(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118103516 A
(43)申请公布日 2024.05.28

(21)申请号 202280068980.1
(22)申请日 2022.10.17
(30)优先权数据
PCT/EP2021/078830 2021.10.18 EP
(85)PCT国际申请进入国家阶段日
2024.04.12
(86)PCT国际申请的申请数据
PCT/EP2022/078870 2022.10.17
(87)PCT国际申请的公布数据
WO2023/066875 EN 2023.04.27
(71)申请人 生物技术欧洲股份公司
地址 德国美因兹
申请人 约翰·古腾堡大学美因兹医学大学转化肿瘤学公司
(72)发明人 M·佩尔科维奇 T·拜塞特
(74)专利代理机构 永新专利商标代理有限公司 72002
专利代理师 曹雯 黄歆
(51)Int.Cl.
C12N 15/86(2006.01)
A61K 39/00(2006.01)
C12N 7/00(2006.01)
A61P 31/14(2006.01)
A61P 35/00(2006.01)

权利要求书3页 说明书74页
序列表（电子公布） 附图10页

(54)发明名称

修饰的可复制RNA和相关组合物及其用途

(57)摘要

本发明涉及含有修饰的核苷酸的可复制RNA,其能够以比不含修饰的核苷酸的相应可复制RNA更高的水平复制和/或翻译,以及这类修饰的可复制RNA分子在各种类型的治疗中的用途。

CN 118103516 A

1.一种5’加帽修饰的可复制RNA分子,其包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了所述分子中的第一个5’尿苷外,所述分子中的至少一个尿苷是修饰的尿苷。

2.权利要求1的分子,其中除了所述分子中的第一个5’尿苷外,所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的尿苷是1mΨ。

3.权利要求1或2的分子,其中除了所述分子中的第一个5’尿苷外,所述分子中的所有尿苷为1mΨ。

4.权利要求1-3中任一项的分子,其中所述分子包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,其包含RNA依赖性RNA聚合酶(复制酶),优选甲病毒复制酶。

5.一种修饰的可复制RNA分子,其包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中所述分子包含序列AUGGCGGA或AUGGGCGG,其中任一序列中的U是尿苷;并且其中所述分子中剩余尿苷中的至少一个是修饰的尿苷。

6.权利要求5的分子,其中所述分子中所有剩余尿苷为1mΨ。

7.权利要求5或6的分子,其中所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的剩余尿苷是1mΨ。

8.权利要求5-7中任一项的分子,其中所述序列AUGGCGGA或AUGGGCGG位于所述分子的非编码区中。

9.权利要求5-8中任一项的分子,其中所述序列AUGGCGGA或AUGGGCGG位于5’调节区中。

10.权利要求5-9中任一项的分子,其中所述序列AUGGCGGA或AUGGGCGG位于5’调节区中的保守序列元件1(CSE1)中。

11.权利要求5-10中任一项的分子,其中所述序列AUGGCGGA或AUGGGCGG位于所述分子的5’末端。

12.权利要求5-11中任一项的分子,其中所述分子包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,其包含RNA依赖性RNA聚合酶(复制酶),优选甲病毒复制酶。

13.权利要求5-12中任一项的分子,其中所述序列AUGGCGGA或AUGGGCGG进一步包含分别在序列AUGGCGGA或AUGGGCGG的5’的额外核苷酸。

14.权利要求13的分子,其中所述额外核苷酸包含额外的ORF和/或控制序列或者形成5’帽结构的一个或多个核苷酸。

15.一种修饰的可复制RNA分子,其包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了5’调节区中包含的保守序列元件1(CSE 1)的10个5’核苷酸内包含的尿苷外,所述分子中的尿苷中的至少一个尿苷是修饰的尿苷。

16.权利要求15的分子,其中除了CSE 1的10个5’核苷酸内包含的尿苷外,所述分子中的所有尿苷为1mΨ。

17.权利要求15或16的分子,其中除了CSE 1的5个5’核苷酸内包含的尿苷外,所述分子中的所有尿苷为1mΨ。

18.权利要求15-17中任一项的分子,其中除了CSE 1的4个5’核苷酸内包含的尿苷外,所述分子中的所有尿苷为1mΨ。

19.权利要求15-18中任一项的分子,其中除了CSE 1的3个5'核苷酸内包含的尿苷外,所述分子中的所有尿苷为1mΨ。

20.权利要求15-19中任一项的分子,其中除了CSE 1的2个5'核苷酸内包含的尿苷外,所述分子中的所有尿苷为1mΨ。

21.权利要求15-20中任一项的分子,其中除了CSE 1中最5'的尿苷外,所述分子中的所有尿苷为1mΨ。

22.权利要求15-21中任一项的分子,其中除了CSE 1中第2位的尿苷外,所述分子中的所有尿苷为1mΨ。

23.权利要求15-22中任一项的分子,其中所述分子包含5'帽,并且CSE1的10个5'核苷酸包括5'帽的任何核苷酸。

24.权利要求15-23中任一项的分子,其中所述分子包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,其包含RNA依赖性RNA聚合酶(复制酶),优选甲病毒复制酶。

25.一种修饰的可复制RNA分子,其包含甲病毒的5'调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了5'调节区中包含的保守序列元件1(CSE1)内包含的最5'的尿苷外,所述分子中的至少一个尿苷是修饰的尿苷。

26.权利要求25的分子,其中除了CSE1内包含的最5'的U外,所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的尿苷是1mΨ。

27.权利要求25或26的分子,其中除了CSE1内包含的最5'的U外,所述分子中的所有尿苷为1mΨ。

28.权利要求25-27中任一项的分子,其中所述分子包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,其包含RNA依赖性RNA聚合酶(复制酶),优选甲病毒复制酶。

29.一种修饰的可复制RNA分子,其包含至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中所述分子中的至少一个尿苷是修饰的尿苷,并且其中所述分子包含具有序列NpppNU的5'帽,其中所述5'帽中的U是未修饰的尿苷,优选其中所述序列是NpppAU。

30.权利要求29的分子,其中所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的尿苷是1mΨ。

31.权利要求29或30的分子,其中所述分子包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,其包含RNA依赖性RNA聚合酶(复制酶),优选甲病毒复制酶。

32.权利要求1-31中任一项的分子,其中所述修饰的尿苷是N1-甲基-假尿苷。

33.权利要求1-32中任一项的分子,其中所述分子进一步包含至少一个修饰的G、C或A。

34.权利要求1-33中任一项的分子,其中所述分子具有修饰的骨架。

35.权利要求34的分子,其中所述骨架包含至少一个硫代磷酸酯键,优选其中所有键是硫代磷酸酯键。

36.权利要求1-35中任一项的分子,其中唯一的核苷酸或核苷修饰是1mΨ。

37.权利要求1-36中任一项的分子,其中所述甲病毒是SFV或VEEV。

38.权利要求1-37中任一项的分子,其中所关注的基因编码抗原(肿瘤、病毒、细菌、真菌、过敏原)或治疗性蛋白或核酸。

39.权利要求1-38中任一项的分子,其中所述5’调节区和编码的RNA依赖性RNA聚合酶源自相同的甲病毒或源自不同的甲病毒。

40.权利要求1-39中任一项的分子,其中所述5’调节区没有起始密码子。

41.权利要求1-40中任一项的分子,其是体外转录的RNA分子。

42.权利要求1-41中任一项的分子,其用脂质配制。

43.一种药物组合物,其包含权利要求1-42中任一项的分子以及药学可接受的载剂或赋形剂。

44.权利要求1-42中任一项的修饰的可复制RNA分子或权利要求43的药物组合物,其用于治疗。

45.一种提高受试者的免疫应答的方法,所述方法包括给药权利要求1-42中任一项的修饰的可复制RNA分子或权利要求43的药物组合物。

46.一种修饰的可复制RNA分子,其用于提高受试者的免疫应答,所述方法包括向所述受试者给药权利要求1-42中任一项的分子或权利要求43的药物组合物。

47.一种治疗受试者的癌症的方法,所述方法包括向所述受试者给药权利要求1-42中任一项的修饰的可复制RNA分子或权利要求43的药物组合物。

48.一种修饰的可复制RNA分子,其用于治疗受试者的癌症的方法,所述方法包括向所述受试者给药权利要求1-42中任一项的分子或权利要求43的药物组合物。

49.一种向缺乏基因功能的受试者提供这种基因功能的方法,所述方法包括向所述受试者给药权利要求1-42中任一项的修饰的可复制RNA分子或权利要求43的药物组合物。

50.一种修饰的可复制RNA分子,其用于向缺乏基因功能的受试者提供这种基因功能的方法,所述方法包括向所述受试者给药权利要求1-42中任一项的分子或权利要求43的药物组合物。

修饰的可复制RNA和相关组合物及其用途

技术领域

[0001] 本发明涉及可复制RNA构建体/分子,其通过包含至少一个修饰的核苷酸如N1-甲基-假尿苷(1mΨ)进行修饰,并且与相应的未修饰的可复制RNA构建体/分子相比能够以相同或相似的水平复制和/或翻译。本发明的可复制RNA构建体是指至少一部分被适当的RNA依赖性RNA聚合酶(复制酶)识别用于复制的区域不含有修饰的核苷酸的可复制RNA构建体。本发明还涉及这类可复制RNA分子在治疗中的用途。

背景技术

[0002] 最近,基于mRNA的疫苗在抗击Covid-19流行病的临床研究中证明了其免疫原性。这些RNA疫苗非常有效,诱导非常强的T细胞免疫应答和高水平的中和抗体(Walsh et al.,2020,N Engl J Med 383:2439-2450;Sahin et al.,2020,Nature 586:594-599)。获得监管部门批准的前两种mRNA疫苗含有化学修饰的核苷酸,N1-甲基-假尿苷(1mΨ),而不是尿苷。这种修饰通过在很大程度上避免刺激先天免疫途径导致干扰素反应,提高了免疫活性(immune competent)细胞中mRNA的翻译(Andries et al.,2015,J Control Release 217:337-344)。

[0003] 这些已批准的RNA疫苗每剂需要30-100μg RNA,并且需要连续两个剂量,间隔数周(初免-加强(prime-boost)方案)。这最终导致需要60-200g RNA来免疫100万人。因此将剂量减少到1μg以下会对为人群提供针对新型病原体的疫苗所需的生产时间产生重大影响。

[0004] 正在研究的有望实现显著减少剂量的疫苗方法是使用自扩增RNA(saRNA)。通过用期望对其免疫应答的抗原替代甲病毒结构基因,可以从甲病毒基因组工程化saRNA。saRNA编码甲病毒复制酶,其具有复制saRNA分子的所有酶促功能,从而导致输入疫苗量的扩增。不幸的是,由saRNA引起的先天免疫应答强烈抑制saRNA疫苗的有效性,这可能是在非人灵长类的saRNACovid-19疫苗的临床前研究中使用了大量saRNA的原因(Erasmus et al.,2020,Sci Transl Med 12:eabc9396)。

[0005] 甲病毒是正链RNA病毒的典型代表。甲病毒的宿主包括广泛的生物,包括昆虫、鱼和哺乳动物,如驯化动物和人。甲病毒在感染细胞的细胞质中复制(甲病毒生命周期的综述参见José et al.,2009,Future Microbiol.4:837-856)。许多甲病毒的总基因组长度通常范围在11,000和12,000个核苷酸之间,并且基因组RNA通常具有5'-帽和3'poly(A)尾。甲病毒的基因组编码非结构蛋白(参与病毒RNA的转录、修饰和复制以及蛋白修饰)和结构蛋白(形成病毒颗粒)。基因组中通常有两个开放阅读框(ORF)。4种非结构蛋白(nsP1-nsP4)通常由基因组5′末端附近开始的第一ORF一起编码,而甲病毒结构蛋白由第二ORF一起编码,所述第二ORF位于第一ORF下游并延伸至基因组的3'末端附近。通常,第一ORF比第二ORF大,比例大约为2:1。

[0006] 在被甲病毒感染的细胞中,只有非结构蛋白从基因组RNA翻译,而结构蛋白可从亚基因组转录物翻译,这是一种类似于真核信使RNA的RNA分子(mRNA;Gould et al.,2010,Antiviral Res.87:111-124)。感染之后,即在病毒生命周期的早期阶段,(+)链基因组RNA

直接充当信使RNA,用于翻译编码非结构多聚蛋白(nsP1234)的开放阅读框。在一些甲病毒中,在nsP3和nsP4的编码序列之间存在乳白型(opal)终止密码子:当翻译终止于opal终止密码子时,产生含有nsP1、nsP2和nsP3的多聚蛋白P123,而在通读这个opal密码子时,产生额外含有nsP4的多聚蛋白P1234(Strauss&Strauss,1994,Microbiol.Rev.58:491-562;Rupp et al.,2015,J.Gen.Virology 96:2483-2500)。nsP1234自蛋白水解切割为片段nsP123和nsP4。多肽nsP123和nsP4结合形成(-)链复制酶复合物,使用(+)链基因组RNA作为模板转录(-)链RNA。通常,在后期,nsP123片段被完全切割为单独的蛋白nsP1、nsP2和nsP3(Shirako&Strauss,1994,J.Virol.68:1874-1885)。所有四种蛋白组装形成(+)链复制酶复合物,使用基因组RNA的(-)链互补物作为模板合成新的(+)链基因组(Kim et al.,2004,Virology 323:153-163,Vasiljeva et al.,2003,J.Biol.Chem.278:41636-41645)。

[0007] 在感染的细胞中,亚基因组RNA以及新基因组RNA由nsP1提供5'-帽(Pettersson et al.,1980,Eur.J.Biochem.105:435-443;Rozanov et al.,1992,J.Gen.Virology 73:2129-2134),由nsP4提供多腺苷酸[poly(A)]尾(Rubach et al.,2009,Virology 384:201-208)。因此,亚基因组RNA和基因组RNA都类似于信使RNA(mRNA)。

[0008] 甲病毒结构蛋白(核心核衣壳蛋白C、包膜蛋白E2和包膜蛋白E1,病毒颗粒的所有组成成分)通常由亚基因组启动子控制的单个开放阅读框编码(Strauss&Strauss,1994,Microbiol.Rev.58:491-562)。亚基因组启动子被顺式作用的甲病毒非结构蛋白识别。特别地,甲病毒复制酶使用基因组RNA的(-)链互补物作为模板合成(+)链亚基因组转录物。(+)链亚基因组转录物编码甲病毒结构蛋白(Kim et al.,2004,Virology 323:153-163,Vasiljeva et al.,2003,J.Biol.Chem.278:41636-41645)。亚基因组RNA转录物作为模板,将编码结构蛋白的开放阅读框翻译为一个多聚蛋白,切割多聚蛋白以产生结构蛋白。在宿主细胞中甲病毒感染的后期,位于nsP2编码序列内的包装信号确保将基因组RNA选择性地包装为出芽病毒粒子,由结构蛋白包装(White et al.,1998,J.Virol.72:4320-4326)。

[0009] 在感染的细胞中,通常仅在感染后的前3-4h观察到(-)链RNA合成,而在后期检测不到,此时仅观察到(+)链RNA(基因组和亚基因组)的合成。根据Frolov et al.,2001,RNA 7:1638-1651,RNA合成调控的主流模型表明对非结构多聚蛋白加工的依赖性:非结构多聚蛋白nsP1234的初始切割产生nsP123和nsP4;nsP4作为RNA依赖性RNA聚合酶(RdRp),对(-)链合成有活性,但是对(+)链RNA的产生效率低下。多聚蛋白nsP123的进一步加工,包括在nsP2/nsP3连接处的切割,改变复制酶的模板特异性,以增加(+)链RNA的合成并减少或终止(-)链RNA的合成。

[0010] 甲病毒RNA的合成还受到顺式作用RNA元件的调控,包括四个保守序列元件(CSE;Strauss&Strauss,1994,Microbiol.Rev.58:491-562;和Frolov,2001,RNA 7:1638-1651)。甲病毒基因组包含四个保守序列元件(CSE),这些元件被认为对于宿主细胞中的病毒RNA复制很重要。在病毒基因组的5'末端或附近发现的CSE 1,被认为是从(-)链模版合成(+)链的启动子。位于CSE 1下游但仍靠近nsP1编码序列内基因组5'末端的CSE 2被认为是启动从基因组RNA模板合成(-)链的启动子(注意,不包含CSE 2的亚基因组RNA转录物并不作为(-)链合成的模板)。CSE 3位于非结构蛋白和结构蛋白编码序列之间的连接区域,是亚基因组转录物有效转录的核心启动子。最后,正好位于甲病毒基因组3'非翻译区中poly(A)序列上游的CSE 4被认为是启动(-)链合成的核心启动子(Joséet al.,2009,Future Microbiol.4:

837-856)。据了解,甲病毒的CSE 4和poly(A)共同发挥作用,以实现有效的(-)链合成(Hardy&Rice,2005,J.Virol.79:4630-4639)。除甲病毒蛋白外,宿主细胞因子(可能是蛋白)也可以与保守序列元件结合。甲病毒基因组的5'复制识别序列不仅参与翻译起始,还包含参与病毒RNA合成的两个保守序列元件CSE 1和CSE 2。对于CSE 1和2的功能,据信二级结构比线性序列更重要(Strauss&Strauss,1994,Microbiol.Rev.58:491-562)。

[0011] 已建议将甲病毒衍生的载体用于将外来遗传信息递送至靶细胞或靶生物体中。在简单的方法中,编码甲病毒结构蛋白的开放阅读框被编码所关注的蛋白的开放阅读框替换。基于甲病毒的反式复制系统依赖于两个单独的核酸分子上的甲病毒核苷酸序列元件:一个核酸分子编码病毒复制酶(通常为多聚蛋白nsP1234),而另一个核酸分子能够被所述复制酶反式复制(因此命名为反式复制系统)。反式复制需要在给定的宿主细胞中同时存在这些核酸分子。能够被复制酶反式复制的核酸分子必须包含某些甲病毒序列元件以允许甲病毒复制酶进行识别和RNA合成。

[0012] 鉴于修饰的mRNA疫苗的成功,修饰saRNA以减少先天免疫应答似乎很有吸引力。然而,已观察到修饰的核苷酸抑制saRNA的复制和翻译(Erasmus et al.,2020,Mol.Ther.Methods Clin.Dev.18:402-414)。迄今为止,还没有发表过为什么RNA修饰与saRNA功能不相容的研究。因此,尽管含有修饰的核苷酸,但本领域仍然需要具有充分复制和/或翻译功能的saRNA分子。本发明满足了这种需求。

发明内容

[0013] 本发明一般涉及提高自复制RNA分子的能力,所述自复制RNA分子也称为复制子或可复制(rRNA),其含有尿嘧啶、腺苷、胞嘧啶和鸟嘌呤以外的核苷酸,以复制和/或翻译来表达编码的蛋白。因此,本发明涉及含有修饰的核苷酸的可复制RNA分子,其中至少一部分被适当的RNA依赖性RNA聚合酶(复制酶)识别用于复制的区域不含有修饰的核苷酸,以及这类修饰的可复制RNA分子在方法中的用途,所述方法用于在细胞中表达蛋白,或者用于提高针对可复制RNA分子编码的蛋白的免疫应答,优选细胞毒性免疫应答,以及用于治疗或预防疾病或病症,其中这样的免疫应答产生/导致这种疾病或病症的治疗或预防。

[0014] 本发明部分基于以下假设:5'-或3'-保守序列元件(CSE)内的RNA二级结构在修饰其中所含核苷酸时改变它们的形状或稳定性。由于RNA复制取决于RNA模板与复制酶的相互作用,因此RNA与复制酶的不正确相互作用会严重影响RNA依赖性RNA转录(和/或翻译)。由于RNA结构取决于核苷酸序列,因此发明人提出,与含有例如1mΨ代替CSE 1中所有尿苷的CSE 1相比,从修饰的RNA的至少一部分中排除修饰的核苷酸可以使与复制酶正确相互作用的结构适应更有效的复制和/或翻译。如本文公开的实验结果所证明的,本发明人表明,除至少一部分CSE 1外含有修饰的核苷酸的可复制RNA导致修饰的可复制RNA的功能改善。

[0015] 在一方面,本发明涉及一种修饰的可复制RNA分子(rRNA),其包含甲病毒的5'调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中所述分子包含序列AUGGCGGA或AUGGGCGG,其中这两个序列AUGGCGGA或AUGGGCGG中的U是尿苷;并且其中所述分子中剩余尿苷中的至少一个是修饰的尿苷,优选N1-甲基-假尿苷(1mΨ)。

[0016] rRNA可以是单链RNA分子,可以被翻译(+-链),并且可以或不可以编码RNA依赖性RNA聚合酶(复制酶),但是包含核苷酸序列,其使得该分子能够由单独提供的复制酶反式复

制或由同一rRNA编码的复制酶顺式复制。rRNA分子也可以称为反式复制子或复制子。此外，rRNA可以包含编码基因序列的野生型或密码子优化的序列，例如，抗原或报告基因，并且rRNA可以包含一个或多个结构元件，所述结构元件针对rRNA在稳定性和翻译效率方面的最大效力进行了优化(5'帽、5'UTR、3'UTR、poly(A)-尾、茎环结构等)。

[0017] 在一实施方案中，rRNA可以包含第二ORF，所述第二ORF编码非结构蛋白1、2、3和4，优选为多聚蛋白(nsP1234)，其在表达和加工时形成RNA依赖性RNA聚合酶(复制酶)，所述复制酶优选为甲病毒复制酶。

[0018] rRNA基本上可以是单链、正链RNA病毒的基因组，例如，甲病毒(alphavirus)、小RNA病毒(picornavirus)、黄病毒(flavivirus)，其任选不编码病毒的功能性结构蛋白。复制酶可以源自自复制的RNA病毒，如甲病毒，例如，塞姆利基森林病毒(Semliki Forest virus)(SFV)、委内瑞拉马脑炎病毒(Venezuelan equine encephalitis virus)(VEEV)、辛德毕斯病毒(Sindbis virus)、东方马脑炎病毒(Eastern equine encephalitis virus)、西方马脑炎病毒(Western equine encephalitis virus)或基孔肯雅病毒(Chikungunya virus)。在一实施方案中，甲病毒可以是SFV、VEEV或辛德毕斯病毒。

[0019] 在一实施方案中，5'调节区可以是没有起始密码子的调节区。在一实施方案中，复制酶和复制酶复制所需的rRNA中的调节序列源自相同的甲病毒或不同的甲病毒。在rRNA还包含编码复制酶的开放阅读框的情况下，通过将复制酶开放阅读框的翻译置于内部核糖体进入位点(IRES)的翻译控制下，可以将复制酶开放阅读框的翻译与5'-末端帽解偶联(uncouple)。在一实施方案中，rRNA可以未加帽。在一实施方案中，rRNA可以具有开放阅读框，用于表达IRES上游的其他基因。

[0020] 在一实施方案中，所述分子中所有剩余尿苷都可以是1mΨ。在一实施方案中，所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的剩余尿苷可以是1mΨ。

[0021] 在一实施方案中，上述序列AUGGCGGA或AUGGGCGG可以位于分子的非编码区中，或者这些序列可以位于5'调节区中。在一实施方案中，序列AUGGCGGA或AUGGGCGG可以位于5'调节区中的保守序列元件1(CSE 1)中。在一实施方案中，序列AUGGCGGA或AUGGGCGG可以位于分子的5'末端。在一实施方案中，这些序列可以进一步包含AUGGCGGA或AUGGGCGG所描述序列5'的额外核苷酸，任选地，其中所述额外核苷酸包含额外的ORF和/或控制序列或者形成5'帽结构的一个或多个核苷酸。

[0022] 在一方面，本发明涉及一种修饰的可复制RNA分子，其包含甲病毒的5'调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF)，其中除了5'调节区中包含的保守序列元件1(CSE 1)的10个5'核苷酸内包含的尿苷外，所述分子中的尿苷中的至少一个尿苷是修饰的尿苷，优选N1-甲基-假尿苷(1mΨ)。在一实施方案中，除了CSE 1的10个5'核苷酸内包含的尿苷外，所述分子中的所有尿苷为1mΨ。在一实施方案中，除了CSE 1的5个5'核苷酸内包含的尿苷外，所述分子中的所有尿苷为1mΨ。在一实施方案中，除了CSE 1的4个5'核苷酸内包含的尿苷外，所述分子中的所有尿苷为1mΨ。在一实施方案中，除了CSE 1的3个5'核苷酸内包含的尿苷外，所述分子中的所有尿苷为1mΨ。在一实施方案中，除了CSE 1的2个5'核苷酸内包含的尿苷外，所述分子中的所有尿苷为1mΨ。在一实施方案中，除了CSE 1中最5'的尿苷外，所述分子中的所有尿苷为1mΨ。

[0023] 在一实施方案中,除了CSE 1中第2位的尿苷外,所述分子中的所有尿苷为1mΨ。

[0024] 在一实施方案中,其中rRNA包含5'帽,CSE 1的10个5'核苷酸包括5'帽的任何核苷酸。在一实施方案中,5'帽可以是G(5')ppp(5')AU。在一实施方案中,5'帽可以是m^7G(5')ppp(5')AU。

[0025] 在一实施方案中,rRNA可以包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,优选为多聚蛋白(nsP1234),其在表达和加工时形成RNA依赖性RNA聚合酶(复制酶),所述复制酶优选为甲病毒复制酶。

[0026] 在一方面,本发明涉及一种修饰的可复制RNA分子,其包含甲病毒的5'调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了5'调节区中包含的保守序列元件1(CSE 1)内包含的最5'的U外,所述分子中的至少一个尿苷是修饰的尿苷,优选N1-甲基-假尿苷(1mΨ)。在一实施方案中,除了CSE 1内包含的最5'的U外,所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的尿苷是1mΨ。在一实施方案中,除了CSE 1内包含的最5'的U外,所述分子中的所有尿苷为1mΨ。在一实施方案中,rRNA可以包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,优选为多聚蛋白(nsP1234),其在表达和加工时形成RNA依赖性RNA聚合酶(复制酶),所述复制酶优选为甲病毒复制酶。

[0027] 在一方面,本发明涉及一种5'加帽修饰的可复制RNA分子,其包含甲病毒的5'调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了所述分子中的第一个5'尿苷外,所述分子中的至少一个尿苷是修饰的尿苷,优选N1-甲基-假尿苷(1mΨ)。在一实施方案中,除了所述分子中的第一个5'尿苷外,所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的尿苷是1mΨ。在一实施方案中,除了所述分子中的第一个5'尿苷外,所述分子中的所有尿苷为1mΨ。在一实施方案中,rRNA可以包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,优选为多聚蛋白(nsP1234),其在表达和加工时形成RNA依赖性RNA聚合酶(复制酶),所述复制酶优选为甲病毒复制酶。

[0028] 在一方面,本发明涉及一种修饰的可复制RNA分子,其包含至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中所述分子中的至少一个尿苷是修饰的尿苷,优选N1-甲基-假尿苷(1mΨ),并且其中所述分子包含具有序列NpppNU的5'帽,其中所述5'帽中的U是未修饰的尿苷。此外,5'帽中的每个N可以是任何核苷酸,无论是修饰的或未修饰的,只要它不是修饰的尿苷,特别不是N1-甲基-假尿苷。在一实施方案中,5'帽具有序列NpppAU,其中A代表修饰或未修饰的腺苷核苷酸。例如,三磷酸键3'的修饰的核苷酸N或A可以具有修饰的核糖结构如2'-O-甲基化核糖(Nm或Am),产生所谓的"Cap 1"。相反,包含具有未甲基化核糖的三磷酸键3'的核苷酸N或A的帽通常被称作"Cap 0"。Nppp优选为鸟苷型核苷酸(Gppp),更优选为修饰的鸟苷核苷酸,如7-甲基-鸟核苷酸(m7Gppp)。在一实施方案中,所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的尿苷是1mΨ。

[0029] 在一实施方案中,所述分子包含第二ORF,所述第二ORF编码非结构蛋白1、2、3和4,其在表达和加工时形成RNA依赖性RNA聚合酶(复制酶),例如甲病毒复制酶。除了rRNA可以包含编码复制酶的第二ORF的实施方案之外,还有许多其他实施方案适用于本发明的任何

方面。例如,这样的实施方案是所述分子进一步包含至少一个修饰的G、C或A。其他示例性实施方案包括所述分子具有修饰的骨架,例如,其中所述修饰的骨架包含至少一个硫代磷酸酯键,或者其中所述骨架中的所有键是硫代磷酸酯键。

[0030] 其他实施方案包括rRNA中唯一的核苷酸或核苷修饰是1mΨ,或者甲病毒可以是SFV或VEEV。

[0031] 在一实施方案中,rRNA分子可以进一步包含一个或多个编码区,例如包含编码所关注基因的序列的一个或多个编码区。例如,编码的所关注基因可以是抗原或治疗性蛋白或核酸或报告基因。在一实施方案中,所述抗原是肿瘤、病毒、细菌或真菌抗原,或者过敏原。

[0032] 在一实施方案中,rRNA可以是稳定的rRNA。

[0033] 在一方面,本发明涉及一种编码本发明的rRNA分子的DNA分子。在一实施方案中,所述rRNA分子或DNA分子可以是线性或环状的。

[0034] 在一实施方案中,rRNA或DNA分子可以用能够与rRNA或DNA分子形成颗粒的试剂配制,例如,所述试剂可以是脂质或聚亚烷基亚胺(polyalkyleneimine)。在各种实施方案中,脂质可以包含阳离子头基,和/或脂质可以是pH响应性脂质,和/或脂质可以是聚乙二醇化脂质。在一实施方案中,所述试剂可以与聚肌氨酸缀合。在一实施方案中,由rRNA或DNA分子以及试剂形成的颗粒可以是基于聚合物的聚合复合物(PLX)或脂质纳米颗粒(LNP),其中所述LNP优选是脂质复合物(LPX)或脂质体。在一实施方案中,所述颗粒可以进一步包含至少一个磷脂酰丝氨酸。在一实施方案中,所述颗粒可以是纳米颗粒,其中(i)纳米颗粒中正电荷的数量不超过纳米颗粒中负电荷的数量,和/或(ii)纳米颗粒具有中性或净负电荷,和/或(iii)纳米颗粒中正电荷与负电荷的电荷比为1.4:1或更少,和/或(iv)纳米颗粒的zeta电位为0或更少。在一实施方案中,纳米颗粒中正电荷与负电荷的电荷比可以在1.4:1和1:8之间,优选在1.2:1和1:4之间。

[0035] 在一实施方案中,rRNA或DNA可以配制或待配制为液体、固体或其组合。在一实施方案中,rRNA或DNA可以配制或待配制用于注射。在一实施方案中,rRNA或DNA可以配制或待配制用于肌肉内给药。在一实施方案中,rRNA或DNA可以配制或待配制为颗粒。在一实施方案中,所述颗粒是脂质纳米颗粒(LNP)或脂质复合物(LPX)颗粒。在一实施方案中,LNP颗粒包含((4-羟基丁基)氮烷二基)双(己烷-6,1-二基)双(2-癸酸己酯)(((4-hydroxybutyl)azanediyl)bis(hexane-6,1-diyl)bis(2-hexyldecanoate))、2-[(聚乙二醇)-2000]-N,N-双十四烷基乙酰胺、1,2-二硬脂酰基-sn-甘油-3-磷酸胆碱和胆固醇。

[0036] 在一实施方案中,rRNA脂质复合物颗粒可以通过混合rRNA与脂质体来获得。在一实施方案中,rRNA脂质复合物颗粒可以通过混合rRNA与脂质来获得。

[0037] 在一实施方案中,rRNA可以配制或待配制为胶体。在一实施方案中,rRNA可以配制或待配制为颗粒,形成胶体的分散相。在一实施方案中,分散相中存在50%或更多、75%或更多或者85%或更多的rRNA。在一实施方案中,rRNA可以配制或待配制为包含rRNA和脂质的颗粒。在一实施方案中,可以通过使溶于水相中的rRNA与溶于有机相中的脂质接触形成颗粒。在一实施方案中,所述有机相可以包含乙醇。在一实施方案中,可以通过使溶于水相中的rRNA与分散于水相中的脂质接触形成颗粒。在一实施方案中,分散于水相中的脂质形成脂质体。

[0038] 在一方面,本发明涉及通过组合本发明的DNA分子以及包含DNA依赖性RNA聚合酶和修饰核苷酸的体外转录混合物,任选在帽的存在下,体外转录本发明的DNA分子。体外转录混合物还包含转录DNA以产生本发明的rRNA分子所必需的所有试剂。

[0039] 适用于任何方面的另一实施方案是所关注的基因编码抗原(肿瘤、病毒、细菌、真菌、过敏原)或治疗性蛋白或核酸,或者5’调节区和编码的RNA依赖性RNA聚合物可以源自相同的甲病毒或可以源自不同的甲病毒。另一个是5’调节区没有起始密码子,或者rRNA可以是体外转录的RNA分子。

[0040] 在一方面,本发明涉及一种药物组合物,其包含本文所述的本发明的rRNA以及药学可接受的载剂或赋形剂。

[0041] 在一方面,本发明涉及本发明的修饰的可复制RNA分子或包含这种rRNA的药物组合物,其用于治疗。

[0042] 在一方面,本发明涉及一种在受试者中提高免疫应答的方法,所述方法包括向受试者给药本发明的修饰的可复制RNA分子或包含这种rRNA的药物组合物。

[0043] 在一方面,本发明涉及一种治疗受试者的癌症的方法,所述方法包括向受试者给药本发明的修饰的可复制RNA分子或包含这种rRNA的药物组合物。

[0044] 在一方面,本发明涉及一种向缺乏基因功能的受试者提供这种基因功能的方法,所述方法包括向受试者给药本发明的修饰的可复制RNA分子或包含这种rRNA的药物组合物。

[0045] 本发明的方面内包括针对RNA分子群体的实施方案,优选可复制RNA分子群体,其中特定修饰的rRNA以群体中存在的所有RNA分子的百分比量存在于群体中。这些群体也可以用于本发明的方法,例如,用于提高免疫应答或治疗癌症的方法。

[0046] 在一实施方案中,rRNA分子群体包含修饰的可复制RNA分子,量为群体中存在的所有rRNA分子的至少1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%,所述修饰的可复制RNA分子是5’加帽修饰的可复制RNA分子,其包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了所述分子中的第一个5’尿苷外,所述分子中的至少一个尿苷是修饰的尿苷。

[0047] 在一实施方案中,rRNA分子群体包含修饰的可复制RNA分子,量为群体中存在的所有rRNA分子的至少1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%,所述修饰的可复制RNA分子包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中所述分子包含序列AUGGCGGA或AUGGGCGG,其中这些序列中的任一序列AUGGCGGA或AUGGGCGG中的U是尿苷;并且其中所述分子中剩余尿苷中的至少一个是修饰的尿苷。

[0048] 在一实施方案中,rRNA分子群体包含修饰的可复制RNA分子,量为群体中存在的所有rRNA分子的至少1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%,所述修饰的可复制RNA分子包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了5’调节区中包含的保守序列

元件1(CSE 1)的10个5’核苷酸内包含的尿苷外,所述分子中的尿苷中的至少一个尿苷是修饰的尿苷。

[0049] 在一实施方案中,rRNA分子群体包含修饰的可复制RNA分子,量为群体中存在的所有rRNA分子的至少1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%,所述修饰的可复制RNA分子包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了5’调节区中包含的保守序列元件1(CSE1)内包含的最5’的尿苷外,所述分子中的至少一个尿苷是修饰的尿苷。

[0050] 在一实施方案中,rRNA分子群体包含修饰的可复制RNA分子,量为群体中存在的所有rRNA分子的至少1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%,所述修饰的可复制RNA分子包含至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中所述分子中的至少一个尿苷是修饰的尿苷,并且其中所述分子包含具有序列NpppNU的5’帽,其中所述5’帽中的U是未修饰的尿苷,优选其中所述序列是NpppAU。

具体实施方式

[0051] 虽然下文详细描述本发明,但是应当理解本发明并不限于本文描述的特定方法、方案和试剂,因为这些可以变化。还应当理解本文使用的术语仅用于描述特定实施方案的目的而不是为了限制本发明的范围,本发明的范围仅受所附权利要求书的限制。除非另有定义,否则本文使用的所有技术和科学术语均具有与本领域普通技术人员通常理解的相同的含义。

[0052] 优选地,本文中使用的术语如“A multilingual glossary of biotechnological terms:(IUPAC Recommendations)”,H.G.W.Leuenberger,B.Nagel,and H.Kölbl,Eds.,Helvetica Chimica Acta,CH-4010Basel,Switzerland,(1995)所述定义。

[0053] 除非另有说明,否则本发明的实施会采用化学、生物化学、细胞生物学、免疫学和重组DNA技术的常规方法,其在所述领域的文献中解释(参见,例如,Molecular Cloning:A Laboratory Manual,2nd Edition,J.Sambrook et al.eds.,Cold Spring Harbor Laboratory Press,Cold Spring Harbor 1989)。

[0054] 在下文中会描述本发明的元素。这些元素用具体实施方案列出,然而,应当理解它们可以以任何方式和任何数量组合以产生另外的实施方案。不同地描述的实例和优选实施方案不应当理解为将本发明仅限于明确描述的实施方案。本说明书应当理解为公开并涵盖组合明确描述的实施方案与任何数量的公开和/或优选的元素的实施方案。此外,除非上下文另有指示,否则本申请中所有描述的元素的任何排列和组合都应当视为由本说明书公开。

[0055] 术语“约”表示近似或接近,并且在本文示出的数值或范围的上下文中优选表示列举或声称的数值或范围的+/-10%。

[0056] 在描述本发明的上下文中(特别是在权利要求的上下文中)使用的术语“一个”和“一种”和“这个”以及相似指称应当理解为覆盖单数和复数,除非在本文中另有指明或与上

下文明显矛盾。本文中值的范围的列举仅为了用作单独提到落在所述范围内的每个不同值的简写方法。除非本文另有说明,否则每个单独的值如其在本文中单独列举地加入本说明书。除非本文另有说明或其他地方显然违背上下文,否则本文描述的所有方法可以以任何合适的顺序进行。本文提供的任何和所有实例或者示例性语言(例如,“如”)的使用仅为了更好地说明本发明,并不对另外要求保护的本发明的范围构成限制。本说明书中没有语言应当理解为表示对于本发明的实施必不可少的任何未要求保护的元素。

[0057] 除非另有明确说明,否则在本文件的上下文中术语“包含”用来表示除了由“包含”引入的列表成员,还可以任选地存在其他成员。但是,作为本发明的具体实施方案,考虑术语“包含”涵盖不存在其他成员的可能性,即,为了这个实施方案的目的,“包含”理解为具有“由…组成”的含义。

[0058] 以通用术语为特征的组分相对量的表示是指所述通用术语涵盖的所有特定变体或成员的总量。如果指定由通用术语定义的某一组分以一定的相对量存在,并且如果该组分进一步表征为通用术语涵盖的特定变体或成员,则意味着没有另外存在通用术语涵盖的其他变体或成员,使得通用术语涵盖的组分的总相对量超过指定的相对量;更优选根本不存在通用术语涵盖的其他变体或成员。

[0059] 在这个说明书的整个正文中引用了几个文件。本文引用的每个文件(包括所有专利、专利申请、科学出版物、制造商的说明、指导等),无论上文或下文,均整体援引加入本文。本文中的任何内容均不应理解为承认本发明无权先于这样的公开。

[0060] 如本文所用的术语如“减少”或“抑制”表示引起水平总体减少优选5%或更大,10%或更大,20%或更大,更优选50%或更大,并且最优选75%或更大的能力。术语“抑制”或相似短语包括完全或基本上完全的抑制,即减少至0或基本上至0。

[0061] 术语如“增加”或“增强”优选涉及增加或增强大约至少10%,优选至少20%,优选至少30%,更优选至少40%,更优选至少50%,甚至更优选至少80%,并且最优选至少100%。

[0062] 术语“净电荷”是指整个物体上的电荷,如化合物或颗粒。

[0063] 具有总净正电荷的离子是阳离子,而具有总净负电荷的离子是阴离子。因此,根据本发明,阴离子是比质子具有更多电子的离子,赋予其净负电荷;而阳离子是比质子具有更少电子的离子,赋予其净正电荷。

[0064] 关于给定化合物或颗粒,术语“带电荷的”、“净电荷”、“带负电荷的”或“带正电荷的”是指给定化合物或颗粒在pH 7.0的水中溶解或悬浮时的净电荷。

[0065] 本发明的术语“核酸”还包括核酸在核苷酸碱基、糖或磷酸酯上的化学衍生,以及含有非天然核苷酸和核苷酸类似物的核酸。在一些实施方案中,核酸是脱氧核糖核酸(DNA)或核糖核酸(RNA)。一般来说,核酸分子或核酸序列是指核酸,其优选是脱氧核糖核酸(DNA)或核糖核酸(RNA)。根据本发明,核酸包含基因组DNA、cDNA、mRNA、病毒RNA、重组制备和化学合成的分子。根据本发明,核酸可以以单链或双链以及线性或共价闭合环状分子的形式存在。

[0066] 根据本发明,“核酸序列”是指核酸中的核苷酸序列,例如,核糖核酸(RNA)或脱氧核糖核酸(DNA)。该术语可以指整个核酸分子(如整个核酸分子的单链)或其部分(例如片段)。

[0067] 根据本发明,术语“RNA”或“RNA分子”涉及包含核糖核苷酸残基并且优选完全或基本上由核糖核苷酸残基组成的分子。术语“核糖核苷酸”涉及在β-D-呋喃核糖基的2’-位具有羟基的核苷酸。术语“RNA”包括双链RNA、单链RNA、分离的RNA如部分或完全纯化的RNA、基本上纯的RNA、合成的RNA以及重组产生的RNA,如修饰的RNA,其通过添加、缺失、替代和/或改变一个或多个核苷酸不同于天然存在的RNA。这类改变可以包括添加非核苷酸物质,如添加至RNA的末端或内部,例如在RNA的一个或多个核苷酸。RNA分子中的核苷酸还可以包含非标准核苷酸,如非天然存在的核苷酸或化学合成的核苷酸或脱氧核苷酸。这些改变的RNA可以称作类似物,特别是天然存在的RNA的类似物。

[0068] 根据本发明,RNA可以是单链或双链的。在本发明的一些实施方案中,优选单链RNA。术语“单链RNA”一般是指没有互补核酸分子(通常没有互补RNA分子)与之结合的RNA分子。单链RNA可以含有自互补序列,其允许部分RNA折叠回来并形成二级结构基序,包括但不限于碱基对、茎、茎环和凸起。单链RNA可以作为负链[(-)链]或正链[(+)链]存在。(+)链是包含或编码遗传信息的链。遗传信息可以是例如编码蛋白的多核苷酸序列。当(+)链RNA编码蛋白时,(+)链可以直接用作翻译(蛋白合成)的模板。(-)链是(+)链的互补物。在双链RNA的情况下,(+)链和(-)链是两个独立的RNA分子,这两个RNA分子相互结合形成双链RNA(“双链体RNA”)。

[0069] RNA的术语“稳定性”涉及RNA的“半衰期”。“半衰期”涉及消除分子的一半活性、量或数量所需要的时间。在本发明的上下文中,RNA的半衰期是所述RNA稳定性的指示。RNA的半衰期可以影响RNA的“表达持续时间”。可以预期具有长半衰期的RNA会在延长的时间内表达。

[0070] 术语“翻译效率”涉及RNA分子在特定时间内提供的翻译产物的量。

[0071] 关于核酸序列,“片段”涉及核酸序列的一部分,即代表在5’-和/或3’-末端缩短的核酸序列的序列。优选地,核酸序列的片段包含来自所述核酸序列的至少80%,优选至少90%、95%、96%、97%、98%或99%的核苷酸残基。在本发明中,优选那些保留RNA稳定性和/或翻译效率的RNA分子片段。

[0072] 关于氨基酸序列(肽或蛋白),“片段”涉及氨基酸序列的一部分,即代表在N-末端和/或C-末端缩短的氨基酸序列的序列。在C-末端缩短的片段(N-末端片段)例如通过缺少开放阅读框3’-末端的截短开放阅读框的翻译可获得。在N-末端缩短的片段(C-末端片段)例如通过缺少开放阅读框5’-末端的截短开放阅读框的翻译可获得,只要截短开放阅读框包含用来起始翻译的起始密码子。氨基酸序列的片段包含例如来自氨基酸序列的至少1%、至少2%、至少3%、至少4%、至少5%、至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%的氨基酸残基。

[0073] 根据本发明,关于例如核酸和氨基酸序列的术语“变体”包括任何变体,特别是突变体、病毒毒株变体、剪接变体、构象、同种型、等位基因变体、物种变体和物种同源物,特别是天然存在的那些。等位基因变体涉及基因正常序列中的改变,其意义通常不清楚。完整基因测序常鉴定给定基因的许多等位基因变体。关于核酸分子,术语“变体”包括简并核酸序列,其中本发明的简并核酸是由于遗传密码的简并性而在密码子序列上不同于参考核酸的核酸。物种同源物是与给定核酸或氨基酸序列具有不同物种来源的核酸或氨基酸序列。病毒同源物是与给定核酸或氨基酸序列具有不同病毒来源的核酸或氨基酸序列。

[0074] 根据本发明,与参考核酸相比,核酸变体包括单个或多个核苷酸缺失、添加、突变、替代和/或插入。缺失包括从参考核酸去除一个或多个核苷酸。添加变体包含一个或多个核苷酸的5'-和/或3'-末端融合,如1、2、3、5、10、20、30、50或更多个核苷酸。在替代的情况下,去除序列中的至少一个核苷酸,并且在其位置插入至少一个其他核苷酸(如颠换和转换)。突变包括无碱基位点、交联位点(crosslinked site)以及化学改变或修饰的碱基。插入包括将至少一个核苷酸添加到参考核酸中。

[0075] 根据本发明,"核苷酸改变"可以是指与参考核酸相比的单个或多个核苷酸缺失、添加、突变、替代和/或插入。在一些实施方案中,"核苷酸改变"选自与参考核酸相比的单核苷酸的缺失、单核苷酸的添加、单核苷酸的突变、单核苷酸的替代和/或单核苷酸的插入。根据本发明,与参考核酸相比,核酸变体可以包含一个或多个核苷酸改变。

[0076] 特定核酸序列的变体优选具有所述特定序列的至少一种功能特性,并且优选在功能上等同于所述特定序列,例如,表现出与特定核酸序列相同或相似特性的核酸序列。

[0077] 如下所述,本发明的一些实施方案的特征尤其在于与其他核酸序列同源的核酸序列。这些同源序列是其他核酸序列的变体。

[0078] 优选地,给定核酸序列与作为所述给定核酸序列变体的核酸序列之间的相同性程度为至少70%,优选至少75%,优选至少80%,更优选至少85%,甚至更优选至少90%或者最优选至少95%、96%、97%、98%或99%。优选对至少约30、至少约50、至少约70、至少约90、至少约100、至少约150、至少约200、至少约250、至少约300或至少约400个核苷酸的区域给出相同性程度。在优选的实施方案中,对参考核酸序列的整个长度给出相同性程度。

[0079] "序列相似性"表示相同或代表保守氨基酸取代的氨基酸的百分比。两个多肽或核酸序列之间的"序列相同性"表示所述序列之间相同的氨基酸或核苷酸的百分比。

[0080] 术语"%相同"特别是指在待比较的两个序列之间的最佳比对中相同的核苷酸的百分比,所述百分比纯粹是统计的,两个序列之间的差异可以随机分布在序列的整个长度上,并且待比较的序列可以包含与参考序列相比的添加或缺失,以获得两个序列之间的最佳比对。通常通过在对片段或"比较窗口"最佳比对之后,比较所述序列来进行两个序列的比较,以便鉴定相应序列的局部区域。用于比较的最佳比对可以手动进行,或者借助Smith and Waterman,1981,Ads App.Math.2:482的局部同源算法,借助Needleman and Wunsch,1970,J.Mol.Biol.48:443的局部同源算法,借助Pearson and Lipman,1988,Proc.Natl Acad.Sci.USA 85:2444的相似性搜索算法或借助使用所述算法的计算机程序(Wisconsin Genetics Software Package,Genetics Computer Group,575Science Drive,Madison,Wis.中的GAP、BESTFIT、FASTA、BLAST P、BLAST N和TFASTA)。

[0081] 通过确定待比较的序列对应的相同位置的数目,用这个数目除以比较的位置数,并将这个结果乘以100,获得百分比相同性。

[0082] 例如,可以使用在网站http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi上可获得的BLAST程序"BLAST 2sequences"。

[0083] 如果两个序列彼此互补,则核酸与另一核酸"能够杂交"或"杂交"。如果两个序列能够彼此形成稳定的双链体,则核酸与另一核酸"互补"。根据本发明,优选在允许多核苷酸之间特异性杂交的条件(严格条件)下进行杂交。例如,Molecular Cloning:A Laboratory Manual,J.Sambrook et al.,Editors,2nd Edition,Cold Spring Harbor Laboratory

press,Cold Spring Harbor,New York,1989或Current Protocols in Molecular Biology,F.M.Ausubel et al.,Editors,John Wiley&Sons,Inc.,New York中描述了严格条件,例如,在杂交缓冲液(3.5x SSC、0.02% Ficoll、0.02%聚乙烯吡咯烷酮、0.02%牛血清白蛋白、2.5mM NaH2PO4(pH 7)、0.5% SDS、2mM EDTA)中于65℃下杂交。SSC是0.15M氯化钠/0.15M柠檬酸钠,pH 7。杂交之后,将已转移DNA的膜洗涤,例如,在室温下以2×SSC洗涤,然后在高达68℃的温度下以0.1-0.5×SSC/0.1×SDS洗涤。

[0084] 百分比互补性表示可以与第二核酸序列形成氢键(例如,Watson-Crick碱基配对)的核酸分子中连续残基的百分比(例如,10个中5、6、7、8、9、10个是50%、60%、70%、80%、90%和100%互补)。“完美互补”或“完全互补”表示核酸序列的所有连续残基会与第二核酸序列中相同数量的连续残基氢键结合。优选地,本发明的互补程度是至少70%,优选至少75%,优选至少80%,更优选至少85%,甚至更优选至少90%,或者最优选至少95%、96%、97%、98%或99%。最优选地,本发明的互补程度是100%。

[0085] 术语“衍生物”包括在核苷酸碱基上、在糖上或在磷酸酯上的核酸的任何化学衍生。术语“衍生物”还包括含有非天然存在的核苷酸和核苷酸类似物的核酸。优选地,核酸的衍生增加其稳定性。

[0086] “衍生自核酸序列的核酸序列”是指作为其衍生核酸的变体的核酸。优选地,当代替RNA分子中的特定序列时,相对于特定序列是变体的序列保留RNA稳定性和/或翻译效率。

[0087] “nt”是核苷酸的缩写;或者对于核苷酸,优选核酸分子中的连续核苷酸。

[0088] 根据本发明,术语“密码子”是指编码核酸中的碱基三联体,其指定在核糖体的蛋白合成过程中接下来会添加哪种氨基酸。

[0089] 术语“转录(transcription)”和“转录(transcribing)”涉及RNA聚合酶读取具有特定核酸序列的核酸分子(“核酸模板”)以使RNA聚合酶产生单链RNA分子的过程。在转录过程中,转录核酸模板中的遗传信息。核酸模板可以是DNA;然而,例如,在从甲病毒核酸模板转录的情况下,模板通常是RNA。随后,可以将转录的RNA翻译为蛋白。根据本发明,术语“转录”包括“体外转录”,其中术语“体外转录”涉及一个过程,其中RNA,特别是mRNA,在不含细胞的系统中体外合成。优选地,将克隆载体用于产生转录物。这些克隆载体一般指定为转录载体,并且根据本发明涵盖在术语“载体”内。克隆载体优选是质粒。根据本发明,RNA优选是体外转录的RNA(IVT-RNA),并且可以通过适当DNA模板的体外转录获得。用于控制转录的启动子可以是任何RNA聚合酶的任何启动子。用于体外转录的DNA模板可以通过克隆核酸,特别是cDNA,并将其引入适当的载体用于体外转录来获得。cDNA可以通过RNA的逆转录获得。

[0090] 转录过程中产生的单链核酸分子通常具有作为模板互补序列的核酸序列。

[0091] 根据本发明,术语“模板”或“核酸模板”或“模板核酸”一般是指可以复制或转录的核酸序列。

[0092] “转录自核酸序列的核酸序列”和类似术语是指核酸序列,在适当情况下作为完整RNA分子的一部分,其是模板核酸序列的转录产物。通常,转录的核酸序列是单链RNA分子。

[0093] 根据本发明,“核酸的3’末端”是指具有游离羟基的末端。在双链核酸,特别是DNA的图示中,3’末端总是在右侧。根据本发明,“核酸的5’末端”是指具有游离磷酸基团的末端。在双链核酸,特别是DNA的图示中,5’末端总是在左侧。

[0094] 5’末端5’--P-NNNNNNN-OH-3’3’末端

[0095] 3'-HO-NNNNNNN-P--5'

[0096] "上游"描述了核酸分子的第一元件相对于该核酸分子的第二元件的相对定位,其中两个元件都包含在同一核酸分子中,并且其中第一元件比该核酸分子的第二元件更靠近核酸分子的5'末端。第二元件则被称为该核酸分子的第一元件的"下游"。位于第二元件"上游"的元件可以同义地称作位于该第二元件的"5'"。对于双链核酸分子,相对于(+)链给出类似"上游"和"下游"的指示。

[0097] 根据本发明,"功能连接"或"功能连接的"涉及功能关系内的连接。如果核酸与另一核酸序列在功能上相关,则它是"功能连接的"。例如,如果启动子影响所述编码序列的转录,则启动子与编码序列是功能连接的。功能连接的核酸通常彼此相邻,在适当的地方被其他核酸序列分开,并且在特定实施方案中,由RNA聚合酶转录得到单个RNA分子(共同转录物)。

[0098] 在特定实施方案中,根据本发明,将核酸与表达控制序列功能连接,所述表达控制序列可以与所述核酸同源或异源。

[0099] 根据本发明,术语"表达控制序列"包含启动子、核糖体结合序列以及控制基因的转录或衍生RNA的翻译的其他控制元件。在本发明的特定实施方案中,可以调节表达控制序列。表达控制序列的精确结构可能因物种或细胞类型而异,但是通常包括分别参与起始转录和翻译的5'-非转录序列以及5'-和3'-非翻译序列。更具体地,5'-非转录表达控制序列包含启动子区,其涵盖用于功能连接的基因的转录控制的启动子序列。表达控制序列还可以包括增强子序列或上游激活子序列。DNA分子的表达控制序列通常包括5'-非转录序列以及5'-和3'-非翻译序列如TATA盒、加帽序列、CAAT序列等。甲病毒RNA的表达控制序列可以包括亚基因组启动子和/或一个或多个保守序列元件。本发明的特定的表达控制序列是甲病毒的亚基因组启动子,如本文所述。

[0100] 本文指定的核酸序列,特别是可转录和编码的核酸序列,可以与任何表达控制序列,特别是启动子组合,这些序列可以与所述核酸序列同源或异源,术语"同源"是指核酸序列也与表达控制序列天然功能连接,而术语"异源"是指核酸序列与表达控制序列不天然功能连接。

[0101] 如果可转录核酸序列,特别是编码肽或蛋白的核酸序列,与表达控制序列彼此共价连接,从而使可转录核酸序列,特别是编码核酸序列的转录或表达受表达控制序列的控制或影响,则它们彼此"功能"连接。如果核酸序列翻译为功能性肽或蛋白,则诱导与编码序列功能连接的表达控制序列会导致所述编码序列的转录,而不会导致编码序列中的移码或编码序列不能翻译为期望的肽或蛋白。

[0102] 术语"启动子"或"启动子区"是指通过为RNA聚合酶提供识别和结合位点来控制转录物(例如包含编码序列的转录物)合成的核酸序列。启动子区可以包括参与调节所述基因转录的其他因子的其他识别或结合位点。启动子可以控制原核或真核基因的转录。启动子可以是"诱导型"的,对诱导物响应启动转录,或者可以是"组成型"的,如果转录不受诱导物的控制。如果没有诱导物,则诱导型启动子仅表达很小程度或根本不表达。在诱导物存在的情况下,基因被"启动(switch on)"或转录水平增加。这通常是由特定转录因子的结合介导的。本发明的特异性启动子是例如甲病毒的亚基因组启动子,如本文所述。其他特异性启动子是例如甲病毒的正链或负链启动子。

[0103] 术语“核心启动子”是指启动子包含的核酸序列。核心启动子通常是正确启动转录所需的启动子的最小部分。核心启动子通常包括转录起始位点和RNA聚合酶的结合位点。

[0104] “聚合酶”一般是指能够催化从单体构建模块合成聚合分子的分子实体。“RNA聚合酶”是能够催化从核糖核苷酸构建模块合成RNA分子的分子实体。“DNA聚合酶”是能够催化从脱氧核糖核苷酸构建模块合成DNA分子的分子实体。对于DNA聚合酶和RNA聚合酶的情况，所述分子实体通常是蛋白或者多种蛋白的组装体或复合物。通常，DNA聚合酶基于模板核酸合成DNA分子，模板核酸通常是DNA分子。通常，RNA聚合酶基于模板核酸合成RNA分子，模版核酸是DNA分子(在这种情况下，RNA聚合酶是DNA依赖性RNA聚合酶，DdRP)，或者是RNA分子(在这种情况下，RNA聚合酶是RNA依赖性RNA聚合酶，RdRP)。

[0105] “RNA依赖性RNA聚合酶”或“RdRP”或“复制酶”是一种催化从RNA模板转录RNA的酶。在甲病毒RNA依赖性RNA聚合酶的情况下，基因组RNA的(-)链互补物和(+)链基因组RNA的顺序合成导致RNA复制。因此RNA依赖性RNA聚合酶同义地称为“RNA复制酶”或简称为“复制酶”。在自然界中，RNA依赖性RNA聚合酶通常由除逆转录病毒外的所有RNA病毒编码。编码RNA依赖性RNA聚合酶的病毒的典型代表是甲病毒。

[0106] 根据本发明，“RNA复制”一般是指基于给定RNA分子(模板RNA分子)的核苷酸序列合成的RNA分子。合成的RNA分子可以例如与模板RNA分子相同或互补。一般来说，RNA复制可以通过DNA中间体的合成发生，也可以直接通过RNA依赖性RNA聚合酶(RdRP)介导的RNA依赖性RNA复制发生。在甲病毒的情况下，RNA复制不是通过DNA中间体发生的，而是由RNA依赖性RNA聚合酶(RdRP)介导的：模板RNA链(第一RNA链)-或其一部分-用作合成第二RNA链的模板，第二RNA链与第一RNA链或其一部分互补。第二RNA链-或其一部分-又可以任选地用作合成第三RNA链的模板，第三RNA链与第二RNA链或其一部分互补。因此，第三RNA链与第一RNA链或其一部分相同。因此，RNA依赖性RNA聚合酶能够直接合成模板的互补RNA链，以及间接合成相同的RNA链(通过互补的中间链)。

[0107] 根据本发明，术语“模板RNA”是指可以被RNA依赖性RNA聚合酶转录或复制的RNA。

[0108] 根据本发明，术语“基因”是指负责产生一种或多种细胞产物和/或实现一种或多种细胞间或细胞内功能的特定核酸序列。更具体地，所述术语涉及核酸部分(通常是DNA；但是在RNA病毒的情况下是RNA)，其包含编码特定蛋白或者功能或结构RNA分子的核酸。

[0109] 如本文所用的“分离的分子”是指基本上不含其它分子如其它细胞物质的分子。术语“分离的核酸”根据本发明表示所述核酸是：(i)体外扩增的，例如通过聚合酶链反应(PCR)，(ii)通过克隆重组产生的，(iii)纯化的，例如通过切割和凝胶电泳分级分离，或者(iv)合成的，例如通过化学合成。分离的核酸是可用于重组技术操作的核酸。

[0110] 术语“载体”在本文中以其最一般的含义使用，包括用于核酸的任何中间媒介物，例如，其使得能够将所述核酸引入原核和/或真核宿主细胞，并且在适当情况下能够整合至基因组中。这类载体优选在细胞中复制和/或表达。载体包括质粒、噬粒、病毒基因组及其部分。

[0111] 在本发明的上下文中术语“重组”表示通过“遗传工程化制备”。优选地，在本发明的上下文中“重组物体”如重组细胞不是天然存在的。

[0112] 如本文所用的术语“天然存在”是指物体可以在自然中发现这一事实。例如，存在于生物体(包括病毒)中且可以分离自自然来源并且尚未被人在实验室中有意修饰的肽或

核酸是天然存在的。术语“在自然界中发现”表示“存在于自然界中”,包括已知的物体以及尚未从自然界发现和/或分离,但是将来可能从自然来源发现和/或分离的物体。

[0113] 根据本发明,术语“表达”以其最一般的含义使用,包括RNA和/或蛋白的产生。它还包括核酸的部分表达。此外,表达可以是瞬时的或稳定的。关于RNA,术语“表达”或“翻译”涉及细胞核糖体中的过程,通过这个过程编码RNA(例如信使RNA)链指导氨基酸序列的组装以产生肽或蛋白。

[0114] 根据本发明,术语“mRNA”表示“信使-RNA”,并且涉及通常通过使用DNA模板产生并编码肽或蛋白的转录物。通常,mRNA包含5’-UTR、蛋白编码区、3’-UTR和poly(A)序列。mRNA可以通过从DNA模板体外转录产生。体外转录方法是技术人员已知的。例如,有各种体外转录试剂盒可商购。根据本发明,mRNA可以通过稳定修饰和加帽进行修饰。

[0115] 如本文所用,术语“poly(A)序列”或“poly(A)尾”是指腺苷酸残基的不间断或间断序列,其通常位于RNA分子的3’末端。不间断序列的特征在于连续的腺苷酸残基。在自然界中,不间断的poly(A)序列是典型的。虽然poly(A)序列通常不在真核DNA中编码,但是在转录后通过不依赖模板的RNA聚合酶在细胞核中的真核转录过程中连接至RNA的游离3’末端,本发明涵盖DNA编码的poly(A)序列。

[0116] 根据本发明,关于核酸分子,术语“一级结构”是指核苷酸单体的线性序列。

[0117] 根据本发明,关于核酸分子,术语“二级结构”是指反映碱基配对的核酸分子的二维表示;例如,在单链RNA分子的情况下,特别是分子内碱基配对。尽管每个RNA分子只有一条多核苷酸链,但是该分子通常特征在于(分子内)碱基对的区域。根据本发明,术语“二级结构”包括结构基序,其包括但不限于碱基对、茎、茎环、凸起、环如内环和多分支环。核酸分子的二级结构可以用二维图(平面图)表示,显示碱基配对(关于RNA分子二级结构的更多详细信息,参见Auber et al.,2006;J.Graph Algorithms Appl.10:329-351)。如本文所用,某些RNA分子的二级结构与本发明有关。

[0118] 根据本发明,核酸分子,特别是单链RNA分子的二级结构是通过使用RNA二级结构预测的网络服务器进行预测来确定的(http://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html)。优选地,根据本发明,关于核酸分子,“二级结构”具体是指由所述预测确定的二级结构。还可以使用MFOLD结构预测(http://unafold.rna.albany.edu/?q=mfold)来进行或确认预测。

[0119] 根据本发明,“碱基对”是二价结构的结构基序,其中两个核苷酸碱基通过碱基上供体和受体位点之间的氢键相互结合。互补碱基A:U和G:C通过碱基上供体和受体位点之间的氢键形成稳定的碱基对;A:U和G:C碱基对称为Watson-Crick碱基对。由碱基G和U(G:U)形成较弱的碱基对(称为Wobble碱基对)。碱基对A:U和G:C称为规范碱基对。其他碱基对如G:U(在RNA中经常出现)和其他罕见碱基对(例如A:C;U:U)称为非规范碱基对。

[0120] 根据本发明,“核苷酸配对”是指两个相互结合的核苷酸,它们的碱基从而形成碱基对(规范或非规范碱基对,优选规范碱基对,最优选Watson-Crick碱基对)。

[0121] 根据本发明,关于核酸分子,术语“茎环”或“发夹”或“发夹环”都可以互换地指核酸分子的特定二级结构,通常是单链核酸分子,如单链RNA。茎环所代表的特定二级结构由包含茎和(末端)环的连续核酸序列组成,也称为发夹环,其中茎由两个相邻的完全或部分互补的序列元件形成;由短序列(例如3-10个核苷酸)分开,形成茎-环结构的环。两个相邻

的完全或部分互补的序列可以定义为例如茎环元件茎1和茎2。当这两个相邻的完全或部分反向互补的序列,例如茎环元件茎1和茎2,彼此形成碱基对时,形成茎环,导致双链核酸序列,在其末端包含由位于茎环元件茎1和茎2之间的短序列形成的未配对环。因此,茎环包含两个茎(茎1和茎2),它们在核酸分子的二价结构水平上彼此形成碱基对,而在核酸分子的一级结构水平上,它们被不属于茎1或茎2的短序列分开。为了说明,茎环的二维表示类似于棒棒糖形结构。茎-环结构的形成需要存在可以折叠回自身以形成成对双链的序列;成对的双链由茎1和茎2形成。成对茎环元件的稳定性通常由长度决定,能够与茎2的核苷酸形成碱基对(优选规范碱基对,更优选Watson-Crick碱基对)的茎1的核苷酸数量,与不能与茎2的核苷酸形成这种碱基对(错配或凸起)的茎1的核苷酸数量。根据本发明,最佳环长度为3-10个核苷酸,更优选4-7个核苷酸,如4个核苷酸、5个核苷酸、6个核苷酸或7个核苷酸。如果给定核酸序列的特征在于茎环,则相应的互补核酸序列通常也特征在于茎环。茎环通常由单链RNA分子形成。例如,甲病毒基因组RNA的5'复制识别序列中存在几个茎环。

[0122] 根据本发明,关于核酸分子的特定二级结构(例如,茎环),"破坏(disruption)"或"破坏(disrupt)"表示所述特定二级结构不存在或改变。通常,二级结构可能因为作为二级结构一部分的至少一个核苷酸的改变的结果而被破坏。例如,茎环可能因为形成茎的一个或多个核苷酸的变化而被破坏,从而不可能进行核苷酸配对。

[0123] 根据本发明,"二级结构破坏的补偿"或"补偿二级结构破坏"是指核酸序列中的一个或多个核苷酸变化;更典型地,它是指核酸序列中的一个或多个第二核苷酸变化,所述核酸序列还包含一个或多个第一核苷酸变化,其特征如下:虽然在不存在一个或多个第二核苷酸变化的情况下,一个或多个第一核苷酸变化导致核酸序列二级结构的破坏,但是一个或多个第一核苷酸变化以及一个或多个第二核苷酸变化的共同出现不会导致核酸的二价结构被破坏。共同出现表示同时存在一个或多个第一核苷酸变化以及一个或多个第二核苷酸变化。通常,一个或多个第一核苷酸变化以及一个或多个第二核苷酸变化一起存在于同一核酸分子中。在一具体实施方案中,补偿二级结构破坏的一个或多个核苷酸变化是补偿一个或多个核苷酸配对破坏的一个或多个核苷酸变化。因此,在一实施方案中,"补偿二级结构破坏"表示"补偿核苷酸配对破坏",即一个或多个核苷酸配对破坏,例如一个或多个茎环内的一个或多个核苷酸配对破坏。可以通过去除至少一个起始密码子来引入一个或多个核苷酸配对破坏。补偿二级结构破坏的一个或多个核苷酸变化中的每一个都是核苷酸变化,可以独立地选自一个或多个核苷酸的缺失、添加、替代和/或插入。在一说明性实例中,当核苷酸配对A:U已被C替代A破坏时(C和U通常不适合形成核苷酸对);那么补偿核苷酸配对破坏的核苷酸变化可以是G替代U,从而使得能够形成C:G核苷酸配对。因此G替代U可以补偿核苷酸配对破坏。在一可选的实例中,当核苷酸配对A:U已被C替代A破坏时;那么补偿核苷酸配对破坏的核苷酸变化可以是A替代C,从而恢复形成原来的A:U核苷酸配对。一般来说,在本发明中,优选那些补偿二级结构破坏的核苷酸变化,其既不恢复原始的核酸序列,也不产生新的AUG三联体。在上面的一组实例中,U到G的替代优于C到A的替代。

[0124] 根据本发明,关于核酸分子,术语"三级结构"是指由原子坐标定义的核酸分子的三维结构。

[0125] 根据本发明,核酸如RNA,例如rRNA,可以编码肽或蛋白。因此,可转录核酸序列或其转录物可以含有编码肽或蛋白的开放阅读框(ORF)。

[0126] 根据本发明,术语“编码肽或蛋白的核酸”表示如果存在于适当环境中,优选在细胞内,核酸可以在翻译过程中指导氨基酸的组装以产生肽或蛋白。优选地,根据本发明的编码RNA能够与细胞翻译机制相互作用,允许编码RNA的翻译以产生肽或蛋白。

[0127] 根据本发明,术语“肽”包含寡肽和多肽,是指包含通过肽键相互连接的2个或更多个、优选3个或更多个、优选4个或更多个、优选6个或更多个、优选8个或更多个、优选10个或更多个、优选13个或更多个、优选16个或更多个、优选20个或更多个以及多达优选50个、优选100个或优选150个连续氨基酸的物质。术语“蛋白”是指大的肽,优选具有至少约151个氨基酸的肽,但是术语“肽”和“蛋白”在本文中通常用作同义词。

[0128] 根据本发明,术语“肽”和“蛋白”包括不仅含有氨基酸组分而且还含有非氨基酸组分如糖和磷酸盐结构的物质,还包括含有键如酯、硫醚或二硫键的物质。

[0129] 根据本发明,术语“起始密码子(initiation codon)”和“起始密码子(start codon)”同义地指RNA分子的密码子(碱基三联体),其可能是由核糖体翻译的第一个密码子。这种密码子通常在真核生物中编码氨基酸甲硫氨酸,在原核生物中编码修饰的甲硫氨酸。在真核生物和原核生物中最常见的起始密码子是AUG。除非本文特别说明是指AUG以外的起始密码子,否则关于RNA分子的术语“起始密码子(initiation codon)”和“起始密码子(start codon)”是指密码子AUG。根据本发明,术语“起始密码子(initiation codon)”和“起始密码子(start codon)”也用来指脱氧核糖核酸的相应碱基三联体,即编码RNA起始密码子的碱基三联体。如果信使RNA的起始密码子是AUG,则编码AUG的碱基三联体是ATG。根据本发明,术语“起始密码子(initiation codon)”和“起始密码子(start codon)”优选是指功能性起始密码子(initiation codon)或起始密码子(start codon),即指被核糖体用作或将用作密码子以开始翻译的起始密码子(initiation codon)或起始密码子(start codon)。RNA分子中可以存在AUG密码子,这些密码子不被核糖体用作密码子来开始翻译,例如,由于密码子到帽的距离很短。术语功能性起始密码子(initiation codon)或起始密码子(start codon)不涵盖这些密码子。

[0130] 根据本发明,术语“开放阅读框的起始密码子(start codon)”或“开放阅读框的起始密码子(initiation codon)”是指在编码序列中用作蛋白合成起始密码子的碱基三联体,例如,在自然界中发现的核酸分子的编码序列中。在RNA分子中,开放阅读框的起始密码子之前通常有5’非翻译区(5’-UTR),尽管这并不是严格要求的。

[0131] 根据本发明,术语“开放阅读框的天然起始密码子(start codon)”或“开放阅读框的天然起始密码子(initiation codon)”是指在天然编码序列中用作蛋白合成起始密码子的碱基三联体。天然编码序列可以是例如自然界中发现的核酸分子的编码序列。在一些实施方案中,本发明提供自然界中发现的核酸分子的变体,其特征在于天然起始密码子(存在于天然编码序列中)已被去除(因此它不存在于变体核酸分子中)。

[0132] 根据本发明,“第一AUG”表示信使RNA分子的最上游AUG碱基三联体,优选被核糖体用作或将用作密码子以开始翻译的信使RNA分子的最上游AUG碱基三联体。因此,“第一ATG”是指编码第一个AUG的编码DNA序列的ATG碱基三联体。在某些情况下,mRNA分子的第一个AUG是开放阅读框的起始密码子,即在核糖体蛋白合成过程中用作起始密码子的密码子。

[0133] 根据本发明,关于核酸变体的某个元件,术语“包括去除”或“特征在于去除”以及类似术语表示与参考核酸分子相比,所述某个元件在核酸变体中没有功能或不存在。不受

限制,去除可以由某个元件的全部或部分的缺失、某个元件的全部或部分的替代或者某个元件的功能或结构特性的改变组成。核酸序列的功能元件的去除要求在包含去除的核酸变体的位置不表现出功能。例如,特征在于去除某个起始密码子的RNA变体要求核糖体蛋白合成不在特征在于去除的RNA变体的位置开始。核酸序列的结构元件的去除要求结构元件不存在于包含去除的核酸变体的位置。例如,特征在于去除某个AUG碱基三联体,即在某个位置去除AUG碱基三联体的RNA变体,可以特征在于例如某个AUG碱基三联体的部分或全部的缺失(例如,ΔAUG),或者用任何一个或多个不同的核苷酸替代某个AUG碱基三联体的一个或多个核苷酸(A、U、G),从而使得到的变体核苷酸序列不包含所述AUG碱基三联体。一个核苷酸的合适替代是将AUG碱基三联体转化为GUG、CUG或UUG碱基三联体,或者转化为AAG、ACG或AGG碱基三联体,或者转化为AUA、AUC或AUU碱基三联体。可以相应地选择更多核苷酸的合适替代。

[0134] 根据本发明,术语“自复制病毒”包括能够在宿主细胞自主复制的RNA病毒。自复制病毒可以具有单链RNA(ssRNA)基因组,包括甲病毒、黄病毒、麻疹病毒(MV)和弹状病毒。甲病毒和黄病毒具有正极性的基因组,而麻疹病毒(MV)和弹状病毒是负链ssRNA。通常,自复制病毒是具有(+)链RNA基因组的病毒,其可以在感染细胞后直接翻译,并且这种翻译提供RNA依赖性RNA聚合酶,然后从感染的RNA产生反义和正义转录物。在下文中,参考甲病毒衍生的载体作为自复制病毒衍生的载体的实例来说明本发明。但是,应当理解本发明并不局限于甲病毒衍生的载体。

[0135] 根据本发明,术语“甲病毒”应从广义上理解,包括任何具有甲病毒特征的病毒颗粒。甲病毒的特征包括存在(+)链RNA,其编码适合在宿主细胞中复制的遗传信息,包括RNA聚合酶活性。描述了许多甲病毒的进一步特征,例如在Strauss&Strauss,1994,Microbiol.Rev.58:491-562中。术语“甲病毒”包括自然界中发现的甲病毒,以及其任何变体或衍生物。在一些实施方案中,在自然界中没有发现的变体或衍生物。

[0136] 在一实施方案中,甲病毒是自然界中发现的甲病毒。通常,自然界中发现的甲病毒对任何一种或多种真核生物都具有感染性,如动物(包括脊椎动物如人,和节肢动物如昆虫)。自然界中发现的甲病毒优选选自由以下组成的组:巴马森林病毒(Barmah Forest virus)复合体(包含巴马森林病毒);东方马脑炎复合体(包含东方马脑炎病毒的七种抗原类型);米德尔堡病毒(Middelburg virus)复合体(包含米德尔堡病毒);恩杜姆病毒(Ndumu virus)复合体(包含恩杜姆病毒);塞姆利基森林病毒复合体(包含贝巴鲁病毒(Bebaru virus)、基孔肯雅病毒、马亚罗病毒(Mayaro virus)及其亚型乌纳病毒(Una virus)、欧尼恩病毒(O'Nyong Nyong virus)及其亚型Igbo-Ora病毒、罗斯河病毒(Ross River virus)及其亚型贝巴鲁病毒、盖塔病毒(Getah virus)、鹭山病毒(Sagiyama virus)、塞姆利基森林病毒及其亚型Me Tri病毒);委内瑞拉马脑炎复合体(包含卡巴斯欧病毒(Cabassou virus)、沼泽地病毒(Everglades virus)、莫斯达斯佩德拉斯病毒(Mosso das Pedras virus)、穆坎布病毒(Mucambo virus)、Paramana病毒、那皮舒纳病毒(Pixuna virus)、里奥内格罗病毒(Rio Negro virus)、特罗卡拉病毒(Trocara virus)及其亚型Bijou Bridge病毒、委内瑞拉马脑炎病毒);西方马脑炎复合体(包含奥拉病毒(Aura virus)、巴班肯病毒(Babanki virus)、孜拉加奇病毒(Kyzylagach virus)、辛德毕斯病毒、奥克尔布病毒(Ockelbo virus)、瓦塔罗阿病毒(Whataroa virus)、博吉河病毒(Buggy Creek virus)、摩

根堡病毒(Fort Morgan virus)、高地J病毒(Highlands J virus)、西方马脑炎病毒);以及一些未分类的病毒,其包括鲑鱼胰腺病病毒;睡眠疾病病毒;南方象海豹病毒;Tonate病毒。更优选地,甲病毒选自塞姆利基森林病毒复合体(包含如上所示的病毒类型,包括塞姆利基森林病毒)、西方马脑炎复合体(包含如上所示的病毒类型,包括辛德毕斯病毒)、东方马脑炎病毒(包含如上所示的病毒类型)、委内瑞拉马脑炎复合体(包含如上所示的病毒类型,包括委内瑞拉马脑炎病毒)。

[0137] 在进一步优选的实施方案中,甲病毒是塞姆利基森林病毒。在可选的另一优选实施方案中,甲病毒是辛德毕斯病毒。在可选的另一优选实施方案中,甲病毒是委内瑞拉马脑炎病毒。

[0138] 在本发明的一些实施方案中,甲病毒不是自然界中发现的甲病毒。通常,自然界中未发现的甲病毒是自然界中发现的甲病毒的变体或衍生物,其与自然界中发现的甲病毒的区别在于核苷酸序列(即基因组RNA)中的至少一个突变。与自然界中发现的甲病毒相比,核苷酸序列中的突变可以选自一个或多个核苷酸的插入、替代或缺失。核苷酸序列中的突变可以与核苷酸序列编码的多肽或蛋白中的突变相关,也可以不相关。例如,自然界中未发现的甲病毒可以是减毒的甲病毒。自然界中未发现的减毒甲病毒是通常在其核苷酸序列中具有至少一个突变的甲病毒,通过该突变可以将其与自然界中发现的甲病毒区分,并且该病毒完全不具有感染性,或者具有感染性但具有较低的致病能力或完全没有致病能力。作为说明性实例,TC83是一种减毒的甲病毒,与自然界中发现委内瑞拉马脑炎病毒(VEEV)不同(McKinney et al.,1963,Am.J.Trop.Med.Hyg.12:597-603)。

[0139] 甲病毒属的成员也可以基于其在人中的相对临床特征进行分类:主要与脑炎相关的甲病毒,以及主要与发热、皮疹和多关节炎相关的甲病毒。

[0140] 术语“甲病毒的”表示在甲病毒中发现的,或者来源于甲病毒或衍生自甲病毒的,例如,通过基因工程化。

[0141] 根据本发明,“SFV”代表塞姆利基森林病毒。根据本发明,“SIN”或“SINV”代表辛德毕斯病毒。根据本发明,“VEE”或“VEEV”代表委内瑞拉马脑炎病毒。

[0142] 根据本发明,术语“甲病毒的(of an alphavirus)”是指源自甲病毒的实体。为了说明,甲病毒的蛋白可以指在甲病毒中发现的蛋白和/或由甲病毒编码的蛋白;甲病毒的核酸序列可以指在甲病毒中发现的核酸序列和/或由甲病毒编码的核酸序列。优选地,“甲病毒的”核酸序列是指“甲病毒基因组的”和/或“甲病毒基因组RNA的”核酸序列。

[0143] 根据本发明,术语“甲病毒RNA”是指甲病毒基因组RNA (即,(+)链)、甲病毒基因组RNA的互补物(即,(-)链)和亚基因组转录物(即(+)链)中的任何一种或多种,或者其中任何一种的片段。

[0144] 根据本发明,“甲病毒基因组”是指甲病毒的基因组(+)链RNA。

[0145] 根据本发明,术语“天然甲病毒序列”和类似术语通常是指天然存在的甲病毒(自然界中发现的甲病毒)的(例如,核酸)序列。在一些实施方案中,术语“天然甲病毒序列”还包括减毒甲病毒的序列。

[0146] 根据本发明,术语“5’复制识别序列”优选是指连续的核酸序列,优选核糖核酸序列,即与自复制病毒基因组如甲病毒基因组的5’片段相同或同源。“5’复制识别序列”是可以被复制酶如甲病毒复制酶识别的核酸序列。术语5’复制识别序列包括天然5’复制识别序

列及其功能等同物,例如,自然界中发现的自复制病毒的5'复制识别序列的功能变体,例如,自然界中发现的甲病毒。根据本发明,功能等同物包括5'复制识别序列的衍生物,其特征在于如本文所述去除至少一个起始密码子。5'复制识别序列是合成甲病毒基因组RNA的(-)链互补物所必需的,并且是基于(-)链模板合成(+)链病毒基因组RNA所必需的。天然5'复制识别序列通常至少编码nsP1的N-末端片段;但是不包含编码nsP1234的整个开放阅读框。鉴于天然5'复制识别序列通常至少编码nsP1的N-末端片段,天然5'复制识别序列通常包含至少一个起始密码子,通常为AUG。在一实施方案中,5'复制识别序列包含甲病毒基因组的保守序列元件1(CSE 1)或其变体以及甲病毒基因组的保守序列元件2(CSE 2)或其变体。5'复制识别序列通常能够形成四个茎环(SL),即SL1、SL2、SL3、SL4。这些茎环的编号开始于5'复制识别序列的5'末端。

[0147] 术语"保守序列元件"或"CSE"是指在甲病毒RNA中发现的核苷酸序列。这些序列被称为"保守的",因为在不同的甲病毒基因组中存在直系同源物,并且不同甲病毒的直系同源CSE优选具有高百分比的序列相同性和/或相似的二级或三级结构。术语CSE包括CSE 1、CSE 2、CSE 3和CSE 4。

[0148] 根据本发明,术语"CSE 1"或"44-nt CSE"同义地指从(-)链模板合成(+)链所需的核苷酸序列。术语"CSE 1"是指(+)链上的序列;CSE 1(在(-)链上)的互补序列作为(+)链合成的启动子发挥功能。优选地,术语CSE 1包括甲病毒基因组的最5'的核苷酸。CSE 1通常形成保守的茎-环结构。不希望受到特定理论的束缚,据信,对于CSE 1,二级结构比一级结构(即线性序列)更重要。在模型病毒辛德毕斯病毒的基因组RNA中,CSE 1由44个核苷酸的连续序列组成,该序列由基因组RNA的最5'的44个核苷酸形成(Strauss&Strauss,1994,Microbiol.Rev.58:491-562)。

[0149] 根据本发明,术语"CSE 2"或"51-nt CSE"同义地指从(+)链模板合成(-)链所需的核苷酸序列。(+)链模板通常是甲病毒基因组RNA或RNA复制子(注意,不包含CSE 2的亚基因组RNA转录物不作为(-)链合成的模板发挥功能)。在甲病毒基因组RNA中,CSE 2通常位于nsP1的编码序列内。在模型甲病毒辛德毕斯病毒的基因组RNA中,51-nt CSE位于基因组RNA的核苷酸位置155-205(Frolov et al.,2001,RNA,vol.7,pp.1638-1651)。CSE 2通常形成两个保守的茎环结构。这些茎环结构被命名为茎环3(SL3)和茎环4(SL4),因为它们分别是甲病毒基因组RNA的第三和第四个保守茎环,从甲病毒基因组RNA的5'末端计数。不希望受到特定理论的束缚,据信,对于CSE 2,二级结构比一级结构(即线性序列)更重要。

[0150] 根据本发明,术语"CSE 3"或"连接序列"同义地指源自甲病毒基因组RNA并且包含亚基因组RNA的起始位点的核苷酸序列。该序列在(-)链中的互补物起到促进亚基因组RNA转录的作用。在甲病毒基因组RNA中,CSE 3通常与编码nsP4 C-末端片段的区域重叠,并延伸至位于编码结构蛋白的开放阅读框上游的短非编码区。

[0151] 根据本发明,术语"CSE 4"或"19-nt保守序列"或"19-nt CSE"同义地指来自甲病毒基因组RNA的核苷酸序列,紧邻甲病毒基因组3′非翻译区中poly(A)序列的上游。CSE 4通常由19个连续核苷酸组成。不希望受到特定理论的束缚,CSE 4被认为是启动(-)链合成的核心启动子(Joséet al.,2009,Future Microbiol.4:837-856);和/或甲病毒基因组RNA的CSE 4和poly(A)尾被认为共同发挥作用,实现高效的(-)链合成(Hardy&Rice,2005,J.Virol.79:4630-4639)。

[0152] 根据本发明,术语“亚基因组启动子”或“SGP”是指核酸序列(例如,编码序列)上游(5')的核酸序列,其通过为RNA聚合酶,通常是RNA依赖性RNA聚合酶,特别是功能性甲病毒非结构蛋白提供识别和结合位点来控制所述核酸序列的转录。SGP可以包括其他因子的其他识别或结合位点。亚基因组启动子通常是正链RNA病毒如甲病毒的遗传元件。甲病毒的亚基因组启动子是病毒基因组RNA中包含的核酸序列。亚基因组启动子一般特征在于它允许在RNA依赖性RNA聚合酶(例如,功能性甲病毒非结构蛋白)的存在下启动转录(RNA合成)。RNA(-)链,即甲病毒基因组RNA的互补物,用作(+)链亚基因组转录物合成的模板,并且(+)链亚基因组转录物的合成通常在亚基因组启动子处或其附近启动。如本文所用的术语“亚基因组启动子”并不局限于包含这种亚基因组启动子的核酸中的任何特定定位。在一些实施方案中,SGP与CSE 3相同或与CSE 3重叠或包含CSE 3。

[0153] 术语“亚基因组转录物”或“亚基因组RNA”同义地指使用RNA分子作为模板(“模板RNA”)进行转录可获得的RNA分子,其中模板RNA包含控制亚基因组转录物转录的亚基因组启动子。亚基因组转录物在RNA依赖性RNA聚合酶,特别是功能性甲病毒非结构蛋白的存在下可获得。例如,术语“亚基因组转录物”可以指使用甲病毒基因组RNA的(-)链互补物作为模板,在甲病毒感染的细胞中制备的RNA转录物。但是,如本文所用,术语“亚基因组转录物”不限于此,还包括通过使用异源RNA作为模板可获得的转录物。例如,通过使用本发明的含有SGP的复制子的(-)链互补物作为模板也可获得亚基因组转录物。因此,术语“亚基因组转录物”可以指通过转录甲病毒基因组RNA片段可获得的RNA分子,以及通过转录本发明的复制子片段可获得的RNA分子。

[0154] 术语“自体”用来描述源自同一受试者的任何事物。例如,“自体细胞”是指源自同一受试者的细胞。将自体细胞引入受试者是有利的,因为这些细胞克服了免疫屏障,否则会导致排斥反应。

[0155] 术语“同种异体”用来描述源自相同物种的不同个体的任何事物。当在一个或多个基因座处的基因不同时,说两个或更多个个体互相是同种异体的。

[0156] 术语“同基因”用来描述源自具有相同基因型的个体或组织的任何事物,即,同一近交系的同卵双胞胎或动物,或者它们的组织或细胞。

[0157] 术语“异源”用来描述由多种不同元素组成的事物。作为实例,将一个个体的细胞引入不同个体构成异源移植。异源基因是源自受试者以外来源的基因。

[0158] 可以用于鉴定序列变化的方法的细胞是任何适当的细胞,其中有或没有任何核苷酸修饰的rRNA可以复制和/或翻译。所述细胞可以是哺乳动物细胞,例如,人细胞。所述细胞可以组成性地表达识别rRNA中存在的序列用于复制的复制酶,或者可以瞬时表达这种复制酶。

[0159] 以下提供本发明各个特征的特定和/或优选变体。本发明还考虑那些特别优选的实施方案,这些实施方案是通过组合本发明的两个或更多个特征描述的两个或更多个特定和/或优选变体而产生的。

[0160] RNA复制子

[0161] 可复制RNA(rRNA)是能够通过RNA依赖性RNA聚合酶(复制酶)复制的RNA,其通过包含可以被复制酶识别的核苷酸序列,从而使RNA被复制。rRNA不一定编码复制酶,因此rRNA可以顺式(通过编码的复制酶)或反式(通过以另一种方式提供的复制酶,例如,单独的编码

复制酶的核酸)复制。术语“RNA复制子”、“复制子”、“rRNA”、“自扩增RNA”、“saRNA”和“可复制RNA分子”可以互换使用。

[0162] 在一实施方案中,可复制RNA(rRNA)分子包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中所述分子包含序列AUGGCGGA或AUGGGCGG,其中这些序列AUGGCGGA或AUGGGCGG中任一个中的U是尿苷;并且其中所述分子中剩余尿苷中的至少一个是N1-甲基-假尿苷(1mΨ)。任选地,所述分子中所有的剩余尿苷可以是1mΨ,或者所述分子中至少5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%或100%的剩余尿苷可以是1mΨ。序列AUGGCGGA或AUGGGCGG可以位于分子的非编码区中,或者序列AUGGCGGA或AUGGGCGG可以位于5’调节区中。AUGGCGGA或AUGGGCGG可以位于5’调节区中的保守序列元件1(CSE 1)中,或者AUGGCGGA或AUGGGCGG可以位于分子的5’末端。AUGGCGGA或AUGGGCGG可以进一步包含AUGGCGGA或AUGGGCGG所描述序列5’的额外核苷酸,任选地,其中所述额外核苷酸包含额外的ORF和/或控制序列或者形成5’帽结构的一个或多个核苷酸。

[0163] 在一实施方案中,修饰的可复制RNA分子包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了5’调节区中包含的保守序列元件1(CSE 1)的10个5’核苷酸内包含的尿苷外,所述分子中的尿苷中的至少一个尿苷是N1-甲基-假尿苷(1mΨ)。任选地,除了CSE 1的10、5、4、3、2个5’核苷酸内包含的尿苷外,所述分子中的所有尿苷为1mΨ。任选地,除了CSE 1中第2位的尿苷外,所述分子中的所有尿苷为1mΨ。

[0164] 在一实施方案中,修饰的可复制RNA分子包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了5’调节区中包含的保守序列元件1(CSE 1)内包含的最5’的U外,所述分子中的至少一个尿苷是N1-甲基-假尿苷(1mΨ)。在一实施方案中,修饰的可复制RNA分子具有5’帽,并且包含甲病毒的5’调节区和至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中除了所述分子中的第一个5’尿苷外,所述分子中的至少一个尿苷是N1-甲基-假尿苷(1mΨ)。所述5’帽可以是G(5’)ppp(5’)AU或m^7G(5’)ppp(5’)AU。

[0165] 在一实施方案中,修饰的可复制RNA分子包含至少一个编码至少一种所关注的基因产物的开放阅读框(ORF),其中所述分子中的至少一个尿苷是N1-甲基-假尿苷(1mΨ),并且其中所述分子包含具有序列NpppNU的5’帽,其中所述5’帽中的U是尿苷。任选地,所述5’帽具有序列NpppAU。

[0166] 在一实施方案中,RNA复制子可以包含内部核糖体进入位点(IRES)和编码来自自复制病毒的功能性非结构蛋白的开放阅读框,其中所述IRES控制功能性非结构蛋白(例如,复制酶)的表达。优选地,RNA复制子含有允许功能性非结构蛋白复制的序列元件。在一实施方案中,自复制病毒是甲病毒,允许功能性非结构蛋白复制的序列元件源自甲病毒。

[0167] 甲病毒复制酶具有加帽酶功能,通常基因组以及亚基因组(+)链RNA被加帽。5’-帽用于保护mRNA免受降解,并且将核糖体亚基以及细胞因子引导至mRNA,以便在mRNA上形成核糖核蛋白复合物,然后可以从附近的起始密码子开始翻译。这个复杂的过程在文献中被广泛描述(Jackson et al.,2010,Nat Rev Mol Biol;Vol 10:113-127)。尽管帽依赖性翻译的机制非常精细和高效,但是细胞有办法完全或部分独立于5’帽启动翻译(Thompson

2012;Trends in Microbiology 20:558-566)。因此,在导致帽依赖性翻译整体下调的细胞应激情况下,细胞仍然可以优先表达选择的基因,通常是在IRES的帮助下。

[0168] 病毒还进化出不同的手段来利用细胞机制翻译病毒基因。由于病毒感染通常被细胞感知,这导致细胞抗病毒反应(干扰素反应;应激反应),因此许多病毒也利用不依赖帽的翻译,特别是RNA病毒。不依赖帽的翻译确保了细胞应激反应时病毒RNA翻译的优势,使病毒有机会完成其生命周期并从感染细胞中释放出来。

[0169] 内部核糖体进入位点(IRES)是形成适当二级结构的RNA序列,其将起始前复合物吸引到翻译起始密码子AUG或其他的附近。文献中描述了四类具有共同特征的IRES。原型IRES是脊髓灰质炎病毒IRES(I型)、脑心肌炎病毒(EMCV)IRES(II型)、丙型肝炎病毒(HCV)IRES(III型)和在双顺反子病毒(dicistroviruses)基因间区域中发现的IRES(IV型)(Thompson,2012;Trends in Microbiology 20:558-566;Lozano et al.,2018;Open Biology 8:180155)。

[0170] I型至III型IRES的共同点是它们在AUG起始密码子处启动翻译,而IV型IRES在非AUG密码子(例如,GCU)处启动翻译。因此I型至III型需要借助eIF2/GTP递送甲硫氨酸的起始子tRNA(eIF2/GTP/Met-tRNAiMet)。在应激下eIF2激酶的激活使eIF2的α亚基磷酸化,这抑制起始于AUG的翻译。因此IV型IRES指导翻译不受eIF2磷酸化的抑制。

[0171] 术语“内部核糖体进入位点”,缩写为“IRES”,涉及将核糖体招募到mRNA的内部区域以不依赖帽的方式启动翻译的RNA元件。IRES通常位于RNA病毒的5’-UTR中。然而,来自双顺反子病毒科(dicistroviridae)家族的病毒mRNA具有两个开放阅读框(ORF),每个ORF的翻译由两个不同的IRES指导。还有人提出,一些哺乳动物细胞mRNA也具有IRES。这些细胞IRES元件被认为位于真核生物mRNA中,所述mRNA编码参与应激存活和其他对存活至关重要的过程的基因。IRES元件的位置通常在5’-UTR中,但是也可以出现在mRNA中的其他位置。

[0172] 术语“内部核糖体进入位点”包括存在于小核糖核酸病毒科的病毒如脊髓灰质炎病毒(PV)和脑心肌炎病毒以及致病性病毒(包括人类免疫缺陷病毒、丙型肝炎病毒(HCV)和口蹄疫病毒)中的IRES。尽管这些病毒IRES含有不同的序列,但是其中许多具有相似的二级结构并通过相似的机制启动翻译。此外,IRES的活性通常需要其他因子的协助,这些因子被称作IRES-反式作用因子(ITAF)。基于翻译起始因子(IF)和ITAF的结构和需要,将病毒IRES分为如本文所述的四种类型。根据本发明,这些IRES类型中的任何一种都是有用的,特别优选IV型IRES。

[0173] I型和II型两组病毒IRES不可以直接与40S小核糖体亚基结合。相反,它们通过不同的ITAF招募40S小核糖体亚基,并且在帽依赖性翻译中需要经典的IF(即,eIF2、eIF3、eIF4A、eIF4B和eIF4G)。I型和II型IRES之间的主要区别是需要40S核糖体扫描,而II型IRES不需要40S核糖体扫描。I型IRES的实例包括在脊髓灰质炎病毒(PV)和鼻病毒中发现的IRES。II型IRES的实例包括在脑心肌炎病毒(EMCV)、口蹄疫病毒(FMDV)和泰勒氏鼠脑脊髓炎病毒(TMEV)中发现的IRES。

[0174] III型IRES可以直接与具有特殊RNA结构的40S小核糖体亚基相互作用,但它们的活性通常需要包括eIF2和eIF3在内的几种IF和起始子Met-tRNAi的协助。实例包括在丙型肝炎病毒(HCV)、猪瘟病毒(CSFV)和猪捷申病毒(PTV)中发现的IRES。

[0175] IV型病毒IRES一般具有很强的活性,可以从非AUG起始密码子启动翻译而无需额

外的ITAF甚至eIF2/Met-tRNAi/GTP三元复合物。这些IRES折叠成紧凑的结构,直接与40S小核糖体亚基相互作用。实例包括在双顺反子病毒如蟋蟀麻痹病毒(CrPV)、斯氏珀蝽肠病毒(plautia staliintestine virus)(PSIV)和托拉综合症病毒(Taura-Syndrom-Virus)(TSV)中发现的IRES。

[0176] 术语“内部核糖体进入位点”还包括在细胞mRNA中发现的IRES,其中许多编码应激反应所需的蛋白,例如在细胞凋亡、有丝分裂、缺氧和营养限制的条件下。基于核糖体招募的机制,细胞IRES可以大致分为两种类型:I型IRES通过结合在顺式元件上的ITAF与核糖体相互作用,例如,RNA结合基序和N-6-甲基腺苷(m6A)修饰,而II型IRES含有与18S rRNA配对以招募核糖体的短顺式元件。

[0177] 在一实施方案中,本文所述的rRNA可以具有修饰的核苷酸/核苷/骨架修饰。如本文所用的术语“RNA修饰”可以指化学修饰,其包括骨架修饰以及糖修饰或碱基修饰。

[0178] 在这种情况下,本文定义的修饰的rRNA分子可以含有核苷酸类似物/修饰,例如,骨架修饰、糖修饰或碱基修饰。与本发明有关的骨架修饰是一种修饰,其中对本文定义的rRNA分子中含有的核苷酸骨架的磷酸酯进行化学修饰。与本发明有关的糖修饰是本文定义的rRNA分子核苷酸糖的化学修饰。此外,与本发明有关的碱基修饰是rRNA分子核苷酸碱基部分的化学修饰。在这种情况下,核苷酸类似物或修饰优选选自适用于转录和/或翻译的核苷酸类似物。

[0179] 糖修饰:可以掺入本文所述修饰的rRNA分子的修饰的核苷和核苷酸可以在糖部分中进行修饰。例如,可以用许多不同的“氧”或“脱氧”取代基修饰或取代2’羟基(OH)。“氧”-2’羟基修饰的实例包括但不限于烷氧基或芳氧基(-OR,例如,R=H、烷基、环烷基、芳基、芳烷基、杂芳基或糖);聚乙二醇(PEG),-O(CH2CH2 O)nCH2CH2 OR;“锁定”核酸(LNA),其中2’羟基通过例如亚甲基桥连接至相同核糖的4’碳;以及氨基(-O-氨基,其中氨基,例如NRR,可以是烷基氨基、二烷基氨基、杂环基、芳基氨基、二芳基氨基、杂芳基氨基或二杂芳基氨基、二乙胺、多氨基)或氨基烷氧基。“脱氧”修饰包括氢、氨基(例如NH2;烷基氨基、二烷基氨基、杂环基、芳基氨基、二芳基氨基、杂芳基氨基、二杂芳基氨基或氨基酸);或者氨基可以通过接头连接至糖,其中所述接头包含原子C、N和O中的一个或多个。糖基还可以含有一个或多个具有与核糖中相应碳相反的立体化学构型的碳。因此,修饰的RNA分子可以包括核苷酸,所述核苷酸含有例如阿拉伯糖作为糖。

[0180] 骨架修饰:磷酸骨架可以在修饰的核苷和核苷酸中进一步修饰,其可以掺入如本文所述修饰的RNA分子。骨架的磷酸基团可以通过用不同取代基取代一个或多个氧原子来进行修饰。此外,修饰的核苷和核苷酸可以包括用本文所述修饰的磷酸完全取代未修饰的磷酸部分。修饰的磷酸基团的实例包括但不限于硫代磷酸、硒代磷酸、硼代磷酸、硼代磷酸酯、氢膦酸、磷酰胺、烷基或芳基膦酸酯以及磷酸三酯。二硫代磷酸酯的两个非连接氧均被硫取代。磷酸接头还可以通过用氮(桥连磷酰胺)、硫(桥连硫代磷酸)和碳(桥连亚甲基-膦酸)取代连接氧来进行修饰。

[0181] 碱基修饰:可以掺入本文所述修饰的rRNA分子的修饰的核苷和核苷酸可以在核碱基部分中进一步修饰。在RNA中发现的核碱基的实例包括但不限于腺嘌呤、鸟嘌呤、胞嘧啶和尿嘧啶。例如,本文所述的核苷和核苷酸可以在主要凹槽表面上进行化学修饰。在一些实施方案中,主要凹槽化学修饰可以包括氨基、硫醇基、烷基或卤素基团。

[0182] 在本发明的特定实施方案中,核苷酸类似物/修饰选自碱基修饰,其优选选自2-氨基-6-氯嘌呤核苷-5'-三磷酸、2-氨基嘌呤-核苷-5'-三磷酸;2-氨基腺苷-5'-三磷酸、2'-氨基-2'-脱氧-胞苷-三磷酸、2-硫代胞苷-5'-三磷酸、2-硫代尿苷-5'-三磷酸、2'-氟胸苷-5'-三磷酸、2'-0-甲基肌苷-5'-三磷酸、4-硫代-尿苷-5'-三磷酸、5-氨基烯丙基胞苷-5'-三磷酸、5-氨基烯丙基尿苷-5'-三磷酸、5-溴胞苷-5'-三磷酸、5-溴尿苷-5'-三磷酸、5-溴-2'-脱氧胞苷-5'-三磷酸、5-溴-2'-脱氧尿苷-5'-三磷酸、5-碘胞苷-5'-三磷酸、5-碘-2'-脱氧胞苷-5'-三磷酸、5-碘尿苷-5'-三磷酸、5-碘-2'-脱氧尿苷-5'-三磷酸、5-甲基胞苷-5'-三磷酸、5-甲基尿苷-5'-三磷酸、5-丙炔基-2'-脱氧胞苷-5'-三磷酸、5-丙炔基-2'-脱氧尿苷-5'-三磷酸、6-氮杂胞苷-5'-三磷酸、6-氮杂尿苷-5'-三磷酸、6-氯嘌呤核苷-5'-三磷酸、7-脱氮-腺苷-5'-三磷酸、7-脱氮鸟苷-5'-三磷酸、8-氮杂腺苷-5'-三磷酸、8-叠氮腺苷-5'-三磷酸、苯并咪唑-核苷-5'-三磷酸、N1-甲基腺苷-5'-三磷酸、N1-甲基鸟苷-5'-三磷酸、N6-甲基腺苷-5'-三磷酸、06-甲基鸟苷-5'-三磷酸、假-尿苷-5'-三磷酸或嘌呤霉素-5'-三磷酸、黄苷-5'-三磷酸。可以特别优选用于碱基修饰的核苷酸,其选自由5-甲基胞苷-5'-三磷酸、7-脱氮鸟苷-5'-三磷酸、5-溴胞苷-5'-三磷酸和假尿苷-5'-三磷酸组成的碱基修饰的核苷酸组。在一些实施方案中,修饰的核苷包括吡啶-4-酮核糖核苷、5-氮杂-尿苷、2-硫代-5-氮杂-尿苷、2-硫代尿苷、4-硫代-假尿苷、2-硫代-假尿苷、5-羟基尿苷、3-甲基尿苷、5-羧甲基-尿苷、1-羧甲基-假尿苷、5-丙炔基-尿苷、1-丙炔基-假尿苷、5-牛磺酸甲基尿苷、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫代尿苷、1-牛磺酸甲基-4-硫代-尿苷、5-甲基-尿苷、1-甲基-假尿苷、4-硫代-1-甲基-假尿苷、2-硫代-1-甲基-假尿苷、1-甲基-1-脱氮-假尿苷、2-硫代-1-甲基-1-脱氮-假尿苷、二氢尿苷、二氢-假尿苷、2-硫代-二氢尿苷、2-硫代-二氢假尿苷、2-甲氧基-尿苷、2-甲氧基-4-硫代-尿苷、4-甲氧基-假尿苷和4-甲氧基-2-硫代-假尿苷。

[0183] 在一些实施方案中,修饰的核苷包括5-氮杂-胞苷、假异胞苷、3-甲基-胞苷、N4-乙酰胞苷、5-甲酰胞苷、N4-甲基胞苷、5-羟甲基胞苷、1-甲基-假异胞苷、吡咯-胞苷、吡咯-假异胞苷、2-硫代-胞苷、2-硫代-5-甲基-胞苷、4-硫代-假异胞苷、4-硫代-1-甲基-假异胞苷、4-硫代-1-甲基-1-脱氮-假异胞苷、1-甲基-1-脱氮-假异胞苷、泽布拉林(zebularine)、5-氮杂-泽布拉林、5-甲基-泽布拉林、5-氮杂-2-硫代-泽布拉林、2-硫代-泽布拉林、2-甲氧基-胞苷、2-甲氧基-5-甲基-胞苷、4-甲氧基-假异胞苷和4-甲氧基-1-甲基-假异胞苷。

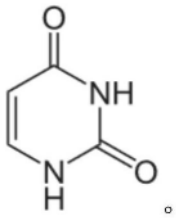
[0184] 在其他实施方案中,修饰的核苷包括2-氨基嘌呤、2,6-二氨基嘌呤、7-脱氮-腺嘌呤、7-脱氮-8-氮杂-腺嘌呤、7-脱氮-2-氨基嘌呤、7-脱氮-8-氮杂-2-氨基嘌呤、7-脱氮-2,6-二氨基嘌呤、7-脱氮-8-氮杂-2,6-二氨基-嘌呤、1-甲基腺苷、N6-甲基腺苷、N6-异戊烯基腺苷、N6-(顺式-羟基异戊烯基)腺苷、2-甲基硫代-N6-(顺式-羟基异戊烯基)腺苷、N6-甘氨酰氨基甲酰腺苷、N6-苏氨酰氨基甲酰腺苷、2-甲基-硫代-N6-苏氨酰氨基甲酰腺苷、N6,N6-二甲基腺苷、7-甲基腺嘌呤、2-甲基硫代-腺嘌呤和2-甲氧基-腺嘌呤。在其他实施方案中,修饰的核苷包括肌苷、1-甲基-肌苷、怀俄苷(wyosine)、怀丁苷(wybutosine)、7-脱氮-鸟苷、7-脱氮-8-氮杂-鸟苷、6-硫代-鸟苷、6-硫代-7-脱氮-鸟苷、6-硫代-7-脱氮-8-氮杂-鸟苷、7-甲基-鸟苷、6-硫代-7-甲基-鸟苷、7-甲基肌苷、6-甲氧基-鸟苷、1-甲基鸟苷、N2-甲基鸟苷、N2,N2-二甲基鸟苷、8-氧代-鸟苷、7-甲基-8-氧代-鸟苷、1-甲基-6-硫代-鸟苷、N2-甲基-6-硫代-鸟苷和N2,N2-二甲基-6-硫代-鸟苷。

[0185] 在一些实施方案中,核苷酸可以在主要凹槽表面上进行修饰,可以包括用甲基或卤素基团取代尿嘧啶C-5上的氢。在具体实施方案中,修饰的核苷是5'-0-(1-硫代磷酸)-腺苷、5'-0-(1-硫代磷酸)-胞苷、5'-0-(1-硫代磷酸)-鸟苷、5'-0-(1-硫代磷酸)-尿苷或5'-0-(1-硫代磷酸)-假尿苷。

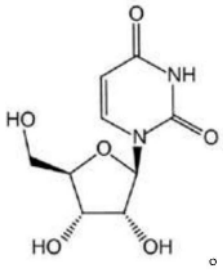
[0186] 在其他实施方案中,修饰的rRNA可以包含选自以下的核苷修饰:6-氮杂-胞苷、2-硫代-胞苷、a-硫代-胞苷、假-异-胞苷、5-氨基烯丙基-尿苷、5-碘-尿苷、N1-甲基-假尿苷、5,6-二氢尿苷、a-硫代-尿苷、4-硫代-尿苷、6-氮杂-尿苷、5-羟基-尿苷、脱氧-胸苷、5-甲基-尿苷、吡咯-胞苷、肌苷、a-硫代-鸟苷、6-甲基-鸟苷、5-甲基-胞苷、8-氧代-鸟苷、7-脱氮-鸟苷、N1-甲基-腺苷、2-氨基-6-氯-嘌呤、N6-甲基-2-氨基-嘌呤、假-异-胞苷、6-氯-嘌呤、N6-甲基-腺苷、a-硫代-腺苷、8-叠氮-腺苷、7-脱氮-腺苷。

[0187] 在某些优选的实施方案中,除本文另有规定,rRNA包含修饰的核苷代替至少一个(例如,每个)尿苷。

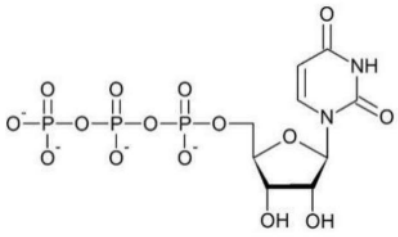
[0188] 如本文所用,术语"尿嘧啶"描述可以在RNA的核酸中出现的核碱基之一。尿嘧啶的结构是:

[0189] 。

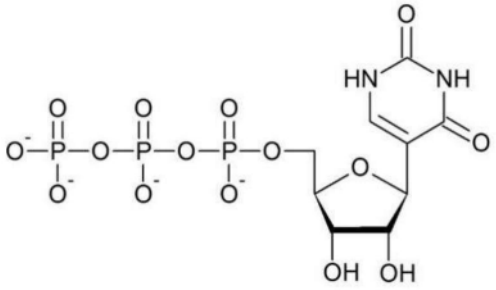
[0190] 如本文所用,术语"尿苷"描述可以在RNA中出现的核苷之一。尿苷的结构是:

[0191] 。

[0192] UTP(尿苷5'-三磷酸)具有以下结构:

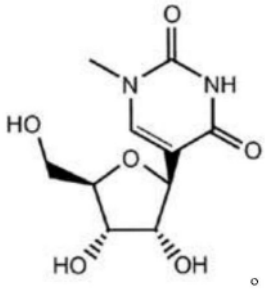
[0193] 。

[0194] 修饰的尿苷也是可以在RNA中出现的核苷之一。一种这样修饰的尿苷是假-UTP(假尿苷5'-三磷酸),其具有以下结构:

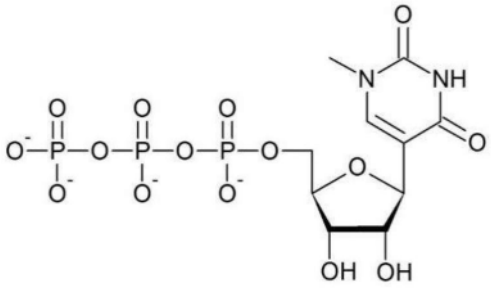
[0195] 

[0196] “假尿苷”是示例性修饰的核苷,其是尿苷的异构体,其中尿嘧啶通过碳-碳键而不是氮-碳糖苷键连接至戊糖环。

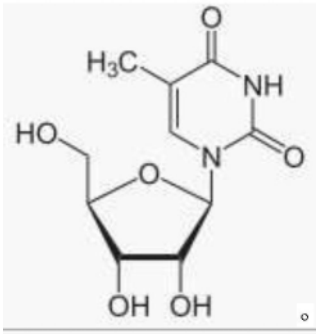
[0197] 另一示例性修饰的尿苷是N1-甲基-假尿苷(1mΨ),其具有以下结构:

[0198] 

[0199] N1-甲基-假-UTP具有以下结构:

[0200] 

[0201] 另一示例性修饰的尿苷是5-甲基-尿苷(m5U),其具有以下结构:

[0202] 

[0203] 在某些优选的实施方案中,本文描述的rRNA中的一个或多个尿苷被修饰的核苷代替。在一些实施方案中,所述修饰的核苷是修饰的尿苷。

[0204] 在某些优选的实施方案中,RNA包含修饰的尿苷代替至少一个尿苷。在一些实施方案中,RNA包含修饰的尿苷代替每个尿苷。

[0205] 在某些优选的实施方案中,修饰的尿苷独立地选自假尿苷(ψ)、N1-甲基-假尿苷(1mΨ)和5-甲基-尿苷(m5U)。在一些实施方案中,修饰的尿苷包含假尿苷(Ψ)。在一些实施方案中,修饰的尿苷包含N1-甲基-假尿苷(1mΨ)。在一些实施方案中,修饰的尿苷包含5-甲

基-尿苷(m5U)。在一些实施方案中,RNA可以包含不止一种类型的修饰的尿苷,并且修饰的尿苷独立地选自假尿苷(Ψ)、N1-甲基-假尿苷(1mΨ)和5-甲基-尿苷(m5U)。在一些实施方案中,修饰的尿苷包含假尿苷(Ψ)和N1-甲基-假尿苷(1mΨ)。在一些实施方案中,修饰的尿苷包含假尿苷(Ψ)和5-甲基-尿苷(m5U)。在一些实施方案中,修饰的尿苷包含N1-甲基-假尿苷(1mΨ)和5-甲基-尿苷(m5U)。在一些实施方案中,修饰的尿苷包含假尿苷(Ψ)、N1-甲基-假尿苷(1mΨ)和5-甲基-尿苷(m5U)。

[0206] 在某些优选的实施方案中,代替rRNA中的一个或多个,例如全部尿苷的修饰的核苷可以是以下修饰的尿苷中的任何一个或多个:3-甲基-尿苷(m^3U)、5-甲氧基-尿苷(mo^5U)、5-氮杂-尿苷、6-氮杂-尿苷、2-硫代-5-氮杂-尿苷、2-硫代-尿苷(s^2U)、4-硫代-尿苷(s^4U)、4-硫代-假尿苷、2-硫代-假尿苷、5-羟基-尿苷(ho^5U)、5-氨基烯丙基-尿苷、5-卤代-尿苷(例如,5-碘-尿苷或5-溴-尿苷)、尿苷5-氧乙酸(cmo^5U)、尿苷5-氧乙酸甲酯($mcmo^5U$)、5-羧甲基-尿苷(cm^5U)、1-羧甲基-假尿苷、5-羧基羟甲基-尿苷(chm^5U)、5-羧基羟甲基-尿苷甲酯($mchm^5U$)、5-甲氧基羰基甲基-尿苷(mcm^5U)、5-甲氧基羰基甲基-2-硫代-尿苷(mcm^5s^2U)、5-氨基甲基-2-硫代-尿苷(nm^5s^2U)、5-甲基氨基甲基-尿苷(mnm^5U)、1-乙基-假尿苷、5-甲基氨基甲基-2-硫代-尿苷(mnm^5s^2U)、5-甲基氨基甲基-2-硒-尿苷(mnm^5se^2U)、5-氨甲酰基甲基-尿苷(ncm^5U)、5-羧甲基氨基甲基-尿苷($cmnm^5U$)、5-羧甲基氨基甲基-2-硫代-尿苷($cmnm^5s^2U$)、5-丙炔基-尿苷、1-丙炔基-假尿苷、5-牛磺酸甲基-尿苷(τm^5U)、1-牛磺酸甲基-假尿苷、5-牛磺酸甲基-2-硫代-尿苷(τm5s2U)、1-牛磺酸甲基-4-硫代-假尿苷)、5-甲基-2-硫代-尿苷(m^5s^2U)、1-甲基-4-硫代-假尿苷($m^1s^4\psi$)、4-硫代-1-甲基-假尿苷、3-甲基-假尿苷($m^3\psi$)、2-硫代-1-甲基-假尿苷、1-甲基-1-脱氮-假尿苷、2-硫代-1-甲基-1-脱氮-假尿苷、二氢尿苷(D)、二氢假尿苷、5,6-二氢尿苷、5-甲基-二氢尿苷(m^5D)、2-硫代-二氢尿苷、2-硫代-二氢假尿苷、2-甲氧基-尿苷、2-甲氧基-4-硫代-尿苷、4-甲氧基-假尿苷、4-甲氧基-2-硫代-假尿苷、N1-甲基-假尿苷、3-(3-氨基-3-羧丙基)尿苷(acp^3U)、1-甲基-3-(3-氨基-3-羧丙基)-假尿苷($acp^3\psi$)、5-(异戊烯基氨基甲基)尿苷(inm^5U)、5-(异戊烯基氨基甲基)-2-硫代-尿苷(inm^5s^2U)、α-硫代-尿苷、2′-O-甲基-尿苷(Um)、5,2′-O-二甲基-尿苷(m^5Um)、2′-O-甲基-假尿苷(ψm)、2-硫代-2′-O-甲基-尿苷(s^2Um)、5-甲氧基羰基甲基-2′-O-甲基-尿苷(mcm^5Um)、5-氨甲酰基甲基-2′-O-甲基-尿苷(ncm^5Um)、5-羧甲基氨基甲基-2′-O-甲基-尿苷($cmnm^5Um$)、3,2′-O-二甲基-尿苷(m^3Um)、5-(异戊烯基氨基甲基)-2′-O-甲基-尿苷(inm^5Um)、1-硫代-尿苷、脱氧胸苷、2'-F-ara-尿苷、2′-F-尿苷、2′-OH-ara-尿苷、5-(2-甲氧羰基乙烯基)尿苷、5-[3-(1-E-丙烯基氨基)尿苷或本领域已知的任何其他修饰的尿苷。

[0207] 在一实施方案中,rRNA包含其他修饰的核苷,或者包含进一步修饰的核苷,例如,修饰的胞苷,如上描述的那些。例如,在一实施方案中,在所述rRNA中,5-甲基胞苷部分或完全,优选完全替代胞苷。在一实施方案中,所述rRNA包含5-甲基胞苷以及选自假尿苷(ψ)、N1-甲基-假尿苷(1mΨ)和5-甲基-尿苷(m5U)中的一种或多种。在一实施方案中,所述rRNA包含5-甲基胞苷和N1-甲基-假尿苷(1mΨ)。在一些实施方案中,所述rRNA包含5-甲基胞苷代替每个胞苷,以及N1-甲基-假尿苷(1mΨ)代替每个尿苷。

[0208] 功能性非结构蛋白

[0209] 术语“非结构蛋白”涉及由病毒编码但不属于病毒颗粒的蛋白。这个术语通常包括

病毒用于自复制的各种酶和转录因子,如RNA复制酶或其他模板定向聚合酶(template-directed polymerase)。术语“非结构蛋白”包括每一种共翻译或翻译后修饰形式,包括非结构蛋白的碳水化合物修饰(如糖基化)和脂质修饰形式,并且优选涉及“甲病毒非结构蛋白”。

[0210] 在一些实施方案中,术语“甲病毒非结构蛋白”是指甲病毒来源的任何一种或多种单独的非结构蛋白(nsP1、nsP2、nsP3、nsP4),或者包含一种以上甲病毒来源的非结构蛋白的多肽序列的多聚蛋白。在一些实施方案中,“甲病毒非结构蛋白”是指nsP123和/或nsP4。在其他实施方案中,“甲病毒非结构蛋白”是指nsP1234。在一实施方案中,由开放阅读框编码的所关注的蛋白由所有nsP1、nsP2、nsP3和nsP4组成,作为一个单一的、任选可切割的多聚蛋白:nsP1234。在一实施方案中,由开放阅读框编码的所关注的蛋白由nsP1、nsP2和nsP3组成,作为一个单一的、任选可切割的多聚蛋白:nsP123。在该实施方案中,nsP4可以是另一种所关注的蛋白,并且可以由另一个开放阅读框编码。

[0211] 在一些实施方案中,非结构蛋白能够形成复合物或缔合物,例如,在宿主细胞中。在一些实施方案中,“甲病毒非结构蛋白”是指nsP123(同义词P123)和nsP4的复合物或缔合物。在一些实施方案中,“甲病毒非结构蛋白”是指nsP1、nsP2和nsP3的复合物或缔合物。在一些实施方案中,“甲病毒非结构蛋白”是指nsP1、nsP2、nsP3和nsP4的复合物或缔合物。在一些实施方案中,“甲病毒非结构蛋白”是指选自nsP1、nsP2、nsP3和nsP4的任何一个或多个的复合物或缔合物。在一些实施方案中,甲病毒非结构蛋白至少包含nsP4。

[0212] 术语“复合物”或“缔合物”是指在空间上接近的两个或更多个相同或不同的蛋白分子。复合物的蛋白优选彼此直接或间接物理或物理化学接触。复合物或缔合物可以由多种不同的蛋白(异源多聚体)和/或一种特定蛋白的多个拷贝(同源多聚体)组成。在甲病毒非结构蛋白的背景下,术语“复合物或缔合物”描述至少两个蛋白分子的集合,其中至少一个是甲病毒非结构蛋白。复合物或缔合物可以由一种特定蛋白的多个拷贝(同源多聚体)和/或多种不同的蛋白(异源多聚体)组成。在多聚体的上下文中,“多”表示不止一个,如两个、三个、四个、五个、六个、七个、八个、九个、十个或十个以上。

[0213] 术语“功能性非结构蛋白”包括具有复制酶功能的非结构蛋白。因此,“功能性非结构蛋白”包括甲病毒复制酶。“复制酶功能”包括RNA依赖性RNA聚合酶(RdRP)的功能,即能够催化基于(+)链RNA模板的(-)链RNA合成的酶,和/或能够催化基于(-)链RNA的(+)链RNA合成的酶。因此,术语“功能性非结构蛋白”可以指使用(+)链(例如基因组)RNA作为模板合成(-)链RNA的蛋白或复合物,使用基因组RNA的(-)链互补物作为模板合成新(+)链RNA的蛋白或复合物,和/或使用基因组RNA的(-)链互补物的片段作为模板合成亚基因组转录物的蛋白或复合物。功能性非结构蛋白还可以具有一种或多种额外的功能,例如,蛋白酶(用于自切割)、复制酶、末端腺苷酰基转移酶(adenylyltransferase)(用于poly(A)尾添加)、甲基转移酶和鸟苷酸转移酶(用于提供具有5’-帽的核酸)、核定位位点、三磷酸酶(Gould et al.,2010,Antiviral Res.87:111-124;Rupp et al.,2015,J.Gen.Virol.96:2483-500)。

[0214] 术语“复制酶”包括RNA依赖性RNA聚合酶。根据本发明,术语“复制酶”包括“甲病毒复制酶”,包括来自天然存在的甲病毒(自然界中发现的甲病毒)的RNA依赖性RNA聚合酶以及来自甲病毒的变体或衍生物(如来自减毒的甲病毒)的RNA依赖性RNA聚合酶。

[0215] 术语“复制酶”包括甲病毒复制酶的所有变体,特别是翻译后修饰的变体、构象、同

种型或同源物,其由甲病毒感染的细胞表达或由已经用编码甲病毒复制酶的核酸转染的细胞表达。此外,术语“复制酶”包括已经产生并且可以通过重组方法产生的所有形式的复制酶。例如,可以通过重组方法制备包含标签的复制酶,所述标签便于在实验室中检测和/或纯化复制酶,例如,myc-标签、HA-标签或寡组氨酸标签(His-标签)。

[0216] 任选地,甲病毒复制酶在功能上还通过与甲病毒保守序列元件1(CSE 1)或其互补序列、保守序列元件2(CSE 2)或其互补序列、保守序列元件3(CSE 3)或其互补序列、保守序列元件4(CSE 4)或其互补序列中的任何一个或多个结合的能力来定义。优选地,复制酶能够与CSE 2[即,(+)链]和/或CSE 4[即,(+)链]结合,或者与CSE 1的互补物[即,(-)链]和/或CSE 3的互补物[即,(-)链]结合。

[0217] 甲病毒复制酶的来源并不局限于任何特定的甲病毒。在一优选实施方案中,甲病毒元件包含来自塞姆利基森林病毒的非结构蛋白,其包括天然存在的塞姆利基森林病毒以及塞姆利基森林病毒的变体或衍生物,如减毒的塞姆利基森林病毒。在一可选的优选实施方案中,甲病毒复制酶包含来自辛德毕斯病毒的非结构蛋白,其包括天然存在的辛德毕斯病毒以及辛德毕斯病毒的变体或衍生物,如减毒的辛德毕斯病毒。在一可选的优选实施方案中,甲病毒复制酶包含来自委内瑞拉马脑炎病毒(VEEV)的非结构蛋白,其包括天然存在的VEEV以及VEEV的变体或衍生物,如减毒的VEEV。在一可选的优选实施方案中,甲病毒复制酶包含来自基孔肯雅病毒(CHIKV)的非结构蛋白,其包括天然存在的CHIKV以及CHIKV的变体或衍生物,如减毒的CHIKV。

[0218] 复制酶还可以包含来自一种以上病毒(例如来自一种以上甲病毒)的非结构蛋白。因此,本发明同样包括含有甲病毒非结构蛋白并具有复制酶功能的异源复合物或缔合物。仅为了说明的目的,复制酶可以包含来自第一种甲病毒的一种或多种非结构蛋白(例如,nsP1、nsP2),以及来自第二种甲病毒的一种或多种非结构蛋白(nsP3、nsP4)。来自一种以上不同甲病毒的非结构蛋白可以由不同的开放阅读框编码,或者可以由单个开放阅读框编码为多聚蛋白,例如,nsP1234。

[0219] 在一些实施方案中,功能性非结构蛋白能够在表达功能性非结构蛋白的细胞中形成膜复制复合物和/或液泡。

[0220] 如果功能性非结构蛋白,即具有复制酶功能的非结构蛋白,由本发明的核酸分子编码,则优选复制子的亚基因组启动子(如果存在)与所述复制酶相容。在这种情况下相容表示复制酶能够识别亚基因组启动子(如果存在)。在一实施方案中,当亚基因组启动子对于衍生复制酶的病毒是天然的时,即这些序列的天然来源是相同病毒时,实现这一点。在一可选实施方案中,亚基因组启动子对于衍生病毒复制酶的病毒不是天然的,条件是病毒复制酶能够识别亚基因组启动子。换句话说,复制酶与亚基因组启动子相容(跨病毒相容性)。关于源自不同甲病毒的亚基因组启动子和复制酶的跨病毒相容性的实例在本领域中是已知的。只要存在跨病毒相容性,亚基因组启动子和复制酶的任何组合都是可能的。从事本发明工作的技术人员可以容易地通过将待测试的复制酶与RNA一起孵育来测试跨病毒相容性,其中RNA具有待测试的亚基因组启动子,在适合从亚基因组启动子合成RNA的条件下。如果制备了亚基因组转录物,则确定亚基因组启动子和复制酶是相容的。跨病毒相容性的各种实例是已知的。

[0221] 复制子可以优选由功能性非结构蛋白复制。特别地,编码功能性非结构蛋白的RNA

复制子可以由复制子编码的功能性非结构蛋白复制。在一优选实施方案中,RNA复制子包含编码功能性甲病毒非结构蛋白的开放阅读框。在一实施方案中,复制子包含编码所关注的蛋白的另一个开放阅读框。这个实施方案特别适合本发明产生所关注的蛋白的一些方法。在一实施方案中,编码所关注的蛋白的另一个开放阅读框位于5'复制识别序列的下游和IRES的上游(以及编码来自自复制病毒的功能性非结构蛋白的开放阅读框的上游)和/或编码来自自复制病毒的功能性非结构蛋白的开放阅读框的下游。位于5'复制识别序列的下游和IRES的上游(以及编码来自自复制病毒的功能性非结构蛋白的开放阅读框的上游)的编码所关注的蛋白的另一个开放阅读框可以表达为融合蛋白,其具有5'复制识别序列编码的序列。位于5'复制识别序列的下游和IRES的上游(以及编码来自自复制病毒的功能性非结构蛋白的开放阅读框的上游)的编码所关注的蛋白的另一个开放阅读框可以或可以不由亚基因组启动子控制。位于编码来自自复制病毒的功能性非结构蛋白的开放阅读框下游的编码一种或多种所关注的蛋白的一个或多个其他开放阅读框一般由亚基因组启动子控制。

[0222] 优选编码功能性非结构蛋白的开放阅读框不与5'复制识别序列重叠。在一实施方案中,编码功能性非结构蛋白的开放阅读框不与亚基因组启动子(如果存在)重叠。其实施方案在WO 2017/162460中公开,援引加入本文。

[0223] 复制所需的序列元件和蛋白编码区的解偶联

[0224] 通用的甲病毒衍生载体难以开发,因为编码nsP1234的开放阅读框与甲病毒基因组的5'复制识别序列(nsP1的编码序列)重叠,并且通常还与包含CSE 3的亚基因组启动子(nsP4的编码序列)重叠。

[0225] 本文所述的RNA复制子一般包含复制酶复制所需的序列元件,特别是5'复制识别序列。在一实施方案中,非结构蛋白的编码序列受IRES的控制,因此IRES位于非结构蛋白编码序列的上游。因此,在一实施方案中,通常与甲病毒非结构蛋白N-末端片段的编码序列重叠的5'复制识别序列位于IRES的上游,并且与非结构蛋白的编码序列不重叠。

[0226] 在一实施方案中,5'复制识别序列的编码序列如nsP1编码序列与位于IRES上游的所关注的基因符合读框地融合。

[0227] 在一实施方案中,5'复制识别序列不编码任何蛋白或其片段,如甲病毒非结构蛋白或其片段。因此,在本发明的RNA复制子中,复制酶复制所需的序列元件和蛋白编码区可以解偶联。与天然病毒基因组RNA(例如,天然甲病毒基因组RNA)相比,解偶联可以通过去除5'复制识别序列中的至少一个起始密码子来实现。

[0228] 因此,rRNA可以包含5'复制识别序列,其中所述5'复制识别序列的特征在于,与天然病毒5'复制识别序列(例如,天然甲病毒5'复制识别序列)相比,它包括去除至少一个起始密码子。

[0229] 与天然病毒5'复制识别序列相比,特征在于包括去除至少一个起始密码子的5'复制识别序列在本文中可以称作"修饰的5'复制识别序列"或"本发明的5'复制识别序列"。如下所述,本发明的5'复制识别序列可以任选地特征在于存在一个或多个额外的核苷酸变化,如通过本发明的方法检测到的那些。

[0230] 能够被复制酶(优选甲病毒复制酶)复制的核酸构建体称为可复制RNA或复制子。根据本发明,术语"复制子"定义一种RNA分子,其可以被RNA依赖性RNA聚合酶复制,在没有DNA中间体的情况下产生一个或多个相同或基本上相同的RNA复制子拷贝。"没有DNA中间

体”表示在形成RNA复制子拷贝的过程中不形成复制子的脱氧核糖核酸(DNA)拷贝或互补物,和/或在形成RNA复制子拷贝的过程中不使用脱氧核糖核酸(DNA)分子或其互补物作为模板。复制酶功能通常由功能性非结构蛋白提供,例如,功能性甲病毒非结构蛋白。

[0231] 根据本发明,术语“可以复制”和“能够复制”一般描述可以制备核酸的一个或多个相同或基本上相同的拷贝。当与术语“复制酶”一起使用时,如在“能够被复制酶复制”中,术语“可以复制”和“能够复制”描述核酸分子(例如RNA复制子)相对于复制酶的功能特征。这些功能特征包括以下至少一项:(i)复制酶能够识别复制子和(ii)复制酶能够充当RNA依赖性RNA聚合酶(RdRP)。优选地,复制酶能够(i)识别复制子和(ii)充当RNA依赖性RNA聚合酶。

[0232] 表述“能够识别”描述了复制酶能够与复制子发生物理联系,并且优选地,复制酶能够与复制子结合,通常是非共价结合。术语“结合”可以表示复制酶具有与保守序列元件1(CSE 1)或其互补序列(如果复制子包含)、保守序列元件2(CSE 2)或其互补序列(如果复制子包含)、保守序列元件3(CSE 3)或其互补序列(如果复制子包含)、保守序列元件4(CSE 4)或其互补序列(如果复制子包含)中的任何一个或多个结合的能力。优选地,复制酶能够与CSE 2[即,(+)链]和/或CSE 4[即,(+)链]结合,或者与CSE 1的互补物[即,(-)链]和/或CSE 3的互补物[即,(-)链]结合。

[0233] 在一实施方案中,表述“能够充当RdRP”表示复制酶能够催化病毒基因组(+)链RNA的(-)链互补物的合成,其中(+)链RNA具有模板功能,和/或复制酶能够催化(+)链病毒基因组RNA的合成,其中(-)链RNA具有模板功能。一般来说,表述“能够充当RdRP”还可以包括复制酶能够催化(+)链亚基因组转录物的合成,其中(-)链RNA具有模板功能,并且其中(+)链亚基因组转录物的合成通常在亚基因组启动子处启动。在一实施方案中,所述病毒是甲病毒。

[0234] 表述“能够结合”和“能够充当RdRP”是指在正常生理条件下的能力。特别地,它们是指细胞内的条件,所述细胞表达功能性非结构蛋白或已用编码功能性非结构蛋白的核酸转染。所述细胞优选是真核细胞。结合的能力和/或充当RdRP的能力可以通过实验测试,例如在无细胞体外系统或真核细胞中。任选地,所述真核细胞是来自代表复制酶来源的特定病毒具有感染性的物种的细胞。例如,当使用来自对人具有感染性的特定病毒的病毒复制酶时,正常生理条件是人细胞中的条件。更优选地,真核细胞(在一实例中,人细胞)来自代表复制酶来源的特定病毒具有感染性的相同组织或器官。

[0235] 如本文所用,“与天然甲病毒序列相比”和类似术语是指为天然甲病毒序列变体的序列。所述变体本身通常不是天然甲病毒序列。

[0236] 在一实施方案中,RNA复制子包含3’复制识别序列。3’复制识别序列是可以被功能性非结构蛋白识别的核酸序列。换句话说,功能性非结构蛋白能够识别3’复制识别序列。优选地,3’复制识别序列位于复制子的3’末端(如果复制子不包含poly(A)尾),或者紧邻poly(A)尾的上游(如果复制子包含poly(A)尾)。在一实施方案中,3’复制识别序列由CSE 4组成或包含CSE 4。

[0237] 在一实施方案中,5’复制识别序列和3’复制识别序列能够在功能性非结构蛋白的存在下指导本发明的RNA复制子的复制。因此,当单独存在或优选一起存在时,这些识别序列在功能性非结构蛋白的存在下指导RNA复制子的复制。

[0238] 优选提供功能性非结构蛋白,其能够识别复制子的5’复制识别序列和3’复制识别

序列。在一实施方案中,当3’复制识别序列对于衍生功能性甲病毒非结构蛋白的甲病毒是天然的时,以及当5’复制识别序列对于衍生功能性甲病毒非结构蛋白的甲病毒是天然的或是衍生功能性甲病毒非结构蛋白的甲病毒的天然5’复制识别序列的变体时,可以实现这一点。天然表示这些序列的自然来源是同一种甲病毒。在一可选实施方案中,5’复制识别序列和/或3’复制识别序列对于衍生功能性甲病毒非结构蛋白的甲病毒不是天然的,条件是功能性甲病毒非结构蛋白能够识别复制子的5’复制识别序列和3’复制识别序列。换句话说,功能性甲病毒非结构蛋白与5’复制识别序列和3’复制识别序列相容。当非天然的功能性甲病毒非结构蛋白能够识别各自的序列或序列元件时,则称功能性甲病毒非结合蛋白是相容的(跨病毒相容性)。(3’/5’)复制识别序列和CSE分别与功能性甲病毒非结构蛋白的任何组合都是可能的,只要存在跨病毒相容性。从事本发明工作的技术人员可以容易地通过将待测试的功能性甲病毒非结构与RNA一起孵育来测试跨病毒相容性,其中RNA具有待测试的3’-和5’复制识别序列,在适合RNA复制的条件下,例如在合适的宿主细胞中。如果发生复制,则确定(3’/5’)复制识别序列和功能性甲病毒非结构蛋白是相容的。

[0239] 去除5’复制识别序列内的至少一个起始密码子提供了几个优点。编码nsP1*(nsP1的N-末端片段)的核酸序列中缺少起始密码子通常会导致nsP1*不被翻译。此外,由于nsP1*不被翻译,编码所关注的蛋白(“GOI 2”)的开放阅读框是核糖体可接近的最上游开放阅读框;因此,当复制子存在于细胞中时,翻译在编码所关注的基因的开放阅读框(RNA)的第一个AUG处开始。

[0240] 至少一个起始密码子的去除可以通过本领域已知的任何合适方法来实现。例如,可以在计算机上设计编码本发明的复制子的合适DNA分子,即特征在于去除起始密码子,随后在体外合成(基因合成);或者,可以通过编码复制子的DNA序列的定点诱变获得合适的DNA分子。在任何情况下,相应的DNA分子可以用作体外转录的模板,从而提供本发明的复制子。

[0241] 与天然5’复制识别序列相比去除至少一个起始密码子不受特别限制,可以选自任何核苷酸修饰,包括一个或多个核苷酸的替代(在DNA水平上,包括起始密码子的A和/或T和/或G的替代);一个或多个核苷酸的缺失(在DNA水平上,包括起始密码子的A和/或T和/或G的缺失),以及一个或多个核苷酸的插入(在DNA水平上,包括起始密码子的A和T之间和/或T和G之间一个或多个核苷酸的插入)。无论核苷酸修饰是替代、插入或缺失,核苷酸修饰都不得导致形成新的起始密码子(作为说明性实例:在DNA水平上,插入不得是ATG的插入)。

[0242] 特征在于去除至少一个起始密码子的RNA复制子的5’复制识别序列(即本发明的修饰的5’复制识别序列)优选是自然界中发现的甲病毒基因组的5’复制识别序列的变体。在一实施方案中,本发明的修饰的5’复制识别序列优选特征在于与自然界中发现的至少一种甲病毒基因组的5’复制识别序列的序列相同性程度为80%或更高,优选85%或更高,更优选90%或更高,甚至更优选95%或更高。

[0243] 在一实施方案中,特征可以在于去除至少一个起始密码子的RNA复制子的5’复制识别序列包含与甲病毒的5’末端(即甲病毒基因组的5’末端)约250个核苷酸同源的序列。在一优选实施方案中,其包含与甲病毒的5’末端(即甲病毒基因组的5’末端)约250-500个,优选约300-500个核苷酸同源的序列。“甲病毒基因组的5’末端”表示从甲病毒基因组的最上游核苷酸开始并包括其在内的核酸序列。换句话说,甲病毒基因组的最上游核苷酸被指

定为核苷酸1,例如,“甲病毒基因组5’末端的250个核苷酸”表示甲病毒基因组的核苷酸1-250。在一实施方案中,RNA复制子的5’复制识别序列的特征在于与自然界中发现的至少一种甲病毒基因组的5’末端至少250个核苷酸的序列相同性程度为80％或更高,优选85％或更高,更优选90％或更高,甚至更优选95％或更高。至少250个核苷酸包括例如250个核苷酸、300个核苷酸、400个核苷酸、500个核苷酸。

[0244] 自然界中发现的甲病毒的5’复制识别序列通常特征在于至少一个起始密码子和/或保守的二级结构基序。例如,塞姆利基森林病毒(SFV)的天然5’复制识别序列包含5个特定AUG碱基三联体。根据Frolov et al.,2001,RNA 7:1638-1651,通过MFOLD分析显示,塞姆利基森林病毒的天然5’复制识别序列预测会形成四个茎环(SL),称为茎环1-4(SL1、SL2、SL3、SL4)。根据Frolov等人,通过MFOLD分析显示,一种不同的甲病毒辛德毕斯病毒的天然5’复制识别序列预测会形成四个茎环:SL1、SL2、SL3、SL4。

[0245] 已知甲病毒基因组的5’末端包含序列元件,这些序列元件使得能够通过功能性甲病毒非结构蛋白复制甲病毒基因组。在本发明的一个实施方案中,RNA复制子的5’复制识别序列包含与甲病毒的保守序列元件1(CSE 1)同源的序列和/或与保守序列元件2(CSE 2)同源的序列。

[0246] 甲病毒基因组RNA的保守序列元件2(CSE 2)通常由SL3和SL4表示,SL3和SL4之前是SL2,其至少包含编码甲病毒非结构蛋白nsP1的第一个氨基酸残基的天然起始密码子。然而,在本说明书中,在一些实施方案中,甲病毒基因组RNA的保守序列元件2(CSE 2)是指从SL2到SL4的区域,并且包含编码甲病毒非结构蛋白nsP1的第一个氨基酸残基的天然起始密码子。在一优选实施方案中,RNA复制子包含CSE 2或与CSE 2同源的序列。在一实施方案中,RNA复制子包含与CSE 2同源的序列,其优选特征在于与自然界中发现的至少一种甲病毒的CSE 2序列的序列相同性程度为80％或更高,优选85％或更高,更优选90％或更高,甚至更优选95％或更高。

[0247] 在一实施方案中,5’复制识别序列包含与甲病毒的CSE 2同源的序列。甲病毒的CSE 2可以包含来自甲病毒的非结构蛋白的开放阅读框的片段。

[0248] 因此,在一实施方案中,RNA复制子的特征在于其包含与来自甲病毒的非结构蛋白的开放阅读框或其片段同源的序列。与非结构蛋白的开放阅读框或其片段同源的序列通常是自然界中发现的甲病毒的非结构蛋白的开放阅读框或其片段的变体。在一实施方案中,与非结构蛋白的开放阅读框或其片段同源的序列优选特征在于与自然界中发现的至少一种甲病毒的非结构蛋白的开放阅读框或其片段的序列相同性程度为80％或更高,优选85％或更高,更优选90％或更高,甚至更优选95％或更高。

[0249] 在一实施方案中,与本发明的复制子包含的非结构蛋白的开放阅读框同源的序列不包含非结构蛋白的天然起始密码子,更优选不包含非结构蛋白的任何起始密码子。在一实施方案中,与CSE 2同源的序列的特征在于与天然甲病毒CSE 2序列相比去除所有起始密码子。因此,与CSE 2同源的序列优选不包含任何起始密码子。

[0250] 当与开放阅读框同源的序列不包含任何起始密码子时,与开放阅读框同源的序列本身不是开放阅读框,因为它不作为翻译的模板。

[0251] 在一实施方案中,5’复制识别序列包含与来自甲病毒的非结构蛋白的开放阅读框或其片段同源的序列,其中与来自甲病毒的非结构蛋白的开放阅读框或其片段同源的序列

的特征在于与天然甲病毒序列相比,它包括去除至少一个起始密码子。

[0252] 在一实施方案中,与来自甲病毒的非结构蛋白的开放阅读框或其片段同源的序列的特征在于其包括至少去除非结构蛋白的开放阅读框的天然起始密码子。优选地,它的特征在于它包括至少去除编码nsP1的开放阅读框的天然起始密码子。

[0253] 天然起始密码子是AUG碱基三联体,当宿主细胞中存在RNA时,宿主细胞中核糖体上的翻译从该碱基三联体开始。换句话说,天然起始密码子是核糖体蛋白合成过程中翻译的第一个碱基三联体,例如,在已经用包含天然起始密码子的RNA接种的宿主细胞中。在一实施方案中,宿主细胞是来自真核物种的细胞,所述真核物种是包含天然甲病毒5'复制识别序列的特定甲病毒的天然宿主。在一实施方案中,宿主细胞是来自细胞系"BHK21[C13](ATCC®CCL10™)"的BHK21细胞,可获得自美国弗吉尼亚州马纳萨斯的美国典型培养物保藏中心。

[0254] 许多甲病毒的基因组已经完全测序并可公开获得,并且由这些基因组编码的非结构蛋白的序列也可公开获得。这样的序列信息允许在计算机上确定天然起始密码子。

[0255] 在一实施方案中,与来自甲病毒的非结构蛋白的开放阅读框或其片段同源的序列的特征在于其包括去除非结构蛋白的开放阅读框的天然起始密码子以外的一个或多个起始密码子。在一实施方案中,所述核酸序列的特征在于额外去除天然起始密码子。例如,除了去除天然起始密码子之外,还可以去除任何一个或两个或三个或四个或四个以上(例如,五个)起始密码子。

[0256] 如果复制子的特征在于去除非结构蛋白的开放阅读框的天然起始密码子,以及任选地去除天然起始密码子以外的一个或多个起始密码子,则与开放阅读框同源的序列本身不是开放阅读框,因为它不作为翻译的模板。

[0257] 优选除了去除天然起始密码子之外,去除的天然起始密码子以外的一个或多个起始密码子优选选自具有启动翻译潜力的AUG碱基三联体。具有启动翻译潜力的AUG碱基三联体可以称作"潜在起始密码子"。给定的AUG碱基三联体是否具有启动翻译的潜力可以通过计算机或基于细胞的体外测定来确定。

[0258] 在一实施方案中,在计算机上确定给定的AUG碱基三联体是否具有启动翻译的潜力:在该实施方案中,检查核苷酸序列,如果AUG碱基三联体是AUGG序列的一部分,优选Kozak序列的一部分,则确定其具有启动翻译的潜力。

[0259] 在一实施方案中,在基于细胞的体外测定中确定给定的AUG碱基三联体是否具有启动翻译的潜力:将特征在于去除天然起始密码子并在去除天然起始密码子的位置下游包含给定AUG碱基三联体的RNA复制子引入宿主细胞。在一实施方案中,宿主细胞是来自真核物种的细胞,所述真核物种是包含天然甲病毒5'复制识别序列的特定甲病毒的天然宿主。在一优选实施方案中,宿主细胞是来自细胞系"BHK21[C13](ATCC®CCL10™)"的BHK21细胞,可获得自美国弗吉尼亚州马纳萨斯的美国典型培养物保藏中心。优选在去除天然起始密码子的位置和给定的AUG碱基三联体之间没有另外的AUG碱基三联体。如果将RNA复制子(其特征在于去除天然起始密码子并包含给定的AUG碱基三联体)引入宿主细胞后,在给定的AUG碱基三联体处启动翻译,则确定给定的AUG碱基三联体具有启动翻译的潜力。可以通过本领域已知的任何合适方法确定翻译是否启动。例如,复制子可以在给定的AUG碱基三联体的下游并与给定的AUG碱基三联体符合读框(in-frame)地编码便于检测翻译产物(如果

有)的标签,例如myc-标签或HA-标签;可以例如通过蛋白印迹确定是否存在具有编码标签的表达产物。在这个实施方案中,优选在给定的AUG碱基三联体和编码标签的核酸序列之间没有另外的AUG碱基三联体。对于一个以上给定的AUG碱基三联体可以单独进行基于细胞的体外测定:在每种情况下,优选在去除天然起始密码子的位置和给定的AUG碱基三联体之间没有另外的AUG碱基三联体。这可以通过去除天然起始密码子的去除位置和给定的AUG碱基三联体之间的所有AUG碱基三联体(如果有)来实现。因此,给定的AUG碱基三联体是去除天然起始密码子的位置下游的第一个AUG碱基三联体。

[0260] 优选地,本发明的RNA复制子的5'复制识别序列的特征在于去除所有潜在起始密码子。因此,根据本发明,5'复制识别序列优选不包含可以翻译为蛋白的开放阅读框。

[0261] 在一实施方案中,本发明的RNA复制子的5'复制识别序列的特征在于二级结构,其等同于病毒基因组RNA的5'复制识别序列的(预测)二级结构。为此,RNA复制子可以包含一个或多个核苷酸变化,以补偿由去除至少一个起始密码子引入的一个或多个茎环内的核苷酸配对破坏。

[0262] 在一实施方案中,本发明的RNA复制子的5'复制识别序列的特征在于二级结构,其等同于甲病毒基因组RNA的5'复制识别序列的二级结构。在一优选实施方案中,本发明的RNA复制子的5'复制识别序列的特征在于预测的二级结构,其等同于甲病毒基因组RNA的5'复制识别序列的预测二级结构。根据本发明,RNA分子的二级结构优选通过RNA二级结构预测的网络服务器http://rna.urmc.rochester.edu/RNAstructureWeb/Servers/Predict1/Predict1.html预测。

[0263] 通过比较特征在于与天然甲病毒5'复制识别序列相比去除至少一个起始密码子的RNA复制子的5'复制识别序列的二级结构或预测二价结构,可以鉴定是否存在核苷酸配对破坏。例如,与天然甲病毒5'复制识别序列相比,在给定位置可以缺少至少一个碱基对,例如茎环,特别是茎环的茎内的碱基对。

[0264] 在一实施方案中,5'复制识别序列的一个或多个茎环未被缺失或破坏。更优选地,茎环3和4未被缺失或破坏。优选地,5'复制识别序列的茎环均未被缺失或破坏。

[0265] 在一实施方案中,去除至少一个起始密码子不破坏5'复制识别序列的二级结构。在一可选实施方案中,去除至少一个起始密码子确实会破坏5'复制识别序列的二级结构。在这个实施方案中,与天然5'复制识别序列相比,去除至少一个起始密码子可能是在给定位置缺少至少一个碱基对的原因,例如茎环内的碱基对。如果与天然5'复制识别序列相比,茎环内部缺少一个碱基对,则确定去除至少一个起始密码子在茎环内引入核苷酸配对破坏。茎环内的碱基对通常是茎环的茎中的碱基对。

[0266] 在一实施方案中,RNA复制子包含一个或多个核苷酸变化,以补偿由去除至少一个起始密码子引入的一个或多个茎环内的核苷酸配对破坏。

[0267] 如果与天然5'复制识别序列相比,去除至少一个起始密码子在茎环内引入核苷酸配对破坏,则可以引入一个或多个核苷酸变化,这些核苷酸变化有望补充核苷酸配对破坏,并且可以将由此获得的二级结构或预测二级结构与天然5'复制识别序列进行比较。

[0268] 基于公知常识和本文的公开,技术人员可以预期某些核苷酸变化会补偿核苷酸配对破坏。例如,如果与天然5'复制识别序列相比,特征在于去除至少一个起始密码子的RNA复制子的给定5'复制识别序列的二级结构或预测二级结构的给定位置上的碱基对被破坏,

则在该位置上恢复碱基对的核苷酸变化(优选不重新引入起始密码子)预期会补偿核苷酸配对破坏。

[0269] 在一实施方案中,复制子的5'复制识别序列不与可翻译核酸序列(即可翻译为肽或蛋白,特别是nsP,特别是nsP1,或者其中任何一种的片段)重叠或不包含可翻译核酸序列。为了使核苷酸序列"可翻译",它需要存在起始密码子:起始密码子编码肽或蛋白的最N-末端氨基酸残基。在一实施方案中,复制子的5'复制识别序列不与编码nsP1的N-末端片段的可翻译核酸序列重叠,或者不包含编码nsP1的N-末端片段的可翻译核酸序列。

[0270] 在某些情况下,RNA复制子包含至少一个亚基因组启动子。在一优选实施方案中,复制子的亚基因组启动子不与可翻译核酸序列(即可翻译为肽或蛋白,特别是nsP,特别是nsP4,或者其中任何一种的片段)重叠或不包含可翻译核酸序列。在一实施方案中,复制子的亚基因组启动子不与编码nsP4的C-末端片段的可翻译核酸序列重叠,或者不包含编码nsP4的C-末端片段的可翻译核酸序列。具有不与可翻译核酸序列(例如可翻译为nsP4的C-末端片段)重叠或不包含可翻译核酸序列的亚基因组启动子的RNA复制子可以通过缺失nsP4的部分编码序列(通常是编码nsP4的N-末端部分的部分)和/或通过去除尚未缺失的nsP4的编码序列部分中的AUG碱基三联体来产生。如果去除nsP4的编码序列或其一部分中的AUG碱基三联体,则去除的AUG碱基三联体优选是潜在起始密码子。或者,如果亚基因组启动子不与编码nsP4的核酸序列重叠,则可以缺失编码nsP4的整个核酸序列。

[0271] 在一实施方案中,RNA复制子不包含编码截短的非结构蛋白(例如,截短的甲病毒非结构蛋白)的开放阅读框。在这个实施方案的背景下,特别优选RNA复制子不包含编码nsP1的N-末端片段的开放阅读框,并且任选不包含编码nsP4的C-末端片段的开放阅读框。nsP1的N-末端片段是截短的甲病毒蛋白;nsP4的C-末端片段也是截短的甲病毒蛋白。

[0272] 在一些实施方案中,本发明的复制子不包含甲病毒基因组5'末端的茎环2(SL2)。根据上文Frolov等人,茎环2是在甲病毒基因组的5'末端发现的保守二级结构,在CSE2上游,但是对于复制是可有可无的。

[0273] 本发明的RNA复制子优选是单链RNA分子。本发明的RNA复制子通常是(+)链RNA分子。在一实施方案中,本发明的RNA分子是分离的核酸分子。本发明的RNA复制子包含至少一个修饰的核苷酸,并且优选包含一个或多个序列变化,特别是通过本文公开的用于鉴定序列变化的方法检测到的序列变化,这些序列变化恢复或改善包含至少一个修饰的核苷酸的rRNA的功能。

[0274] 至少一个编码至少一种所关注的基因产物的开放阅读框

[0275] 在一实施方案中,本发明的RNA复制子包含至少一个编码所关注的基因产物(如所关注的肽或所关注的蛋白)的开放阅读框。优选地,所关注的蛋白由异源核酸序列编码。编码所关注的肽或蛋白的基因同义地称为"所关注的基因"或"转基因"。在各种实施方案中,所关注的肽或蛋白由异源核酸序列编码。根据本发明,术语"异源"是指核酸序列在功能上或结构上不与病毒核酸序列(例如,甲病毒核酸序列)天然连接的事实。

[0276] 本发明的复制子可以编码单个多肽或多种多肽。多种多肽可以被编码为单个多肽(融合多肽)或分隔的多肽。在一些实施方案中,本发明的复制子可以包含一个以上的开放阅读框,每个可以独立地选择是否受亚基因组启动子的控制。或者,多聚蛋白或融合多肽包含由2A自切割肽(例如来自口蹄疫病毒2A蛋白)或蛋白酶切割位点或内含肽分隔的单个多

肽。

[0277] 所关注的蛋白可以例如选自报告蛋白、药学活性肽或蛋白、细胞内干扰素(IFN)信号传导的抑制物。根据本发明,所关注的蛋白优选不包括来自自复制病毒的功能性非结构蛋白,例如,功能性甲病毒非结构蛋白。

[0278] 报告蛋白

[0279] 在一实施方案中,开放阅读框编码报告蛋白,例如,细胞表面表达的蛋白如CD90。在该实施方案中,开放阅读框包含报告基因。可以选择某些基因作为报告基因,因为它们赋予表达它们的细胞或生物体的特征可以很容易地鉴定和测量,或者因为它们是选择性标记。报告基因通常用作某个基因是否已被细胞或生物体群体吸收或表达的指示。优选地,报告基因的表达产物是在视觉上可检测的。常见的视觉上可检测的报告蛋白通常具有荧光或发光蛋白。特定报告基因的实例包括编码水母绿色荧光蛋白(GFP)的基因,该基因使表达它的细胞在蓝光下发出绿光,萤光素酶,其催化与萤光素的反应产生光,以及红色荧光蛋白(RFP)。这些特定报告基因中任何一个的变体都是可能的,只要变体具有视觉上可检测的特性。例如,eGFP是GFP的点突变变体。报告蛋白实施方案特别适合用于测试表达。

[0280] 药学活性基因产物,如肽或蛋白或核酸

[0281] 根据本发明,在一实施方案中,rRNA包含药学活性rRNA或由药学活性rRNA组成。“药学活性RNA”可以是编码药学活性肽或蛋白的RNA。优选地,本发明的RNA复制子编码药学活性肽或蛋白或者其他基因产物。优选地,开放阅读框编码药学活性肽或蛋白。优选地,RNA复制子包含编码药学活性肽或蛋白的开放阅读框,任选受亚基因组启动子的控制。

[0282] 当以治疗有效量给药至受试者时,“药学活性肽或蛋白”对受试者的病况或疾病状态具有积极或有利的影响。优选地,药学活性肽或蛋白具有治疗或姑息(palliative)特性,并且可以将其给药以改善、减轻、缓解、逆转、延迟疾病或病症的一种或多种症状的发生或者减轻疾病或病症的一种或多种症状的严重程度。药学活性肽或蛋白可以具有预防特性,并且可以用来延迟疾病的发生或者减轻这种疾病或病理状态的严重程度。术语“药学活性肽或蛋白”包括完整的蛋白或多肽,并且还可以指其药学活性片段。它还可以包括肽或蛋白的药学活性类似物。术语“药学活性肽或蛋白”包括为抗原的肽和蛋白,即所述肽或蛋白在受试者中引发免疫应答,其可以是治疗性的或者部分或完全保护性的。

[0283] 在一实施方案中,药学活性肽或蛋白是或包含免疫活性化合物或抗原或表位。

[0284] 根据本发明,术语“免疫活性化合物”涉及改变免疫应答的任何化合物,优选通过诱导和/或抑制免疫细胞的成熟,诱导和/或抑制细胞因子生物合成,和/或通过刺激B细胞产生抗体来改变体液免疫。在一实施方案中,免疫应答包括刺激抗体应答(通常包括免疫球蛋白G(IgG))。免疫活性化合物具有有效的免疫刺激活性,包括但不限于抗病毒和抗肿瘤活性,并且还可以下调免疫应答的其他方面,例如将免疫应答从Th2免疫应答转移,这可用于治疗广泛的Th2介导的疾病。

[0285] 根据本发明,术语“抗原”或“免疫原”涵盖任何会引发免疫应答的物质。特别地,“抗原”涉及与抗体或T-淋巴细胞(T-细胞)特异性反应的任何物质。根据本发明,术语“抗原”包括包含至少一个表位的任何分子。优选地,本发明的上下文中的抗原是一种分子,任选在加工之后,其诱导免疫反应,优选是抗原特异性的。根据本发明,可以使用任何合适的抗原,其是免疫反应的候选物,其中免疫反应可以是体液以及细胞免疫反应。在本发明的实

施方案的背景下,抗原优选由细胞呈递,优选由抗原呈递细胞呈递,在MHC分子的背景下,这导致针对抗原的免疫反应。抗原优选是对应于或源自天然存在的抗原的产物。这类天然存在的抗原可以包括或可以源自过敏原、病毒、细菌、真菌、寄生虫和其他传染物和病原体,或者抗原也可以是肿瘤抗原。根据本发明,抗原可以对应于天然存在的产物,例如,病毒蛋白或其部分。在优选的实施方案中,抗原是表面多肽,即天然展示在细胞、病原体、细菌、病毒、真菌、寄生虫、过敏原或肿瘤表面上的多肽。抗原可以引发针对细胞、病原体、细菌、病毒、真菌、寄生虫、过敏原或肿瘤的免疫应答。

[0286] 术语“病原体”是指能够在生物体(优选脊椎动物生物体)中引起疾病的致病性生物材料。病原体包括微生物如细菌、单细胞真核生物体(原生动物)、真菌以及病毒。

[0287] 术语“表位”、“抗原肽”、“抗原表位”、“免疫原性肽”和“MHC结合肽”在本文中可互换使用,是指分子如抗原中的抗原决定簇,即免疫系统识别的免疫活性化合物的一部分或片段,例如,T细胞识别的免疫活性化合物的一部分或片段,特别是在MHC分子的背景下呈递时。蛋白的表位优选包含所述蛋白的连续或不连续部分,并且优选长度为5-100、优选5-50、更优选8-30、最优选10-25个氨基酸,例如,表位可以优选长度为9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24或25个氨基酸。根据本发明,表位可以与MHC分子结合,如细胞表面上的MHC分子,因此可以是“MHC结合肽”或“抗原肽”。术语“主要组织相容性复合物”和缩写“MHC”包括MHC I类和MHC II类分子,并且涉及存在于所有脊椎动物中的基因复合物。MHC蛋白或分子对于免疫反应中淋巴细胞与抗原呈递细胞或病变细胞之间的信号传导很重要,其中MHC蛋白或分子结合肽并呈递它们用于T细胞受体识别。MHC编码的蛋白在细胞表面上表达,并且向T细胞展示自身抗原(来自细胞自身的肽片段)和非自身抗原(例如,入侵微生物的片段)。优选这样的免疫原性部分与MHC I类或II类分子结合。如本文所用,如果使用本领域已知的任何测定可检测到这样的结合,则称免疫原性部分“结合”MHC I类或II类分子。术语“MHC结合肽”涉及与MHC I类和/或MHC II类分子结合的肽。在I类MHC/肽复合物的情况下,结合肽通常长8-10个氨基酸,尽管更长或更短的肽可能是有效的。在II类MHC/肽复合物的情况下,结合肽通常长10-25个氨基酸,特别是长13-18个氨基酸,而更长或更短的肽可能是有效的。

[0288] 在一实施方案中,本发明的所关注的蛋白包含适用于靶生物体疫苗接种的表位。本领域技术人员会知道免疫生物学和疫苗接种的原理之一是基于以下事实:通过用抗原免疫生物体来产生对疾病的免疫保护性反应,其与待治疗的疾病在免疫学上相关。根据本发明,抗原选自自身抗原和非自身抗原。非自身抗原优选是细菌抗原、病毒抗原、真菌抗原、过敏原或寄生虫抗原。优选抗原包含能够在靶生物体中引发免疫应答的表位。例如,所述表位可以引发针对细菌、病毒、真菌、寄生虫、过敏原或肿瘤的免疫应答。

[0289] 在一些实施方案中,非自身抗原是细菌抗原。在一些实施方案中,抗原引发针对细菌的免疫应答,所述细菌感染动物,包括鸟、鱼和哺乳动物,包括驯化动物。优选地,针对其引发免疫应答的细菌是致病菌。

[0290] 在一些实施方案中,非自身抗原是病毒抗原。例如,病毒抗原可以是来自病毒表面蛋白的肽,例如衣壳多肽或刺突多肽,如来自冠状病毒。在一些实施方案中,抗原引发针对病毒的免疫应答,所述病毒感染动物,包括鸟、鱼和哺乳动物,包括驯化动物。优选地,针对其引发免疫应答的病毒是致病性病毒。

[0291] 在一些实施方案中,非自身抗原是来自真菌的多肽或蛋白。在一些实施方案中,抗原引发针对真菌的免疫应答,所述真菌感染动物,包括鸟、鱼和哺乳动物,包括驯化动物。优选地,针对其引发免疫应答的真菌是致病性真菌。

[0292] 在一些实施方案中,非自身抗原是来自单细胞真核寄生虫的多肽或蛋白。在一些实施方案中,抗原引发针对单细胞真核寄生虫的免疫应答,优选致病性单细胞真核寄生虫。致病性单细胞真核寄生虫可以例如来自疟原虫属(Plasmodium),例如恶性疟原虫(P.falciparum)、间日疟原虫(P.vivax)、三日疟原虫(P.malariae)或卵形疟原虫(P.ovale),来自利什曼原虫属,或者来自锥虫属(Trypanosoma),例如克氏锥虫(T.cruzi)或布氏锥虫(T.brucei)。

[0293] 在一些实施方案中,非自身抗原是致敏多肽或致敏蛋白。致敏蛋白或致敏多肽适用于过敏原免疫疗法,也称作减敏(hypo-sensitization)。

[0294] 在一些实施方案中,抗原是自身抗原,特别是肿瘤抗原。肿瘤抗原及其测定是技术人员已知的。

[0295] 在本发明的上下文中,术语“肿瘤抗原”或“肿瘤相关抗原”涉及在正常条件下在有限数量的组织和/或器官中或者在特定发育阶段特异性地表达的蛋白,例如,肿瘤抗原可以在正常条件下在胃组织中,优选在胃粘膜中,在生殖器官中,例如,在睾丸中,在滋养层组织中,例如,在胎盘中,或者在生殖系细胞中特异性地表达,并且在一种或多种肿瘤或癌组织中表达或异常表达。在这个上下文中,“有限数量”优选表示不超过3个,更优选不超过2个。本发明的上下文中的肿瘤抗原包括,例如,分化抗原,优选细胞类型特异性分化抗原,即在正常条件下在某个分化阶段在某种细胞类型中特异性地表达的蛋白,癌症/睾丸抗原,即在正常条件下在睾丸中并且有时在胎盘中特异性地表达的蛋白,以及生殖系特异性抗原。在本发明的上下文中,肿瘤抗原优选与癌细胞的细胞表面相关,并且优选在正常组织中不表达或仅很少表达。优选地,肿瘤抗原或肿瘤抗原的异常表达鉴定癌细胞。在本发明的上下文中,受试者(例如,患有癌症疾病的患者)中的癌细胞表达的肿瘤抗原,优选是所述受试者的自身蛋白。在优选的实施方案中,本发明的上下文中的肿瘤抗原在正常条件下在非必需的组织或器官(即,当被免疫系统损伤时不导致受试者死亡的组织或器官)中特异性地表达,或者在免疫系统不可接近或仅很难接近的身体器官或结构中特异性地表达。优选地,肿瘤抗原的氨基酸序列在正常组织中表达的肿瘤抗原和癌组织中表达的肿瘤抗原之间相同。

[0296] 可以用于本发明的肿瘤抗原的实例是p53,ART-4,BAGE,β-连环蛋白/m,Bcr-abL CAMEL,CAP-1,CASP-8,CDC27/m,CDK4/m,CEA,紧密连接蛋白家族的细胞表面蛋白,如CLAUDIN-6、CLAUDIN-18.2和CLAUDIN-12,c-MYC,CT,Cyp-B,DAM,ELF2M,ETV6-AML1,G250,GAGE,GnT-V,Gap100,HAGE,HER-2/neu,HPV-E7,HPV-E6,HAST-2,hTERT(或hTRT),LAGE,LDLR/FUT,MAGE-A,优选MAGE-A1、MAGE-A2、MAGE-A3、MAGE-A4、MAGE-A5、MAGE-A6、MAGE-A7、MAGE-A8、MAGE-A9、MAGE-A10、MAGE-A11或MAGE-A12,MAGE-B,MAGE-C,MART-1/Melan-A,MC1R,肌球蛋白/m,MUC1,MUM-1、-2、-3,NA88-A,NF1,NY-ESO-1,NY-BR-1,p190最小BCR-abL(p190 minor BCR-abL),Pm1/RARa,PRAME,蛋白酶3,PSA,PSM,RAGE,RU1或RU2,SAGE,SART-1或SART-3,SCGB3A2,SCP1,SCP2,SCP3,SSX,SURVIVIN,TEL/AML1,TPI/m,TRP-1,TRP-2,TRP-2/INT2,TPTE以及WT。特别优选的肿瘤抗原包括CLAUDIN-18.2(CLDN18.2)和CLAUDIN-6(CLDN6)。

[0297] 在一些实施方案中,不要求药学活性肽或蛋白是引发免疫应答的抗原。合适的药学活性蛋白或肽可以选自细胞因子和免疫系统蛋白如免疫活性化合物(例如,白细胞介素、集落刺激因子(CSF)、粒细胞集落刺激因子(G-CSF)、粒细胞-巨噬细胞集落刺激因子(GM-CSF)、促红细胞生成素、肿瘤坏死因子(TNF)、干扰素、整合素、地址素、选择素(seletin)、归巢受体、T细胞受体、嵌合抗原受体(CAR)、免疫球蛋白)、激素(胰岛素、甲状腺激素、儿茶酚胺、促性腺激素、营养激素、催乳素、催产素、多巴胺、牛促生长素、瘦素等)、生长激素(例如,人生长激素)、生长因子(例如,表皮生长因子、神经生长因子、胰岛素样生长因子等)、生长因子受体、酶(组织纤溶酶原激活剂、链激酶、胆固醇生物合成或降解、类固醇生成酶、激酶、磷酸二酯酶、甲基化酶、去甲基化酶、脱氢酶、纤维素酶、蛋白酶、脂肪酶、磷脂酶、芳香化酶、细胞色素、腺苷酸或鸟苷酸环化酶、神经氨酸酶等)、受体(类固醇激素受体、肽受体)、结合蛋白(生长激素或生长因子结合蛋白等)、转录和翻译因子、肿瘤生长抑制蛋白(例如,抑制血管生成的蛋白)、结构蛋白(如胶原蛋白、丝素蛋白、纤维蛋白原、弹性蛋白、微管蛋白、肌动蛋白和肌球蛋白)、血液蛋白(凝血酶、血清白蛋白、因子VII、因子VIII、胰岛素、因子IX、因子X、组织纤溶酶原激活剂、蛋白C、血管性血友病因子、抗凝血酶III、葡糖脑苷脂酶、促红细胞生成素粒细胞集落刺激因子(GCSF)或修饰的因子VIII、抗凝剂等。在一实施方案中,本发明的药学活性蛋白是参与调节淋巴稳态的细胞因子,优选是参与并优选诱导或增强T细胞发育、引发、扩增、分化和/或存活的细胞因子。在一实施方案中,细胞因子是白细胞介素,例如IL-2、IL-7、IL-12、IL-15或IL-21。

[0298] 干扰素(IFN)信号传导的抑制

[0299] 由开放阅读框编码的另一种合适的所关注的蛋白是干扰素(IFN)信号传导的抑制物。虽然已报道引入RNA进行表达的细胞的生存力会降低,特别是如果用RNA多次转染细胞,但是发现IFN抑制剂增强表达RNA的细胞的生存力(WO 2014/071963 A1)。优选地,所述抑制物是IFN I型信号传导的抑制物。防止细胞外IFN与IFN受体的结合并抑制细胞中的胞内IFN信号传导可以使RNA在细胞中稳定表达。可选地;或者额外地,防止细胞外IFN与IFN受体的结合并抑制抑制细胞内IFN信号传导可增强细胞的存活,特别是如果用RNA重复转染细胞。不希望受理论束缚,设想细胞内IFN信号传导可以导致翻译和/或RNA降解的抑制。这可以通过抑制一种或多种IFN诱导的抗病毒活性效应蛋白来解决。IFN诱导的抗病毒活性效应蛋白可以选自RNA依赖性蛋白激酶(PKR)、2',5'-寡腺苷酸合成酶(OAS)和RNaseL。抑制细胞内IFN信号传导可以包括抑制PKR依赖性途径和/或OAS依赖性途径。合适的所关注的蛋白是能够抑制PKR依赖性途径和/或OAS依赖性途径的蛋白。抑制PKR依赖性途径可以包括抑制eIF2-α磷酸化。抑制PKR可以包括用至少一种PKR抑制物处理细胞。PKR抑制物可以是PKR的病毒抑制物。PKR的优选病毒抑制物是痘苗病毒E3。如果肽或蛋白(例如E3、K3)是为了抑制细胞内IFN信号传导,则优选肽或蛋白的细胞内表达。痘苗病毒E3是一种25kDa dsRNA结合蛋白(由基因E3L编码),其结合并隔离(sequester)dsRNA以防止PKR和OAS的激活。E3可以直接与PKR结合并抑制其活性,导致eIF2-α的磷酸化减少。其他合适的IFN信号传导抑制物是单纯庖疹病毒ICP34.5、流感病毒NS1、托斯卡纳病毒NSs、家蚕核型多角体病毒(Bombyx mori nucleopolyhedrovirus)PK2和HCV NS34A。

[0300] rRNA分子中至少一个开放阅读框的位置

[0301] rRNA复制子适用于表达一个或多个编码所关注的肽或所关注的蛋白的基因,任选

在亚基因组启动子的控制下。各种实施方案是可能的。RNA复制子上可以存在一个或多个开放阅读框,每个编码所关注的肽或所关注的蛋白。RNA复制子的最上游开放阅读框称作“第一开放阅读框”。在一实施方案中,编码所关注的蛋白的第一开放阅读框位于5’复制识别序列的下游和IRES的上游(以及编码来自自复制病毒的功能性非结构蛋白的开放阅读框)。在一些实施方案中,“第一开放阅读框”是RNA复制子的唯一开放阅读框。任选地,第一开放阅读框的下游可以存在一个或多个其他开放阅读框。第一开放阅读框下游的一个或多个其他开放阅读框可以称作“第二开放阅读框”、“第三开放阅读框”等,按照它们在第一开放阅读框下游出现的顺序(5’至3’)。在一实施方案中,编码一种或多种所关注的蛋白的一个或多个其他开放阅读框位于编码来自自复制病毒的功能性非结构蛋白的开放阅读框下游,并且优选由亚基因组启动子控制。优选地,每个开放阅读框包含起始密码子(碱基三联体),通常是AUG(在RNA分子中),对应于ATG(在相应的DNA分子中)。

[0302] 如果复制子包含3’复制识别序列,则优选所有开放阅读框位于3’复制识别序列的上游。

[0303] 在一些实施方案中,复制子的至少一个开放阅读框受亚基因组启动子的控制,优选甲病毒亚基因组启动子。甲病毒亚基因组启动子非常高效,因此适用于高水平的异源基因表达。优选地,亚基因组启动子是甲病毒中亚基因组转录物的启动子。这意味着亚基因组启动子是甲病毒天然的启动子,并且优选控制所述甲病毒中编码一种或多种结构蛋白的开放阅读框的转录。或者,亚基因组启动子是甲病毒亚基因组启动子的变体;任何在宿主细胞中作为亚基因组RNA转录启动子的变体都是合适的。如果复制子包含亚基因组启动子,则优选复制子包含保守序列元件3(CSE 3)或其变体。

[0304] 优选地,在亚基因组启动子控制下的至少一个开放阅读框位于亚基因组启动子的下游。优选地,亚基因组启动子控制包含开放阅读框转录物的亚基因组RNA的产生。

[0305] 在一些实施方案中,第一开放阅读框受亚基因组启动子的控制。在一实施方案中,当第一开放阅读框受亚基因组启动子的控制时,第一开放阅读框编码的基因可以从复制子及其亚基因组转录物(后者在功能性甲病毒非结构蛋白的存在下)表达。每个在亚基因组启动子控制下的一个或多个其他开放阅读框可以存在于可以在亚基因组启动子控制下的第一开放阅读框的下游。由一个或多个其他开放阅读框(例如由第二开放阅读框)编码的基因可以从一个或多个亚基因组转录物翻译,每个都受亚基因组启动子的控制。例如,RNA复制子可以包含亚基因组启动子,其控制编码第二所关注的蛋白的转录物的产生。

[0306] 在其他实施方案中,第一开放阅读框不受亚基因组启动子的控制。在一实施方案中,当第一开放阅读框不受亚基因组启动子的控制时,第一开放阅读框编码的基因可以从复制子表达。每个在亚基因组启动子控制下的一个或多个其他开放阅读框可以存在于第一开放阅读框的下游。由一个或多个其他开放阅读框编码的基因可以从亚基因组转录物表达。

[0307] 在包含本发明的复制子的细胞中,复制子可以通过功能性非结构蛋白扩增。此外,如果复制子包含一个或多个在亚基因组启动子控制下的开放阅读框,则预期一个或多个亚基因组转录物将由功能性非结构蛋白制备。

[0308] 如果复制子包含一个以上编码所关注的蛋白的开放阅读框,则优选每个开放阅读框编码不同的蛋白。例如,第二开放阅读框编码的蛋白与第一开放阅读框编码的蛋白不同。

本发明的可复制RNA分子的其他特征

[0309] 本发明的RNA分子可以任选地具有进一步的特征,例如5'-帽、5'-UTR、3'-UTR、poly(A)序列和/或密码子使用的适应性以优化RNA分子的翻译和/或稳定,如下所述。

[0310] 帽

[0311] 在一些实施方案中,本发明的复制子包含5'-帽。

[0312] 术语"5'-帽"、"帽"、"5'-帽结构"、"帽结构"同义地用来指在一些真核生物初级转录物如前体信使RNA的5'末端发现的二核苷酸。5'-帽是一种结构,其中(任选修饰的)鸟苷通过5'至5'三磷酸键(或在某些帽类似物的情况下修饰的三磷酸键)与mRNA分子的第一个核苷酸结合。该术语可以指常规的帽或帽类似物。

[0313] "包含5'-帽的RNA"或"具有5'-帽的RNA"或"用5'-帽修饰的RNA"或"加帽的RNA"是指包含5'-帽的RNA。例如,提供具有5'-帽的RNA可以通过在所述5'-帽的存在下体外转录DNA模板实现,其中所述5'-帽共转录地掺入产生的RNA链,或者可以例如通过体外转录产生RNA,并且可以利用加帽酶,例如痘苗病毒的加帽酶,将5'-帽转录后连接至RNA。在加帽的RNA中,(加帽的)RNA分子的第一个碱基的3'位置通过磷酸二酯键连接至RNA分子的后续碱基("第二个碱基")的5'位置。

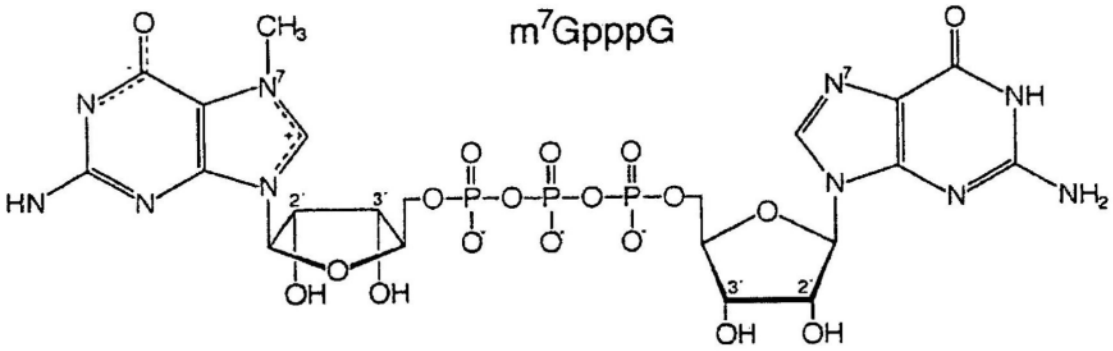
[0314] 在一实施方案中,RNA复制子包含5'-帽。在一实施方案中,RNA复制子不包含5'-帽。

[0315] 术语"常规5'-帽"是指天然存在的5'-帽,优选7-甲基鸟苷帽。在7-甲基鸟苷帽中,帽的鸟苷是修饰的鸟苷,其中修饰由7-位置的甲基化组成。

[0316] 在本发明的上下文中,术语"5'-帽类似物"是指一种分子结构,其类似于常规5'-帽,但经过修饰以在连接时具有稳定RNA的能力,优选在体内和/或细胞中。帽类似物不是常规5'-帽。

[0317] 对于真核mRNA的情况,一般描述5'-帽参与mRNA的高效翻译:一般来说,在真核生物中,翻译仅在信使RNA(mRNA)分子的5'末端开始,除非存在内部核糖体进入位点(IRES)。真核细胞能够在细胞核转录过程中提供具有5'-帽的RNA:新合成的mRNA通常用5'-帽结构修饰,例如当转录物的长度达到20-30个核苷酸时。首先,5'末端核苷酸pppN(ppp代表三磷酸;N代表任何核苷)在细胞中通过具有RNA 5'-三磷酸酶和鸟苷酸转移酶活性的加帽酶转化为5'GpppN。随后GpppN可以通过具有(鸟嘌呤-7)-甲基转移酶活性的第二种酶在细胞中甲基化,形成单甲基化的m^7GpppN帽。在一实施方案中,本发明中使用的5'-帽是天然的5'-帽。

[0318] 在本发明中,天然的5'-帽二核苷酸通常选自非甲基化的帽二核苷酸(G(5')ppp(5')N;也称为GpppN)和甲基化的帽二核苷酸(((m^7G(5')ppp(5')N;也称为m^7GpppN)。m^7GpppN(其中N为G)由下式表示:

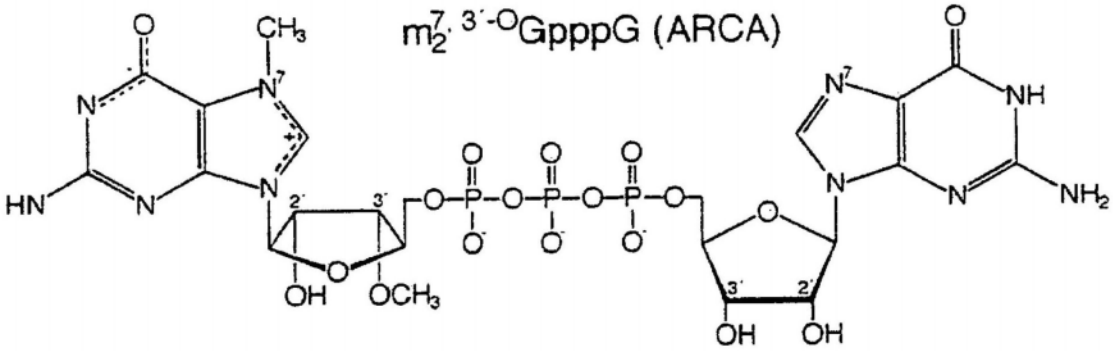
[0319] 

[0320] 本发明的加帽RNA可以在体外制备,因此不依赖于宿主细胞中的加帽机制。在体外制备加帽RNA的最常用方法是在所有四种核糖核苷三磷酸和帽二核苷酸如m^7G(5')ppp(5')G(也称为m^7GpppG)的存在下用细菌或噬菌体RNA聚合酶转录DNA模板。RNA聚合酶通过m^7GpppG鸟苷部分的3'-OH对下一个模板核苷三磷酸(pppN)的α-磷酸进行亲核攻击来启动转录,产生中间体m^7GpppGpN(其中N是RNA分子的第二个碱基)。在体外转录过程中,通过将帽与GTP的摩尔比设置在5和10之间来抑制竞争性GTP起始产物pppGpN的形成。

[0321] 在本发明的优选实施方案中,5'-帽(如果存在)是5'-帽类似物。如果RNA是通过体外转录获得的,例如是体外转录的RNA(IVT-RNA),则这些实施方案特别合适。最初已经描述了帽类似物促进通过体外转录大规模合成RNA转录物。

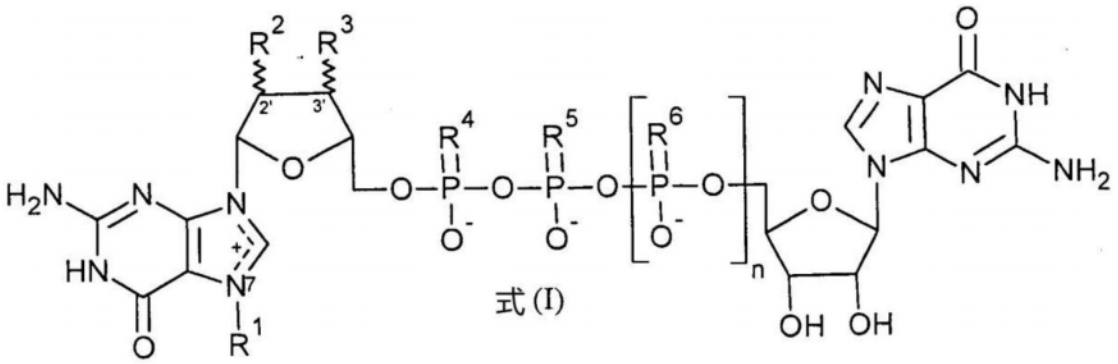
[0322] 对于信使RNA,迄今为止已普遍描述了一些帽类似物(合成帽),并且它们都可以在本发明的背景下使用。理想情况下,选择与更高的翻译效率和/或增加的体内降解抗性和/或增加的体外降解抗性相关的帽类似物。

[0323] 优选地,使用的帽类似物只能以一个方向掺入RNA链中。Pasquinelli et al.,1995,RNA J.1:957-967)证明在体外转录过程中,噬菌体RNA聚合酶使用7-甲基鸟苷单元启动转录,其中约40-50%具有帽的转录物具有相反方向的帽二核苷酸(即,初始反应产物是Gpppm^{7}GpN)。与具有正确帽的RNA相比,具有反向帽的RNA在将核酸序列翻译为蛋白方面不起作用。因此,期望以正确的方向掺入帽,即产生具有基本上对应于m^7GpppGpN等的结构的RNA。研究表明,甲基化鸟苷单元的2'-或3'-OH基团的取代抑制帽-二核苷酸的反向整合(Stepinski et al.,2001,RNA J.7:1486-1495;Peng et al.,2002,Org.Lett.24:161-164)。在这类"抗反向帽类似物"的存在下合成的RNA比在常规5'-帽m^7GpppG的存在下体外转录的RNA更高效地翻译。为此,描述了其中甲基化鸟苷单元的3'OH基团被OCH_3取代的一种帽类似物,例如,Holtkamp et al.,2006,Blood 108:4009-4017(7-甲基(3'-O-甲基)GpppG;抗反向帽类似物(ARCA))。根据本发明,ARCA是合适的帽二核苷酸。

[0324] 

[0325] 在一实施方案中,本发明的RNA基本上不易脱帽。这一点很重要,因为一般来说,由

引入培养的哺乳动物细胞的合成mRNA产生的蛋白量受到mRNA自然降解的限制。mRNA降解的一种体内途径始于去除mRNA帽。这种去除由异源二聚体焦磷酸酶催化,其含有调节亚基(Dcp1)和催化亚基(Dcp2)。催化亚基在三磷酸桥的α和β磷酸基团之间切割。在本发明中,可以选择或存在对该类型的切割不敏感或不太敏感的帽类似物。为此目的,合适的帽类似物可以选自根据式(I)的帽二核苷酸:

[0326] 

[0327] 其中R^1选自任选取代的烷基、任选取代的烯基、任选取代的炔基、任选取代的环烷基、任选取代的杂环基、任选取代的芳基和任选取代的杂芳基。

[0328] R^2和R^3独立地选自H、卤素、OH和任选取代的烷氧基,或者R^2和R^3一起形成O-X-O,其中X选自任选取代的CH_2、CH_2CH_2、$CH_2CH_2CH_2$、$CH_2CH(CH_3)$,和

[0329] $C(CH_3)_2$,或者R^2与R^2连接的环的4'位的氢原子结合形成-O-CH_2-或-CH_2-O-,

[0330] R^5选自S、Se和BH_3,

[0331] R^4和R^6独立地选自O、S、Se和BH_3。

[0332] n为1、2或3。

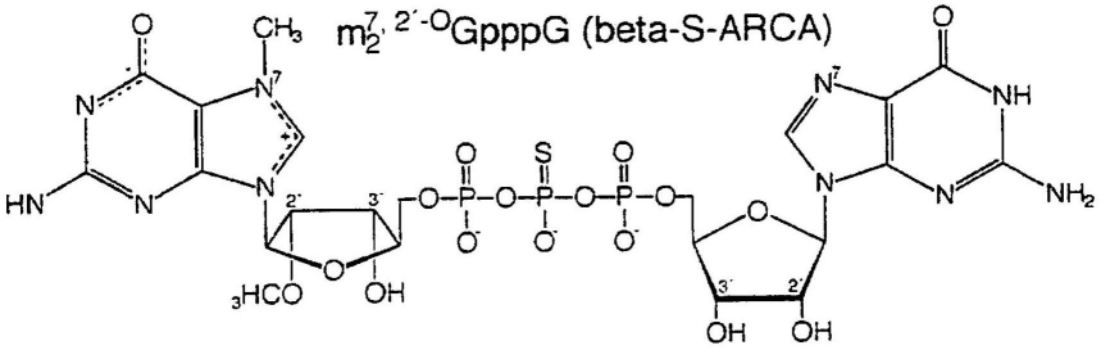
[0333] R^1、R^2、R3、R^4、R^5、R^6的优选实施方案在WO 2011/015347 A1中公开,并且可以在本发明中相应地选择。

[0334] 例如,在一实施方案中,本发明的RNA包含硫代磷酸-帽-类似物。硫代磷酸-帽-类似物是特定的帽类似物,其中三磷酸链中的三个非桥连O原子之一被S原子取代,即式(I)中的R^4、R^5或R^6之一是S。Kowalska et al.,2008,RNA,14:1119-1131已经描述了硫代磷酸-帽-类似物,作为不期望的脱帽过程的解决方案,从而增加RNA在体内的稳定性。特别地,在5'-帽的β-磷酸基团上用硫原子取代氧原子导致对Dcp2的稳定。在本发明优选的该实施方案中,式(I)中的R^5是S;并且R^4和R^6是O。

[0335] 在另一实施方案中,本发明的RNA包含硫代磷酸-帽-类似物,其中RNA 5'-帽的硫代磷酸修饰与"抗反向帽类似物"(ARCA)修饰组合。WO 2008/157688 A2中描述了相应的ARCA-硫代磷酸-帽-类似物,它们全部可以用于本发明的RNA。在该实施方案中,式(I)中R^2或R^3中至少一个不是OH,优选R^2或R^3中的一个是甲氧基(OCH_3),R^2或R^3中的另一个优选是OH。在一优选实施方案中,β-磷酸基团上的一个氧原子被硫原子取代(因此式(I)中的R^5是S;并且R^4和R^6是O)。据信ARCA的硫代磷酸修饰确保α、β和γ硫代磷酸基团在翻译和脱帽机制中精确定位在帽结合蛋白的活性位点内。这些类似物中至少有一些基本上对焦磷酸酶Dcp1/Dcp2具有抗性。据描述硫代磷酸修饰的ARCA对eIF4E的亲和力远高于缺少硫代磷酸基团的相应ARCA。

[0336] 在本发明中特别优选的相应帽类似物,即$m_{2'}^{7,2'-O}Gpp_spG$,被称为β-S-ARCA(WO

2008/157688A2;Kuhn et al.,2010,Gene Ther.17:961-971)。因此,在本发明的一个实施方案中,用β-S-ARCA修饰本发明的RNA。β-S-ARCA由以下结构表示:

[0337] 

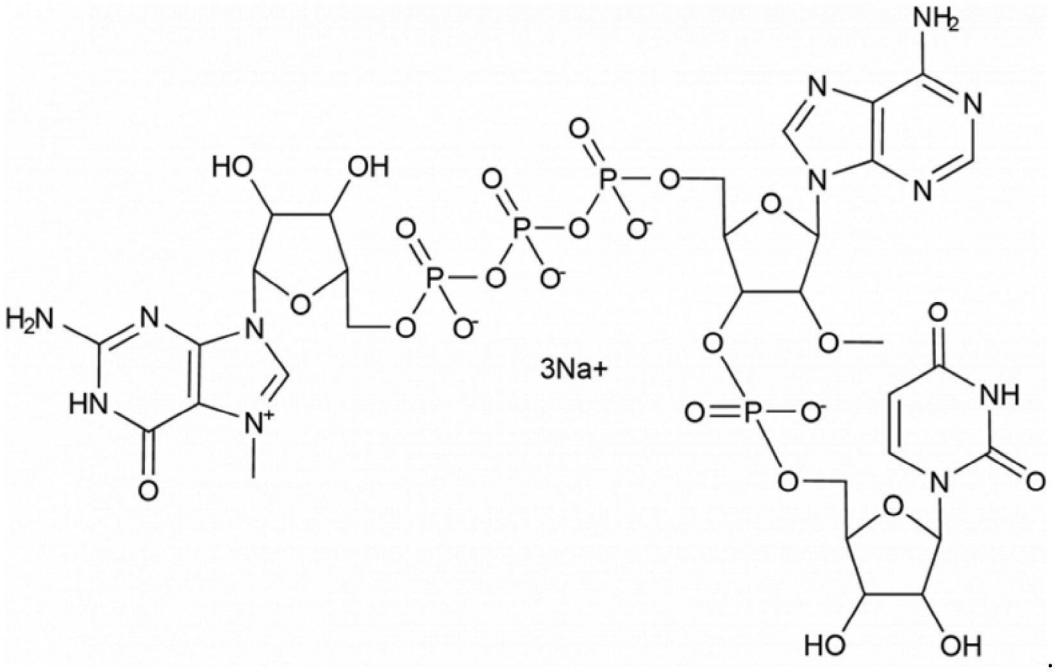
[0338] 一般来说,在桥连磷酸酯上用硫原子代替氧原子会产生硫代磷酸酯非对映体,基于它们在HPLC中的洗脱模式,将其命名为D1和D2。简而言之,“β-S-ARCA的D1非对映体”或“β-S-ARCA(D1)”是β-S-ARCA的非对映体,与β-S-ARCA的D2非对映体(β-S-ARCA(D2))相比其在HPLC柱上首先洗脱,因此表现出较短的保留时间。WO 2011/015347A1中描述了通过HPLC确定立体化学构型。

[0339] 在本发明的第一个特别优选的实施方案中,用β-S-ARCA(D2)非对映体修饰本发明的RNA。β-S-ARCA的两种非对映体对核酸酶的敏感性不同。研究表明,携带β-S-ARCA的D2非对映体的RNA几乎完全抵抗Dcp2切割(与在未修饰的ARCA 5'-帽存在下合成的RNA相比,仅6%切割),而具有β-S-ARCA(D1)5'-帽的RNA对Dcp2切割表现出中间敏感性(71%切割)。研究进一步表明,抗Dcp2切割的稳定性增加与哺乳动物细胞中蛋白表达的增加有关。特别地,研究表明携带β-S-ARCA(D2)帽的RNA在哺乳动物细胞中比携带β-S-ARCA(D1)帽的RNA更高效地翻译。因此,在本发明的一个实施方案中,本发明的RNA用根据式(I)的帽类似物修饰,其特征在于在包含式(I)中的取代基R^5的P原子处的立体化学构型,其对应于β-S-ARCA的D2非对映体的P_β原子处的立体化学构型。在该实施方案中,式(I)中的R^5是S;并且R^4和R^6是O。此外,式(I)中R^2或R^3中至少一个优选不是OH,优选R^2或R^3中的一个是甲氧基(OCH3),R^2或R^3中的另一个优选是OH。

[0340] 在第二个特别优选的实施方案中,用β-S-ARCA(D1)非对映体修饰本发明的RNA。这个实施方案特别适用于将加帽的RNA转移至未成熟的抗原呈递细胞中,如用于疫苗接种目的。已经证明,在将分别加帽的RNA转移至未成熟的抗原呈递细胞中后,β-S-ARCA(D1)非对映体特别适用于增加RNA的稳定性,提高RNA的翻译效率,延长RNA的翻译,增加RNA的总蛋白表达,和/或增加针对由所述RNA编码的抗原或抗原肽的免疫应答(Kuhn et al.,2010,Gene Ther.17:961-971)。因此,在本发明的一个可选实施方案中,本发明的RNA用根据式(I)的帽类似物修饰,其特征在于在包含式(I)中的取代基R^5的P原子处的立体化学构型,其对应于β-S-ARCA的D1非对映体的P_β原子处的立体化学构型。WO 2011/015347 A1和Kuhn et al.,2010,Gene Ther.17:961-971中描述了相应的帽类似物及其实施方案。WO 2011/015347 A1中描述的任何帽类似物,其中包含取代基R^5的P原子处的立体化学构型对应于β-S-ARCA的D1非对映体的P_β原子的立体化学构型,可以用于本发明。优选地,式(I)中R^5是S;并且R^4和R^6是O。此外,式(I)中R^2或R^3中至少一个优选不是OH,优选R^2或R^3中的一个是甲氧基(OCH3),R^2或R^3中的另一个优选是OH。

[0341] 在一实施方案中,本发明的RNA用根据式(I)的5'-帽结构修饰,其中任何一个磷酸基团被硼磷酸基团或硒代磷酸(phosphoroselenoate)基团取代。这类帽在体外和体内都具有增加的稳定性。任选地,相应化合物具有2'-O-或3'-O-烷基(其中烷基优选是甲基);相应帽类似物命名为BH_3-ARCA或Se-ARCA。特别适用于mRNA加帽的化合物包括β-BH_3-ARCA和β-Se-ARCA,如WO 2009/149253 A2所述。对于这些化合物,优选式(I)中包含取代基R^5的P原子处的立体化学构型,其对应于β-S-ARCA的D1异构体的P_β原子处的立体化学构型。

[0342] 在一实施方案中,5'帽可以是具有以下结构的三核苷酸AU(cap1):

[0343] 
NH2
N
N
N
N
HO
OH
O
P
O
O
O

优选存在3'-UTR,其有活性以提高RNA复制子的翻译效率和/或稳定性。

[0352] 关于第一核苷酸序列(例如UTR),术语"有活性以提高翻译效率"和/或"有活性以提高稳定性"表示第一核酸序列能够在与第二核酸序列的共同转录物中修饰所述第二核酸序列的翻译效率和/或稳定性,从而使所述翻译效率和/或稳定性与没有所述第一核酸序列的所述第二核酸序列的翻译效率和/或稳定性相比提高。

[0353] 在一实施方案中,本发明的RNA复制子包含5'-UTR和/或3'-UTR,其与衍生功能性甲病毒非结构蛋白的甲病毒是异源的或非天然的。这允许根据期望的翻译效率和RNA稳定性设计非翻译区。因此,异源或非天然UTR允许高度的灵活性,与天然甲病毒UTR相比,这种灵活性是有利的。

[0354] 优选地,本发明的RNA复制子包含非病毒来源的5'-UTR和/或3'-UTR;特别是非甲病毒来源的。在一实施方案中,RNA复制子包含源自真核5'-UTR的5'-UTR和/或源自真核3'-UTR的3'-UTR。

[0355] 本发明的5'-UTR可以包含一种以上核酸序列的任何组合,任选地通过接头隔开。本发明的3'-UTR可以包含一种以上核酸序列的任何组合,任选地通过接头隔开。

[0356] 本发明的术语"接头"涉及添加在两个核酸序列之间以连接所述两个核酸序列的核酸序列。关于接头序列没有特别的限制。

[0357] 3'-UTR通常具有200-2000个核苷酸的长度,例如500-1500个核苷酸。免疫球蛋白mRNA的3'-非翻译区相对较短(少于约300个核苷酸),而其他基因的3'-非翻译区相对较长。例如,tPA的3'-非翻译区的长度为约800个核苷酸,因子VIII的3'-非翻译区的长度为约1800个核苷酸,促红细胞生成素的3'-非翻译区的长度为约560个核苷酸。哺乳动物mRNA的3'-非翻译区通常具有称作AAUAAA六核苷酸序列的同源区。这个序列可能是poly(A)连接信号,通常位于poly(A)连接位点上游10-30个碱基处。3'-非翻译区可以含有一个或多个反向重复,这些反向重复可以折叠产生茎-环结构,其充当核糖核酸外切酶的屏障或与已知增加RNA稳定性的蛋白(例如RNA-结合蛋白)相互作用。

[0358] 人β-珠蛋白3'-UTR,特别是人β-珠蛋白3'-UTR的两个连续相同拷贝,有助于高转录物稳定性和翻译效率(Holtkamp et al.,2006,Blood 108:4009-4017)。因此,在一实施方案中,本发明的RNA复制子包含人β-珠蛋白3'-UTR的两个连续相同拷贝。因此,它在5'→3'方向上包含:(a)任选存在的5'-UTR;(b)开放阅读框;(c)3'-UTR;所述3'-UTR包含人β-珠蛋白3'-UTR的两个连续相同拷贝,其片段,或者人β-珠蛋白3'-UTR的变体或其片段。

[0359] 在一实施方案中,本发明的RNA复制子包含3'-UTR,其有活性以提高翻译效率和/或稳定性,但其不是人β-珠蛋白3'-UTR,其片段,或者人β-珠蛋白3'-UTR的变体或其片段。

[0360] 在一实施方案中,本发明的RNA复制子包含5'-UTR,其有活性以提高翻译效率和/或稳定性。

[0361] Poly(A)序列

[0362] 在一些实施方案中,本发明的复制子包含3'-poly(A)序列。如果复制子包含保守序列元件4(CSE 4),则复制子的3'-poly(A)序列优选存在于CSE 4的下游,最优选直接与CSE 4相邻。

[0363] 根据本发明,在一实施方案中,poly(A)序列包含至少20、优选至少26、优选至少40、优选至少80、优选至少100以及优选多达500、优选多达400、优选多达300、优选多达200、

特别是多达150个A核苷酸,特别是约120个A核苷酸或者基本上由至少20、优选至少26、优选至少40、优选至少80、优选至少100以及优选多达500、优选多达400、优选多达300、优选多达200、特别是多达150个A核苷酸,特别是约120个A核苷酸组成或者由至少20、优选至少26、优选至少40、优选至少80、优选至少100以及优选多达500、优选多达400、优选多达300、优选多达200、特别是多达150个A核苷酸,特别是约120个A核苷酸。在这种情况下,"基本上由…组成"表示poly(A)序列中的大多数核苷酸,通常"poly(A)序列"中核苷酸数量的至少50%,优选至少75%是A核苷酸(腺苷酸),但是允许剩余的核苷酸是除A核苷酸以外的核苷酸,如U核苷酸(尿苷酸)、G核苷酸(鸟苷酸)或C核苷酸(胞苷酸)。在这种情况下,"由…组成"表示poly(A)序列中的所有核苷酸,即poly(A)序列中核苷酸数量的100%是A核苷酸。术语"A核苷酸"或"A"是指腺苷酸。

[0364] 事实上,已证实约120个A核苷酸的3'poly(A)序列对转染的真核细胞中的RNA水平以及从存在于3'poly(A)序列上游(5')的开放阅读框翻译的蛋白水平具有有益影响(Holtkamp et al.,2006,Blood,vol.108,pp.4009-4017)。

[0365] 在甲病毒中,至少11个连续腺苷酸残基或至少25个连续腺苷酸残基的3'poly(A)序列被认为对负链的高效合成很重要。特别地,在甲病毒中,至少25个连续腺苷酸残基的3'poly(A)序列被认为与保守序列元件4(CSE 4)一起发挥作用,促进(-)链的合成(Hardy&Rice,2005,J.Virol.79:4630-4639)。

[0366] 本发明提供在RNA转录期间,即在制备体外转录的RNA期间,基于在与编码链互补的链中包含重复dT核苷酸(脱氧胸苷酸)的DNA模板连接的3'poly(A)序列。编码poly(A)序列的DNA序列(编码链)称作poly(A)盒。

[0367] 在本发明的一个优选实施方案中,DNA编码链中存在的3'poly(A)盒基本上由dA核苷酸组成,但是被具有均等分布的四种核苷酸(dA、dC、dG和dT)的随机序列中断。这样的随机序列的长度可以为5-50,优选10-30,更优选10-20个核苷酸。WO 2016/005004A1中公开了这样的盒。WO 2016/005004 A1中公开的任何poly(A)盒均可以用于本发明。基本上由dA核苷酸组成,但是被具有均等分布的四种核苷酸(dA、dC、dG、dT)且具有例如5-50个核苷酸长度的随机序列中断的poly(A)盒,在DNA水平上,在大肠杆菌(E.coli)中表现出质粒DNA的恒定增殖,并且在RNA水平上仍与支持RNA稳定性和翻译效率方面的有益特性相关。

[0368] 因此,在本发明的一个优选实施方案中,本文所述RNA分子中含有的3'poly(A)序列基本上由A核苷酸组成,但是被具有均等分布的四种核苷酸(A、C、G、U)的随机序列中断。这样的随机序列的长度可以为5-50,优选10-30,更优选10-20个核苷酸。

[0369] 密码子使用

[0370] 一般来说,遗传密码的简并性会允许RNA序列中存在的某些密码子(编码氨基酸的碱基三联体)被其他密码子(碱基三联体)替代,同时保持相同的编码能力(这样替换密码子编码与被替换的密码子相同的氨基酸)。在本发明的一些实施方案中,RNA(rRNA)分子包含的开放阅读框的至少一个密码子不同于开放阅读框来源物种中相应开放阅读框中的相应密码子。在该实施方案中,开放阅读框的编码序列被称为"调整的(adpated)"或"修饰的(modified)"。复制子包含的开放阅读框的编码序列可以被调整。

[0371] 例如,当开放阅读框的编码序列被调整时,可以选择常用密码子:WO 2009/024567A1描述了核酸分子编码序列的调整,包括用更常用的密码子替代稀有密码子。由于

密码子使用的频率取决于宿主细胞或宿主生物体,因此这种类型的调整适合使核酸序列合适在特定宿主细胞或宿主生物体中表达。一般来说,更常用的密码子通常在宿主细胞或宿主生物体中更高效地翻译,尽管并不总是需要调整开放阅读框的所有密码子。

[0372] 例如,当开放阅读框的编码序列被调整时,可以通过选择每个氨基酸具有最高富含GC含量的密码子来改变G(鸟苷酸)残基或C(胞苷酸)残基的含量。据报道具有富含GC的开放阅读框的RNA分子具有降低免疫激活以及改善RNA的翻译和半衰期的潜力(Thess et al.,2015,Mol.Ther.23:1457-1465)。

[0373] 特别地,可以根据需要调整非结构蛋白的编码序列。这种自由是可能的,因为编码非结构蛋白的开放阅读框不与复制子的5'复制识别序列重叠。

[0374] 本发明的实施方案的安全特征

[0375] 在本发明中优选以下特征,单独或以任何合适的组合:

[0376] 本发明的复制子不会形成颗粒。这意味着,用本发明的复制子接种宿主细胞后,宿主细胞不会产生病毒颗粒,如下一代病毒颗粒。在一实施方案中,本发明的RNA复制子完全不含编码任何病毒结构蛋白的遗传信息,例如,甲病毒结构蛋白,如核心核衣壳蛋白C、包膜蛋白P62和/或包膜蛋白E1。优选地,本发明的复制子不包含病毒包装信号,例如,甲病毒包装信号。例如,SFV的nsP2的编码区中包含的甲病毒包装信号(White et al.,1998,J.Virol.72:4320-4326)可以例如通过缺失或突变来去除。去除甲病毒包装信号的合适方法包括nsP2编码区的密码子使用的调整。遗传密码的简并性可以允许缺失包装信号的功能而不影响编码的nsP2的氨基酸序列。

[0377] DNA

[0378] 本发明还提供一种DNA,其包含编码本发明的RNA复制子的核酸序列。

[0379] 优选地,所述DNA是双链的。

[0380] 在一优选实施方案中,所述DNA是质粒。如本文所用,术语"质粒"一般涉及染色体外遗传物质的构建体,通常是环状DNA双链体,其可以独立于染色体DNA进行复制。

[0381] 本发明的DNA可以包含启动子,其可以被DNA依赖性RNA聚合酶识别。这允许在体内或体外转录编码的RNA,例如本发明的RNA。IVT载体可以以标准化的方式用作体外转录的模板。本发明优选的启动子的实例是SP6、T3或T7聚合酶的启动子。

[0382] 在一实施方案中,本发明的DNA是分离的核酸分子。

[0383] 制备RNA的方法

[0384] 本发明的RNA分子可通过体外转录获得。本发明特别关注体外转录的RNA(IVT-RNA)。IVT-RNA可通过从核酸分子(特别是DNA分子)转录获得。本发明的DNA分子适用于这类目的,特别是如果包含可以被DNA依赖性RNA聚合酶识别的启动子。

[0385] 本发明的rRNA可以在体外合成。这允许将帽类似物添加到体外转录反应中。通常,poly(A)尾由DNA模板上的poly-(dT)序列编码。或者,可以在转录后酶促实现加帽和poly(A)尾添加。

[0386] 体外转录方法是技术人员已知的。例如,如WO 2011/015347 A1中提到的,各种体外转录试剂盒是可商购的。

[0387] 试剂盒

[0388] 本发明还提供一种试剂盒,其包含本发明的RNA复制子。

[0389] 在一实施方案中,试剂盒的组成部分作为单独的实体存在。例如,试剂盒的一个组成部分可以存在于一个实体中,试剂盒的另一个组成部分可以存在于不同的实体中。例如,开放或封闭的容器是合适的实体。优选封闭容器。使用的容器应当优选是不含RNAse的或基本上不含RNAse的。

[0390] 在一实施方案中,本发明的试剂盒包含用于接种细胞的RNA和/或向人或动物受试者给药的RNA。

[0391] 本发明的试剂盒任选包含标签或其他形式的信息元素,例如电子数据载体。标签或信息元素优选包含说明,例如印刷的说明或任选可打印的电子形式的说明。说明可以提及试剂盒的至少一种合适的可能用途。

[0392] 药物组合物

[0393] 本文所述的RNA复制子可以以药物组合物的形式存在。本发明的药物组合物可以包含至少一种本发明的核酸分子。本发明的药物组合物包含药学可接受的稀释剂和/或药学可接受的赋形剂和/或药学可接受的载剂和/或药学可接受的媒介物。药学可接受的载剂、媒介物、赋形剂或稀释剂的选择不受特别限制。可以使用本领域已知的任何合适的药学可接受的载剂、媒介物、赋形剂或稀释剂。

[0394] 在本发明的一个实施方案中,药物组合物可以进一步包含溶剂如水性溶剂或可以保持rRNA完整性的任何溶剂。在一优选实施方案中,药物组合物是包含RNA的水性溶液。所述水性溶液可以任选包含溶质,例如盐。

[0395] 在本发明的一实施方案中,药物组合物为冷冻干燥组合物的形式。冷冻干燥组合物可通过冷冻干燥相应的水性组合物获得。

[0396] 在一实施方案中,药物组合物包含至少一种阳离子实体。一般来说,阳离子脂质、阳离子聚合物和其他带正电荷的物质可以与带负电荷的核酸形成复合物。可以通过与阳离子化合物(优选聚阳离子化合物,例如阳离子或聚阳离子肽或蛋白)复合(complexation)来稳定本发明的RNA。在一实施方案中,本发明的药物组合物包含至少一种阳离子分子,其选自鱼精蛋白、聚乙烯亚胺、聚L-赖氨酸、聚L-精氨酸、组蛋白或阳离子脂质。

[0397] 根据本发明,阳离子脂质是阳离子两亲分子,例如包含至少一个亲水和亲油部分的分子。阳离子脂质可以是单阳离子的或多阳离子的。阳离子脂质通常具有亲脂部分,如固醇,酰基或二酰基链,并且具有总的净正电荷。脂质的头部基团通常携带正电荷。阳离子脂质优选具有1-10价的正电荷,更优选1-3价的正电荷,并且更优选1价的正电荷。阳离子脂质的实例包括但不限于1,2-二-O-十八烯基-3-三甲基铵丙烷(DOTMA);二甲基双十八烷基铵(DDAB);1,2-二油酰基-3-三甲基铵-丙烷(DOTAP);1,2-二油酰基-3-二甲基铵丙烷(DODAP);1,2-二酰氧基-3-二甲基铵丙烷;1,2-二烷氧基-3-二甲基铵丙烷;双十八烷基二甲基氯化铵(DODAC)、1,2-二豆蔻酰基丙基-1,3-二甲基羟基乙基铵(1,2-dimyristoyloxypropyl-1,3-dimethylhydroxyethyl ammonium)(DMRIE)以及2,3-二油酰基氧基-N-[2(精胺羧酰胺)乙基]-N,N-二甲基-1-正丙基胺三氟乙酸(2,3-dioleoyloxy-N-[2(spermine carboxamide)ethyl]-N,N-dimethyl-1-propanamium trifluoroacetate)(DOSPA)。阳离子脂质还包括具有叔胺基团的脂质,包括1,2-二亚油酰氧基-N,N-二甲基-3-氨基丙烷(DLinDMA)。阳离子脂质适用于如本文所述在脂质制剂中配制RNA,如脂质体、乳剂和脂质复合物。通常,正电荷由至少一种阳离子脂质贡献,负电荷由RNA贡献。在一实施方案

中,除阳离子脂质外,药物组合物包含至少一种辅助脂质。辅助脂质可以是中性或阴离子脂质。辅助脂质可以是天然脂质,如磷脂,或天然脂质的类似物,或完全合成的脂质,或脂质样分子,与天然脂质没有相似性。在药物组合物包括阳离子脂质和辅助脂质的情况下,阳离子脂质比中性脂质的摩尔比可以考虑制剂的稳定性等适当确定。

[0398] 在一实施方案中,本发明的药物组合物包含鱼精蛋白。根据本发明,鱼精蛋白可用作阳离子载剂物质。术语"鱼精蛋白"是指富含精氨酸并发现在动物如鱼的精细胞中代替体细胞组蛋白特别与DNA结合的任何各种相对低分子量的强碱性蛋白。特别地,术语"鱼精蛋白"是指在鱼精子中发现的蛋白,其是强碱性的,在水中可溶,加热不凝固,并且包含多个精氨酸单体。根据本发明,如本文所用的术语"鱼精蛋白"包含获得自或源自天然或生物来源的任何鱼精蛋白氨基酸序列,包括其片段以及所述氨基酸序列或其片段的多聚体形式。此外,该术语涵盖(合成的)多肽,所述多肽是人工的,特别设计用于特定目的,并且不可以分离自天然或生物来源。

[0399] 在一些实施方案中,本发明的组合物可以包含一种或多种佐剂。可以将佐剂添加至疫苗以刺激免疫系统的应答:佐剂本身通常不提供免疫。示例性佐剂包括但不限于以下:无机化合物(例如明矾、氢氧化铝、磷酸铝、磷酸钙氢氧化物);矿物油(例如石蜡油),细胞因子(例如IL-1、IL-2、IL-12);免疫刺激性多核苷酸(如RNA或DNA;例如,含CpG的寡核苷酸);皂苷(例如来自皂树、大豆、美远志(Polygala senega)的植物皂苷);油乳剂或脂质体;聚氧乙烯醚和聚氧乙烯酯制剂;聚磷腈(PCPP);胞壁酰肽;咪唑并喹啉酮(imidazoquinolone)化合物;缩氨基硫脲(thiosemicarbazone)化合物;Flt3配体(WO 2010/066418A1);或本领域技术人员已知的任何其他佐剂。给药本发明的RNA的优选佐剂是Flt3配体(WO 2010/066418 A1)。当Flt3配体与编码抗原的RNA一起给药时,可以观察到抗原特异性CD8+T细胞的强烈增加。

[0400] 本发明的药物组合物可以进行缓冲(例如,用乙酸盐缓冲液、柠檬酸盐缓冲液、琥珀酸盐缓冲液、Tris缓冲液、磷酸盐缓冲液)。

[0401] 含RNA的颗粒

[0402] 在一些实施方案中,由于无保护RNA的不稳定性,以复合或包封形式提供本发明的RNA分子是有利的。本发明提供相应的药物组合物。特别地,在一些实施方案中,本发明的药物组合物包括含核酸的颗粒,优选含RNA的颗粒。相应的药物组合物称作颗粒(particulate)制剂。在本发明的颗粒制剂中,颗粒包含本发明的核酸以及适用于递送核酸的药学可接受的载剂或药学可接受的媒介物。例如,含核酸的颗粒可以是蛋白颗粒的形式或含脂质颗粒的形式。合适的蛋白或脂质称作颗粒形成剂。先前已描述蛋白颗粒和含脂质颗粒适合以颗粒形式递送甲病毒RNA(例如Strauss&Strauss,1994,Microbiol.Rev.58:491-562)。特别地,甲病毒结构蛋白(例如由辅助病毒提供)是以蛋白颗粒形式递送RNA的合适载剂。

[0403] 在一实施方案中,本发明的颗粒制剂是纳米颗粒制剂。在该实施方案中,本发明的组合物包含纳米颗粒形式的本发明的核酸。纳米颗粒制剂可以通过各种方案和各种复合化合物获得。脂质、聚合物、低聚物或两亲分子是纳米颗粒制剂的典型成分。

[0404] 如本文所用,术语"纳米颗粒"是指具有使得颗粒适合全身,特别是肠胃外给药,特别是核酸,的直径的任何颗粒,通常直径为1000纳米(nm)或更小。在一实施方案中,纳米颗

粒的平均直径范围为约50nm至约1000nm,优选约50nm至约400nm,优选约100nm至约300nm,如约150nm至约200nm。在一实施方案中,纳米颗粒的直径范围为约200至约700nm,约200至约600nm,优选约250至约550nm,特别是约300至约500nm或约200至约400nm。

[0405] 在一实施方案中,通过动态光散射测量,本文所述纳米颗粒的多分散指数(PI)为0.5或更小,优选0.4或更小,或者甚至更优选0.3或更小。"多分散指数"(PI)是对颗粒混合物中单个颗粒(如脂质体)的均匀或非均匀尺寸分布的测量,指示混合物中颗粒分布的宽度。例如,可以如WO 2013/143555 A1所述确定PI。

[0406] 如本文所用,术语"纳米颗粒制剂"或相似术语是指含有至少一种纳米颗粒的任何颗粒制剂。在一些实施方案中,纳米颗粒组合物是纳米颗粒的均匀集合。在一些实施方案中,纳米颗粒组合物是含脂质的药物制剂,如脂质体制剂或乳剂。

[0407] 含脂质的药物组合物

[0408] 在一实施方案中,本发明的药物组合物包含至少一种脂质。优选地,至少一种脂质是阳离子脂质。所述含脂质的药物组合物包含本发明的核酸。在一实施方案中,本发明的药物组合物包含包封在囊泡(例如脂质体)中的RNA。在一实施方案中,本发明的药物组合物包含乳剂形式的RNA。在一实施方案中,本发明的药物组合物包含rRNA与阳离子化合物的复合物,从而形成例如所谓的脂质复合物或聚合复合物。囊泡如脂质体内RNA的包封不同于例如脂质/RNA复合物。例如当RNA例如与预先形成的脂质体混合时,可获得脂质/RNA复合物。

[0409] 在一实施方案中,本发明的药物组合物包含包封在囊泡中的rRNA。这种制剂是本发明的特定颗粒制剂。囊泡是一种脂质双层,卷成球形壳,包围着一个小空间,并将该空间与囊泡外的空间隔开。通常,囊泡内的空间是水性空间,即包含水。通常,囊泡外的空间是水性空间,即包含水。脂质双层由一种或多种脂质(囊泡形成脂质)形成。包围囊泡的膜是层状相,类似于质膜。本发明的囊泡可以是多层囊泡、单层囊泡或它们的混合物。当包封在囊泡中时,rRNA通常与任何外部介质分离。因此,它以受保护的形式存在,在功能上等同于天然甲病毒中的受保护形式。合适的囊泡是颗粒,特别是纳米颗粒,如本文所述。

[0410] 例如,RNA(rRNA)可以包封在脂质体中。在该实施方案中,药物组合物是脂质体制剂或包含脂质体制剂。包封在脂质体内通常保护RNA免受RNase消化。脂质体可能包括一些外部RNA(例如在它们的表面上),但是至少一半的RNA(理想情况下是全部)包封在脂质体的核心内。

[0411] 脂质体是微小的脂质囊泡,其通常具有一个或多个形成囊泡的脂质双层,如磷脂,并且能够包封药物,例如RNA。在本发明的上下文中可以采用不同类型的脂质体,包括但不限于多层囊泡(MLV)、小单层囊泡(SUV)、大单层囊泡(LUV)、空间稳定脂质体(SSL)、多囊泡囊泡(multivesicular vesicle)(MV)和大多囊泡囊泡(large multivesicular vesicle)(LMV)以及本领域已知的其他双层形式。脂质体的大小和层数(lamellarity)取决于制备方式。有几种其他形式的超分子组织,其中脂质可以存在于水性介质中,包含层状相、六角形和倒六角形相、立方相、胶束、由单层组成的反胶束。这些相还可以与DNA或RNA组合获得,并且与RNA和DNA的相互作用可能实质上影响相状态。这类相可以存在于本发明的纳米颗粒RNA制剂中。

[0412] 脂质体可以使用技术人员已知的标准方法形成。相应的方法包括逆向蒸发法、乙醇注射法、脱水-复水法、超声或其他合适的方法。脂质体形成后,可以改变脂质体的大小以

获得具有基本上均匀大小范围的脂质体群体。

[0413] 在本发明的一个优选实施方案中,rRNA存在于包括至少一种阳离子脂质的脂质体中。相应的脂质体可以由单一脂质或脂质混合物形成,条件是使用至少一种阳离子脂质。优选的阳离子脂质具有能够质子化的氮原子;优选地,这类阳离子脂质是具有叔胺基团的脂质。具有叔胺基团的特别合适的脂质是1,2-二亚油酰氧基-N,N-二甲基-3-氨基丙烷(DLinDMA)。在一实施方案中,本发明的RNA存在于如WO 2012/006378 A1所述的脂质体制剂中:具有包封包括RNA的水性核心的脂质双层的脂质体,其中脂质双层包含pKa在5.0-7.6范围内的脂质,其优选具有叔胺基团。具有叔胺基团的优选阳离子脂质包括DLinDMA(pKa 5.8),并且一般在WO 2012/031046 A2中进行描述。根据WO 2012/031046A2,包含相应化合物的脂质体特别适合包封RNA,因此适合RNA的脂质体递送。在一实施方案中,本发明的RNA存在于脂质体制剂中,其中脂质体包括至少一种阳离子脂质,其头基包括至少一个能够被质子化的氮原子(N),其中脂质体和RNA的N:P比在1:1和20:1之间。根据本发明,"N:P比"是指含脂质颗粒(例如脂质体)中阳离子脂质中的氮原子(N)与包含的RNA中的磷酸原子(P)的摩尔比,如WO 2013/006825A1所述。1:1和20:1之间的N:P比与脂质体的净电荷和向脊椎动物细胞递送RNA的效率有关。

[0414] 在一实施方案中,本发明的rRNA存在于脂质体制剂中,所述脂质体制剂包含至少一种脂质,所述脂质包括聚乙二醇(PEG)部分,其中RNA包封在聚乙二醇化脂质体内,使得PEG部分存在于脂质体的外部,如WO 2012/031043 A1和WO 2013/033563 A1所述。

[0415] 在一实施方案中,本发明的rRNA存在于脂质体制剂中,其中脂质体的直径在60-180nm的范围内,如WO 2012/030901 A1所述。

[0416] 在一实施方案中,本发明的rRNA存在于脂质体制剂中,其中含rRNA的脂质体具有接近零或负值的净电荷,如WO 2013/143555 A1中公开的。

[0417] 在其他实施方案中,本发明的rRNA以乳剂的形式存在。先前已经描述乳剂用于将核酸分子如rRNA分子递送至细胞。本文优选水包油乳剂。相应的乳剂颗粒包括油核和阳离子脂质。更优选的是阳离子水包油乳剂,其中本发明的RNA复合至乳剂颗粒。乳剂颗粒包含油核和阳离子脂质。阳离子脂质可以与带负电荷的rRNA相互作用,从而将rRNA锚定至乳剂颗粒。在水包油乳剂中,乳剂颗粒分散在水性连续相中。例如,乳剂颗粒的平均直径通常可以为约80nm-180nm。在一实施方案中,本发明的药物组合物是阳离子水包油乳剂,其中乳剂颗粒包含油核和阳离子脂质,如WO 2012/006380 A2所述。本发明的rRNA可以以包含阳离子脂质的乳剂形式存在,其中乳剂的N:P比为至少4:1,如WO 2013/006834 A1所述。本发明的rRNA可以以阳离子脂质乳剂的形式存在,如WO 2013/006837 A1所述。特别地,组合物可以包含与阳离子水包油乳剂颗粒复合的rRNA,其中油/脂质的比例为至少约8:1(摩尔:摩尔)。

[0418] 在其他实施方案中,本发明的药物组合物包含脂质复合物形式的RNA。术语"脂质复合物"或"RNA脂质复合物"是指脂质和核酸如RNA的复合物。脂质复合物可以由阳离子(带正电荷)脂质体和阴离子(带负电荷)核酸形成。阳离子脂质体还可以包括中性"辅助"脂质。在最简单的情况下,通过将核酸与脂质体按照一定的混合方案混合,自发形成脂质复合物,但是也可以应用各种其他方案。据了解,带正电荷的脂质体和带负电荷的核酸之间的静电相互作用是脂质复合物形成的驱动力(WO 2013/143555 A1)。在本发明的一个实施方案中,RNA脂质复合物的净电荷接近零或负值。已知RNA和脂质体的电中性或带负电荷的脂质复合

物在全身给药后导致脾树突细胞(DC)中的大量RNA表达,并且与已报道的带正电荷的脂质体和脂质复合物的毒性升高无关(参见WO 2013/143555A1)。因此,在本发明的一个实施方案中,本发明的药物组合物包含纳米颗粒形式的RNA,优选脂质复合物纳米颗粒,其中(i)纳米颗粒中正电荷的数量不超过纳米颗粒中负电荷的数量,和/或(ii)纳米颗粒具有中性或净负电荷,和/或(iii)纳米颗粒中正电荷与负电荷的电荷比为1.4:1或更少,和/或(iv)纳米颗粒的zeta电位为0或更少。如WO 2013/143555A1所述,zeta电位是胶体系统中电动电势(electrokinetic potential)的科学术语。在本发明中,(a)zeta电位和(b)纳米颗粒中阳离子脂质与RNA的电荷比都可以如WO 2013/143555A1中公开的那样计算。总之,如WO 2013/143555 A1中公开的,作为具有限定粒径的纳米颗粒脂质复合物制剂的药物组合物,其中颗粒的净电荷接近零或负值,是本发明的上下文中的优选药物组合物。

[0419] 在一实施方案中,本文描述的核酸如rRNA以脂质纳米颗粒(LNP)的形式给药。LNP可以包含能够形成颗粒的任何脂质,一个或多个核酸分子附着在所述脂质上,或者一个或多个核酸分子包封在所述脂质中。

[0420] 在一实施方案中,LNP包含一种或多种阳离子脂质,以及一种或多种稳定脂质。稳定脂质包括中性脂质和聚乙二醇化的脂质。

[0421] 在一实施方案中,LNP包含阳离子脂质、中性脂质、类固醇、聚合物缀合的脂质;以及包封在脂质纳米颗粒内或与脂质纳米颗粒缔合的RNA。

[0422] 在一实施方案中,LNP占阳离子脂质的40-55摩尔%、40-50摩尔%、41-49摩尔%、41-48摩尔%、42-48摩尔%、43-48摩尔%、44-48摩尔%、45-48摩尔%、46-48摩尔%、47-48摩尔%或47.2-47.8摩尔%。在一实施方案中,LNP占阳离子脂质的约47.0、47.1、47.2、47.3、47.4、47.5、47.6、47.7、47.8、47.9或48.0摩尔%。

[0423] 在一实施方案中,中性脂质以5-15摩尔%、7-13摩尔%或9-11摩尔%的浓度存在。在一实施方案中,中性脂质以约9.5、10或10.5摩尔%的浓度存在。

[0424] 在一实施方案中,类固醇以30-50摩尔%、35-45摩尔%或38-43摩尔%的浓度存在。在一实施方案中,类固醇以约40、41、42、43、44、45或46摩尔%的浓度存在。

[0425] 在一实施方案中,LNP包含1-10摩尔%、1-5摩尔%或1-2.5摩尔%的聚合物缀合的脂质。

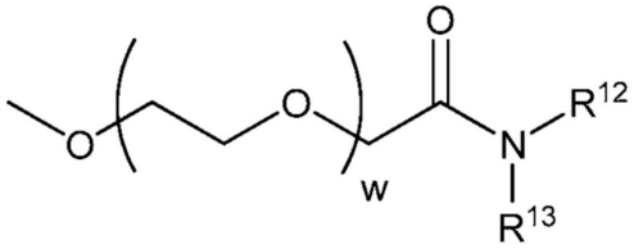
[0426] 在一实施方案中,LNP包含40-50摩尔%的阳离子脂质;5-15摩尔%的中性脂质;35-45摩尔%的类固醇;1-10摩尔%的聚合物缀合的脂质;以及包封在脂质纳米颗粒内或与脂质纳米颗粒缔合的RNA。

[0427] 在一实施方案中,基于脂质纳米颗粒中存在的脂质的总摩尔确定摩尔%。

[0428] 在一实施方案中,中性脂质选自DSPC、DPPC、DMPC、DOPC、POPC、DOPE、DOPG、DPPG、POPE、DPPE、DMPE、DSPE和SM。在一实施方案中,中性脂质选自DSPC、DPPC、DMPC、DOPC、POPC、DOPE和SM。在一实施方案中,中性脂质是DSPC。

[0429] 在一实施方案中,类固醇是胆固醇。

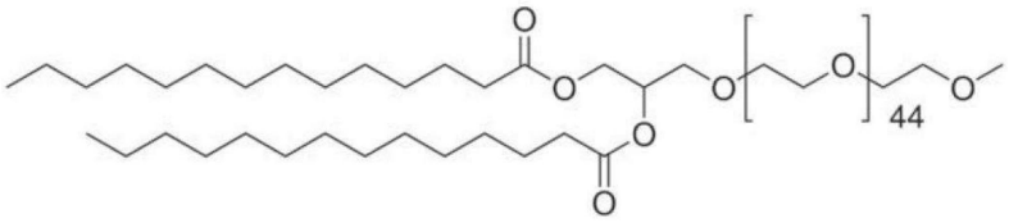
[0430] 在一实施方案中,聚合物缀合的脂质是聚乙二醇化脂质。在一实施方案中,聚乙二醇化脂质具有以下结构:

[0431] 

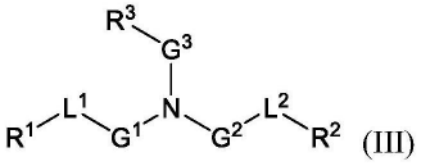
[0432] 或者其药学可接受的盐、互变异构体或立体异构体,其中:

[0433] R^{12}和R^{13}各自独立地是含有10-30个碳原子的直链或分支的、饱和或不饱和的烷基链,其中烷基链任选被一个或多个酯键中断;并且w的平均值为30-60。在一实施方案中,R^{12}和R^{13}各自独立地是含有12-16个碳原子的直链、饱和的烷基链。在一实施方案中,w的平均值为40-55。在一实施方案中,平均w为约45。在一实施方案中,R^{12}和R^{13}各自独立地是含有约14个碳原子的直链、饱和的烷基链,并且w的平均值为约45。

[0434] 在一实施方案中,聚乙二醇化脂质是DMG-PEG 2000,例如,具有以下结构:

[0435] 

[0436] 在一些实施方案中,LNP的阳离子脂质组分具有式(III)的结构:

[0437] 

[0438] 或者其药学可接受的盐、互变异构体、前药或立体异构体,其中:

[0439] L^1或L^2之一是-O(C=O)-、-(C=O)O-、-C(=O)-、-O-、-S(O)$_x$-、-S-S-、C(=O)S-、SC(=O)-、-NR^aC(=O)-、-C(=O)NR^a-、NR^aC(=O)NR^a、-OC(=O)NR^a-或-NR^aC(=O)O-,并且L^1或L^2中的另一个是-O(C=O)-、-(C=O)O-、-C(=O)-、-O-、-S(O)$_x$-、-SS-、-C(=O)S-、SC(=O)-、-NR^aC(=O)-、-C(=O)NR^a-、NR^aC(=O)NR^a、-OC(=O)NR^a-或-NR^aC(=O)O-或直接的键;

[0440] G^1和G^2各自独立地是未取代的C_1-C_{12}亚烷基或C_1-C_{12}亚烯基;

[0441] G^3是C_1-C_{24}亚烷基、C_1-C_{24}亚烯基、C_3-C_8环亚烷基、C_3-C_8环亚烯基;

[0442] R^a是H或C_1-C_{12}烷基;

[0443] R^1和R^2各自独立地是C_6-C_{24}烷基或C_6-C_{24}烯基;

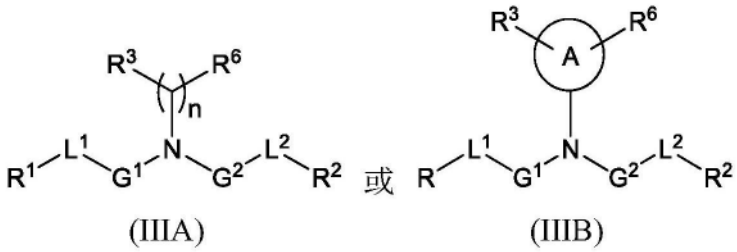
[0444] R^3是H、OR^5、CN、-C(=O)OR^4、-OC(=O)R^4或-NR^5C(=O)R^4;

[0445] R^4是C_1-C_{12}烷基;

[0446] R^5是H或C_1-C_6烷基;并且

[0447] x是0、1或2。

[0448] 在式(III)的一些前述实施方案中,脂质具有以下结构(IIIA)或(IIIB)之一:

[0449] 

[0450] 其中：

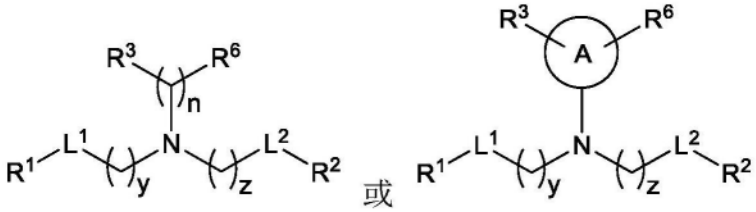
[0451] A是3-8元环烷基或环亚烷基环；

[0452] 在每次出现时，R^6独立地是H、OH或C_1-C_{24}烷基；

[0453] n是1-15的整数。

[0454] 在式(III)的一些前述实施方案中，脂质具有结构(IIIA)，而在其他实施方案中，脂质具有结构(IIIB)。

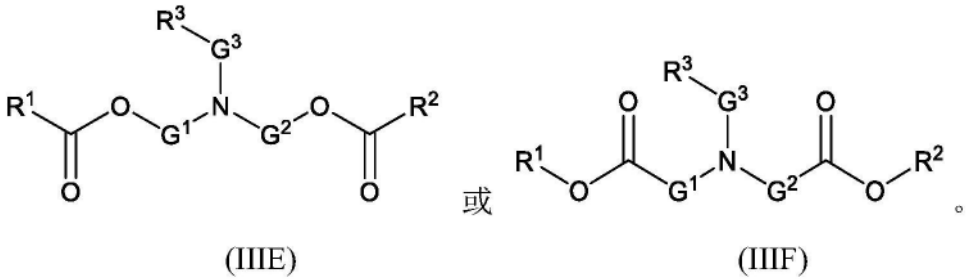
[0455] 在式(III)的其他实施方案中，脂质具有以下结构(IIIC)或(IIID)之一：

[0456] 

[0457] 其中y和z各自独立地是1-12的整数。

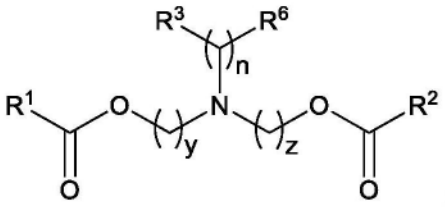
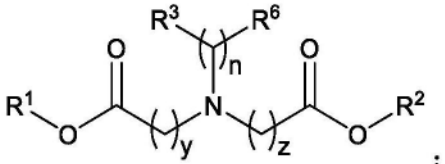
[0458] 在式(III)的任何前述实施方案中，L^1或L^2之一是-O(C=O)-。例如，在一些实施方案中，L^1和L^2各自是-O(C=O)-。在前述任一项的一些不同实施方案中，L^1和L^2各自独立地是-(C=O)O-或-O(C=O)-。例如，在一些实施方案中，L^1和L^2各自是-(C=O)O。

[0459] 在式(III)的一些不同实施方案中，脂质具有以下结构(IIIE)或(IIIF)之一：

[0460] 

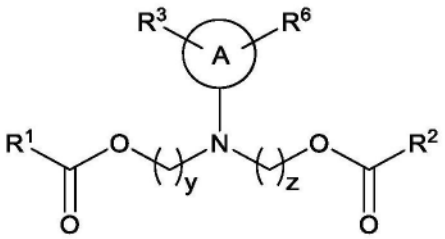
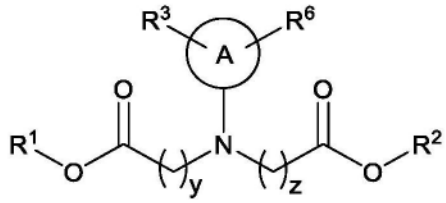
[0461] 在式(III)的一些前述实施方案中，脂质具有以下结构(IIIG)、(IIIH)、(IIII)或(IIIJ)之一：

[0462]

(IIIG) (IIIH)

[0463]


或


(IIII) (IIIJ)

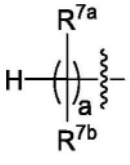
[0464] 在式(III)的一些前述实施方案中,n是2-12的整数,例如2-8或2-4。例如,在一些实施方案中,n是3、4、5或6。在一些实施方案中,n是3。在一些实施方案中,n是4。在一些实施方案中,n是5。在一些实施方案中,n是6。

[0465] 在式(III)的一些其他前述实施方案中,y和z各自独立地是2-10的整数。例如,在一些实施方案中,y和z各自独立地是4-9或4-6的整数。

[0466] 在式(III)的一些前述实施方案中,R^6是H。在其他前述实施方案中,R^6是C_1-C_{24}烷基。在其他实施方案中,R^6是OH。

[0467] 在式(III)的一些实施方案中,G^3是未取代的。在其他实施方案中,G^3是取代的。在各种不同实施方案中,G^3是线性C_1-C_{24}亚烷基或线性C_1-C_{24}亚烯基。

[0468] 在式(III)的一些其他前述实施方案中,R^1或R^2或两者是C_6-C_{24}烯基。例如,在一些实施方案中,R^1和R^2各自独立地具有以下结构:

[0469] ,

[0470] 其中:

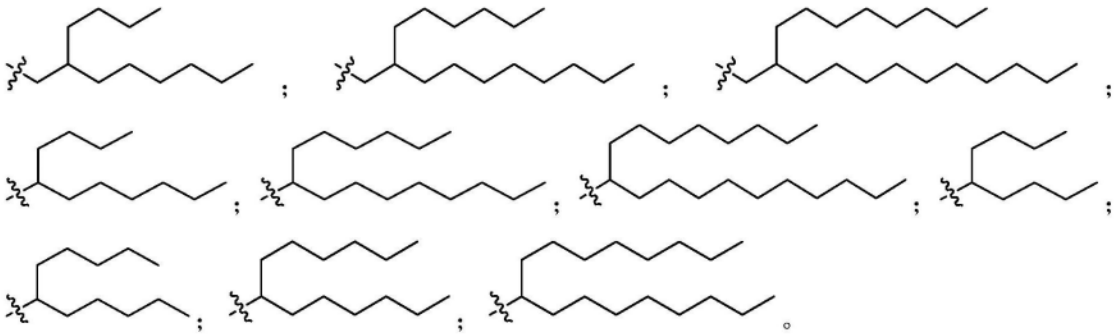
[0471] 在每次出现时,R^{7a}和R^{7b}独立地是H或C_1-C_{12}烷基;并且

[0472] a是2-12的整数,

[0473] 其中各自选择R^{7a}、R^{7b}和a,从而R^1和R^2各自独立地包含6-20个碳原子。例如,在一些实施方案中,a是5-9或8-12的整数。

[0474] 在式(III)的一些前述实施方案中,R^{7a}的至少一次出现是H。例如,在一些实施方案中,在每次出现时,R^{7a}是H。在前述其他不同实施方案中,R^{7b}的至少一次出现是C_1-C_8烷基。例如,在一些实施方案中,C_1-C_8烷基是甲基、乙基、正丙基、异丙基、正丁基、异丁基、叔丁基、正己基或正辛基。

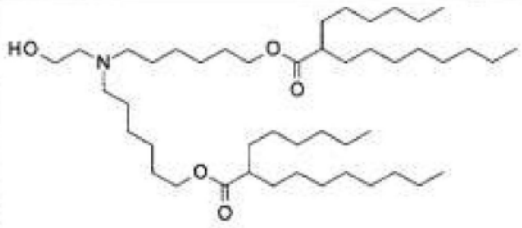
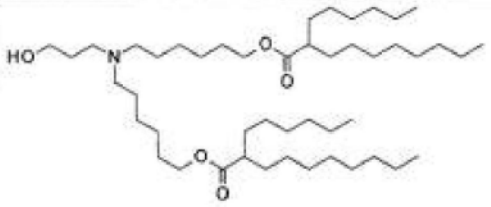
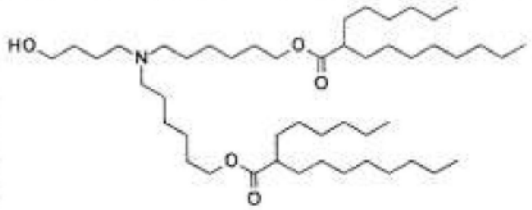
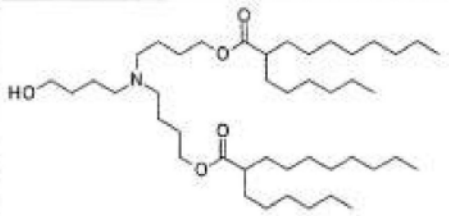
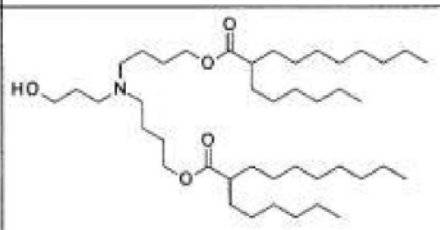
[0475] 在式(III)的不同实施方案中,R^1或R^2或两者具有以下结构之一:

[0476] 

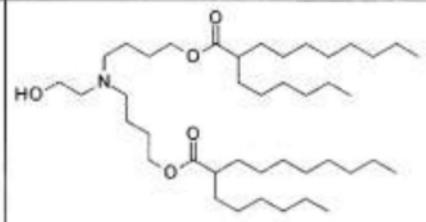
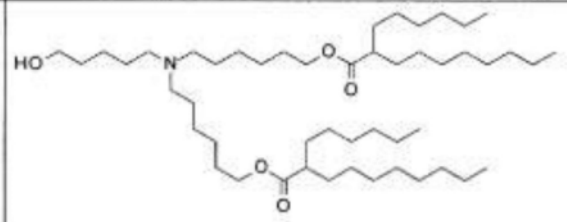
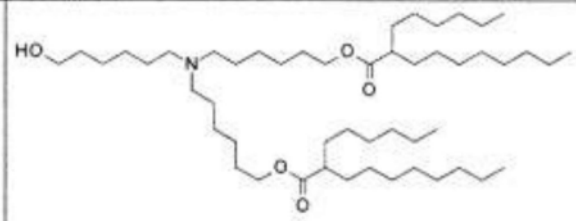
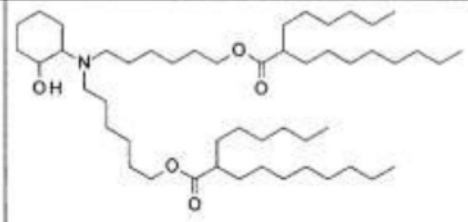
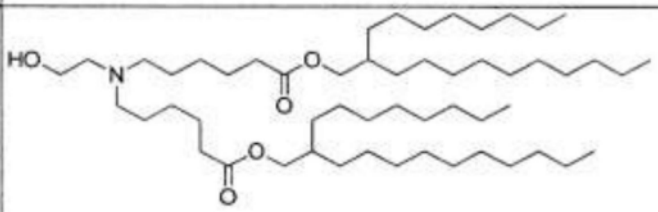
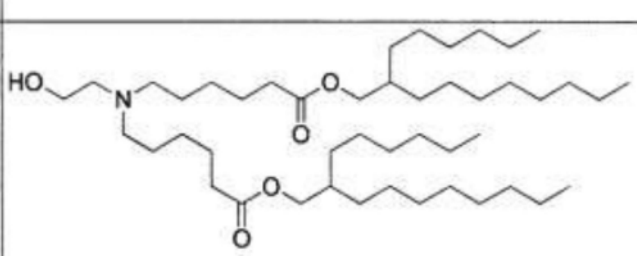
[0477] 在式(III)的一些前述实施方案中,R^3是OH、CN、-C(=O)OR^4、-OC(=O)R^4或-NHC(=O)R^4。在一些实施方案中,R^4是甲基或乙基。

[0478] 在各种不同实施方案中,式(III)的阳离子脂质具有下表所示的结构之一。式(III)的代表性化合物。

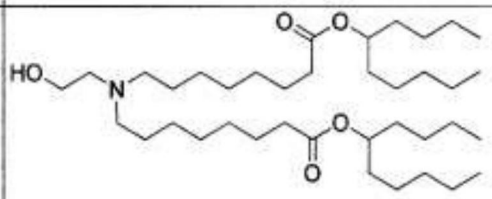
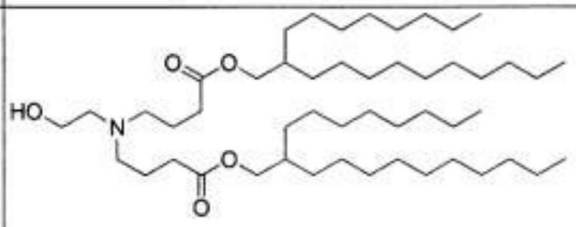
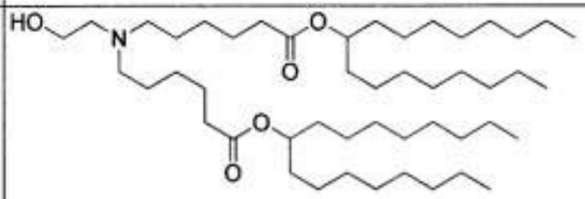
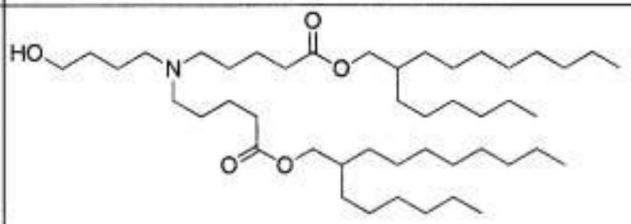
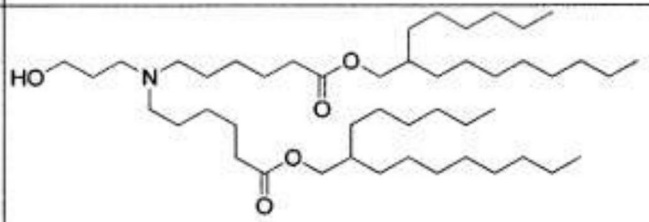
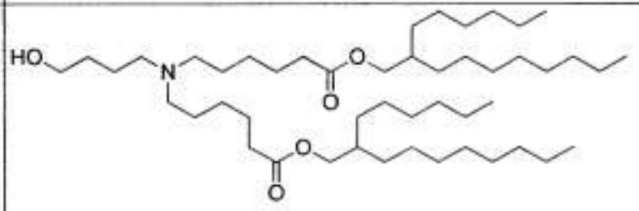
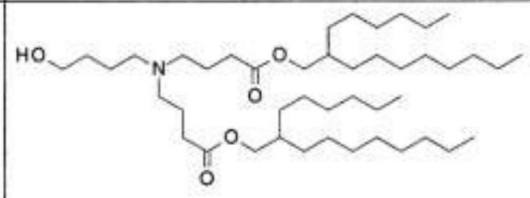
[0479]

编号	结构
III-1	
III-2	
III-3	
III-4	
III-5	

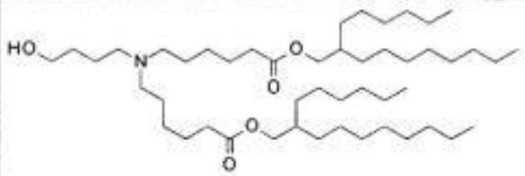
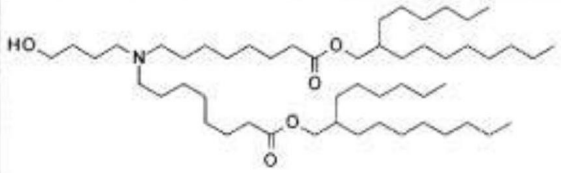
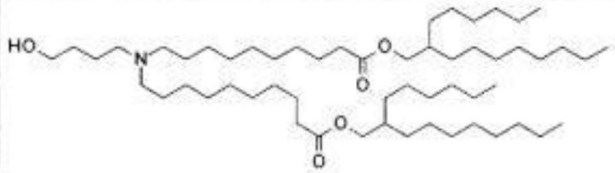
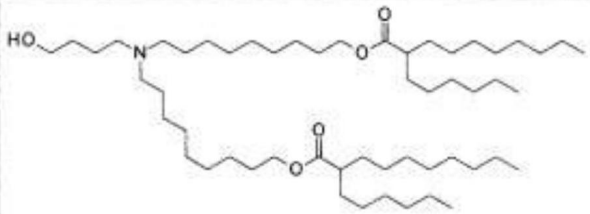
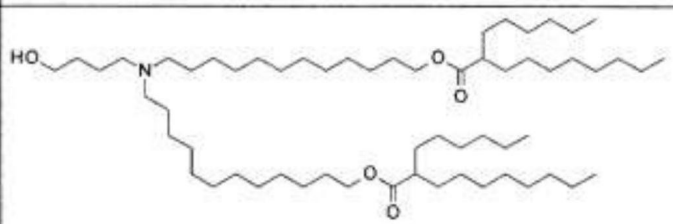
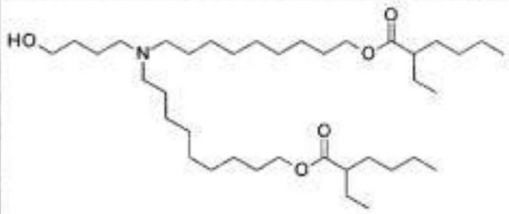
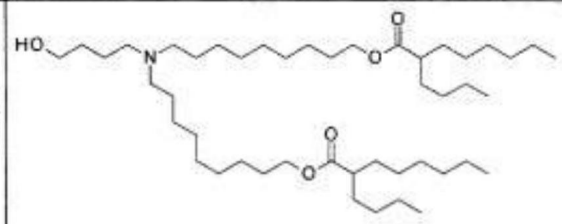
[0480]

编号	结构
III-6	
III-7	
III-8	
III-9	
III-10	
III-11	

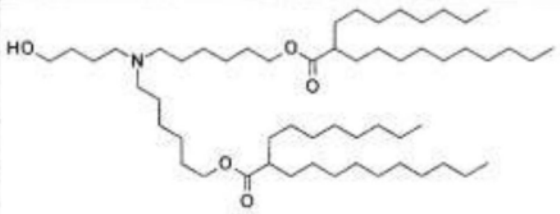
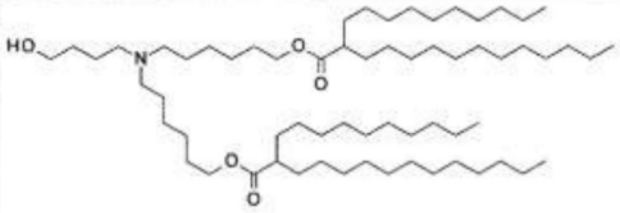
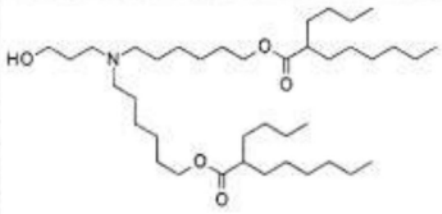
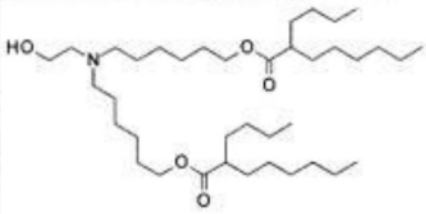
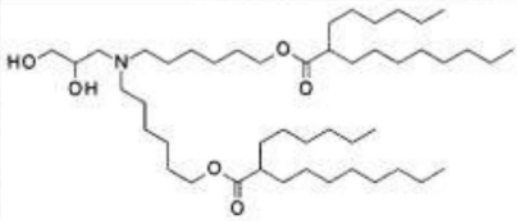
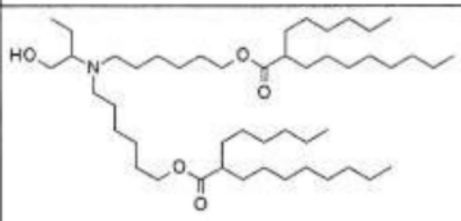
[0481]

编号	结构
III-12	
III-13	
III-14	
III-15	
III-16	
III-17	
III-18	

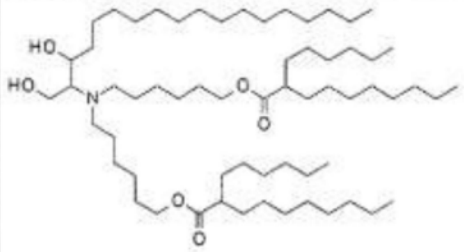
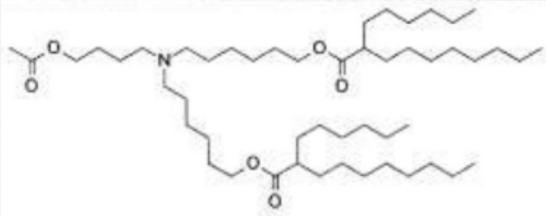
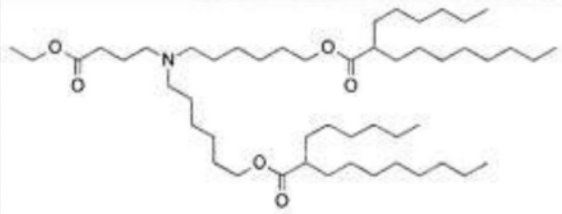
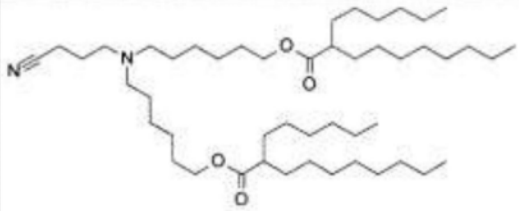
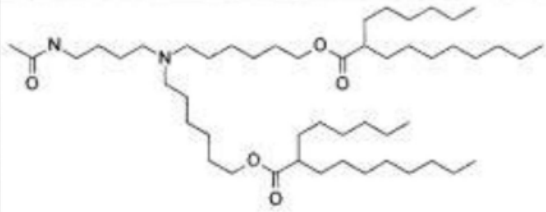
[0482]

编号	结构
III-19	
III-20	
III-21	
III-22	
III-23	
III-24	
III-25	

[0483]

编号	结构
III-26	
III-27	
III-28	
III-29	
III-30	
III-31	

[0484]

编号	结构
III-32	
III-33	
III-34	
III-35	
III-36	

[0485] 在一些实施方案中,LNP包含式(III)的脂质、RNA、中性脂质、类固醇和聚乙二醇化脂质。在一些实施方案中,式(III)的脂质是化合物III-3。在一些实施方案中,中性脂质是DSPC。在一些实施方案中,类固醇是胆固醇。在一些实施方案中,聚乙二醇化脂质是ALC-0159。

[0486] 在一些实施方案中,阳离子脂质以约40至约50摩尔百分比的量存在于LNP中。在一实施方案中,中性脂质以约5至约15摩尔百分比的量存在于LNP中。在一实施方案中,类固醇以约35至约45摩尔百分比的量存在于LNP中。在一实施方案中,聚乙二醇化脂质以约1至约10摩尔百分比的量存在于LNP中。

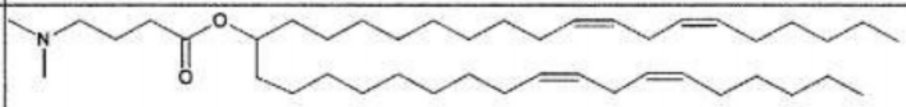
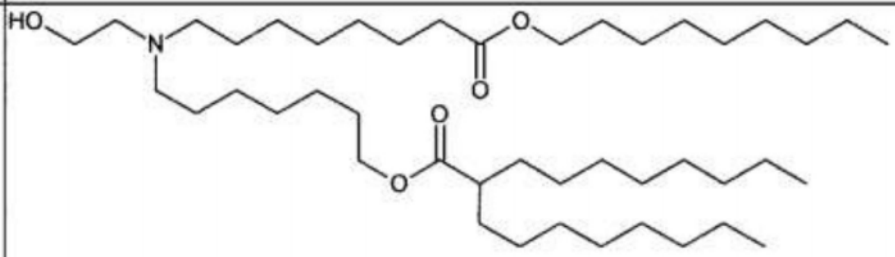
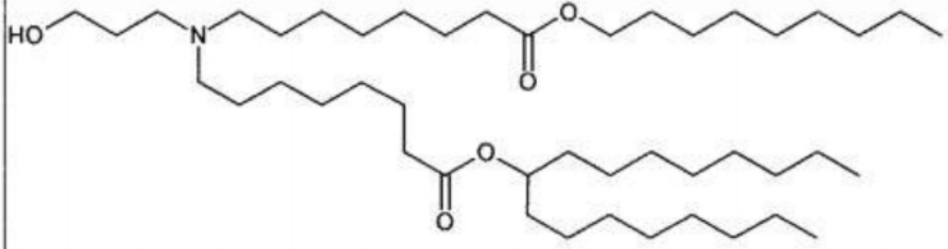
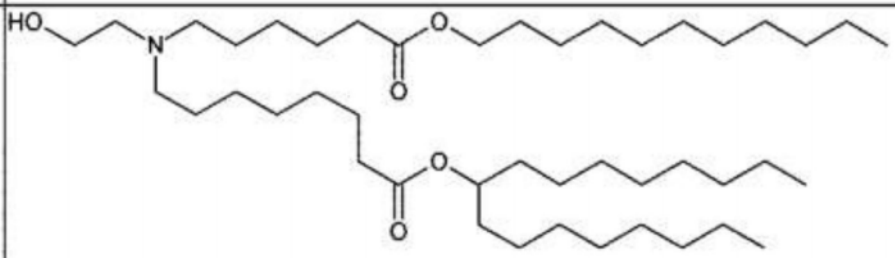
[0487] 在一些实施方案中,LNP包含约40至约50摩尔百分比量的化合物III-3,约5至约15

摩尔百分比量的DSPC,约35至约45摩尔百分比量的胆固醇,以及约1至约10摩尔百分比量的ALC-0159。

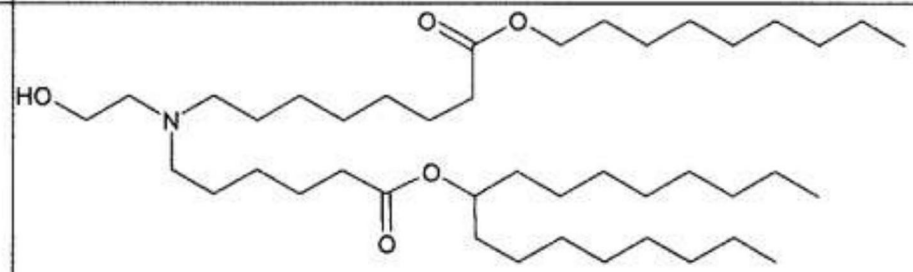
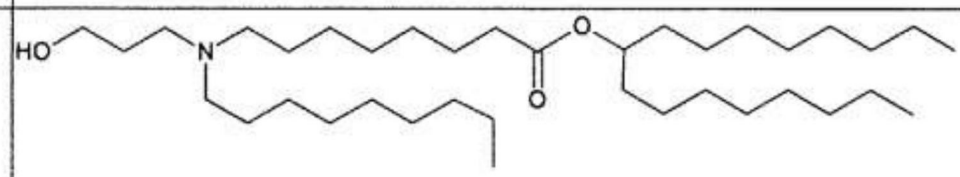
[0488] 在一些实施方案中,LNP包含约47.5摩尔百分比量的化合物III-3,约10摩尔百分比量的DSPC,约40.7摩尔百分比量的胆固醇,以及约1.8摩尔百分比量的ALC-0159。

[0489] 在各种不同实施方案中,阳离子脂质具有下表所示的结构之一。

[0490]

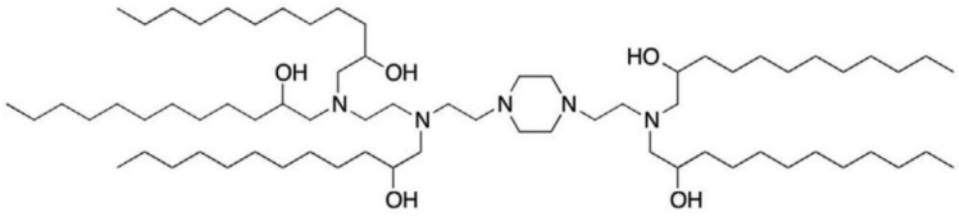
编号	结构
A	
B	
C	
D	

[0491]

编号	结构
E	
F	

[0492] 在一些实施方案中,LNP包含上表所示的阳离子脂质,例如,式(B)或式(D)的阳离子脂质,特别是式(D)的阳离子脂质,RNA,中性脂质,类固醇和聚乙二醇化脂质。在一些实施方案中,中性脂质是DSPC。在一些实施方案中,类固醇是胆固醇。在一些实施方案中,聚乙二醇化脂质是DMG-PEG 2000。

[0493] 在一实施方案中,LNP包含阳离子脂质,其是可电离的脂质样材料(类脂质(lipidoid))。在一实施方案中,阳离子脂质具有以下结构:

[0494] 

[0495] N/P值优选是至少约4。在一些实施方案中,N/P值范围为4-20、4-12、4-10、4-8或5-7。在一实施方案中,N/P值是约6。

[0496] 本文所述的LNP的平均直径在一实施方案中可以是约30nm至约200nm或约60nm至约120nm。

[0497] RNA靶向

[0498] 本公开的一些方面涉及本文公开的rRNA(例如,编码疫苗抗原和/或免疫刺激剂的RNA)的靶向递送。

[0499] 在一实施方案中,本公开涉及靶向肺。如果给药的RNA是编码疫苗抗原的RNA,则特别优选靶向肺。例如,通过吸入,通过给药可以配制为如本文所述的颗粒如脂质颗粒的RNA,可以将RNA递送至肺。

[0500] 在一实施方案中,本公开涉及靶向淋巴系统,特别是次级淋巴器官,更具体地是脾。如果给药的RNA是编码疫苗抗原的RNA,则特别优选靶向淋巴系统,特别是次级淋巴器官,更具体地是脾。

[0501] 在一实施方案中,靶细胞是脾细胞。在一实施方案中,靶细胞是抗原呈递细胞如脾

中的专职抗原呈递细胞。在一实施方案中,靶细胞是脾中的树突细胞。

[0502] “淋巴系统”是循环系统的一部分,并且是免疫系统的重要部分,其包含携带淋巴的淋巴管网络。淋巴系统由淋巴器官、淋巴管的传导网络和循环淋巴组成。初级或中枢淋巴器官从未成熟的祖细胞产生淋巴细胞。胸腺和骨髓构成初级淋巴器官。包括淋巴结和脾在内的次级或外周淋巴器官维持成熟的幼稚淋巴细胞并启动适应性免疫应答。

[0503] 可以通过所谓的脂质复合物制剂将RNA递送至脾,其中RNA结合至包含阳离子脂质以及任选存在的额外的或辅助脂质的脂质体以形成可注射的纳米颗粒制剂。脂质体可以通过将脂质在乙醇中的溶液注射至水或合适的水相中来获得。RNA脂质复合物颗粒可以通过将脂质体与RNA混合来制备。WO 2013/143683中描述了靶向脾的RNA脂质复合物颗粒,其援引加入本文。已发现具有净负电荷的RNA脂质复合物颗粒可以用来优先靶向脾组织或脾细胞如抗原呈递细胞,特别是树突细胞。因此,在给药RNA脂质复合物颗粒之后,在脾中发生RNA积累和/或RNA表达。因此,本公开的RNA脂质复合物颗粒可以用于在脾中表达RNA。在一实施方案中,给药RNA脂质复合物颗粒之后,在肺和/或肝中不发生或基本上不发生RNA积累和/或RNA表达。在一实施方案中,给药RNA脂质复合物颗粒之后,在抗原呈递细胞如脾中的专职抗原呈递细胞中发生RNA积累和/或RNA表达。因此,本公开的RNA脂质复合物颗粒可以用于在这类抗原呈递细胞中表达RNA。在一实施方案中,抗原呈递细胞是树突细胞和/或巨噬细胞。

[0504] 本公开的RNA脂质复合物颗粒的电荷是至少一种阳离子脂质中存在的电荷和RNA中存在的电荷的总和。电荷比是至少一种阳离子脂质中存在的正电荷与RNA中存在的负电荷的比例。通过以下等式计算至少一种阳离子脂质中存在的正电荷与RNA中存在的负电荷的电荷比:电荷比=[(阳离子脂质浓度(摩尔))*(阳离子脂质中正电荷的总数)]/[(RNA浓度(摩尔))*(RNA中负电荷的总数)]。

[0505] 本文所述的靶向脾的RNA脂质复合物颗粒在生理pH下优选具有净负电荷,如正电荷与负电荷的电荷比为约1.9:2至约1:2,或约1.6:2-约1:2,或约1.6:2-约1.1:2。在具体实施方案中,在生理pH下RNA脂质复合物颗粒中正电荷与负电荷的电荷比为约1.9:2.0、约1.8:2.0、约1.7:2.0、约1.6:2.0、约1.5:2.0、约1.4:2.0、约1.3:2.0、约1.2:2.0、约1.1:2.0或约1:2.0。

[0506] 可以通过在将RNA优先递送至肝或肝组织的制剂中向受试者给药编码免疫刺激剂的RNA来向受试者提供免疫刺激剂。优选将RNA递送至这样的靶器官或组织,特别地,如果期望表达大量的免疫刺激剂和/或如果期望或需要全身存在免疫刺激剂,特别是大量的。

[0507] RNA递送系统对肝有天生的偏好。这涉及基于脂质的颗粒,阳离子和中性纳米颗粒,特别是脂质纳米颗粒如脂质体、纳米胶束和生物共轭物中的亲脂性配体。肝脏积累是由肝血管系统或脂质代谢(脂质体以及脂质或胆固醇缀合物)的不连续性质引起的。

[0508] 为了将RNA体内递送至肝,药物递送系统可以用于通过防止RNA降解将其运送至肝。例如,由聚(乙二醇)(PEG)包被的表面和含mRNA的核心组成的聚合复合物纳米胶束是可用的系统,因为纳米胶束在生理条件下提供优秀的RNA体内稳定性。此外,由密集的PEG栅(palisade)组成的聚合复合物纳米胶束表面提供的隐形特性有效规避宿主免疫防御。

[0509] 用于靶向肝的合适免疫刺激剂的实例是参与T细胞增殖和/或维持的细胞因子。合适的细胞因子的实例包括IL2或IL7,其片段和变体,以及这些细胞因子、片段和变体的融合

蛋白,如PK延长的细胞因子。

[0510] 在另一实施方案中,编码免疫刺激剂的RNA可以在将RNA优先递送至淋巴系统(特别是次级淋巴器官,更特别是脾)的制剂中给药。优选将免疫刺激剂递送至这样的靶组织,特别地,如果期望在这种器官或组织中存在免疫刺激剂(例如,为了诱导免疫应答,特别是在T-细胞引发期间或为了激活常驻免疫细胞需要免疫刺激剂如细胞因子的情况下),但是不期望免疫刺激剂全身存在,特别是大量存在(例如,因为免疫刺激剂具有全身毒性)。

[0511] 合适的免疫刺激剂的实例是参与T细胞引发的细胞因子。合适的细胞因子的实例包括IL12、IL15、IFN-α或IFN-β,其片段和变体,以及这些细胞因子、片段和变体的融合蛋白,如PK延长的细胞因子。

[0512] 产生蛋白的方法

[0513] 本发明还提供一种在细胞中产生所关注的蛋白的方法,所述方法包括以下步骤:

[0514] (a)获得本发明的RNA复制子,其包含编码所关注的蛋白的开放阅读框,以及

[0515] (b)将RNA复制子接种至细胞中。

[0516] 在所述方法的各种实施方案中,RNA复制子如上文对本发明的RNA复制子所定义,只要RNA复制子包含编码所关注的蛋白的开放阅读框,任选包含编码功能性非结构蛋白的开放阅读框,并且可以通过功能性非结构蛋白复制,其中rRNA可以包含至少一个修饰的核苷酸以及一个或多个调节序列中的点突变,其恢复或改善修饰的rRNA的功能。

[0517] 可以接种一种或多种核酸分子的细胞可以称作“宿主细胞”。根据本发明,术语“宿主细胞”是指可以用外源核酸分子转化或转染的任何细胞。术语“细胞”优选是完整细胞,即具有完整膜的尚未释放其正常胞内组分如酶、细胞器或遗传物质的细胞。完整细胞优选是活细胞,即能够进行其正常代谢功能的活细胞。根据本发明,术语“宿主细胞”包括原核(例如大肠杆菌)或真核细胞(例如人和动物细胞、植物细胞、酵母细胞和昆虫细胞)。特别优选哺乳动物细胞如来自人、小鼠、仓鼠、猪、驯化动物(包括马、牛、绵羊和山羊)以及灵长类的细胞。细胞可以源自多种组织类型,并且包括原代细胞和细胞系。具体实例包括角质形成细胞、外周血白细胞、骨髓干细胞和胚胎干细胞。在其他实施方案中,宿主细胞是抗原呈递细胞,特别是树突细胞、单核细胞或巨噬细胞。核酸可以以单个或多个拷贝的形式存在于宿主细胞中,并且在一实施方案中,在宿主细胞中表达。

[0518] 所述细胞可以是原核细胞或真核细胞。原核细胞在本文中适用于例如本发明的DNA的增殖,而真核细胞在本文中适用于例如复制子的开放阅读框的表达。

[0519] 在本发明的方法中,可以使用任何本发明的RNA复制子、或本发明的试剂盒、或本发明的药物组合物。RNA可以以药物组合物的形式使用,或者作为裸RNA使用,例如用于电穿孔。

[0520] 在本发明的在细胞中产生蛋白的方法中,所述细胞可以是抗原呈递细胞,并且所述方法可以用于表达编码抗原的RNA。为此,本发明可以涉及将编码抗原的RNA引入抗原呈递细胞如树突细胞。为了转染抗原呈递细胞如树突细胞,可以使用包含编码抗原的RNA的药物组合物。

[0521] 在一实施方案中,在细胞中产生蛋白的方法是体外方法。在一实施方案中,在细胞中产生蛋白的方法不包括通过手术或治疗从人或动物受试者移出细胞。

[0522] 在这个实施方案中,可以将本发明接种的细胞施用于受试者,以便在受试者中产

生蛋白并向受试者提供蛋白。就受试者而言,所述细胞可以是自体的、同基因的、同种异体的或异源的。

[0523] 在其他实施方案中,在细胞中产生蛋白的方法中的细胞可以存在于受试者如患者中。在这些实施方案中,在细胞中产生蛋白的方法是体内方法,其包括向受试者给药RNA分子。

[0524] 在这方面,本发明还提供一种在受试者中产生所关注的蛋白的方法,所述方法包括以下步骤:

[0525] (a)获得本发明的RNA复制子,其包含编码所关注的蛋白的开放阅读框,以及

[0526] (b)向受试者给药RNA复制子。

[0527] 在所述方法的各种实施方案中,RNA复制子如上文对本发明的RNA复制子所定义,只要RNA复制子包含编码所关注的蛋白的开放阅读框,任选包含编码功能性非结构蛋白的开放阅读框,并且可以通过功能性非结构蛋白复制,其中rRNA可以包含至少一个修饰的核苷酸以及一个或多个调节序列中的点突变,其恢复或改善修饰的rRNA的功能。

[0528] 任何本发明的RNA复制子、或本发明的试剂盒、或本发明的药物组合物可以用于本发明的在受试者中产生蛋白的方法。例如,在本发明的方法中,RNA可以以药物组合物的形式使用,例如如本文所述,或者作为裸RNA使用。

[0529] 鉴于向受试者给药的能力,每个本发明的RNA复制子、或本发明的试剂盒、或本发明的药物组合物都可以称作“药物”等。本发明预见到可提供本发明的RNA复制子、试剂盒、药物组合物以用作药物。所述药物可以用于治疗受试者。“治疗”表示向受试者给药如本文所述的化合物或组合物或其他实体。该术语包括通过疗法治疗人体或动物体的方法。

[0530] 上述药物通常不包含DNA,因此与现有技术中描述的DNA疫苗相比具有额外的安全特征(例如WO 2008/119827 A1)。

[0531] 本发明的可选医学用途包括本发明在细胞中产生蛋白的方法,其中所述细胞可以是抗原呈递细胞如树突细胞,然后将所述细胞引入受试者。例如,可以将编码药学活性蛋白如抗原的RNA离体引入(转染入)抗原呈递细胞,例如取自受试者的抗原呈递细胞,并且可以将任选离体克隆增殖的抗原呈递细胞重新引入相同或不同的受试者。可以使用本领域已知的任何方法将转染的细胞重新引入受试者。

[0532] 可以向有此需要的受试者给药本发明的药物。本发明的药物可以用于治疗受试者的预防性方法以及治疗性方法。

[0533] 本发明的药物以有效量给药。“有效量”是指单独或与其他剂量一起足以引起反应或期望效果的量。在治疗受试者的某种疾病或某种病况的情况下,期望的效果是抑制疾病进展。这包括疾病进展的减速,特别是疾病进展的中断。治疗疾病或病况中的期望效果也可以是延迟疾病爆发或抑制疾病爆发。

[0534] 有效量取决于治疗的病况,疾病的严重程度,患者的个体参数,包括年龄、生理状况、尺寸和体重,治疗的持续时间,伴随疗法的类型(如果存在),给药的具体模式以及其他因素。

[0535] 疫苗接种

[0536] 术语“免疫”或“疫苗接种”一般是指出于治疗或预防原因治疗受试者的过程。治疗,特别是预防性治疗,是或包括优选旨在诱导或增强受试者的免疫应答(例如针对一种或

多种抗原)的治疗。根据本发明,如果希望通过使用如本文所述的rRNA来诱导或增强免疫应答,则可以通过rRNA引发或增强免疫应答。在一实施方案中,本发明提供一种预防性治疗,其是或包括优选受试者的疫苗接种。本发明的一个实施方案,其中复制子编码作为所关注的蛋白的药学活性肽或蛋白,其是免疫活性化合物或抗原,特别适用于疫苗接种。

[0537] 先前已经描述RNA用于针对包括病原体或癌症的外来物质的疫苗接种(Ulmer et al.,2012,Vaccine 30:4414-4418最近综述)。与现有技术中的常用方法相反,本发明的复制子是用于有效疫苗接种的特别合适的元件,因为它能够通过如本文所述的功能性甲病毒非结构蛋白进行复制。本发明的疫苗接种可以用于例如诱导对弱免疫原性蛋白的免疫应答。在本发明的RNA疫苗的情况下,蛋白抗原从未暴露于血清抗体,而是在RNA翻译后由转染的细胞本身产生。因此,过敏反应应当不是问题。因此本发明允许患者重复免疫而没有过敏反应的风险。

[0538] 在涉及本发明的疫苗接种的方法中,向受试者给药本发明的药物,特别是如果需要治疗患有涉及抗原的疾病或有患涉及抗原的疾病风险的受试者。

[0539] 在涉及本发明的疫苗接种的方法中,由本发明的复制子编码的所关注的蛋白编码例如细菌抗原,针对其产生免疫应答,或编码病毒抗原,针对其产生免疫应答,或编码癌症抗原,针对其产生免疫应答,或编码单细胞生物体的抗原,针对其产生免疫应答。可以通过已知的标准方法评估疫苗接种的效力,如通过测量来自生物体的抗原特异性IgG抗体。在涉及本发明的过敏原特异性免疫疗法的方法中,由本发明的复制子编码的所关注的蛋白编码与过敏相关的抗原。过敏原特异性免疫疗法(也称作减敏)定义为向具有一种或多种过敏的生物体给药优选增加剂量的过敏原疫苗,以达到与随后暴露于致病性过敏原相关的症状缓解的状态。可以通过已知的标准方法评估过敏原特异性免疫疗法的效力,如通过测量来自生物体的过敏原特异性IgG和IgE抗体。

[0540] 可以向受试者给药本发明的药物,例如用于受试者的治疗,包括受试者的疫苗接种。

[0541] 术语“受试者”涉及脊椎动物,特别是哺乳动物。例如,本发明的上下文中的哺乳动物是人,非人灵长类,驯养哺乳动物如狗、猫、绵羊、牛、山羊、猪、马等,实验室动物如小鼠、大鼠、兔、豚鼠等,以及圈养动物如动物园的动物。术语“受试者”还涉及非哺乳类脊椎动物如鸟类(特别是驯养鸟类如鸡、鸭、鹅、火鸡)和鱼类(特别是养殖鱼类,例如鲑鱼或鲶鱼)。如本文所用的术语“动物”还包括人。

[0542] 在一些实施方案中优选向驯养动物如狗、猫、兔、豚鼠、仓鼠、绵羊、牛、山羊、猪、鸡、鸭、鹅、火鸡或野生动物如狐狸给药。例如,本发明的预防性疫苗接种可以适用于对动物群体(例如在农业中)或野生动物群体进行疫苗接种。其他圈养的动物群体,如宠物或动物园的动物,都可以进行疫苗接种。

[0543] 给药模式

[0544] 本发明的药物可以以任何合适的途径应用于受试者。

[0545] 例如,药物可以全身给药,例如静脉内(i.v.)、皮下(s.c.)、皮内(i.d.)或吸入。

[0546] 在一实施方案中,本发明的药物施用于肌肉组织,如骨骼肌,或皮肤,例如皮下。一般认为将RNA转移至皮肤或肌肉中会导致大量持续的局部表达,同时强烈诱导体液和细胞免疫应答(Johansson et al.,2012,PLoS.One.7:e29732;Geall et al.,2012,

Proc.Natl.Acad.Sci.U.S.A 109:14604-14609)。

[0547] 向肌肉组织或皮肤给药的替代方法包括但不限于:皮内、鼻内、眼内、腹腔内、静脉内、间质、口颊(buccal)、透皮和舌下给药。皮内和肌肉内给药是两种优选途径。

[0548] 给药可以通过各种方式实现。在一实施方案中,本发明的药物通过注射给药。在一优选实施方案中,注射是通过针。无针注射可以用作替代方法。

[0549] 通过图和实施例详细描述和说明本发明,其仅用于说明目的而不是表示限制。由于描述和实施例,技术人员容易理解同样包括在本发明内的其他实施方案。

附图说明

[0550] 图1A-1B.(A)载体设计(未按比例绘制)。由甲病毒基因组构建的自扩增RNA(saRNA)。saRNA是一种双顺反子RNA,具有编码甲病毒RNA依赖性RNA聚合酶(复制酶)的5'-开放阅读框(ORF)和编码所关注的基因的3'-开放阅读框(ORF)。ORF以AUG-起始密码子标记。saRNA的编码区的侧翼有病毒5'-和3'-非翻译区(vUTR)。此外,它含有由RNA复制所需的保守序列元件(CSE)组成的调节区(包含基因组正链-启动子(5'-复制识别序列RRS)的CSE 1&2,核心基因组负链启动子CSE 4以及亚基因组启动子CSE 3)。值得注意的是,5'-和3'-CSE协作启动负链合成。复制酶的ORF的起始密码子在5'-调节区内。(B)saRNA的5'末端。saRNA可以通过体外转录(IVT)产生,使用常规核苷酸ATP、CTP、GTP和UTP,或者ATP、CTP、GTP和N1-甲基-假-UTP。对于saRNA的共转录加帽,可以使用帽类似物GpppG或GpppAU。根据帽类似物的选择,倒数第二个U(下划线)被1mΨ交换,当该核苷酸代替UTP用于IVT时。显示源自SFV和VEEV的saRNA的5'末端没有帽。

[0551] 图2A-2H.建立saRNA复制的可能性。使用常规核苷酸ATP、CTP、GTP以及UTP(黑色柱)或N1-甲基-假-UTP(白色柱),通过IVT产生在5'末端具有GA或AU的编码GFP的saRNA。用指定剂量转染细胞,24小时后通过流式细胞术评估GFP阳性细胞的百分比。三个独立实验的平均值与标准差在以下图中示出:(A)用SFV衍生的saRNA电穿孔BHK-21细胞。(B)用VEEV衍生的saRNA电穿孔BHK-21细胞。(C)用SFV衍生的saRNA脂质体转染(lipofect)BHK-21细胞。(D)用VEEV衍生的saRNA脂质体转染BHK-21细胞。(E)用SFV衍生的saRNA电穿孔HFF细胞。(F)用VEEV衍生的saRNA电穿孔HFF细胞。(G)用SFV衍生的saRNA脂质体转染HFF细胞。(H)用VEEV衍生的saRNA脂质体转染HFF细胞。

[0552] 实施例

[0553] 材料和方法

[0554] 实施例中使用以下材料和方法。

[0555] 质粒的克隆、体外转录、RNA纯化:使用标准技术克隆质粒。实施例1中描述了本发明的实施例中使用的单个质粒的克隆细节。简而言之,两个质粒编码基于委内瑞拉马脑炎病毒特立尼达驴株(Trinidad donkey strain)的saRNA(VEEV;登录号L01442)。将增强型绿色荧光蛋白和可分泌型纳米萤光素酶的融合基因(GFP-SecNLuc)插入亚基因组启动子的下游。对于使用T7噬菌体聚合酶的体外转录,其中一个质粒包括病毒AU-5'末端上游的G。基于塞姆利基森林病毒克隆4(SFV4),克隆了两个类似的质粒用于saRNA。

[0556] 对于体外转录,通过在poly-A下游进行限制性消化将质粒线性化,以用作T7 RNA-聚合酶的模板。如以前所述进行RNA合成和纯化(Holtkamp et al.,2006,Blood 108:4009-

4017;Kuhn et al.,2010,Gene Ther.17:961-971)。根据需要,用N1-甲基-假-UTP交换UTP。

[0557] 通过分光光度法评估纯化RNA的质量,并且在5200片段分析仪(Advanced Analytical)上进行分析。实施例中转染至细胞中的所有RNA是体外转录的RNA(IVT-RNA)。

[0558] 细胞培养:除非另有说明,否则所有生长培养基、抗生素和其他补充剂均由Life Technologies/Gibco提供。胎牛血清(FCS)购自Sigma-Aldrich。将获得自System Bioscience的人包皮成纤维细胞(HFF,新生儿)在含有15% FCS、1%非必需氨基酸、1mM丙酮酸钠的最低必需培养基(MEM)中于37℃下培养。使细胞在37℃下于平衡至5% CO2的加湿环境中生长。使BHK21细胞(ATCC;CCL10)在补充了10% FCS的伊格尔最低必需培养基中生长。

[0559] RNA转移至细胞中:对于电穿孔,将15nM、3nM或0.6nM rRNA与32,000个细胞混合,最终体积为62.5μl/mm比色皿间隙尺寸。使用方波电穿孔装置(BTX ECM 830,Harvard Apparatus,美国马萨诸塞州霍利斯顿)在室温下进行电穿孔。对于使用的细胞类型,应用以下设置:HFF(500V/cm,1个24毫秒(ms)的脉冲);BHK21(750V/cm,1个16ms的脉冲)。

[0560] 使用Lipofectamine MessengerMAX按照制造商的说明(Life Technologies,德国达姆施塔特)进行RNA脂质体转染。将HFF和BHK21细胞以25,000个细胞/cm^2生长面积铺板,24h后用250ng/cm^2 RNA和1μl/cm^2 MessengerMAX、50ng/cm2 RNA和0.2μl/cm2 MessengerMAX或者10ng/cm2 RNA和0.04μl/cm^2 MessengerMAX转染。

[0561] 萤光素酶测定:为了评估转染细胞中萤光素酶的表达,将转染的细胞在96-孔黑色微孔板中铺板(Nunc,德国朗根塞尔博德)中。用Bright-Glo萤光素酶测定系统根据制造商的说明进行萤火虫萤光素酶的检测。使用酶标仪Infinite M200(Tecan Group,瑞士门内多夫)测量生物发光。在给定时间点测定的萤光素酶活性随时间作图,并且通过梯形法则计算曲线下面积。萤光素酶阴性细胞用于评估背景信号。

[0562] 流式细胞术(CD90.1;GFP):对于流式细胞术,不对表达GFP的细胞进行染色。使用BD FACS Canto II流式细胞仪测量GFP荧光。使用配套的FACSDiva软件或FlowJo软件进行数据分析。

[0563] 实施例1载体设计

[0564] 自扩增(可复制)RNA(saRNA或rRNA)构建自委内瑞拉马脑炎病毒(VEEV,Genbank登录号L01443)和塞姆利基森林病毒(SFV;克隆SFV4)的甲病毒基因组。甲病毒saRNA一般特征在于以下必不可少的结构域和编码区(图1A)。关于编码区,最常见形式的saRNA是具有两个开放阅读框(ORF)的双顺反子RNA。5'-ORF编码甲病毒非结构多聚蛋白(nsP),由4个亚基组成,具有RNA依赖性RNA转录所必需的所有酶促功能。nsP逐步成熟,它构建了一种与细胞膜相关的蛋白复合物(所谓的复制酶),在细胞膜上形成所谓的小球(spherule),作为RNA扩增的隔室(compartment)。位于复制酶ORF下游并通过亚基因组启动子(SGP)与之隔开的第二ORF承担所关注的基因(通常是抗原)的遗传信息。在甲病毒物种之间保守的许多RNA-结构域(保守序列元件;CSE)控制与复制酶的相互作用,从而控制RNA复制。在5'末端,2个CSE位于与复制酶的ORF重叠的较长高度结构化区域(称作5'-复制识别序列;RRS)内。这个区域比其他更下游的CSE(CSE3=亚基因组启动子和CSE4=3'-末端CSE)更长和更加结构化。此外,saRNA是5'-加帽和3'-多聚腺苷酸化的。除了能够实现复制酶的高效翻译外,poly-A尾的前大约15个腺苷是3'CSE的一部分。

[0565] 为了使用噬菌体聚合酶T7体外转录模板双链DNA,在对应于病毒衍生的5'末端的cDNA序列上游紧接着插入G:C碱基对,因为T7以G开始RNA合成。为了实现共转录加帽,使用含有7-甲基鸟苷酸和G的帽类似物(GpppG)。为了产生具有病毒AU 5'末端的saRNA,去除G:C碱基对,并且使用用于共转录加帽的帽类似物,其含有7-甲基鸟苷酸帽和AU二核苷酸(GpppAU),使能使用T7启动。因此,如果在IVT使用GpppG帽类似物以及ATP、GTP、CTP和N1-甲基-假-UTP,则病毒基因组第2位的保守U会变为1mΨ,但如果使用GpppAU则不会(图1B)。

[0566] 实施例2保守的倒数第二个尿苷改善1mΨ修饰的saRNA的表达

[0567] 在我们的研究中,saRNA(rRNA)是由VEEV和SFV工程化的。使用未修饰的ATP、GTP和CTP以及UTP或N1-甲基-假-UTP,在合成的帽类似物(β-S-ARCA,AU(cap1))的存在下进行体外转录。将BHK-21细胞以及原代人包皮成纤维细胞(HFF)用3种不同剂量的各种saRNA电穿孔或脂质体转染。为了监测转染细胞群中的saRNA表达,saRNA编码增强型绿色荧光蛋白(GFP)。转染后24h,通过流式细胞术定量GFP阳性细胞。

[0568] 据观察,当甲病毒基因组的倒数第二个U(penultimate)不被1mΨ替代时,通过GFP表达反映的用修饰的saRNA建立saRNA复制的可能性始终更大。转基因表达的这种改善与细胞类型、转染方法以及甲病毒来源或saRNA的剂量无关(图2A至2H)。由于BHK-21细胞具有缺陷的先天免疫,而HFF相反表现出强烈的抗病毒应答,因此这些数据为以下结论提供了证据:GFP表达的改善不是先天免疫刺激减少的结果,而是RNA复制改善的结果。

[0569] 本文参考的文件和研究的引用并不意味着承认任何前述内容是相关的现有技术。关于这些文件内容的所有陈述是基于申请人可获得的信息,并且不构成对这些文件内容正确性的任何承认。

[0570] 提出以下描述以使本领域普通技术人员能够制备和使用各种实施方案。特定设备、技术和应用的描述仅作为实例提供。对本文描述的实例的各种修改对于本领域普通技术人员是明显的,并且在不背离各种实施方案的精神和范围的情况下,本文中定义的一般原理可以应用于其他实例和应用。因此,各种实施方案不是为了限制本文描述和示出的实例,而是与符合权利要求的范围一致。

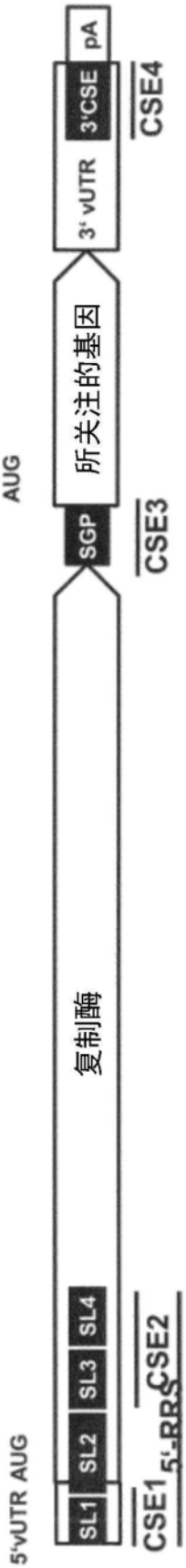

5'vUTR
AUG
SL1
SL2
SL3
SL4
CSE1
5'-RRS
CSE2
复制酶
SGP
CSE3
AUG
所关注的基因
3' vUTR
3'CSE
pA
CSE4

图1A

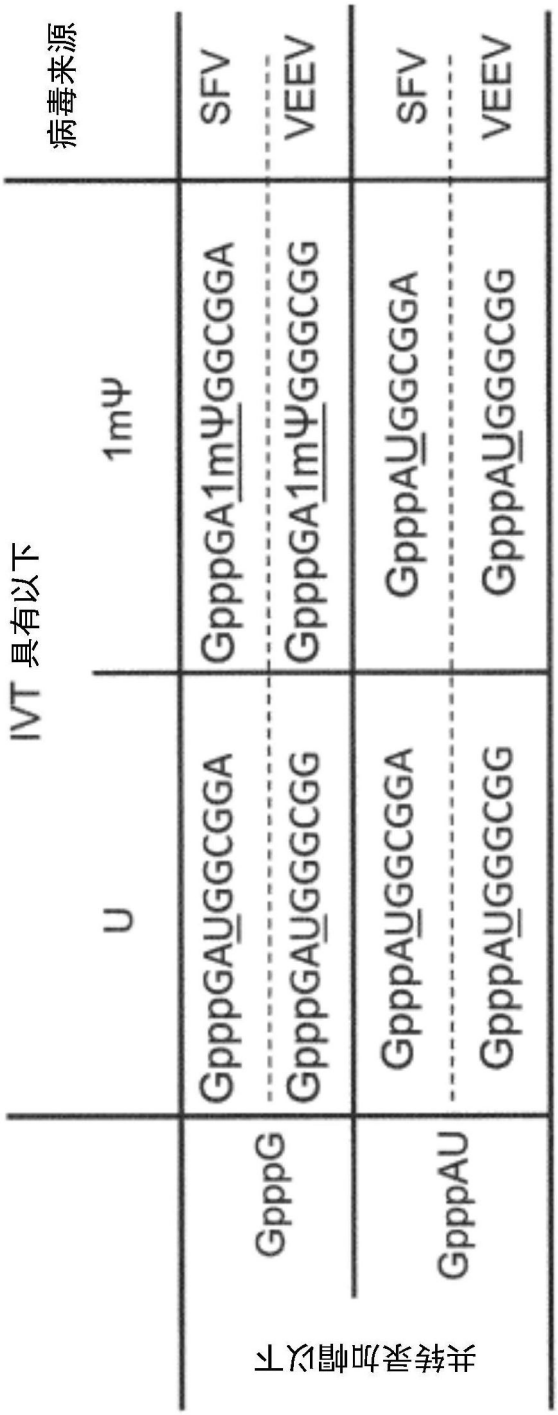

共转录加帽以下
GpppG
GpppAU
IVT 具有以下
U
1mΨ
病毒来源
GpppGAUGGCGGA
GpppGAUGGGCGG
GpppGA1mΨGGCGGA
GpppGA1mΨGGGCGG
GpppAUGGCGGA
GpppAUGGGCGG
GpppAUGGCGGA
GpppAUGGGCGG
SFV
VEEV
SFV
VEEV

图1B

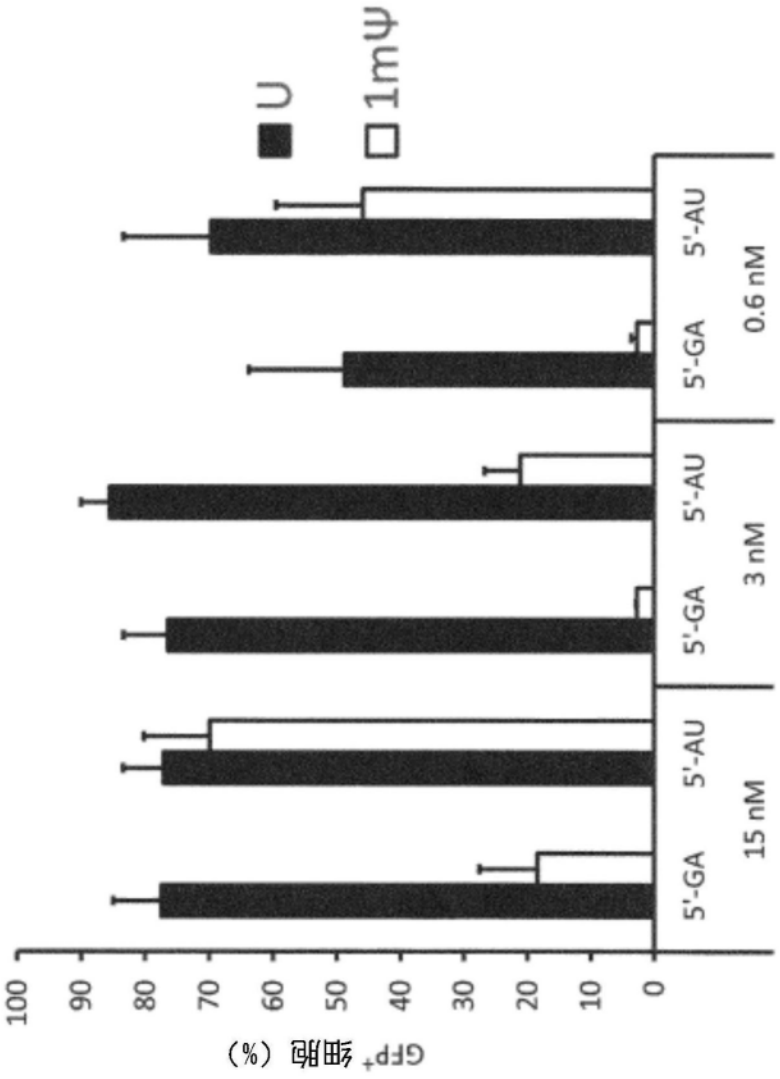

U
1mΨ
5'-AU
5'-GA
0.6 nM
5'-AU
5'-GA
3 nM
5'-AU
5'-GA
15 nM
100
90
80
70
60
50
40
30
20
10
0
GFP+ 细胞 (%)

图2A

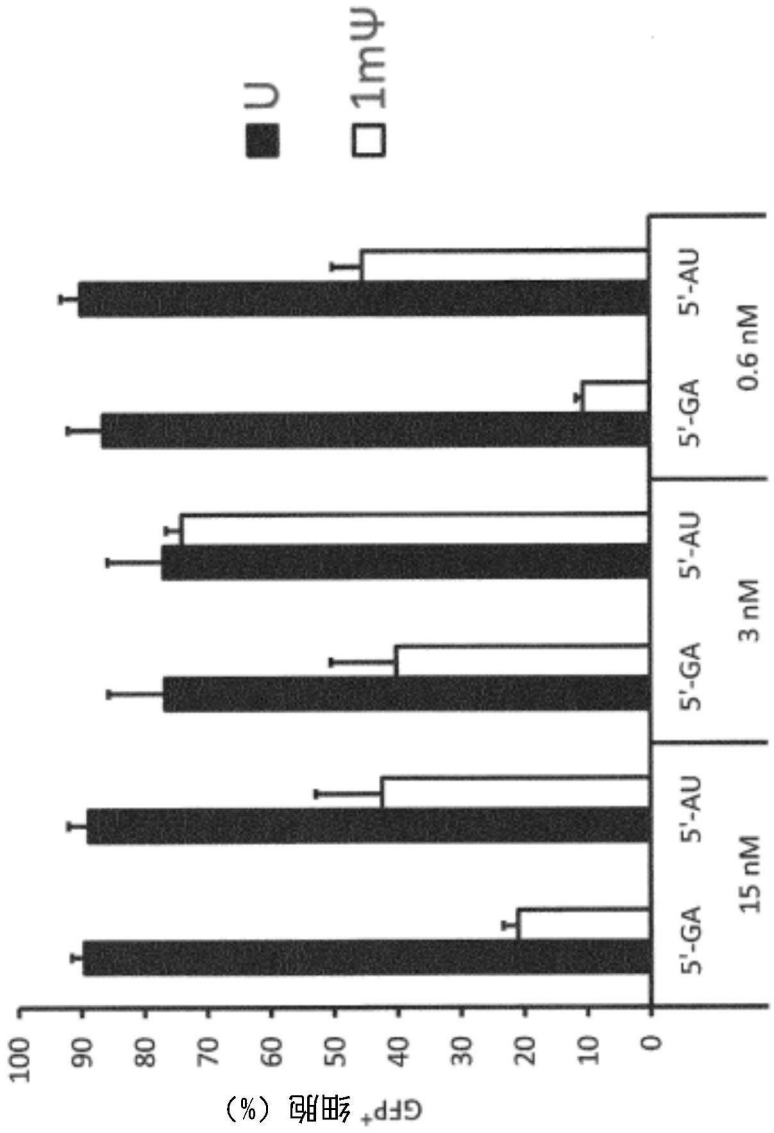


图2B

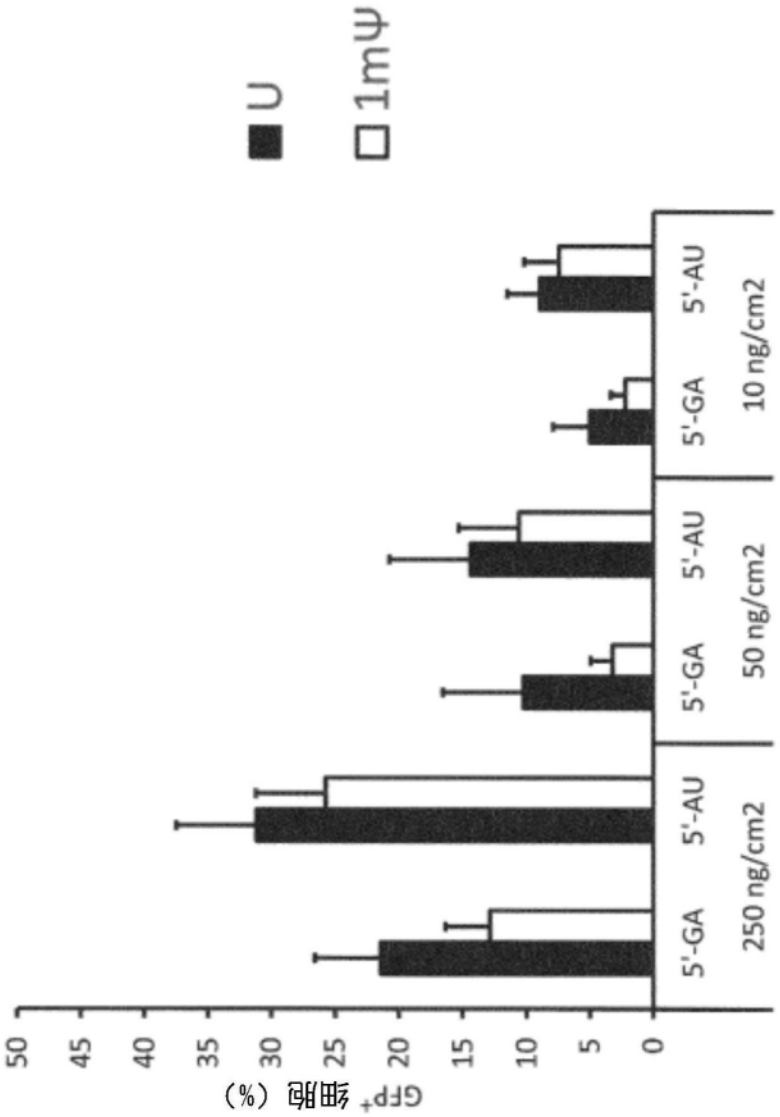

U
1mΨ
GFP+ 细胞 (%)
50
45
40
35
30
25
20
15
10
5
0
5'-GA
5'-AU
250 ng/cm2
5'-GA
5'-AU
50 ng/cm2
5'-GA
5'-AU
10 ng/cm2

图2C

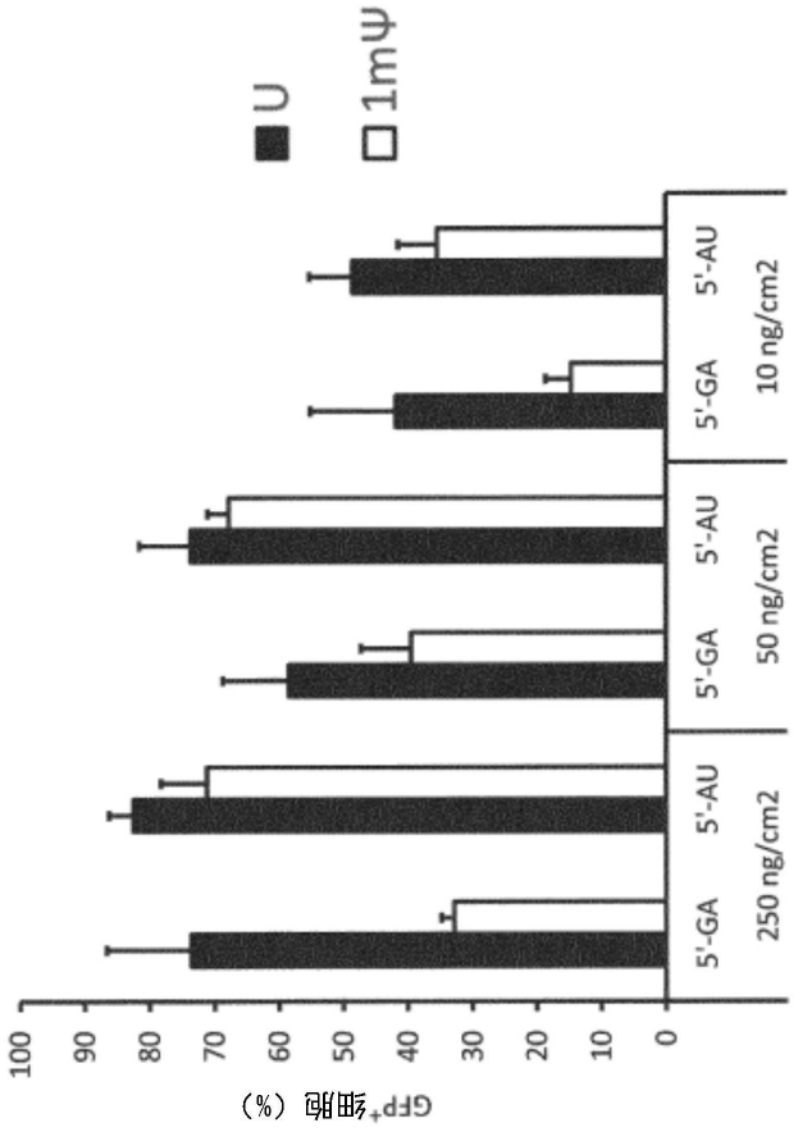

U
1mΨ
GFP+细胞 (%)
100
90
80
70
60
50
40
30
20
10
0
5'-GA
5'-AU
250 ng/cm2
5'-GA
5'-AU
50 ng/cm2
5'-GA
5'-AU
10 ng/cm2

图2D

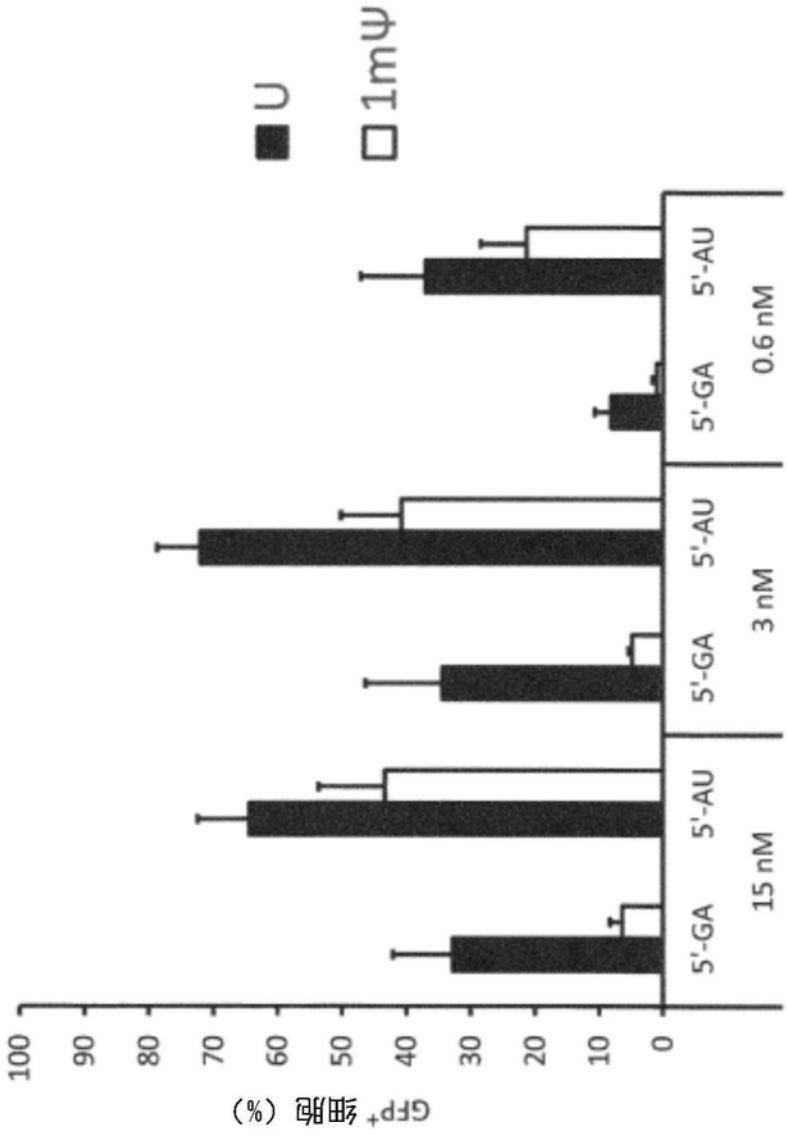


图2E

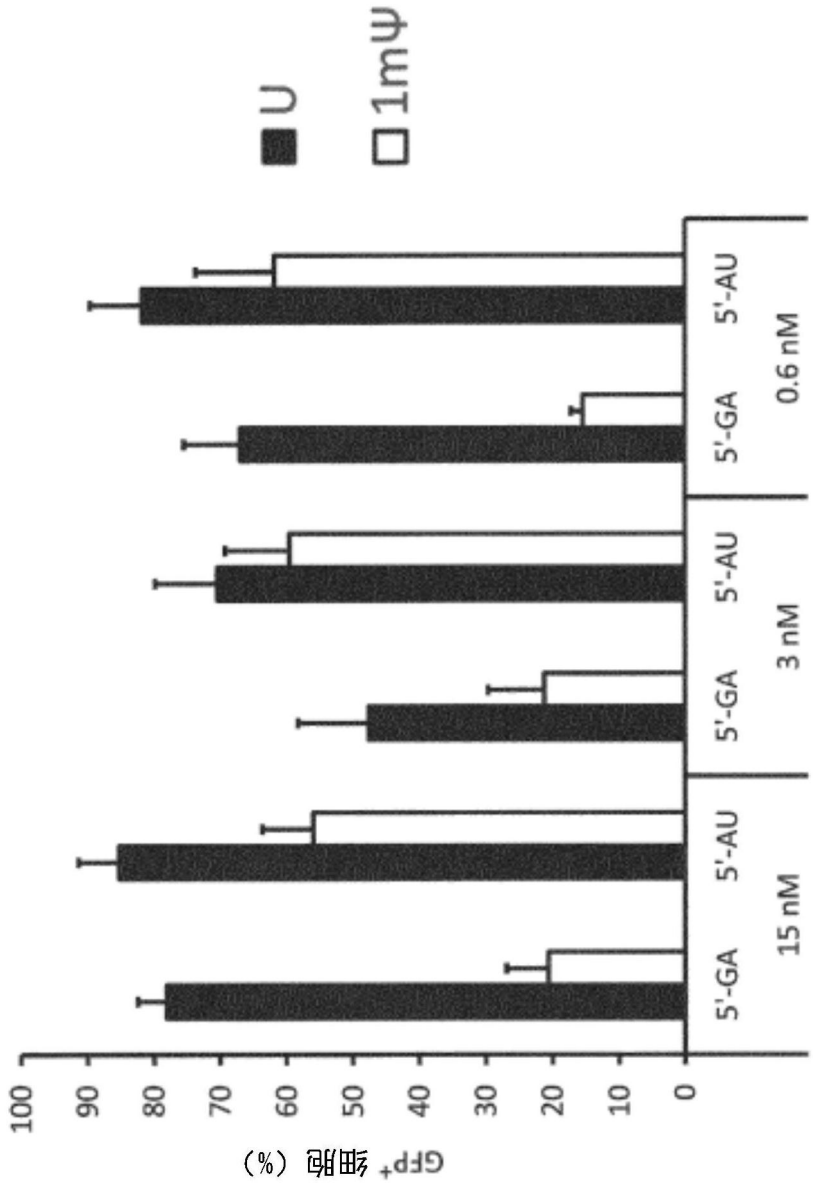


图2F

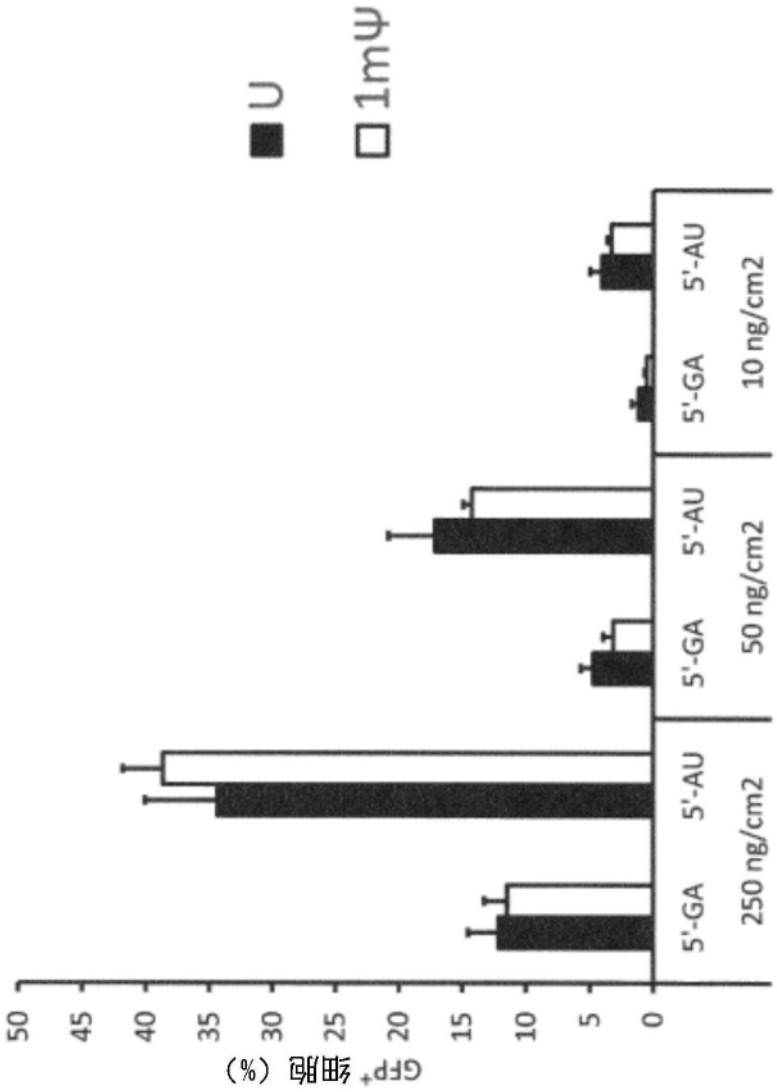

U
1mΨ
GFP+ 细胞 (%)
0
5
10
15
20
25
30
35
40
45
50
5'-GA
5'-AU
250 ng/cm2
5'-GA
5'-AU
50 ng/cm2
5'-GA
5'-AU
10 ng/cm2

图2G

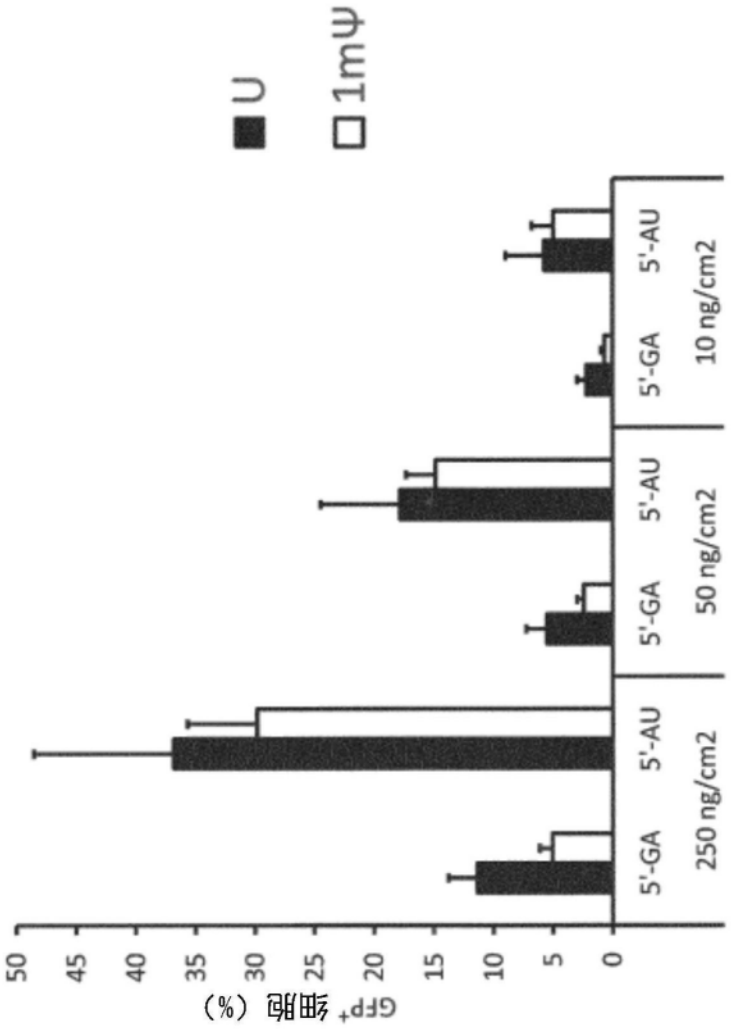

U
1mΨ
GFP+ 细胞 (%)
50
45
40
35
30
25
20
15
10
5
0
5'-GA
5'-AU
250 ng/cm2
5'-GA
5'-AU
50 ng/cm2
5'-GA
5'-AU
10 ng/cm2

图2H