



(12)发明专利申请

(10)申请公布号 CN 109154001 A

(43)申请公布日 2019.01.04

(21)申请号 201780029249.7

(74)专利代理机构 北京纪凯知识产权代理有限公司 11245

(22)申请日 2017.03.30

代理人 王永伟 丁秀云

(30)优先权数据

16163264.1 2016.03.31 EP

16177094.6 2016.06.30 EP

(51)Int.Cl.

C12N 15/67(2006.01)

(85)PCT国际申请进入国家阶段日

2018.11.12

(86)PCT国际申请的申请数据

PCT/EP2017/057592 2017.03.30

(87)PCT国际申请的公布数据

W02017/167910 EN 2017.10.05

(71)申请人 埃泽瑞斯公司

地址 德国普拉内格

(72)发明人 C·普兰克 C·鲁道夫

M·K·阿内嘉 L·威斯

权利要求书2页 说明书49页

序列表17页 附图20页

(54)发明名称

新型最小UTR序列

(57)摘要

描述了可转录成携带新型UTR序列的mRNA的DNA分子,所述UTR序列结合了极短并且同时允许含有其的RNA分子的高翻译效率的优势。此外,描述了包含这种DNA分子的载体和包含这种载体的宿主细胞。而且,描述了含有这种UTR的相应RNA分子。此外,描述了包含所述RNA分子和任选地药学上可接受的载体的药物组合物以及所述UTR用于将RNA分子的编码区翻译成所述编码区编码的多肽或蛋白质的应用。

1. DNA分子,其可转录成mRNA,所述DNA分子包含具有下列元件的一条链:

(a) 编码区,在其5'端包括起始密码子,所述编码区编码多肽;和

(b) 直接在所述编码序列上游的序列,其选自:

(b1) R₁-CGCCACC (SEQ ID NO:1);

或这样的序列,其中在所述序列中,SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代;和

(b2) R₁-CNGCCACC (SEQ ID NO:2),其中SEQ ID NO:2的位置2处的核苷酸N是选自T、G、C或A的核苷酸;

或这样的序列,其中在所述序列中,SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代,

其中R₁是启动子,其被DNA依赖性RNA聚合酶识别;

或包含互补链,

其中被DNA依赖性RNA聚合酶识别的所述启动子选自:

(i) 被T7DNA依赖性RNA聚合酶识别的TAATACGACTCACTATAGGGAGA (SEQ ID NO:3);

(ii) 被T3DNA依赖性RNA聚合酶识别的AATTAACCCTCACTAAAGGGAGA (SEQ ID NO:4);

(iii) 被SP6DNA依赖性RNA聚合酶识别的ATTTAGGTGACACTATAGAAG (SEQ ID NO:5);和

(iv) 被K11DNA依赖性RNA聚合酶识别的AATTAGGGCACACTATAGGGA (SEQ ID NO:6)。

2. 根据权利要求1所述的DNA分子,其中所述SEQ ID NO:2的位置2处的所述核苷酸N是选自T、G或C的核苷酸并且其中核苷酸N不是A。

3. 根据权利要求2所述的DNA分子,其中SEQ ID NO:2的位置2处的所述核苷酸N是T。

4. 载体,其包含权利要求3所述的DNA分子。

5. 宿主细胞,其包含权利要求4所述的载体。

6. 组合物,其包含:

根据权利要求1到3中任一项所述的DNA分子、根据权利要求4所述的载体或根据权利要求5所述的宿主细胞。

7. RNA分子,其包含

(a) 编码区,在其5'端包括起始密码子,所述编码区编码多肽;和

(b) 直接在所述编码序列上游的UTR,其选自:

(b1) 序列R₂-CGCCACC (SEQ ID NO:1)的UTR,

或这样的序列,其中在所述UTR序列中,SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代;和

(b2) 序列R₂-CNGCCACC (SEQ ID NO:2)的UTR,其中SEQ ID NO:2的位置2处的核苷酸N是选自U、G、C或A的核苷酸,或这样的序列,其中在所述UTR序列中,SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代,

其中R₂是这样的RNA序列,其对应于以其中DNA依赖性RNA聚合酶起始RNA合成的核苷酸开始的启动子区域的部分,

其中R₂选自:

(i) GGGAGA (SEQ ID NO:7);

(ii) GGGAGA (SEQ ID NO:8) ;

(iii) GAAG (SEQ ID NO:9) ;和

(iv) GGGA (SEQ ID NO:10) ;和

其中所述RNA分子在所述3'端包含聚腺苷酸尾。

8. 根据权利要求7所述的RNA分子,其中所述SEQ ID NO:2的位置2处的所述核苷酸N是选自U、G或C的核苷酸并且其中核苷酸N不是A。

9. 根据权利要求8所述的RNA分子,其中SEQ ID NO:2的位置2处的所述核苷酸N是U。

10. 根据权利要求7到9中任一项所述的RNA分子,其中所述聚腺苷酸尾的长度为至少120个核苷酸。

11. 核酸分子,其编码权利要求7到10中任一项所述的RNA分子。

12. 载体,其包含权利要求11所述的核酸分子。

13. 宿主细胞,其包含权利要求12所述的载体。

14. 药物组合物,其包含根据权利要求7到16中任一项所述的RNA分子、根据权利要求11所述的核酸分子、根据权利要求12所述的载体或根据权利要求13所述的宿主细胞和任选地药学上可接受的载体。

15. 权利要求14所述的药物组合物,其用于基于RNA的疗法。

16. 试剂盒,其包含根据权利要求1到3中任一项所述的DNA分子、根据权利要求7到10中任一项所述的RNA分子、根据权利要求11所述的核酸分子、根据权利要求4或12所述的载体或根据权利要求5或13所述的宿主细胞。

17. 如在权利要求7b中限定的UTR在用于将RNA分子的编码区翻译成所述编码区编码的多肽或蛋白质的应用。

新型最小UTR序列

[0001] 本发明涉及可转录成携带(具有, harbouring)新型UTR序列的mRNA的DNA分子, 所述UTR序列结合了作为极短的和同时允许含有其的RNA分子的高翻译效率的优势。此外, 本发明涉及包含这种DNA分子的载体和涉及包含这种载体的宿主细胞。而且, 本发明涉及含有这种UTR的相应RNA分子。此外, 本发明涉及包含所述RNA分子和任选地药学上可接受载体的药物组合物以及涉及用于将RNA分子的编码区翻译成所述编码区编码的多肽或蛋白质的所述UTR的应用。

[0002] 近年来, 信使RNA (mRNA) 作为新的药物实体已变得越来越相关。与基于DNA的基因疗法相反, mRNA不必被运输至细胞核而是在细胞质中被直接翻译成蛋白质(J Control Release, 2011, 150:238-247, 和Eur J Pharm Biopharm, 2009, 71:484-489)。这使得mRNA在避免潜在的插入突变发生——DNA基因药品的不太可能却又存在的风险——方面更加安全。因而, mRNA疗法逐渐变成广泛多种医疗适应症中具有前景的基因和蛋白质替代疗法的可选方案(J Control Release, 2011, 150:238-247; Eur J Pharm Biopharm, 2009, 71:484-489; Nat Biotech, 2011, 29:154-157, 和Nat Rev Genet, 2011, 12:861-874)。然而, 不得不克服常规mRNA的强免疫原性以及有限的稳定性从而进一步建立其临床适用性。对此, mRNA的稳定性和具体地mRNA的翻译速率是所设想的医疗应用的必要参数, 因为其决定了, 例如, mRNA药物的剂量和给药间隔。

[0003] 若干策略在增加稳定性和减少由施用于细胞或有机体的mRNA引发的免疫应答两者上证明是成功的。在这些中包括化学修饰的核苷酸; Current Opinion in Drug Discovery and Development, 2007, 10:523. Kormann等已表明用2-硫尿苷和5-甲基-胞苷仅替换25%的尿苷残基和胞苷残基足以增加mRNA的稳定性以及减少在体外外部施用的mRNA引发的先天免疫的激活(WO2012/0195936 A1; WO2007024708A2)。

[0004] 并且, 已经报道mRNA中的非翻译区(UTR)在调控mRNA的稳定性和mRNA的翻译两者中起到关键作用。已知UTR通过其与RNA结合蛋白的相互作用影响翻译的起始、延长、和终止, 以及mRNA的稳定性和胞内定位(Briefings in Bioinformatics, 2000, 1:236-249和Cold Spring Harbor Monograph Archive, 2007, 48:87-128)。取决于UTR内的特定动机, 其可增强或减少mRNA的周转(turnover) (Cell. Mol. Life Sci., 2012, 69:3613-3634; Nucleic Acids Research, 2005, 33:D141-D146; Science, 2005, 309:1514-1518和Current Protein & Peptide Science, 2012, 13:294-304)。最近, 公布了有关mRNA的半衰期和相应UTR序列的数据(Nucleic Acids Research, 2011, 39:556-566和Nucleic acids research, 37, e115)。

[0005] UTR是mRNA分子的区段, 其在mRNA的起始密码子上游和终止密码子下游, 即, 不被翻译的序列。这些区域与编码区一起转录, 并因此是外显的——由于其存在于成熟mRNA中。mRNA的起始密码子上游的UTR被称为5'UTR, 以及一经转录, 尤其携带与启动子的(剩余3')部分对应的序列以及所谓的Kozak序列。

[0006] Kozak共有序列(Kozak consensus sequence, Kozak consensus)或Kozak序列是已知存在于真核mRNA中并具有共有序列(gcc) gccRccAUGG的序列。Kozak共有序列在翻译过程的起始中起主要作用。该序列是以使其卓著的人的名字——Marilyn Kozak——命名。核

糖体在翻译起始位点——mRNA分子从其编码蛋白质——识别mRNA分子中的该序列。核糖体需要该序列或其可能的变化来起始翻译。该序列由记号(gcc)gccRccAUGG识别,其从各种来源(全部中的大约699种)概括了Kozak分析的数据,如下:小写字母表示其中碱基仍可以变化的位置处最常见的碱基;大写字母指示高度保守的碱基,即“AUGG”序列是恒定的或几乎——如果可能——不变的,“R”表示总会在此位置观察到嘌呤(腺嘌呤或鸟嘌呤)(其中Kozak声称的腺嘌呤将变得更频繁);和括号中的序列((gcc))的显著性(significance)不确定。

[0007] Kozak共有序列最初被定义为ACCAUGG是由于对前胰岛素原基因翻译的起始密码子(AUG,其中A在本上下文中限定了位置+1)周围的点突变的分析。699种脊椎动物mRNA的更详细的突变发生导致了共有序列GCCGCCACCAUGG,其中AUG起始密码子上游位置-3处的A还可以是G(Nucleic Acids Res.,1987,15(20):8125-8148)。对真核细胞中前胰岛素原和 α 球蛋白翻译的研究揭示了位置-3处的嘌呤(通常为A)对于有效的翻译起始是关键,并且如果该嘌呤缺失,则位置+4处的G是关键(J.Cell Biol.,1989,108:229-41)。由mRNA分子合成的蛋白质的量强烈地依赖于Kozak元件的序列:编码蛋白质的N端甲硫氨酸的AUG起始密码子是最重要的。对于强共有序列,位置+4(G)和-3(A或G)处的核苷酸必须都匹配共有序列。适当的(充足的,adequate)共有序列仅具有这两个位点中的一个,而弱共有序列既不满足位置+4处的要求也不满足位置-3处的要求。在-1和-2处的两个胞苷残基并非那么保守(Cell,1986,44(2):283-92),而位置-6处的G对于翻译的起始是重要的(Br.J.Haematol.,2004,124(2):224-31)。

[0008] 虽然现有技术中已经存在用于增加mRNA的稳定性、减少由施用于细胞或有机体的mRNA引发的免疫应答并提高表达效率(即,转录和/或翻译效率)的所述手段和方法,但是由于表达效率对于所设想的医疗应用是关键参数,因为其决定了例如mRNA药物的剂量和给药间隔,并最终决定终产物(即,编码的肽或蛋白质)的生物利用率,因此对于改进——特别是关于提高表达效率(即,转录和/或翻译效率)的另外的或替代的方法——仍然存在需要。同时,存在对于以下的恒定需要:进一步减少生产mRNA药物的成本、增加生产的mRNA分子的产量和增加用于实际转基因(即,用于编码期望多肽的编码区)的生产的mRNA分子中的可用空间。

[0009] 本申请通过提供权利要求中限定的实施方式解决了这个需要。

[0010] 具体地,本申请惊奇地发现将UTR序列的尺寸减小至“最小UTR”序列是可能的,进而减少用于生产mRNA药物的成本、增加生产的mRNA分子的产量和增加用于实际转基因(即,用于编码期望多肽的编码区)的生产的mRNA分子中的可用空间。而且,同时,相比常规UTR序列,该最小UTR序列惊奇地保持或甚至提高了表达速率,同时已发现在该最小UTR序列中的修饰甚至增加mRNA分子的表达速率。

[0011] 此发现导致提供权利要求中表征的实施方式,具体地导致提供允许生成携带这种“最小UTR”序列的RNA分子的DNA分子以及提供相应RNA分子。

[0012] 在第一方面,在DNA水平上描述相应分子,同时在下面进一步在第二方面中在RNA水平上描述相应分子。

[0013] 因此,在第一方面,本发明涉及DNA分子,其可转录成mRNA,其中所述DNA分子包含具有下列元件的一条链:

[0014] (a) 编码区,包括其5'端处的起始密码子,编码多肽;和

[0015] (b) 直接在所述编码序列上游的选自以下的序列:

[0016] (b1) R₁-CGCCACC (SEQ ID NO:1);

[0017] 或这样的序列,其中所述序列中在SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代;和

[0018] (b2) R₁-CNGCCACC (SEQ ID NO:2),其中SEQ ID NO:2的位置2处的核苷酸N是选自T、G、C或A的核苷酸;

[0019] 或这样的序列,其中在所述序列中SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代,

[0020] 其中R₁是启动子,其被DNA依赖性RNA聚合酶识别;

[0021] 或包含互补链。

[0022] DNA序列如果其序列与翻译成蛋白质的信使RNA拷贝的序列相同则被称为“有义”。相对的序列——互补链被称为“反义”序列。本发明的DNA分子在上文(a)和(b)中通过参考有义链被定义,而相应互补的反义链可容易地被技术人员根据碱基配对原则确定。

[0023] 本发明的DNA分子是可转录成mRNA分子的DNA分子。转录是基因表达的第一步,其中DNA分子的特定区段通过酶(RNA聚合酶)被拷贝成mRNA分子。在转录期间,RNA聚合酶阅读DNA序列,其产生互补、反向平行(anti-parallel)的RNA链,被称为初级转录物。

[0024] 两条DNA链中仅一条充当用于转录的模板。在转录期间DNA依赖性RNA聚合酶从3'端到5'端(3'→5')阅读DNA的反义链。互补RNA以相反方向——以5'→3'方向——产生,匹配有义链的序列,除了将胸腺嘧啶转变为尿嘧啶外。此方向性是因为RNA聚合酶只能将核苷酸添加到生长中的mRNA链的3'端。DNA的非模板有义链被称为编码链,因为其序列与新产生的RNA转录物相同(除了胸腺嘧啶取代为尿嘧啶)。当表示DNA序列时,这是由常规并且在本发明的上下文中使用的链。

[0025] 本发明的DNA分子可以是双链或单链或部分双链和部分单链。

[0026] 本发明的DNA分子包含两种主要组件(模块,modules)(也称为“项目”),即,(a)编码区,编码多肽并且在其5'端包括起始密码子,和(b)上文(b1)或(b2)中限定的直接在所述编码区上游的序列。这种DNA分子,在转录时,导致具有赋予上述优势的极短的UTR序列的mRNA。

[0027] 另外,本发明的DNA分子优选地包含这样的序列,其在转录成mRNA时导致UTR在编码区的下游。因此,本发明的DNA分子优选地携带这样的编码区以及序列,其一经转录,就导致生成的mRNA分子中的(5'和3')非翻译区(UTR)。

[0028] 根据本发明使用的术语“在其5'端包括起始密码子的编码区”涉及由密码子组成的DNA序列,其被DNA依赖性RNA聚合酶转录成mRNA分子,其中相应mRNA分子可根据“遗传密码”提供的信息被核糖体译码并翻译成蛋白质。编码区通常以在其5'端的起始密码子开始并以终止密码子结束。总体上,起始密码子是ATG三联体(对应于RNA水平上的AUG三联体),并且终止密码子是TAA、TAG或TGA(对应于RNA水平上的UAA、UAG、或UGA)。除了编码蛋白质外,编码区的部分可作为外显子剪接增强子或外显子剪接沉默子充当前mRNA中的调控序列。根据本发明使用的编码多肽或蛋白质的基因编码区还称为编码序列或CDS(来自编码DNA序列),并且是基因的DNA或RNA的部分,由外显子组成,编码多肽或蛋白质。mRNA中的编

码区的侧翼(flanked)为5'非翻译区(5'UTR)和3'非翻译区(3'UTR),其也是外显子的部分。而且,mRNA分子可进一步包含所谓的5'帽和聚腺苷酸尾。5'帽、5'UTR、3'UTR和聚腺苷酸尾是mRNA分子的不翻译成蛋白质的区域。

[0029] 根据本发明使用的术语“非翻译区”或“UTR”涉及mRNA的在起始密码子上游和在终止密码子下游的不被翻译的区段,并因此分别被称作5端(five prime)非翻译区(5'UTR)和3端非翻译区(3'UTR)。这些区域与编码区一起转录并因此是外显的——因为它们出现在成熟mRNA中。

[0030] 如本发明中使用的,3'非翻译区(3'-UTR)涉及信使RNA(mRNA)的直接(immediately)跟随翻译终止密码子的区段。3'UTR可在3'非翻译区内包含调控区域,其已知影响mRNA的多聚腺苷化作用和稳定性。许多3'-UTR还含有富含AU的元件(ARE)。并且,3'-UTR可优选地含有序列AAUAAA,其指导将被称为聚腺苷酸尾的数百个腺嘌呤残基添加到mRNA转录物的末端。

[0031] 5'非翻译区(5'UTR)(还称为前导序列或前导RNA)是mRNA的直接在起始密码子上游的区域。5'UTR在转录起始位点处开始并以编码区的起始密码子(通常为mRNA中的AUG)前的一个核苷酸(nt)结束。在真核生物中,5'UTR的长度通常为100至数千个核苷酸长但有时在真核生物中也存在较短的UTR。

[0032] 在本发明中,启动子和编码区之间的序列(如上文(b1)或(b2)中定义的)是极短的,并且一经转录就导致具有很短的“最小”UTR序列的mRNA分子。

[0033] DNA分子中的一个组件,即,“在其5'端包括起始密码子的编码多肽的编码区”(组件(a))不被具体地限定,并可以是任何期望的将在给定细胞中表达的编码区。因此,该组件可以是编码期望多肽(即,期望的最终产物)的编码区。本发明不限于“在其5'端包括起始密码子的编码多肽的编码区”,因为编码区的本质取决于将在细胞中产生的期望产物。这种编码区还可以是不同于已知天然序列并含有突变(即点突变、插入突变、缺失和其组合)的核苷酸序列。而且,这种编码区可部分地或完全地是从将用作组件(a)的天然序列衍生的密码子优化的序列。密码子优化是一种通过提高衍生自感兴趣基因的mRNA的翻译效率来最大化蛋白质表达的技术。已知天然基因不随机利用可用的密码子,但显示对于相同氨基酸的特定密码子的某种偏好。因此,由于遗传密码的简并性——一个氨基酸可由若干密码子编码——将感兴趣基因的核苷酸序列转变成相同种属或另一种属的一组优选的密码子。

[0034] 如上所述,组件(a)不被具体地限定并且可以是任何将在给定细胞中表达的期望编码区。因此,在本发明的上下文中,应当理解“编码区”是指任何多脱氧核糖核苷酸(polydesoxyribonucleotide)分子,如果被导入细胞,则可转录成能够翻译成多肽/蛋白质或其片段的mRNA分子。术语“多肽”和“蛋白质”在此包括任意种类的氨基酸序列,即,两个或更多个氨基酸的链,其通过肽键各自连接且还包括肽和融合蛋白。

[0035] 在优选实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”含有脱氧核糖核苷酸(desoxyribonucleotide)序列,编码多肽/蛋白质或其片段——其功能在细胞中或在细胞附近是需要的或是有益的,例如,蛋白质,其缺失或缺陷型引发疾病或病、其提供可减轻或防止疾病或病,或这样的蛋白质,其可在细胞中或在其附近促进对身体有益的过程。编码区可含有完全蛋白质或其功能变体的序列。此外,编码区的脱氧核糖核苷酸序列可编码充当因子、诱导物、调控物、刺激物或酶、或其功能片段的蛋白质,其中这种蛋白质是这

样的蛋白质:其功能是必需的以便治疗病症,具体地代谢紊乱,或以便在体内起始过程,例如新血管、新组织等的形成。在此,应当理解功能性变体意指其在细胞中可承担蛋白质的功能的片段,所述蛋白质在细胞中的功能是需要或或其缺失或缺陷形式是致病的。

[0036] 在优选实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”编码具有治疗作用或预防作用的治疗活性或药理学活性的多肽或蛋白质。因此,可转录成mRNA、包含所述“在其5'端包括起始密码子的编码多肽的编码区”的本发明DNA分子可用于核酸疗法和相关应用中。在本上下文中,根据本发明,本发明的DNA分子转录和翻译成mRNA和进一步成多肽或蛋白质可意图补偿或补充内源基因表达,特别是在其中内源基因有缺陷或沉默,导致基因不表达产物、基因表达的产物不足或缺陷或功能异常的情况下,例如具有多种代谢和遗传性疾病如举例而言(to name a few)囊性纤维化、血友病或肌肉萎缩症的情况。本发明的DNA分子转录和翻译成mRNA和进一步成多肽或蛋白质还可意图使表达的产物与任何内源细胞过程例如基因表达的调控、信号传导和其它细胞过程相互作用或干扰任何内源细胞过程例如基因表达的调控、信号传导和其它细胞过程。本发明的DNA分子转录和翻译成mRNA和进一步成多肽或蛋白质还可意图在其中转染或转导的细胞驻留(resides)或被驻留的有机体的情况下导致免疫应答。实例是对抗原呈递细胞例如树突细胞遗传修饰以便使它们呈递抗原用于接种目的。另一实例是本发明的DNA分子转录和翻译成mRNA和进一步成多肽或蛋白质,其中所述编码区编码细胞因子。这可能例如在肿瘤中是期望的以便引起肿瘤特异性免疫应答。而且,本发明的DNA分子转录和翻译成mRNA和进一步成多肽或蛋白质还可意图在体内或先体外后体内(离体,ex vivo)产生用于细胞疗法的瞬时基因修饰细胞,例如用于再生医疗的修饰的T-细胞或前体或干细胞或其它细胞。

[0037] 在其它优选实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”可编码这样的蛋白质,其与生长过程和血管发生有关,其在例如受控再生中是必需的,并且随后可通过导入根据本发明的RNA分子而特异地形成。这可在例如生长过程中或对于治疗骨缺陷、组织缺陷以及在植入和移植的情况中是有用的。

[0038] 如上所述,本发明的包含“在其5'端包括起始密码子的编码多肽的编码区”的DNA分子和具体地相应转录的RNA分子可适当地用于任何情况,其中这样的多肽或蛋白质将被提供给身体,所述多肽或蛋白质会天然地存在于身体中但是由于基因缺陷或疾病而不存在或以缺陷形式或过少量存在。已知蛋白质和编码其的基因,其缺乏或缺陷与疾病相关。可根据本发明使用编码完整多肽或蛋白质的编码区的各自的完整版。

[0039] 单一基因突变引起的多种遗传病症是已知的并且是用于mRNA治疗方案的候选。单基因突变引起的病症,如囊性纤维化、血友病和多种其它病症对于某种性状将在子孙中出现的可能性而言可能是显性的或隐性的。尽管显性的等位基因在仅具有一个拷贝等位基因的个体中显示表型,但对于隐性的等位基因,个体必须具有两个拷贝——来自每一个亲本的一个以变得明显(manifest)。相比之下,由两个或更多个基因引起的多基因病和各自疾病的显示经常是流畅的并且与环境因素相关。多基因病的实例是高血压、升高的胆固醇水平、癌症、神经变性疾病、精神病和其它。并且在这些情况下,代表这些基因中的一种或多种的治疗性mRNA对那些患者可能是有益的。而且,遗传病不是必然从亲本基因向下传递,而是也可以由新突变引起。并且在这些情况下,代表正确基因序列的治疗性mRNA对患者可能是有益的。

[0040] 目前人类基因和遗传病连同其各自基因和其表型描述的22,993个条目的线上目录可在ONIM(Online Mendelian Inheritance in Man)网页(<http://onim.org>)上获得;各自的序列都可获自Uniprot数据库(<http://www.uniprot.org>)。作为非限制性实例,下表1列出了一些先天性疾病,及相应基因(一个或多个)。由于细胞信号途径的高度相互作用,某些基因的突变引起多种致病症状,其中仅仅特征性症状列于表1中。

[0041] 在本发明的一些实施方式中,治疗性蛋白质选自表1中列出的细胞蛋白质。因此,本发明的DNA分子可编码治疗性细胞蛋白质,其中编码的治疗性蛋白质是表1中列出的蛋白质或其同系物。

[0042] 在本发明的另外的实施方式中,治疗性蛋白质选自表1中列出的分泌蛋白。因此,本发明的DNA分子可编码治疗性融合蛋白,其中编码的治疗性蛋白质或其同系物是表1中列出的蛋白质或其同系物,而第二蛋白质是允许分泌治疗性蛋白质的信号肽。信号肽是短的、典型地5-30个氨基酸长的氨基酸序列,存在于所述治疗性蛋白质的N端并通过某种细胞器(即内质网、高尔基体或内体)将融合蛋白引向细胞的分泌途径。因此,这种融合蛋白从细胞或从细胞器分泌或在细胞区室或在细胞表面处插入到细胞膜中(例如,多次跨越的跨膜蛋白(multi-spanning trans-membrane))。

[0043] 因此,在本发明的优选实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”(组件(a))可编码,但不限于,以下引起疾病、易于患病或预防疾病的基因。这种可治疗(或预防)的病症的非限制性实例包括其中所述多肽、蛋白质或肽选自下表1中概括的那些。

[0044] 在一些实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”可转录和翻译成部分长度或全长的蛋白质,其包含的细胞活性水平等于或大于天然蛋白质的细胞活性水平。在一些实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”编码具有治疗作用或预防作用的治疗活性或药学活性的多肽、蛋白质或肽,其中所述多肽、蛋白质或肽选自下表1中概括的那些。“在其5'端包括起始密码子的编码多肽的编码区”可用于表达部分长度或全长的蛋白质,其具有的细胞活性水平等于或小于天然蛋白质的细胞活性水平。这可允许RNA分子的施用可被指示的疾病的治疗。

[0045] 表1:人类基因和遗传病的非限制性实例

[0046]

疾病	病理	基因, 遗传性
血液疾病		
范康尼贫血 (Fanconi Anemia)	贫血和嗜中性白细胞减少症, DNA 修复机制受影响的证据	FANCA, 常染色体隐性
血友病-A	异常出血	凝血因子 VIII, X-染色体隐性
血友病-B	异常出血	凝血因子 IX, X-染色体隐性
遗传性球形红细胞症(各种类型)	球形红细胞(球形红细胞, spherocytes)	锚蛋白 (ANK1)
阵发性夜间血红蛋白尿	贫血和尿中出现血液	PIG-A, X-染色体
迟发性皮肤卟啉症	血红素的过量产生, 铁超负荷	尿卟啉原脱羧酶 (UROD), 常染色体隐性
严重联合免疫缺陷病 (SCID)	由于损伤的 DNA 合成 体液免疫和细胞免疫中的 严重免疫缺陷	腺苷脱氨酶, 常染色体隐性, IL-2R- γ 、JAK3、(IL-7R- α 、 RAG1/2、Artemis、CD3 δ 、 CD3 ϵ
镰状细胞贫血症	异常血红蛋白 (HbS)	β -血红蛋白 (HB), 常染色体 隐性
地中海贫血 (α -和 β 形式)	α -或 β 血红蛋白的缺乏 导致贫血症	HBA1 和/或 HBA2 的缺失,
维勒布兰德病 (Von Willebrand disease) (已知三种类型, III 型最严重)	异常出血, 类似于血友病 A 和 B 的出血	常染色体显性和隐性的形式

[0047]

癌症		
恶性黑色素瘤	P16 突变导致成纤维细胞失控增生	细胞周期蛋白依赖性激酶抑制剂 2 (Cyclie dependant kinase inhibitor 2) (CDKN2)
神经纤维瘤 (2 种类型)	听觉神经上的良性肿瘤导致耳聋	NF1、NF2, 常染色体显性
耳聋 (耳)		
耳聋	听力损失	耳聋-1A (DFNB1), 常染色体隐性
潘德雷德综合征 (Pendred syndrome)	听力损失	潘蛋白 (Pendrin) (PDS), 常染色体隐性
心脏		
运动失调性毛细血管扩张症	DNA 损伤修复受干扰,	ATM,
动脉粥样硬化	血液胆固醇增加	apoE,
LQT 综合征 (长 QT)	钾通道缺陷	LQT1 和其它基因
希-林二氏综合征 (Von-Hippel Lindau Syndrome)	血管的异常生长, 可导致癌症	VHL, 常染色体显性
William' s Beuren 综合征	弹性蛋白的缺失导致血管缺陷, 主动脉瓣上狭窄	弹性蛋白和 LIM 激酶基因的缺失
代谢紊乱与糖原贮积病		
肾上腺脑白质营养不良 (Adrenoleukodystrophy)	受干扰的脂肪酸运输和代谢	ABCD1, X-染色体
尿黑酸尿 (Alkaptonuria)	氮代谢缺陷, 当暴露于氧时尿变黑	尿黑酸氧化酶 (Homogentisic oxidase), 常染色体隐性
I 型糖尿病	受干扰的胰岛素生成	IDDM1、IDDM2、GCK、...
半乳糖血症	半乳糖代谢的失调	半乳糖-1-磷酸尿苷酰转移酶基因 (GALT), 常染色体隐性
高歇病 (Gauche disease)	脂肪代谢紊乱	葡糖脑苷脂酶
葡萄糖半乳糖苷酶吸收不良	受干扰的葡萄糖和半乳糖从肠腔运输出去导致腹泻	SGLT1, 常染色体隐性
I 型糖原贮积病, 冯·基尔克氏病 (Von-Gierke' s disease)	肝脏和肾脏中葡萄糖的积累	葡萄糖-6-磷酸酶, 常染色体隐性
II 型糖原贮积病, 蓬珀病 (Pompe' s disease)	肝脏、心脏、骨骼肌中糖原的积累, 心肥大	α -1-葡萄糖苷酶, 常染色体隐性
III 型糖原贮积病, 柯里病 (Cori' s disease)	肝脏、心脏、骨骼肌中糖原的积累, 肝肿大	脱支酶, 常染色体隐性
V 型糖原贮积病, 麦卡德尔疾病	不能利用肌肉细胞中的糖	肌磷酸化酶, 常染色体隐性

[0048]

(McArdle' s disease)	原	
葡萄糖-6-磷酸脱氢酶	不能保持谷胱甘肽导致溶血性贫血	G6PD, X-染色体隐性
遗传性血色素沉着症(4种类型)	由于肠中过量铁吸收导致身体(尤其是肝脏)中铁过量	血色素沉着症(HFE)
高胱氨酸尿	氮代谢缺陷	胱硫醚(Cystathione)合成酶缺陷, 常染色体隐性
累-奈二氏综合征(Lesh Nyhan Syndrome)	尿酸积累导致痛风、尿结石和肌肉损失	HPRT1, X-染色体
枫糖尿症	氨基酸代谢缺陷导致 α -酮酸积累, 并在不治疗的情况下在最初数月内死亡	支链 α 脱氢酶(BCKDH)
门克斯综合征(Menkes' Syndrome)	吸收铜的能力降低, 导致在不治疗的情况下在婴儿期死亡	ATP7A, X-染色体隐性
肥胖	体重升高	多基因的, 升高的瘦蛋白水平可能起作用
苯丙酮尿	不能将苯丙氨酸分解成酪氨酸, 导致智力迟钝	苯丙氨酸羟化酶(PAH), 常染色体隐性
丹吉尔病(Tangier disease)	血浆高密度脂蛋白水平降低	ATP结合盒1基因(ABCA1)
泽耳韦格综合征(Zellweger Syndrome)(导致婴儿死亡)	血液中铁和铜的水平高	PXR1(过氧化物酶体表面上的受体)
威尔逊病	铜在脑和肝脏中积累	ATP7B(P-型ATP酶), 常染色体隐性
肌骨骼系统		
软骨发育不全	由于软骨细胞增殖缓慢导致具有大头的矮小身材	成纤维细胞生长因子受体3(FGF3R),
夏-马-图三氏综合征及其更严重形式的代-索二氏综合征	四肢中的肌肉退化	由不同基因突变引起的不同形式, 常染色体隐性和X-染色体
柯克凯内综合征(Cockayne Syndrome)(2种类型)	早熟衰老症和矮小身材, 缺失“运行中(on the fly)”DNA修复	8型切除修复交叉互补蛋白(ERCC8)
软骨外胚层发育不良	骨畸形和多指(趾)	EVC, 常染色体隐性
畸形发育不良(Diastrophic dysplasia)(DTD)	畸形的手, 硫酸盐运载体缺陷	DTDST基因
杜兴肌肉萎缩症	肌肉组织扩张伴随后续功能丧失	DMD, X-染色体隐性

[0049]

进行性骨化性纤维发育不良	异位骨形成	NOG, BMP, 常染色体显性
弗里德赖希共济失调 (Friedreich's ataxia)	心脏扩张和肌肉协调的进行性丧失	弗氏共济失调蛋白 (Frataxin), 常染色体隐性
磷酸酶过少症 (Hypophosphatasia)	生成影响矿化过程的异常版本的碱性磷酸酶	ALPL, 常染色体隐性
马方综合征	由于肌原纤维蛋白缺乏导致结缔组织病	肌原纤维蛋白 1 (FBN), 常染色体显性
肌强直性营养不良 (Myotonic dystrophy) (在成年早期发作)	骨骼肌细胞中的蛋白质激酶缺陷	肌强直性营养不良蛋白激酶 (DMPK), 常染色体显性
成骨不全 (各种类型)	I 型胶原蛋白形成缺陷导致出生后多发性骨折	COL1A1、COL1A2
普 - 韦二氏综合征 (Prader-Willi Syndrome)	肌紧张 (muscle tone) 降低和智力迟钝	由于染色体 15 上的缺失导致 SNRPN (小核糖核蛋白 N) 缺失
神经元与脑		
阿尔茨海默病	淀粉状蛋白生成增加, 不能进行性记忆事实	多基因的, PS1、PS2、...
肌萎缩性侧索硬化 (ALS) (多种形式)	运动神经元细胞的进行性退化 (清除超氧自由基缺陷)	超氧化物歧化酶 1 (SOD1), 涉及多种基因
安格耳曼 (天使人, Angelman) 综合征	智力迟钝伴有不适当的笑	染色体 15 上的基因组印迹
丙酮酸脱氢酶	如果不治疗, 就会神经性缺陷	丙酮酸脱氢酶, 常染色体隐性
雷弗素姆病 (Refsum disease)	植烷酸的积累导致周围神经病	植烷酰 CoA (Phytanoyl-CoA) 羟化酶 (PHYH), 常染色体隐性
雷特综合征 (Rett's syndrome)	年龄在 6 到 18 个月之间智力迟钝伴有发育停滞	甲基 -CpG- 结合蛋白 -2 (MECP2), X-染色体显性
泰 - 萨二氏病 (Tay-Sachs disease) (多种形式的严重性)	GM2 神经节苷脂的分解受干扰导致神经损伤	HEXA (β -氨基己糖苷酶 A) (β -hexosaminidase A), 常染色体隐性
拉福拉病 (LaFora disease)	癫痫的攻击性形式	EPM2A, 常染色体隐性
特发性震颤 (可变形式)	震动失控	ETM1、ETM2, 常染色体显性
脆性 X 染色体综合征	FMR1 RNA 结合蛋白缺乏, 智力迟钝	由于 5' UTR 区域中 CGG 扩增导致 FMR1 基因不表达
亨廷顿病	进行性痴呆成年发作	HTT (亨廷顿蛋白), 常染色体显性
肠		
巴雷特综合征 (Bartter's)	肾病	肾氯通道 B 基因 (CLCNKB),

[0050]

syndrome) (3 种类型)		常染色体隐性
多囊肾 (2 类型)	肾病	PDK1、PDK2, 常染色体显性, 已知也是常染色体隐性 (ARPKD)
肺		
α_1 -抗胰蛋白酶	由于弹性蛋白酶释放失控导致缺陷肺泡	SERPINA1, 常染色体共显性
哮喘	气道的慢性炎症性疾病	多基因
囊性纤维化	由于 Cl ⁻ 离子运输缺陷导致粘液过度粘稠	CFTR (囊性纤维化跨膜传导调节蛋白), 常染色体隐性
表面活性剂代谢障碍 (各种类型)	新生儿体重正常, 但不能膨胀 (inflate)	ATP- 结合性盒型转运蛋白 (ABCA3)
原发性纤毛运动障碍 (Primary ciliary dyskinesia)	由于纤毛功能缺陷/缺失导致粘液过度粘稠	DNAI1、CCNO、CCDC40 等
溶酶体贮积病		
法布里病 (Fabry's disease)	除其它外, 神经酰胺三己糖苷酶积累导致皮肤损伤	α -半乳糖苷酶 A, X-染色体隐性
高歇病 I 型: 成人形式 (在治疗下寿命正常) II 型: 婴儿形式 (1 岁前死亡) III 型: 青少年形式 (童年早期发作, 没有 II 型严重)	葡萄糖脑苷脂 (神经节苷脂、鞘脂) 积累	葡萄糖脑苷脂酶, 常染色体隐性
亨特综合征 (Hunter's syndrome)	粘多糖积累	L- 艾杜糖硫酸 (iduronosulfat) 硫酸酯酶, X-染色体隐性
胡尔勒综合征 (Hurler's Syndrome) (10 岁前死亡)	粘多糖积累	α -L-艾杜糖苷酸酶, 常染色体隐性
尼-皮二氏病 (Niemann-Pick Disease) (三种不同形式 A、B、C)	胆固醇从溶酶体释放缺陷, 鞘磷脂积累	鞘磷脂酶, 常染色体隐性
泰-萨二氏病 (Tay-Sachs Disease) (4 岁前死亡)	神经细胞中 G _{M2} 神经节苷脂积累	己糖胺酶 A, 常染色体隐性
皮肤		
白化病	氮代谢缺陷	酪氨酸酶缺乏, 常染色体隐性
白化病, 眼皮肤的, II 型	黑色素生物合成减少	OCA2, 常染色体隐性
埃-丹二氏综合征	常见的膈疝, 视网膜脱落	胶原蛋白合成的多种缺陷

[0051]

(Ehlers-Danlos Syndrome)(各种类型)		
大疱性表皮松解症 (各种类型, 包括单纯性 EB、连接的 EB、营养不良的 EB 和金德勒综合征 (Kindler syndrome))	保持角化细胞结构稳定性或角化细胞粘附底层真皮的缺陷	大疱性表皮松解黄斑型 (EBM)、进行性大疱性表皮松解 3 型 (EBR3)、假接合性 (pseudo-junctional) 大疱性表皮松解 4 型 (EBR4)、桥粒斑蛋白 (DSP)、血小板亲和蛋白-1 (Plakophilin-1) (PKP1)、肌酸 (角蛋白, kreatin) (KRT5、KRT14)、网蛋白 (PLEC)、ITGA6、整合蛋白亚单位 (ITGB4)、层粘连蛋白亚单位 (LAMA3、LAMP3、LAMB3、LAMC2)、胶原蛋白 (COL17A1、COL7A1 (常染色体显性)、FERMT1, 常染色体隐性
哈尔脱努帕病 (Hartnup's disease)	胃肠道中色氨酸摄入缺陷, 光敏皮肤	SLC6A19, 常染色体隐性
遗传性出血性毛细血管扩张, 奥-韦-朗三氏综合征	皮肤和黏膜的毛细血管扩张	内皮糖蛋白 (Endoglin) (ENG), 常染色体显性
高胆固醇血症, 家族性	结合低密度脂蛋白的血清胆固醇升高, 在皮肤中积累和动脉硬化	低密度脂蛋白受体 (LDLR)、载脂蛋白 B (APOB), 常染色体显性
着色性干皮病	UV 暴露导致皮肤缺陷和色素瘤	DNA 修复缺陷, 常染色体隐性
男性型光秃	干扰皮肤中睾酮转变为二氢睾酮	5- α -还原酶
遗传性肝病		
氨基酸代谢紊乱	干扰分解氨基酸酪氨酸和苯丙氨酸的多步过程	FAH、TAT、HPD, 常染色体隐性
β -地中海贫血中间物 (intermedia)	缺少成熟红血细胞	HBB, 常染色体隐性
克-纳二氏综合征 (Crigler-Najjar Syndrome)	缺乏葡糖苷酸化, 其中胆红素不能溶解在水中	UGT1A1, 常染色体隐性
脂肪酸氧化障碍	缺乏长链脂肪酸和极长链脂肪酸加工导致昏睡和低血糖症	HADHA、ACADVL 常染色体隐性
果糖代谢紊乱	糖原异生受损导致低血糖	FBP1、ALDOB, 常染色体隐性
半乳糖血症	缺乏半乳糖的加工	GALT、GALK1、GALE, 常染色

[0052]

		体隐性
糖原贮积病	干扰葡萄糖6-磷酸和糖原分解导致糖原积累以及异常糖原分子导致细胞损伤	G6PC、SLC37A4、AGL、GBE1, 常染色体隐性
血红素生物合成紊乱	尿卟啉原脱羧酶的减少导致称为卟啉的化合物积累引起肝脏中的毒性水平	UROD 常染色体显性、ALAS2 伴 X (X-linked) 显性、ALAD 常染色体隐性
脂质代谢 (运输) 紊乱	缺少功能蛋白, 其阻止胆固醇和其它脂质运动, 导致它们在细胞中积累	NPC1、NPC2 常染色体隐性, LDLR, 常染色体显性
金属代谢紊乱	铁和铜的贮积和运输紊乱导致在组织和器官中积累	ATP7B、HAMP、HFE、HFE2, 常染色体隐性
有机酸紊乱 (酸尿 (Acidurias) / 酸血症 (Acidemias))	干扰若干蛋白质构件 (building blocks) (氨基酸)、某些脂质、和胆固醇的分解	BCKDHA、BCKDHB、和 DBT, PCCA 和 PCCB、MUT、MMAA、MMAB、MMADHC、MCEE、IVD、MCCC1 或 MCCC2, 常染色体隐性
1 型原发性尿草酸盐过多 (Primary Hyperoxaluria Type I)	干扰乙醛酸盐分解导致肾损伤	AGXT、GRHPR, 常染色体隐性
进行性家族性肝内胆汁阻塞	肝脏细胞中胆汁酸积累导致肝损伤	ATP8B1, 常染色体隐性
血小板活性紊乱 (Thrombocyte activity disorder)	酶活性缺乏干扰出血和凝血之间的常规平衡	ADAMTS13, 常染色体隐性
尿素循环紊乱	尿素循环紊乱, 其导致高氨血症的形成	OTC (伴 X 疾病)、CPS1、ASS1 和 SLC25A13、ASL, 常染色体隐性

[0053] 上表1显示了基因缺陷导致疾病的基因的实例,所述疾病可用由本发明的DNA分子转录的RNA分子治疗,其中DNA分子(和相应转录的RNA分子)包含“在其5'端包括起始密码子的编码多肽的编码区”——其编码上述公开的缺陷基因的完整版蛋白质或其功能片段。在特别优选实施方式中,可提及例如影响肺的遗传性疾病,如SPB(表面活性蛋白B)缺乏、ABCA3缺乏、囊性纤维化和 α 1-抗胰蛋白酶缺乏,或影响浆蛋白(例如先天性血色素沉着症(铁调素缺乏)、血栓形成性(thrombotic)血小板减少性紫癜(TPP、ADAMTS13缺乏)并导致凝血缺陷(例如血友病a和b)和补体(complement)缺陷(例如蛋白质C缺乏)的遗传性疾病,免疫缺陷诸如例如SCID(由不同基因例如:RAG1、RAG2、JAK3、IL7R、CD45、CD3 δ 、CD3 ϵ 中的突变所引起)或由于缺乏腺苷脱氨酶的缺陷例如(ADA-SCID)、脓性肉芽肿(septic granulomatosis)(例如由gp-91-phox基因、p47-phox基因、p67-phox基因或p33-phox基因的突变引起)和贮积病如高歇病、法伯尔疾病、克腊比病(Krabbe's disease)、MPS I、MPS II(亨特综合征)、MPS VI、II型糖原贮积病或粘多糖贮积病(mucopolysaccharidoses)。

[0054] 包含“在其5'端包括起始密码子的编码肽的编码区”的本发明可用于的其它疾病包括以下疾病:例如SMN1相关的脊髓性肌萎缩(SMA);肌萎缩性侧索硬化(ALS);GALT相关的半乳糖血症;囊性纤维化(CF);SLC3A1相关的疾病,包括胱氨酸尿;COL4A5相关的疾病,包括

Alport综合征;半乳糖脑苷脂酶缺陷;伴X肾上腺脑白质营养不良和肾上腺脊神经病;弗里德赖希共济失调;佩-梅二氏病;TSC1和TSC2相关的结节性硬化症;桑菲列普综合征B(MPS IIIB);CTNS相关的胱氨酸病;FMR1相关的疾病,包括脆性X染色体综合征、脆性X染色体相关的震动/共济失调综合征和脆性X染色体卵巢功能早衰综合征;普-韦二氏综合征;遗传性出血性毛细血管扩张(AT);C1型尼-皮二氏病;神经元蜡样质脂褐质沉积症相关的疾病,包括幼年神经元蜡样质脂褐质沉积症(JNCL)、幼年Batten病、Santavuori-Haltia病、詹-比二氏病、以及PTT-1和TPP1缺陷;EIF2B1、EIF2B2、EIF2B3、EIF2B4和EIF2B5相关的伴有中枢神经系统低髓鞘化/白质消失(central nervous system hypomyelination/vanishing white matter)的儿童共济失调;CACNA1A和CACNB4相关的2型发作性共济失调;MECP2相关的疾病,包括典型雷特综合征、MECP2相关的严重的新生儿脑病和PPM-X综合征;CDKL5相关的非典型雷特综合征;肯尼迪病(SBMA);伴有皮层下梗塞的刻缺蛋白-3(Notch-3)相关的脑常染色体显性动脉病和脑白质病(CADASIL);SCN1A和SCN1B相关的疾病发作;聚合酶G相关的疾病,包括Alpers-Huttenlocher综合征、POLG相关的感觉共济失调神经病、构音障碍、和眼肌瘫痪(ophthalmoparesis)、和伴有线粒体DNA缺失的常染色体显性和隐性的进行性外眼肌麻痹;伴X肾上腺发育不全;伴X血中丙球蛋白贫乏;法布里病;和威尔逊病。

[0055] 在所有这些疾病中,蛋白质(例如酶)有缺陷,其可通过使用本发明的DNA分子转录的RNA的治疗来治疗——其使得由缺陷基因或其功能片段编码的蛋白质可用(有效, available)。转录物替换疗法/酶替换疗法不影响基本的(潜在的, underlying)遗传缺失,但增加患者缺乏(deficient)的酶的浓度。作为实例,在蓬珀病中,转录物替换疗法/酶替换疗法替换了有缺陷的溶酶体酶——酸性 α -葡萄糖苷酶(GAA)。

[0056] 因此,根据本发明的组件(a)的可以由“在其5'端包括起始密码子的编码多肽的编码区”编码的蛋白质的非限制性实例是红细胞生成素(EPO)、生长激素(生长激素(somatotropin), hGH)、囊性纤维化跨膜传导调节物(CFTR)、生长因子如GM-SCF、G-CSF、MPS、蛋白质C、铁调素、ABCA3和表面活性蛋白B。可用根据本发明的RNA治疗的疾病的进一步实例是血友病A/B、法布里病、CGD、ADAMTS13、胡尔勒病、X染色体介导的A- γ -球蛋白血症、腺苷脱氨酶相关的免疫缺乏和新生儿中的呼吸窘迫综合征,其与SP-B有关联。特别优选地,根据本发明的DNA分子的“在其5'端包括起始密码子的编码多肽的编码区”含有表面活性蛋白B(SP-B)的序列或红细胞生成素的序列。可被根据本发明的DNA分子的“在其5'端包括起始密码子的编码多肽的编码区”编码的蛋白质的进一步实例是生长因子,如人生长激素hGH、BMP-2或血管生成因子。

[0057] 可选地,核酸可编码全长抗体或较小的抗体(例如,重链和轻链两者)从而赋予受试者免疫力。在另一实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”可编码功能性单克隆抗体或多克隆抗体,其可用于靶向和/或钝化生物靶标(例如,刺激性细胞因子如肿瘤坏死因子)。类似地,“在其5'端包括起始密码子的编码多肽的编码区”可编码例如用于治疗II型膜性增生性肾小球肾炎或急性溶血性尿毒综合征的功能性抗肾病因子抗体,或可选地可编码用于治疗VEGF介导的疾病例如癌症的抗血管内皮细胞生长因子(VEGF)抗体。

[0058] 组件(a),即“在其5'端包括起始密码子的编码多肽的编码区”,可以是编码可用于基因组编辑技术中的多肽或蛋白质的编码区。基因组编辑是一种利用核酸酶在有机体的基

基因组中将DNA插入、缺失或替换的遗传工程。这些核酸酶在基因组中期望的位置处产生位点特异性双链断裂(DSB)。所诱发的双链断裂被非同源末端接合或同源重组修复,导致在基因组中靶向突变,进而“编辑”基因组。本领域已知利用不同多肽或蛋白质的多种基因组编辑系统,即,例如,CRISPR-Cas系统、大范围核酸酶、锌指核酸酶(ZFN)和基于类转录激活因子效应物的核酸酶(TALEN)。Trends in Biotechnology,2013,31(7),397-405中回顾了用于基因组工程的方法

[0059] 因此,在优选实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”含有脱氧核糖核苷酸序列——其编码Cas(CRISPR相关蛋白质)蛋白质家族的多肽或蛋白质,优选地Cas9(CRISPR相关蛋白质9)。Cas蛋白质家族的蛋白质,优选地Cas9,可用于基于CRISPR/Cas9的方法和/或CRISPR/Cas9基因组编辑技术。Nat.Biotechnol.,2014,32(4):347-355中回顾了用于基因组编辑、调控和靶向的CRISPR-Cas系统。

[0060] 在另一优选实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”含有编码大范围核酸酶的脱氧核糖核苷酸序列。大范围核酸酶是脱氧核糖核酸内切酶,其与“常规的”脱氧核糖核酸内切酶相反,识别大的识别位点(例如,12到40个碱基对的双链DNA序列)。结果,在任意给定基因组中各自的位点仅少数次发生,优选地仅一次。大范围核酸酶因此被认为是最特异的天然存在的限制性酶,并因此是基因组编辑技术中合适的工具。

[0061] 在另一优选实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”含有编码锌指核酸酶(ZFN)的脱氧核糖核苷酸序列。ZFN是通过将锌指DNA结合结构域融合到DNA裂解结构域产生的人工限制性酶。锌指结构域可被改造(engineered)以靶向特异的期望DNA序列并且这使得锌指核酸酶能够靶向复合基因组内的独特序列。通过利用内源DNA修复机制,ZFN可用于精确改变高等有机体的基因组,并因此是基因组编辑技术中合适的工具。

[0062] 在另一优选实施方式中,“在其5'端包括起始密码子的编码多肽的编码区”含有编码类转录激活因子效应物核酸酶(TALEN)的脱氧核糖核苷酸序列。TALEN是可被改造以切割DNA的特异性序列的限制性酶。TALEN是融合蛋白,其中TAL效应物DNA结合结构域融合到核酸酶的DNA裂解结构域。类转录激活因子效应物(TALE)可被改造以在实践中结合任意期望DNA序列。因此,在结合核酸酶时,可在特异的期望位置处切割DNA。

[0063] 本发明的DNA分子包含直接位于编码序列上游、作为第二组件(b)的序列。

[0064] 更具体地,本发明的DNA分子包含直接在所述编码序列上游的组件(b),其中所述组件(b)是选自以下的序列:

[0065] (b1) R₁-CGCCACC (SEQ ID NO:1);

[0066] 或这样的序列,其中在所述序列中,SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代;和(b2) R₁-CNGCCACC (SEQ ID NO:2),其中SEQ ID NO:2的位置2处的核苷酸N是选自T、G、C或A的核苷酸;

[0067] 或这样的序列,其中在所述序列中,SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代,

[0068] 其中R₁是启动子,其被DNA依赖性RNA聚合酶识别。

[0069] 上文项目(b)中限定的序列不具体地限于上述特异性序列而且还可涉及序列(一个或多个),其相比于这种序列显示核苷酸(一个或多个)添加(一个或多个),其中另外的核

核苷酸(一个或多个)可添加在上述序列(一个或多个)中R₁的5'端。另外的核苷酸(一个或多个)包含多达0(无变化)、1、2、3、4、5、6、7、8、9或10个核苷酸,优选地多达20个核苷酸的多核苷酸链。更优选地,11、12、13、14、15、16、18、或19个核苷酸被添加在5'端。甚至更优选地多达30个核苷酸添加在5'端。

[0070] 由于启动子R₁上游的核苷酸添加不会改变本发明的UTR(一个或多个)的上述功能特性,核苷酸添加的长度还可为多达40、50、60、70、80、90、或甚至100个核苷酸或甚至更多,多达200、300、400或500个核苷酸。

[0071] 如上所述,双链DNA分子包含两条反向平行链,其中一条链被称为“有义”链,如果其序列与翻译成蛋白质的信使RNA拷贝的序列相同。相对、互补的链上的序列被称为“反义”序列。因此,本发明的DNA分子不仅涉及与包含具有上述元件(a)和(b)的一条链的mRNA对应的上述DNA分子,还涉及包含互补链(即可转录成mRNA的反义链)的DNA分子。本发明的DNA分子的这种互补链通过参考可容易地根据碱基配对原则确定的反义链进行限定。

[0072] 本发明的DNA分子在组件(b)中还包含被DNA依赖性RNA聚合酶识别的启动子R₁。优选地,所述启动子R₁直接连接到项目(b1)或(b2)中限定的剩余序列,即,不出现任何插入(intervening)的核苷酸。

[0073] 不具体限定被DNA依赖性RNA聚合酶识别的启动子R₁的性质。可使用任何启动子(及其变体),只要相应DNA依赖性RNA聚合酶能够识别各自的序列。本领域已知多种RNA聚合酶(还称为DNA依赖性RNA聚合酶并经常缩写为RNAP或RNAPol)。这些酶能够产生初级转录物RNA。如上文概括的,DNA依赖性RNA聚合酶在被称为转录的过程中能够利用DNA作为模板合成RNA链。DNA依赖性RNA聚合酶在称为启动子的特异性DNA序列处起始转录。然后其产生与模板DNA链互补的RNA链。将核苷酸添加到RNA链的过程称为延伸(延长,elongation)。因此,在本发明的上下文中,术语“识别”优选地不仅意为DNA依赖性RNA聚合酶能够特异性探测/结合其相应启动子序列R₁。该术语还指代DNA依赖性RNA聚合酶能够起始转录然后在延伸期间产生RNA分子的能力。

[0074] 技术人员可通过本领域已知的方法确定给定的DNA依赖性RNA聚合酶是否能够识别各自的启动子。而且,通过利用用于评价蛋白质/DNA相互作用的熟知方法,给定的DNA依赖性RNA聚合酶的相应(未知)启动子序列R₁可以被识别,反之亦然。

[0075] 因此,DNA依赖性RNA聚合酶识别/结合其启动子R₁的能力,并优选地,起始转录的能力可通过本领域已知方法确定如,例如,Journal of Biological Chemistry,1993,268(26):19299-19304中描述的,尽管Journal of Biological Chemistry,2005,280(52):42477-42485中回顾了发现多种DNA依赖性RNA聚合酶)。

[0076] 在优选的实施方式中,被DNA依赖性RNA聚合酶识别的启动子R₁是噬菌体启动子。

[0077] 仅作为实例,本领域已知T7DNA依赖性RNA聚合酶识别序列TAATACGACTCACTATAGGG AGA (SEQ ID NO:3)、T3DNA依赖性RNA聚合酶识别序列AATTAACCCTCACTAAAGGGGAGA (SEQ ID NO:4)、SP6DNA依赖性RNA聚合酶识别序列ATTTAGGTGACACTATAGAAG (SEQ ID NO:5)以及K11DNA依赖性RNA聚合酶识别序列AATTAGGGCACACTATAGGGA (SEQ ID NO:6)。然而,这些实例仅出于示例目的提供,因为本发明不限于这些启动子及相应的DNA依赖性RNA聚合酶。事实上,可以使用任何启动子(及其变体),只要相应的DNA依赖性RNA聚合酶,优选地噬菌体DNA依赖性RNA聚合酶,能够识别各自的序列。

[0078] 在优选的实施方式中, R₁选自:

[0079] (i) TAATACGACTCACTATAGGGAGA (SEQ ID NO:3) 或这样的序列, 其与SEQ ID NO:3相比显示1到6个取代, 并且其被T7DNA依赖性RNA聚合酶识别;

[0080] (ii) AATTAACCCTCACTAAAGGGAGA (SEQ ID NO:4) 或这样的序列, 其与SEQ ID NO:4相比显示1到6个取代, 并且其被T3DNA依赖性RNA聚合酶识别;

[0081] (iii) ATTTAGGTGACACTATAGAAG (SEQ ID NO:5) 或这样的序列, 其与SEQ ID NO:5相比显示1到6个取代, 并且其被SP6DNA依赖性RNA聚合酶识别; 和

[0082] (iv) AATTAGGGCACACTATAGGGA (SEQ ID NO:6) 或这样的序列, 其与SEQ ID NO:6相比显示1到6个取代, 并且其被K11DNA依赖性RNA聚合酶识别。

[0083] 在另一优选实施方式中, 序列可以是显示1到3、4或5个取代的序列, 只要相应序列仍能够分别被T7、T3、SP6和K11DNA依赖性RNA聚合酶识别。在更优选的实施方式中, 序列可以是显示1到2个取代的序列, 只要相应序列仍能够分别被T7、T3、SP6和K11DNA依赖性RNA聚合酶识别。最优选地, 序列可以是显示1个取代的序列, 只要相应序列仍能够分别被T7、T3、SP6和K11DNA依赖性RNA聚合酶识别。

[0084] 在其它实施方式中, 被DNA依赖性RNA聚合酶识别的启动子序列R₁不具体地限于以下的任何序列: SEQ ID NO:3、SEQ ID NO:4、SEQ ID NO:5或SEQ ID NO:6或相比于其显示1至6个取代的序列但也可以是显示1至7、8、9、10、11或12个取代的序列, 只要相应序列仍能够分别被T7、T3、SP6和K11DNA依赖性RNA聚合酶识别。

[0085] 在优选的实施方式中, 从TAATACGACTCACTATAGGGAGA (SEQ ID NO:3)、AATTAACCCTCACTAAAGGGAGA (SEQ ID NO:4)、ATTTAGGTGACACTATAGAAG (SEQ ID NO:5) 或AATTAGGGCACACTATAGGGA (SEQ ID NO:6) 的序列中的上述取代(一个或多个)中排除上述SEQ ID NO:3到6的序列中位置11到12处5个核苷酸“CACTA”处的取代, 因为这5个核苷酸在四个序列中是保守的。

[0086] 在另一优选实施方式中, 从SEQ ID NO:3到6的序列中的上述取代(一个或多个)中排除上述SEQ ID NO:3到6的序列中位置4处核苷酸“T”处的取代, 因为该核苷酸在四个序列中是保守的。

[0087] 在另一优选实施方式中, 从SEQ ID NO:3到6的序列中的上述取代(一个或多个)中排除上述SEQ ID NO:3到6的序列中位置5处核苷酸“A”处的取代, 因为该核苷酸在四个序列中是保守的。

[0088] 在另一优选实施方式中, 从SEQ ID NO:3到6的序列中的上述取代(一个或多个)中排除上述SEQ ID NO:3到6的序列中位置18处核苷酸“G”处的取代, 因为该核苷酸在四个序列中是保守的。

[0089] T7、T3、SP6和K11DNA依赖性RNA聚合酶识别/结合其启动子R₁的能力可通过上文概括的本领域已知方法确定。

[0090] 在更优选的实施方式中, 本发明的DNA分子是包含直接在所述编码序列上游的组件(b1)的DNA分子, 其中在所述组件(b1)中, SEQ ID NO:2的位置2处的核苷酸N是选自T、G或C的核苷酸, 并且其中核苷酸N不是A。

[0091] 在甚至更优选的实施方式中, SEQ ID NO:2的位置2处的所述核苷酸N是T。

[0092] 在优选的实施方式中, 本发明的DNA分子是其中直接跟在起始密码子下游的核苷

酸不是核苷酸G的DNA分子。在另一优选的实施方式中,本发明的DNA分子是其中直接跟在起始密码子下游的核苷酸是选自A、T和C的核苷酸的DNA分子。

[0093] 在甚至更优选的实施方式中,本发明的DNA分子是包含上文限定的组件 (b1) 的DNA分子,其中所述组件 (b1) 是其中SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代的序列;和/或SEQ ID NO:1的位置5处的A被G取代并且其中直接跟在起始密码子下游的核苷酸是选自A、T和C的核苷酸的序列。

[0094] 在另一甚至更优选的实施方式中,本发明的DNA分子是包含上文限定的组件 (b2) 的DNA分子,其中所述组件 (b2) 是其中SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代的序列;和/或SEQ ID NO:2的位置6处的A被G取代并且其中直接跟在起始密码子下游的核苷酸是选自A、T和C的核苷酸的序列。

[0095] 在分子生物学和遗传学中,DNA分子中的上游和下游都指代相对位置。在本发明的上下文中,上游朝向DNA分子的有义链的5'端而下游朝向该分子的3'端。

[0096] 因此,在本发明中,上文项目 (b) 中限定的序列直接位于项目 (a) 的编码区上游,更具体地,直接位于编码区的起始密码子上游。因此,在本上下文中“直接上游”意为在项目 (b) 中限定的序列和以起始密码子起始的编码序列之间不存在进一步的核苷酸(一个或多个)。因此,以起始密码子起始的编码区与上文项目 (b) 中限定的序列直接相邻。

[0097] 本发明的DNA分子可以(例如,在体内或体外系统中)以重组方式或通过本领域技术人员已知的方法(例如,通过PCR反应或在化学反应中)以合成方式产生/合成。

[0098] 本发明的DNA分子优选地是重组核酸分子,即,其由在该重组中非天然存在的元件组成。本发明的核酸分子可以是合成的或半合成的。

[0099] DNA分子可以以(上文项目 (a) 和 (b) 中分别限定的) 组件 (a) 和 (b) 的融合DNA序列的形式存在,即,通过结合含有所述组件的至少两个核苷酸序列形成的(融合)DNA分子。一般,如下文将更详细解释的,这可通过将cDNA克隆到允许转录成RNA分子的表达载体中实现。因此,本发明的DNA分子可以是融合DNA序列,即,是通过以下形成的嵌合分子:通过来自一个核苷酸的磷酸基团结合到另一核苷酸上的3'碳使两个或更多个多核苷酸连接、在一个组件的各自末端和另一分子的末端之间形成磷酸二酯键。以此方式,含有所述至少两个组件的DNA分子以DNA分子的形式连接在一起。一经在框架中(in frame)克隆,这种重组DNA分子然后可转录成编码所述蛋白质、多肽或酶分子的其相应RNA核酸序列。

[0100] 根据本发明的DNA分子可通过标准分子生物学技术(参见例如Sambrook等, *Molecular Cloning, A laboratory manual*, 第2版, 1989) 导入载体, 优选地表达载体。术语“载体(vector)”诸如本发明的含义中的“表达载体”或“克隆载体”被理解为在细胞内独立于染色体DNA进行复制并用作将遗传材料携带至细胞中的载体(vehicle)的环状、双链单元的DNA,可在其中复制和/或表达(即,转录成RNA并翻译成氨基酸序列)。含有外源DNA的载体被称为重组DNA。载体本身一般是典型地由插入物(insert)(例如,本发明的核酸分子/DNA分子)和充当载体“主链”的较大序列构成的DNA序列。在本发明的含义中,质粒最常在细菌中发现并用于重组DNA研究中以在细胞之间转移基因,并因此是本发明的含义中所使用的“载体”的亚群。

[0101] 可将进一步的调控序列添加到本发明的DNA分子对于本领域技术人员是显而易见的。例如,可以使用允许诱导表达的转录增强子和/或序列。适合的诱导系统是例如调控四

环素的基因表达——例如Gossen和Bujard在Proc.Natl.Acad.Sci.USA 89(1992),5547-5551)中和Gossen在Trends Biotech.12(1994),58-62中描述的,或是地塞米松可诱导的基因表达系统——例如Crook在EMBO J.8(1989),513-519中描述的。

[0102] 本发明还涉及载体,优选地表达载体,其包含本发明的DNA分子。

[0103] 本发明的载体可为,例如,质粒、粘粒、病毒、噬菌体或例如在遗传工程中常规使用的另一载体,以及可包含进一步的基因,例如标记基因,其允许在适合的宿主细胞中和在适合的条件选择所述载体。

[0104] 本发明的DNA分子优选地还含有聚腺苷酸信号,通过添加聚腺苷酸尾来确保转录的终止和转录物的稳定。

[0105] 本发明的DNA分子和载体可被设计用于直接导入或用于通过脂质体、病毒载体(例如腺病毒的、逆转录病毒的)、电穿孔、射击(ballistic)(例如基因枪)或其它递送系统导入到细胞中。另外,杆状病毒系统可用作本发明的核酸分子的真核表达系统。

[0106] 本发明还涉及包含本发明的载体的宿主细胞。因此,本发明涉及用本发明的载体转染或转化的宿主或携带本发明的载体的非人类宿主,即涉及用根据本发明的DNA分子或用包含这种DNA分子的载体基因修饰(基因改造,genetically modified)的宿主细胞或宿主。术语“基因修饰(基因改造,genetically modified)”意为宿主细胞或宿主包含除其天然基因组之外的被导入细胞或宿主中或导入其祖先/亲本中的一个的根据本发明的DNA分子或载体。DNA分子或载体可存在于基因修饰宿主细胞或宿主中,或者作为在基因组外的独立分子,优选地作为能够复制的分子,或其可稳定地整合到宿主细胞或宿主的基因组中。用根据本发明的载体转化宿主细胞可通过标准方法进行,例如Sambrook和Russell(2001)在Molecular Cloning:A Laboratory Manual,CSH Press,Cold Spring Harbor,NY,USA; Methods in Yeast Genetics,A Laboratory Course Manual,Cold Spring Harbor Laboratory Press,1990中描述的。在满足所使用的特定宿主细胞的要求的营养培养基中培养宿主细胞,具体地涉及pH值、温度、盐浓度、曝气、抗生素、维生素、微量元素等。

[0107] 本发明的宿主细胞可以是任意的原核细胞或真核细胞。适合的原核细胞是通常用于克隆的原核细胞如,大肠杆菌(E.coli)或枯草杆菌(Bacillus subtilis)。并且,真核细胞包括,例如,真菌细胞或动物细胞。适合的真菌细胞的实例是酵母细胞,优选地酵母属的酵母细胞以及最优选地酿酒酵母属中的酵母细胞。适合动物细胞是,例如,昆虫细胞、脊椎动物细胞,优选地哺乳动物细胞,诸如例如HEK293、NS0、CHO、COS-7、MDCK、U2-OSHeLa、NIH3T3、MOLT-4、Jurkat、PC-12、PC-3、IMR、NT2N、Sk-n-sh、CaSki、C33A。本领域已知的进一步适合的细胞系可从细胞系贮藏所获得,诸如例如,Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH(DSMZ)或美国模式培养物保藏所(American Type Culture Collection)(ATCC)。根据本发明,进而设想初级细胞/细胞培养物可起到宿主细胞的作用。所述细胞具体地衍生自昆虫(如果蝇属或蜚蠊属的昆虫)或哺乳动物(如人类、猪、小鼠或大鼠)。所述宿主细胞还可包括来自和/或衍生自细胞系如成神经细胞瘤细胞系的细胞。上述初级细胞是本领域熟知的并尤其包括初级星形胶质细胞、(混合的)脊椎培养物或海马培养物。

[0108] 本发明还涉及包括本发明的DNA分子、本发明的载体或本发明的宿主细胞的组合物。

[0109] 在第二方面中,本发明涉及RNA分子,其包含

[0110] (a) 编码区,在5'端包含起始密码子,编码多肽;和

[0111] (b) 直接在所述编码序列上游的UTR,其选自:

[0112] (b1) 序列R₂-CGCCACC (SEQ ID NO:1) 的UTR,

[0113] 或这样的序列,其中在所述UTR序列中,SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代;和

[0114] (b2) 序列R₂-CNGCCACC (SEQ ID NO:2) 的UTR,其中SEQ ID NO:2的位置2处的核苷酸N是选自U、G、C或A的核苷酸,或这样的序列,其中在所述UTR序列中,SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代,

[0115] 其中R₂是对应于启动子区域的部分的RNA序列,以DNA依赖性RNA聚合酶起始RNA合成的核苷酸起始。

[0116] 根据本发明使用的核糖核酸(RNA)分子涉及被组装为称为G、A、U、和C的核苷酸的链的聚合分子。RNA中的每一个核苷酸都含有核糖,具有编号为1'到5'的碳。含氮碱基附接到1'位置,一般地,腺嘌呤(A)、胞嘧啶(C)、鸟嘌呤(G)、或尿嘧啶(U)。在聚合RNA分子中,磷酸基团附接到一个核糖的3'位置和下一个核糖的5'位置。因此,聚合RNA分子中的核苷酸相互共价连接,其中来自一个核苷酸的磷酸基团结合到随后核苷酸上的3'碳,进而形成磷酸二酯键。因此,RNA链具有5'端和3'端,因而以核糖环上的碳来命名。根据惯例,上游和下游涉及RNA转录发生的5'到3'方向。优选地,RNA分子是信使RNA(mRNA)分子。mRNA是将来自DNA的遗传信息传递到核糖体的RNA分子的大家族,其中,RNA分子规定了基因表达的蛋白质产物的氨基酸序列。在RNA聚合酶转录初级转录物mRNA(称为前mRNA)后,加工的、成熟mRNA翻译成氨基酸的聚合物:蛋白质,如分子生物学的中心法则中概括的。如在DNA中,mRNA遗传信息存在于核苷酸的序列中,其排列成每一个都由三个碱基构成的密码子。每一个密码子编码特定氨基酸,除了终止蛋白质合成的终止密码子。

[0117] 本发明的RNA分子包含如上文项目(a)和(b)中限定的两个主要组件。另外,本发明的RNA分子优选地在其3'端包含UTR。因此,本发明的RNA分子对于其结构来说类似于天然存在的“正常”mRNA分子,携带编码区以及(5'和3')非翻译区(UTR)和任选地聚腺苷酸尾。

[0118] 根据本发明使用的术语“在其5'端包括起始密码子的编码区”涉及由密码子构成的序列,其由核糖体根据遗传密码提供的信息进行译码并翻译成蛋白质。编码区通常以其5'端的起始密码子开始并以终止密码子终止。总体上,起始密码子是AUG三联体,而终止密码子是UAA、UAG、或UGA。除了编码蛋白质,编码区的部分可充当前mRNA中的调控序列,作为外显子剪接增强子或外显子剪接沉默子。根据本发明使用的编码多肽或蛋白质的基因编码区还称为编码序列或CDS(来自编码DNA序列)并且是基因的DNA或RNA的部分——由编码多肽或蛋白质的外显子构成。mRNA中的编码区的侧翼是5'非翻译区(5'UTR)和3'非翻译区(3'UTR),其也是外显子的部分。而且,mRNA分子可进一步包含所谓的5'帽和聚腺苷酸尾。5'帽、5'UTR、3'UTR和聚腺苷酸尾是mRNA分子的不翻译成蛋白质的区域。

[0119] 根据本发明使用的术语“非翻译区”或“UTR”涉及mRNA的在起始密码子上游和终止密码子下游的不翻译的区段,并因此分别称为5端非翻译区(5'UTR)和3端非翻译区(3'UTR)。这些区域与编码区一起被转录并因此是外显的,因为它们存在于成熟mRNA中。

[0120] 如本发明中使用的,3'非翻译区(3'-UTR)涉及直接跟随翻译终止密码子的信使RNA(mRNA)的区段。3'UTR可在3'非翻译区内包含调控区域,其已知影响多聚腺苷化作用和mRNA的稳定性。许多3'-UTR还含有富含AU的元件(ARE)。并且,3'-UTR可优选地含有序列AAUAAA——其指导将称为聚腺苷酸尾的数百个腺嘌呤残基添加到mRNA转录物的末端。

[0121] 如本发明中使用的,5'非翻译区(5'UTR)(还称为前导序列或前导RNA)是直接起始密码子上游的mRNA区域。5'UTR在转录起始位点开始并在编码区的起始密码子(通常为AUG)前的一个核苷酸(nt)终止。在真核生物中,5'UTR的长度一般为100到数千个核苷酸长,并且有时在真核生物中也存在较短的UTR。

[0122] 在本发明中,5'UTR是极短的,因为其是本发明用来提供最小UTR序列的目标。

[0123] 本发明的RNA分子可还含有聚腺苷酸尾。聚腺苷酸尾是通过称为多聚腺苷化作用的过程添加到前mRNA的3'端的腺嘌呤核苷酸的长序列(通常为数百)。该尾促进从细胞核输出以及翻译,并防止mRNA降解。多聚腺苷化作用是将聚腺苷酸尾添加到信使RNA。聚腺苷酸尾由多个腺苷单磷酸构成;换言之,它是仅具有腺嘌呤碱基的RNA的延伸。在真核细胞中,多聚腺苷化作用是产生用于翻译的成熟信使RNA(mRNA)的过程的部分。

[0124] RNA分子的一个组件,即,“在其5'端包括起始密码子的编码多肽的编码区”(组件(a))不具体地限制,并可以是将在给定细胞中表达的任何期望的编码区。关于术语“在其5'端包括起始密码子的编码多肽的编码区”(组件(a))的优选实施方式,加上必要的修改,同样适用于本发明的RNA分子,如上面在本发明的DNA分子的上下文中列出的。

[0125] 本发明的RNA分子包含直接在所述编码序列上游的组件(b),其中所述组件(b)是选自以下的UTR:

[0126] (b1)序列R₂-CGCCACC(SEQ ID NO:1)的UTR,

[0127] 或这样的序列,其中在所述UTR序列中,SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代;和

[0128] (b2)序列R₂-CNGCCACC(SEQ ID NO:2)的UTR,其中SEQ ID NO:2的位置2处的核苷酸N是选自U、G、C或A的核苷酸,或这样的序列,其中在所述UTR序列中,SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代,

[0129] 其中R₂是对应于启动子区域的部分的RNA序列,以DNA依赖性RNA聚合酶起始RNA合

[0130] 成的核苷酸开始。

[0131] 不具体限制R₂的性质。可以使用对应于以DNA依赖性RNA聚合酶起始RNA合成的核苷酸开始的启动子区域的部分的任何RNA序列。技术人员能够容易地确定从DNA依赖性RNA聚合酶起始RNA合成的核苷酸开始的启动子区域的那些部分。该RNA序列R₂是对应于被转录的启动子部分的启动子序列,即,其实际上一经转录就存在于转录物中。

[0132] 在优选实施方式中,启动子R₂是对应于以噬菌体衍生的DNA依赖性RNA聚合酶起始RNA合成的核苷酸开始的启动子区域的部分的RNA序列。

[0133] 在优选实施方式中,启动子R₂是对应于以T7DNA依赖性RNA聚合酶、T3DNA依赖性RNA聚合酶、SP6DNA依赖性RNA聚合酶或K11DNA依赖性RNA聚合酶起始RNA合成的核苷酸开始的启动子区域的部分的RNA序列。

[0134] 为了对此示例,作为非限制性实例,R₂是以下启动子序列中的标下划线的

(underlined) 序列:TAATACGACTCACTATAGGGAGA (SEQ ID NO:3;即,启动子被T7DNA依赖性RNA聚合酶识别),AATTAACCCTCACTAAAGGGAGA (SEQ ID NO:4;即,启动子被T3DNA依赖性RNA聚合酶识别),ATTTAGGTGACACTATAGAAG (SEQ ID NO:5;即,启动子被SP6DNA依赖性RNA聚合酶识别)和AATTAGGGCACACTATAGGGA (SEQ ID NO:6;即,启动子被K11DNA依赖性RNA聚合酶识别)。标下划线的序列对应于DNA依赖性RNA聚合酶起始RNA合成并因此一经转录就实际存在于RNA分子(即,在转录物中)中的各自的启动子的部分。

[0135] 相比于序列R₂-CGCCACC (SEQ ID NO:1)的UTR或相比于序列R₂-CNGCCACC (SEQ ID NO:2)的UTR,具有任意以上取代的UTR序列(一个或多个)可导致显示分别与包含序列R₂-CGCCACC (SEQ ID NO:1)的UTR的RNA分子和包含序列R₂-CNGCCACC (SEQ ID NO:2)的UTR的RNA分子相同或相似的,优选地较高翻译效率的RNA分子。包含本文所述UTR的给定RNA分子的翻译效率可由技术人员通过本领域已知以及在下文中描述的方法来确定。

[0136] 翻译效率是在细胞内mRNA翻译成多肽或蛋白质的速率。给定mRNA的翻译效率被测量为每个mRNA每时间单位翻译的蛋白质或多肽的数量。翻译是其中细胞核糖体产生蛋白质的过程并且是技术人员熟知的。简略地,在翻译中,从DNA转录产生的信使RNA(mRNA)经核糖体译码从而产生特定氨基酸链或多肽或蛋白质。

[0137] 因此,与相同的却分别携带上文限定的R₂-CGCCACC (SEQ ID NO:1)或R₂-CNGCCACC (SEQ ID NO:2)的UTR的给定RNA的翻译效率相比,携带具有任意以上取代的修饰的UTR序列的给定RNA分子的翻译效率优选地相同或更高。因此,每时间单位每个RNA翻译的由携带具有任意以上取代的修饰的UTR序列的RNA分子的编码区编码的蛋白质或多肽的数量至少是相同于,或优选地,高于每个RNA每时间单位翻译的由分别携带上文中限定的R₂-CGCCACC (SEQ ID NO:1)或R₂-CNGCCACC (SEQ ID NO:2)的UTR的RNA分子的编码区编码的蛋白质或多肽的数量。

[0138] 在本发明的上下文中,翻译效率优选地是在细胞内在某个时间点mRNA翻译成蛋白质的速率,与在所述细胞中在相同时间点编码各自蛋白质的mRNA的量有关。因此,翻译效率是在细胞内在某个时间点翻译成蛋白质的mRNA与编码各自的蛋白质的mRNA的量的商。两个参数,即,翻译成蛋白质的mRNA以及编码各自的蛋白质的mRNA的量,可通过本领域已知方法确定。作为非限制性实例,在细胞内翻译成蛋白质的mRNA的量可以确定,例如通过流式细胞仪(FC)来确定,同时编码各自蛋白质的mRNA的量可通过例如qPCR来测量。

[0139] 上文项目(b)中限定的UTR(一个或多个)不具体地限于上述特定序列,而且还可涉及包含相比于这种序列显示(a)核苷酸(一个或多个)添加(一个或多个)的序列的(a)UTR序列(一个或多个),其中可在上述UTR(一个或多个)的5'端添加另外的核苷酸(一个或多个)。另外的核苷酸(一个或多个)包括多达0(无变化)、1、2、3、4、5、6、7、8、9或10个核苷酸的多核苷酸链,优选地多达20个核苷酸的多核苷酸链。更优选地,在5'端添加11、12、13、14、15、16、18、或19个核苷酸。甚至更优选地,在5'端添加多达30个核苷酸的多核苷酸链。

[0140] 根据基本原理,核苷酸的添加可能不改变各自的UTR(一个或多个)的上述功能特性,核苷酸添加的长度还可为多达40、50、60、70、80、90、或甚至100个核苷酸或甚至更多,多达200、300、400或500个核苷酸,只要这些序列具有与上文项目(b)中限定的UTR相似的能力(就上述翻译效率而言)。

[0141] 在优选实施方式中,上文项目(b1)中限定的UTR具有的最大长度为11、12或13个核

苷酸。优选地,上文项目 (b1) 中限定的UTR具有的最大长度为13个核苷酸,如果R₂是GGGAGA (SEQ ID NO:7) 或GGGAGA (SEQ ID NO:8)。

[0142] 优选地,上文项目 (b1) 中限定的UTR具有的最大长度为11个核苷酸,如果R₂是GAAG (SEQ ID NO:9) 或GGGA (SEQ ID NO:10)。

[0143] 在另一优选实施方式中,上文项目 (b2) 中限定的UTR具有的最大长度为12、13或14个核苷酸。优选地,上文项目 (b2) 中限定的UTR具有的最大长度为14个核苷酸,如果R₂是GGGAGA (SEQ ID NO:7) 或GGGAGA (SEQ ID NO:8)。

[0144] 优选地,上文项目 (b2) 中限定的UTR具有的最大长度为12个核苷酸,如果R₂是GAAG (SEQ ID NO:9) 或GGGA (SEQ ID NO:10)。

[0145] 含有上述UTR (一个或多个) 的本发明的RNA分子可以通过本领域技术人员已知的方法以重组方式 (例如,在体内系统或体外系统中) 或以合成方式产生/合成。

[0146] RNA的体外转录通常需要含有双链启动子区域的线性DNA模板,在该双链启动子区域DNA依赖性RNA聚合酶结合并起始RNA合成,同时编码区可以是双链或单链。在线性DNA模板含有单链编码区的情况下,编码区的反义链 (即,DNA依赖性聚合酶所阅读的链) 是模板的部分。常见的DNA依赖性RNA聚合酶是T7聚合酶、T3聚合酶、SP6聚合酶和K11聚合酶。SEQ ID NO:3到6中显示了其各自的启动子的全部序列。

[0147] 用于体外转录的转录模板包括,例如,由RNA前体合成的cDNA模板、PCR产生的模板、化学合成的寡核苷酸和质粒构建物。多种广泛使用的质粒克隆载体携带噬菌体聚合酶启动子——其位于多克隆位点的每一侧从而允许转录插入多克隆位点中的核苷酸序列的任一链。通常使用的克隆载体包括例如Invitrogen的pCRII、Promega的pGEM和Stratagene的pBluescript载体。Ambion的pTRIPLEscript载体家族含有串联的全部三种噬菌体聚合酶启动子 (在多克隆位点的相同侧), 允许任意这三种聚合酶——SP6、T7或T3被使用。

[0148] 本发明的RNA分子可通过本领域技术人员已知的方法在体内系统中以重组方式产生。

[0149] 可选地,可利用例如体外转录系统在体外系统中产生本发明的RNA分子。体外转录系统是常见已知的并通常需要含有“编码”RNA分子的DNA序列的纯化的线性DNA模板,其中所述DNA序列处于适当启动子的控制下。而且,体外转录系统通常还需要核糖核苷三磷酸、包含DTT和镁离子的缓冲系统、和为将DNA序列体外转录成本发明的相应RNA分子提供酶活的适当的RNA聚合酶。

[0150] 并且, RNA分子可化学合成,例如,利用固相支持和标准技术在自动化核苷酸序列合成仪上通过常规化学合成或通过化学合成各自的DNA-序列并随后在体外或体内对其转录。

[0151] 根据上文,本发明提供了RNA分子/多核糖核酸分子,优选地修饰的多核糖核酸分子,其中所述RNA分子的一个组件,即,“在其5'端包括起始密码子的编码区” (组件 (a)), 编码多肽。术语核酸和多核苷酸可交换使用并包括包含核苷酸聚合物的任意化合物和/或物质。术语核苷酸包括脱氧核苷酸和核糖核苷酸。术语核糖核酸和多核糖核苷酸可交换使用,并且在某些实施方式中,包括包含核苷酸聚合物的任意化合物和/或物质,其中大于50%的核苷酸是核糖核苷酸。在某些实施方式中,多核糖核苷酸包含核苷酸聚合物,其中大于60%、70%、75%、80%、90%、大于95%、大于99%或100%的核苷酸是核糖核苷酸。其中一

个或多个核苷酸是修饰的核苷酸的多核糖核苷酸可称为修饰多核糖核苷酸。然而，术语多核糖核苷酸可包括修饰的多核糖核苷酸。

[0152] RNA分子/多核糖核苷酸的序列可衍生自，例如，包含感兴趣基因的遗传信息的任何合适的核酸。核酸的实例包括基因组DNA、RNA、或来自包含感兴趣基因(一个或多个)的任何细菌细胞或古细菌细胞的cDNA。多核苷酸可衍生自携带突变基因和多态性的核酸。本发明的RNA分子/多核糖核苷酸包含不具体限制的序列并可包含作为组件A在给定细胞中表达的任何期望的编码区。在优选实施方式中，所述序列可以是上文概括的编码期望多肽/蛋白质的编码区。优选地，与上文一致，RNA分子/多核糖核苷酸进一步包含位于组件A的起始密码子上游(5')的非翻译序列、位于组件A的终止密码子下游(3')的非翻译序列、或位于组件A的起始密码子上游(5')的非翻译序列和位于组件A的终止密码子下游(3')的非翻译序列两者。在优选的实施方式中，本发明的RNA分子/多核糖核苷酸可以是修饰的RNA分子/多核糖核苷酸。

[0153] 除了四种经典的核糖核苷酸，即，腺苷、鸟苷、胞苷和尿苷之外，还存在这些核酸碱基中的每一种的多种类似物。有时贯穿并在文献中，这些类似物或包括一种或多种这些类似物的RNA分子/多核糖核苷酸被称为修饰的(例如，修饰的核苷酸或修饰的核糖核苷酸)。一些类似物不同于上述规范的核酸碱基，但也存在于自然界。其它类似物是非天然存在的。考虑类似物的任一类型。

[0154] 在某些实施方式中，本发明的RNA分子/多核糖核苷酸包括核苷酸类似物(例如，多核糖核苷酸包括修饰的多核糖核苷酸)。下文提供了示例性核苷酸类似物(例如，U的类似物；C的类似物；A的类似物；G的类似物)。另外，在某些实施方式中，本公开的RNA分子/多核糖核苷酸或其它核酸还可(除此之外或可选地)在磷酸二酯主链中或在核酸碱基之间的连接中包含修饰。可形成本公开的RNA分子/多核糖核苷酸的部分或全部的示例性核酸包括，但不限于，核糖核酸(RNA)、脱氧核糖核酸(DNA)、苏阿糖核酸(TNA)、乙二醇核酸(GNA)、肽核酸(PNA)、锁定核酸(LNA，包括具有 β -D-核糖构型的LNA、具有 α -L-核糖构型(LNA的非对映异构体)的 α -LNA、具有2'-氨基功能化的2'-氨基-LNA、和具有2'-氨基功能化的2'-氨基- α -LNA)或其杂交物。

[0155] 在某些实施方式中，修饰可以在核酸/多核苷酸分子的一个或多个核苷(一个或多个)上或主链上。在某些实施方式中，修饰可在核苷和主链连接两者上。在某些实施方式中，修饰可在体外改造成多核苷酸。在某些实施方式中，修饰的核糖核苷酸/核苷酸还可通过共价修饰经典的/天然的核糖核苷酸/核苷酸以转录后方式合成。

[0156] 本发明的RNA分子/多核糖核苷酸可以是修饰的RNA分子/多核糖核苷酸，并且在某些实施方式中，可包含嘌呤类似物和/或嘧啶类似物。在某些实施方式中，本发明的修饰的RNA分子/多核糖核苷酸包含嘧啶类似物，例如尿苷类似物和/或胞苷类似物。在某些实施方式中，本发明的修饰的RNA分子/多核糖核苷酸包含尿苷类似物和胞苷类似物。在某些实施方式中，修饰的RNA分子/多核糖核苷酸不包含腺苷类似物和/或鸟苷类似物。在某些实施方式中，RNA分子/多核糖核苷酸包含单一类型的尿苷类似物和单一类型的胞苷类似物(例如，一种类型的类似物，不是类似物的单一分子——单一类似物可以以本文所述任意若干百分数存在)。在其它实施方式中，RNA分子/多核糖核苷酸包含多于一种类型的尿苷和/或胞苷类似物，和任选地且如果存在，一种或多种腺苷和/或鸟苷类似物(或没有两者之一或两

者)。

[0157] 在某些情况下,修饰的尿苷(例如,尿苷类似物)选自2-硫尿苷、5'-甲基尿苷、假尿苷、5-碘尿苷(I5U)、4-硫尿苷(S4U)、5-溴尿苷(Br5U)、2'-甲基-2'-脱氧尿苷(U2'm)、2'-氨基-2'-脱氧尿苷(U2'NH₂)、2'-叠氮-2'-脱氧尿苷(U2'N₃)、和2'-氟-2'-脱氧尿苷(U2'F)。在某些情况下,修饰的胞苷(例如,胞苷类似物)选自5-甲基胞苷、3-甲基胞苷、2-硫-胞苷、2'-甲基-2'-脱氧胞苷(C2'm)、2'-氨基-2'-脱氧胞苷(C2'NH₂)、2'-氟-2'-脱氧胞苷(C2'F)、5-碘胞苷(I5C)、5-溴胞苷(Br5C)和2'-叠氮-2'-脱氧胞苷(C2'N₃)。注意当指代类似物时,前述还指代以其5'三磷酸盐形式的类似物。在某些实施方式中,胞苷类似物是5-碘胞苷而尿苷类似物是5-碘尿苷。

[0158] 在一些实施方式中,RNA分子/多核糖核苷酸是修饰的RNA分子/多核糖核苷酸。在某些情况下,与非修饰的(或未修饰的)RNA分子/多核糖核苷酸相比,修饰的RNA分子/多核糖核苷酸的稳定性至少高25%。在某些情况下,与非修饰的RNA分子/多核糖核苷酸相比,修饰的RNA分子/多核糖核苷酸的稳定性可至少高30%、至少高35%、至少高40%、至少高45%、至少高50%、至少高55%、至少高60%、至少高65%、至少高70%、至少高75%、至少高80%、至少高85%、至少高90%、或至少高95%。在某些实施方式中,在体内测量稳定性。在某些实施方式中,在体外测量稳定性。在某些实施方式中,稳定性通过测量多核糖核苷酸的半衰期来定量。

[0159] 本发明的RNA分子/多核糖核苷酸可具有以相同形式修饰的核苷酸或其它不同修饰的核苷酸的混合物。修饰的核苷酸可具有天然存在或非天然存在于信使RNA中的修饰。可使用各种修饰的核苷酸的混合物。例如,RNA分子/多核糖核苷酸内一个或多个修饰的核苷酸可具有天然修饰,而另外部分具有不在mRNA中天然发现的修饰。另外,一些修饰的核苷酸可具有碱基修饰,而其它修饰的核苷酸具有糖修饰。以相同方式,可能所有修饰都是碱基修饰或所有修饰都是糖修饰或其任意合适的混合。在一些情况下,修饰的RNA分子/多核糖核苷酸的稳定性可通过改变修饰的多核糖核苷酸内修饰的碱基的性质来选择性优化。

[0160] 表2:U类似物的非限制性实例

[0161]

名称	碱基修饰(5'位置)	糖修饰(2'位置)	天然地在 mRNA 中
5-甲基尿苷(m5U)	CH ₃	-	否
5-碘尿苷(I5U)	I	-	否
5-溴尿苷(Br5U)	Br	-	否

[0162]

2-硫尿苷 (S2U)	S (在位置 2 中)	-	否
4-硫尿苷 (S4U)	S (在位置 4 中)	-	否
2'-甲基-2'-脱氧尿苷 (U2'm)	-	CH ₃	是
2'-氨基-2'-脱氧尿苷 (U2'NH ₂)	-	NH ₂	否
2'-叠氮-2'-脱氧尿苷 (U2'N ₃)	-	N ₃	否
2'-氟-2'-脱氧尿苷 (U2'F)	-	F	否

[0163] 表3:C类似物的非限制性实例

[0164]

名称	碱基修饰 (5' 位置)	糖修饰 (2' 位置)	天然地在 mRNA 中
5-甲基胞苷 (m5C)	CH ₃	-	是
5-碘胞苷 (I5C)	I	-	否
5-溴胞苷 (Br5C)	Br	-	否
2-硫胞苷 (S2C)	S (在位置 2 中)	-	否
2'-甲基-2'-脱氧胞苷 (C2'm)	-	CH ₃	是
2'-氨基-2'-脱氧胞苷 (C2'NH ₂)	-	NH ₂	否
2'-叠氮-2'-脱氧胞苷 (C2'N ₃)	-	N ₃	否
2'-氟-2'-脱氧胞苷 (C2'F)	-	F	否

[0165] 表4:A类似物的非限制性实例

[0166]

名称	碱基修饰 (5' 位置)	糖修饰 (2' 位置)	天然地在 mRNA 中
N6-甲基腺苷 (m6A)	CH ₃ (在位置 6 中)	-	是
N1-甲基腺苷 (m1A)	CH ₃ (在位置 1 中)	-	否
2'-O-甲基腺苷 (A2'm)	-	CH ₃	是
2'-氨基-2'-脱氧腺苷 (A2'NH ₂)	-	NH ₂	否
2'-叠氮-2'-脱氧腺苷	-	N ₃	否

[0167]

(A2'N3)			
2'-氟-2'-脱氧腺苷(A2'F)	-	F	否

[0168] 表5:G类似物的非限制性实例

[0169]

名称	碱基修饰 (5' 位置)	糖修饰 (2' 位置)	天然地在 mRNA 中
N1-甲基鸟苷 (m1G)	CH ₃ (在位置 1 中)		否
2'-O-甲基鸟苷 (G2'm)	-	CH ₃	是
2'-氨基-3'-脱氧鸟苷 (G2'NH ₂)	-	NH ₂	否
2'-叠氮-2'-脱氧鸟苷 (G2'N3)	-	N ₃	否
2'-氟-2'-脱氧鸟苷 (G2'F)	-	F	否

[0170] 在某些实施方式中,类似物(例如,修饰的核苷酸)可选自以下:吡啶-4-酮核糖核苷、5-碘尿苷、5-碘胞苷、5-氮杂-尿苷、2'-氨基-2'-脱氧胞苷、2'-氟-2'-脱氧胞苷、2-硫-5-氮杂-尿苷、2-硫尿苷、4-硫-假尿苷、2-硫-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-尿苷、1-牛磺酸甲基-4-硫-尿苷、5-甲基-尿苷、1-甲基-假尿苷、4-硫-1-甲基-假尿苷、2-硫-1-甲基-假尿苷、1-甲基-1-脱氮-假尿苷、2-硫-1-甲基-1-脱氮-假尿苷、二氢尿苷、二氢假尿苷、2-硫-二氢尿苷、2-硫-二氢假尿苷、2-甲氧基尿苷、2-甲氧基-4-硫-尿苷、4-甲氧基-假尿苷、4-甲氧基-2-硫-假尿苷、5-氮杂-胞苷、假异胞苷(pseudoisocytidine)、3-甲基-胞苷、N4-乙酰胞苷、5-甲酰胞苷、5-甲基胞苷、N4-甲基胞苷、5-羟基甲基胞苷、1-甲基-假异胞苷、吡咯并-胞苷(pyrrolo-cytidine)、吡咯并-假异胞苷、2-硫-胞苷、2-硫-5-甲基-胞苷、4-硫-假异胞苷、4-硫-1-甲基-假异胞苷、4-硫-1-甲基-1-脱氮-假异胞苷、1-甲基-1-脱氮-假异胞苷、zebularine、5-氮杂-zebularine、5-甲基-zebularine、5-氮杂-2-硫-zebularine、2-硫-zebularine、2-甲氧基-胞苷、2-甲氧基-5-甲

基-胞苷、4-甲氧基-假异胞苷、4-甲氧基-1-甲基-假异胞苷、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-异戊烯腺苷、N6-(顺-羟基异戊烯基)腺苷、2-甲基硫-N6-(顺-羟基异戊烯基)腺苷、N6-甘氨酸氨基甲酰腺苷、N6-苏氨酸氨基甲酰腺苷、2-甲基硫-N6-苏氨酸氨基甲酰腺苷、N6,N6-二甲基腺苷、7-甲基腺嘌呤、2-甲基硫-腺嘌呤、2-甲氧基-腺嘌呤、肌苷、1-甲基-肌苷、γ-苷、wybutosine、7-脱氮-鸟苷、7-脱氮-8-氮杂-鸟苷、6-硫-鸟苷、6-硫-7-脱氮-鸟苷、6-硫-7-脱氮-8-氮杂-鸟苷、7-甲基-鸟苷、6-硫-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-鸟苷、7-甲基-8-氧-鸟苷、1-甲基-6-硫-鸟苷、N2-甲基-6-硫-鸟苷、和N2,N2-二甲基-6-硫-鸟苷。

[0171] 在某些实施方式中,本发明的修饰的RNA分子/多核糖核苷酸不包括假尿苷。在某些实施方式中,本发明的修饰的RNA分子/多核糖核苷酸不包括5-甲基胞苷。在某些实施方式中,本发明的修饰的RNA分子/多核糖核苷酸不包括5-甲基尿苷。在某些实施方式中,本发明的修饰的RNA分子/多核糖核苷酸包括U类似物和C类似物,其中这种U类似物可以都是相同类似物或可以是不同类似物(例如,多于一种类型的类似物),并且其中这种C类似物可以都是相同类似物或可以是不同类似物(例如,多于一种类型的类似物)。在某些实施方式中,本发明的修饰的RNA分子/多核糖核苷酸不包括腺苷类似物和鸟苷类似物。

[0172] 如本文详细描述,当RNA分子/多核糖核苷酸包含修饰的多核糖核苷酸时,类似物可以以某种核苷酸比例存在于化合物中(例如,给定百分数的给定核碱基可以是类似物,如本文所述)。

[0173] 包含至少一个修饰的核苷酸的RNA分子/多核糖核苷酸是修饰的RNA分子/多核糖核苷酸。在某些实施方式中,至少约5%的修饰的RNA分子/多核糖核苷酸包括修饰的或非天然存在的(例如,的类似物或修饰的)腺苷、胞苷、鸟苷、或尿苷,如本文所述的类似物核苷酸。在一些情况下,至少约5%、10%、15%、20%、25%、30%、40%、45%、50%的修饰的RNA分子/多核糖核苷酸包括修饰的或非天然存在的(例如,的类似物或修饰的)腺苷、胞苷、鸟苷、或尿苷。在一些情况下,至多约50%、45%、40%、35%、30%、25%、20%、15%、10%、5%的修饰的RNA分子/多核糖核苷酸包括修饰的或非天然存在的腺苷、胞苷、鸟苷、或尿苷。

[0174] 在优选的实施方式中,本发明的RNA分子含有修饰的和未修饰的核苷酸的组合。优选地,本发明的RNA分子含有W02011/012316中所述的修饰的和未修饰的核苷酸的组合。这种RNA分子还称为并且商业化为“SNIM®-RNA”。W02011/012316中所述的RNA分子据报道显示增加的稳定性和减少的免疫原性。在优选的实施方式中,在这种修饰的RNA分子中,5到50%的胞苷核苷酸和5到50%的尿苷核苷酸是修饰的。含有腺苷和含有鸟苷的核苷酸可以是未修饰的。腺苷和鸟苷核苷酸可以是未修饰的或部分修饰的,并且它们优选地以未修饰形式存在。优选地,10到35%的胞苷和尿苷核苷酸是修饰的,并且特别优选地,修饰的胞苷核苷酸的含量在7.5到25%的范围中,而修饰的尿苷核苷酸的含量在7.5到25%的范围中。已发现实际上相对低的含量,例如每一种仅10%的修饰的胞苷和尿苷核苷酸能获得期望的特性。特别优选地是,修饰的胞苷核苷酸是5-甲基胞苷残基,而修饰的尿苷核苷酸是2-硫尿苷残基。最优选地,修饰的胞苷核苷酸的含量和修饰的尿苷核苷酸的含量分别是25%。

[0175] 在某些其它实施方式中,在这种修饰的RNA分子/多核糖核苷酸分子中,5到50%的

胞苷是C的类似物而5到50%的尿苷是U的类似物。在某些实施方式中,在这种修饰的多核糖核苷酸分子中,5到40%的胞苷是C的类似物而5到40%的尿苷是U的类似物。在某些实施方式中,在这种修饰的RNA分子/多核糖核苷酸分子中,5到30%的胞苷是C的类似物而5到30%的尿苷是U的类似物。在某些实施方式中,在这种修饰的RNA分子/多核糖核苷酸分子中,10到30%的胞苷是C的类似物而10到30%的尿苷是U的类似物。在某些实施方式中,在这种修饰的多核糖核苷酸分子中,5到20%的胞苷是C的类似物而5到20%的尿苷是U的类似物。在某些实施方式中,在这种修饰的RNA分子/多核糖核苷酸分子中,5到10%的胞苷核苷酸和5到10%的尿苷核苷酸是修饰的。在某些实施方式中,在这种修饰的RNA分子/多核糖核苷酸分子中,25%的胞苷核苷酸和25%的尿苷核苷酸是修饰的。在某些实施方式中,含有腺苷和含有鸟苷的核苷酸可以是未修饰的。在某些实施方式中,腺苷和鸟苷核苷酸可以是未修饰的或部分修饰的,并且它们优选地以未修饰形式存在。

[0176] 如上指出的,在某些实施方式中,U的类似物指代单一类型的U的类似物。在某些实施方式中,U的类似物指代两种或更多种类型的U的类似物。在某些实施方式中,C的类似物指代单一种类型的C的类似物。在某些实施方式中,C的类似物指代两种或更多种类型的C的类似物。

[0177] 在某些实施方式中,RNA分子/多核糖核苷酸中是胞苷类似物的胞苷的百分数与RNA分子/多核糖核苷酸中是尿苷类似物的尿苷的百分数不同。在某些实施方式中,胞苷类似物的百分数低于尿苷类似物的百分数。如上指出的,这可能是在腺苷和鸟苷的类似物的存在或不存在的条件下,但在某些实施方式中,是在腺类似物苷和鸟苷类似物不存在的条件下。在某些实施方式中,本公开的多核糖核苷酸包含少于15%、少于10%、少于5%或少于2%的腺苷类似物、鸟苷类似物或两者。

[0178] 在某些实施方式中,本发明的RNA分子/多核糖核苷酸包含胞苷类似物和尿苷类似物,并且5到20%的胞苷是胞苷类似物而25到45%的尿苷是尿苷类似物。换言之,RNA分子/多核糖核苷酸包含修饰的和未修饰的胞苷以及修饰的和未修饰的尿苷,并且5到20%的胞苷包括胞苷类似物同时25到45%的尿苷包括尿苷类似物。在其它实施方式中,RNA分子/多核糖核苷酸包含5到10%的胞苷类似物和30到40%的尿苷类似物,例如7-9%的胞苷类似物——例如约7、7.5或8%,和例如32-38%的尿苷类似物——例如约33、34、35、36%。

[0179] 在某些实施方式中,可以使用本文所述的任何尿苷类似物和胞苷类似物,任选地不包括假尿苷。在某些实施方式中,胞苷类似物包括5-碘胞苷或由5-碘胞苷组成(例如,在组成的情况下,其就是所使用的单一类似物种类)以及尿苷类似物包括5-碘尿苷或由5-碘尿苷组成(例如,在组成的情况下,其就是所使用的单一类似物种类)。

[0180] 在任意前述的某些实施方式中,给定核苷酸的类似物的百分数指代输入百分数(例如,起始反应中类似物的百分数,例如起始体外转录反应)。在任意前述的某些实施方式中,给定核苷酸的类似物的百分数指代输出(例如,合成或转录的化合物中的百分数)。

[0181] 本发明的RNA分子/多核糖核苷酸分子可通过本领域技术人员已知的方法在体内系统中以重组方式产生,其将在下文进一步更详细地描述。

[0182] 可选地,本发明的修饰的多核糖核苷酸分子可在体外系统中产生,利用例如在下文进一步更详细地描述的体外转录系统。能够产生RNA分子/多核糖核苷酸的体外转录系统需要修饰的和未修饰的核苷三磷酸盐的输入混合物以产生具有本发明的期望特性的修饰

的RNA分子/多核糖核苷酸。在某些实施方式中,在这种输入混合物中5到50%的胞苷是胞苷类似物以及在这种输入混合物中5到50%的尿苷是尿苷类似物。在某些实施方式中,在这种输入混合物中5到40%的胞苷是胞苷类似物以及在这种输入混合物中5到40%的尿苷是尿苷类似物。在某些实施方式中,在这种混合物中5到30%的胞苷是胞苷类似物以及在这种输入混合物中5到30%的尿苷是尿苷类似物。在某些实施方式中,在这种混合物中5到30%的胞苷是胞苷类似物以及在这种混合物中10到30%的尿苷是尿苷类似物。在某些实施方式中,在这种输入混合物中5到20%的胞苷是胞苷类似物以及在这种输入混合物中5到20%的尿苷是尿苷类似物。在某些实施方式中,在这种输入混合物中5到10%的胞苷是胞苷类似物以及在这种输入混合物中5到10%的尿苷是尿苷类似物。在某些实施方式中,在这种输入混合物中25%的胞苷是胞苷类似物以及在这种输入混合物中25%的尿苷是尿苷类似物。在某些实施方式中,输入混合物不包含腺苷和/或鸟苷类似物。在其它实施方式中,任选地,输入混合物包含一种或多种腺苷和/或鸟苷类似物(或没有两者之一或两者)。

[0183] 在某些实施方式中,输入混合物中是胞苷类似物的胞苷的百分数与输入混合物中是尿苷类似物的尿苷的百分数不相同。在某些实施方式中,输入混合物中胞苷类似物的百分数低于输入混合物中尿苷类似物的百分数。如上指出的,这可能是在输入混合物中腺苷和鸟苷类似物存在或不存在的条件下,但在某些实施方式中,是在输入混合物中腺苷类似物和鸟苷类似物不存在的条件下。

[0184] 在某些实施方式中,产生本发明的RNA分子/多核糖核苷酸的体外转录系统的核苷酸的输入混合物包含胞苷类似物和尿苷类似物,并且5到20%的输入混合物的胞苷是胞苷类似物而25到45%的输入混合物的尿苷是尿苷类似物。换言之,输入混合物包含修饰的和未修饰的胞苷以及修饰的和未修饰的尿苷,并且5到20%的输入混合物的胞苷包括胞苷类似物同时25到45%的输入混合物的尿苷包括尿苷类似物。在其它实施方式中,输入混合物包含5到10%的胞苷类似物和30到40%的尿苷类似物,例如7-9%的胞苷类似物——例如7、7.5或8%,和例如32-38%的尿苷类似物——例如33、34、35、36%。

[0185] 在某些实施方式中,可以使用本文所述的任意尿苷类似物和胞苷类似物,任选地不包括假尿苷。在某些实施方式中,胞苷类似物包含5-碘胞苷或由5-碘胞苷组成(例如,其就是所使用的单一C类似物种类)以及尿苷类似物包含5-碘尿苷或由5-碘尿苷组成(例如,其就是所使用的单一U类似物)。

[0186] 上表中描述了示例性类似物。应当理解,对于编码期望多肽(组件(a))的修饰的多核糖核苷酸,类似物和修饰的水平(除非以其它方式说明)被认为横跨编码期望多肽(组件(a))的整个多核糖核苷酸,包括5'和3'非翻译区(例如,修饰的水平基于体外转录反应中类似物的输入比例使得类似物可并入所转录的位置处)。

[0187] 并且,修饰的RNA分子/多核糖核苷酸分子可化学合成,例如,利用固相支持和标准技术在自动化核苷酸序列合成仪上通过常规化学合成或通过化学合成各自的DNA序列并随后在体外或体内对其转录。

[0188] 在分子生物学和遗传学中,上游和下游均指代RNA分子中的相对位置。在本发明的上下文中,上游朝向RNA分子的5'端而下游朝向分子的3'端。

[0189] 因此,在本发明中,上文项目(b)中限定的UTR直接位于项目(a)的编码区上游,更具体地,直接位于编码区的起始密码子上游。因此,在本上下文中“直接上游”意为项目(b)

中限定的UTR和以起始密码子起始的编码序列之间不存在进一步的核苷酸(一个或多个)。因此,以起始密码子起始的编码区直接直接与所述UTR序列相邻。

[0190] RNA分子可以以组件(a)和(b)(上文项目(a)和(b)中分别限定的)的融合RNA序列的形式存在,即,(融合)RNA分子,其由通过结合编码所述组件的至少两个核苷酸序列产生的杂合基因的表达形成。一般地,如将在下文进一步更详细地解释,这可通过将cDNA克隆到允许转录成RNA分子的表达载体中来实现。因此,编码本发明的RNA分子的DNA分子可以是融合DNA序列,即,嵌合分子,其通过来自一个核苷酸的磷酸基团结合到另一核苷酸上的3'碳使两个或更多个多核苷酸连接——在一个组件的各自末端和另一分子的末端之间形成磷酸二酯键——而形成。以此方式,编码所述至少两个组件的上述DNA分子以DNA分子的形式连接在一起。这种重组DNA分子然后转录成其相应的RNA核酸序列。

[0191] 在一个优选的实施方式中,R₂选自:

[0192] (i) GGGAGA (SEQ ID NO:7);

[0193] (ii) GGGAGA (SEQ ID NO:8);

[0194] (iii) GAAG (SEQ ID NO:9);和

[0195] (iv) GGGG (SEQ ID NO:10)。

[0196] 在优选的实施方式中,包含序列R₂-CNGCCACC (SEQ ID NO:2)的RNA分子是这样的RNA分子,其中SEQ ID NO:2的位置2处的核苷酸N是选自U、G或C的核苷酸并且其中核苷酸N不是A。

[0197] 在另一优选实施方式中,在SEQ ID NO:2的位置2处的所述核苷酸N是U。

[0198] 在优选的实施方式中,本发明的RNA分子是这样的RNA分子,其中直接跟在起始密码子下游的核苷酸不是核苷酸G。在另一优选的实施方式中,本发明的RNA分子是这样的RNA分子,其中直接跟在起始密码子下游的核苷酸是选自A、U和C的核苷酸。

[0199] 在甚至更优选的实施方式中,本发明的RNA分子是包含上文限定的组件(b1)的RNA分子,其中所述组件(b1)是这样的序列:其中SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代并且其中直接跟在起始密码子下游的核苷酸是选自A、U和C的核苷酸。

[0200] 在另一甚至更优选的实施方式中,本发明的RNA分子是包含上文限定的组件(b2)的RNA分子,其中所述组件(b2)是这样的序列:其中SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代并且其中直接跟在起始密码子下游的核苷酸是选自A、U和C的核苷酸。

[0201] 如上所述,Kozak共有序列(gcc)gccRccAUGG尤其可以是关于位置-3处(即,从起始密码子AUG上游第3个核苷酸)由“R”表示的核苷酸的变体,只要该位置是嘌呤(即,腺嘌呤或鸟嘌呤)。在上述UTR中,与该位置对应的核苷酸限定为“A”。然而,本发明还涉及包含在该位置具有“G”的相应UTR的RNA分子。

[0202] 因此,在优选的实施方式中,本发明的RNA分子含有上文项目(b1)中限定的UTR,其中在所述(b1)的UTR序列中,SEQ ID NO:1的位置5处的A被G取代;或其中在所述(b2)的UTR序列中,SEQ ID NO:2的位置6处的A被G取代。

[0203] 如上所述,本发明的RNA分子还可携带聚腺苷酸尾。如本文使用的,聚腺苷酸尾涉及位于RNA的3'端的腺嘌呤核苷酸序列。聚腺苷酸尾通常通过被称为多聚腺苷化作用的过

程添加到RNA的3'端。因此,本发明涉及任意上述RNA,其中RNA分子在3'端包含聚腺苷酸尾。

[0204] 不具体限制聚腺苷酸尾的长度。但是,在优选的实施方式中,本发明的RNA分子在3'端包含聚腺苷酸尾,其中聚腺苷酸尾的长度为至少50、60、70、80、90、100或110个核苷酸。在更优选的实施方式中,本发明的RNA分子在3'端包含聚腺苷酸尾,其中聚腺苷酸尾的长度为至少120个核苷酸。在其它优选的实施方式中,本发明的RNA分子在3'端包含聚腺苷酸尾,其中聚腺苷酸尾的长度为至少150、200、250、300、350、400、500、600、700、800、900或1000个核苷酸。

[0205] 如果本发明的RNA分子是通过本文进一步在下文描述的体外转录方法产生,聚腺苷酸尾位于RNA的3'端、与RNA分子的3'端的UTR相邻,同时在体外转录聚腺苷酸尾下游前将携带本发明的RNA分子的质粒线性化以确保体外转录的RNA分子含有所述聚腺苷酸尾。

[0206] 如上所述,本发明的RNA分子可以以组件(a)和(b)的融合RNA序列的形式存在,即,(融合)RNA分子,其由通过结合编码所述组件的至少两个核苷酸序列产生的杂合基因的转录形成。一般地,这是通过将cDNA克隆到允许整个RNA分子转录的表达载体中来实现。已知用于产生融合构建物的各种方法包括核酸合成、杂交和/或扩增以产生“编码”本发明的RNA分子的合成双链核酸分子。这种双链核酸分子(即,DNA分子)在一条链(即,在编码链或有义链上)上携带与本发明的RNA分子对应的DNA序列,并因此,“编码”本发明的RNA分子。换言之,这种双链核酸/DNA分子在一条链上包含与上文限定的转录的本发明的RNA分子对应的遗传信息。在本发明的上下文中术语“编码(coding)”或“编码(encoding)”不仅以其常规含义使用,即,涉及编码蛋白质的基因的DNA(并且,因此,可翻译成多肽或蛋白质氨基酸序列的遗传信息)。然而,就本发明而言,在这样的构建物中:其中编码组件(a)和(b)的单独的DNA序列“融合”或连接到单一(嵌合)DNA分子中,该构建物还包含不翻译成蛋白质的构件(即,组件(b))。不过,与组件(b)对应的DNA序列为5'UTR'结构提供信息,即“密码”,并因此,在本发明中术语“编码”还涉及可被表达的(即,转录的)UTR的遗传信息,如果例如存在于双链核酸分子中的话。因此,本发明的上下文中的术语“编码”虽然通常仅用于涉及蛋白质的编码/表达,但其将以这种方式理解:核酸分子可转录成携带编码蛋白质或多肽(即,组件(a))的部分和“编码”UTR(即,组件(b))的部分的相应RNA分子,其中在表达时后者表示最终产物,因为UTR不翻译成蛋白质或多肽。这种双链核酸可通过标准分子生物学技术插入表达载体(参见,例如Sambrook等,Molecular Cloning,A laboratory manual,第二版,1989)。在本发明的含义中术语“载体”诸如“表达载体”或“克隆载体”理解为环形、双链单元的DNA,其在细胞内独立于染色体DNA进行复制并被用作载体以将遗传材料携带到细胞中,在细胞中其可复制和/或表达(即,转录成RNA并翻译成氨基酸序列)。含有外源DNA的载体被称为重组DNA。载体本身通常是这样的DNA序列:其典型地由插入物(即,不翻译成蛋白质的组件(b)和编码区组件(a))和充当载体“主链”的较大序列构成。在本发明的含义中质粒最常发现于细菌中并用于重组DNA研究中以将基因在细胞之间转移,并因此是用于本发明的含义中的“载体”的亚群。

[0207] 因此,本发明还涉及编码本发明的RNA分子的核酸分子。

[0208] 核酸是,例如DNA,编码本发明的RNA分子的两个主要组件(即,组件(a)和组件(b))。本发明的上述核酸分子优选地是重组核酸分子。本发明的核酸分子可以是合成的或是半合成的。

[0209] 对本领域技术人员显而易见的是可将进一步的调控序列添加到编码RNA分子的本发明的核酸分子。例如,可以利用转录增强子和/或允许诱导表达的序列。适合的诱导系统是例如调控四环素的基因表达——例如由Gossen和Bujard, Proc. Natl. Acad. Sci. USA 89 (1992), 5547-5551)和Gossen, Trends Biotech. 12 (1994), 58-62描述的,或是地塞米松诱导的基因表达系统——例如由Crook, EMBO J. 8 (1989), 513-519描述的。

[0210] 本发明还涉及载体,优选地表达载体,其包含本发明的核酸分子。

[0211] 关于包含编码本发明的RNA分子的核酸分子的载体,加上必要的修改,同样适用,如在上文限定的包含本发明的DNA分子的载体的上下文中在上文列出的。

[0212] 本发明还涉及包含本发明的载体的宿主细胞。因此,本发明涉及用本发明的载体或携带本发明的载体的非人类宿主转染或转化的宿主,即涉及用根据本发明的核酸分子或用包含这种核酸分子的载体进行基因修饰的宿主细胞或宿主。

[0213] 关于包括包含编码本发明的RNA分子的核酸分子的载体的宿主细胞,加上必要的修改,同样适用,如在上文限定的包括包含本发明的DNA分子的载体的宿主细胞的上下文中在上文列出的。

[0214] 本发明还涉及产生本发明的RNA分子的方法:通过在培养基中培养携带编码本发明的单独组件或本发明的整个RNA分子的表达载体的宿主细胞,并从宿主细胞或培养基中回收RNA分子。本发明还可涉及产生本发明的RNA分子的方法,包括培养本发明的宿主细胞并任选地从培养物中回收RNA分子。

[0215] 回收和/或后续纯化本发明RNA分子的方法是本领域技术人员已知的。

[0216] 本发明还涉及通过本领域技术人员已知的方法在体外反应中产生本发明的RNA分子的方法。更具体地,本发明的RNA分子可利用体外转录系统在体外产生。体外转录系统通常是已知的并且通常需要如上概括的含有“编码”组件(b)和组件(a)的DNA序列的纯化的线性DNA模板,其中所述DNA序列处于适当启动子的控制下。而且,体外转录系统通常还需要核糖核苷酸三磷酸盐、包含DTT和镁离子的缓冲系统、和为将DNA序列体外转录成本发明的RNA分子提供酶活性的适当的RNA聚合酶。

[0217] 通常用于利用体外转录产生RNA分子的方法是本领域技术人员熟知的并且是在,例如, Mol. Biol. 703 (2011): 29-41中描述的方法。

[0218] 如上所述,假使本发明的RNA分子是通过本文进一步在下文描述的体外转录方法产生的,则上述聚腺苷酸尾可能是本发明的RNA分子的部分(并且不必最初就位于克隆载体上)并位于RNA的3'端,例如与RNA分子的3'端的UTR相邻。假使本发明的RNA分子是通过体外转录方法产生的,在体外转录前将携带本发明的RNA分子的质粒在聚腺苷酸尾下游线性化以便确保体外转录的RNA分子含有所述聚腺苷酸尾。

[0219] 可选地,本发明的RNA分子还可以化学合成,例如,利用固相支持和标准技术在自动化核苷酸序列合成仪上通过常规化学合成。

[0220] 本发明还涉及通过本领域技术人员已知和如上概括的方法在体外反应中产生本发明的RNA分子并从反应中回收RNA分子的方法。

[0221] 回收和/或随后纯化本发明的RNA分子的方法是本领域技术人员已知的。

[0222] 本发明的RNA分子可容易地用于本领域已知的体外翻译系统中用于高效表达组件(a)的编码区编码的任何期望多肽或蛋白质。

[0223] 体外翻译系统是本领域已知的并可直接与本发明的RNA分子一起使用。可选地,这些体外翻译系统可与上述体外转录系统结合使用。用于体外转录和/或体外翻译的相应的无细胞系统是已知的和可获得的。用于蛋白质合成的这些无细胞系统(也被称为体外蛋白质合成或缩写为CFPS),允许使用生物学机构而不使用活细胞来表达/生成多肽或蛋白质。在这些系统中,体外蛋白质合成环境不受细胞壁或维持细胞生存力所必需的稳态条件的限制并能够直接进入(access)和控制翻译环境,这对包括以下的多种应用是有利的:蛋白质生成的优化、蛋白质复合物的优化、研究蛋白质合成、并入非天然氨基酸、高通量筛选、和合成生物学。无细胞反应的常见构件包括细胞提取物、能量来源、氨基酸供给、辅因子诸如镁、以及编码期望多肽或蛋白质的DNA或RNA。细胞提取物可通过裂解感兴趣细胞并将细胞壁、DNA基因组、和其它杂物(碎片,debris)离心出去而获得。剩余物是必需的细胞机构,包括核糖体、氨酰基-tRNA合成酶、翻译起始和延长因子、核酸酶等。在用于由DNA起始合成多肽或蛋白质的无细胞系统中(即,在包括体外转录和体外翻译的步骤的系统中),通常使用两种DNA,即,质粒或线性表达模板(LET)。在用于由RNA起始合成多肽或蛋白质的无细胞系统中(即,在仅包括体外翻译步骤的系统中),可直接使用RNA。这些体外无细胞反应需要能量来源(其通常通过含有所需能量来源的单独混合物来提供),连同供给添加到用于反应的提取物的氨基酸。常见的能量来源是磷酸烯醇丙酮酸(phosphoenol pyruvate)、乙酰磷酸、和磷酸肌酸。通常使用的常见的细胞提取物由大肠杆菌(ECE)、兔网织红细胞(RRL)、小麦胚芽(WGE)、和昆虫细胞(ICE)制得。所有这些提取物都可商业获得。

[0224] 因此,本发明还涉及本发明的RNA分子用于体外翻译所述RNA分子中含有的编码区编码的期望多肽或蛋白质的应用。

[0225] 关于本发明的RNA分子的这种应用的优选实施方式,加上必要的修改,同样适用,如上文限定的RNA分子的上下文中在上面列出的。

[0226] 上文限定的RNA分子具体用于医疗环境(医疗设施,medical settings)并用于治疗某些疾病,和特别地,用于基于RNA的疗法。因此,本发明还涉及包含本发明的RNA分子、本发明的核酸分子、本发明的载体或本发明的宿主细胞以及任选地药学上可接受的载体的药物组合物。

[0227] 本文所用的术语“治疗”等总体上意指获得期望的药理和/或生理效果。因此,本发明的治疗可涉及治疗某种疾病的(急性)状态但在完全或部分预防疾病或其症状方面也可涉及预防性治疗。优选地,术语“治疗”将理解为就部分或完全治愈疾病和/或归因于疾病的副作用和/或症状而言是治疗性的。在此方面,“急性”意指受试者显示了疾病的症状。换言之,待治疗的受试者实际需要治疗,并且在本发明的上下文中术语“急性治疗”涉及在疾病发作或疾病爆发后实际治疗疾病所采取的措施。治疗还可以是预防性(prophylactic)或防止性(预防,preventive)治疗,即,针对疾病预防采取措施,例如,以便防止感染和/或疾病发作。

[0228] 本发明的药物组合物可通过技术人员已知的大范围的各种施用形式来施用。施用可以是全身地、局部地、经口、通过气溶胶,包括但不限于片剂、针注射、使用吸入器、霜剂、泡沫、凝胶、洗剂和膏剂。

[0229] 如上所述,本发明涉及药物组合物,包括有效量的根据上文的本发明的RNA分子(或核酸分子、载体或宿主细胞)和至少一种药学上可接受的赋形剂或载体。

[0230] 赋形剂或载体是与活性组分——即,本发明的RNA分子(或核酸分子、载体或宿主细胞)一同配制的非活性物质,用于填充(增强,bulking-up)含有有效力的活性组分的制剂的目的。赋形剂经常被称为“填充剂(bulking agent)”、“填料(filler)”或“稀释剂”。当生产剂型时,填充允许方便且准确地分配药物物质。它们还可用于各种治疗增强的目的,例如促进药物吸收或溶解,或其它药物代谢动力学考虑。除了有助于体外稳定性诸如防止在(经过,over)期望的货架期而变性之外,赋形剂还可用于生产过程,例如关于通过促进粉末流动性或非粘性特性从而有助于处理有关的活性物质。适当赋形剂的选择还取决于施用路径和剂型,以及活性组分和其它因素。

[0231] 因此,包含有效量的本发明的RNA分子(或核酸分子、载体或宿主细胞)的药物组合物可以是固态、液态或气态形式,并且尤其可以是粉末(一个或多个)、片剂(一个或多个)、溶液(一个或多个)或气溶胶(一个或多个)形式。优选地,所述药物组合物任选地包含药学上可接受的载体和/或稀释剂。

[0232] 适合的药学载体、赋形剂和/或稀释剂的实例是本领域熟知的并且包括磷酸盐缓冲液、水、乳液,如油/水乳液、各种类型的湿润剂、无菌溶液等。包含这种载体的组合物可通过熟知的常规方法配制。这些药物组合物可以以适合的剂量,即,以技术人员通过本领域已知方法可容易地确定的“有效量”施用给受试者。剂量方案将由主治医师和临床因素来确定。如医疗领域熟知的,用于任意一个患者的剂量取决于多种因素,包括患者或受试者的尺寸(size)、身体表面积、年龄、待施用的具体化合物、性别、施用的时间和路径、总体的健康、和同时施用的其它药物。

[0233] 因此,优选地,本发明的RNA分子(或核酸分子、载体或宿主细胞)包括在有效量中。术语“有效量”指代在药物组合物将施用的受试者中足以引起可检测的治疗响应的量。根据上文,药物组合物中本发明的RNA分子(或核酸分子、载体或宿主细胞)的含量没有被限制,只要其对上述治疗有用,但优选地每总组合物按重量计含有0.000001-10%。此外,本文所述RNA分子(或核酸分子、载体或宿主细胞)优选地用于载体中。总体上,适当量的药学上可接受的盐被用于载体中从而使组合物等渗。载体的实例包括但不限于盐水、林格氏溶液和葡萄糖溶液。优选地,可接受的赋形剂、载体、或稳定剂在所用的剂量和浓度上是无毒的,包括缓冲液诸如柠檬酸盐、磷酸盐、和其它有机酸;形成盐的反离子,如钠和钾;低分子量(>10个氨基酸残基)多肽;蛋白质,例如血清白蛋白、或明胶;亲水聚合物,例如聚乙烯吡咯烷酮;氨基酸诸如组氨酸、谷氨酰胺、赖氨酸、天冬酰胺、精氨酸、或甘氨酸;碳水化合物,包括葡萄糖、甘露糖、或糊精;单糖;二糖;其它糖,例如蔗糖、甘露醇、海藻糖或山梨醇;螯合剂,例如EDTA;非离子表面活性剂,例如吐温、普兰尼克(Pluronic)或聚乙二醇;抗氧化剂,包括甲硫氨酸、抗坏血酸和生育酚;和/或防腐剂,例如十八烷基二甲基苄基氯化铵;六甲氯铵;氯化苄甲乙氧铵;苯酚、丁醇或苄醇;烷基对羟基本甲酸酯,例如甲基或丙基对羟基本甲酸酯;邻苯二酚;间苯二酚;环己醇;3-戊醇;和间甲酚。Remington's Pharmaceutical Sciences,第17版,1985,Mack Publishing Co.中更详细地描述了适合的载体及其制剂

[0234] 可通过周期性评估来监测治疗进展。本发明的RNA分子(或核酸分子、载体或宿主细胞)或本发明的药物组合物可以在无菌水溶液或非水溶液、悬浮液、和乳液以及霜剂和栓剂中。非水溶剂的实例是丙二醇、聚乙二醇、植物油诸如橄榄油、以及有机酯诸如油酸乙酯。

水性载体包括水、醇溶液/水溶液、乳液或悬浮液,包括盐水和缓冲介质。还可以存在防腐剂和其它添加剂,诸如例如,抗微生物剂、抗氧化剂、螯合剂、和惰性气体及类似物。并且,本发明的药物组合物根据药物组合物的预期用途可包含进一步的试剂。所述试剂可以是,例如,聚氧乙烯失水山梨醇单月桂酸酯在市场上可以以商业名称吐温获得,适用于药物组合物的预期用途的丙二醇、EDTA、柠檬酸盐、蔗糖以及其它试剂是本领域技术人员熟知的。

[0235] 根据本发明,术语“药物组合物”涉及施用给患者——优选地,人类患者——的组合物。

[0236] 本发明的药物组合物可用于基于RNA的疗法。如上所述,包含“编码多肽的编码区”的本发明的RNA分子可用于基于RNA的疗法,其中“编码多肽的编码区”编码具有治疗作用或预防作用的治疗活性或药理学活性的多肽或蛋白质。因此,在优选的实施方式中,本发明的药物组合物可用于基于RNA的疗法以治疗或预防上表1中陈述的疾病。因此,根据本发明的基于RNA的疗法可用于治疗或预防上表1中陈述的疾病。

[0237] 因此,本发明的药物组合物可用于以下情况的基于RNA的疗法中:如果其中上述表1中所述基因缺陷导致疾病,该疾病然后可用本发明的RNA分子通过转录物替换疗法/酶替换疗法来治疗或预防,其中RNA分子包含编码补偿所公开的缺陷基因的完整版蛋白质或其功能片段的“多肽的编码区”。在特别优选的实施方式中,本发明的药物组合物可用于基于RNA的疗法中以治疗或预防溶酶体疾病如高歇病、法布里病、MPS I、MPS II(亨特综合征)、MPS VI以及糖原贮积病诸如例如I型糖原贮积病(冯·基尔克氏病(von Giercke's disease))、II型(蓬珀病)、III型(柯里病)、IV型(安德森病)、V型(麦卡德尔病)、VI型(赫斯病)、VII型(Tauri's病)、VIII型、IX型、X型、XI型(Fanconi-Bickel综合征)、XII型、或O型。转录物替换疗法/酶替换疗法有益地不影响基本的遗传缺失,但增加了患者缺乏的酶的浓度。作为实例,在蓬珀病中,转录物替换疗法/酶替换疗法替代了缺陷的溶酶体酶——酸性 α -葡萄糖苷酶(GAA)。

[0238] 在其它优选的实施方式中,本发明的药物组合物可用于根据本发明的基于RNA的疗法中,其中“编码多肽的编码区”编码具有治疗作用或预防作用的治疗活性或药理学活性的多肽、蛋白质或肽,其中所述多肽、蛋白质或肽选自表1中概括的基因所编码的组。

[0239] 在其它优选的实施方式中,根据本发明的基于RNA的疗法可用于治疗癌症、心血管疾病、病毒感染、免疫功能失调、自身免疫疾病、神经障碍、遗传代谢紊乱或遗传疾病或其中细胞中产生的蛋白质或蛋白质片段对患者可具有有益效果的任何疾病。癌症的实例包括头颈癌、乳腺癌、肾癌、膀胱癌、肺癌、前列腺癌、骨癌、脑癌、宫颈癌、肛门癌、结肠癌、结直肠癌、阑尾癌、眼癌、胃癌、白血病、淋巴瘤、肝癌、皮肤癌、卵巢癌、阴茎癌、胰腺癌、睾丸癌、甲状腺癌、阴道癌、外阴癌、子宫内膜癌、心脏癌和肉瘤。

[0240] 心血管疾病的实例包括动脉粥样硬化、冠心病、肺心疾病(pulmonary heart disease)和心肌病。

[0241] 免疫功能失调和自身免疫疾病的实例包括,但不限于,风湿病、多发性硬化症和哮喘。

[0242] 病毒感染的实例包括,但不限于,具有人类免疫缺陷病毒、单纯疱疹病毒、人乳头状瘤病毒以及乙型肝炎病毒和丙型肝炎病毒的感染。

[0243] 神经障碍的实例包括,但不限于,帕金森病、多发性硬化、和痴呆。

[0244] 遗传代谢紊乱的实例包括,但不限于,高歇病和苯丙酮尿症。

[0245] 本发明还涉及基于RNA的疗法的方法。因此,本发明涉及通过基于RNA的疗法治疗疾病诸如癌症、心血管疾病、病毒感染、免疫功能失调、自身免疫疾病、神经障碍、遗传代谢紊乱或遗传疾病的方法。关于治疗方法的优选实施方式,加上必要的修改,同样适用,如用于上文限定的基于RNA的疗法中的RNA分子或药物组合物的上下文中如上列出的。

[0246] 在本发明中,在优选的实施方式中,受试者是哺乳动物诸如狗、猫、猪、牛、羊、马、啮齿动物,例如,大鼠、小鼠、和豚鼠、或灵长类动物,例如,大猩猩、黑猩猩、和人类。在最优选的实施方式中,受试者是人类。

[0247] 如上所述,上文限定的RNA分子具体用于医疗环境以及用于治疗某种疾病,并具体地,用于基于RNA的疗法中。因此,本发明还涉及包含本发明的RNA分子、核酸分子、载体或宿主细胞以及任选地药学上可接受的载体的药物组合物。

[0248] 但是,在RNA疗法中,通常期望在某阶段沉默RNA分子的效应。

[0249] 这可以,例如,通过利用与本发明的UTR序列互补的核酸链通过利用RNAi (RNA干扰) 机制完成。事实上,本发明的最小UTR的小尺寸使得该方案可行,因为这些UTR不形成二级或三级结构并且它们不存在于正常细胞中。因此,这种UTR序列的互补链可利用本发明的药物组合物在治疗上述疾病或在上述基于RNA的疗法之后有益地用于医疗环境,进而沉默本发明的治疗性RNA分子。

[0250] 因此,在本发明的上下文中还设想了RNAi方案用于制备使本发明的治疗性RNA分子的效应沉默的药物组合物。

[0251] 术语“RNA干扰”或“抑制RNA”(RNAi/iRNA)描述了使用双链RNA靶向特定mRNA用以降解,进而沉默其表达。优选的抑制RNA分子可选自双链RNA(dsRNA)、RNAi、siRNA、shRNA和stRNA。匹配基因序列的dsRNA在体外合成并导入细胞。dsRNA还可以以表达有义和反义方向的靶基因序列的载体的形式导入细胞,例如以发夹mRNA的形式。有义和反义序列还可由单独的载体表达,借此单独的反义和有义分子一经其表达就形成双链RNA。本领域已知在某些场合中由于细胞中的内部扩增机制,有义方向的序列的表达或甚至启动子序列的表达足以导致dsRNA以及随后引起siRNA。因此,将根据本发明使用导致编码区编码的多肽或蛋白质的活性降低的所有手段和方法。例如,有义构建物、反义构建物、发夹构建物、有义和反义分子和其组合可用于产生/导入这些siRNA。将dsRNA注入(feeds into)自然但仅部分理解的过程,该过程包含高度保守的核酸酶切丁酶(nuclease dicer),其将dsRNA前体分子切割成短的干扰RNA(siRNA)。WO 02/055693,Wei (2000) Dev. Biol. 15:239-255;La Count (2000) Biochem. Paras. 111:67-76;Baker (2000) Curr. Biol. 10:1071-1074;Svoboda (2000) Development 127:4147-4156或Marie (2000) Curr. Biol. 10:289-292中尤其描述了siRNA(一个或多个)的产生和制备以及抑制靶基因的表达的方法。然后这些siRNA构建了RNA诱导沉默复合体(RISC)的序列特异性部分——破坏与沉默触发子(silencing trigger)同源的信使RNA的多复合核酸酶)。Elbashir (2001) EMBO J. 20:6877-6888显示21个核苷酸的RNA的双链体(duplexes)可用于细胞培养以干扰哺乳动物细胞中的基因表达。已知哺乳动物细胞中siRNA非常有效地介导RNAi,但稳定细胞系或非人类转基因动物的产生受到限制。然而,可利用新一代载体以便稳定表达,例如短的发夹RNA(shRNA)。Brummelkamp (2002) Science 296:550-553中尤其显示了哺乳动物细胞中siRNA的稳定表达。Paul (2002)

Nat. Biotechnol. 20:505-508还记录了人类细胞中小的干扰RNA的有效表达。Yu (2002) PNAS 99:6047-6052还显示了通过在哺乳细胞中表达短的干扰RNA和发夹RNA的RNA干扰。用于基因沉默的shRNA方案是本领域熟知的并且可包括利用st(小时序small temporal) RNA; 尤其参见Paddison (2002) Genes Dev. 16:948-958。这些方案可以是基于载体的, 例如pSUPER载体, 或者可使用RNApol III载体, 尤其如在Yu (2002), loc. cit.; Miyagishi (2002), loc. cit. 或Brummelkamp (2002), loc. cit. 中示例的。设想本发明的调控序列以与基于pSUPER或RNApol III载体的系统相似的方式(fashion)使用。

[0252] 推断并构建siRNA的方法是本领域已知的并被描述在Elbashir (2002) Methods 26:199-213中, 在siRNA的商业供应商的互联网网站上, 例如Qiagen GmbH (<https://www1.qiagen.com/GeneGlobe/Default.aspx>); Dharmacon (www.dharmacon.com); Xeragon Inc. (<http://www.dharmacon.com/Default.aspx>), 和Ambion (www.ambion.com), 或在TomTuschl的研究小组的网站上 (<http://www.rockefeller.edu/labheads/tuschl/sirna.html>)。另外, 从给定的mRNA序列推断siRNA的程序可在线获得(例如http://www.ambion.com/techlib/misc/siRNA_finder.html或<http://katahdin.cshl.org:9331/RNAi/html/rnai.html>)。2-nt的3'突出端(延伸物, overhang)中的尿苷残基可被2'脱氧胸苷取代而不丧失活性, 其显著降低RNA合成的成本并且还可在应用于哺乳动物细胞时增强siRNA双链体的抗性(Elbashir (2001) loc. cit.)。siRNA还可以利用T7或其它RNA聚合酶以酶促方式合成(Donze (2002) Nucleic Acids Res 30:e46)。介导有效RNA干扰(esRNA)的短RNA双链体也可通过用大肠杆菌RNase III水解来产生(Yang (2002) PNAS 99:9942-9947)。并且, 已开发表达载体以在真核细胞中表达小发夹RNA环连接的双链siRNA(例如(Brummelkamp (2002) Science 296:550-553))。所有这些构建物可在上文所命名的程序的帮助下开发。另外, 并入序列分析程序或单独售卖的可商业获得的序列预测工具, 例如www.oligoEngine.com (Seattle, WA) 提供的siRNA设计工具(Design Tool)可用于siRNA序列预测。

[0253] 因此, 特异性干扰RNA可根据本发明用作本发明RNA分子编码区编码的多肽或蛋白质的表达和/或功能的拮抗剂/沉默剂(silencers)。这些siRNA通过互补/反义链和有义链形成, 借此反义/有义链优选地包含至少10个, 更优选地至少12个, 更优选地至少14个, 更优选地至少16个, 更优选地至少18个, 更优选地至少19、20、21或22个核苷酸。在甚至更优选的实施方式中, 反义/有义链优选地包含25个或更多个核苷酸。

[0254] 如上所述, 制备根据本发明的将使用的siRNA的方法是本领域熟知的。基于本文提供的教导, 本领域技术人员有能力不仅容易地制备这种siRNA而且能够评价siRNA是否能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。本文设想上述siRNA导致携带编码多肽或蛋白质的编码区和UTR组件的本发明RNA分子的降解, 并因此导致本发明RNA分子的编码区编码的多肽或蛋白质的多肽/蛋白质水平降低。

[0255] 因此, 本发明涉及与如上文所述的本发明的UTR互补的RNA分子。

[0256] 在优选的实施方式中, 与本发明的UTR互补的所述RNA分子包括序列CAUGGUGCGUCUCCC (SEQ ID NO:11)或这样的序列, 其SEQ ID NO:11相比显示1至4个取代。

[0257] 在另一优选实施方式中, 与本发明的UTR互补的所述RNA分子包括序列CAUGGUGCNGUCUCCC (SEQ ID NO:12), 其中SEQ ID NO:12的位置10处的核苷酸N是选自U、G、

C或A的核苷酸,或这样的序列,其与SEQ ID NO:12相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0258] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUUGGCGUCUCCC (SEQ ID NO:13),或这样的序列,其与SEQ ID NO:13相比显示1至4个取代。

[0259] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUUGGCGUCUCCC (SEQ ID NO:14),其中SEQ ID NO:14的位置10处的核苷酸N是选自U、G、C或A的核苷酸同时更优选为A,或这样的序列,其与SEQ ID NO:14相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0260] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGCGGCGUCUCCC (SEQ ID NO:15),或这样的序列,其与SEQ ID NO:15相比显示1至4个取代。

[0261] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGCGGCGUCUCCC (SEQ ID NO:16),其中SEQ ID NO:16的位置10处的核苷酸N是选自U、G、C或A的核苷酸,或这样的序列,其与SEQ ID NO:16相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0262] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUCGGCGUCUCCC (SEQ ID NO:17),或这样的序列,其与SEQ ID NO:17相比显示1至4个取代。

[0263] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUCGGCGUCUCCC (SEQ ID NO:18),其中SEQ ID NO:18的位置10处的核苷酸N是选自U、G、C或A的核苷酸同时更优选为A,或这样的序列,其与SEQ ID NO:18相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0264] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGUGGCGUCCC (SEQ ID NO:19),或这样的序列,其与SEQ ID NO:19相比显示1至4个取代。

[0265] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGUGGCGUCCC (SEQ ID NO:20),其中SEQ ID NO:20的位置10处的核苷酸N是选自U、G、C或A的核苷酸,或这样的序列,其与SEQ ID NO:20相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0266] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUUGGCGUCCC (SEQ ID NO:21),或这样的序列,其与SEQ ID NO:21相比显示1至4个取代。

[0267] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUUGGCGUCCC (SEQ ID NO:22),其中SEQ ID NO:22的位置10处的核苷酸N是选自U、G、C或A的核苷酸同时更优选为A,或这样的序列,其与SEQ ID NO:22相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0268] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGCGGCGUCCC (SEQ ID NO:23),或这样的序列,其与SEQ ID NO:23相比显示1至4个取代。

[0269] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGCGGCGUCCC (SEQ ID NO:24),其中SEQ ID NO:24的位置10处的核苷酸N是选自U、G、C

或A的核苷酸,或这样的序列,其与SEQ ID NO:24相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0270] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUCGGCGUCCC (SEQ ID NO:25),或这样的序列,其与SEQ ID NO:25相比显示1至4个取代。

[0271] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUCGGCNGUCCC (SEQ ID NO:26),其中SEQ ID NO:26的位置10处的核苷酸N是选自U、G、C或A的核苷酸同时更优选为A,或这样的序列,其与SEQ ID NO:26相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0272] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGUGGCGCUUC (SEQ ID NO:27),或这样的序列,其与SEQ ID NO:27相比显示1至4个取代。

[0273] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGUGGCNGCUUC (SEQ ID NO:28),其中SEQ ID NO:28的位置10处的核苷酸N是选自U、G、C或A的核苷酸,或这样的序列,其与SEQ ID NO:28相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0274] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUUGGCGCUUC (SEQ ID NO:29),或这样的序列,其与SEQ ID NO:29相比显示1至4个取代。

[0275] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUUGGCNGCUUC (SEQ ID NO:30),其中SEQ ID NO:30的位置10处的核苷酸N是选自U、G、C或A的核苷酸同时更优选为A,或这样的序列,其与SEQ ID NO:30相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0276] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGCGGCGCUUC (SEQ ID NO:31),或这样的序列,其与SEQ ID NO:31相比显示1至4个取代。

[0277] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUGGCGGCNGCUUC (SEQ ID NO:32),其中SEQ ID NO:32的位置10处的核苷酸N是选自U、G、C或A的核苷酸,或这样的序列,其与SEQ ID NO:32相比显示1到4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0278] 在优选的实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUCGGCGCUUC (SEQ ID NO:33),或这样的序列,其与SEQ ID NO:33相比显示1至4个取代。

[0279] 在另一优选实施方式中,与本发明的UTR互补的所述RNA分子包括序列CAUCUCGGCNGCUUC (SEQ ID NO:34),其中SEQ ID NO:34的位置10处的核苷酸N是选自U、G、C或A的核苷酸同时更优选为A,或这样的序列,其与SEQ ID NO:34相比显示1至4个取代并且能够拮抗/抑制/沉默本发明RNA分子的编码区编码的多肽或蛋白质。

[0280] 在另一优选实施方式中,本发明涉及选自SEQ ID NO:11至34的RNA分子,其在5'端具有另外的核苷酸(一个或多个)——延伸超过与起始密码子互补的三联体,并且其与本发明RNA分子编码区编码的期望多肽或蛋白质的序列互补。优选地,包含与本发明的UTR序列互补的上述序列的互补序列(即,选自SEQ ID NO:11至34的RNA分子)优选地包含至少15个、更优选地至少16个、更优选地至少17个、更优选地至少18个、更优选地至少19个、更优选地至少20、21、22、23或24个核苷酸。在甚至更优选的实施方式中,这些序列包含25、30、35、40或更多个核苷酸。可以期望增加5'端的长度以便增加互补序列的特异性进而防止不期望的

副作用。

[0281] 在另一优选实施方式中,本发明不仅涉及任意上述RNA分子而且涉及选自SEQ ID NO:11到34的RNA分子——其包含与上述RNA分子多达5%、10%、20%或30%的错配。并且, RNA分子可如上文所述进行化学修饰。

[0282] 本发明还涉及包含本发明的DNA分子、本发明的RNA分子、本发明的核酸分子、本发明的载体或本发明的宿主细胞的试剂盒。关于优选实施方式,加上必要的修改,同样适用,如根据本发明的DNA分子、RNA分子、核酸分子、载体或宿主细胞的上下文中列出的。有利地,本发明的试剂盒进一步包含,任选地缓冲液(一种或多种)、存储溶液和/或执行上述和下述应用和方法所需的剩余试剂或材料。并且,本发明的试剂盒的部分可单独包装在小瓶或瓶子中或与容器或多容器单元组合进行包装。本发明的试剂盒可有利地使用,尤其用于实施本发明的方法或用于制备本发明的RNA分子并可用于本文提到的多种应用中,例如,上文和下文概括的应用中。可包括在试剂盒中的另外的构件是给针对其用途而使用试剂盒的人的说明书。试剂盒的制造优选地遵循本领域技术人员已知的标准程序。

[0283] 本发明还涉及如上文所述的UTR的应用,用于将RNA分子的编码区翻译成所述编码区编码的多肽或蛋白质。

[0284] 在更优选的实施方式中,本发明还涉及如上文所述的UTR的应用,用于提高将RNA分子的编码区翻译成所述编码所编码的多肽或蛋白质的效率。

[0285] 关于应用的优选实施方式,加上必要的修改,同样适用,如本发明的RNA分子的上下文中列出的。

[0286] 在优选的实施方式中,本发明涉及由以下项目1到20表征的下述:

[0287] 1. DNA分子,其可转录成mRNA,包含具有下列元件的一条链:

[0288] (a) 编码区,在其5'端包括起始密码子,编码多肽;和

[0289] (b) 直接在所述编码序列上游的的序列,其选自:

[0290] (b1) R₁-CGCCACC (SEQ ID NO:1);

[0291] 或这样的序列,其中在所述序列中,SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代;和

[0292] (b2) R₁-CNGCCACC (SEQ ID NO:2),其中SEQ ID NO:2的位置2处的核苷酸N是选自T、G、C或A的核苷酸;

[0293] 或这样的序列,其中在所述序列中,SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代,

[0294] 其中R₁是启动子,其被DNA依赖性RNA聚合酶识别;

[0295] 或包含互补链。

[0296] 2. 根据项目1的DNA分子,其中被DNA依赖性RNA聚合酶识别的启动子选自:

[0297] (i) TAATACGACTCACTATAGGGAGA (SEQ ID NO:3) 或这样的序列,其与SEQ ID NO:3相比显示1至6个取代并且其被T7DNA依赖性RNA聚合酶识别;

[0298] (ii) AATTAACCCTCACTAAAGGGAGA (SEQ ID NO:4) 或这样的序列,其与SEQ ID NO:4相比显示1至6个取代并且其被T3DNA依赖性RNA聚合酶识别;

[0299] (iii) ATTTAGGTGACACTATAGAAG (SEQ ID NO:5) 或这样的序列,其与SEQ ID NO:5相比显示1至6个取代并且其被SP6DNA依赖性RNA聚合酶识别;和

- [0300] (iv) AATTAGGGCACA CTATAGGGA (SEQ ID NO:6) 或这样的序列,其与SEQ ID NO:6相比显示1至6个取代并且其被K11DNA依赖性RNA聚合酶识别。
- [0301] 3. 根据项目1或2的DNA分子,其中SEQ ID NO:2的位置2处的核苷酸N是选自T、G或C的核苷酸并且其中核苷酸N不是A。
- [0302] 4. 根据项目3的DNA分子,其中SEQ ID NO:2的位置2处的所述核苷酸N是T。
- [0303] 5. 载体,其包含项目4的DNA分子。
- [0304] 6. 宿主细胞,其包含项目5的载体。
- [0305] 7. 组合物,其包含:
- [0306] 根据项目1到4中任一项的DNA分子、根据项目5的载体或根据项目6的宿主细胞。
- [0307] 8. RNA分子,其包含
- [0308] (a) 编码区,在其5'端包括起始密码子,编码多肽;和
- [0309] (b) 直接在所述编码序列上游的的UTR,其选自:
- [0310] (b1) 序列R₂-CGCCACC (SEQ ID NO:1) 的UTR,
- [0311] 或这样的序列,其中在所述UTR序列中,SEQ ID NO:1的位置6处的C被A取代并且SEQ ID NO:1的位置7处的C被G取代;和/或SEQ ID NO:1的位置5处的A被G取代;和
- [0312] (b2) 序列R₂-CNGCCACC (SEQ ID NO:2) 的UTR,其中SEQ ID NO:2的位置2处的核苷酸N是选自U、G、C或A的核苷酸,或这样的序列,其中在所述UTR序列中,SEQ ID NO:2的位置7处的C被A取代并且SEQ ID NO:2的位置8处的C被G取代;和/或SEQ ID NO:2的位置6处的A被G取代,
- [0313] 其中R₂是与以DNA依赖性RNA聚合酶起始RNA合成的核苷酸起始的启动子区域的部分对应的RNA序列。
- [0314] 9. 根据项目8的RNA分子,其中R₂选自:
- [0315] (i) GGGAGA (SEQ ID NO:7);
- [0316] (ii) GGGAGA (SEQ ID NO:8);
- [0317] (iii) GAAG (SEQ ID NO:9);和
- [0318] (iv) GGGGA (SEQ ID NO:10)。
- [0319] 10. 根据项目8或9的RNA分子,其中SEQ ID NO:2的位置2处的核苷酸N是选自U、G或C的核苷酸并且其中核苷酸N不是A。
- [0320] 11. 根据项目10的RNA分子,其中SEQ ID NO:2的位置2处的所述核苷酸N是U。
- [0321] 12. 根据项目8到11中任一项的RNA分子,其中RNA分子在3'端包含聚腺苷酸尾。
- [0322] 13. 根据项目8到12中任一项的RNA分子,其中聚腺苷酸尾的长度为至少120个核苷酸。
- [0323] 14. 核酸分子,其编码项目8到13中任一项的RNA分子。
- [0324] 15. 载体,其包含项目14的核酸分子。
- [0325] 16. 宿主细胞,其包含项目15的载体。
- [0326] 17. 药物组合物,其包含根据项目8到13中任一项的RNA分子、根据项目14的核酸分子、根据项目15的载体或根据项目16的宿主细胞以及任选地药学上可接受的载体。
- [0327] 18. 项目17的药物组合物,其用于基于RNA的疗法。
- [0328] 19. 试剂盒,其包含根据项目1到4中任一项的DNA分子、根据项目8到13中任一项的

RNA分子、根据项目14的核酸分子、根据项目5或15的载体或根据项目6或16的宿主细胞。

[0329] 20. 项目8 (b) 中限定的UTR在用于将RNA分子的编码区翻译成所述编码区编码的多肽或蛋白质中的应用。

[0330] 图1:显示了携带“最小UTR”序列连同本发明中使用的各自的荧光素酶报道构建物 (reporter constructs) 名称的序列。序列携带T7启动子的部分和起始密码子ATG所跟随的Kozak元件的部分。包括TATA序列的前10个碱基以及随后的6个碱基 (GGGAGA) 是T7启动子衍生的序列而在起始密码子ATG上游的剩余碱基属于Kozak元件 (GCCACC)。“Sp30”是30个核苷酸的随机序列。序列No.9中的标下划线的序列是来自具有30个核苷酸长度的人 α 珠蛋白 (“hAg”) 的5'UTR序列。

[0331] 图1中所示序列1至9分别对应于SEQ ID NO:37到45。

[0332] 图2A

[0333] 和B:显示了“最小UTR”中额外的“C”是必要的 (图1中的序列No.1和No.2)。在96孔板中分别以20,000个细胞/孔和40,000个细胞/孔的密度接种人肺泡上皮细胞系 (A549) 和人肝细胞癌细胞系 (HepG2)。接种后24小时,利用Lipofectamine2000用不同的编码荧光素酶的SNIM RNA构建物 (图1中的序列1和2) 转染细胞。在转染后24小时测量荧光素酶表达。数值表示3次重复的平均值 \pm SD并针对经GraphPad Prism分析的转染剂量和数据绘图。在A549和HepG2细胞中,C的缺失导致表达降低。因此该额外的C包括在所有进一步的构建物的设计中。

[0334] 图3:显示了所示单独核苷酸的影响并证明了在A549转染的细胞中额外的“C”和Kozak元件之间的距离的影响。转染的细胞和进行的荧光素酶试验在材料和方法下进行了描述。由于较高剂量远离了线性范围,只有高达62,5ng/孔的剂量响应在此呈现。来自人 α 珠蛋白的5'UTR用作阳性对照。用携带分别来自图1的序列3-8的SNIM RNA分子进行转染实验。在96孔板中以20,000个细胞/孔的密度接种人肺泡上皮细胞系 (A549)。接种后24小时,利用Lipofectamine2000用不同的编码荧光素酶的SNIM RNA构建物 (来自图1的序列No.3-8) 转染细胞。在转染后24小时测量荧光素酶表达。将数值针对经GraphPad Prism分析的转染剂量和数据绘图。数值表示3次重复的平均值 \pm SD。

[0335] 在肺泡上皮细胞系 (A549) 中,在C和Kozak元件 (图1中的序列No.3) 之间插入额外的“A”导致表达显著降低 (图3)。在C和Kozak元件 (图1中的序列No.4) 之间插入单一“T”导致表达水平与用作阳性对照的人 α 珠蛋白5'UTR所获得的表达水平相当。

[0336] 图4:显示了所示单独核苷酸的影响并证明了HepG2转染的细胞中额外的“C”和Kozak元件之间的距离的影响。转染的细胞和进行的荧光素酶试验在材料和方法下进行了描述。由于较高剂量远离了线性范围,只有高达62,5ng/孔的剂量响应在此呈现。用来自图1的序列No.3-8进行转染实验。在96孔板中以40,000个细胞/孔的密度接种肝细胞癌细胞系 (HepG2)。接种后24小时,利用Lipofectamine2000用不同的编码荧光素酶的SNIM RNA构建物 (来自图1的序列No.3-8) 转染细胞。在转染后24小时测量荧光素酶表达。将数值针对经GraphPad Prism分析的转染剂量和数据绘图。数值表示3次重复的平均值 \pm SD。在两种细胞系 (A549细胞 (图3) 和HepG2 (图4)) 中,在C和Kozak元件 (来自图1的序列No.3) 之间插入额外的“A”导致表达显著降低 (图3和4)。在两种细胞中,在C和Kozak元件 (来自图1的序列No.4) 之间插入单一“T”导致表达水平与用作阳性对照的人 α 珠蛋白5'UTR所获得的表达水

平相当。在HepG2细胞中,序列No.1(图1)也是同等有效的。

[0337] 图5:显示了TISU元件对A549细胞中荧光素酶表达的影响。针对所选编码荧光素酶的构建物进行详细的剂量响应和曲线拟合。基于来自图2-4的先前数据,将TISU元件与序列4(图1)组合,其含有两种期望的属性:(C在T7启动子和Kozak元件之间以及额外的T在C和Kozak元件之间从而获得来自图1的序列No.9)。

[0338] 在96孔板中分别以20,000个细胞/孔和40,000个细胞/孔的密度接种人肺泡上皮细胞系(A549)(图5A和B)和人肝细胞癌细胞系(HepG2)(图5C和D)。接种后24小时,利用Lipofectamine2000用不同的编码荧光素酶的SNIM RNA构建物转染细胞。在转染后24和48小时测量荧光素酶表达(图5E)。将数值针对经GraphPad Prism分析的转染剂量和数据绘图。如所示,用不同的编码荧光素酶的mRNA转染A459(A、B)细胞和HepG2(C、D)细胞。在转染后24小时(A、C)和48(B、D)小时测量荧光素酶活性。数值表示3次重复的平均值±SD。

[0339] 在两种细胞系中以及在两种测量时间点上,用含有TISU元件的荧光素酶构建物获得了显著更高的表达(图5A-D)。

[0340] 图6:显示了TISU元件对荧光素酶在A549细胞(图6A)中和在HepG2细胞中(图6B)表达的影响。在96孔板中分别以20,000个细胞/孔和40,000个细胞/孔的密度接种人肺泡上皮细胞系(A549)和人肝细胞癌细胞系(HepG2)。接种后24小时,利用Lipofectamine2000用不同的编码荧光素酶的SNIM RNA构建物转染细胞(X-轴显示了96孔板的每个孔SNIM RNA的ng量)。在转染后24小时测量荧光素酶表达。将数值针对经GraphPad Prism分析的转染剂量和数据绘图。如所示,用不同的编码荧光素酶的mRNA转染A549(A)细胞和HepG2(B)细胞。

[0341] 图7:显示了用不同的编码荧光素酶的mRNA构建物在小鼠中进行的体内实验的结果。将图7中所示荧光素酶构建物(各自的UTR序列元件,参见图1)在Balb/c小鼠(雌性,6-8周)中进行体内测试。对于这组实验,已经显示增强转基因表达的另外的UTR元件(国际公开号WO 2012/170930 A1)也针对其效率进行了测试。含有该UTR元件的荧光素酶构建物被表示为Luc2-SUSA。20μg各自的SNIM-RNA与LF-44复合并静脉注射到Balb/c小鼠中。在注射后6小时利用IVIS成像系统进行体内成像,并将量化为光子/秒/cm²/sr的数值绘图。图7A中显示了整个动物成像的结果,以及图7B(肝)、7C(肺)、7D(脾)中分别显示了整个器官的成像结果。

[0342] 取自动物的器官在液氮中冷冻并均质。在Tris-HCl裂解缓冲液中裂解细胞并测量荧光素酶活性。图7E(肝)、7F(肺)、7G(脾)中分别显示了结果。

[0343] TISU元件的插入导致与先前公开的5'和3'UTR(国际公开号WO 2012/170930 A1)相比表达更高。在C和Kozak(来自图1的序列No.4)之间添加单一T导致观察到的表达水平与人α珠蛋白UTR(来自图1的序列No.8)相当。将TISU元件添加到序列No.4(图1)中进一步增加了表达(来自图1的序列No.9)。惊奇地发现,人α珠蛋白UTR的影响未发现是序列特异性的。随机30个核苷酸序列支持的表达水平与人α珠蛋白5'UTR类似。基于细胞系中的体外结果和小鼠中的体内实验,序列No.1、4、7和9(图1)被提议为携带“最小UTR”的序列的有前景的候选者用于转录物疗法。这些最小UTR序列在体外转录期间对RNA产量没有负面影响,并且产生的mRNA与含有现有技术UTR的状态的mRNA相比翻译更有效得多。

[0344] 图8:显示了白细胞计数(WBC)(图8A)、红细胞(RBC)(图8B)、血小板(图8C)、血红蛋白(图8D)和血细胞比容(hematicrit)(图8E)的数值,其来自具有不同的编码荧光素酶的

mRNA构建物的小鼠。实验基本上如图7中描述而进行,并且血液参数通过利用Sysmex KX-21N™自动血液分析仪(Automated Hematology Analyzer)(IL,美国)分析。

[0345] 图9:显示了使用含有TISU元件、编码人EPO的mRNA的表达实验与含有来自自己已知支持非常高EPO表达的(国际公开号WO 2012/170930 A1:图1和2)(SUSAUTR)的5'和3'UTR、编码人EPO的mRNA的表达实验的比较。

[0346] 在96孔板中分别以20,000个细胞/孔和40,000个细胞/孔的密度接种人肺泡上皮细胞系(A549)和人肝细胞癌细胞系(HepG2)。接种后24小时,利用Lipofectamine2000用250ng不同的编码EPO的SNIM RNA构建物转染细胞。在转染后24小时通过ELISA(来自R&D Systems(MN,美国)的人类红细胞生成素(Human Erythropoietin)Quantikine IVD ELISA试剂盒)定量EPO的量并通过GraphPad Prism分析数据。数值表示3次重复的平均值±SD。

[0347] 图10:显示了使用人OTC的表达实验。对于人OTC,将来自含有TISU元件、编码hOTC的mRNA的表达与含有5'人α珠蛋白UTR的编码hOTC的mRNA(已知其与迄今为止已知的所有其他组合相比产生的表达最高)的表达进行比较。

[0348] 在96孔板中接种人肝细胞癌细胞系(HepG2)并在接种后24小时,利用Lipofectamine2000用不同的编码hOTC的SNIM RNA构建物转染细胞。转染后24h,将细胞裂解并利用蛋白质印迹定量OTC的量。

[0349] 含有hAg和TISU元件两者的编码hOTC的SNIM RNA都导致类似水平的hOTC表达(图10A)。粘着斑蛋白被用作持家蛋白(housekeeper)并且对条带强度定量并用作内部定量标准(图10B)。

[0350] 图11:存在于序列7中随机30个核苷酸长的间隔区(左)和存在于序列8中人α珠蛋白的5'UTR(右)的预测的二级结构。

[0351] 本发明的其它方面和优势将通过以下实施例进行描述,所述实施例出于示例目的并以非限制的方式提供。本申请中引用的每一出版物、专利、专利申请或其它文件在此通过引用以其全部内容并入。

实施例

[0352] I. 材料和方法

[0353] 质粒载体

[0354] 各自的5'UTR序列连同密码子优化的荧光素酶序列通过GeneScriptG(NJ,美国)合成并克隆在pUC57-Kan(GeneScript)中。在EPO(密码子优化的人红细胞生成素)和OTC(密码子优化的人鸟氨酸转氨甲酰酶)的情况下,编码序列荧光素酶基因分别被EPO(SEQ ID No:35)基因和OTC(SEQ ID No:36)基因的编码序列替换。图1中显示了构建物中使用的UTR序列连同各自的荧光素酶报道构建物的名称。

[0355] mRNA产生

[0356] 为了产生体外转录的mRNA(IVT mRNA),质粒通过BstBI消化(digestion)而线性化并通过氯仿提取和醇沉淀而纯化。纯化的线性质粒用作模板用于利用RiboMax Large Scale RNA production System-T7(Promega,德国)的体外转录。将抗逆转帽类似物(Anti-Reverse Cap Analog)(ARCA)添加到反应混合物中以产生5'加帽的mRNA,并且mRNA被聚腺苷酸化(Thermo Scientific)从而产生3'聚腺苷酸尾。

[0357] 另外对于生成SNIM mRNA,将化学修饰的核苷酸即甲基-CTP和硫-UTP (Jena Bioscience, 德国) 添加到终浓度为7.57mM:5.68mM:5.68mM:1.89mM:1.89mM:1.21mM的ATP:CTP:UTP:甲基-CTP:硫-UTP:GTP中。将全部IVT混合物在37°C下温育2小时,然后用DNaseI在37°C下进行DNA消化20分钟。用醋酸铵(终浓度为2.5M)沉淀RNA并用70%EtOH洗涤。洗涤步骤进行两次。最后,在无RNase的水中重悬RNA小球(pellet)。在1%琼脂糖凝胶上验证所有mRNA。转录的RNA被化学修饰,在其中约25%的尿苷残基是2-硫尿苷(s2U)以及约25%的胞苷残基是5-甲基胞苷(m5C)。图1中给出了UTR的序列。

[0358] 体外转染

[0359] 在96孔板中分别以20,000个细胞/孔和40,000个细胞/孔的密度接种人肺泡上皮细胞系(A549)和人肝细胞癌细胞系(HepG2)。接种后24小时,利用商业转染试剂LipofectamineTM2000以每1 μ g mRNA 2.5 μ l LipofectamineTM2000的比例用不同的编码荧光素酶的SNIM RNA构建物转染细胞(图2-6中的X-轴显示96孔板的每个孔SNIM RNA的ng量)。复合物形成按如下准备:将LipofectamineTM2000和mRNA分别在OptiMEM转染介质中稀释以将每一个都加到多达45 μ l的总体积。将这些混合物在室温下温育5分钟。然后将LipofectamineTM2000溶液与mRNA溶液混合,随后在室温下再温育20分钟。在37°C下将细胞在90 μ l总转染体积中温育一小时(5%CO₂水平)。然后将转染介质移除并用PBS洗涤细胞。随后,用含有10%FBS的Leibovitz's L-15培养基再次温育细胞。

[0360] 细胞培养

[0361] 人肺泡腺癌细胞系(A549, ATCC CCL-185)生长在补充有10%FBS的Ham's F12K培养基中。人肝细胞癌细胞系(HepG2, ATCC HB-8065)培养在补充有10%胎牛血清的DMEM培养基中。所有细胞系都以5%CO₂水平生长在潮湿气氛中。

[0362] 生物发光测量

[0363] 萤火虫荧光素酶(FFL)是一种常见的报道蛋白,其非内源地存在于哺乳动物中并可容易地通过发光成像检测。荧光素酶催化荧光素与氧的反应,该反应导致生物发光发射。

[0364] 在96孔板中分别以20,000个细胞/孔和40,000个细胞/孔的密度接种人肺泡上皮细胞系(A549)和人肝细胞癌细胞系(HepG2)。接种后24小时,利用Lipofectamine2000用不同的编码荧光素酶的SNIM RNA构建物转染细胞(X-轴显示96孔板的每一个孔SNIM RNA的ng量)。在转染后24小时测量生物发光。将数值针对经GraphPad Prism分析的转染剂量和数据点进行绘图。

[0365] 为了定量均一的组织裂解物中荧光素酶表达(expersiiion),从动物取下器官,在液氮中冷冻,均质化,并在裂解缓冲液(25mM Tris-HCl, pH 7.5, 具有0.1%曲拉通-X100)中裂解细胞。

[0366] 动物

[0367] 六到八周大的雌性BALB/c小鼠获自Janvier, Route Des Chênes SecsBP5, F-53940Legenet St. Isle, 法国,并维持在特定的无病原体的条件下。在实验前至少七天使小鼠适应动物设施的环境。所有动物程序都经地方伦理委员会批准和控制,并且根据德国动物生命保护法的指导方针实施。

[0368] 类脂质体(Lipidoid)制剂

[0369] 用mRNA如下配制类脂质体:将C12-(2-3-2)、DOPE、Chol和DSPE-PEG2k(重量比3.6:

0.18:0.76:1) 溶解在乙醇中并以10.5的脂质/mRNA重量比迅速注射到柠檬酸盐缓冲溶液(10mM柠檬酸,150mM NaCl,pH=4.5)中以产生20%最终乙醇浓度并针对水透析(dialyzed),所述柠檬酸盐缓冲溶液包含编码萤火虫荧光素酶的经化学修饰的mRNA。产生的类脂质体/mRNA复合物导致带正电的纳米颗粒($92.6 \pm 0.7\text{nm}$; $21.0 \pm 0.2\text{mV}$)并被静脉注射到受限制小鼠的尾静脉中。在第二实验中,在静脉注射前用PBS调节类脂质体/mRNA复合物,这导致几乎不带电的纳米颗粒($91.5 \pm 0.6\text{nm}$; $-0.7 \pm 0.2\text{mV}$)。

[0370] 利用体内生物发光成像测量小鼠中的Luc活性

[0371] 在施用后二十四小时,通过腹膜内注射美托咪定($11.5\mu\text{g}/\text{kg BW}$)、咪达唑仑(midazolame) ($115\mu\text{g}/\text{kg BW}$)和芬太尼($1.15\mu\text{g}/\text{kg BW}$)使小鼠麻醉。经静脉注射施用D-荧光素底物(每只小鼠 $3\text{mg}/100\mu\text{l PBS}$)。10分钟后,利用IVIS 100成像系统(Xenogen, Alameda,美国)和照相机设置:Bin(HS)、视场10(field of view 10)、f1f制光圈(f1f-stop)、高分辨率进仓(high-resolution binning)和曝光时间5分钟来测量生物发光。利用活体图像软件(Living Image Software)2.50版(Xenogen,Alameda,美国)定量并分析信号。

[0372] OTC蛋白的蛋白质印迹分析

[0373] 将冷冻盘解冻并在盘中直接进行细胞裂解。利用补充有蛋白酶抑制剂(完全、无EDTA,罗氏诊断(Roche Diagnostics),德国)和DNase(DNase I溶液(2500U/mL), (赛墨飞(Thermo Fisher),美国)的裂解缓冲液(25mM TRIS,0.1%曲拉通-X 100,西格玛欧德里奇(Sigma-Aldrich),德国)溶解蛋白质。溶解后,将样本与NuPage® LDS样品缓冲液和样品还原剂(Sample Reducing Agent)(Thermo Fisher,美国)混合并在70°C加热10分钟。用XCell4SureLock™ Midi、Bio-Rad Criterion™ System(Thermo Fisher,美国)在NuPAGE 10%Bis-Tris Midi凝胶上利用15μL裂解物进行凝胶电泳。利用TransBlot® Turbo™ 转移系统(伯乐(Biorad),德国)转移蛋白质30分钟。转移后,用NET-明胶将膜封闭30分钟,之后于4°C下将膜用在NET-明胶(OTC多克隆抗体(中心),AP6928c-ABBiocat,德国)中1:2000稀释的初级抗体(primary antibody)过夜温育。在用NET-明胶三次洗涤步骤之后,加入以1:10,000在NET-明胶中稀释的辣根过氧化物酶缀合的二级抗体(山羊抗兔IgG-HRP,sc-2004,圣克鲁斯生物技术(Santa Cruz Biotechnology),美国)在RT下持续1小时。用NET-明胶再洗涤膜三次,直到用化学发光底物试剂盒(LuminataCrescendo Western HRP底物,默克密理博(Merck Millipore),德国)使信号可视化 and 利用ChemiDoc™ MP系统(Biorad,德国)可视化。

[0374] 材料

[0375] FBS、Leibovitz's L-15培养基(Gibco)、Lipofectamine™2000、和OptiMEM(Gibco)购自德国Invitrogen。无菌PBS是内部制备的。Ham's F-12K、DMEM、和胰蛋白酶-EDTA购自德国c.c.pro GmbH。

[0376] II. 结果

[0377] II.a 细胞培养实验

[0378] 图2A和B显示T7启动子和Kozak元件之间的额外的“C”是关键。该碱基缺失导致两种被比较的细胞类型的表达都降低。对于两种构建物(来自图1的序列No.1和No.2),为了方便比较,将完整的剂量范围和线性范围(不包括的数值:分析不包括高于62,5ng/孔的剂

量)单独呈现。在A549和HepG2细胞两者中,C的缺失导致表达降低。因此该额外的C包括在所有进一步的构建物的设计中。

[0379] 基于A549和HepG2细胞中获得的结果,用含有额外的“C”的构建物进行进一步实验(序列号1:T7Luc2)。

[0380] 图3和图4。

[0381] 序列1被用作模板并用于该序列,单一核苷酸(A、T、G或C:分别来自图1的序列号3-6)、或随机序列——30个核苷酸长并且无任何可预测的二级结构(序列7)或来自人 α 珠蛋白(序列8)的5'UTR被并入到被研究的“C”和Kozak元件之间。

[0382] 转染细胞并进行荧光素酶试验,在材料和方法下进行了描述。由于较高剂量远离了线性范围,只有高达62,5ng/孔的剂量响应在此呈现。来自人 α 珠蛋白的5'UTR用作阳性对照。

[0383] 为了概括上述结果,图1到4显示了T7启动子和Kozak元件之间的额外的“C”对于通过利用最低要求的(极简的,minimalistic)5'UTR而获得高蛋白质表达是关键。缺失所述核苷酸导致表达降低。在额外的“C”和Kozak元件之间添加额外的“A”负面影响表达。当在该位置添加嘧啶碱基并且最优选地“T”时,获得与用来自hAg的5'UTR所观察到的水平相当的水平。

[0384] 随后,进行另外的实验以:

[0385] -在与最佳工作序列(序列9)结合时阐明TISU元件的影响,并且

[0386] -确定来自hAg的5'UTR的影响是否是序列特异性影响或5'帽和起始密码子之间的距离是否重要。

[0387] 图5显示了TISU元件对荧光素酶在A549细胞中表达的影响。与图1所示序列No.4相比,“TISU元件”在图1所示序列No.9中并入了“AG”而非“CC”。A549细胞(图5A和B)以及HepG2细胞(图5C和D)显示在转染后24小时(A、C)和48(B、D)小时,使用含有TISU元件连同来自序列No.1的“C”和在该“C”与Kozak元件之间的另外的“T”的荧光素酶构建物显著更高的荧光素酶表达。

[0388] 图6显示与图5相同实验的结果,但其中添加了含有30个核苷酸随机序列的5'UTR以允许并排比较人 α 珠蛋白UTR(来自图1的序列8)和相同长度的随机序列(来自图1的序列7)。如所示,在用SNIM RNA转染后24小时,在HepG2细胞(图6A)和A549细胞(图6B)中测量荧光素酶表达。

[0389] 图9显示在用各自的SNIM RNA转染A549和HepG2细胞后,使用含有TISU元件、编码hEPO的mRNA的表达实验与来自(国际公开号W0 2012/170930A1:图1和2)(SUSAUTR)、被用作标准、含有5'和3'UTR、编码hEPO的mRNA的表达实验相比较的结果。在转染后24小时通过ELISA定量EPO的量。数值表示3次重复的平均值 \pm SD。

[0390] 在人A549细胞中,并入TISU元件导致表达比并入5'和3'UTR所得到的表达更高(图9A)。在HepG2细胞中观察到相当的表达水平(图9B)。这是尤其令人惊讶的,因为并入SUSA 5'和3'UTR使得RNA比根据本发明的UTR长约200个核苷酸。

[0391] 图10显示了使用人OTC的表达实验。为了比较,将含有TISU元件、编码hOTC的mRNA与含有5'人 α 珠蛋白UTR、编码hOTC的mRNA(已知其与迄今为止已知的的所有其他组合相比产生最高的表达)进行比较。

[0392] 用不同的编码hOTC的SNIM RNA构建物转染HepG2细胞,裂解24小时后并通过蛋白质印迹法定量OTC的量。

[0393] 含有hAg和TISU元件、编码hOTC的SNIM RNA导致类似的hOTC表达水平(图10A)。粘着斑蛋白被用作管家蛋白并利用光密度法比较条带强度(图10B)。

[0394] II. b小鼠中Luc2构建物的IV应用

[0395] 结果示于图7和图8中。

[0396] 下述构建物已在小鼠中用于IV应用中:

[0397] Luc2(+8+A)

[0398] Luc2(+8+T)

[0399] Luc2(+8+T)+TISU

[0400] Luc2-hAg

[0401] Luc2-Sp30

[0402] Luc2-SUSA UTRs

[0403] 20 μ g各自的SNIM-RNA与LF-44复合并IV注射到Balb/c小鼠中。作为另外的对照,还生成了Luc2序列:其侧面有在5'端的人CMV增强子(Luc2-SUSA)和在3'端的人类生长激素3'UTR。在该构建物中被用作UTR的序列已从Shire的专利(WO 2012/170930 A1:序列ID 1/图1)获取。

[0404] 在注射后6小时利用IVIS成像系统进行体内成像,并将量化为光子/秒/cm²/sr的数值绘图。图7A中显示了整个动物成像的结果,以及图7B(肝)、7C(肺)、7D(脾)中分别显示了整个器官的成像结果。

[0405] 取自动物的器官在液氮中冷冻、均质化、裂解,并测量荧光素酶活性。图7E(肝)、7F(肺)、7G(脾)中分别显示了结果。

[0406] 通过利用Sysmex KX-21NTM自动血液分析仪分析动物的血液参数:白细胞计数(WBC)(图8A)、红细胞(RBC)(图8B)、血小板(图8C)、血红蛋白(图8D)和血细胞比容(hematocrit)(图8E)的数值,其来自具有不同的编码荧光素酶的mRNA构建物的小鼠,没有显示显著的差异。

[0407] 图11:出现在序列7中随机30个核苷酸长的间隔区(左)和出现在序列8中相同长度的人 α 珠蛋白的5'UTR(右)的预测的二级结构。虽然两个序列的二级结构甚至都不相似,但是它们导致相似的表达水平(图6A和6B)——与T7Luc2(+8+T)-TISU相比,两者同样地低。

序列表

<110> 埃泽瑞斯公司
 <120> 新型最小UTR序列
 <130> Y2004 PCT S3
 <150> EP 16 16 3264.1
 <151> 2016-03-31
 <150> EP 16 17 7094.6
 <151> 2016-06-30
 <160> 45
 <170> SIPOSequenceListing 1.0
 <210> 1
 <211> 7
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (7)
 <223> /mol_type="未指定的DNA" /注意="最小UTR1" /有机体="人工序列"
 <400> 1
 cgccacc 7
 <210> 2
 <211> 8
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (8)
 <223> /mol_type="未指定的DNA" /注意="最小UTR2" /有机体="人工序列"
 <220>
 <221> 变异
 <222> (2)
 <223> /替代="A" /替代="T" /替代="C" /替代="G"
 <220>
 <221> misc_feature
 <222> (2) .. (2)
 <223> n i s a , c , g , t o r u
 <400> 2
 cngccacc 8

<210> 3
 <211> 23
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (23)
 <223> /mol_type="未指定的DNA" /注意="T7启动子" /有机体="人工序列"
 <400> 3
 taatacgact cactataggg aga 23
 <210> 4
 <211> 23
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (23)
 <223> /mol_type="未指定的DNA" /注意="T3启动子" /有机体="人工序列"
 <400> 4
 aattaaccct cactaaaggg aga 23
 <210> 5
 <211> 21
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (21)
 <223> /mol_type="未指定的DNA" /注意="SP6启动子" /有机体="人工序列"
 <400> 5
 atttaggtga cactatagaa g 21
 <210> 6
 <211> 21
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (21)
 <223> /mol_type="未指定的DNA" /注意="K11启动子" /有机体="人工序列"
 <400> 6

aattagggca cactataggg a	21
<210> 7	
<211> 6	
<212> DNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (6)	
<223> /mol_type="未指定的DNA" /注意="T7启动子" /有机体="人工序列"	
<400> 7	
gggaga	6
<210> 8	
<211> 6	
<212> DNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (6)	
<223> /mol_type="未指定的DNA" /注意="T3启动子" /有机体="人工序列"	
<400> 8	
gggaga	6
<210> 9	
<211> 4	
<212> DNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (4)	
<223> /mol_type="未指定的DNA" /注意="SP6启动子" /有机体="人工序列"	
<400> 9	
gaag	4
<210> 10	
<211> 4	
<212> DNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (4)	
<223> /mol_type="未指定的DNA" /注意="K11启动子" /有机体="人工序列"	

<400> 10	
ggga	4
<210> 11	
<211> 16	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (16)	
<223> /mol_type="未指定的RNA" /注意="最小UTR1的反向互补序列" /有机体="人 工序列"	
<400> 11	
caugguggcg ucucce	16
<210> 12	
<211> 17	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (17)	
<223> /mol_type="未指定的RNA" /注意="最小UTR2的反向互补序列" /有机体="人 工序列"	
<220>	
<221> 变异	
<222> (10)	
<223> /替代="G" /替代="C" /替代="U" /替代="A"	
<220>	
<221> misc_feature	
<222> (10) .. (10)	
<223> n is a, c, g, t or u	
<400> 12	
caugguggcn gucucce	17
<210> 13	
<211> 16	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (16)	

<223> /mol_type="未指定的RNA" /注意="包含TISU元件的最小UTR1的反向互补序列" /有机体="人工序列"

<400> 13

caucuuggcg ucucce

16

<210> 14

<211> 17

<212> RNA

<213> 人工序列()

<220>

<221> 来源

<222> (1) .. (17)

<223> /mol_type="未指定的RNA" /注意="包含TISU元件的最小UTR2的反向互补序列" /有机体="人工序列"

<220>

<221> 变异

<222> (10)

<223> /替代="G" /替代="C" /替代="A" /替代="U"

<220>

<221> misc_feature

<222> (10) .. (10)

<223> n i s a , c , g , t o r u

<400> 14

caucuuggcn gucucce

17

<210> 15

<211> 16

<212> RNA

<213> 人工序列()

<220>

<221> 来源

<222> (1) .. (16)

<223> /mol_type="未指定的RNA" /注意="可选的最小UTR1的反向互补序列" /有机体="人工序列"

<400> 15

cauggcggcg ucucce

16

<210> 16

<211> 17

<212> RNA

<213> 人工序列()

<220>

<221> 来源
 <222> (1) .. (17)
 <223> /mol_type="未指定的RNA" /注意="可选的最小UTR2的反向互补序列" /有机体="人工序列"
 <220>
 <221> 变异
 <222> (10)
 <223> /替代="U" /替代="A" /替代="G" /替代="C"
 <220>
 <221> misc_feature
 <222> (10) .. (10)
 <223> n is a, c, g, t or u
 <400> 16
 cauggcggcn gucucucc 17
 <210> 17
 <211> 16
 <212> RNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (16)
 <223> /mol_type="未指定的RNA" /注意="包含TISU元件的可选最小UTR1的反向互补序列" /有机体="人工序列"
 <400> 17
 caucucggcg ucucucc 16
 <210> 18
 <211> 17
 <212> RNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (17)
 <223> /mol_type="未指定的RNA" /注意="包含TISU元件的可选最小UTR2的反向互补序列" /有机体="人工序列"
 <220>
 <221> 变异
 <222> (10)
 <223> /替代="U" /替代="G" /替代="C" /替代="A"
 <220>

<221> misc_feature	
<222> (10) .. (10)	
<223> n is a, c, g, t or u	
<400> 18	
caucucggcn gucucucc	17
<210> 19	
<211> 14	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (14)	
<223> /mol_type="未指定的RNA" /注意="最小UTR1的反向互补序列" /有机体="人 工序列"	
<400> 19	
caugguggcg uccc	14
<210> 20	
<211> 15	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (15)	
<223> /mol_type="未指定的RNA" /注意="最小UTR2的反向互补序列" /有机体="人 工序列"	
<220>	
<221> 变异	
<222> (10)	
<223> /替代="A" /替代="C" /替代="G" /替代="U"	
<220>	
<221> misc_feature	
<222> (10) .. (10)	
<223> n is a, c, g, t or u	
<400> 20	
caugguggcn guccc	15
<210> 21	
<211> 14	
<212> RNA	
<213> 人工序列()	

<220>
 <221> 来源
 <222> (1) .. (14)
 <223> /mol_type="未指定的RNA" /注意="包含TISU元件的最小UTR1的反向互补序列" /有机体="人工序列"
 <400> 21
 caucuuggcg uccc 14
 <210> 22
 <211> 15
 <212> RNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (15)
 <223> /mol_type="未指定的RNA" /注意="包含TISU元件的最小UTR2的反向互补序列" /有机体="人工序列"
 <220>
 <221> 变异
 <222> (10)
 <223> /替代="A" /替代="C" /替代="G" /替代="U"
 <220>
 <221> misc_feature
 <222> (10) .. (10)
 <223> n is a, c, g, t or u
 <400> 22
 caucuuggcn guccc 15
 <210> 23
 <211> 14
 <212> RNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (14)
 <223> /mol_type="未指定的RNA" /注意="可选的最小UTR1的反向互补序列" /有机体="人工序列"
 <400> 23
 cauggcggcg uccc 14
 <210> 24
 <211> 15

<212> RNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (15)
 <223> /mol_type="未指定的RNA" /注意="可选的最小UTR2的反向互补序列" /有机体="人工序列"
 <220>
 <221> 变异
 <222> (10)
 <223> /替代="A" /替代="C" /替代="G" /替代="U"
 <220>
 <221> misc_feature
 <222> (10) .. (10)
 <223> n i s a , c , g , t o r u
 <400> 24
 cauggcggcn guccc 15
 <210> 25
 <211> 14
 <212> RNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (14)
 <223> /mol_type="未指定的RNA" /注意="包含TISU元件的可选最小UTR1的反向互补序列" /有机体="人工序列"
 <400> 25
 caucucggcg uccc 14
 <210> 26
 <211> 15
 <212> RNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (15)
 <223> /mol_type="未指定的RNA" /注意="包含TISU元件的可选最小UTR2的反向互补序列" /有机体="人工序列"
 <220>
 <221> 变异

<222> (10)	
<223> /替代="A" /替代="C" /替代="G" /替代="U"	
<220>	
<221> misc_feature	
<222> (10) .. (10)	
<223> n is a, c, g, t or u	
<400> 26	
caucucggcn guccc	15
<210> 27	
<211> 14	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (14)	
<223> /mol_type="未指定的RNA" /注意="最小UTR1的反向互补序列" /有机体="人 工序列"	
<400> 27	
caugguggcg cuuc	14
<210> 28	
<211> 15	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (15)	
<223> /mol_type="未指定的RNA" /注意="最小UTR2的反向互补序列" /有机体="人 工序列"	
<220>	
<221> 变异	
<222> (10)	
<223> /替代="A" /替代="C" /替代="G" /替代="U"	
<220>	
<221> misc_feature	
<222> (10) .. (10)	
<223> n is a, c, g, t or u	
<400> 28	
caugguggcn gcuuc	15
<210> 29	

<211> 14	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (14)	
<223> /mol_type="未指定的RNA" /注意="包含TISU元件的最小UTR1的反向互补序列" /有机体="人工序列"	
<400> 29	
caucuuggcg cuuc	14
<210> 30	
<211> 15	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (15)	
<223> /mol_type="未指定的RNA" /注意="包含TISU元件的最小UTR2的反向互补序列" /有机体="人工序列"	
<220>	
<221> 变异	
<222> (10)	
<223> /替代="A" /替代="C" /替代="G" /替代="U"	
<220>	
<221> misc_feature	
<222> (10) .. (10)	
<223> n is a, c, g, t or u	
<400> 30	
caucuuggcn gcuuc	15
<210> 31	
<211> 14	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (14)	
<223> /mol_type="未指定的RNA" /注意="可选的最小UTR1的反向互补序列" /有机体="人工序列"	
<400> 31	

cauggcggcg cuuc	14
<210> 32	
<211> 15	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (15)	
<223> /mol_type="未指定的RNA" /注意="可选的最小UTR2的反向互补序列" /有机体="人工序列"	
<220>	
<221> 变异	
<222> (10)	
<223> /替代="A" /替代="C" /替代="G" /替代="U"	
<220>	
<221> misc_feature	
<222> (10) .. (10)	
<223> n is a, c, g, t or u	
<400> 32	
cauggcggcn gcuuc	15
<210> 33	
<211> 14	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (14)	
<223> /mol_type="未指定的RNA" /注意="包含TISU元件的可选最小UTR1的反向互补序列" /有机体="人工序列"	
<400> 33	
caucucggcg cuuc	14
<210> 34	
<211> 15	
<212> RNA	
<213> 人工序列()	
<220>	
<221> 来源	
<222> (1) .. (15)	
<223> /mol_type="未指定的RNA" /注意="包含TISU元件的可选最小UTR2的反向互补"	

序列” /有机体=“人工序列”

<220>

<221> 变异

<222> (10)

<223> /替代=“A” /替代=“C” /替代=“G” /替代=“U”

<220>

<221> misc_feature

<222> (10) .. (10)

<223> n is a, c, g, t or u

<400> 34

caucucggcn gcuuc

15

<210> 35

<211> 637

<212> DNA

<213> 人工序列()

<220>

<221> 来源

<222> (1) .. (637)

<223> /mol_type=“未指定的DNA” /注意=“T7-TISU-hEPO (密码子优化的)” /有机体=“人工序列”

<400> 35

```
gtgactagat ctaatacga ctactatag ggagactgcc aagatgggcg tgcacgaatg 60
tcctgcttgg ctgtggctgc tgctgagcct gctgtctctg cctctgggac tgcctgtgct 120
gggagcccct cctagactga tctgcgacag ccgggtgctg gaaagatacc tgctggaagc 180
caaagaggcc gagaacatca ccaccggctg cgccgagcac tgcagcctga acgagaatat 240
caccgtgccc gacaccaaag tgaacttcta cgcttggga ggatggaag tgggccagca 300
ggctgtggaa gtgtggcagg gactggccct gctgagcga gctgtgctga gaggacaggc 360
tctgctcgtg aacagcagcc agccttggga gcctctgcag ctgcacgtgg acaaggccgt 420
gtctggcctg agaagcctga ccacactgct gagagccctg ggggccaga aagaggccat 480
ctctccacct gatgccgct ctgccgccc tctgagaacc atcaccgccc acaccttcag 540
aaagctgttc cgggtgtaca gcaacttct gcggggcaag ctgaagctgt acacaggcga 600
ggcctgccgg accggcgata gataattcga agtgact 637
```

<210> 36

<211> 1120

<212> DNA

<213> 人工序列()

<220>

<221> 来源

<222> (1) .. (1120)

<223> /mol_type="未指定的DNA" /注意="T7-TISU-hOTC (密码子优化的)" /有机体="人工序列"

<400> 36

```
gtgactgcta gctaatacga ctactatag ggagactgcc aagatgctgt tcaacctgcg 60
gatcctgctg aacaacgccg cttccggaa cggccacaac ttcattggtgc gcaacttcag 120
atgcggccag cccctgcaga acaaggtgca gctgaagggc agggacctgc tgacctgaa 180
gaacttcacc ggcaagaga tcaagtacat gctgtggctg agcgccgacc tgaagttccg 240
gatcaagcag aagggcgagt acctgcccct gctgcagggc aagtctctgg gcatgatctt 300
cgagaagcgg agcaccggga cccggctgtc taccgagaca ggatttgccc tgctgggagg 360
ccacccttgc tttctgacca cccaggatat ccacctgggc gtgaacgaga gcctgaccga 420
cacagccaga gtgctgagca gcatggccga tgccgtgctg gccagagtgt acaagcagag 480
cgacctggac accctggcca aagaggccag catecccatc atcaacggcc tgtccgacct 540
gtaccacccc atccagatcc tggccgacta cctgacctg caggaacact acagctccct 600
gaagggcctg aactgagct ggatcggcga cggcaacaac atcctgcact ctatcatgat 660
gagcgccgcc aagttcggca tgcatttgcg ggccgccacc cccaagggt atgagcctga 720
tgccagcgtg accaagctgg ccgagcagta cgccaaagag aacggcacca agctgctgct 780
gaccaacgac cctctggaag ccgcccacgg cggcaatgtg ctgatcaccg atacctggat 840
cagcatgggc caggaagagg aaaagaagaa gcggctgcag gccttccagg gctaccaagt 900
gaccatgaag accgcaaaag tggccgccag cgactggacc ttctgcaact gcctgcccag 960
aaagcccga gaggtggacg acgaggtgtt ctacagcccc cggtccttgg tgtttcccga 1020
ggccgagaac cggaagtgga ccatcatggc tgtgatggtg tctctgctga ccgactactc 1080
ccccagctg cagaagccca agttctgaag cgctgtgact 1120
```

<210> 37

<211> 20

<212> DNA

<213> 人工序列()

<220>

<221> 来源

<222> (1) .. (20)

<223> /mol_type="未指定的DNA" /注意="T7Luc2" /有机体="人工序列"

<400> 37

tatagggaga cgccaccatg 20

<210> 38

<211> 19

<212> DNA

<213> 人工序列()

<220>

<221> 来源

<222> (1) .. (19)

<223> /mol_type="未指定的DNA" /注意="T7Luc2 (Δ C)" /有机体="人工序列"
 <400> 38
 tatagggaga gccaccatg 19
 <210> 39
 <211> 21
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (21)
 <223> /mol_type="未指定的DNA" /注意="T7Luc2 (+8+A)" /有机体="人工序列"
 <400> 39
 tatagggaga cagccacat g 21
 <210> 40
 <211> 21
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (21)
 <223> /mol_type="未指定的DNA" /注意="T7Luc2 (+8+T)" /有机体="人工序列"
 <400> 40
 tatagggaga ctgccacat g 21
 <210> 41
 <211> 21
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (21)
 <223> /mol_type="未指定的DNA" /注意="T7Luc2 (+8+G)" /有机体="人工序列"
 <400> 41
 tatagggaga cggccacat g 21
 <210> 42
 <211> 21
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源

<222> (1) .. (21)
 <223> /mol_type="未指定的DNA" /注意="T7Luc2 (+8+C)" /有机体="人工序列"
 <400> 42
 tatagggaga cgcaccacat g 21
 <210> 43
 <211> 21
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (21)
 <223> /mol_type="未指定的DNA" /注意="T7Luc2 (+8+Sp30)" /有机体="人工序列"
 <220>
 <221> 变异
 <222> (12)
 <223> /替代="30个核苷酸的随机序列的序列"
 <220>
 <221> misc_feature
 <222> (12) .. (12)
 <223> n is a, c, g, t or u
 <400> 43
 tatagggaga cngccacat g 21
 <210> 44
 <211> 50
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (50)
 <223> /mol_type="未指定的DNA" /注意="T7Luc2 (+8+hAg)" /有机体="人工序列"
 <400> 44
 tatagggaga ctcttctgtt cccacagac tcagagagaa cgccacatg 50
 <210> 45
 <211> 21
 <212> DNA
 <213> 人工序列()
 <220>
 <221> 来源
 <222> (1) .. (21)

<223> /mol_type="未指定的DNA" /注意="T7Luc2 (+8+T)+TISU" /有机体="人工序列"

<400> 45

tatagggaga ctgccaagat g

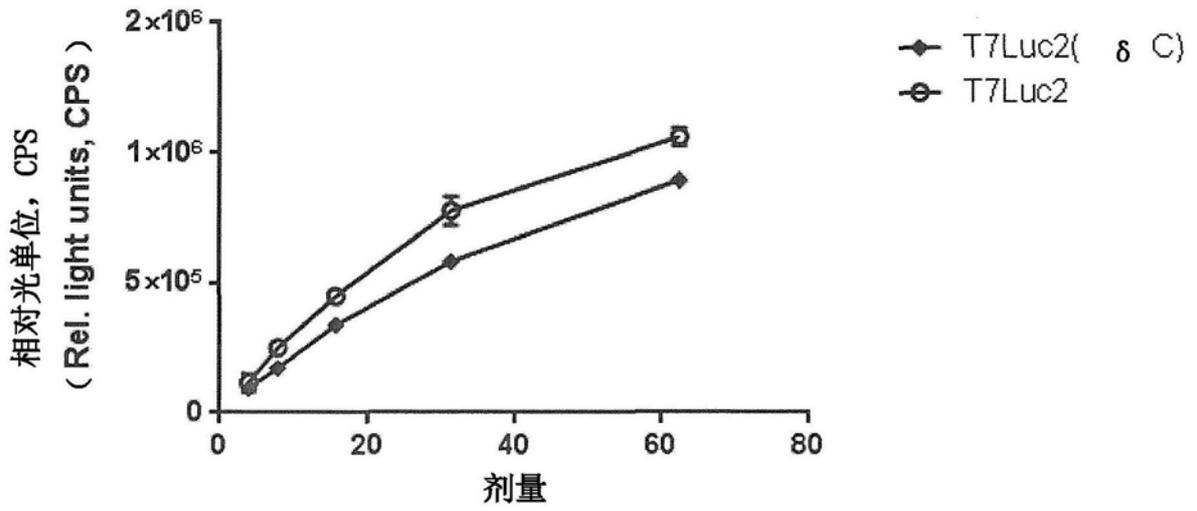
21

1.) ...TATAGGGAGACGCCACCATG	T7Luc2 (SEQ ID NO:37)
2.) ...TATAGGGAGAGCCACCATG	T7Luc2 (Δ C) (SEQ ID NO:38)
3.) ...TATAGGGAGACAGCCACCATG	T7Luc2 (+8+A) (SEQ ID NO:39)
4.) ...TATAGGGAGACTGCCACCATG	T7Luc2 (+8+T) (SEQ ID NO:40)
5.) ...TATAGGGAGACGGCCACCATG	T7Luc2 (+8+G) (SEQ ID NO:41)
6.) ...TATAGGGAGACCGCCACCATG	T7Luc2 (+8+C) (SEQ ID NO:42)
7.) ...TATAGGGAGACSp30GCCACCATG	T7Luc2 (+8+Sp30) (SEQ ID NO:43)
8.) ... TATAGGGAGACT <u>CTTCTGGTCCCCAC AGACTCAGAGAGAACGCCACCATG</u>	T7Luc2 (+8+hAg) (SEQ ID NO:44)
9.) ...TATAGGGAGACTGCCAAGATG	T7Luc2 (+8+T)+TISU (SEQ ID NO:45)

Sp30: 在序列No. 8)中标下划线的30个核苷酸的随机序列: 来自人 α 珠蛋白 (hAg) 的5' UTR (长度为30个核苷酸)

图1

A549细胞-不包括的数值



A549细胞-24hrs

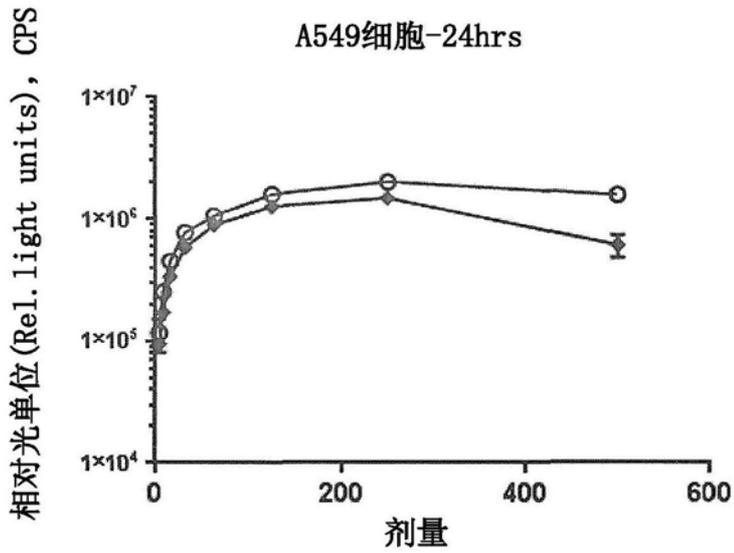


图2A

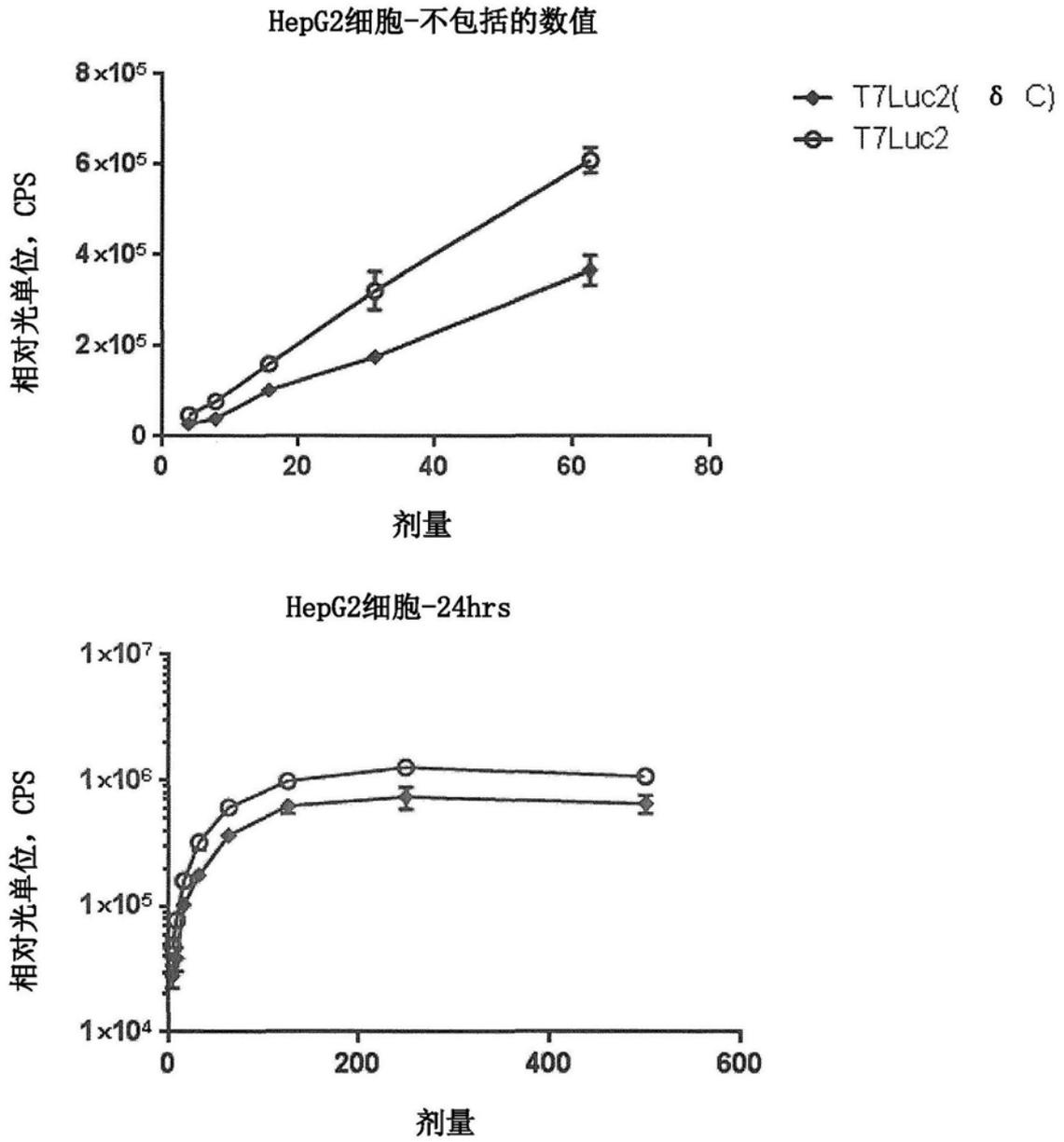


图2B

A549细胞-不包括的数值

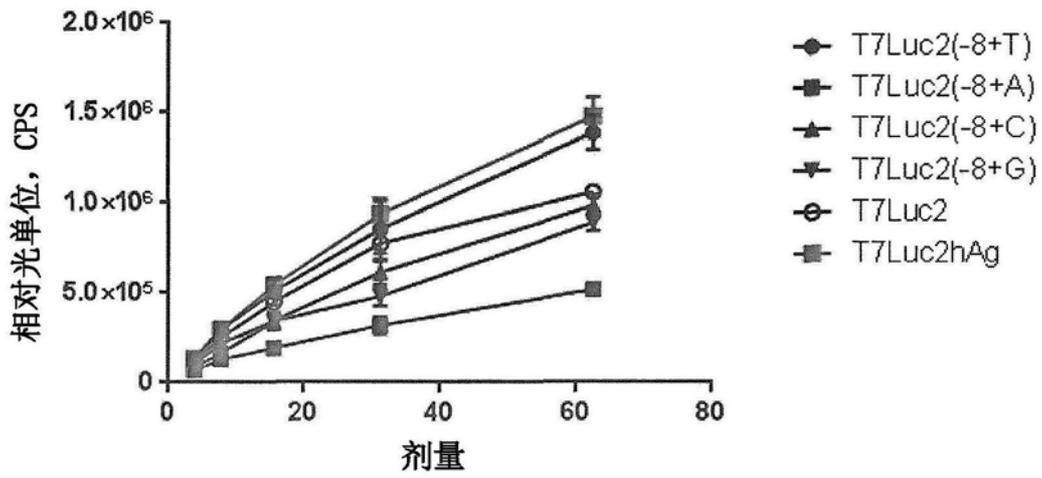


图3

HepG2细胞-不包括的数值

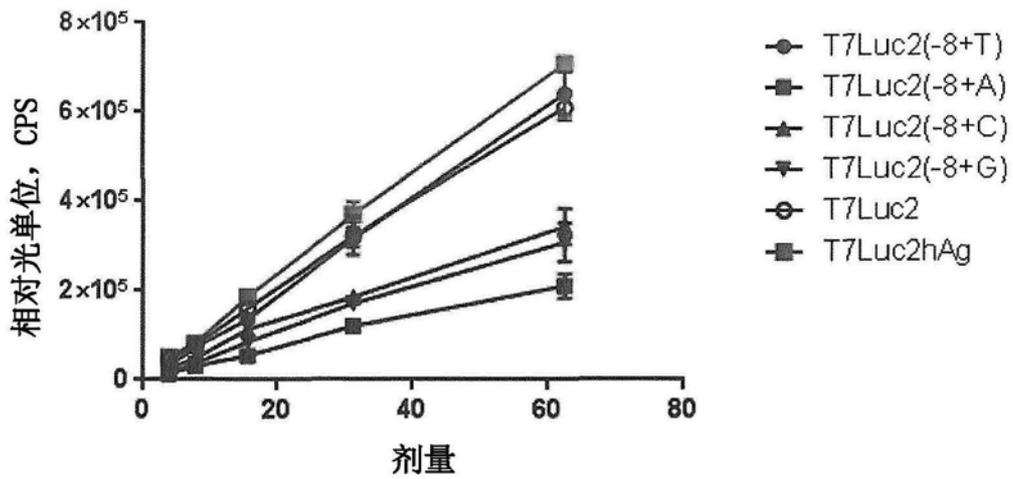


图4

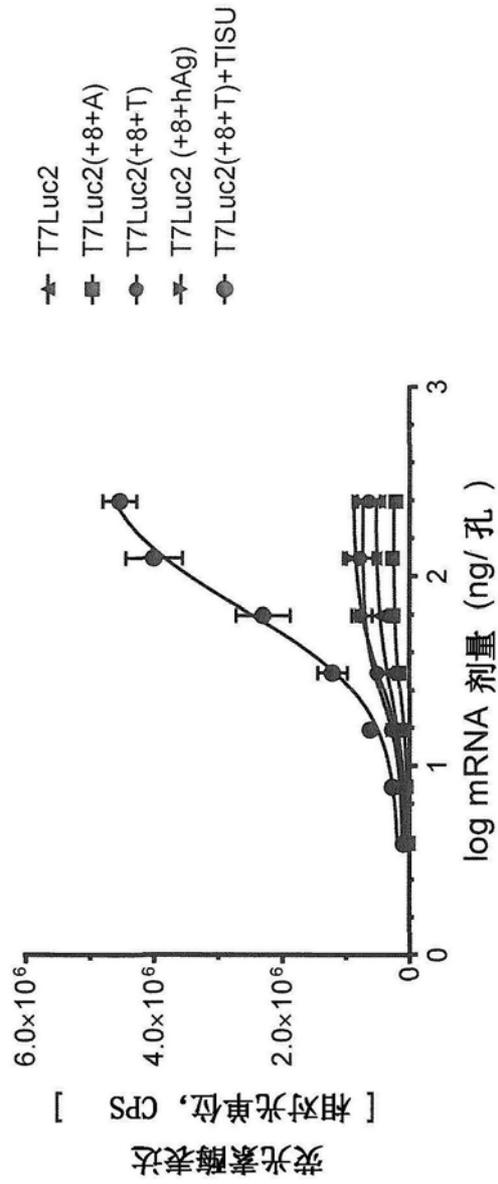


图5A

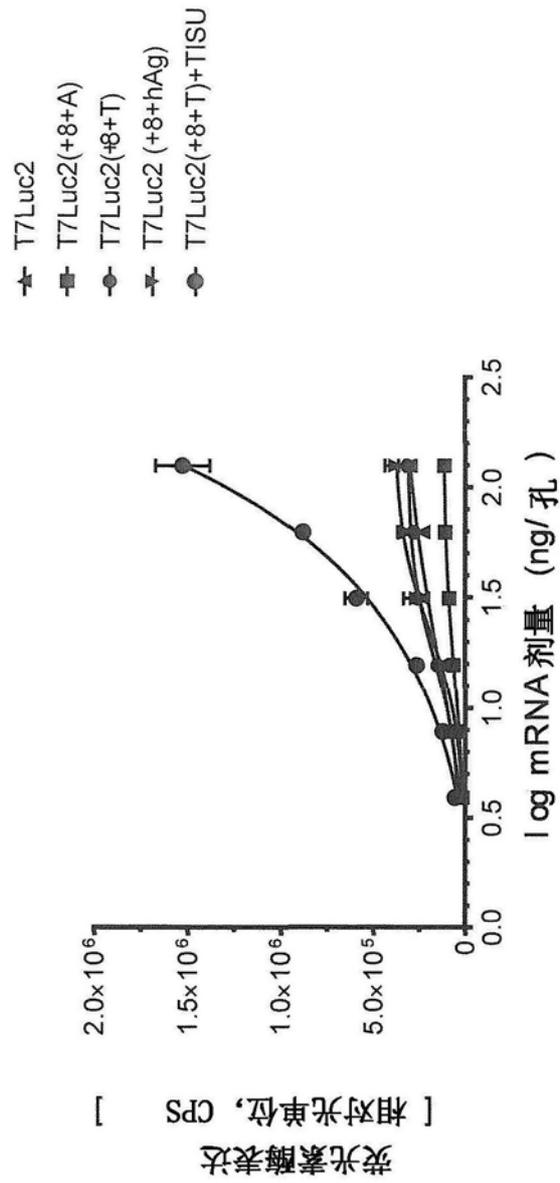


图5B

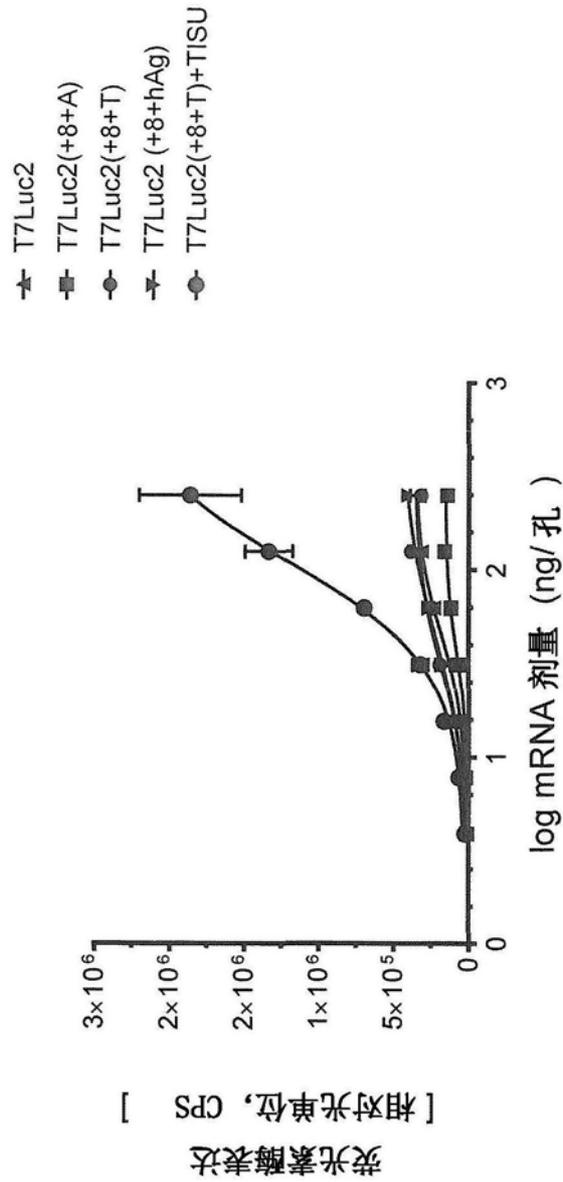


图5C

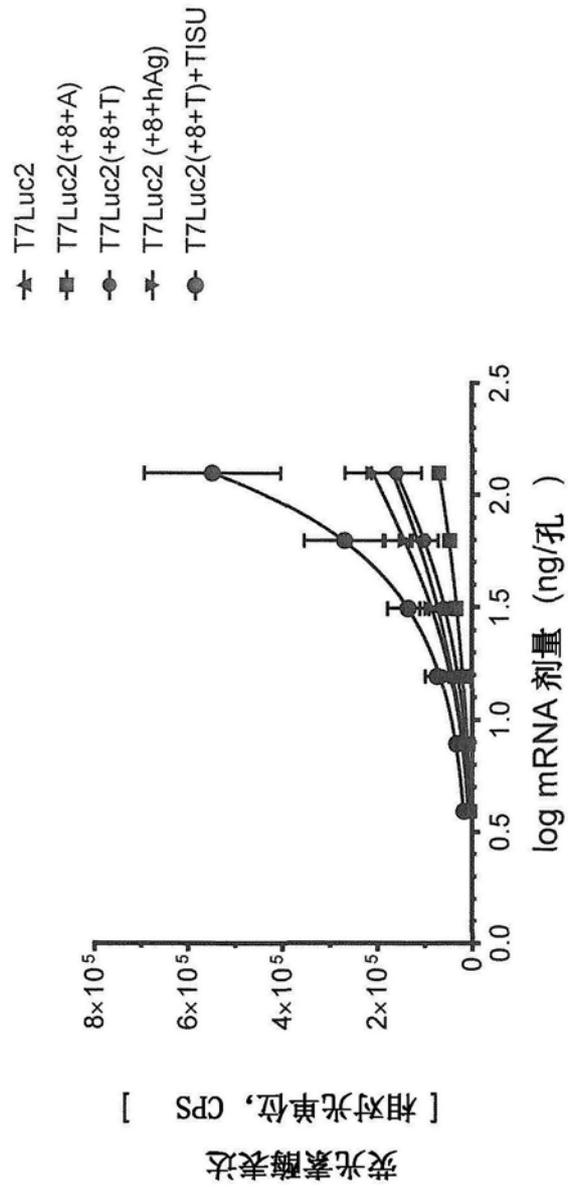


图5D

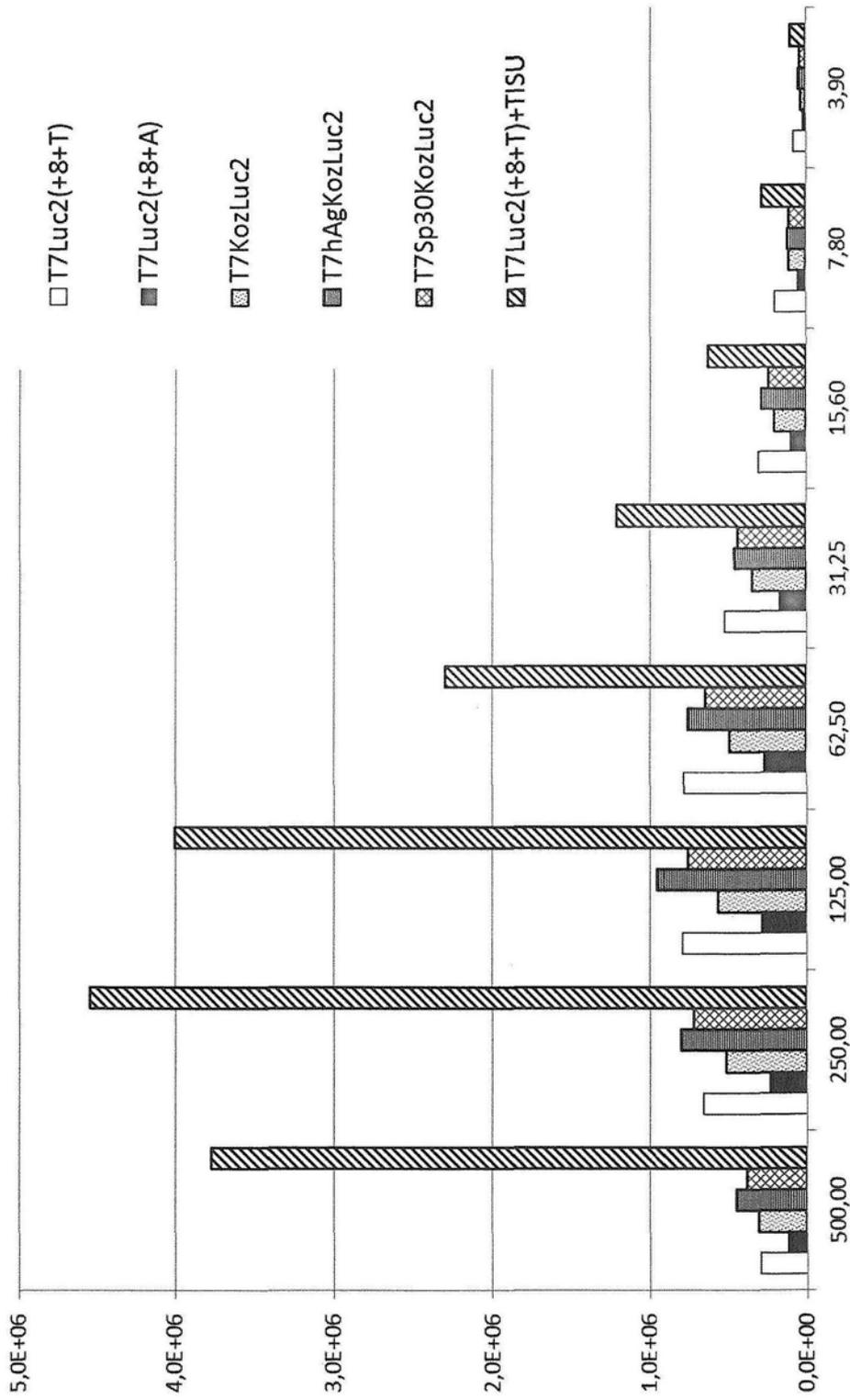


图6A

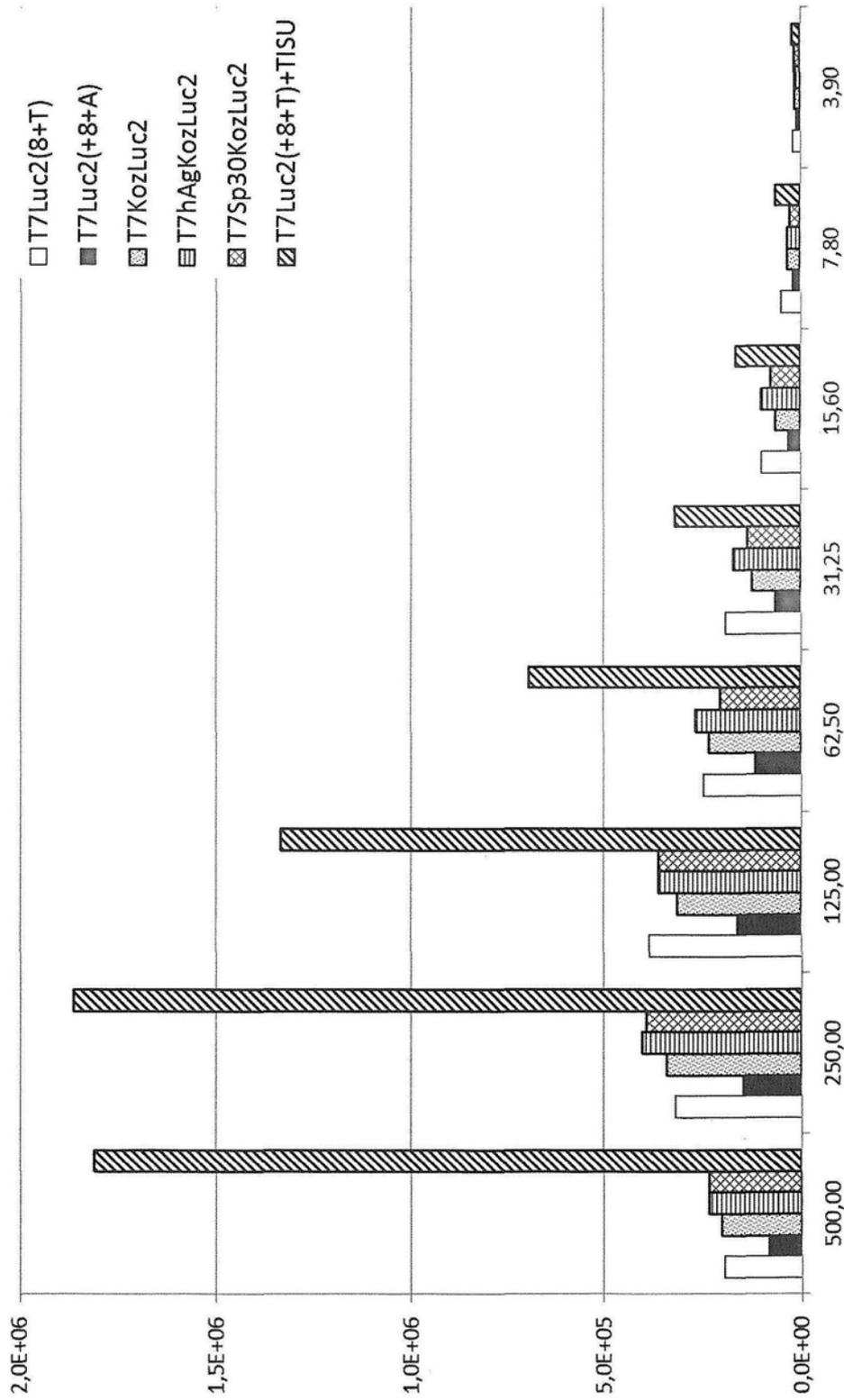


图6B

动物, Ivis

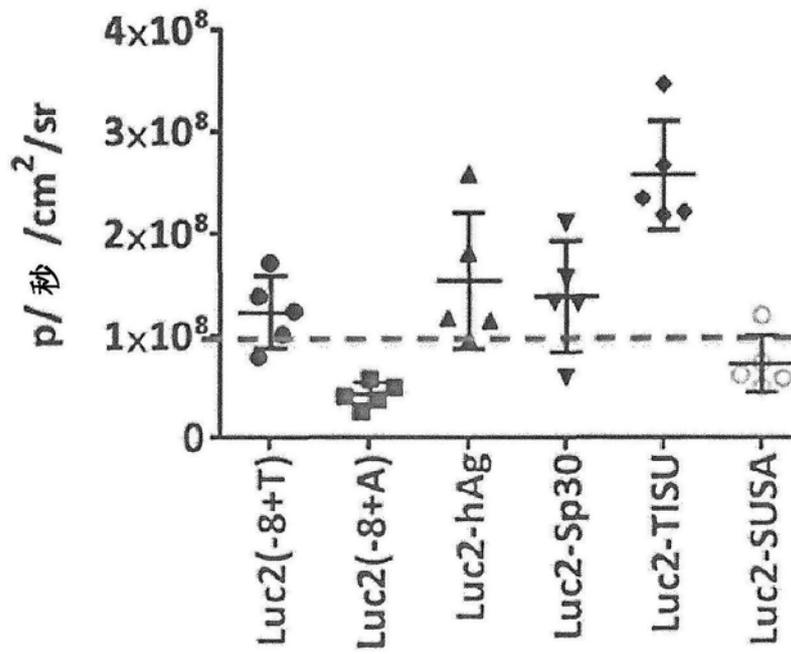


图7A

肝, Ivis

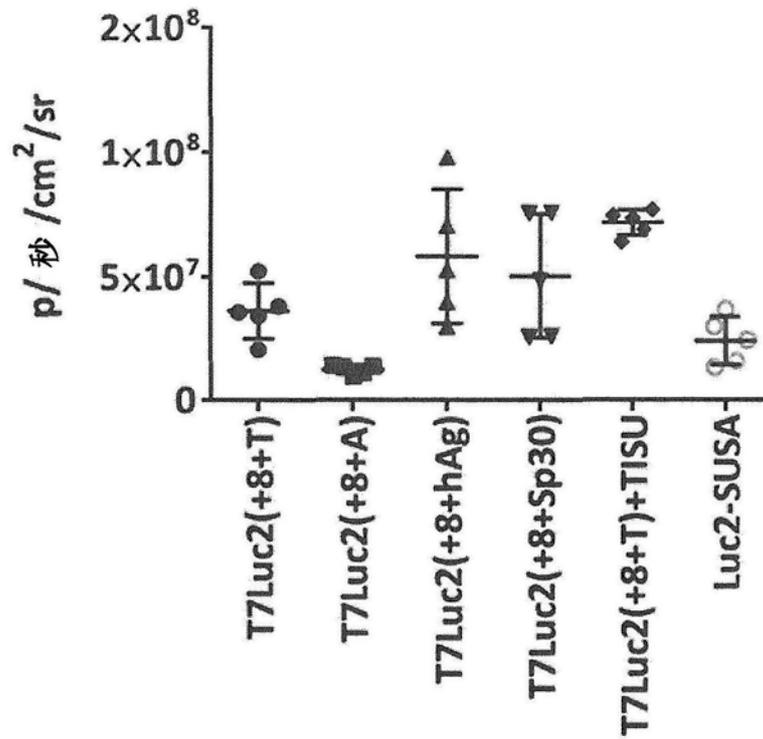


图7B

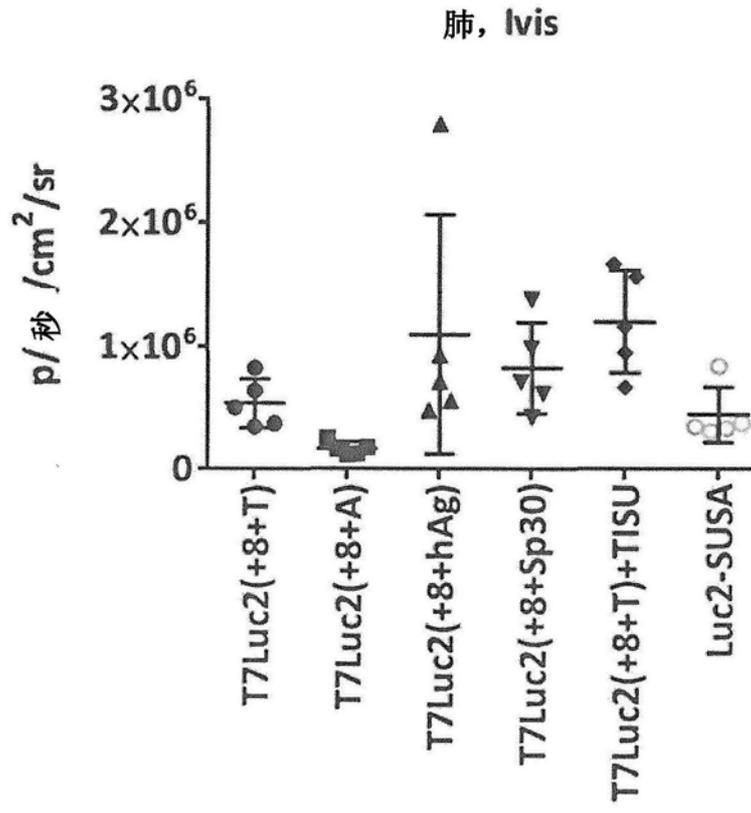


图7C

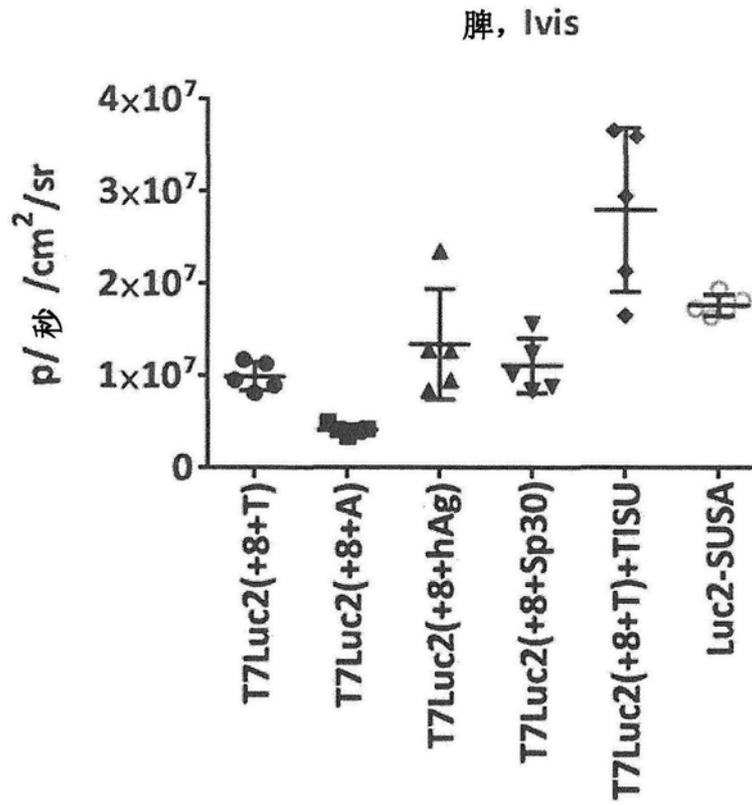


图7D

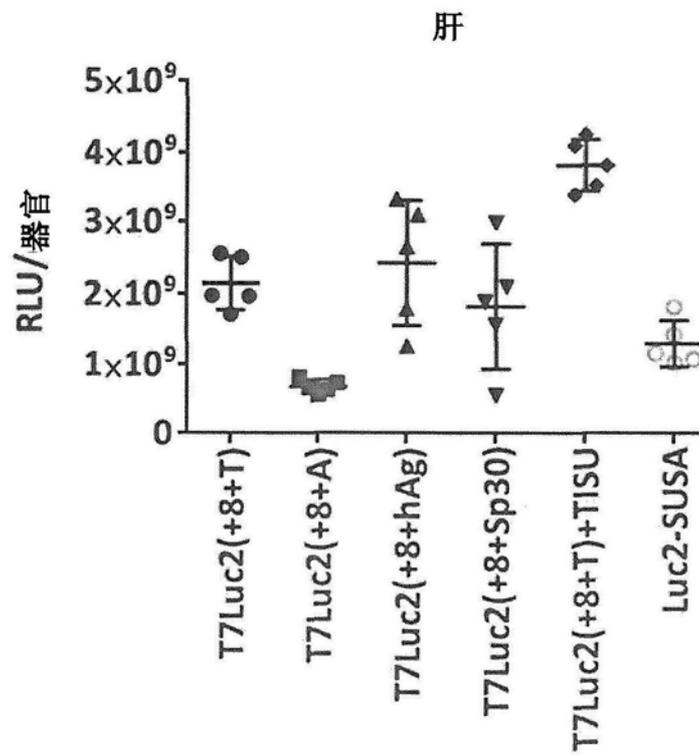


图7E

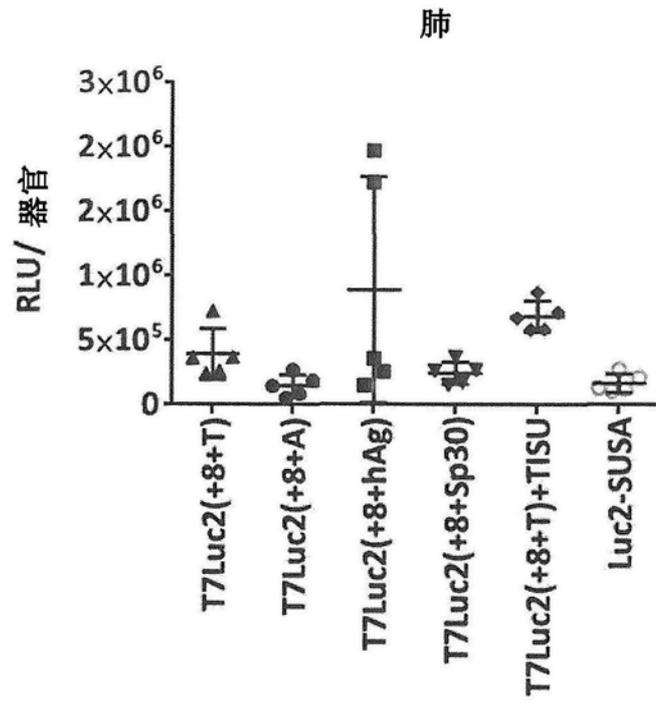


图7F

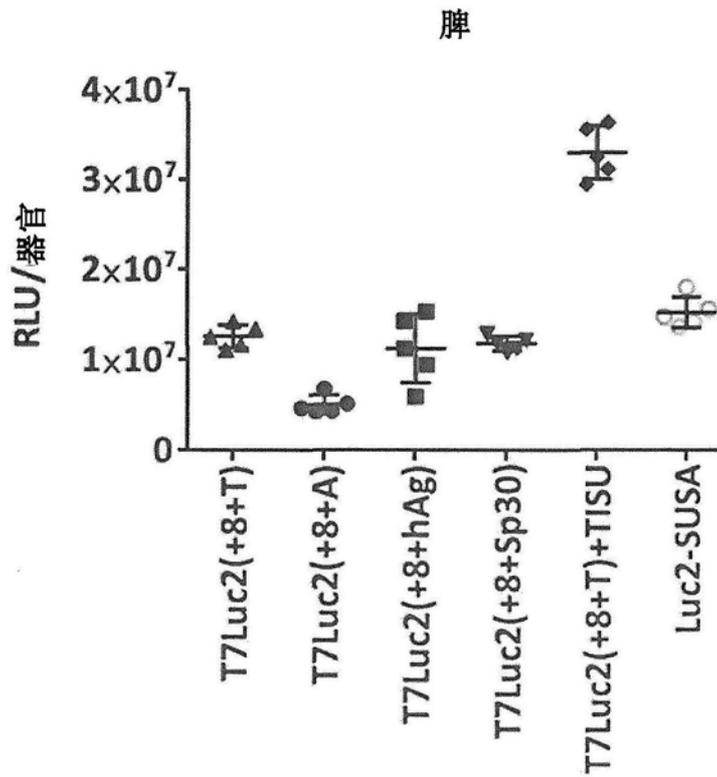


图7G

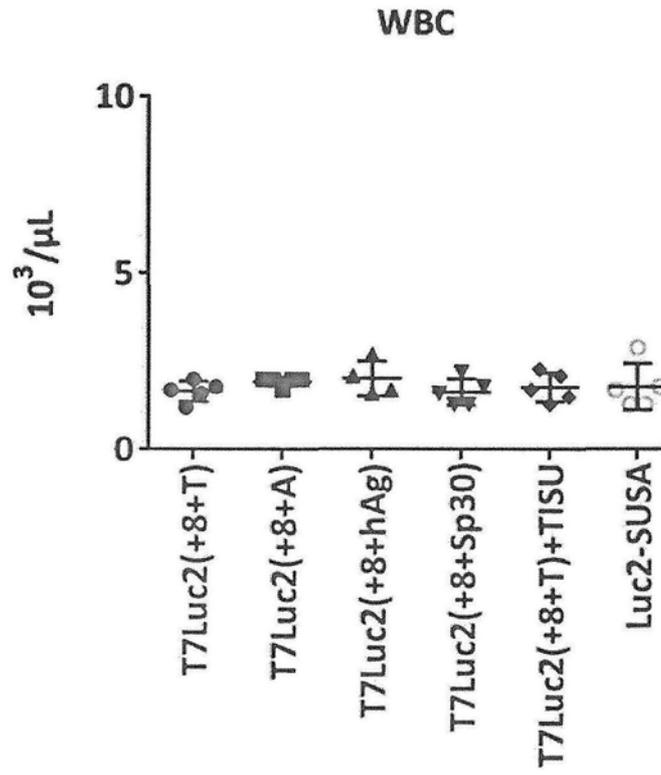


图8A

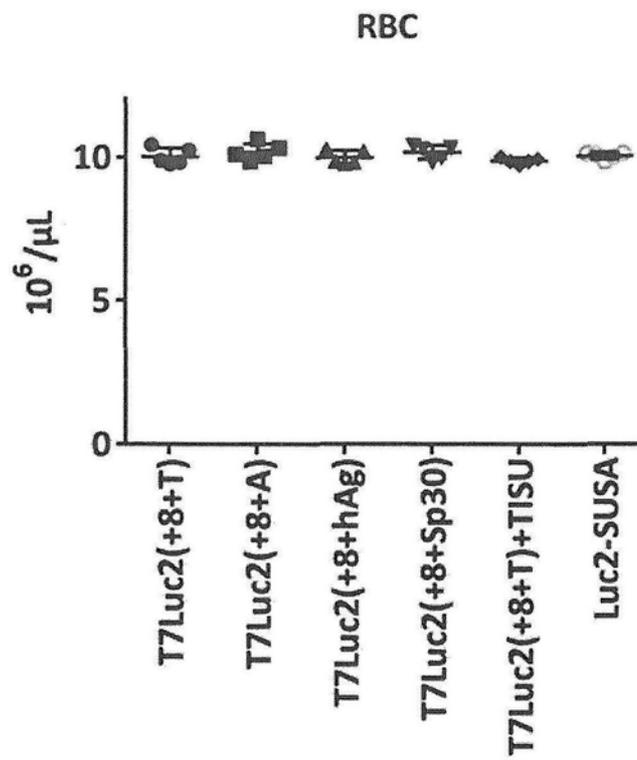


图8B

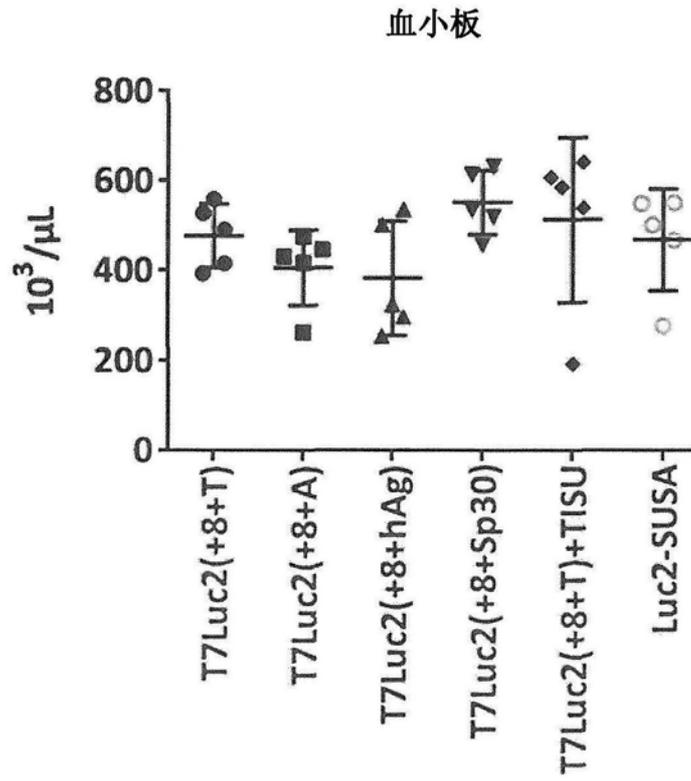


图8C

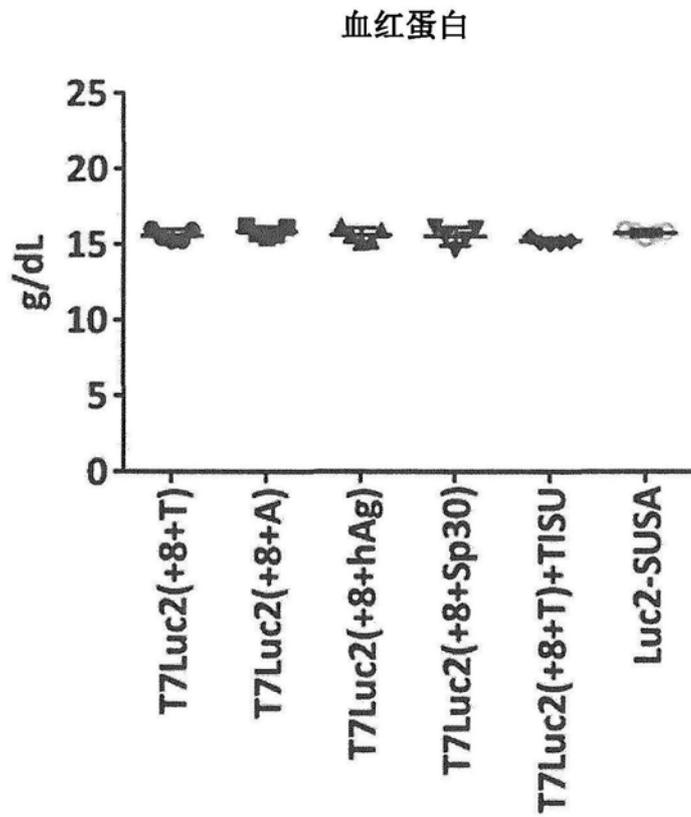


图8D

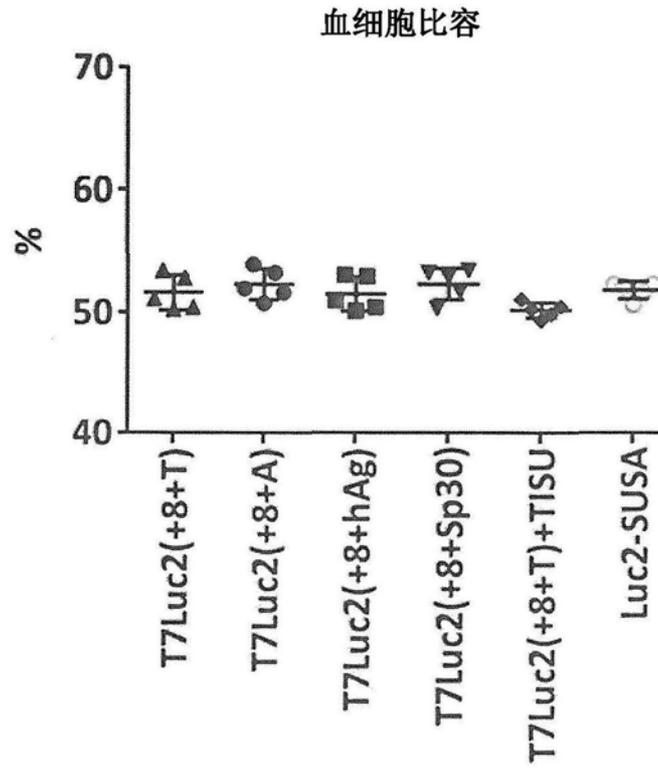


图8E

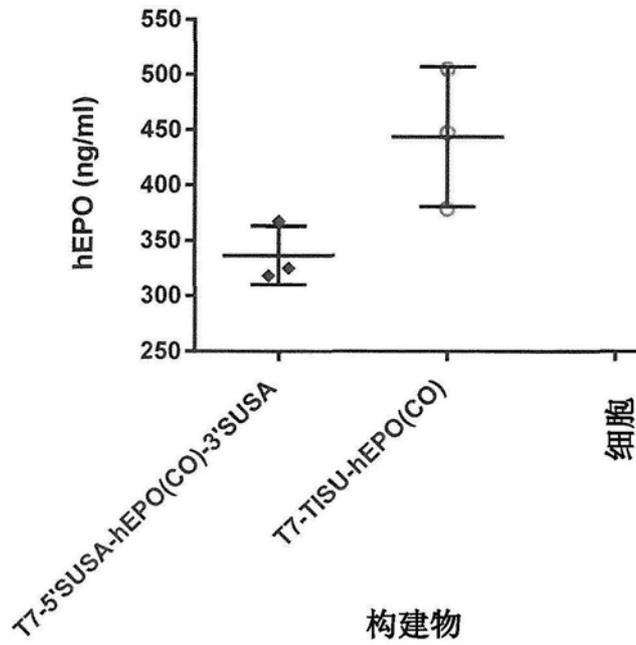


图9A

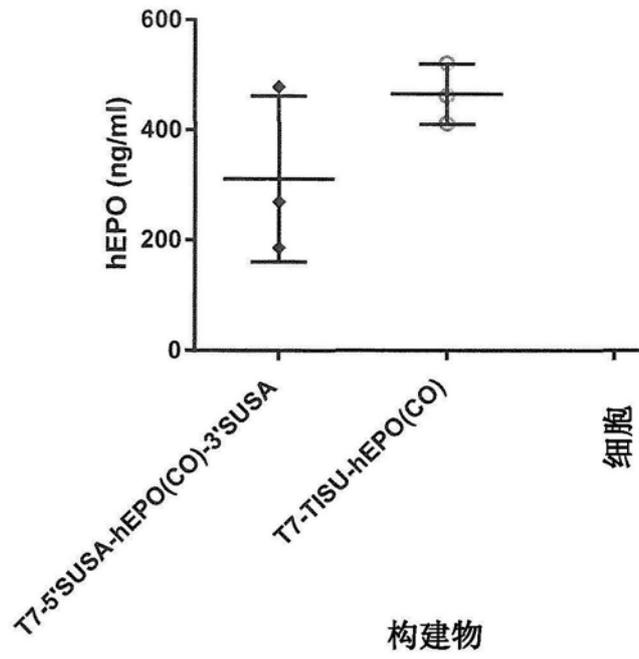


图9B

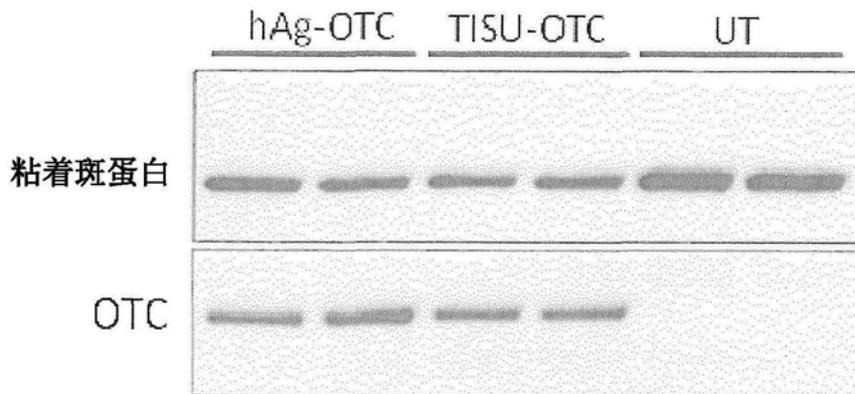


图10A

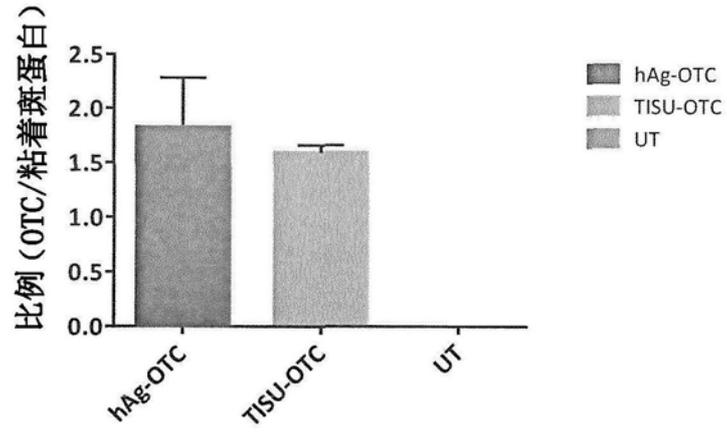


图10B

