[0387] ADAR2氨基酸序列;NCBI参考序列NG_052015.1

[0388] LSNGGGGGPGRKRPLEEGSNGHSKYRLKKRRKTPGPVLPKNALMQLNEIKPGLQYTLLSQ TGPVHAPLFVMSVEVNGQVFEGSGPTKKKAKLHAAEKALRSFVQFPNASEAHLAMGRTLS VNTDFTSDQADFPDTLFNGFETPDKAEPPFYVGSNGDDSFSSSGDLSLSASPVPASLAQP PLPVLPPFPPPSGKNPVMILNELRPGLKYDFLSESGESHAKSFVMSVVVDGQFFEGSGRN KKLAKARAAQSALAAIFNLHLDQTPSRQPIPSEGLQLHLPQVLADAVSRLVLGKFGDLTD NFSSPHARRKVLAGVVMTTGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISR RSLLRFLYTQLELYLNNKDDQKRSIFQKSERGGFRLKENVQFHLYISTSPCGDARIFSPH EPILEGSRSYTQAGVQWCNHGSLQPRPPGLLSDPSTSTFQGAGTTEPADRHPNRKARGQL RTKIESGEGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLLSIFVEPI YFSSIILGSLYHGDHLSRAMYQRISNIEDLPPLYTLNKPLLSGISNAEARQPGKAPNFSV NWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGKVPSHLLRSKITKPNVYHE SKLAAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQFSLTP (SEQ ID NO: 2)

[0389] SEQ ID NO:3:GluA2的氨基酸序列;NCBI参考序列:NM_000826.3.

[0390] MQKIMHISVL LSPVLWGLIF GVSSNSIQIG GLFPRGADQE YSAFRVGMVQ FSTSEFRLTP HIDNLEVANS FAVTNAFCSQ FSRGVYAIFG FYDKKSVNTI

[0391]

TSFCGTLHVS FITPSFPTDG THPFVIQMRP DLKGALLSLI EYYQWDKFAY
LYDSDRGLST LQAVLDSAAE KKWQVTAINV GNINNDKKDE MYRSLFQDLE
LKKERRVILD CERDKVNDIV DQVITIGKHV KGYHYIIANL GFTDGDLLKI
QFGGANVSGF QIVDYDDSLV SKFIERWSTL EEKEYPGAHT TTIKYTSALT
YDAVQVMTEA FRNLRKQRIE ISRRGNAGDC LANPAVPWGQ GVEIERALKQ
VQVEGLSGNI KFDQNGKRIN YTINIMELKT NGPRKIGYWS EVDKMVVTLT
ELPSGNDTSG LENKTVVVTT ILESPYVMMK KNHEMLEGNE RYEGYCVDLA
AEIAKHCGFK YKLTIVGDGK YGARDADTKI WNGMVGELVY GKADIAIAPL
TITLVREEVI DFSKPFMSLG ISIMIKKPQK SKPGVFSFLD PLAYEIWMCI
VFAYIGVSVV LFLVSRFSPY EWHTEEFEDG RETQSSESTN EFGIFNSLWF
SLGAFMQQGC DISPRSLSGR IVGGVWWFFT LIIISSYTAN LAAFLTVERM
VSPIESAEDL SKQTEIAYGT LDSGSTKEFF RRSKIAVFDK MWTYMRSAEP
SVFVRTTAEG VARVRKSKGK YAYLLESTMN EYIEQRKPCD TMKVGGNLDS
KGYGIATPKG SSLRNAVNLA VLKLNEQGLL DKLKNKWWYD KGECGSGGGD
SKEKTSALSL SNVAGVFYIL VGGLGLAMLV ALIEFCYKSR AEAKRMKVAK
NAQNINPSSS QNSQNFATYK EGYNVYGIES VKI

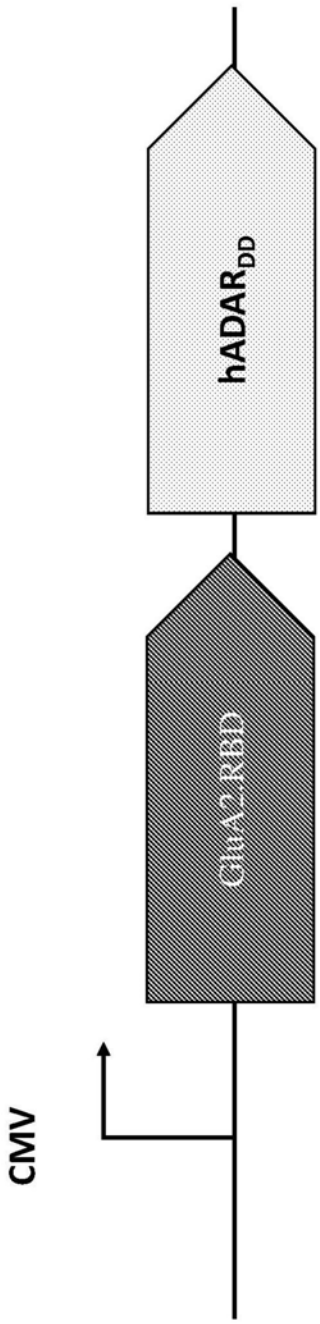


图1A

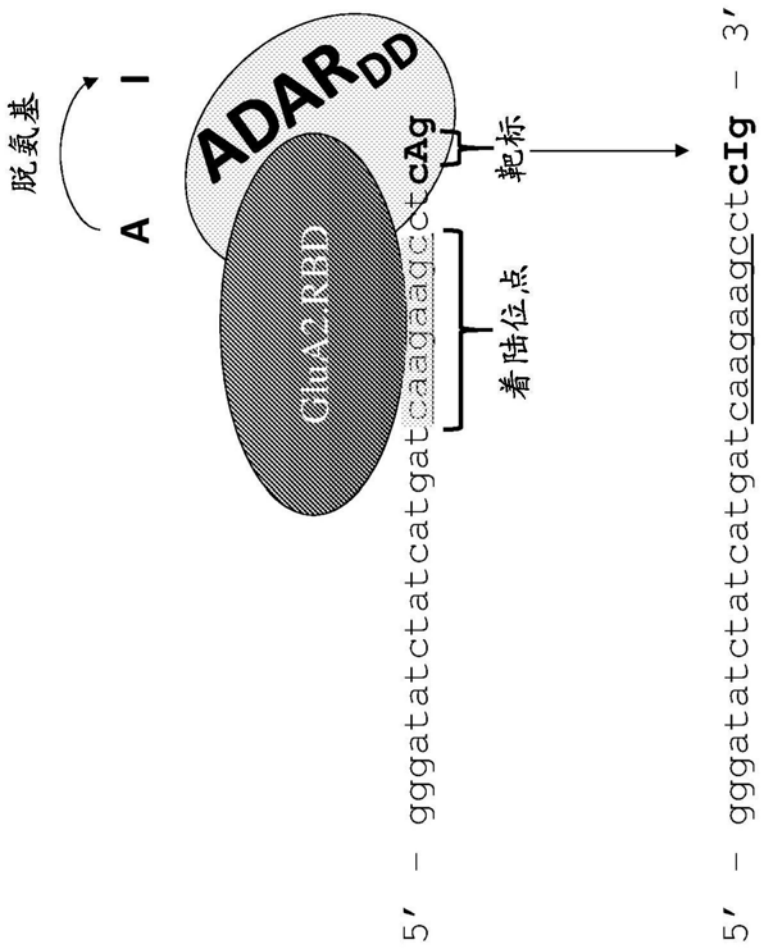

脱氨基
A
I
ADARDD
GluA2.RNB
靶标
着陆位点
5′ - gggatatctatcatgatcaagaagcctcAg
5′ - gggatatctatcatgatcaagaagcctcIg - 3′

图1B

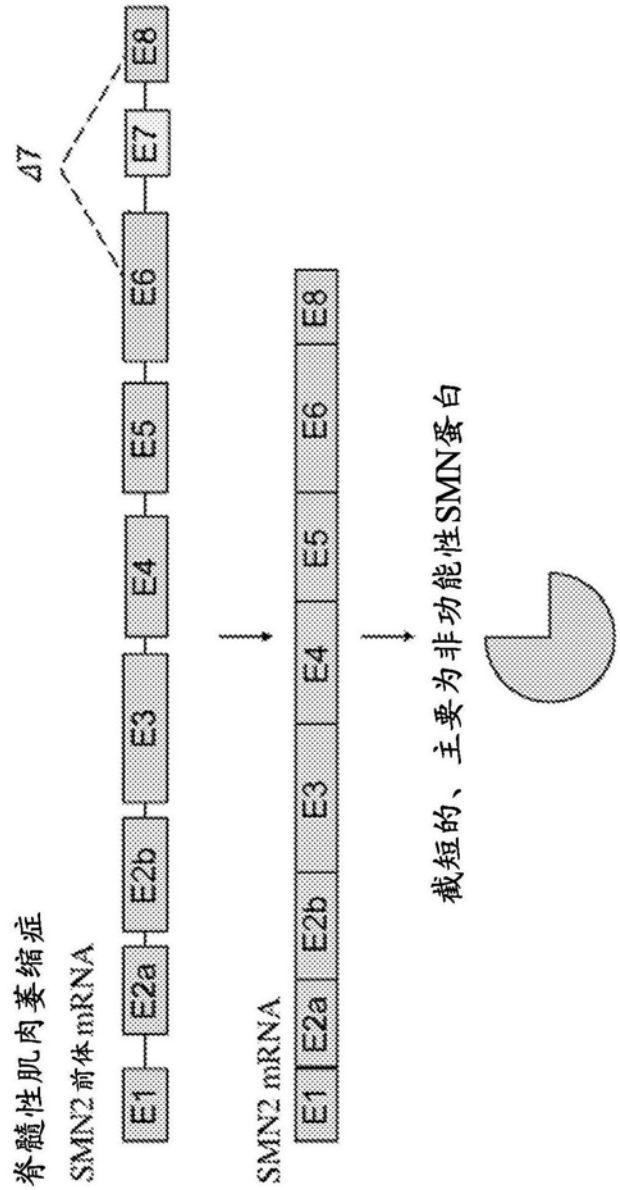

脊髓性肌肉萎缩症
SMN2 前体 mRNA
E1
E2a
E2b
E3
E4
E5
E6
Δ7
E7
E8
SMN2 mRNA
E1
E2a
E2b
E3
E4
E5
E6
E8
截短的、主要为非功能性SMN蛋白

图2

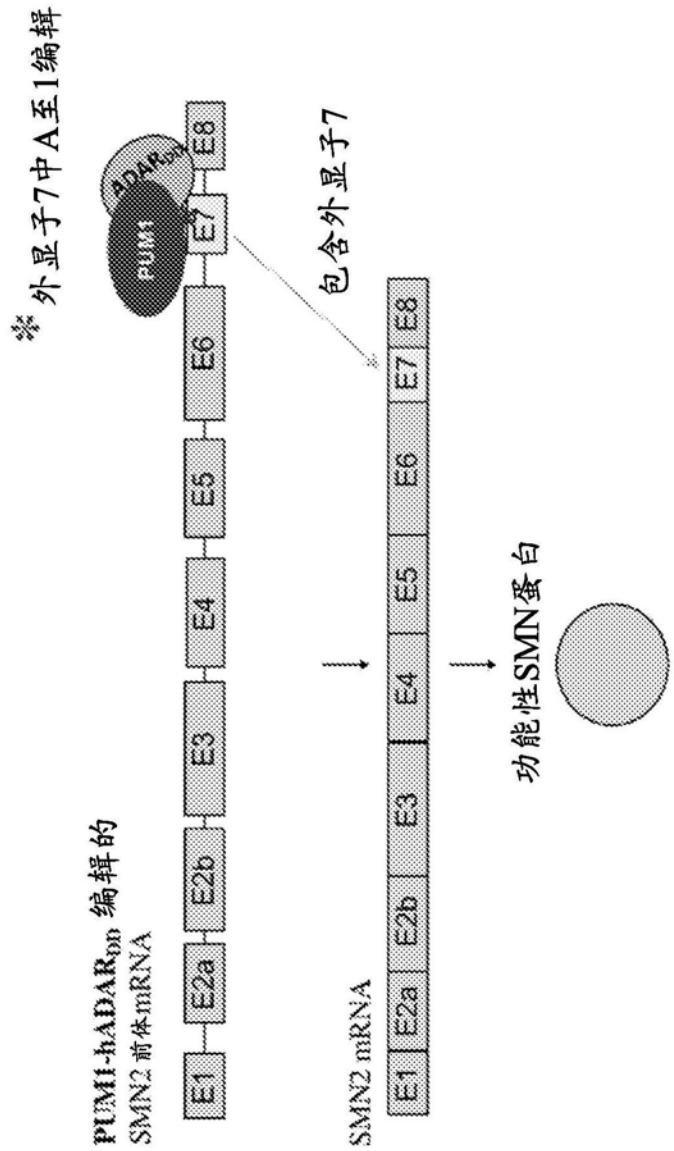

※外显子7中A至I编辑
PUM1-hADAR$_{DD}$ 编辑的
SMN2 前体mRNA
E1
E2a
E2b
E3
E4
E5
E6
E7
E8
PUM1
ADAR$_{DD}$
包含外显子7
SMN2 mRNA
E1
E2a
E2b
E3
E4
E5
E6
E7
E8
功能性SMN蛋白

图3

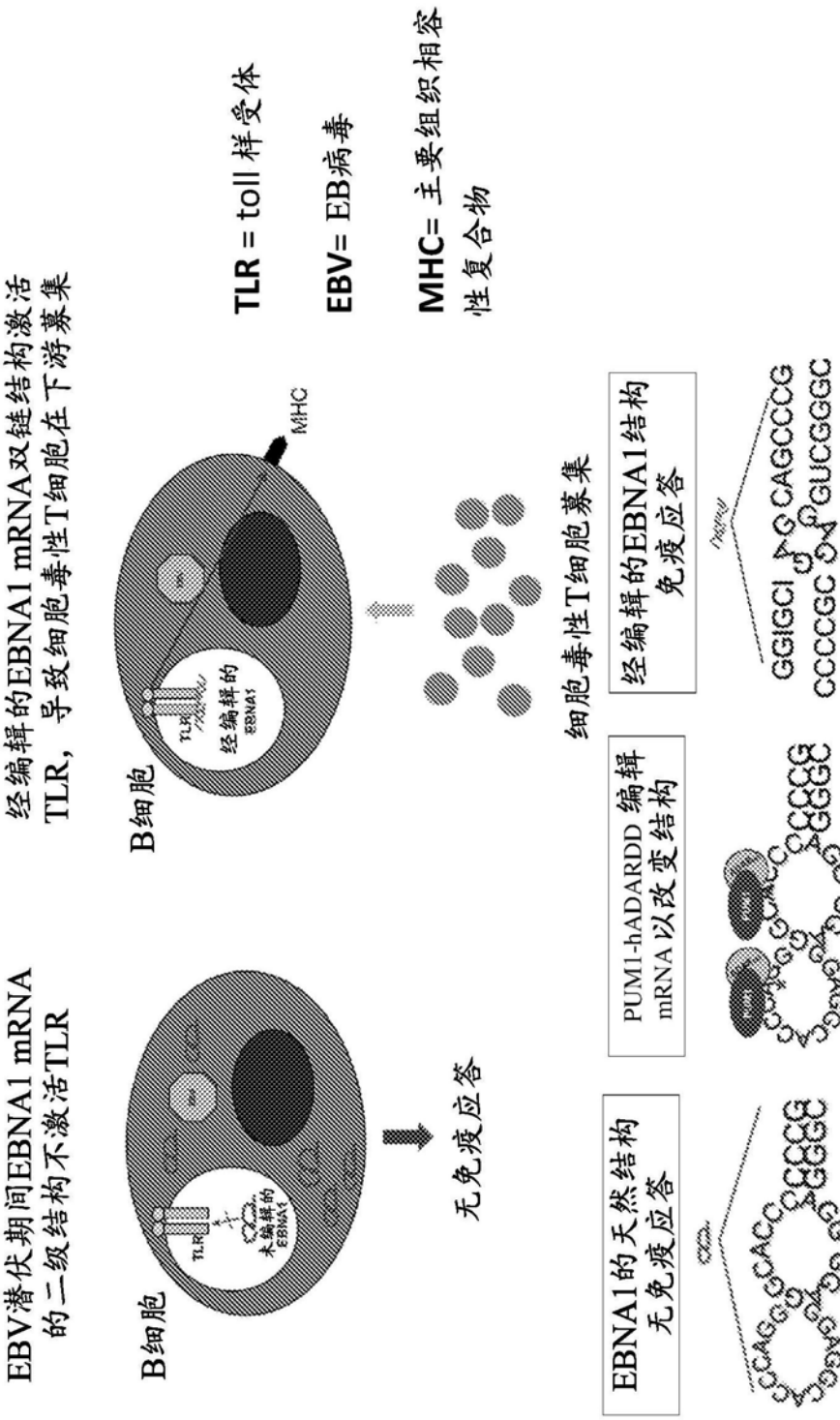

EBV潜伏期间EBNA1 mRNA
的二级结构不激活TLR
B细胞
TLR
未编辑的
EBNA1
无免疫应答
EBNA1的天然结构
无免疫应答
PUM1-hADARDD 编辑
mRNA 以改变结构
经编辑的EBNA1 mRNA双链结构激活
TLR，导致细胞毒性T细胞在下游募集
B细胞
TLR
经编辑的
EBNA1
MHC
细胞毒性T细胞募集
经编辑的EBNA1结构
免疫应答
GGIGCI
CCCCGC
CAGCCCG
GUCGGGC
TLR = toll 样受体
EBV= EB病毒
MHC= 主要组织相容
性复合物

图4