



(12) 发明专利申请

(10) 申请公布号 CN 116262917 A

(43) 申请公布日 2023.06.16

(21) 申请号 202111523630.3

A61K 31/7088 (2006.01)

(22) 申请日 2021.12.14

(71) 申请人 成都凌泰氮生物技术有限公司
地址 610000 四川省成都市高新区成都天府国际生物城(双流区生物城中路二段18号)

(72) 发明人 张晓霞 曾秒

(74) 专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 罗文锋 初明明

(51) Int. Cl.

C12N 15/11 (2006.01)

C12N 15/113 (2010.01)

C12N 15/115 (2010.01)

A61K 47/26 (2006.01)

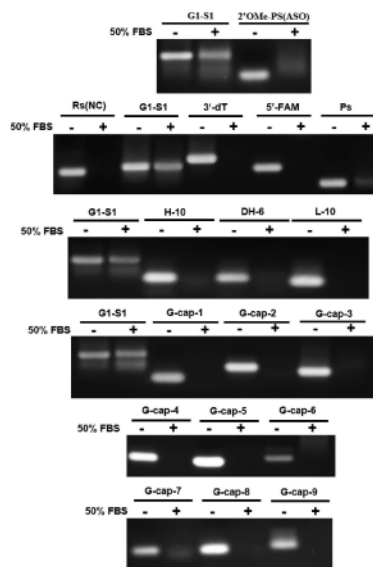
权利要求书2页 说明书22页
序列表29页 附图9页

(54) 发明名称

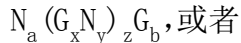
含有G-四链体的核酸

(57) 摘要

本发明涉及含有G-四链体的核酸。本发明公开了一种用于提高目的核酸稳定性的具有G-四链体结构的核酸分子,其特征在于包含: $N_a(G_xN_y)_zG_b$,或者选自SEQ ID NO:67、81-82、93和97中的任一个核苷酸序列其中:G代表鸟嘌呤;x每次出现时独立地为选自2-13的整数;N每次出现时独立地代表任意碱基;y每次出现时独立地为选自1-19的整数;z为选自3-15的整数;a为选自0-8的整数;以及b为选自0和2-10的整数,其中当z=3时,b不为0;其中当N=G且 N_y 包含GT时,y不为9;当N=G且 N_y 包含TGA时,y不为4、15或16。



1. 一种用于提高目的核酸稳定性的具有G-四链体结构的核酸分子,其特征在于包含:



选自SEQ ID NO:67、81-82、93和97中的任一个核苷酸序列;

其中:

G代表鸟嘌呤;

x每次出现时独立地为选自2-13的整数;

N每次出现时独立地代表任意碱基;

y每次出现时独立地为选自1-19的整数,优选1-10,更优选1-6;

z为选自3-15的整数,优选3-10,更优选3-6;

a为选自0-8的整数,优选0-6,更优选0-3;以及

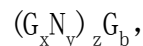
b为选自0和2-10的整数,优选2-9,其中当z=3时,b不为0;

其中当N=G且 N_y 包含GT时,y不为9;当N=G且 N_y 包含TGA时,y不为4、15或16;

其中当所述具有G-四链体结构的核酸分子包含SEQ ID NO:67的核苷酸序列时,所述核酸分子的G-四链体结构由2分子的该核苷酸序列形成;以及

其中当所述具有G-四链体结构的核酸分子包含SEQ ID NO:81或82的核苷酸序列时,所述核酸分子的G-四链体结构由4分子的该核苷酸序列形成。

2. 根据权利要求1所述的具有G-四链体结构的核酸分子,其特征在于包含下式:



其中:

G代表鸟嘌呤;

x每次出现时独立地为选自2-13的整数;

N每次出现时独立地代表任意碱基;

y每次出现时独立地为选自1-19的整数,优选1-10,更优选1-6;

z为选自3-15的整数,优选3-10,更优选3-6;以及

b为选自0和2-10的整数,优选2-9,其中当z=3时,b不为0;以及

其中当N=G且 N_y 包含GT时,y不为9;当N=G且 N_y 包含TGA时,y不为4或15。

3. 根据权利要求1或2所述的具有G-四链体结构的核酸分子,其特征在于其核苷酸序列选自SEQ ID NO: 3、41-66、68-80、83-92、94-96和98-131中的任一个。

4. 一种经保护的核酸,其特征在于所述目的核酸在其5'端、3'端或5'端和3'端两者由权利要求1-3中任一项的具有G-四链体结构的核酸分子缀合或者将权利要求1-3中任一项的具有G-四链体结构的核酸分子插入所述目的核酸中,并且所述目的核酸的稳定性得到增加。

5. 根据权利要求4所述的经保护的核酸,其特征在于所述目的核酸长度为 ≥ 8 nt,优选8-5000 nt;优选地,所述目的核酸为基因敲低用核酸、基因敲除用核酸、基因活化用核酸、基因修饰用核酸、基因编辑用核酸、基因调控用核酸、蛋白表达用核酸、蛋白调控用核酸、生物检测用核酸或核酸药物;更优选地,所述目的核酸为寡核苷酸、单链DNA、双链DNA、单链RNA、双链RNA、mRNA或ncRNA。

6. 根据权利要求5所述的经保护的核酸,其特征在于所述寡核苷酸为反义寡核苷酸ASO或核酸适配体。

7. 根据权利要求5所述的经保护的核酸,其特征在于所述ncRNA为miRNA、siRNA、saRNA、piRNA、lncRNA、circRNA、它们的片段或其它调控性RNA。

8. 一种药物组合物,其特征在于所述药物组合物包含权利要求4-7中任一项所述的经保护的核酸和药学上可接受的载体。

9. 一种用于增加目的核酸的稳定性和/或活性的方法,其特征在于将权利要求1-3中任一项的具有G-四链体结构的核酸分子缀合至所述目的核酸的5'端、3'端或5'端和3'端两者,或者插入所述目的核酸中。

10. 根据权利要求9的方法,其特征在于所述目的核酸长度为 ≥ 8 nt,优选8-5000 nt;优选地,所述目的核酸为基因敲低用核酸、基因敲除用核酸、基因活化用核酸、基因修饰用核酸、基因编辑用核酸、基因调控用核酸、蛋白表达用核酸、蛋白调控用核酸、生物检测用核酸或核酸药物;更优选地,所述目的核酸为寡核苷酸(例如反义寡核苷酸ASO或核酸适配体)、单链DNA、双链DNA、单链RNA、双链RNA、mRNA或ncRNA(例如miRNA、siRNA、saRNA、piRNA、lncRNA、circRNA、它们的片段或其它调控性RNA)。

含有G-四链体的核酸

技术领域

[0001] 本发明涉及生物医药和生物技术领域,特别是涉及一种含有G-四链体的核酸。

背景技术

[0002] 通过利用核酸与蛋白质或其它靶标结合的能力,可以开发多种核酸工具或核酸药物,包括了mRNA药物、反义寡核苷酸(Antisense oligonucleotide, ASO)和核酸适配体(aptamer)及调控性RNA等。其中,ASO通过序列特异地与靶基因RNA结合而抑制该基因表达,被开发为基因敲低的方法和应用于核酸药物领域。核酸适配体可以与靶标(蛋白质、小分子、细胞等)结合,发挥类似“化学抗体”的作用,被应用于生物检测和核酸药物领域。但是,这类核酸均存在易被核酸酶(核酸外切酶、核酸内切酶)降解、与靶标亲和力不稳定、定位及运输困难的问题等。

[0003] 为保护核酸不被降解,已开发了多种方法来对抗核酸酶的降解,包括化学修饰方法和非化学修饰方法两类。目前被广泛应用的方法主要是化学修饰方法,包括磷酸盐骨架修饰、糖修饰、磷酸修饰和碱基修饰等,用于增强核酸的稳定性。但是化学修饰的核酸药物存在潜在的安全性问题,比如通过激活toll样受体引起免疫应答、导致细胞毒性(如诱导凋亡而引起蛋白尿)、影响凝血功能和导致血小板减少等。

[0004] 目前报道的非化学修饰的方法甚少,在1989年报道了能够形成极稳定的二级结构的短链寡核苷酸,1993年Coulson课题组报道了一种热力学稳定的发夹结构hairpin(GCGAAAGC)可以保护寡核苷酸不被3'端核酸外切酶降解,并对该hairpin进行优化,得到min-hairpin(GCGAAGC)。但是该方法的保护效果十分有限。在2002年有专利报道(US7022832 B2)将寡核苷酸的5'端或/和3'端设计为发夹结构,在10%的热失活的血清中的可以起到保护效果,在动物中利用纳米材料的辅助实现基因敲低,但保护效果有限,目前并没有得到广泛应用。在2002年有专利报道(US6121434 A)连续G碱基(其中G的个数为0-10)可以保护寡核苷酸,但该方法通常情况下是要和硫代修饰进行联用,因此,目前也没有得到广泛应用。

[0005] G-四链体(G-quadruplex)是由富含串联重复鸟嘌呤(G)的DNA或RNA折叠形成的高级结构。G-四分体(G-quartet)是四链体的结构单元,由Hoogsteen氢键连接4个G形成环状平面,两层或更多层的G-四分体通过 π - π 堆积形成G-四链体,其是一类具有稳定二级结构的核酸。1991年Za11ler课题组报道了在 K^+ 存在的条件下,折叠成G-四链体结构的端粒DNA不能成为端粒酶的作用底物,并能够抑制端粒酶对端粒的延伸。G-四链体形成的稳定结构或许可以通过与蛋白结合,增强在血清中的稳定性。然而,现有技术尚未提供具有高稳定性和活性的核酸药物。

[0006] 因此,对提高核酸稳定性和/或活性的物质和方法存在未被满足的需求。

发明内容

[0007] 本发明基于以下发现:特定序列的具有G-四链体结构的核酸分子能够用于保护核

酸,增强其在血清中的稳定性,并提高其活性。

[0008] 在第一方面,本公开内容提供了一种用于提高目的核酸稳定性的具有G-四链体结构的核酸分子,其特征包含于包含下式: $N_a(G_xN_y)_zG_b$,或者选自SEQ ID NO:67、81-82、93和97中的任一个核苷酸序列;

其中:

G代表鸟嘌呤;

x每次出现时独立地为选自2-13的整数;

N每次出现时独立地代表任意碱基;

y每次出现时独立地为选自1-19的整数例如1-13的整数,优选1-10,更优选1-6;

z为选自3-15的整数,优选3-10,更优选3-6;

a为选自0-8的整数,优选0-6,更优选0-3;以及

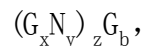
b为选自0和2-10的整数,优选2-9,其中当z=3时,b不为0;

其中当N=G且 N_y 包含GT时,y不为9;当N=G且 N_y 包含TGA时,y不为4、15或16;

其中当所述具有G-四链体结构的核酸分子包含SEQ ID NO:67的核苷酸序列时,所述核酸分子的G-四链体结构由2分子的该核苷酸序列形成;以及

其中当所述具有G-四链体结构的核酸分子包含SEQ ID NO:81或82的核苷酸序列时,所述核酸分子的G-四链体结构由4分子的该核苷酸序列形成。

[0009] 在一些实施方案中,所述具有G-四链体结构的核酸分子包含下式:



其中:

G代表鸟嘌呤;

x每次出现时独立地为选自2-13的整数;

N每次出现时独立地代表任意碱基;

y每次出现时独立地为选自1-19的整数例如1-13的整数,优选1-10,更优选1-6;

z为选自3-15的整数,优选3-10,更优选3-6;以及

b为选自0和2-10的整数,优选2-9,其中当z=3时,b不为0;以及

其中当N=G且 N_y 包含GT时,y不为9;当N=G且 N_y 包含TGA时,y不为4或15。

[0010] 在一些优选的实施方案中,所述的具有G-四链体结构的核酸分子包含下式或如下式所示: $(G_xN_y)_zG_b$,其中:

G代表鸟嘌呤;

x每次出现时独立地为选自2-5的整数,优选为3或4;

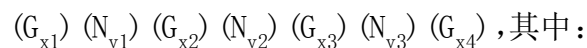
N每次出现时独立地代表T碱基或A碱基或C碱基,优选为T碱基;

y每次出现时独立地为选自1-4的整数,优选1-3;

z为3、4和5的整数倍的整数,优选为3-4的整数倍的整数,最优选为3;

b为选自1-5的整数,优选0-6,更优选0-3。

[0011] 在一些更优选的实施方案中,所述的具有G-四链体结构的核酸分子包含下式或如下式所示:



G代表鸟嘌呤;

x_1-x_4 独立地为2-10的整数,可以相同或不相同;优选地, x_1-x_4 独立地为2-5的整数,可以相同或不相同;更优选地, $x_1=x_2=x_3=x_4=3$;

N为T碱基或A碱基或C碱基,优选为T碱基;

y_1-y_3 独立地为1-5的整数,可以相同或不相同;优选地, y_1-y_3 为1-3的整数,可以相同或不相同;更优选地, $y_1=2, y_2=3$,以及 $y_3=2$ 。

[0012] 在一些实施方案中,所述具有G-四链体结构的核酸分子包含选自SEQ ID NO: 3、41-66、68-80、83-92、94-96和98-131中的任一个的核苷酸序列或由其组成。在一些实施方案中,所述具有G-四链体结构的核酸分子可不包含或不为SEQ ID NO: 3、42-48和110-131中任一个所示的核苷酸序列,或者在具有G-四链体结构的核酸分子包含或为SEQ ID NO: 3、42-48和110-131中任一个所示的核苷酸序列时,在所述核酸分子的3'端缀合一个或多个(多达10个例如2、3、4、5、6、7、8、9或10个)T碱基。

[0013] 在一些实施方案中,所述目的核酸为DNA或RNA。在一些实施方案中,所述目的核酸为修饰的或未经修饰的。在一些实施方案中,所述核酸分子为DNA或RNA。在一些实施方案中,所述具有G-四链体结构的核酸分子为修饰的或未经修饰的例如未经硫代修饰的。在一些实施方案中,G-四链体为分子内、双分子或四分子的G-四链体。

[0014] 在一些实施方案中,所述目的核酸长度为8-5000 nt,例如8-4000nt、8-3000nt、8-2000nt、8-1000nt、8-500 nt、8-200 nt、8-150 nt、8-100 nt、8-80 nt、8-50 nt、8-40 nt、8-30 nt、8-20 nt、8-10 nt、30-180 nt、20-100 nt、30-50 nt等。

[0015] 在一些实施方案中,所述目的核酸为基因敲低用核酸、基因敲除用核酸、基因活化用核酸、基因修饰用核酸、基因编辑用核酸、基因调控用核酸、蛋白调控用核酸、蛋白表达用核酸、生物检测用核酸或核酸药物。

[0016] 在一些实施方案中,所述目的核酸为寡核苷酸、单链DNA、双链DNA、单链RNA、双链RNA、mRNA或ncRNA(非编码RNA)。

[0017] 在一些实施方案中,所述寡核苷酸为反义寡核苷酸ASO或核酸适配体。

[0018] 在一些实施方案中,所述ncRNA为miRNA(微小RNA)、siRNA(小干扰RNA)、saRNA(小激活RNA)、piRNA(与Piwi蛋白相作用的RNA)、lncRNA(长非编码RNA)、circRNA(环状RNA)、它们的片段或其它调控性RNA。

[0019] 在一些实施方案中,可在本发明的具有G-四链体结构的核酸分子缀合一个或多个(多达10个)T碱基而不影响其保护效果或活性。

[0020] 在第二方面,本公开内容提供了一种经保护的核酸,其特征在于所述目的核酸在其5'端、3'端或5'端和3'端两者由第一方面的具有G-四链体结构的核酸分子缀合或者将第一方面的具有G-四链体结构的核酸分子插入所述目的核酸中,并且所述目的核酸的稳定性得到增加。

[0021] 在一些实施方案中,所述的缀合为化学共价连接,优选为通过磷酸二酯键连接。

[0022] 在一些实施方案中,构建所述经保护的核酸的方式为:化学合成法、基于PCR原理的基因工程方法或生物合成方法。

[0023] 在一些实施方案中,所述目的核酸为DNA或RNA。在一些实施方案中,所述经保护的核酸为修饰的或未经修饰的。

[0024] 在一些实施方案中,所述目的核酸长度为8-5000 nt,例如8-4000nt、8-3000nt、8-

2000nt、8-1000nt、8-500 nt、8-200 nt、8-150 nt、8-100 nt、8-80 nt、8-50 nt、8-40 nt、8-30 nt、8-20 nt、8-10 nt、30-180 nt、20-100 nt、30-50 nt等。在一些实施方案中,所述目的核酸为基因敲低用核酸、基因敲除用核酸、基因活化用核酸、基因修饰用核酸、基因编辑用核酸、基因调控用核酸、蛋白表达用核酸、蛋白调控用核酸、生物检测用核酸或核酸药物。在一些实施方案中,所述目的核酸为寡核苷酸、单链DNA、双链DNA、单链RNA、双链RNA、mRNA或ncRNA(非编码RNA)。在一些实施方案中,所述寡核苷酸为反义寡核苷酸ASO或核酸适配体。在一些实施方案中,所述ncRNA为miRNA(微小RNA)、siRNA(小干扰RNA)、saRNA(小激活RNA)、piRNA(与Piwi蛋白相作用的RNA)、lncRNA(长非编码RNA)、circRNA(环状RNA)、它们的片段或其它调控性RNA。

[0025] 在第三方面,本公开内容提供了一种药物组合物,其特征在于所述药物组合物包含第二方面的经保护的核酸和药学上可接受的载体。

[0026] 在第四方面,本公开内容提供了一种用于增加目的核酸的稳定性和/或活性的方法,其特征在于将第一方面的具有G-四链体结构的核酸分子缀合至所述目的核酸的5'端、3'端或5'端和3'端两者,或者插入所述目的核酸中。

[0027] 在一些实施方案中,所述的缀合为化学共价连接,优选为通过磷酸二酯键连接。

[0028] 在一些实施方案中,构建所述经保护的核酸的方式为:化学合成法、基于PCR原理的基因工程方法或生物合成方法。

[0029] 在一些实施方案中,所述目的核酸为DNA或RNA。在一些实施方案中,所述经保护的核酸为修饰的或未经修饰的。

[0030] 在一些实施方案中,所述目的核酸长度为 ≥ 8 nt,例如8-5000 nt如8-4000nt、8-3000nt、8-2000nt、8-1000nt、或8-500 nt,优选8-200 nt,例如8-150 nt、8-100 nt、8-80 nt、8-50 nt、8-40 nt、8-30 nt、8-20 nt、8-10 nt、30-180 nt、20-100 nt、30-50 nt等。在一些实施方案中,所述目的核酸为基因敲低用核酸、基因敲除用核酸、基因活化用核酸、基因修饰用核酸、基因编辑用核酸、基因调控用核酸、蛋白表达用核酸、蛋白调控用核酸、生物检测用核酸或核酸药物。在一些实施方案中,所述目的核酸为寡核苷酸、单链DNA、双链DNA、单链RNA、双链RNA、mRNA或ncRNA。在一些实施方案中,所述寡核苷酸为反义寡核苷酸ASO或核酸适配体。在一些实施方案中,所述ncRNA为miRNA、siRNA、saRNA、piRNA、lncRNA、circRNA、它们的片段或其它调控性RNA。

[0031] 下列描述和实施例详细阐述了本发明的实施方案。要理解的是,本发明不限于本文所述的具体实施方案并因此可改动。本领域技术人员将认识的是,存在本发明的许多变动和修改,所述变动和修改均包含在其范围之内。

附图说明

[0032] 上述仅是本发明技术方案的概述,为了能够更清楚了解本发明的技术手段,以下结合附图与具体实施方式对本发明作进一步的详细说明。

[0033] 图1显示了含有G1的寡核苷酸药物在50%血清中的稳定性;

图2显示了含有G1的核酸药物与目前常用的核酸保护方法保护的核酸药物的稳定性对比结果图;

图3显示了在5'端或者3'端含有G1的核酸药物在50%血清中的稳定性;

图4显示了含有DNA G-四链体的核酸药物的稳定性；

图5显示了基于G1改造的G-四链体对核酸药物稳定性的影响；

图6显示了含有DNA G-四链体的不同长度的核酸药物在50%血清中的稳定性(仅代表性的提供G1 (SEQ ID NO: 3)的电泳结果)；

图7显示了含有DNA G-四链体的不同序列的核酸药物在50%血清中的稳定性(仅代表性的提供G1 (SEQ ID NO: 3)的电泳结果)；

图8显示了含有RNA G-四链体的不同序列的RNA药物在50%血清中稳定性(仅代表性的提供G1 (SEQ ID NO: 3)的电泳结果)；

图9显示了含有G-四链体的长链RNA药物在血清中的稳定性(仅代表性的提供G1 (SEQ ID NO: 3)的电泳结果)；

图10显示了含有RNA G-四链体的核酸药物嵌合体以及含有DNA G-四链体的核酸药物嵌合体药物在血清中的稳定性(仅代表性的提供G1 (SEQ ID NO: 3)的电泳结果)；

图11显示了含有G-四链体的寡核苷酸药物在细胞内对靶标RNA的敲低效果图(仅代表性的提供G1 (SEQ ID NO: 3)的电泳结果)；和

图12显示了采用含有G-四链体的靶向PCSK9的反义寡核苷酸在动物体内敲低PCSK9 (a)和降低LDL-c (b)的结果图(仅代表性的提供G1 (SEQ ID NO: 3)的电泳结果)。

具体实施方式

[0034] 参考用于说明的示例应用在下文中描述本发明的数个方面。应当理解的是,陈述许多具体细节、关系和方法来提供对本发明的充分理解。然而,在相关领域的普通技术人员将容易地认识到,可在不含一个或多个具体细节的情况下实施本发明或者可用其他方法来实施本发明。

[0035] 现有技术的保护核酸药物不被降解的方法包括化学修饰方法和非化学修饰方法两类,但均有其局限性。本申请的发明人首创性地利用具有特定序列的G-四链体对核酸药物进行保护,显著地增强了其在血清中的稳定性和/或活性。

[0036] 相比于现有技术,本发明具有如下有益效果:

1. 本发明核酸分子和方法能够在无需任何修饰的情况下增强核酸在血清中的稳定性;

2. 本发明的核酸分子和方法可以增强反义寡核苷酸的活性,在细胞水平和整体动物水平验证了ASO敲低RNA的能力,说明本发明可以应用于基因敲低和核酸药物领域;

3. 经验证,相对于现有常规的化学修饰的保护方法和非化学修饰的保护方法,本发明的核酸药物在50%的血清中具有稳定性高的优点。

[0037] 本文所用的术语仅以描述具体的实施方案为目的而不意图限制本发明。除非上下文另有明确指示,否则本文所用的单数形式“一”、“一个”和“所述”也意图包括复数形式。此外,开放式的表述“包括”和“包含”解释为还可以含有没有述及的结构组成部分或方法步骤,但需要注意的是,该开放式的表述也涵盖仅由所述的组分和方法步骤组成的情形(即涵盖了封闭式表述“由……组成”的情形)。

[0038] 如全文所用,范围用作描述该范围内的每个数值和所有数值的简写形式。范围内的任何数值例如整数都可选做该范围的终点。例如范围2-13的整数表示整数2、3、4、5、6、

7、8、9、10、11、12和13及它们形成的任何子范围例如2-10、3-12、4-9等。

[0039] 本说明书中提及的所有科学技术术语具有与本领域技术人员通常理解的含义相同的含义,在冲突的情况下以本说明书中的定义为准。为了使本发明的描述更易于理解,以下对一些术语进行解释说明。

[0040] 在本发明中,术语“核酸”、“核酸分子”、“寡核苷酸药物”和“核酸药物”可互换使用,指任何DNA、RNA或DNA/RNA嵌合体,并且可以为寡核苷酸或多核苷酸,并且可以是未修饰的RNA或DNA或者修饰的RNA或DNA。该术语包括但不限于单链和双链的DNA、为单链区和双链区的混合物的DNA、单链和双链的RNA以及为单链区和双链区的混合物的RNA、包含可以是单链或双链或者是单链区和双链区的混合物的DNA和RNA的杂合分子。

[0041] 在一些实施方案中,核酸可包含一个或多个经修饰的核苷酸、经修饰的键等。经修饰的键或核苷酸间键的实例包括硫代磷酸酯、二硫代磷酸酯等。在一些实施方案中,所述核苷酸包括磷衍生物。可连接到本发明的修饰的核苷酸中的糖或糖类似物部分的磷衍生物(或经修饰的磷酸基),可为单磷酸酯、二磷酸酯、三磷酸酯、烷基磷酸酯、链烷磷酸酯、硫代磷酸酯等。上述磷酸酯类似物的制备,以及它们掺入到核苷酸、修饰核苷酸和核酸中本身也为已知的且无需在此描述。

[0042] 对于核酸修饰

(1)核酸结构

在一些实施方案中,本发明核酸(例如,靶向DNA的RNA)包含一个或多个修饰(例如,碱基修饰、骨架修饰等)以对核酸提供新的或增强的特征(例如,改进的稳定性)。如本领域中所已知,核苷为碱基-糖组合。核苷的碱基部分通常为杂环碱基。此类杂环碱基的两个最常见类别为嘌呤和嘧啶。核苷酸为还包括共价连接至核苷的糖部分的磷酸酯基团的核苷。对于包括呋喃戊糖的那些核苷,磷酸酯基团可连接至糖的2'、3'或5'羟基部分。在形成核酸中,磷酸酯基团共价连接彼此相邻的核苷以形成线性聚合化合物。反过来,线性聚合化合物的各端可进一步连接以形成环状化合物,然而,线性化合物通常为适合的。另外,线性化合物可具有内部核苷酸碱基互补性并且因此可以为了产生完全或部分双链化合物的方式折叠。在核酸内,磷酸酯基团通常称为形成核酸的核苷间骨架。RNA和DNA的正常键或骨架为3'至5'的磷酸二酯键。

[0043] (2)修饰的骨架和修饰的核苷间键

含有修饰的适合的核酸的实例包括含有修饰的骨架或非天然的核苷间键的核酸。核酸(具有修饰的骨架)包括在骨架中保留磷原子的那些和在骨架中不具有磷原子的那些。

[0044] 其中含有磷原子的适合的修饰的核酸包括例如,硫代磷酸酯、手性硫代磷酸酯、二硫代磷酸酯、磷酸三酯、氨基烷基磷酸三酯、甲基和其它烷基磷酸酯包括3'-亚烷基磷酸酯、5'-亚烷基磷酸酯和手性磷酸酯、膦酸盐、氨基磷酸酯包括3'-氨基氨基磷酸酯和氨基烷基氨基磷酸酯、二氨基磷酸酯、硫羰氨基磷酸酯、硫羰烷基磷酸酯、硫羰烷基磷酸三酯、具有正常3'-5'键的硒代磷酸酯和硼代磷酸酯、这些的2'-5'连接类似物以及具有反极性的那些,其中一个或多个核苷酸间键为3'至3'、5'至5'或2'至2'键。具有反极性的适合的核酸在最3'核苷酸间键处包含单个3'至3'键,即可为无碱基(核碱基丢失或其被羟基替代)的单个反转核苷残基。还包括各种盐(例如像钾或钠)、混合盐和游离酸形式。

[0045] 在一些实施方案中,本发明核酸包含一个或多个硫代磷酸酯键和/或杂原子核苷

间键,具体为 $-\text{CH}_2-\text{NH}-\text{O}-\text{CH}_2-$ 、 $-\text{CH}_2-\text{N}(\text{CH}_3)-\text{O}-\text{CH}_2-$ (称为亚甲基(甲基亚氨基)或MMI骨架)、 $-\text{CH}_2-\text{O}-\text{N}(\text{CH}_3)-\text{CH}_2-$ 、 $-\text{CH}_2-\text{N}(\text{CH}_3)-\text{N}(\text{CH}_3)-\text{CH}_2-$ 以及 $-\text{O}-\text{N}(\text{CH}_3)-\text{CH}_2-\text{CH}_2-$ (其中天然磷酸二酯核苷间键表示为 $-\text{O}-\text{P}(=\text{O})(\text{OH})-\text{O}-\text{CH}_2-$)。

[0046] 其他的骨架修饰还包括例如吗啉代骨架结构的核酸。例如,在一些实施方案中,本发明核酸包含替代核糖环的6元吗啉代环。在这些实施方案的一些中,二氨基磷酸酯或其它非磷酸二酯核苷间键替代磷酸二酯键。

[0047] 不包括磷原子的适合的修饰的多核苷酸骨架具有通过短链烷基或环烷基核苷间键、混合杂原子和烷基或环烷基核苷间键或一个或多个短链杂原子或杂环核苷间键形成的骨架。这些包括具有以下的那些:吗啉代键(部分地从核苷的糖部分中形成);硅氧烷骨架;硫化物、亚砷和砷骨架;甲酰乙酰基和硫代甲酰乙酰基骨架;亚甲基甲酰乙酰基和硫代甲酰乙酰基骨架;核糖乙酰基(riboacetyl)骨架;含烯烃的骨架;氨基磺酸酯骨架;亚甲基亚胺基和亚甲基胍基骨架;磺酸酯和磺酰胺骨架;酰氨骨架;以及具有混合N、O、S和 CH_2 组成部分的其它骨架。

[0048] 另一种骨架修饰包括锁定核酸(LNA),其中2'-羟基连接至糖环的4'碳原子从而形成2'-C、4'-C-氧基亚甲基键,从而形成双环糖部分。链可为亚甲基($-\text{CH}_2-$) (桥联2'氧原子和4'碳原子的基团),其中n为1或2(Singh等,Chem. Commun., 1998, 4, 455-456)。LNA和LNA类似物显现出与互补DNA和RNA具有非常高的双链体热稳定性($T_m = +3^\circ\text{C}$ 至 $+10^\circ\text{C}$)、朝向3'-核酸外切降解的稳定性和良好的溶解性质。

[0049] 现有技术中已描述了LNA单体腺嘌呤、胞嘧啶、鸟嘌呤、5-甲基-胞嘧啶、胸腺嘧啶和尿嘧啶的合成和制备连同其低聚化以及核酸识别性质(Koshkin等,Tetrahedron, 1998, 54, 3607-3630)。

[0050] (3) 修饰的糖部分

本发明核酸还可包括一个或多个取代的糖部分。适合的多核苷酸包含选自以下的糖取代基:OH;F;O-、S-或N-烷基;O-、S-或N-烯基;O-、S-或N-炔基;或O-烷基-O-烷基,其中烷基、烯基和炔基可为取代或未取代的C1至C10烷基或C2至C10烯基和炔基。特别适合的为 $\text{O}((\text{CH}_2)_n)_m\text{CH}_3$ 、 $\text{O}(\text{CH}_2)_n\text{OCH}_3$ 、 $\text{O}(\text{CH}_2)_n\text{NH}_2$ 、 $\text{O}(\text{CH}_2)_n\text{CH}_3$ 、 $\text{O}(\text{CH}_2)_n\text{ONH}_2$ 和 $\text{O}(\text{CH}_2)_n\text{N}((\text{CH}_2)_n\text{CH}_3)_2$,其中n和m为1至约10。其它适合的多核苷酸包含选自以下的糖取代基:C1至C10低级烷基、取代的低级烷基、烯基、炔基、烷芳基、芳烷基、O-烷芳基或O-芳烷基、SH、SCH₃、OCN、Cl、Br、CN、CF₃、OCF₃、SOCH₃、SO₂CH₃、ONO₂、NO₂、N₃、NH₂、杂环烷基、杂环烷芳基、氨基烷基氨基、多烷基氨基、取代的甲硅烷基、RNA裂解基团、报道基团、嵌入剂、用于改进核酸的药物代谢动力学性质的基团或用于改进核酸的药效性质的基团以及具有类似性质的其它取代基。适合的修饰包括2'-甲氧基乙氧基(2'-O-CH₂CH₂OCH₃,又称为2'-O-(2-甲氧基乙基)或2'-MOE)(Martin等,Helv.Chim.Acta, 1995, 78, 486-504),即烷氧基烷氧基。另一适合的修饰包括2'-二甲基氨基氧基乙氧基,即 $\text{O}(\text{CH}_2)_2\text{ON}(\text{CH}_3)_2$ 基团,又称为2'-DMAOE,如在下文的实施例所述;和2'-二甲基氨基乙氧基乙氧基(在本领域中又称为2'-O-二甲基-氨基-乙氧基-乙基或2'-DMAEOE),即 $2'-\text{O}-\text{CH}_2-\text{O}-\text{CH}_2-\text{N}(\text{CH}_3)_2$ 。

[0051] 其它适合的糖取代基包括甲氧基(-O-CH₃)、氨基丙氧基(-OCH₂CH₂CH₂NH₂)、烯丙基(-CH₂-CH=CH₂)、O-烯丙基(-O-CH₂-CH=CH₂)以及氟(F)。2'-糖取代基可处于阿拉伯糖(上)位或核糖(下)位。适合的2'-阿拉伯糖修饰为2'-F。还可在低聚化合物上的其它位置上

做出类似的修饰,具体地在3'末端核苷上或在2'-5'连接的核酸中的糖的3'位置和5'末端核苷酸的5'位置。低聚化合物还可具有替代呋喃戊糖的糖模拟物如环丁基部分。

[0052] (4) 碱基修饰和取代

本发明核酸还可包括核碱基(在本领域中常常简称为“碱基”)修饰或取代。如本文所使用,“未修饰”或“天然”核碱基包括嘌呤碱基腺嘌呤(A)和鸟嘌呤(G)以及嘧啶碱基胸腺嘧啶(T)、胞嘧啶(C)和尿嘧啶(U)。修饰的核碱基包括其它合成和天然的核碱基如5-甲基胞嘧啶(5-me-C)、5-羟甲基胞嘧啶、黄嘌呤、次黄嘌呤、2-氨基腺嘌呤、腺嘌呤和鸟嘌呤的6-甲基衍生物和其它烷基衍生物、腺嘌呤和鸟嘌呤的2-丙基衍生物和其它烷基衍生物、2-硫尿嘧啶、2-硫胸腺嘧啶和2-硫胞嘧啶、5-卤代尿嘧啶和胞嘧啶、5-丙炔基(-C=C-CH₃)尿嘧啶和胞嘧啶以及嘧啶碱基的其它炔基衍生物、6-偶氮基尿嘧啶、胞嘧啶和胸腺嘧啶、5-尿嘧啶(假尿嘧啶)、4-硫尿嘧啶、8-卤代、8-氨基、8-巯基、8-硫烷基、8-羟基和其它8-取代的腺嘌呤和鸟嘌呤、5-卤代(具体为5-溴)、5-三氟甲基和其它5-取代的尿嘧啶和胞嘧啶、7-甲基鸟嘌呤和7-甲基腺嘌呤、2-F-腺嘌呤、2-氨基-腺嘌呤、8-氮杂鸟嘌呤和8-氮杂腺嘌呤、7-脱氮鸟嘌呤和7-脱氮腺嘌呤以及3-脱氮鸟嘌呤和3-脱氮腺嘌呤。另外的修饰的核碱基包括三环嘧啶如吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮)、吩噻嗪胞苷(1H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮)、G-夹如取代的吩噻嗪胞苷(例如9-(2-氨基乙氧基)-H-嘧啶并(5,4-b)(1,4)苯并噻嗪-2(3H)-酮)、咪唑胞苷(2H-嘧啶并(4,5-b)吡啶-2-酮)、吡啶并咪唑胞苷(H-吡啶并(3',2':4,5)吡咯并(2,3-d)嘧啶-2-酮)。

[0053] 杂环碱基部分还可包括其中嘌呤或嘧啶碱基被其它杂环替代的那些,例如7-脱氮腺嘌呤、7-脱氮鸟苷、2-氨基吡啶和2-吡啶酮。另外的核碱基包括公开于The Concise Encyclopedia Of Polymer Science And Engineering,第858-859页,Kroschwitz,J.I.编著John Wiley&Sons,1990中的那些、由Angewandte Chemie,International Edition,1991,30,613公开的那些以及由Sanghvi,Y.S.,第15章,Antisense Research and Applications,第289-302页,Crooke,S.T.和Lebleu,B.编著,CRC Press,1993公开的那些。这些核碱基中的某些有用于增加低聚化合物的结合亲和力。这些包括5-取代的嘧啶、6-氮杂嘧啶以及N-2、N-6和O-6取代的嘌呤,包括2-氨基丙基腺嘌呤、5-丙炔基尿嘧啶和5-丙炔基胞嘧啶。5-甲基胞嘧啶取代已显示出使核酸双链体稳定性增加0.6°C-1.2°C(Sanghvi等编著,Antisense Research and Applications,CRC Press,Boca Raton,1993,第276-278页)并且例如当与2'-O-甲氧基乙基糖修饰组合时为适合的碱基取代。

[0054] (5) 缀合物

本发明核酸的另一种可能的修饰涉及将增强核酸的活性、细胞分布或细胞吸收的一个或多个部分或缀合物化学连接至多核苷酸。这些部分或缀合物可包括共价结合如伯羟基或仲羟基的官能团的缀合物基团。缀合物基团包括但不限于嵌入剂、报道分子、多胺、聚酰胺、聚乙二醇、聚醚、增强低聚物药效性质的基团以及增强低聚物药物代谢动力学性质的基团。适合的缀合物基团包括但不限于胆固醇、脂质、磷脂、生物素、吩噻嗪、叶酸酯、菲啶、葱醌、吡啶、荧光素、罗丹明、香豆素以及染料。增强药效性质的基团包括改进吸收、增强对降解的抗性和/或加强与靶核酸的序列特异性杂交的基团。增强药物代谢动力学性质的基团包括改进本发明核酸的吸收、分布、代谢或排泄的基团。

[0055] 缀合物部分包括但不限于:脂质部分如胆固醇部分(Letsinger等,

Proc. Natl. Acad. Sci. USA, 1989, 86, 6553-6556)、胆酸 (Manoharan等, Bioorg. Med. Chem. Lett., 1994, 4, 1053-1060)、硫醚例如己基-S-三苯甲基硫醇 (Manoharan等, Ann. N.Y. Acad. Sci., 1992, 660, 306-309; Manoharan等, Bioorg. Med. Chem. Lett., 1993, 3, 2765-2770)、硫代胆固醇 (Oberhauser等, Nucl. Acids Res., 1992, 20, 533-538)、脂肪族链例如十二烷二醇或十一烷基残基 (Saison-Behmoaras等, EMBO J., 1991, 10, 1111-1118; Kabanov等, FEBS Lett., 1990, 259, 327-330; Svinarchuk等, Biochimie, 1993, 75, 49-54)、磷脂例如二-十六烷基-外消旋甘油或1,2-二-0-十六烷基-外消旋甘油-3-H-磷酸三乙基铵盐 (Manoharan等, Tetrahedron Lett., 1995, 36, 3651-3654; Shea等, Nucl. Acids Res., 1990, 18, 3777-3783)、多胺或或聚乙二醇链 (Manoharan等, Nucleosides & Nucleotides, 1995, 14, 969-973)、或金刚烷乙酸 (Manoharan等, Tetrahedron Lett., 1995, 36, 3651-3654)、棕榈基部分 (Mishra等, Biochim. Biophys. Acta, 1995, 1264, 229-237)、或十八烷基胺或己基氨基-羰基-氧基胆固醇部分 (Crooke等, J. Pharmacol. Exp. Ther., 1996, 277, 923-937)。

[0056] 本文所用的“未修饰的”或“天然的”核苷酸包括腺嘌呤 (A)、鸟嘌呤 (G)、胸腺嘧啶 (T)、胞嘧啶 (C) 和尿嘧啶 (U)。修饰核苷酸包括在天然核酸中仅稀少或短暂地存在的核苷酸, 例如, 次黄嘌呤、6-甲基腺嘌呤、5-Me嘧啶、特别地5-甲基胞嘧啶 (也称为5-甲基-2' 脱氧胞嘧啶且常常在本领域中称为5-Me-C)、5-羟甲基胞嘧啶 (HMC)、糖基HMC和龙胆二糖基HMC, 以及合成核苷酸, 例如, 2-氨基腺嘌呤、2-(甲氨基)腺嘌呤、2-(咪唑基烷基)腺嘌呤、2-(氨烷基氨基)腺嘌呤或其他杂取代的烷基腺嘌呤、2-硫尿嘧啶、2-硫胸腺嘧啶、5-溴尿嘧啶、5-羟基甲基尿嘧啶、8-氮杂鸟嘌呤、7-脱氮杂鸟嘌呤、N6(6-氨基己基)腺嘌呤和2,6-二氨基嘌呤。可包括本领域已知的“通用的”碱基, 例如, 肌苷。

[0057] 本文使用的术语“G-四链体”是由富含串联重复鸟嘌呤 (G) 的DNA或RNA折叠形成的高级结构。G-四分体是四链体的结构单元, 由Hoogsteen氢键连接4个G形成环状平面, 两层或更多层的G-四分体通过 π - π 堆积形成G-四链体, 其是一类具有稳定二级结构的核酸 (参见 Zahler, A. M., Williamson, J. R., Cech, T. R. & Prescott, D. M. Inhibition of telomerase by G-quartet DNA structures. *Nature* 350, 718-720 (1991), 其通过引用并入本文用于所有目的)。G-四链体通过阳离子, 特别是钾的存在进一步稳定, 阳离子位于每对四分体之间的中心通道中。G-四链体可以由DNA、RNA、LNA和PNA形成, 并且可以是分子内、双分子或四分子的G-四链体。取决于形成四分体的股线或股线的部分的方向, 结构可描述为平行或反平行 (参见例如 Parkinson G N, Lee M P H, Neidle S, Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* 417 (6891): 876-880 (2002); Wang Y, Patel D J, Solution structure of the human telomeric repeat d[AG₃(T₂AG₃)₃] G-tetraplex. *Structure* 1(4): 263-282 (1993); 和 Dai J. Carver M., Yang D., Polymorphism of human telomeric quadruplex structures. *Biochimie*. 90(8): 1172-1183 (2008), 其通过引用并入本文用于所有目的)。

[0058] 在一些实施方案中, 本发明的具有G-四链体结构的核酸分子包含: $N_a(G_x N_y)_z G_b$, 或者选自SEQ ID NO: 67, 81-82, 93和97中的任一个核苷酸序列; 其中: G代表鸟嘌呤; x每次出现时独立地为选自2-13的整数; N每次出现时独立地代表任意碱基; y每次出现时独立地为

选自1-19的整数例如1-13的整数,优选1-10,更优选1-6; z为选自3-15的整数,优选3-10,更优选3-6; a为选自0-8的整数,优选0-6,更优选0-3; 以及b为选自0和2-10的整数,优选2-9,其中当z=3时, b不为0; 其中当N=G且N_y包含GT时, y不为9; 当N=G且N_y包含TGA时, y不为4、15或16; 其中当所述核酸分子包含SEQ ID NO:67的核苷酸序列时, 所述核酸分子的G-四链体结构由2分子的该核苷酸序列形成; 以及其中当所述核酸分子包含SEQ ID NO:81或82的核苷酸序列时, 所述核酸分子的G-四链体结构由4分子的该核苷酸序列形成。

[0059] 在一些实施方案中, x每次出现时独立地选自整数2、3、4、5、6、7、8、9、10、11、12和13及它们形成的任何子范围例如2-10、3-12、4-9等。在一些实施方案中, y每次出现时独立地选自整数1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18和19及它们形成的任何子范围例如2-10、3-13、1-15、5-10、2-13等。在一些实施方案中, z选自整数3、4、5、6、7、8、9、10、11、12、13、14和15及它们形成的任何子范围例如3-9、4-12、5-10、7-15、6-14等。在一些实施方案中, a选自0、1、2、3、4、5、6、7和8及它们形成的任何子范围例如2-8、0-7、3-6、4-7等。在一些实施方案中, b选自0、2、3、4、5、6、7、8、9和10及它们形成的任何子范围例如2-8、3-9、4-6、5-10等。在一些实施方案中, 本发明的G-四链体可为分子内、双分子或四分子的G-四链体。在一些实施方案中, 本发明的G-四链体可包含阳离子, 例如一价阳离子如钾、钠离子等以及二价阳离子例如镁、钙、锌、铜离子等。

[0060] 本发明的核酸分子以及经本发明核酸分子保护的核酸例如核酸药物可用于各种应用, 例如用于进行基因敲低、基因敲除、基因活化、基因修饰、基因编辑、基因调控、蛋白表达、蛋白调控或生物检测或用于作为核酸药物。因此, 本发明预期本发明的核酸分子以及经本发明核酸分子保护的核酸在这方面的应用, 例如用于制备用于上述应用的药物的用途等。

[0061] 在本发明中, 术语“缀合”和“连接”可互换使用, 指通过化学键连接。

[0062] 在本文中, 本发明所述的核酸在血清中的稳定性, 是指核酸在各种浓度的血清例如50%的血清中的稳定性。在本文中, 稳定性的提高是指与未经本发明核酸分子保护或用其它方法保护的核酸相比, 所述目的核酸的稳定性例如在血清中的稳定性得到提高。在本文中, 目的核酸活性的提高是指与未经本发明核酸分子保护或用其它方法保护的核酸相比, 经本发明的核酸分子保护后目的核酸的生物活性得到提高。在一些实施方案中, 所述生物活性选自基因敲低活性、基因敲除活性、基因活化活性、基因修饰活性、基因编辑活性、基因调控活性、蛋白表达活性、蛋白调控活性或生物检测活性或用于作为核酸药物的活性等。在一些实施方案中, 目的核酸的活性等同或类似于未经本发明核酸分子保护或用其它方法保护的核酸的活性, 但其稳定性得到提高。

[0063] 本文使用的术语“药学上可接受的载体”旨在包括与药物给予相容的任何和所有溶剂、分散介质、包衣剂、抗细菌和抗真菌剂、等渗剂和吸收延迟剂等。合适的载体描述于最新版本Remington's Pharmaceutical Sciences (本领域的标准参考文本), 其通过引用结合于本文中。这种载体或稀释剂的优选实例包括但不限于水、盐水、林格氏溶液、右旋糖溶液和5%人血清白蛋白。这种介质和试剂的使用为本领域熟知的。除非任何常规介质或试剂与本文提供的物质不相容, 否则考虑在组合物中使用它们。

[0064] 以下以具体实施例来说明本发明的方案, 但并非是对本发明的技术方案的限制, 本领域技术人员应该理解, 依然可以对发明进行修改或等同替换, 而不脱离本发明的精神

和范围的任何修改或局部替换,其均应涵盖在本发明的保护范围之内。

实施例

[0065] 除非另外说明,否则本文实施例所用的材料均市购获得,用于进行实验的各种具体实验方法均为本领域常规的实验方法或者按照制造商所建议的步骤和条件,并能由本领域技术人员根据需要常规地确定。此外,本说明书实施例中所用血清均为Gibco特级胎牛血清。降解体系包含:核酸药物、1×PBS、50%血清和水。所用电泳条件均为:180V,20min;上样量为:2 μg-5 μg。除非另外说明,否则本说明书实施例中所用的用于进行保护的序列均为G-四链体。除非另外说明,否则本说明书实施例中的G-四链体均添加至受保护的序列的两端。除非另外说明,否则本说明书实施例中所用的序列均为脱氧核糖核苷酸序列。

[0066] 实施例1:评估含有G1的寡核苷酸药物在50%血清中的稳定性

在需要被保护的脱氧核糖核苷酸(S1)的5'端和3'端分别添加本发明的特定DNA序列(G1),并在50%血清中孵育不同时间:24 h、48 h和72 h,然后通过3%琼脂糖凝胶电泳分析结果,其中选取随机序列Rs作为阴性对照是因为寡核苷酸本身长度较短,很容易降解,因此选择了与保护后的序列长度基本一致的随机对照序列,使其电泳结果图几乎在一条直线,方便对照。结果如图1所示,不添加本发明的特定序列的随机序列Rs作为阴性对照(NC)在6 h时基本完全降解,而G1-S1在72h依然存在没被降解的完整的DNA序列,说明含有G1的寡核苷酸药物在50%的血清中具有良好的稳定性。发明人也发现,采用S1作为阴性对照的对比效果更明显。

[0067] 在实施例1中,使用的序列的序列信息如下:

S1:TTTGAATGTAGAGATGCGGTGGTTT (SEQ ID NO: 1)

Rs: ACCCGACCTCTTCTATCTGGACCCGACCGTCTCTTTTTTGGAGCCCACACTCTACTCGAC (SEQ ID NO: 2)

G1:GGGTTGGGTTTGGGTTGGG (SEQ ID NO: 3)。

[0068] 实施例2:比较含有G1的核酸药物与目前常用的核酸保护方法保护的核酸药物的稳定性

为评估含有G1的核酸药物的稳定性,发明人选择已经报道的核酸保护方法(化学修饰和非化学修饰)进行对比,设置了4个对照组,其中对照组1是化学修饰的,包括2'OMe-Ps、3'-dT、5'-FAM和Ps;对照组2为非化学修饰的,包括G-cap序列修饰(G-cap-1、G-cap-2、G-cap-3、G-cap-4、G-cap-5、G-cap-6)、mini-hairpin序列修饰及发夹序列修饰(H-10、DH-6、L-10);对照组3(G-cap-7、G-cap-8、G-cap-9)为根据专利(US20020151512A1)内容对G-cap进行了化学修饰(硫代修饰,以*表示)得到的序列;对照组4为用已报道专利(US20210139893A1)中的保护方法保护的序列(P1-P20)。将上述修饰施加于S1核酸药物或将上述序列添加到S1核酸药物两端,并分别在50%血清中孵育24 h,然后通过3%琼脂糖凝胶电泳分析结果。结果如图2所示,除本发明方法保护的核酸外,采用化学修饰和非化学修饰方法保护的核酸在50%血清中孵育24 h后,基本全部被降解,而含G1的核酸存在大量的未降解的完整的DNA序列,说明含G1的核酸在50%的血清中具有稳定性高的优点。

[0069] 在实施例2中,使用的序列的序列信息如下:

S1:TTTGAATGTAGAGATGCGGTGGTTT (SEQ ID NO: 1)

H10:TGAACACGCCATGTCGATTCTTTAGAATCGACA (SEQ ID NO: 4)
DH6:TGTTTCATCTGAACACGCCATGTCGATTCTTTAGAATC (SEQ ID NO: 5)
L-10:GCGCTTATGAACACGCCATGTCGATTCTTAAGCGC (SEQ ID NO: 6)
G-cap-1:GGGGCTCCTGGAGCGGGGCACAC (SEQ ID NO: 7)
G-cap-2:GGGGTCGACACCCAATTCTGAAAATGGATAA (SEQ ID NO: 8)
G-cap-3:GGGGACACCCAATTCTGAAAATGGGG (SEQ ID NO: 9)
G-cap-4:GGGAGGTCCCTGTTCCGGGCGCCAGGGG (SEQ ID NO: 10)
G-cap-5:GGGGACACCCAATTCTGAAAATGGGG (SEQ ID NO: 11)
G-cap-6:GGGGGAAGGAGGAGGATGAGGGGG (SEQ ID NO: 12)
G-cap-7:G*G*G*AGGT*CC*C*TGT*T*CGGGCGC*CAG*G*G*G (SEQ ID NO: 13)
G-cap-8:G*G*G*GACACCCAATTCTGAAAATG*G*G*G (SEQ ID NO: 14)
G-cap-9:G*G*G*GAAGGAGGAGGAT*GAGGG*G*G (SEQ ID NO: 15)
P1:TTGGGGTTAGGGTTAGGGTTAGGGATGAATGTAGAGATGCGGTGGTTGGGGTTAGGGTTAGGGTT
AGGGA (SEQ ID NO: 16)
P2:TTGGGTGGGTGGGTGGGTGGAATGTAGAGATGCGGTGGTTGGGTGGGTGGGTGGGT (SEQ ID
NO: 17)
P3:GGGGTGGGAGGAGGGTTGAATGTAGAGATGCGGTGGGGGTGGGAGGAGGGT (SEQ ID NO:
18)
P4:TCCACGCACAGTTGAATGTAGAGATGCGGTGGTCCACGCACAGT (SEQ ID NO: 19)
P5:TTTCCACGCACAGTTGAATGTAGAGATGCGGTGGTTTCCACGCACAGT (SEQ ID NO: 20)
P6:GCTTTCCACGCACAGTTGAATGTAGAGATGCGGTGGGCTTTCCACGCACAGT (SEQ ID NO:
21)
P7:ACGCTTTCCACGCACAGTTGAATGTAGAGATGCGGTGGACGCTTTCCACGCACAGT (SEQ ID
NO: 22)
P8:CTACGCTTTCCACGCACAGTTGAATGTAGAGATGCGGTGCTACGCTTTCCACGCACAGT (SEQ
ID NO: 23)
P9:TTTCCACGTGAATGTAGAGATGCGGTGGTTTCCACG (SEQ ID NO: 24)
P10:CTACGCTGAATGTAGAGATGCGGTGGCTACGC (SEQ ID NO: 25)
P11:CTACGCTTTGAATGTAGAGATGCGGTGGCTACGCTT (SEQ ID NO: 26)
P12:CTACGCTTTCTGAATGTAGAGATGCGGTGGCTACGTTTC (SEQ ID NO: 27)
P13:CTACGCTTTCCATGAATGTAGAGATGCGGTGGCTACGCTTTCCA (SEQ ID NO: 28)
P14:CTACGCTTTCCACGTGAATGTAGAGATGCGGTGGCTACGCTTTCCACG (SEQ ID NO: 29)
P15:CTACGCTTTCCACGCATGAATGTAGAGATGCGGTGGCTACGCTTTCCACGCA (SEQ ID NO:
30)
P16:CTACGCTTTCCACGCACATGAATGTAGAGATGCGGTGGCTACGCTTTCCACGCACA (SEQ ID
NO: 31)
P17:CTTTCCACGCTGAATGTAGAGATGCGGTGGCTTTCCACGC (SEQ ID NO: 32)
P18:GCTTTCCACGCATGAATGTAGAGATGCGGTGGGCTTTCCACGCA (SEQ ID NO: 33)
P19:CGCTTTCCACGCACTGAATGTAGAGATGCGGTGGCGCTTTCCACGCAC (SEQ ID NO: 34)

P20:ACGCTTTCCACGCACATGAATGTAGAGATGCGGTGGACGCTTTCCACGCACA (SEQ ID NO: 35)

其中*代表Ps修饰。

[0070] 实施例3:评估在5'端或者3'端含有G1的核酸药物在50%血清中的稳定性

为评估在5'端或者3'端含有G1的核酸在50%血清中稳定性,发明人选择对不同长度的核酸(S2、S3和S4)的5'端、3'端及两端均添加G1,其中S2的长度是25 nt,S3的长度是8 nt,S4的长度是13 nt。将各序列分别在50%血清中孵育24h,然后通过16%聚丙烯酰胺凝胶分析结果。结果如图3所示,其中当仅在5'端含有G1时,稳定效果不理想,而仅在3端含有G1时,具有一定的稳定性。另外,核酸长度越短,仅3'端含有G1的稳定性越好。当G1在两端(5'+3')时,稳定效果最好。以上结果说明5'端或者3'端含有G1的核酸均具有稳定性,两端(5'+3')均含有G1时,稳定性最强。

[0071] 另外,对将G1包含在核酸序列内部时核酸序列的稳定性进行了评估。设计了两类不同的序列(G1-S和G1-D),将上述序列分别在50%血清中孵育不同时间,然后通过3%琼脂糖凝胶电泳分析结果。结果如图3显示,在6 h依然存在没被降解的完整的DNA序列,说明当G1在核酸序列内部时,依然具有一定的稳定性。

[0072] 在实施例3中,使用的序列的序列信息如下:

S2:TTTGCCTTTAGGATTCTAGACATTT (SEQ ID NO: 36)

S3:AGACATTT (SEQ ID NO: 37)

S4:ATTCTAGACATTT (SEQ ID NO: 38)

G1-S: TTTGCCTTTAGGAGGGTTGGGTTGGGTTGGGTTCTAGACATTT (SEQ ID NO: 39)

G1-D: TTTGCCTTTAGGAGGGTTGGGTTGGGTTGGGTTGCCTTTAGGATTCTAGACATTTGGGTTGGGTTGGGTTCTAGACATTT (SEQ ID NO: 40)。

[0073] 实施例4:评估含有DNA G-四链体的核酸药物的稳定性

发明人以G1 (SEQ ID NO:3)为基础进一步构建了9种不同的G-四链体序列,将其分别添加在核酸S1的两端,并在血清中孵育24 h;另外,选取了60条不同的G-四链体序列,将其添加在S1的3'端,并在血清中孵育3 h,以不添加G-四链体的随机序列作为阴性对照(DRs),然后分别通过3%琼脂糖凝胶电泳分析结果。结果如图4所示,其中具有不同序列的核酸药物在血清中降解程度存在差异,但均存在没被降解的完整的DNA序列,表明含有DNA G-四链体的核酸药物在血清中均具备一定的稳定性。

[0074] 在实施例4中,使用的序列的序列信息如下:

G2: GGGTTTGGGTGGGTTTGGG (SEQ ID NO: 41)

G3: GGGTGGGTGGGTGGGT (SEQ ID NO: 42)

G4: GGTGGTGGTGGTTGTGGTGGTGGTGG (SEQ ID NO: 43)

G5: GTGGGGCATTGTGGGTGGGTGTGG (SEQ ID NO: 44)

G6: AGGGTTAGGGTTAGGGTTAGGG (SEQ ID NO: 45)

G7: GGTTGGTGTGGTTGG (SEQ ID NO: 46)

G8: GGGTAGGGCGGGTTGGGG (SEQ ID NO: 47)

G9: GGTGGTGGTGGTTGTGGTGGTGGTGGG (SEQ ID NO: 48)

G10: GGTGGTGGTGGTTGTGGTGGTGGTGGGGTGGTGGTGGTTGTGGTGGTGGTGG (SEQ ID NO:

49)

- 1:GGTTAGGTTAGGTTAGG (SEQ ID NO: 50)
- 2:GGGGTTTTGGGGT TTTGGGGTTTTGGGG (SEQ ID NO: 51)
- 3:GGCTTAGGCTTAGGCTTAGG (SEQ ID NO: 52)
- 4:GGGTTTAGGGTTTAGGGTTTAGGG (SEQ ID NO: 53)
- 5:AGGGTTAGGGTTAGGGTTAGGGT (SEQ ID NO: 54)
- 6:GGGCTAGGGCTAGGGCTAGGG (SEQ ID NO: 55)
- 7:TGAGGGTGGGTAGGGTGGGTAA (SEQ ID NO: 56)
- 8:GGGAGGGCGCTGGGAGGAGGG (SEQ ID NO: 57)
- 9:GGGCGGGCGCGAGGGAGGGG (SEQ ID NO: 58)
- 10:ACAGGGGTGTGGGGACAGGGGTGTGGGG (SEQ ID NO: 59)
- 11:GCGGGCGGCTCGGGCGCGGG (SEQ ID NO: 60)
- 12:GGGAGGGTTGGGGTGGG (SEQ ID NO: 61)
- 13:GGGCAGGGAGGGAAGTGGG (SEQ ID NO: 62)
- 14:GGGAGGGTTGGGGTGGGT (SEQ ID NO: 63)
- 15:GGGCAGGGAGGGAAGTGGGT (SEQ ID NO: 64)
- 16:GGGCGGTGTGGGAAGAGGGAAGAGGGG (SEQ ID NO: 65)
- 17:GGTTGGTGTGGTTGGT (SEQ ID NO: 66)
- 18:GGGCTTTTGGGC (SEQ ID NO: 67)
- 19:TTTAAGGGTTAGGGTTAGGGTTAGGG (SEQ ID NO: 68)
- 20:TGAGGGTGGTGGGGTGGGGAAGG (SEQ ID NO: 69)
- 21:TGAGGGTGGGTAGGGTGGGTAAT (SEQ ID NO: 70)
- 22:TGAGGGTGGGAGGGTGGGGAAGGA (SEQ ID NO: 71)
- 23:GGGCGGGGAGGAATTGGGCGGG (SEQ ID NO: 72)
- 24:AGGGAGGGCGCTGGGAGGAGGG (SEQ ID NO: 73)
- 25:CGGGCGGGCGCGAGGGAGGGT (SEQ ID NO: 74)
- 26:CGGGCGGGCGCTAGGGAGGGT (SEQ ID NO: 75)
- 27:GGAGGAGGAGGA (SEQ ID NO: 76)
- 28:GGTTGGTGTGGTTGGTT (SEQ ID NO: 77)
- 29:GGTGGTGGTGGTTGTGGTGGTGGTGGT (SEQ ID NO: 78)
- 30:GGGCGGGCGGGCGGGC (SEQ ID NO: 79)
- 31:TGTGGGGTGGACGGGCCGGGTAGA (SEQ ID NO: 80)
- 32:TTGGGTT (SEQ ID NO: 81)
- 33:TGGGAG (SEQ ID NO: 82)
- 34:GGGGTGGGAGGAGGGT (SEQ ID NO: 83)
- 35:GGTTGGTGTGGTTGGTTT (SEQ ID NO: 84)
- 36:ATGGGGTCGGGCGGGCCGGGTGTC (SEQ ID NO: 85)
- 37:GTGGGTAGGGCGGGTTGG (SEQ ID NO: 86)
- 38:GGGTTAGGGTTAGGGTAGGG (SEQ ID NO: 87)

39:CCTGGGGGAGTATTGCGGAGGAAGG (SEQ ID NO: 88)
 40:AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGG (SEQ ID NO: 89)
 41:AGGGCGGTGTGGGAAGAGGGAAGAGGGGGAGGCAG (SEQ ID NO: 90)
 42:CGGTCGCTCCGTGTGGCTTGGGTTGGGTGTGCCAGTGAC (SEQ ID NO: 91)
 43:GGGGGTGGGAGGGTAGGCCTTAGGTTTCTGA (SEQ ID NO: 92)
 44:CGCCTGATTAGCGATACTCAGCGTTGGGGGGGGGGG (SEQ ID NO: 93)
 45:GGGCGTGGTGGGTGGGGTACTAATAATGTGCGTTTG (SEQ ID NO: 94)
 46:AGCGGGCATATGGTGGTGGGTGGTATGGTC (SEQ ID NO: 95)
 47:AACACATAGGTTTGGTTAGGTTGGTTGGTTGAATTA (SEQ ID NO: 96)
 48:TTGCGCGTTAATTGGGGGGGTGGGTGGGTT (SEQ ID NO: 97)
 49:GGTGGTGGGGGGGTTGGTAGGGTGTCTTC (SEQ ID NO: 98)
 50:GATCGGGTGTGGGTGGCGTAAAGGGAGCATCGGACA (SEQ ID NO: 99)
 51:GGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGG (SEQ ID NO: 100)
 52:GGGGCGCTTATGGGGAGGGTGGGGAGGGTGGGGAAGGTGGGGAGGAG (SEQ ID NO: 101)
 53:CCCCCGGGGCGGGCCGGGGCGGGGTCCCGGCGGGGCGGAGCCATG (SEQ ID NO: 102)
 54:GGCGGGGGGGGGGGGGCGGGGGCGGGGGCGGGGAGGGGCGGGCGG (SEQ ID NO: 103)
 55:CGCGGAGGGGCGGGCGGGAGGAAGGGGGCGGGAGCGGGGCTGTGG (SEQ ID NO: 104)
 56:GAGGAGCGGGTAGGGGCGGGCGGGGCGGGGCGGTCCAGGGGTGGG (SEQ ID NO: 105)
 57:CGCTCGGGCGCGGGGAGGGGAGAGGGGGCGGGAGCGCGCCCGCTC (SEQ ID NO: 106)
 58:AGAGGAGGAGGAGGTACGGAGGAGGAGAGAAGGAGGAGGAGGAAA (SEQ ID NO: 107)
 59:GGGGCGCTTATGGGGAGGGTGGGTAGGGTGGTAAGGTGGGGAGGAG (SEQ ID NO: 108)
 60:GGGGCGCTTATGGGGAGGGTGGGTAGGGTGGGAAGGTGGGGAGGAG (SEQ ID NO: 109)
 DRs:GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGCAAGCTGCCCGTGCCTGGCCC (SEQ ID NO: 166)。

[0075] 实施例5:评估含有基于G1改造的富含G的序列的核酸药物的稳定性

发明人选择对G1序列进行不同改造,将改造后的序列分别添加在核酸S1的两端,评估含有基于G1改造的富含G的序列的核酸药物的稳定性。将上述序列分别在50%血清中孵育24 h,然后通过3%琼脂糖凝胶电泳分析结果。结果如图5所示,其中含有基于G1改造的富含G的序列的核酸药物均具备稳定性。

[0076] 在实施例5中,使用的序列的序列信息如下:

G1-1: GGAAGGGAGGGAAGGG (SEQ ID NO: 110)
 G1-2: GGGTTGGGTGGGTGGG (SEQ ID NO: 111)
 G1-3: GGGCCGGGCGGGCCGGG (SEQ ID NO: 112)
 G1-4: GGGTTGGGTGGGTGGG (SEQ ID NO: 113)
 G1-5: GGGTAGGGCGGGTTGGG (SEQ ID NO: 114)
 G1-6: GGGTGGGTGGGTGGG (SEQ ID NO: 115)
 G1-7: GGGTGGGTGGGTGGG (SEQ ID NO: 116)
 G1-8: GGGTGGGTTGGGTGGG (SEQ ID NO: 117)
 G1-9: GGGTTGGGTTGGGTTGGG (SEQ ID NO: 118)

G1-10: GGGTTTTGGGTTTTGGGTTTTGGG (SEQ ID NO: 119)
 G1-11: GGGTTTTGGGTTTTGGGTTTTGGG (SEQ ID NO: 120)
 G1-12: GGGGTTGGGGTTGGGGTTGGGG (SEQ ID NO: 121)
 G1-13: GGGGGTTGGGGTTGGGGTTGGGG (SEQ ID NO: 122)
 G1-14: GGGGGGTTGGGGGTTGGGGGTTGGGGG (SEQ ID NO: 123)
 G1-15: GGGGGGGTTGGGGGGTTGGGGGGTTGGGGGG (SEQ ID NO: 124)
 G1-16: GGGGGGGGTTGGGGGGGTTGGGGGGGTTGGGGGGG (SEQ ID NO: 125)
 G1-17: GGGGGGGGGTTGGGGGGGGTTGGGGGGGGTTGGGGGGGG (SEQ ID NO: 126)
 G1-18: GGGGGGGGGGTTGGGGGGGGGTTGGGGGGGGGTTGGGGGGGGG (SEQ ID NO:

127)

G1-19: GGGTAGGGCGGGTTGGGTAGGG (SEQ ID NO: 128)
 G1-20: GGGTAGGGCGGGTTGGGTAGGGCGGG (SEQ ID NO: 129)
 G1-21: GGGTAGGGCGGGTTGGGTAGGGCGGGTTGGG (SEQ ID NO: 130)
 G1-22: GGGTAGGGCGGGTTGGGGGTAGGGCGGGTTGGG (SEQ ID NO: 131)。

[0077] 实施例6:评估含有DNA G-四链体的不同长度的核酸药物在50%血清中的稳定性
 评估了含有DNA G-四链体(此处采用G-四链体(G1~G10以及G1-1~G1-22))的不同长度的核酸药物在50%血清中的稳定性。选择不同长度的核酸,将上述DNA G-四链体分别添加在其两端。所选长度分别为16-20nt、25nt、30nt、35nt、50nt、80nt和200nt,其中长度为35nt和50nt的核酸分别为CD 44和c-Met的核酸适配体。将上述序列在50%血清中孵育24 h,以不添加DNA G-四链体的随机序列作为阴性对照(Rs/Control),然后通过3%琼脂糖凝胶电泳分析结果。结果如图6所示,其中不同长度的核酸在血清中的降解程度存在差异,但均存在没被降解的完整的DNA序列,结合实施例3中对8nt核酸的保护能力,说明含有DNA G-四链体的不同长度核酸药物均具有稳定性。

[0078] 在实施例6中,使用的序列的序列信息如下:

16nt: AATGTAGAGATGCGGT (SEQ ID NO: 132)
 17nt: AATGTAGAGATGCGGTG (SEQ ID NO: 133)
 18nt: GAATGTAGAGATGCGGTG (SEQ ID NO: 134)
 19nt: GAATGTAGAGATGCGGTGG (SEQ ID NO: 135)
 20nt: TGAATGTAGAGATGCGGTGG (SEQ ID NO: 136)
 25nt: TTTGAATGTAGAGATGCGGTGTTT (SEQ ID NO: 137)
 30nt: TTTAGATTATATGTCATACCTCCATTGGTT (SEQ ID NO: 138)
 33nt: TTTGGACGGTGTTAAACGAAAGGGGACGACCTT (SEQ ID NO: 139)
 35nt: GGATAGGGATTCTGTTGGTCGGCTGGTTGGTATCC (SEQ ID NO: 140)
 50nt: ATCAGGCTGGATGGTAGCTCGGTTCGGGTGGGTGGGTTGGCAAGTCTGAT (SEQ ID NO:

141)

80nt: GCTGTGTGACTCCTGCAAAGTGTGGACAACCTCCACGGAGGAATCCCGTATCTAAAGGTG
 CAGCTGTATCTTGTCTCC (SEQ ID NO: 142)

200nt:

AACACCGGTTTGCCATGTGTTGCCATGTGTATGTGGGAGACGGTCGGGTCCAGATATTCGTATCTGT

CGAGTAGAGTGTGGGCTCCCCACATACTCTGATGATCCGAGACGGTCGGGTCCAGATATTCGTATCTGTCGAGTAG
AGTGTGGGCTCGGATCATTGCAACTCTGATGATCCTCTATCTATCTGTCGAG (SEQ ID NO: 143)。

[0079] 实施例7:评估含有DNA G-四链体的不同序列的核酸药物在50%血清中的稳定性

选择不同序列的核酸S5-S13,评估了含有DNA G-四链体(此处采用G-四链体(G1~G10以及G1-1~G1-22))的不同序列的核酸药物在50%血清中的稳定性。将上述序列在50%血清中孵育不同时间,以不使用该方法进行保护的随机序列作为阴性对照(RS),然后通过3%琼脂糖凝胶电泳分析结果。结果如图7所示,其中含有DNA G-四链体的不同序列的核酸药物在50%血清中降解程度存在差异,但均存在没被降解的完整的DNA序列,结合实施例1中对S1的保护能力,说明含有DNA G-四链体的不同序列的核酸药物在50%血清中均具有稳定性。

[0080] 在实施例7中,使用的序列的序列信息如下:

RS: GGATAGGGATTCTGTTGGTCCGGCTGGTTGGTATCC (SEQ ID NO: 140)

S5: TGCCTTTAGGATTCTAGACA (SEQ ID NO: 144)

S6: ATGGAGGTATGACATATAAT (SEQ ID NO: 145)

S7: TACATTGCCTCTTCATT (SEQ ID NO: 146)

S8: TTCCGAATAAACTCCAGGC (SEQ ID NO: 147)

S9: AAGCAAAACAGGTCTAGAA (SEQ ID NO: 148)

S10: GTGTCTAGGAGATACACCT (SEQ ID NO: 149)

S11: TTCCGAATAAACTCCAGGCT (SEQ ID NO: 150)

S12: AAGCAAAACAGGTCTAGAAT (SEQ ID NO: 151)

S13: GTGTCTAGGAGATACACCTT (SEQ ID NO: 152)。

[0081] 实施例8:评估含有RNA G-四链体的不同序列的RNA药物在50%血清中的稳定性

评估了含有RNA G-四链体的不同序列的RNA药物在50%血清中的稳定性。将G-四链体(经转换为RNA的G1~G10以及G1-1~G1-22)添加在RNA药物的两端,直接体外转录为完整的RNA(RNA-1和RNA-2),然后在50%血清中分别孵育3 h、6 h和12 h,其中采用不含有G-四链体的RNA作为对照(Rs-RNA);然后通过3%琼脂糖凝胶电泳分析结果。结果如图8所示,Rs-RNA在血清中孵育3h后,基本全部被降解,而RNA-1和RNA-2在血清中孵育12 h后,依然存在大量的完整的RNA序列,说明含有RNA G-四链体的不同序列的RNA药物在50%血清中均具有稳定性。

[0082] 在实施例8中,使用的序列的序列信息如下:

RS-RNA:

UCUAUCUAUCUGUCGAGUAGAGUGUGGGCUCUUUUGGAAGAAACUGUGGCACUUCGGUGCCAG (SEQ ID NO: 153)

RNA1:

GGGUAGGGCGGUUGGGUUUGCCUUUAGGAUUCUAGACAUUUGGGUAGGGCGGUUGGG (SEQ ID NO: 154)

RNA2: GGGUUGGGUUUGGGUUUGCCUUUAGGAUUCUAGACAUUUGGGUUGGGUUUGGGUUGGG

(SEQ ID NO: 155)。

[0083] 实施例9:评估含有RNA G-四链体序列的长链RNA药物在血清中的稳定性

评估了含有RNA G-四链体(经转换为RNA的G1~G10以及G1-1~G1-22)序列的长链

RNA (包括mRNA和lncRNA)在血清中的稳定性,其中选择长度为5000nt和2000nt的RNA作为靶标,含有RNA G-四链体的长链RNA命名为L-RNA,将其在10%血清中分别孵育10 min和30 min,其中采用不含有RNA G-四链体的长链RNA作为阴性对照(NC);然后通过3%琼脂糖凝胶电泳分析结果。结果如图8所示,NC在血清中孵育10 min后,基本全部被降解,而L-RNA在血清中孵育10 min后,依然存在大量的完整的RNA序列,说明含有RNA G-四链体长链RNA药物在血清中具有稳定性。

[0084] 在实施例9中,使用的序列的序列信息如下:

L-RNA:

```
GGGUUGGGUUUGGGUUGGGUUAUUUUUAAGAGCUGUGGAGUUCUAAAAUAUCAACCAUGGCACUUUCU
CCUGACCCCUUCCCUAGGGGAUUUCAGGAUUGAGAAUUUUUCCAUCGAGCCUUUUAAAAUUGUAGGACUUGUUC
CUGUGGGCUUCAGUGAUGGGAUAGUACACUUCACUCAGAGGCAUUUGCAUCUUAAAAUAAUUUCUAAAAGCCUCU
AAAGUGAUCAGUGCCUUGAUGCCAACUAAGGAAUUUGUUUAGCAUUGAAUCUCUGAAGGCUCUAUGAAAGGAAUA
GCAUGAUGUGCUGUUAGAAUCAGAUGUUACUGCUAAAAUUACAUGUUGUGAUGUAAAUUGUGUAGAAAACCAUUA
AAUCAUUCAAAAUAAUAAACUAAAAAAAAUUAGAGAAUGUAUACUUUUAGAAAAGCUGUCUCCUUAUUUAAAAU
AGUGUUUGUCUGUAGUUCAGUGUUGGGGCAAUUCUUGGGGGGAUUCUUCUAAUCUUUCAGAAACUUUGUCUGCG
AACACUCUUUAAUGGACCAGAUCAGGAUUUGAGCGGAAGAACGAAUGUAACUUUAAGGCAGGAAAGACAAAUUUA
UUCUUCAUAAAGUGAUGAGCAUUAUAAUUAUCCAGGCACAUGGCAAUAGAGGCCUCUAAAAUAGGAAUAAUAAAC
CUCUUAGACAGGUGGGAGAUUAUGAUCAGAGUAAAAGGUAAUACACAUUUUUUUCCAGAAAGUCAGGGGUCUAU
AAAUUGACAGUGAUUAGAGUAAUACUUUUUCACAUUUCCAAAGUUUGCAUGUUAACUUUAAAUGCUUACAAUCUUA
GAGUGGUAGGCAAUGUUUACACUAAUUGACCUUAAUAGGGAAGGGAGGGGGUGCCUGUGGGGUUUUAAAGAAUUU
UCCUUUGCAGAGGCAUUUCAUCCUUAUGAAGCCAUUCAGGAUUUUGAAUUGCAUUGAGUGCUUGGCUCUCCUU
CUGUUCUAGUGAGUGUAUGAGACCUUGCAGUGAGUUUAUCAGCAUACUAAAAUUUUUUUCCUGGAAUUUGGAGGG
AUGGGAGGAGGGGGUGGGGCUUACUUGUUGUAGCUUUUUUUUUUUUACAGACUUCACAGAGAAUGCAGUUGUCUU
GACUUCAGGUCUGUCUGUUCUGUUGGCAAGUAAAUGCAGUACUGUUCUGAUCCCGCUGCUAUUAGAAUGCAUUGUG
AAACGACUGGAGUAUGAUUAAAAGUUGUGUUCCCCAUUGCUUGGAGUAGUGAUUGUUAAGGAAAAAUCCAGCUG
AGUGAUAAAAGGCUGAGUGUUGAGGAAUUUCUGCAGUUUUAAGCAGUCGUUUUGUGAUUGAAGCUGAGUACAUUU
UGCUGGUGUAUUUUUAGGUAAAUGCUUUUUGUUCAUUUUCUGGUGGUGGGAGGGGACUGAAGCCUUUAGUCUUUC
CAGAUACAACUUAAAAUCAGUGACAAGAAACAUUCCAAACAAGCAACAGUCUUAAGAAUUAAACUGGCAAGUG
GAAAUGUUUAAACAGUUCAGUGAUCUUUAGUGCAUUGUUUUAUGUGUGGGUUUCUCUCUCCCCUCCCUUGGUCUUA
UUCUUACAUGCAGGAACACUCAGCAGACACACGUAUGCGAAGGGCCAGAGAAGCCAGACCCAGUAAGAAAAAUAG
CCUAAUUACUUUAAAAUAAACCAACAUUCCAUUUUAAAUGUGGGGAUUUGGGAACCACUAGUUCUUUCAGAUGGUAU
UCUUCAGACUUAAGAAGGAGCUUCCAGUUGAAUUCACCAGUGGACAAAAUAGAGGAAAACAGGUGAACAAAGCUUUU
CUGUAUUUACAUACAAAGUCAGAUCAUUGGACAAUAGUAUUGAAUAGAUUUCAGCUUUUAGCUGGAGUAACU
GGCAUGUGAGCAAACUGUGUUGGCGUGGGGGUGGAGGGGUGAGGUGGGCGCUAAGCCUUUUUUUAAAGAUUUUCAG
GUACCCUCACUAAAGGCACCGAAGGCUUAAAGUAGGACAACCAUGGAGCCGGGUUGGGUUUGGGUUGGG (SEQ
ID NO: 156)
```

NC:

```
UUUUUUUAAGAGCUGUGGAGUUCUAAAAUAUCAACCAUGGCACUUUCUCCUGACCCCUUCCCUAGGG
GAUUUCAGGAUUGAGAAUUUUUCCAUCGAGCCUUUUAAAAUUGUAGGACUUGUCCUGUGGGCUUCAGUGAUGG
```

GAUAGUACACUUCACUCAGAGGCAUUUGCAUCUUUAAAUAUUUCUUAAAAGCCUCUAAAAGUGAUCAGUGCCUUGA
 UGCCAACUAAGGAAAUUUGUUUAGCAUUGAAUCUCUGAAGGCUCUAUGAAAAGGAAUAGCAUGAUGUCUGUUAGAA
 UCAGAUGUUACUGCUAAAAUUUACAUGUUGUGAUGUAAAUUGUGUAGAAAACCAUAAAUCAUUCAAAAUAUAAA
 CUAUUUUUAUUAGAGAAUGUAUACUUUUAGAAAGCUGUCUCCUUAUUUAAAUAUUUUAGUGUUUGUCUGUAGUUA
 GUGUUGGGGCAAUCUUGGGGGGGAUUCUUCUCUAAUCUUUCAGAAAACUUUGUCUGCGAACACUCUUUAAUGGACCA
 GAUCAGGAUUUGAGCGGAAGAACGAAUGUAACUUUAAGGCAGGAAAGACAAAUUUUUAUUCUUCAUAAAAGUGAUGAG
 CAUAAAUAUUUCCAGGCACAUGGCAAUAGAGGCCUCUAAAUAAGGAAUAAAUAACCUCUUAAGACAGGUGGGAGA
 UUAUGAUCAGAGUAAAAGGUAUUACACAUUUUAUUUCCAGAAAGUCAGGGGUCUAUAAAUUGACAGUGAUUAGAG
 UAAUACUUUUUACAUUUCCAAGUUUGCAUGUUAACUUUAAAUGCUUACAAUCUUAAGAGUGGUAGGCAAUGUUUU
 ACACUAUUGACCUUAUUAUAGGGAAGGGAGGGGGUGCCUGUGGGGUUUUAAAAGAAUUUCCUUGCAGAGGCAUUUC
 AUCCUUAUGAAGCAUUCAGGAUUUUGAAUUGCAUUAUGAGUGCUUGGCUCUCCUUCUGUUCUAGUGAGUGUAUG
 AGACCUUGCAGUGAGUUUAUCAGCAUACUCAAAAUUUUUCCUGGAAUUUGGAGGGAUGGGAGGAGGGGGUGGGG
 CUUACUUGUUGAGCUUUUUUUUUUUUACAGACUUCACAGAGAAUGCAGUUGUCUUGACUUCAGGUCUGUCUGUU
 CUGUUGGCAAGUAAAUGCAGUACUGUUCUGAUCCCGCUGCUAUUAGAAUGCAUUGUGAAACGACUGGAGUAUGAUU
 AAAAGUUGUGUCCCCAAUGCUUGGAGUAGUGAUUGUUAAGGAAAAAAUCCAGCUGAGUGAAUAAAGGCUGAGUGU
 UGAGGAAAUUCUGCAGUUUAAGCAGUCGUUUUUGUGAUUGAAGCUGAGUACAUUUUGCUGGUGUAUUUUUAGGU
 AAAAUGCUUUUUGUCAUUUCUGGUGGGGAGGGGACUGAAGCCUUUAGUCUUUCCAGAUGCAACCUUAAAAUC
 AGUGACAAGAAACAUUCCAACAAGCAACAGUCUUAAGAAUUAACUGGCAAGUGGAAUUGUUAAACAGUUCA
 GUGAUCUUUAGUGCAUUGUUUAUGUGUGGGUUUCUCUCUCCCCUCCUUGGUCUAAAUCUUAUCAUGCAGGAACAC
 UCAGCAGACACACGUAUGCGAAGGGCCAGAGAAGCCAGACCAGUAAGAAAAAAUAGCCUAAUUACUUUAAAUAAA
 CCAAACAUUCCAUUUUAAAUGUGGGGAUUGGGAACCACUAGUUCUUUCAGAUGGUAUUUCUUCAGACUAUAGAAGGA
 GCUUCCAGUUGAAUUCACCAGUGGACAAAUGAGGAAAACAGGUGAACAAGCUUUUUCUGUAUUUACAUACAAAGU
 CAGAUCAGUUAUGGGACAAUAGUAUUGAAUAGAUUUCAGCUUUAUGCUGGAGUAACUGGCAUGUGAGCAAACUGUG
 UUGGCGUGGGGGUGGAGGGGUGAGGUGGGCGCUAAGCCUUUUUUAAGAUUUUUCAGGUACCCUCACUAAAGGCA
 CCGAAGGCUUAAAGUAGGACAACCAUGGAGCC (SEQ ID NO: 157)。

[0085] 实施例10:评估含有RNA G-四链体的核酸药物嵌合体以及含有DNA G-四链体的核酸药物嵌合体在血清中的稳定性

评估了含有DNA G-四链体(G1~G10以及G1-1~G1-22)的核酸药物嵌合体的稳定性:将各DNA G-四链体序列分别添加在RS1 (RNA) 的两端,分别构成G1-RS1至G1-22-RS1,从而形成DNA和RNA的嵌合体,将其在血清中孵育15min,其中采用不含有DNA G-四链体的核酸嵌合体作为对照(Rs-D+R);然后通过3%琼脂糖凝胶电泳分析结果。评估了含有RNA G-四链体(以G1为代表,RG1)的核酸药物嵌合体的稳定性:将RG1序列分别添加在S1的两端,构成RG1-S1,从而形成RNA和DNA的嵌合体,将其在血清中孵育30min、2h和4h,其中采用不含有RNA G-四链体的嵌合体作为对照(Rs-R+D);然后通过3%琼脂糖凝胶电泳分析结果。结果如图10所示,其中含有G-四链体的核酸药物嵌合体在血清中均具有一定的稳定性。

[0086] 在实施例10中,使用的序列的序列信息如下:

G1-RS1:GGGTTGGGTTTGGGTTGGGUGAAUGUAGAGAUGC GGUGGGGGTTGGGTTTGGGTTGGG
 (SEQ ID NO: 158)

Rs-D+R:

TCTATCTATCTGTGTCGAGTAUGAAUGUAGAGAUGCGGUGAAACTGTGGCACTTCGGTGC (SEQ ID NO: 159)

RG1-S1:

GGGUUGGGUUUGGGUUGGGTGAATGTAGAGATGCGGTGGGGUUUGGGUUUGGG (SEQ ID NO: 160)

RS-R+D:

UCUAUCUAUCUGUCGAGUATGAATGTAGAGATGCGGTAGAAACUGUGGCACUUCGGUGA (SEQ ID NO: 161)。

[0087] 实施例11:评估含有G-四链体的寡核苷酸药物在细胞内对靶标RNA的敲低效果

1.采用Transeasy转染G-四链体(G1~G10以及G1-1~G1-22)保护的ASO (Su1、Su2、MA1、MA2、PC1和PC2):将适量的ASO加入50 μ L的Opti-MEM进行稀释,同时将4.8 μ L的Transeasy用50 μ L的Opti-MEM进行稀释,轻轻吹打混匀后室温放置5 min。然后将DNA稀释液加入到Transeasy稀释液当中形成混合液,室温放置孵育20min。同时将12孔板的细胞换成无双抗的培养基,将混合液滴加到细胞中,轻柔混匀,5~6 h后更换成含血清和双抗的完全培养基。

[0088] 2.细胞总RNA的提取:由于RNA易降解,且空气和实验容器中存在大量外源性RNA酶,因此在操中应使用无RNA酶的枪头和管子。实验过程中所用的玻璃器皿需要在160 $^{\circ}$ C高温下烘烤4h以上;塑料器皿需经0.5 M的NaOH溶液处理10 min,再用无RNA酶的水清洗干净后高压灭菌。

[0089] (1)待贴壁细胞密度约80%的时候可进行收取,弃掉培养基,然后用预冷的1 \times PBS润洗细胞一次。

[0090] (2)向6孔板的细胞中加入1mL的Trizol,12孔板的细胞用量减半。用移液器吹打混匀。

[0091] (3)每1mL的Trizol中加入200 μ L氯仿,盖上管盖,用力摇晃使之充分混匀,并在室温下静止2min,然后4 $^{\circ}$ C,12000g离心10min。

[0092] (4)吸取大约400 μ L的水相层至一新的EP管中,加入500 μ L的异丙醇混匀后,在-80 $^{\circ}$ C低温冰箱中放置半小时以上,然后4 $^{\circ}$ C,12000g离心15min。

[0093] (5)离心后可看见管底有白色沉淀,此为RNA。倒掉上清,加入1mL 75%的乙醇并重悬沉淀,然后4 $^{\circ}$ C,8500g离心10min。

[0094] (6)吸去乙醇,小心不要吸到沉淀。管子在室温下静置,待乙醇挥发完后,加入50 μ L无RNA酶H₂O,并于42 $^{\circ}$ C孵育2 min,使得RNA充分溶解。然后取出少量RNA 测其浓度,并用TBE电泳检测RNA的完整性。

[0095] 3.实时荧光定量PCR:

(1)从-20 $^{\circ}$ C冰箱中取出cDNA,室温溶解,然后加入180 μ L的ddH₂O稀释10倍。

[0096] (2)qPCR反应体系如下:

2×qPCR SYRB Green Mix	20 μL
上游引物(10 μM)	0.5 μL
下游引物(10 μM)	0.5 μL
cDNA 模板	4 μL
ddH ₂ O	5 μL

PCR反应条件为:95℃ 2min;95℃ 5s,58℃ 5s,72℃ 15s (此程序为40个循环);95℃ 15s;60℃ 60s;溶解曲线60℃~95℃,每上升0.5℃吸收一次荧光信号;95℃ 10s。

[0097] (3) 利用ABI step one plus定量仪进行PCR,根据结果运用 $\Delta\Delta C_t$ 算法分析实验数据如图11。

[0098] 其中所述的ASO针对的靶标RNA分别是*Survivin*、*Malat1*和*PCSK9*等三个基因,针对每个基因分别设计了两条相应的ASO,命名分别为Su1、Su2、MA1、MA2、PC1和PC2,分别在细胞水平测试其敲低能力,其中Su1、Su2、MA1和MA2是在两种细胞内进行测试,包括A549细胞和HeLa细胞;PC1和PC2是在97H细胞进行测试。同时设置为未转染组(Mock)作为阴性对照。结果显示,含有G-四链体的寡核苷酸药物在细胞内具有良好的敲低能力。

[0099] 在实施例11中,使用的序列的序列信息如下:

Su1: GAATGTAGAGATGCGGTGG (SEQ ID NO: 135)
 Su2: TGAGGGCGAATCAAATCCATC (SEQ ID NO: 162)
 MA1: TGCCTTTAGGATTCTAGACA (SEQ ID NO: 144)
 MA2: ATGGAGGTATGACATATAAT (SEQ ID NO: 145)
 MA3: ACATTGCCTCTTCATT (SEQ ID NO: 163)
 PC1: GTCTGTGGAAGCG (SEQ ID NO: 164)
 PC2: GCCTGTCTGTGGAAGCGGGT (SEQ ID NO: 165)。

[0100] 实施例12:采用含有G-四链体的靶向PCSK9的反义寡核苷酸在动物体内敲低PCSK9 (a) 和降低LDL (b) 的能力

为了评估采用含有G-四链体(G1~G10以及G1-1~G1-22)的反义寡核苷酸在动物体内的有效性,发明人设计了靶向PCSK9的ASO序列,同时设计全硫代修饰的PS-ASO作为对照实验,以对比两种不同保护方法在动物体内的有效性。注射前,将实验所需的“ssDNA”用生理盐水稀释。将小鼠分组称重并编号:PS-ASO (0.8μmol/kg) 和LT-001 (在PCSK9的ASO序列两端添加G1所形成的序列) (0.8μmol/kg) 共2组。注射方式采用小鼠尾静脉高压注射,使用2ml空针针管和4.5号针头进行注射,注射稀释好的ssDNA生理盐水溶液。每只小鼠注射的生理盐水溶液总量为其体重的8%~10% (体积/体重),约2ml总体积,注射过程在5~8秒内完成。分别在注射前后3天每天进行剪尾采血,而后是第5天、第30天进行剪尾采血。

[0101] 采用ELISA方法分别对小鼠血浆中的PCSK9和LDL-c (低密度脂蛋白胆固醇) 含量进行检测,考察采用该方法保护的靶向PCSK9反义寡核苷酸在动物体内敲低PCSK9的效果。结果如图12所示,其中注射LT-001的小鼠与注射PS-ASO相比,PCSK9和LDL-c的值显著下降,并且与PS-ASO相比,下降更多;说明采用含有G-四链体的靶向PCSK9反义寡核苷酸在动物体内

具有良好地敲低PCSK9和降低LDL含量的能力。

[0102] 实施例中所指的靶向PCSK9反义寡核苷酸的序列信息:GTCTGTGGAAGCG (SEQ ID NO: 164);PS-ASO为该序列的全硫代磷酸脂的修饰体。

[0103] 以上所述,仅是本发明的较佳实施例而已,并非对本发明作任何形式上的限制,本领域技术人员利用上述揭示的技术内容做出些许简单修改、等同变化或修饰,均落在本发明的保护范围内。

<110>	成都凌泰氮生物技术有限公司	
<120>	含有G-四链体的核酸	
<130>	CPCH2162766N	
<160>	166	
<170>	PatentIn version 3.3	
<210>	1	
<211>	25	
<212>	DNA	
<213>	人工序列	
<400>	1	
	tttgaatgta gagatgceggt ggttt	25
<210>	2	
<211>	59	
<212>	DNA	
<213>	人工序列	
<400>	2	
	acccgacctc ttctatctgg acccgaccgt ctcttttttg agcccacact ctactcgac	59
<210>	3	
<211>	19	
<212>	DNA	
<213>	人工序列	
<400>	3	
	gggttgggtt tgggttggg	19
<210>	4	
<211>	33	
<212>	DNA	
<213>	人工序列	
<400>	4	
	tgaacacgcc atgtcgattc tttagaatcg aca	33
<210>	5	
<211>	37	
<212>	DNA	
<213>	人工序列	
<400>	5	
	tgttcatctg aacacgccat gtcgattett tagaatc	37
<210>	6	
<211>	35	
<212>	DNA	

<213> 人工序列	
<400> 6	
gcgcttatga acacgccatg tcgattctta agcgc	35
<210> 7	
<211> 23	
<212> DNA	
<213> 人工序列	
<400> 7	
ggggctcctg gagcggggca cac	23
<210> 8	
<211> 31	
<212> DNA	
<213> 人工序列	
<400> 8	
ggggtcgaca cccaattctg aaaatggata a	31
<210> 9	
<211> 26	
<212> DNA	
<213> 人工序列	
<400> 9	
ggggacaccc aattctgaaa atgggg	26
<210> 10	
<211> 27	
<212> DNA	
<213> 人工序列	
<400> 10	
gggaggtccc tgttcgggcg ccagggg	27
<210> 11	
<211> 26	
<212> DNA	
<213> 人工序列	
<400> 11	
ggggacaccc aattctgaaa atgggg	26
<210> 12	
<211> 24	
<212> DNA	
<213> 人工序列	
<400> 12	
gggggaagga ggaggatgag gggg	24

<210>	13	
<211>	27	
<212>	DNA	
<213>	人工序列	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代修饰	
<220>		
<221>	misc_feature	
<222>	(7) .. (7)	
<223>	硫代修饰	
<221>	misc_feature	
<222>	(9) .. (10)	
<223>	硫代修饰	
<221>	misc_feature	
<222>	(13) .. (14)	
<223>	硫代修饰	
<221>	misc_feature	
<222>	(21) .. (21)	
<223>	硫代修饰	
<221>	misc_feature	
<222>	(24) .. (26)	
<223>	硫代修饰	
<400>	13	
	gggaggtccc tgttcgggcg ccagggg	27
<210>	14	
<211>	26	
<212>	DNA	
<213>	人工序列	
<220>		
<221>	misc_feature	
<222>	(1) .. (3)	
<223>	硫代修饰	
<221>	misc_feature	
<222>	(23) .. (25)	
<223>	硫代修饰	
<400>	14	
	ggggacaccc aattctgaaa atgggg	26

<210>	15	
<211>	24	
<212>	DNA	
<213>	人工序列	
<220>		
<221>	misc_feature	
<222>	(1) .. (2)	
<223>	硫代修饰	
<220>		
<221>	misc_feature	
<222>	(17) .. (17)	
<223>	硫代修饰	
<221>	misc_feature	
<222>	(22) .. (23)	
<223>	硫代修饰	
<400>	15	
	gggggaagga ggaggatgag gggg	24
<210>	16	
<211>	70	
<212>	DNA	
<213>	人工序列	
<400>	16	
	ttggggttag ggtagggtt agggatgaat gtagagatgc ggtggttggg gtagggtta	60
	gggtaggga	70
<210>	17	
<211>	56	
<212>	DNA	
<213>	人工序列	
<400>	17	
	ttgggtgggt ggtgggtt aatgtagaga tgcggtggtt ggtgggtgg gtgggt	56
<210>	18	
<211>	52	
<212>	DNA	
<213>	人工序列	
<400>	18	
	ggggtgggag gaggttgaa ttagagatg cggtggggg tggaggagg gt	52
<210>	19	
<211>	44	
<212>	DNA	

<213> 人工序列	
<400> 19	
tccacgcaca gttgaatgta gagatgcggt ggtccacgca cagt	44
<210> 20	
<211> 48	
<212> DNA	
<213> 人工序列	
<400> 20	
tttccacgca cagttgaatg tagagatgcg gtggtttcca cgcacagt	48
<210> 21	
<211> 52	
<212> DNA	
<213> 人工序列	
<400> 21	
gctttccacg cacagttgaa tgtagagatg cgggtggctt tccacgcaca gt	52
<210> 22	
<211> 56	
<212> DNA	
<213> 人工序列	
<400> 22	
acgctttcca cgcacagttg aatgtagaga tgcggtggac gctttccacg cacagt	56
<210> 23	
<211> 59	
<212> DNA	
<213> 人工序列	
<400> 23	
ctacgctttc cacgcacagt tgaatgtaga gatgcggtgc tacgctttcc acgcacagt	59
<210> 24	
<211> 36	
<212> DNA	
<213> 人工序列	
<400> 24	
tttccacgtg aatgtagaga tgcggtggtt tccacg	36
<210> 25	
<211> 32	
<212> DNA	
<213> 人工序列	
<400> 25	
ctacgctgaa tgtagagatg cgggtggctac gc	32

<210>	26	
<211>	36	
<212>	DNA	
<213>	人工序列	
<400>	26	
	ctacgctttg aatgtagaga tgcggtggct acgctt	36
<210>	27	
<211>	40	
<212>	DNA	
<213>	人工序列	
<400>	27	
	ctaccgtttc tgaatgtaga gatgcggtgg ctaccgtttc	40
<210>	28	
<211>	44	
<212>	DNA	
<213>	人工序列	
<400>	28	
	ctacgctttc catgaatgta gagatgcggt ggctacgctt tcca	44
<210>	29	
<211>	48	
<212>	DNA	
<213>	人工序列	
<400>	29	
	ctacgctttc cacgtgaatg tagagatgcg gtggctacgc tttccacg	48
<210>	30	
<211>	52	
<212>	DNA	
<213>	人工序列	
<400>	30	
	ctacgctttc cacgcatgaa thtagatg cggtggctac gctttccacg ca	52
<210>	31	
<211>	56	
<212>	DNA	
<213>	人工序列	
<400>	31	
	ctacgctttc cacgcacatg aatgtagaga tgcggtggct acgctttcca cgcaca	56
<210>	32	
<211>	40	
<212>	DNA	

<213> 人工序列	
<400> 32	
ctttccacgc tgaatgtaga gatgcggtgg ctttccacgc	40
<210> 33	
<211> 44	
<212> DNA	
<213> 人工序列	
<400> 33	
gctttccacg catgaatgta gagatgcggt ggctttcca cgca	44
<210> 34	
<211> 48	
<212> DNA	
<213> 人工序列	
<400> 34	
cgctttccac gcaactgaatg tagagatgcg gtggcgcttt ccacgcac	48
<210> 35	
<211> 52	
<212> DNA	
<213> 人工序列	
<400> 35	
acgctttcca cgcacatgaa thtagatg cggtggacgc tttccacgca ca	52
<210> 36	
<211> 25	
<212> DNA	
<213> 人工序列	
<400> 36	
tttgccttta ggattctaga cattt	25
<210> 37	
<211> 8	
<212> DNA	
<213> 人工序列	
<400> 37	
agacattt	8
<210> 38	
<211> 13	
<212> DNA	
<213> 人工序列	
<400> 38	
attctagaca ttt	13

<210>	39	
<211>	44	
<212>	DNA	
<213>	人工序列	
<400>	39	
	tttgccttta ggagggttgg gtttgggttg gtttctagac attt	44
<210>	40	
<211>	88	
<212>	DNA	
<213>	人工序列	
<400>	40	
	tttgccttta ggagggttgg gtttgggttg ggtttgcctt taggattcta gacatttggg	60
	ttgggttttg gttgggttct agacattt	88
<210>	41	
<211>	19	
<212>	DNA	
<213>	人工序列	
<400>	41	
	gggtttgggt gggtttggg	19
<210>	42	
<211>	16	
<212>	DNA	
<213>	人工序列	
<400>	42	
	gggtgggtgg gtgggt	16
<210>	43	
<211>	26	
<212>	DNA	
<213>	人工序列	
<400>	43	
	ggtggtggtg gttgtggtgg tggtag	26
<210>	44	
<211>	24	
<212>	DNA	
<213>	人工序列	
<400>	44	
	gtggggcatt gtgggtgggt gtgg	24
<210>	45	
<211>	22	

<212> DNA	
<213> 人工序列	
<400> 45	
agggttaggg ttagggtag gg	22
<210> 46	
<211> 15	
<212> DNA	
<213> 人工序列	
<400> 46	
ggttggtgtg gttgg	15
<210> 47	
<211> 19	
<212> DNA	
<213> 人工序列	
<400> 47	
gggtagggcg gggttgggg	19
<210> 48	
<211> 29	
<212> DNA	
<213> 人工序列	
<400> 48	
ggtggtggtg gttgtggtgg gtgggtggg	29
<210> 49	
<211> 52	
<212> DNA	
<213> 人工序列	
<400> 49	
ggtggtggtg gttgtggtgg tggtaggggtg gtggtggttg tggtaggtgg gg	52
<210> 50	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 50	
ggttaggtta ggttagg	17
<210> 51	
<211> 28	
<212> DNA	
<213> 人工序列	
<400> 51	

ggggttttgg ggttttgggg ttttgggg	28
<210> 52	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 52	
ggcttaggct taggcttagg	20
<210> 53	
<211> 24	
<212> DNA	
<213> 人工序列	
<400> 53	
gggttcaggg ttcagggttc aggg	24
<210> 54	
<211> 23	
<212> DNA	
<213> 人工序列	
<400> 54	
agggttaggg ttagggtag ggt	23
<210> 55	
<211> 21	
<212> DNA	
<213> 人工序列	
<400> 55	
gggctagggc tagggctagg g	21
<210> 56	
<211> 22	
<212> DNA	
<213> 人工序列	
<400> 56	
tgagggtggg tagggtgggt aa	22
<210> 57	
<211> 21	
<212> DNA	
<213> 人工序列	
<400> 57	
gggagggcgc tgggaggagg g	21
<210> 58	
<211> 20	

<212> DNA	
<213> 人工序列	
<400> 58	
gggcgggCGC gagggagggg	20
<210> 59	
<211> 28	
<212> DNA	
<213> 人工序列	
<400> 59	
acaggggtgt ggggacaggg gtgtgggg	28
<210> 60	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 60	
gcgggCGGCT cgggCGCGGG	20
<210> 61	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 61	
gggagggttg gggtggg	17
<210> 62	
<211> 19	
<212> DNA	
<213> 人工序列	
<400> 62	
gggcaggag ggaactggg	19
<210> 63	
<211> 18	
<212> DNA	
<213> 人工序列	
<400> 63	
gggagggttg gggtgggt	18
<210> 64	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 64	

gggcagggag ggaactgggt	20
<210> 65	
<211> 27	
<212> DNA	
<213> 人工序列	
<400> 65	
gggcggtgtg ggaagagggg agagggg	27
<210> 66	
<211> 16	
<212> DNA	
<213> 人工序列	
<400> 66	
ggttggtgtg gttggt	16
<210> 67	
<211> 12	
<212> DNA	
<213> 人工序列	
<400> 67	
gggcttttgg gc	12
<210> 68	
<211> 26	
<212> DNA	
<213> 人工序列	
<400> 68	
tttaagggtt agggtaggg ttaggg	26
<210> 69	
<211> 24	
<212> DNA	
<213> 人工序列	
<400> 69	
tgagggtggt gaggtgggg aagg	24
<210> 70	
<211> 23	
<212> DNA	
<213> 人工序列	
<400> 70	
tgagggtggg tagggtgggt aat	23
<210> 71	
<211> 24	

<212> DNA	
<213> 人工序列	
<400> 71	
tgagggtggg agggtgggga agga	24
<210> 72	
<211> 23	
<212> DNA	
<213> 人工序列	
<400> 72	
gggcgcggga ggaattgggc ggg	23
<210> 73	
<211> 22	
<212> DNA	
<213> 人工序列	
<400> 73	
agggagggcg ctgggaggag gg	22
<210> 74	
<211> 21	
<212> DNA	
<213> 人工序列	
<400> 74	
cgggcgggcg cgagggaggg t	21
<210> 75	
<211> 21	
<212> DNA	
<213> 人工序列	
<400> 75	
cgggcgggcg ctagggaggg t	21
<210> 76	
<211> 12	
<212> DNA	
<213> 人工序列	
<400> 76	
ggaggaggag ga	12
<210> 77	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 77	

ggttggtgtg gttggtt	17
<210> 78	
<211> 27	
<212> DNA	
<213> 人工序列	
<400> 78	
ggtggtggtg gttgtggtgg tgggtgt	27
<210> 79	
<211> 16	
<212> DNA	
<213> 人工序列	
<400> 79	
gggcggggcgg gcgggc	16
<210> 80	
<211> 25	
<212> DNA	
<213> 人工序列	
<400> 80	
tgtgggggtg gacgggccgg gtaga	25
<210> 81	
<211> 7	
<212> DNA	
<213> 人工序列	
<400> 81	
ttgggtt	7
<210> 82	
<211> 6	
<212> DNA	
<213> 人工序列	
<400> 82	
tgggag	6
<210> 83	
<211> 16	
<212> DNA	
<213> 人工序列	
<400> 83	
ggggtgggag gaggggt	16
<210> 84	
<211> 18	

<212> DNA	
<213> 人工序列	
<400> 84	
ggttggtgtg gttggttt	18
<210> 85	
<211> 24	
<212> DNA	
<213> 人工序列	
<400> 85	
atggggtcgg gcgggccggg tgtc	24
<210> 86	
<211> 18	
<212> DNA	
<213> 人工序列	
<400> 86	
gtgggtaggg cgggttgg	18
<210> 87	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 87	
gggttagggt taggtaggg	20
<210> 88	
<211> 25	
<212> DNA	
<213> 人工序列	
<400> 88	
cctgggggag tattgcggag gaagg	25
<210> 89	
<211> 32	
<212> DNA	
<213> 人工序列	
<400> 89	
agggcggtgt gggaaagaggg aagaggggga gg	32
<210> 90	
<211> 35	
<212> DNA	
<213> 人工序列	
<400> 90	

agggcgggtgt ggggaagaggg aagaggggga ggcag	35
<210> 91	
<211> 39	
<212> DNA	
<213> 人工序列	
<400> 91	
cggtcgctcc gtgtggcttg gttgggtgt ggcagtgac	39
<210> 92	
<211> 31	
<212> DNA	
<213> 人工序列	
<400> 92	
gggggtggga gggtaggcct taggtttctg a	31
<210> 93	
<211> 37	
<212> DNA	
<213> 人工序列	
<400> 93	
cgcctgatta gcgatactca gcgttggggg ggggggg	37
<210> 94	
<211> 36	
<212> DNA	
<213> 人工序列	
<400> 94	
gggcgtgggtg ggtgggttac taataatgtg cgtttg	36
<210> 95	
<211> 30	
<212> DNA	
<213> 人工序列	
<400> 95	
agcgggcata tgggtgtggg tggatatggtc	30
<210> 96	
<211> 36	
<212> DNA	
<213> 人工序列	
<400> 96	
aacacatagg tttggttagg ttggttggtt gaatta	36
<210> 97	
<211> 30	

<212> DNA	
<213> 人工序列	
<400> 97	
ttgcgcgtta attggggggg tgggtgggtt	30
<210> 98	
<211> 30	
<212> DNA	
<213> 人工序列	
<400> 98	
ggtggtgggg ggggttggtta ggggtgtcttc	30
<210> 99	
<211> 36	
<212> DNA	
<213> 人工序列	
<400> 99	
gatcgggtgt ggggtggcgta aaggagcat cggaca	36
<210> 100	
<211> 45	
<212> DNA	
<213> 人工序列	
<400> 100	
gggttagggg tagggtagg gttagggtta ggttagggg taggg	45
<210> 101	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 101	
ggggcgctta tggggagggt ggggagggtg gggaaggtgg ggaggag	47
<210> 102	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 102	
ccccccgggg cgggccgggg gcgggggtccc ggcggggcgg agccatg	47
<210> 103	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 103	

ggcgggggggg gggggggcggg ggcggggggcg ggggaggggc gcggcgg	47
<210> 104	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 104	
cgcggagggg cgggcgcggg aggaaggggg cgggagcggg gctgtgg	47
<210> 105	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 105	
gaggagcggg taggggcggg gcggggcggg ggcggtccag gggtggg	47
<210> 106	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 106	
cgctcgggcg cgcggggagg ggagaggggg cgggagcgcg cccgctc	47
<210> 107	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 107	
agaggaggag gaggtcacgg aggaggagga gaaggaggag gagga	47
<210> 108	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 108	
ggggcgctta tggggagggt gggtagggtg ggtaaggtgg ggaggag	47
<210> 109	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 109	
ggggcgctta tggggagggt gggtagggtt gggaaggtgg ggaggag	47
<210> 110	
<211> 17	

<212> DNA	
<213> 人工序列	
<400> 110	
gggaagggag ggaaggg	17
<210> 111	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 111	
gggttgggtg ggttggg	17
<210> 112	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 112	
gggccgggcg ggccggg	17
<210> 113	
<211> 18	
<212> DNA	
<213> 人工序列	
<400> 113	
gggttgggtt gggttggg	18
<210> 114	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 114	
ggtagggcg ggttggg	17
<210> 115	
<211> 15	
<212> DNA	
<213> 人工序列	
<400> 115	
gggtgggtgg gtggg	15
<210> 116	
<211> 16	
<212> DNA	
<213> 人工序列	
<400> 116	

gggtgggttg ggtggg	16
<210> 117	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 117	
gggtgggttt ggtggg	17
<210> 118	
<211> 21	
<212> DNA	
<213> 人工序列	
<400> 118	
gggtttgggt ttgggtttg g	21
<210> 119	
<211> 24	
<212> DNA	
<213> 人工序列	
<400> 119	
gggttttggg ttttgggtt tggg	24
<210> 120	
<211> 27	
<212> DNA	
<213> 人工序列	
<400> 120	
gggtttttgg gtttttgggt ttttggg	27
<210> 121	
<211> 23	
<212> DNA	
<213> 人工序列	
<400> 121	
ggggttgggg tttggggtt ggg	23
<210> 122	
<211> 27	
<212> DNA	
<213> 人工序列	
<400> 122	
gggggttggg gtttggggg ttggggg	27
<210> 123	
<211> 31	

<212> DNA	
<213> 人工序列	
<400> 123	
ggggggttgg ggggtttggg gggttggggg g	31
<210> 124	
<211> 35	
<212> DNA	
<213> 人工序列	
<400> 124	
gggggggttg ggggggtttg ggggggttgg ggggg	35
<210> 125	
<211> 39	
<212> DNA	
<213> 人工序列	
<400> 125	
gggggggggtt ggggggggtt tgggggggggt tgggggggg	39
<210> 126	
<211> 43	
<212> DNA	
<213> 人工序列	
<400> 126	
gggggggggtt tggggggggg tttggggggg ggttgggggg ggg	43
<210> 127	
<211> 47	
<212> DNA	
<213> 人工序列	
<400> 127	
gggggggggg ttgggggggg ggtttggggg gggggttggg ggggggg	47
<210> 128	
<211> 22	
<212> DNA	
<213> 人工序列	
<400> 128	
gggtagggcg ggttgggtag gg	22
<210> 129	
<211> 26	
<212> DNA	
<213> 人工序列	
<400> 129	

gggtagggcg ggttgggtag ggcggg	26
<210> 130	
<211> 31	
<212> DNA	
<213> 人工序列	
<400> 130	
gggtagggcg ggttgggtag ggcgggttg g	31
<210> 131	
<211> 34	
<212> DNA	
<213> 人工序列	
<400> 131	
gggtagggcg ggttgggggg tagggcgggt tggg	34
<210> 132	
<211> 16	
<212> DNA	
<213> 人工序列	
<400> 132	
aatgtagaga tgcggt	16
<210> 133	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 133	
aatgtagaga tgcggtg	17
<210> 134	
<211> 18	
<212> DNA	
<213> 人工序列	
<400> 134	
gaatgtagag atgcggtg	18
<210> 135	
<211> 19	
<212> DNA	
<213> 人工序列	
<400> 135	
gaatgtagag atgcggtgg	19
<210> 136	
<211> 20	

<212> DNA	
<213> 人工序列	
<400> 136	
tgaatgtaga gatgcbgtgg	20
<210> 137	
<211> 25	
<212> DNA	
<213> 人工序列	
<400> 137	
tttgaatgta gagatgcbgt ggttt	25
<210> 138	
<211> 30	
<212> DNA	
<213> 人工序列	
<400> 138	
tttagattat atgcatacc tccattggtt	30
<210> 139	
<211> 33	
<212> DNA	
<213> 人工序列	
<400> 139	
tttggacggg gttaaacgaa aggggacgac ctt	33
<210> 140	
<211> 35	
<212> DNA	
<213> 人工序列	
<400> 140	
ggatagggat tctgttggtc ggctgggttg tatcc	35
<210> 141	
<211> 50	
<212> DNA	
<213> 人工序列	
<400> 141	
atcaggctgg atggtagctc ggtcggggtg ggtgggttg caagtctgat	50
<210> 142	
<211> 80	
<212> DNA	
<213> 人工序列	
<400> 142	

gctgtgtgac tcctgcaaag tgtggacaac ttcccacgga ggaattcccg tatctaaagg	60
tgcagctgta tcttgtctcc	80
<210> 143	
<211> 200	
<212> DNA	
<213> 人工序列	
<400> 143	
aacaccggtt tgccatgtgt tgccatgtgt atgtggggag acggtcgggt ccagatattc	60
gtatctgtcg agtagagtgt gggctcecca catactctga tgatccgaga cggtcgggtc	120
cagatattcg tatctgtcga gtagagtgtg ggctcggatc attcatggca actctgatga	180
tcctctatct atctgtcgag	200
<210> 144	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 144	
tgcccttagg attctagaca	20
<210> 145	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 145	
atggaggat gacatataat	20
<210> 146	
<211> 17	
<212> DNA	
<213> 人工序列	
<400> 146	
tacattgcct cttcatt	17
<210> 147	
<211> 19	
<212> DNA	
<213> 人工序列	
<400> 147	
ttccgaataa actccaggc	19
<210> 148	
<211> 19	
<212> DNA	
<213> 人工序列	

<400> 148	
aagcaaaaca ggtctagaa	19
<210> 149	
<211> 19	
<212> DNA	
<213> 人工序列	
<400> 149	
gtgtctagga gatacacct	19
<210> 150	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 150	
ttccgaataa actccaggct	20
<210> 151	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 151	
aagcaaaaca ggtctagaat	20
<210> 152	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 152	
gtgtctagga gatacacctt	20
<210> 153	
<211> 63	
<212> RNA	
<213> 人工序列	
<400> 153	
ucuaucuauc ugucgaguag agugugggcu cuuuuggaag aaacuguggc acuucggugc	60
cag	63
<210> 154	
<211> 59	
<212> RNA	
<213> 人工序列	
<400> 154	
ggguagggcg gguuggguuu gccuuuagga uucuagacau uuggguaggg cggguuggg	59

<210>	155								
<211>	63								
<212>	RNA								
<213>	人工序列								
<400>	155								
	ggguuggguu	uggguugggu	uugccuuuag	gauucuagac	auuuggguug	gguuuggguu			60
	ggg								63
<210>	156								
<211>	2038								
<212>	RNA								
<213>	人工序列								
<400>	156								
	ggguuggguu	uggguugggu	uauuuuuuag	agcuguggag	uucuuaaaua	ucaaccaugg			60
	cacuuucucc	ugaccccuuc	ccuaggggau	uucaggauug	agaaauuuu	ccaucgagcc			120
	uuuuuaaaa	uguaggacu	guuccugugg	gcuucaguga	ugggauagua	cacuucacuc			180
	agaggcauuu	gcaucuuuaa	auaaauucuu	aaaagccucu	aaagugauca	gugccuugau			240
	gccacaauag	gaaauuuuu	uagcauuuag	ucucugaagg	cucuaugaaa	ggaauagcau			300
	gaugugcugu	uagaaucaaga	uguuacugcu	aaaauuuaca	uguugugaug	uaaaauugugu			360
	agaaaaccu	uaaaucuuuc	aaaauaauaa	acuauuuuu	uuagagaau	uauacuuuu			420
	gaaagcuguc	uccuuuuuu	aaauaaaauag	uguuuugucug	uaguucagug	uuggggcaau			480
	cuuggggggg	auucuuucuc	aaucuuucag	aaacuuuguc	ugcgaacacu	cuuuuaugga			540
	ccagaucagg	auuugagcgg	aagaacgaau	guaacuuua	ggcaggaaag	acaaauuuu			600
	uucuucauaa	agugaugagc	auauaauaau	uccaggcaca	uggcaauaga	ggcccucuaa			660
	auaaggaaua	auaaccucu	uagacaggug	ggagauuau	aucagagua	aagguaauu			720
	cacauuuuau	uuccagaaag	ucaggggucu	auaaaauugac	agugauuaga	gaaauacuuu			780
	uucacauuuc	caaaguuuuc	auguuuacuu	uaaaugcuua	caucuuuaga	gugguaggca			840
	auguuuuaca	cuauugaccu	uauauaggga	agggaggggg	ugccuguggg	guuuuaaaga			900
	auuuuccuuu	gcagaggcau	uucuuucuu	augaagccau	ucaggauuuu	gaaauugcau			960
	ugagugcuug	gcucuuccuu	cuguucuuu	gaguguuaga	gaccuugcag	ugaguuuuac			1020
	agcauacuca	aaaauuuuuu	ccuggaauuu	ggagggauug	gaggaggggg	uggggcuuac			1080
	uuguuguagc	uuuuuuuuuu	uuuacagacu	ucacagagaa	ugcaguuguc	uugacuucag			1140
	gucugucugu	ucuguuggca	aguaaaugca	guacuguucu	gaucccgcug	cuuuuagaau			1200
	gcauugugaa	acgacuggag	uaugauuaaa	aguuguguuc	cccaaugcuu	ggaguaguga			1260
	uuguugaagg	aaaaaaucua	gcugagugau	aaaggcugag	uguugaggaa	auuucugcag			1320
	uuuuuagcag	ucguuuuuuu	gauugaagcu	gaguacuuuu	ugcuggugua	uuuuuaggua			1380
	aaaugcuuuu	uguucuuuuu	uggugguggg	aggggacuga	agccuuuagu	cuuuuccaga			1440
	ugcaaccuuu	aaaucaguga	caagaaacau	uccaaacaag	caacagucuu	caagaaauu			1500
	aacuggcaag	uggaaauguu	uaaacaguuc	agugaucuuu	agugcauuu	uuauuguggg			1560
	guuucucucu	ccccuccuu	ggucuuuuuu	cuuacaugca	ggaacacuca	gcagacacac			1620

guaugcgaag	ggccagagaa	gccagacca	guaagaaaa	auagccuauu	uacuuuaaa	1680
aaaccaaaaca	uuccauuuua	aaugugggga	uugggaacca	cuaguucuuu	cagaugguau	1740
ucuucagacu	auagaaggag	cuuccaguug	aaucaccag	uggacaaaau	gaggaaaaca	1800
ggugaacaag	cuuuuucugu	auuuacauac	aaagucagau	caguuauagg	acaauaguau	1860
ugaauagauu	ucagcuuuau	gcuggaguua	cuggcaugug	agcaaacugu	guuggcgugg	1920
ggguggaggg	gugagguggg	cgcuagccu	uuuuuaaga	uuuuucaggu	accccucacu	1980
aaaggcaccg	aaggcuuaaa	guaggacaac	cauggagccg	gguuggguuu	ggguuggg	2038
<210>	157					
<211>	2000					
<212>	RNA					
<213>	人工序列					
<400>	157					
uuuuuuuuuaa	gagcugugga	guucuuuuuu	aucaaccaug	gcacuuucuc	cugaccccuu	60
cccuaggggga	uuucaggauu	gagaaauuuu	uccaucgagc	cuuuuuuaaa	uuguaggacu	120
uguuccugug	ggcuucagug	augggauagu	acacuucacu	cagaggcauu	ugcaucuuua	180
aaauuuuuuu	uaaaagccuc	uaaagugauc	agugccuuga	ugccaacuaa	ggaaauuugu	240
uuagcauuga	aucucugaag	gcucuaugaa	aggaauagca	ugaugugcug	uuagaauagc	300
auguuacugc	uaaaauuuac	auguugugau	guaaauugug	uagaaaacca	uuaaaucauu	360
caaaauaaua	aacuauuuuu	auuagagaau	guauacuuuu	agaaagcugu	cuccuuuuuu	420
aaauaaaaua	guguuuugucu	guaguucagu	guuggggcaa	ucuuuggggg	gauucuuuc	480
uaaucuuuca	gaaacuuugu	cugcgaacac	ucuuuaaugg	accagaucag	gauuugagcg	540
gaagaacgaa	uguaacuuua	aggcaggaaa	gacaaauuuu	auucuucaua	aagugaugag	600
cauauaauaa	uuccaggcac	auggcaauag	aggcccucua	aaauaggaau	aaauaaccuc	660
uuagacaggu	gggagauuuu	gaucagagua	aaagguaauu	acacuuuuu	uuuccagaaa	720
gucagggguc	uauaaauga	cagugauuag	aguaauacuu	uuucacauuu	caaaguuug	780
cauguuaacu	uuaaaugcuu	acaaucuuag	agugguaggc	aauguuuuac	acuauugacc	840
uuauauaggg	aaggaggagg	gugccugugg	gguuuuaaag	aauuuuccuu	ugcagaggca	900
uuucauccuu	caugaagcca	uucaggauuu	ugaauugcau	augagugcuu	ggcucuuccu	960
ucuguucuaag	ugaguguaug	agaccuugca	gugaguuuau	cagcauacuc	aaaauuuuuu	1020
uccuggaauu	uggagggaug	ggaggagggg	guggggcuua	cuuguuguag	cuuuuuuuuu	1080
uuuuacagac	uucacagaga	augcaguugu	cuugacuuca	ggucugucug	uucuguuggc	1140
aaguaaaugc	aguacuguuu	ugaucuccgc	gcuaauagaa	ugcauuguga	aacgacugga	1200
guaugauuaa	aaguuguguu	ccccaaugcu	uggaguagug	auuguugaag	gaaaaauucc	1260
agcugaguga	uaaaggcuga	guguugagga	aauuucugca	guuuuaagca	gucguauuug	1320
ugaauugaagc	ugaguacauu	uugcuggugu	auuuuuaggu	aaaauugcuu	uuguucauuu	1380
cugguggugg	gaggggacug	aagccuuuag	ucuuuuuccg	augcaaccuu	aaaauucagug	1440
acaagaaaca	uuccaaacaa	gcaacagucu	ucaagaaauu	aaacuggcaa	guggaaaugu	1500
uuaaacaguu	cagugaucuu	uagugcauug	uuuaugugug	gguuucucuc	uccccucccu	1560
uggucuuauu	ucuuacaugc	aggaacacuc	agcagacaca	cguaugcgaa	gggccagaga	1620

agccagaccc	aguaagaaaa	aauagccuau	uuacuuuaaa	uaaaccaaac	auuccauuuu	1680
aaaugugggg	auugggaacc	acuaguucuu	ucagauggua	uucuucagac	uauagaagga	1740
gcuuccaguu	gaauucacca	guggacaaaa	ugaggaaaac	aggugaacaa	gcuuuuucug	1800
uauuuacaua	caaagucaga	ucaguuauug	gacaauagua	uugaauagau	uucagcuuaa	1860
ugcuggagua	acuggcaugu	gagcaaacug	uguuggcgug	ggguggagg	ggugaggugg	1920
gcgcuagacc	uuuuuuuaag	auuuuucagg	uaccccucac	uaaaggcacc	gaaggcuuaa	1980
aguaggacaa	ccauggagcc					2000
<210>	158					
<211>	58					
<212>	DNA					
<213>	人工序列					
<400>	158					
gggttgggtt	tgggttgggu	gaauguagag	augcgguggg	ggttgggttt	gggttggg	58
<210>	159					
<211>	58					
<212>	DNA					
<213>	人工序列					
<400>	159					
tctatctatc	tgtcgagtau	gaauguagag	augcggugaa	actgtggcac	ttcgggtgc	58
<210>	160					
<211>	59					
<212>	DNA					
<213>	人工序列					
<400>	160					
ggguuggguu	uggguugggt	gaatgtagag	atgcggtggg	gguuggguuu	ggguuuggg	59
<210>	161					
<211>	59					
<212>	DNA					
<213>	人工序列					
<400>	161					
ucuaucuauc	ugucgaguat	gaatgtagag	atgcggtaga	aacuguggca	cuucgguga	59
<210>	162					
<211>	21					
<212>	DNA					
<213>	人工序列					
<400>	162					
tgaggcgaa	tcaaatccat	c				21
<210>	163					
<211>	16					

<212> DNA	
<213> 人工序列	
<400> 163	
acattgcctc ttcatt	16
<210> 164	
<211> 13	
<212> DNA	
<213> 人工序列	
<400> 164	
gtctgtggaa gcg	13
<210> 165	
<211> 20	
<212> DNA	
<213> 人工序列	
<400> 165	
gcctgtctgt ggaagcgggt	20
<210> 166	
<211> 57	
<212> DNA	
<213> 人工序列	
<400> 166	
ggcaagctga ccctgaagtt catctgcacc accggcaagc tgcccgtgcc ctggccc	57

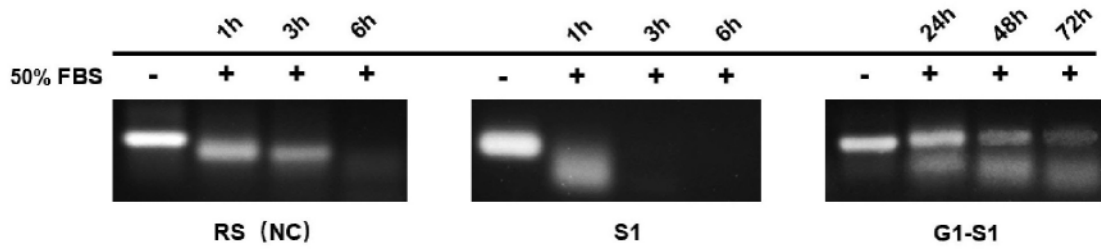


图 1

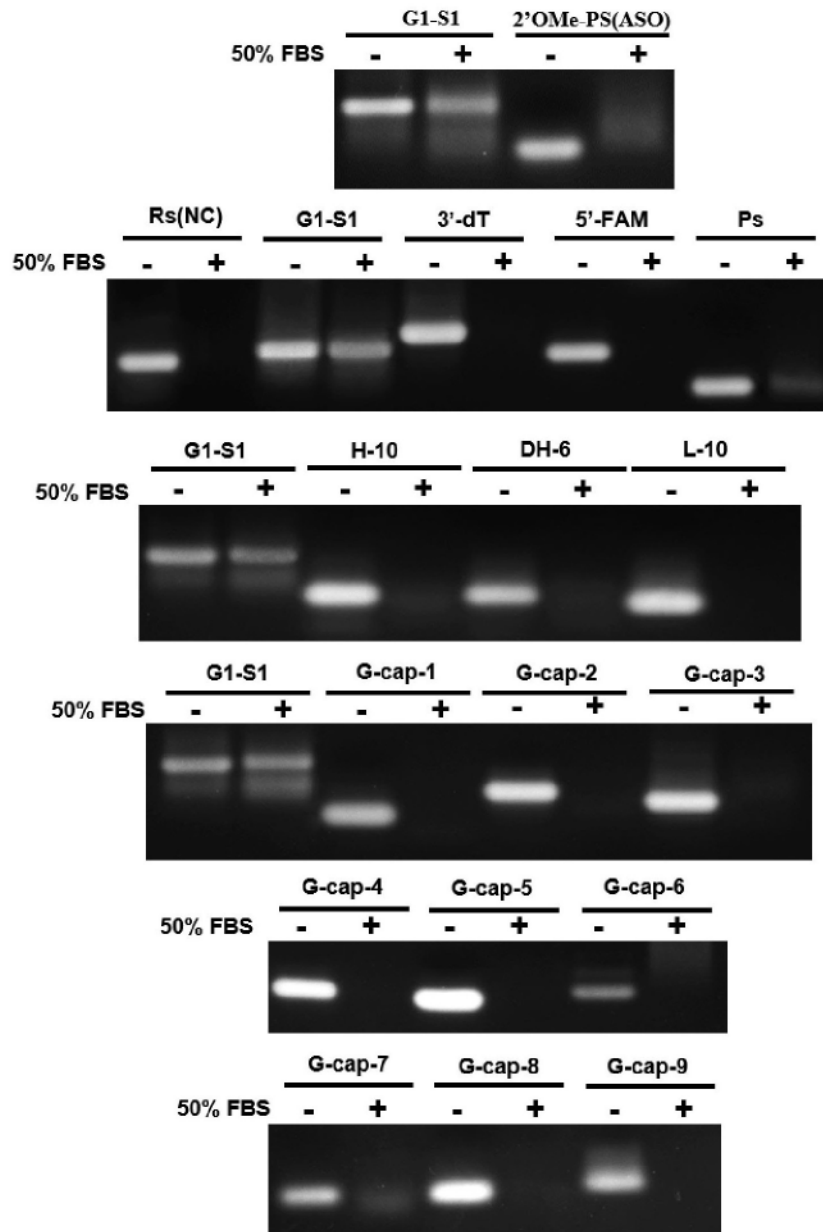


图 2

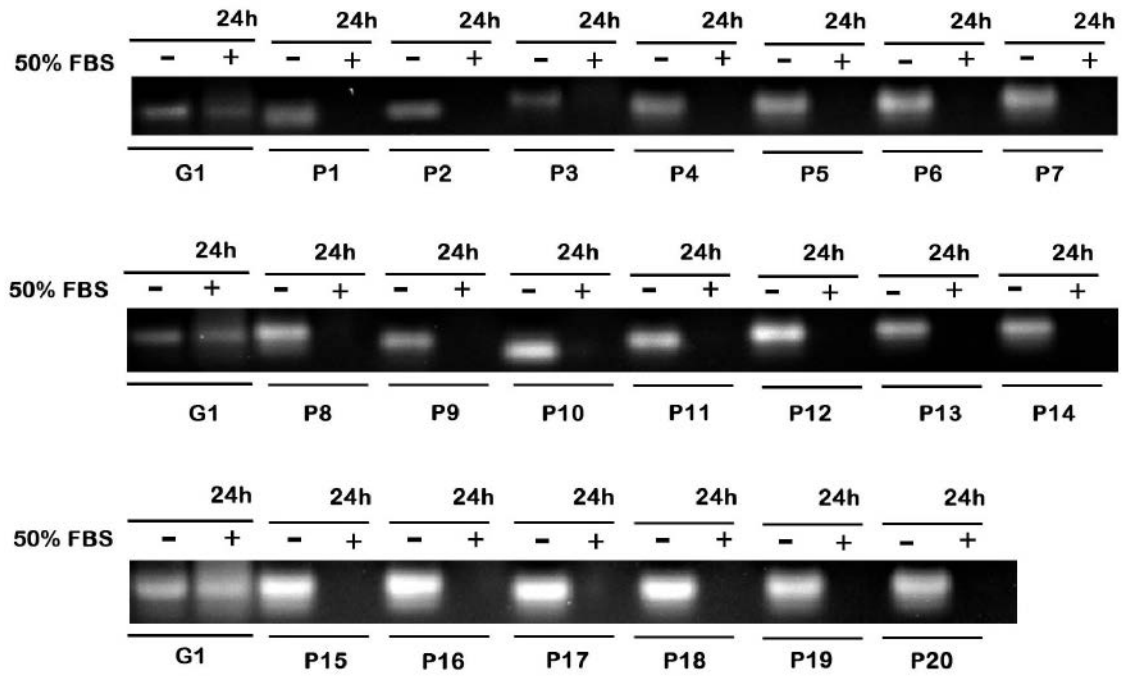


图 2(续)

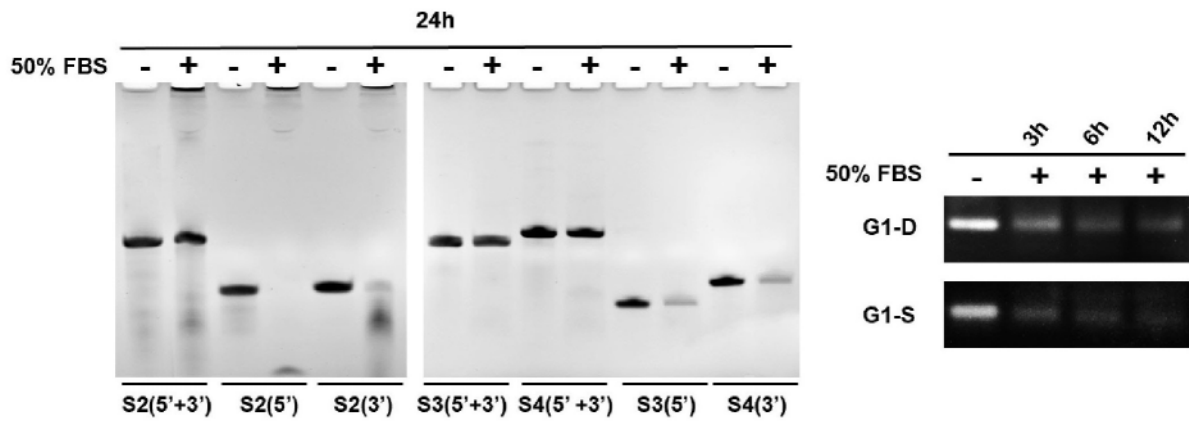


图 3

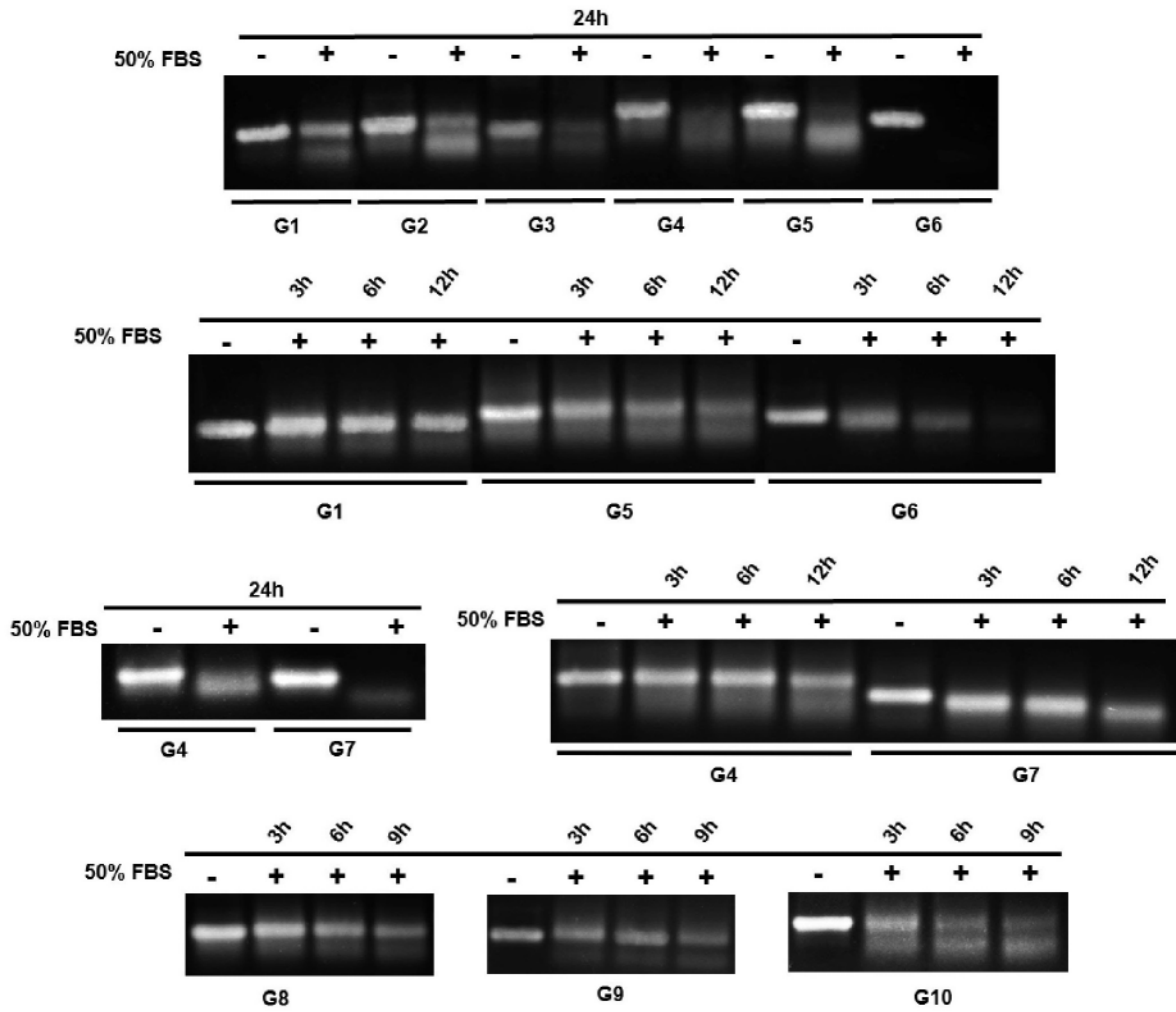


图 4

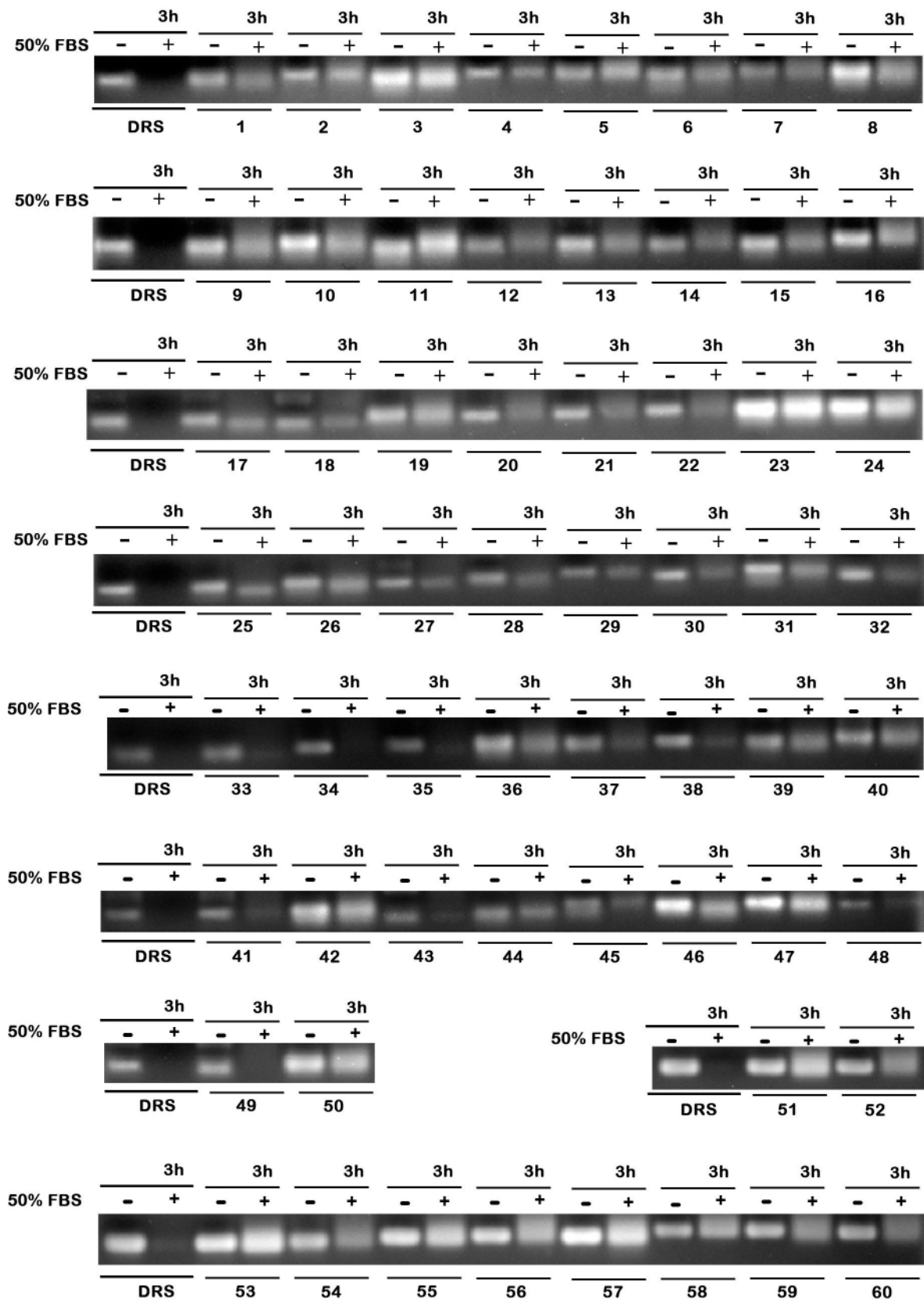


图 4(续)

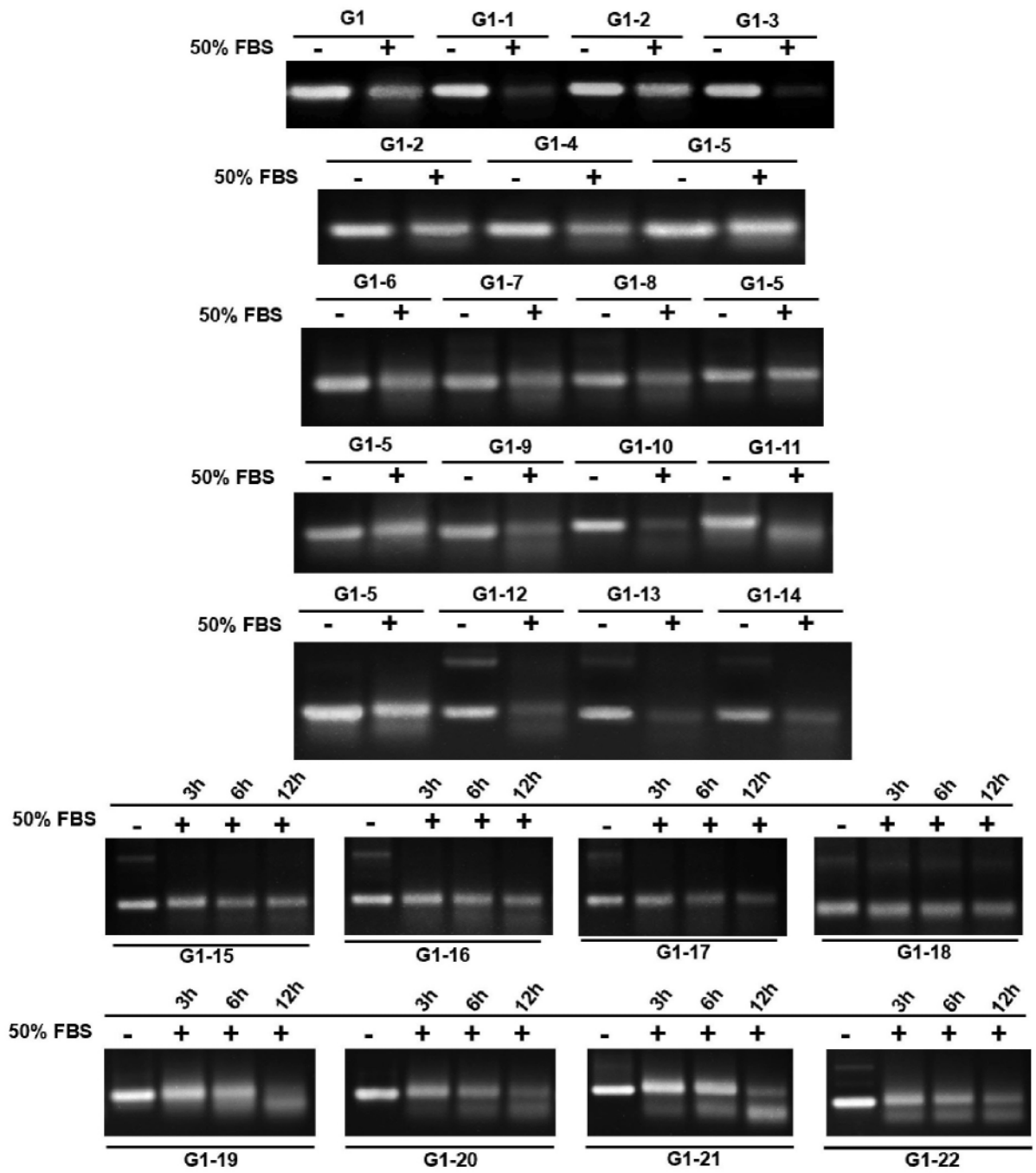


图 5

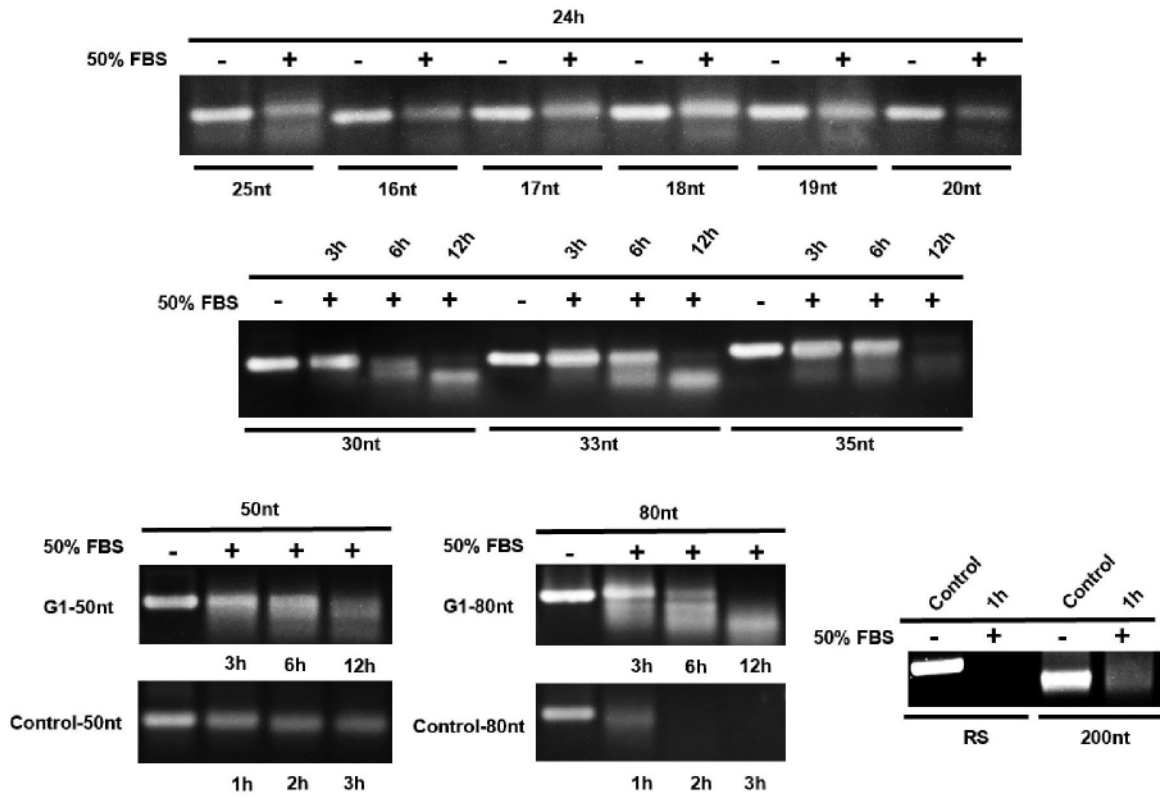


图 6

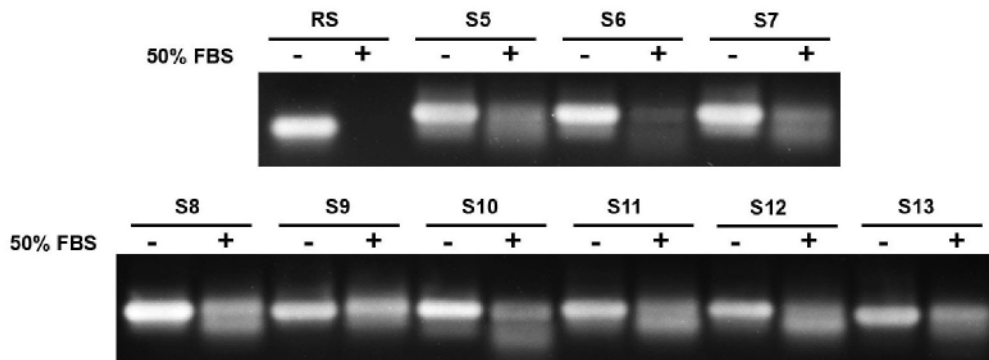


图 7

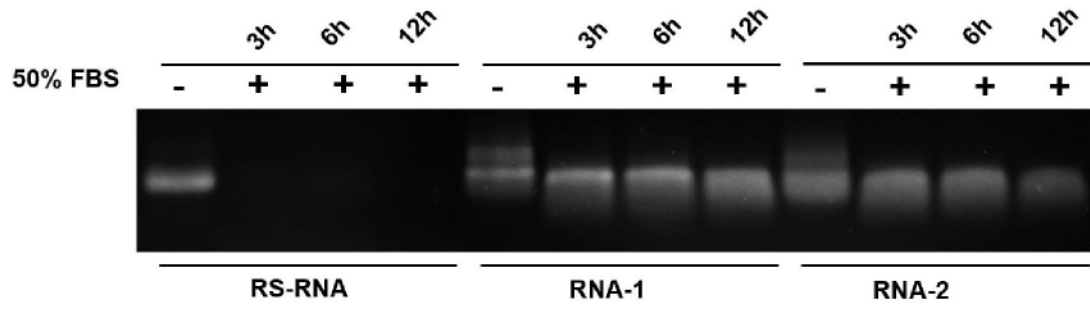


图 8

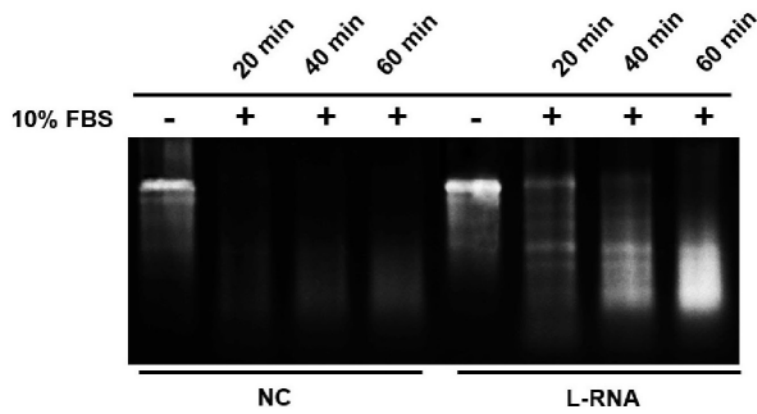


图 9

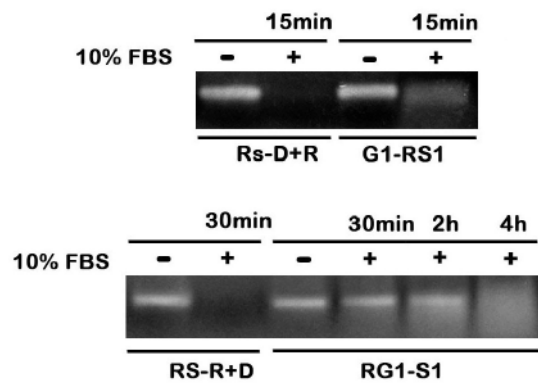


图 10

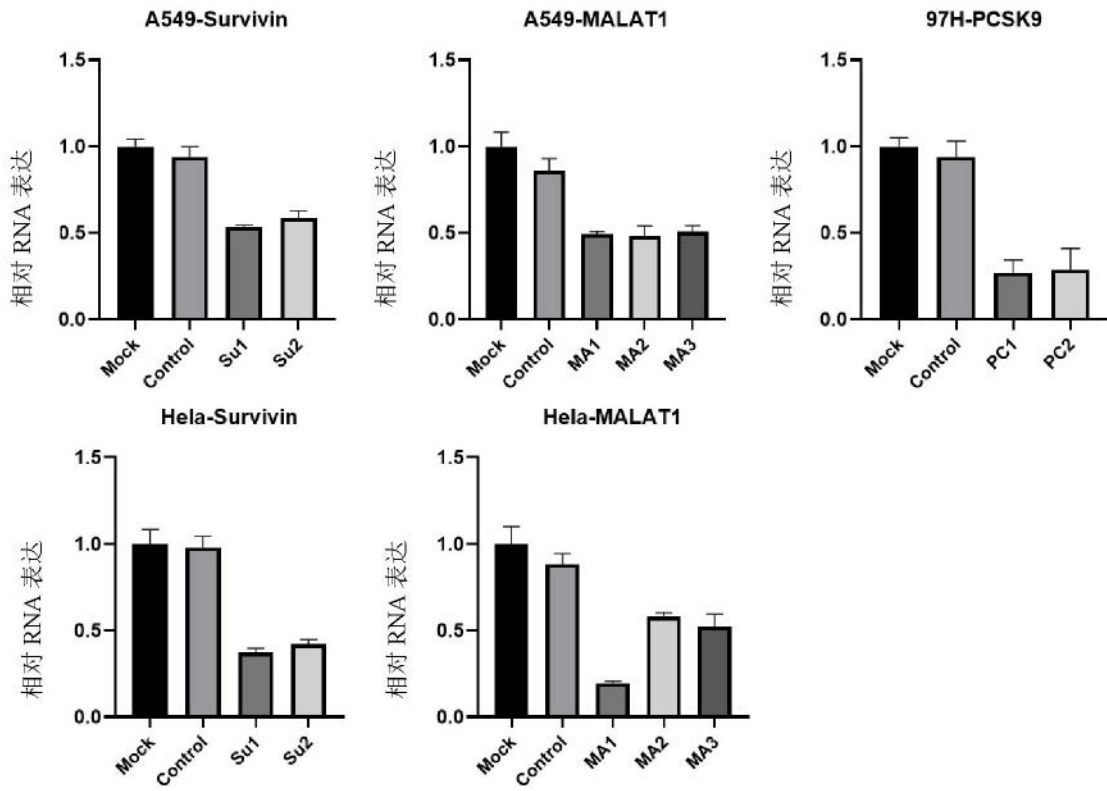
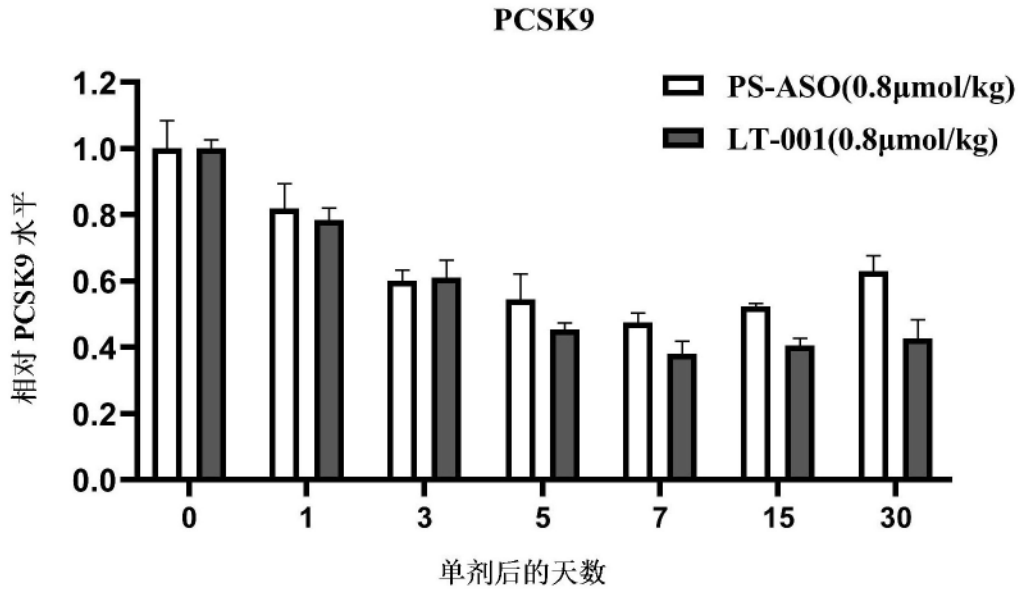
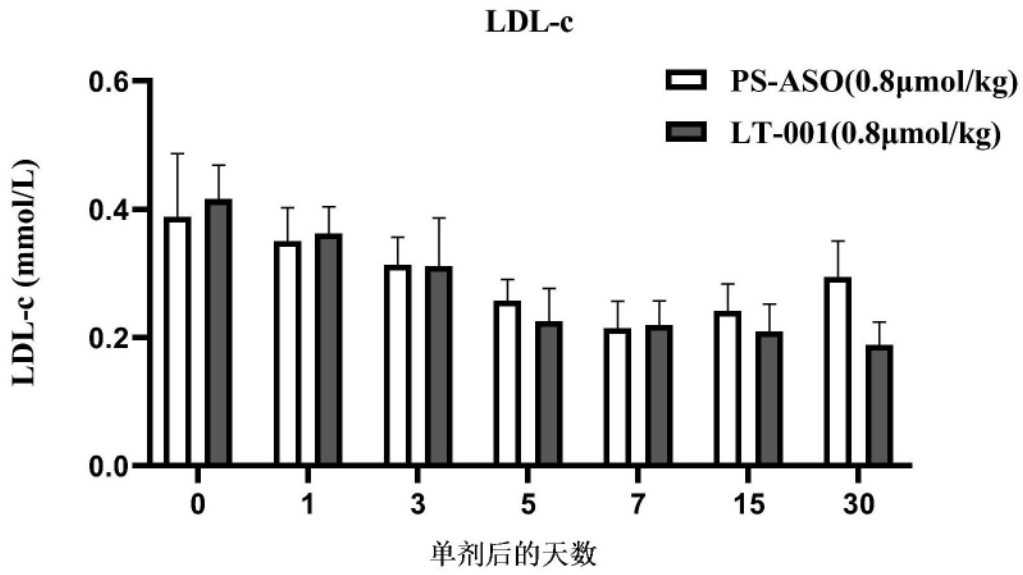


图 11



(a)



(b)

图 12