(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118401553 A
(43)申请公布日 2024.07.26

(21)申请号 202280060769.5
(22)申请日 2022.07.08
(30)优先权数据
63/220,030 2021.07.09 US
(85)PCT国际申请进入国家阶段日
2024.03.07
(86)PCT国际申请的申请数据
PCT/US2022/073534 2022.07.08
(87)PCT国际申请的公布数据
WO2023/283619 EN 2023.01.12
(71)申请人 达因疗法公司
地址 美国
(72)发明人 科迪·A·德雅尔丹 基姆·唐
詹姆斯·麦克斯维根
罗梅什·R·苏布拉马尼亚
蒂莫西·威登
穆罕默德·T·卡塔纳尼
布伦丹·奎因 约翰·纳吉姆
(74)专利代理机构 北京集佳知识产权代理有限公司 11227
专利代理师 张福誉 刘振佳
(51)Int.Cl.
C07K 16/28(2006.01)
A61K 39/395(2006.01)
A61K 47/68(2006.01)
A61P 21/00(2006.01)
C12N 15/113(2006.01)

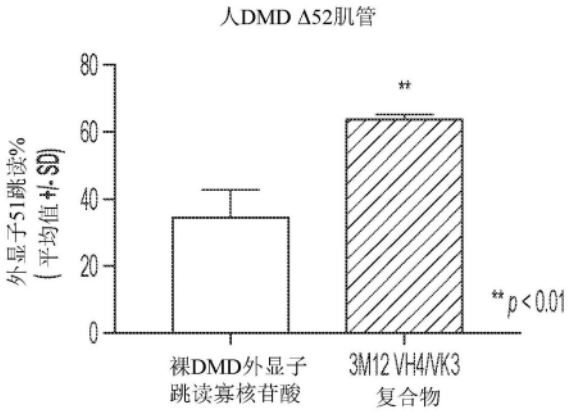
权利要求书4页 说明书126页
序列表（电子公布） 附图1页

(54)发明名称

肌肉靶向复合物及其用于治疗肌养蛋白病的用途

(57)摘要

本公开内容的一些方面涉及包含与分子载荷共价连接的肌肉靶向剂的复合物。在一些实施方案中,所述肌肉靶向剂与肌细胞上的内化细胞表面受体特异性结合。在一些实施方案中,所述分子载荷促进功能性肌养蛋白的表达或活性。在一些实施方案中,所述分子载荷是寡核苷酸,例如反义寡核苷酸,例如引起由突变体DMD等位基因表达的mRNA中的外显子跳读的寡核苷酸。



CN 118401553 A

1.复合物,其包含与寡核苷酸共价连接的抗转铁蛋白受体1(TfR1)抗体,所述寡核苷酸被配置成用于诱导DMD前体mRNA中外显子51的跳读,其中所述寡核苷酸包含与SEQ ID NO:160至383中任一者的至少8个连续核苷酸互补的互补区。

2.权利要求1所述的复合物,其中所述抗TfR1抗体包含:

(i)SEQ ID NO:33的重链互补决定区1(CDR-H1)、SEQ ID NO:34的重链互补决定区2(CDR-H2)、SEQ ID NO:35的重链互补决定区3(CDR-H3)、SEQ ID NO:36的轻链互补决定区1(CDR-L1)、SEQ ID NO:37的轻链互补决定区2(CDR-L2)和SEQ ID NO:32的轻链互补决定区3(CDR-L3);

(ii)SEQ ID NO:7的CDR-H1、SEQ ID NO:8的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3;

(iii)SEQ ID NO:7的CDR-H1、SEQ ID NO:20的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3;

(iv)SEQ ID NO:7的CDR-H1、SEQ ID NO:24的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3;

(v)SEQ ID NO:51的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3;

(vi)SEQ ID NO:64的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3;或者

(vii)SEQ ID NO:67的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3。

3.权利要求1或权利要求2所述的复合物,其中所述抗TfR1抗体包含:

(i)含有与SEQ ID NO:76具有至少85%同一性的氨基酸序列的重链可变区(VH);和/或含有与SEQ ID NO:75具有至少85%同一性的氨基酸序列的轻链可变区(VL);

(ii)含有与SEQ ID NO:69具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85%同一性的氨基酸序列的VL;

(iii)含有与SEQ ID NO:71具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85%同一性的氨基酸序列的VL;

(iv)含有与SEQ ID NO:72具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85%同一性的氨基酸序列的VL;

(v)含有与SEQ ID NO:73具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:74具有至少85%同一性的氨基酸序列的VL;

(vi)含有与SEQ ID NO:73具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:75具有至少85%同一性的氨基酸序列的VL;

(vii)含有与SEQ ID NO:76具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:74具有至少85%同一性的氨基酸序列的VL;

(viii)含有与SEQ ID NO:77具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:78具有至少85%同一性的氨基酸序列的VL;

(ix)含有与SEQ ID NO:79具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:80具有至少85%同一性的氨基酸序列的VL;或者

(x)含有与SEQ ID NO:77具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:80具有至少85%同一性的氨基酸序列的VL。

4.权利要求1至3中任一项所述的复合物,其中所述抗TfR1抗体包含:

(i)含有SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL;

(ii)含有SEQ ID NO:69的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

(iii)含有SEQ ID NO:71的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

(iv)含有SEQ ID NO:72的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

(v)含有SEQ ID NO:73的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL;

(vi)含有SEQ ID NO:73的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL;

(vii)含有SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL;

(viii)含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:78的氨基酸序列的VL;

(ix)含有SEQ ID NO:79的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL;或者

(x)含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL。

5.权利要求1至4中任一项所述的复合物,其中所述抗TfR1抗体是Fab片段、Fab'片段、F(ab')2片段、scFv、Fv或全长IgG。

6.权利要求5所述的复合物,其中所述抗TfR1抗体是Fab片段。

7.权利要求6所述的复合物,其中所述抗TfR1抗体包含:

(i)含有与SEQ ID NO:101具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:90具有至少85%同一性的氨基酸序列的轻链;

(ii)含有与SEQ ID NO:97具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:85具有至少85%同一性的氨基酸序列的轻链;

(iii)含有与SEQ ID NO:98具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:85具有至少85%同一性的氨基酸序列的轻链;

(iv)含有与SEQ ID NO:99具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:85具有至少85%同一性的氨基酸序列的轻链;

(v)含有与SEQ ID NO:100具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:89具有至少85%同一性的氨基酸序列的轻链;

(vi)含有与SEQ ID NO:100具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:90具有至少85%同一性的氨基酸序列的轻链;

(vii)含有与SEQ ID NO:101具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:89具有至少85%同一性的氨基酸序列的轻链;

(viii)含有与SEQ ID NO:102具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:93具有至少85%同一性的氨基酸序列的轻链;

(ix)含有与SEQ ID NO:103具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:95具有至少85%同一性的氨基酸序列的轻链;或者

(x)含有与SEQ ID NO:102具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:95具有至少85%同一性的氨基酸序列的轻链。

8.权利要求6或权利要求7所述的复合物,其中所述抗TfR1抗体包含:

(i)含有SEQ ID NO:101的氨基酸序列的重链;和含有SEQ ID NO:90的氨基酸序列的轻链;

(ii)含有SEQ ID NO:97的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

(iii)含有SEQ ID NO:98的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

(iv)含有SEQ ID NO:99的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

(v)含有SEQ ID NO:100的氨基酸序列的重链;和含有SEQ ID NO:89的氨基酸序列的轻链;

(vi)含有SEQ ID NO:100的氨基酸序列的重链;和含有SEQ ID NO:90的氨基酸序列的轻链;

(vii)含有SEQ ID NO:101的氨基酸序列的重链;和含有SEQ ID NO:89的氨基酸序列的轻链;

(viii)含有SEQ ID NO:102的氨基酸序列的重链;和含有SEQ ID NO:93的氨基酸序列的轻链;

(ix)含有SEQ ID NO:103的氨基酸序列的重链;和含有SEQ ID NO:95的氨基酸序列的轻链;或者

(x)含有SEQ ID NO:102的氨基酸序列的重链;和含有SEQ ID NO:95的氨基酸序列的轻链。

9.权利要求1至8中任一项所述的复合物,其中所述抗TfR1抗体不与转铁蛋白受体1的转铁蛋白结合位点特异性结合,和/或其中所述抗TfR1抗体不抑制转铁蛋白与转铁蛋白受体1的结合。

10.权利要求1至9中任一项所述的复合物,其中所述寡核苷酸包含所述DMD前体mRNA的剪接特征的至少4个连续核苷酸的互补区。

11.权利要求10所述的复合物,其中所述剪接特征是所述DMD前体mRNA的外显子51中的外显子剪接增强子(ESE),任选地其中所述ESE包含SEQ ID NO:860至894中任一者的序列。

12.权利要求10所述的复合物,其中所述剪接特征是分支点、剪接供体位点或剪接受体位点,任选地其中所述剪接特征跨越所述DMD前体mRNA的外显子50与内含子50的连接处、在其内含子50中、跨越其内含子50与外显子51的连接处、跨越其外显子51与内含子51的连接处、在其内含子51中、或跨越其内含子51与外显子52的连接处,并且还任选地其中所述剪接特征包含SEQ ID NO:855至859和895至898中任一者的序列。

13.权利要求1至9中任一项所述的复合物,其中所述寡核苷酸包含SEQ ID NO:160至383中任一者的互补序列或包含SEQ ID NO:384至831中任一者的序列,其中每个胸腺嘧啶碱基(T)可独立且任选地用尿嘧啶碱基(U)替换,并且每个U可独立且任选地用T替换。

14.权利要求1至13中任一项所述的复合物,其中所述寡核苷酸包含一个或更多个磷酸二酰胺吗啉代,任选地其中所述寡核苷酸是磷酸二酰胺吗啉代寡聚物(PMO)。

15.权利要求1至14中任一项所述的复合物,其中所述抗TfR1抗体通过可切割接头与所述寡核苷酸共价连接,任选地其中所述可切割接头包含缬氨酸-瓜氨酸序列。

16.权利要求1至15中任一项所述的复合物,其中所述抗TfR1抗体通过与所述抗体的赖氨酸残基或半胱氨酸残基的缀合而与所述寡核苷酸共价连接。

17.靶向DMD的寡核苷酸,其中所述寡核苷酸包含SEQ ID NO:160至383中任一者的互补区,任选地其中所述互补区包含与SEQ ID NO:160至383中任一者互补的至少15个连续核苷。

18.权利要求17所述的寡核苷酸,其中所述寡核苷酸包含SEQ ID NO:384至831中任一者的至少15个连续核苷,任选地其中所述寡核苷酸包含SEQ ID NO:384至831中任一者的序列,其中每个胸腺嘧啶碱基(T)可独立且任选地用尿嘧啶碱基(U)替换,并且每个U可独立且任选地用T替换。

19.将寡核苷酸递送至细胞的方法,所述方法包括使所述细胞与权利要求1至16中任一项所述的复合物或者权利要求17或权利要求18所述的寡核苷酸接触。

20.促进细胞中肌养蛋白的表达或活性的方法,所述方法包括使所述细胞与有效促进寡核苷酸内化至所述细胞的量的权利要求1至16中任一项所述的复合物或者权利要求18或权利要求19所述的寡核苷酸接触,任选地其中所述细胞是肌细胞。

肌肉靶向复合物及其用于治疗肌养蛋白病的用途

[0001] 相关申请

[0002] 本申请根据35 U.S.C.§119(e)要求以下的优先权:2021年7月9日提交的标题为“MUSCLE TARGETING COMPLEXES AND USES THEREOF FOR TREATING DYSTROPHINOPATHIES”的美国临时申请序列No.63/220030,其内容通过引用整体并入本文。

技术领域

[0003] 本申请涉及用于将分子载荷(例如,寡核苷酸)递送至细胞的靶向复合物及其用途,特别是与疾病治疗相关的用途。

[0004] 参考电子序列表

[0005] 电子序列表(D082470066WO00-SEQ-COB.xml;大小:1,203,807字节;以及创建日期:2022年7月7日)的内容通过引用整体并入本文。

背景技术

[0006] 肌养蛋白病(dystrophinopathy)是由编码肌养蛋白(dystrophin)的基因中的突变引起的一组独特的神经肌肉疾病。肌养蛋白病包括迪谢内肌营养不良(Duchenne muscular dystrophy)、贝克肌营养不良(Becker muscular dystrophy)和X连锁的扩张型心肌病。DMD基因(“DMD”),其编码肌养蛋白,是包含79个外显子和约260万总碱基对的大基因。DMD中的许多突变(包括外显子移码、缺失、替换和重复突变)能够降低功能性肌养蛋白的表达,导致肌养蛋白病。靶向人DMD的外显子的数种药剂已被美国食品和药物管理局(Food and Drug Administration,FDA)批准,包括casimersen、viltolarsen、golodirsen和依特立生(eteplirsen)。其中,依特立生靶向外显子51。

发明内容

[0007] 根据一些方面,本公开内容提供了靶向肌细胞以用于将分子载荷递送至这些细胞的复合物,以及可在其中使用的分子载荷。在一些实施方案中,本文中提供的复合物特别可用于递送提高或恢复功能性肌养蛋白之表达或活性的分子载荷。在一些实施方案中,复合物包含基于寡核苷酸的分子载荷,其通过框内外显子跳读机制或抑制终止密码子,例如通过促进DMD外显子51的跳读,来促进功能性肌养蛋白的表达。在一些实施方案中,本文中提供的分子载荷可用于促进DMD序列中的外显子跳读,例如DMD外显子51的跳读。因此,在一些实施方案中,本文中提供的复合物包含与肌细胞表面上的受体特异性结合以用于将分子载荷递送至肌细胞的肌肉靶向剂(例如,肌肉靶向抗体)。在一些实施方案中,复合物通过受体介导的内化被摄取到细胞中,然后分子载荷可被释放以在细胞内部执行功能。例如,经改造以递送寡核苷酸的复合物可释放寡核苷酸,使得该寡核苷酸可促进肌细胞中功能性肌养蛋白的表达(例如,通过外显子跳读机制,例如通过促进DMD外显子51的跳读)。在一些实施方案中,通过对连接复合物的寡核苷酸和肌肉靶向剂的共价接头进行内体切割而释放该寡核苷酸。本文中提供的复合物和分子载荷可用于治疗具有突变DMD基因(例如适合于外显子51

跳读的突变DMD基因)的对象。

[0008] 根据一些方面,本文中提供了复合物,其包含与寡核苷酸共价连接的抗转铁蛋白受体1(transferrin receptor 1,TfR1)抗体,所述寡核苷酸被配置成用于诱导DMD前体mRNA(DMD pre-mRNA)中外显子51的跳读,其中所述寡核苷酸包含与SEQ ID NO:160至383中任一者的至少8个连续核苷酸互补的互补区。

[0009] 在一些实施方案中,所述抗TfR1抗体包含:

[0010] (i)SEQ ID NO:33的重链互补决定区1(heavy chain complementarity determining region 1,CDR-H1)、SEQ ID NO:34的重链互补决定区2(heavy chain complementarity determining region 2,CDR-H2)、SEQ ID NO:35的重链互补决定区3(heavy chain complementarity determining region 3,CDR-H3)、SEQ ID NO:36的轻链互补决定区1(light chain complementarity determining region 1,CDR-L1)、SEQ ID NO:37的轻链互补决定区2(light chain complementarity determining region 2,CDR-L2)和SEQ ID NO:32的轻链互补决定区3(light chain complementarity determining region 3,CDR-L3);

[0011] (ii)SEQ ID NO:7的CDR-H1、SEQ ID NO:8的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3;

[0012] (iii)SEQ ID NO:7的CDR-H1、SEQ ID NO:20的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3;

[0013] (iv)SEQ ID NO:7的CDR-H1、SEQ ID NO:24的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3;

[0014] (v)SEQ ID NO:51的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3;

[0015] (vi)SEQ ID NO:64的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3;或者

[0016] (vii)SEQ ID NO:67的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3。

[0017] 在一些实施方案中,所述抗TfR1抗体包含:

[0018] (i)含有与SEQ ID NO:76具有至少85%同一性的氨基酸序列的重链可变区(heavy chain variable region,VH);和/或含有与SEQ ID NO:75具有至少85%同一性的氨基酸序列的轻链可变区(light chain variable region,VL);

[0019] (ii)含有与SEQ ID NO:69具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85%同一性的氨基酸序列的VL;

[0020] (iii)含有与SEQ ID NO:71具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85%同一性的氨基酸序列的VL;

[0021] (iv)含有与SEQ ID NO:72具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85%同一性的氨基酸序列的VL;

[0022] (v)含有与SEQ ID NO:73具有至少85%同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:74具有至少85%同一性的氨基酸序列的VL;

[0023] (vi)含有与SEQ ID NO:73具有至少85%同一性的氨基酸序列的VH;和/或含有与

SEQ ID NO:75具有至少85％同一性的氨基酸序列的VL;

[0024] (vii)含有与SEQ ID NO:76具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:74具有至少85％同一性的氨基酸序列的VL;

[0025] (viii)含有与SEQ ID NO:77具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:78具有至少85％同一性的氨基酸序列的VL;

[0026] (ix)含有与SEQ ID NO:79具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:80具有至少85％同一性的氨基酸序列的VL;或者

[0027] (x)含有与SEQ ID NO:77具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:80具有至少85％同一性的氨基酸序列的VL。

[0028] 在一些实施方案中,所述抗TfR1抗体包含:

[0029] (i)含有SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL;

[0030] (ii)含有SEQ ID NO:69的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

[0031] (iii)含有SEQ ID NO:71的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

[0032] (iv)含有SEQ ID NO:72的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

[0033] (v)含有SEQ ID NO:73的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL;

[0034] (vi)含有SEQ ID NO:73的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL;

[0035] (vii)含有SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL;

[0036] (viii)含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:78的氨基酸序列的VL;

[0037] (ix)含有SEQ ID NO:79的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL;或者

[0038] (x)含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL。

[0039] 在一些实施方案中,所述抗TfR1抗体是Fab片段、Fab'片段、F(ab')2片段、scFv、Fv或全长IgG。

[0040] 在一些实施方案中,所述抗TfR1抗体是Fab片段。

[0041] 在一些实施方案中,所述抗TfR1抗体包含:

[0042] (i)含有与SEQ ID NO:101具有至少85％同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:90具有至少85％同一性的氨基酸序列的轻链;

[0043] (ii)含有与SEQ ID NO:97具有至少85％同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:85具有至少85％同一性的氨基酸序列的轻链;

[0044] (iii)含有与SEQ ID NO:98具有至少85％同一性的氨基酸序列的重链;和/或含有

与SEQ ID NO:85具有至少85%同一性的氨基酸序列的轻链;

[0045] (iv)含有与SEQ ID NO:99具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:85具有至少85%同一性的氨基酸序列的轻链;

[0046] (v)含有与SEQ ID NO:100具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:89具有至少85%同一性的氨基酸序列的轻链;

[0047] (vi)含有与SEQ ID NO:100具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:90具有至少85%同一性的氨基酸序列的轻链;

[0048] (vii)含有与SEQ ID NO:101具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:89具有至少85%同一性的氨基酸序列的轻链;

[0049] (viii)含有与SEQ ID NO:102具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:93具有至少85%同一性的氨基酸序列的轻链;

[0050] (ix)含有与SEQ ID NO:103具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:95具有至少85%同一性的氨基酸序列的轻链;或者

[0051] (x)含有与SEQ ID NO:102具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:95具有至少85%同一性的氨基酸序列的轻链。

[0052] 在一些实施方案中,所述抗TfR1抗体包含:

[0053] (i)含有SEQ ID NO:101的氨基酸序列的重链;和含有SEQ ID NO:90的氨基酸序列的轻链;

[0054] (ii)含有SEQ ID NO:97的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

[0055] (iii)含有SEQ ID NO:98的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

[0056] (iv)含有SEQ ID NO:99的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

[0057] (v)含有SEQ ID NO:100的氨基酸序列的重链;和含有SEQ ID NO:89的氨基酸序列的轻链;

[0058] (vi)含有SEQ ID NO:100的氨基酸序列的重链;和含有SEQ ID NO:90的氨基酸序列的轻链;

[0059] (vii)含有SEQ ID NO:101的氨基酸序列的重链;和含有SEQ ID NO:89的氨基酸序列的轻链;

[0060] (viii)含有SEQ ID NO:102的氨基酸序列的重链;和含有SEQ ID NO:93的氨基酸序列的轻链;

[0061] (ix)含有SEQ ID NO:103的氨基酸序列的重链;和含有SEQ ID NO:95的氨基酸序列的轻链;或者

[0062] (x)含有SEQ ID NO:102的氨基酸序列的重链;和含有SEQ ID NO:95的氨基酸序列的轻链。

[0063] 在一些实施方案中,所述抗TfR1抗体不与转铁蛋白受体1的转铁蛋白结合位点特异性结合,和/或所述抗TfR1抗体不抑制转铁蛋白与转铁蛋白受体1的结合。

[0064] 在一些实施方案中,所述寡核苷酸包含所述DMD前体mRNA的剪接特征的至少4个连

续核苷酸的互补区。

[0065] 在一些实施方案中,所述剪接特征是所述DMD前体mRNA的外显子51中的外显子剪接增强子(exonic splicing enhancer,ESE),任选地其中所述ESE包含SEQ ID NO:860至894中任一者的序列。

[0066] 在一些实施方案中,所述剪接特征是分支点、剪接供体位点或剪接受体位点,任选地其中所述剪接特征跨越所述DMD前体mRNA的外显子50与内含子50的连接处、在其内含子50中、跨越其内含子50与外显子51的连接处、跨越其外显子51与内含子51的连接处、在其内含子51中、或跨越其内含子51与外显子52的连接处,并且还任选地其中所述剪接特征包含SEQ ID NO:855至859和895至898中任一者的序列。

[0067] 在一些实施方案中,所述寡核苷酸包含SEQ ID NO:160至383中任一者的互补序列或包含SEQ ID NO:384至831中任一者的序列,其中每个胸腺嘧啶碱基(T)可独立且任选地用尿嘧啶碱基(U)替换,并且每个U可独立且任选地用T替换。

[0068] 在一些实施方案中,所述寡核苷酸包含一个或更多个磷酸二酰胺吗啉代,任选地其中所述寡核苷酸是磷酸二酰胺吗啉代寡聚物(phosphorodiamidate morpholino oligomer,PMO)。

[0069] 在一些实施方案中,所述抗TfR1抗体通过可切割接头与所述寡核苷酸共价连接,任选地其中所述可切割接头包含缬氨酸-瓜氨酸序列。

[0070] 在一些实施方案中,所述抗TfR1抗体通过与所述抗体的赖氨酸残基或半胱氨酸残基的缀合而与所述寡核苷酸共价连接。

[0071] 根据一些方面,本文中提供了靶向DMD的寡核苷酸,其中所述寡核苷酸包含SEQ ID NO:160至383中任一者的互补区,任选地其中所述互补区包含与SEQ ID NO:160至383中任一者互补的至少15个连续核苷。

[0072] 在一些实施方案中,所述寡核苷酸包含SEQ ID NO:384至831中任一者的至少15个连续核苷,任选地其中所述寡核苷酸包含SEQ ID NO:384至831中任一者的序列,其中每个胸腺嘧啶碱基(T)可独立且任选地用尿嘧啶碱基(U)替换,并且每个U可独立且任选地用T替换。

[0073] 根据一些方面,本文中提供了将寡核苷酸递送至细胞的方法,该方法包括使细胞与本文中公开的复合物或本文中公开的寡核苷酸接触。

[0074] 根据一些方面,本文中提供了促进细胞中肌养蛋白的表达或活性的方法,该方法包括使细胞与有效促进寡核苷酸内化至所述细胞的量的本文中公开的复合物或本文中公开的寡核苷酸接触,任选地其中所述细胞是肌细胞。

[0075] 在一些实施方案中,对象具有适合于外显子51跳读的DMD基因。

[0076] 在一些实施方案中,DMD蛋白是截短的DMD蛋白。

附图说明

[0077] 图1示出了这样的数据,其表明在迪谢内肌营养不良患者肌管中,包含与DMD外显子跳读寡核苷酸缀合的抗TfR1 Fab(3M12 VH4/Vκ3)的缀合物与裸DMD外显子跳读寡核苷酸相比导致外显子跳读增强。

具体实施方式

[0078] 本公开内容的一些方面涉及这样的认识:尽管某些分子载荷(例如,寡核苷酸、肽、小分子)可在肌细胞中具有有益作用,但是已证明有效地靶向这样的细胞具有挑战。因此,如本文中所述,本公开内容提供了包含与分子载荷共价连接的肌肉靶向剂的复合物,以克服这样的挑战。在一些实施方案中,所述复合物特别可用于递送调节(例如,促进)肌养蛋白(例如,截短的肌养蛋白)或DMD(例如,突变DMD等位基因)的表达或活性的分子载荷。在一些实施方案中,本文中提供的复合物可包含促进肌养蛋白或DMD的表达和活性的寡核苷酸,例如通过促进框内外显子跳读和/或抑制过早终止密码子。例如,复合物可包含诱导DMD RNA(例如,前体mRNA)的外显子跳读的寡核苷酸,例如诱导外显子51跳读的寡核苷酸。在一些实施方案中,可使用表达一种或更多种促进肌养蛋白或DMD的正常表达和活性的蛋白质的合成核酸载荷(例如,DNA或RNA载荷)。

[0079] 迪谢内肌营养不良是X连锁的肌肉病症,其由位于Xp21上的DMD基因中的一个或更多个突变引起。肌养蛋白通常在肌膜处形成肌养蛋白相关的糖蛋白复合物(dystrophin-associated glycoprotein complex,DGC),其将肌肉肌节结构与胞外基质连接,并保护肌膜免于收缩诱导的损伤。在患有迪谢内肌营养不良的患者中,肌养蛋白通常不存在并且肌纤维通常由于机械性过度伸展而受损。DMD基因中的突变与两种类型的肌营养不良:迪谢内肌营养不良和贝克肌营养不良相关,这取决于翻译阅读框是丢失还是维持。贝克肌营养不良是迪谢内肌营养不良的一种临床上较温和的形式,并且其特征在于其特征与迪谢内肌营养不良相似。在一些实施方案中,由寡核苷酸诱导的外显子跳读(例如,使用本文中提供的复合物递送的)可用于恢复突变DMD等位基因的阅读框,导致产生具有足够功能以改善肌肉功能的截短的肌养蛋白。在一些实施方案中,这样的外显子跳读将迪谢内肌营养不良表型转化为较温和的贝克肌营养不良表型。

[0080] 下面提供了本公开内容的另一些方面,包括对限定的术语的描述。

[0081] I.定义

[0082] 施用:本文中使用的术语“施用”意指以生理和/或(例如,和)药理学上可用的方式向对象提供复合物(例如,以治疗对象中的病症)。

[0083] 大约:本文中使用的术语“大约”或“约”,当应用于一个或更多个目的值时,是指类似于陈述的参考值的值。在某些实施方案中,术语“大约”或“约”是指落入陈述的参考值的任一方向上(大于或小于)15%、14%、13%、12%、11%、10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、或更小以内的值的范围,除非另有说明或在其他情况下从上下文中明显看出(除非这样的数字超过可能值的100%)。

[0084] 抗体:本文中使用的术语“抗体”是指包含至少一个免疫球蛋白可变结构域或至少一个抗原决定簇(例如,与抗原特异性结合的互补位(paratope))的多肽。在一些实施方案中,抗体是全长抗体。在一些实施方案中,抗体是嵌合抗体。在一些实施方案中,抗体是人源化抗体。然而,在一些实施方案中,抗体是Fab片段、Fab’片段、F(ab’)2片段、Fv片段或scFv片段。在一些实施方案中,抗体是来源于骆驼科抗体的纳米抗体(nanobody)或来源于鲨鱼抗体的纳米抗体。在一些实施方案中,抗体是双抗体(diabody)。在一些实施方案中,抗体包含具有人种系序列的框架。在另一个实施方案中,抗体包含选自IgG、IgG1、IgG2、IgG2A、IgG2B、IgG2C、IgG3、IgG4、IgA1、IgA2、IgD、IgM和IgE恒定结构域的重链恒定结构域。在一些

实施方案中,抗体包含重(H)链可变区(在本文中简称为VH)和/或(例如,和)轻(L)链可变区(在本文中简称为VL)。在一些实施方案中,抗体包含恒定结构域,例如Fc区。免疫球蛋白恒定结构域是指重链或轻链恒定结构域。人IgG重链和轻链恒定结构域氨基酸序列及其功能变异是已知的。关于重链,在一些实施方案中,本文中所述的抗体的重链可以是alpha(α)、delta(Δ)、epsilon(ε)、gamma(γ)或mu(μ)重链。在一些实施方案中,本文中所述的抗体的重链可包含人alpha(α)、delta(Δ)、epsilon(ε)、gamma(γ)或mu(μ)重链。在一个具体实施方案中,本文中所述的抗体包含人γ1CH1、CH2和/或(例如,和)CH3结构域。在一些实施方案中,VH结构域的氨基酸序列包含人gamma(γ)重链恒定区的氨基酸序列,例如本领域已知的任何。已在本领域中描述了人恒定区序列的一些非限制性实例,例如,参见美国专利No.5,693,780和Kabat E A et al.,(1991)同上。在一些实施方案中,VH结构域包含与本文中提供的任何可变链恒定区具有至少70%、75%、80%、85%、90%、95%、98%或至少99%同一性的氨基酸序列。在一些实施方案中,对抗体进行修饰,例如,通过糖基化、磷酸化、SUMO化(sumoylation)和/或(例如,和)甲基化进行修饰。在一些实施方案中,抗体是与一个或更多个糖或碳水化合物分子缀合的糖基化抗体。在一些实施方案中,一个或更多个糖或碳水化合物分子通过N-糖基化、O-糖基化、C-糖基化、糖基磷脂酰肌醇化(GPI锚定附着)和/或(例如,和)磷酸糖基化(phosphoglycosylation)与抗体缀合。在一些实施方案中,一个或更多个糖或碳水化合物分子是单糖、二糖、寡糖或聚糖。在一些实施方案中,一个或更多个糖或碳水化合物分子是支化的寡糖或支化的聚糖。在一些实施方案中,一个或更多个糖或碳水化合物分子包含甘露糖单元、葡萄糖单元、N-乙酰葡糖胺单元、N-乙酰半乳糖胺单元、半乳糖单元、岩藻糖单元或磷脂单元。在一些实施方案中,抗体是包含多肽的构建体,所述多肽包含与接头多肽或免疫球蛋白恒定结构域连接的一个或更多个本公开内容的抗原结合片段。接头多肽包含通过肽键连接的两个或更多个氨基酸残基,并且用于连接一个或更多个抗原结合部分。已报道了接头多肽的一些实例(参见,例如,Holliger,P.,et al.(1993)Proc.Natl.Acad.Sci.USA90:6444-6448;Poljak,R.J.,et al.(1994)Structure 2:1121-1123)。另外,抗体可以是更大的免疫黏附分子的一部分,所述免疫黏附分子通过抗体或抗体部分与一个或更多个其他蛋白质或肽的共价或非共价缔合而形成。这样的免疫黏附分子的一些实例包括使用链霉亲和素核芯区域来制备四聚体scFv分子(Kipriyanov,S.M.,et al.(1995)Human Antibodies and Hybridomas6:93-101),以及使用半胱氨酸残基、标记肽和C端多组氨酸标签来制备二价和生物素化的scFv分子(Kipriyanov,S.M.,et al.(1994)Mol.Immunol.31:1047-1058)。

[0085] 分支点:本文中使用的术语“分支点”或“分支位点”是指基因或前体mRNA的内含子内的核酸序列基序,其参与将前体mRNA剪接成mRNA(即从前体mRNA中去除内含子),并且可称为剪接特征。分支点通常位于从内含子的3’末端起的18至40个核苷酸处,并且包含腺嘌呤,但在其他方面上序列相对不受限制。分支点的常见序列基序是YNYYRAY、YTRAC和YNYTRAY,其中Y是嘧啶,N是任意核苷酸,R是任意嘌呤,并且A是腺嘌呤。在剪接期间,前体mRNA在内含子的5’末端处被切割,其然后通过鸟嘌呤和腺嘌呤之间的酯交换键合分别从5’末端和分支点与下游的分支点区连接,以形成环状套索结构。

[0086] CDR:本文中使用的术语“CDR”是指抗体可变序列内的互补决定区。典型的抗体分子包含重链可变区(VH)和轻链可变区(VL),其通常参与抗原结合。VH区和VL区可进一步细

分为高变区,也称为“互补决定区”(“complementarity determining region,CDR”),其中散布有更保守的称为“框架区”(“framework region,FR”)的区域。每个VH和VL通常由三个CDR和四个FR构成,从氨基端到羧基端按以下顺序排列:FR1、CDR1、FR2、CDR2、FR3、CDR3、FR4。可使用本领域已知的方法,例如通过Kabat定义、IMGT定义、Chothia定义、AbM定义和/或(例如,和)接触定义(所有这些都是本领域公知的)来精确鉴定框架区和CDR的范围。参见,例如Kabat,E.A.,et al.(1991)Sequences of Proteins of Immunological Interest,Fifth Edition,U.S.Department of Health and Human Services,NIH Publication No.91-3242;IMGT®,the international ImMunoGeneTics information system®www.imgt.org,Lefranc,M.-P.et al.,Nucleic Acids Res.,27:209-212(1999);Ruiz,M.et al.,Nucleic Acids Res.,28:219-221(2000);Lefranc,M.-P.,Nucleic Acids Res.,29:207-209(2001);Lefranc,M.-P.,Nucleic Acids Res.,31:307-310(2003);Lefranc,M.-P.et al.,In Silico Biol.,5,0006(2004)[Epub],5:45-60(2005);Lefranc,M.-P.et al.,Nucleic Acids Res.,33:D593-597(2005);Lefranc,M.-P.et al.,Nucleic Acids Res.,37:D1006-1012(2009);Lefranc,M.-P.et al.,Nucleic Acids Res.,43:D413-422(2015);Chothia et al.,(1989)Nature 342:877;Chothia,C.et al.(1987)J.Mol.Biol.196:901-917;Al-lazikani et al(1997)J.Molec.Biol.273:927-948;以及Almagro,J.Mol.Recognit.17:132-143(2004)。还参见bioinf.org.uk/abs。本文中使用的CDR可以是指由本领域已知的任何方法定义的CDR。具有相同CDR的两种抗体意指这两种抗体的该CDR的氨基酸序列相同,如通过相同的方法(例如IMGT定义)确定的。

[0087] 重链和轻链的每个可变区中有三个CDR,对于每个可变区分别称为CDR1、CDR2和CDR3。本文中使用的术语“CDR组”是指出现在单个可变区内的能够结合抗原的三个CDR的组。这些CDR的确切边界已根据不同的系统进行了不同的定义。Kabat描述的系统(Kabat et al.,Sequences of Proteins of Immunological Interest(National Institutes of Health,Bethesda,Md.(1987)和(1991))不仅提供了适用于抗体的任何可变区的明确的残基编号系统,而且还提供了定义三个CDR的精确残基边界。这些CDR可被称为Kabat CDR。CDR的子部分可被指定为L1、L2和L3或H1、H2和H3,其中“L”和“H”分别指定轻链和重链区域。这些区域可称为Chothia CDR,其具有与Kabat CDR重叠的边界。Padlan(FASEB J.9:133-139(1995))和MacCallum(J Mol Biol 262(5):732-45(1996))已经描述了定义与Kabat CDR重叠的CDR的其他边界。另一些CDR边界定义可能并不严格遵循上述系统之一,但仍与Kabat CDR重叠,尽管可根据预测或者根据特定残基或残基的组或甚至整个CDR不会显著影响抗原结合的实验发现来缩短或延长它们。本文中使用的方法可利用根据这些系统中的任何一个定义的CDR。表1中提供了CDR定义系统的一些实例。

[0088] 表1.CDR定义

[0089]

	IMGT[1]	Kabat[2]	Chothia[3]
CDR-H1	27-38	31-35	26-32
CDR-H2	56-65	50-65	53-55
CDR-H3	105-116/117	95-102	96-101
CDR-L1	27-38	24-34	26-32

CDR-L2	56-65	50-56	50-52
CDR-L3	105-116/117	89-97	91-96

[0090] [1] IMGT®, the international ImMunoGeneTics information system®, imgt.org, Leffanc,M.-P.et al.,Nucleic Acids Res.,27:209-212(1999)

[0091] [2]Kabat et al.(1991)Sequences of Proteins of Immunological Interest, Fifth Edition,U.S.Department of Health and Human Services,NIH Publication No.91-3242

[0092] [3]Chothia et al.,J.Mol.Biol.196:901-917(1987))

[0093] CDR接枝抗体(CDR-grafted antibody):术语“CDR接枝抗体”是指包含来自一个物种的重链和轻链可变区序列但是其中VH和/或(例如,和)VL的一个或更多个CDR区的序列被另一物种的CDR序列替代的抗体,例如具有鼠重链和轻链可变区并且其中一个或更多个鼠CDR(例如,CDR3)已被人CDR序列替代的抗体。

[0094] 嵌合抗体:术语“嵌合抗体”是指包含来自一个物种的重链和轻链可变区序列和来自另一物种的恒定区序列的抗体,例如具有与人恒定区连接的鼠重链和轻链可变区的抗体。

[0095] 互补:本文中使用的术语“互补”是指在两个核苷或两组核苷之间精确配对的能力。特别地,互补是表征氢键配对引起两个核苷或两组核苷之间结合的程度的术语。例如,如果寡核苷酸的一个位置处的碱基能够与靶核酸(例如,mRNA)的相应位置处的碱基进行氢键合,则认为在该位置处碱基彼此互补。碱基配对可包括规范的沃森-克里克碱基配对和非沃森-克里克碱基配对(例如,Wobble碱基配对和Hoogsteen碱基配对)二者。例如,在一些实施方案中,对于互补碱基配对,腺苷型碱基(A)与胸苷型碱基(T)或尿嘧啶型碱基(U)互补,胞嘧啶型碱基(C)与鸟苷型碱基(G)互补,并且通用碱基例如3-硝基吡咯或5-硝基吲哚可与任何A、C、U或T杂交并被认为是互补的。肌苷(I)在本领域中也被认为是通用碱基,并且被认为与任何A、C、U或T互补。

[0096] 保守氨基酸替换:本文中使用的“保守氨基酸替换”是指不改变进行氨基酸替换的蛋白质的相对电荷或尺寸特征的氨基酸替换。可以根据本领域普通技术人员已知的用于改变多肽序列的方法来制备变体,所述方法例如可见于汇编这样的方法的参考文献中:例如Molecular Cloning:A Laboratory Manual,J.Sambrook,et al.,eds.,Fourth Edition, Cold Spring Harbor Laboratory Press,Cold Spring Harbor,New York,2012,或Current Protocols in Molecular Biology,F.M.Ausubel,et al.,eds.,John Wiley& Sons,Inc.,New York。氨基酸的保守替换包括在以下组内的氨基酸之间进行的替换:(a)M、I、L、V;(b)F、Y、W;(c)K、R、H;(d)A、G;(e)S、T;(f)Q、N;和(g)E、D。

[0097] 共价连接:本文中使用的术语“共价连接”是指两个或更多个分子通过至少一个共价键连接在一起的特征。在一些实施方案中,两个分子可以通过充当分子之间的接头的单键例如二硫键或二硫桥共价连接在一起。然而,在一些实施方案中,两个或更多个分子可以通过充当接头的分子共价连接在一起,该接头通过多个共价键将两个或更多个分子连接在一起。在一些实施方案中,接头可以是可切割接头。然而,在一些实施方案中,接头可以是不可切割接头。

[0098] 交叉反应性:如本文中使用以及在靶向剂(例如,抗体)的情况下,术语“交叉反应

性”是指物质能够以相似亲和力或亲合力与相似类型或类别的超过一种抗原(例如,多个同源物、旁系同源物或直系同源物的抗原)特异性结合的性质。例如,在一些实施方案中,对相似类型或类别的人和非人灵长类抗原(例如,人转铁蛋白受体和非人灵长类转铁蛋白受体)具有交叉反应性的抗体能够以相似亲和力或亲合力与人抗原和非人灵长类抗原结合。在一些实施方案中,抗体对相似类型或类别的人抗原和啮齿动物抗原具有交叉反应性。在一些实施方案中,抗体对相似类型或类别的啮齿动物抗原和非人灵长类抗原具有交叉反应性。在一些实施方案中,抗体对相似类型或类别的人抗原、非人灵长类抗原和啮齿动物抗原具有交叉反应性。

[0099] DMD:本文中使用的术语“DMD”是指编码肌养蛋白的基因,肌养蛋白是在肌细胞中,特别是在肌纤维中桥接内部细胞骨架和胞外基质的肌养蛋白-糖蛋白复合物的关键组分。DMD中的缺失、重复和点突变可导致肌养蛋白病,例如迪谢内肌营养不良、贝克肌营养不良或心肌病。替代启动子使用和选择性剪接导致该基因的许多独特的转录物变体和蛋白质同种型。在一些实施方案中,肌养蛋白基因(DMD或DMD基因)可以是人基因(基因ID:1756)、非人灵长类基因(例如,基因ID:465559)或啮齿动物基因(例如,基因ID:13405;基因ID:24907)。另外,已经表征了编码不同蛋白质同种型的多种人转录物变体(例如,如以以下GenBank RefSeq登录号进行注释:NM_000109.3、NM_004006.2、NM_004009.3、NM_004010.3和NM_004011.3)。

[0100] DMD等位基因:本文中使用的术语“DMD等位基因”是指DMD基因的替代形式(例如,野生型或突变体形式)中的任一种。在一些实施方案中,DMD等位基因可编码保留其正常和典型功能的肌养蛋白。在一些实施方案中,DMD等位基因可包含一种或更多种导致肌营养不良的突变。导致迪谢内肌营养不良的常见突变涉及肌养蛋白等位基因中存在的79个外显子中的一个或更多个(例如外显子8、外显子23、外显子41、外显子44、外显子45、外显子50、外显子51、外显子52、外显子53或外显子55)的移码、缺失、替换和重复突变。DMD突变的另外的实例公开在例如Flanigan KM,et al.,Mutational spectrum of DMD mutations in dystrophinopathy patients:application of modern diagnostic techniques to a large cohort.Hum Mutat.2009Dec;30(12):1657-66中,其内容通过引用整体并入本文。

[0101] 肌养蛋白病:本文中使用的术语“肌养蛋白病”是指由一种或更多种突变DMD等位基因引起的肌肉疾病。肌养蛋白病包括疾病(从轻到严重)的谱系,其包括迪谢内肌营养不良、贝克肌营养不良和DMD相关扩张型心肌病(DMD-associated dilated cardiomyopathy,DCM)。在一些实施方案中,在谱系的一端,肌养蛋白病在表型上与肌酸磷酸激酶(CK)的血清浓度的无症状增加和/或(例如,和)伴随有肌红蛋白尿的肌肉痉挛相关。在一些实施方案中,在谱系的另一端,肌养蛋白病在表型上与进行性肌肉疾病相关,该进行性肌肉疾病当骨骼肌最初受影响时通常归类为迪谢内或贝克肌营养不良,并且当心脏最初受影响时被归类为DMD相关扩张型心肌病(DCM)。迪谢内肌营养不良的症状包括肌肉损失或变性、肌肉功能减弱、舌和小腿肌肉的假性肥大、神经系统异常的高风险和寿命缩短。迪谢内肌营养不良与在线人类孟德尔遗传(Online Mendelian Inheritance in Man)(OMIM)Entry#310200相关。贝克肌营养不良与OMIM Entry#300376相关。扩张型心肌病与OMIM Entry X#302045相关。

[0102] 外显子剪接增强子(ESE):本文中使用的术语“外显子剪接增强子”或“ESE”是指基

因、前体mRNA或mRNA的外显子内指导或增强将前体mRNA剪接为mRNA的核酸序列基序,例如如Blencowe et al.,Trends Biochem Sci 25,106-10.(2000)中所述,所述文献通过引用并入本文。ESE可称为剪接特征。ESE可指导或增强剪接例如以从基因转录物中去除一个或更多个内含子和/或一个或更多个外显子。ESE基序的长度通常为6至8个核碱基。SR蛋白(例如,由基因SRSF1、SRSF2、SRSF3、SRSF4、SRSF5、SRSF6、SRSF7、SRSF8、SRSF9、SRSF10、SRSF11、SRSF12、TRA2A或TRA2B编码的蛋白质)通过其RNA识别基序区与ESE结合以促进剪接。ESE基序可通过多种方法鉴定,所述方法包括Cartegni et al.,Nucleic Acids Research,2003,Vol.31,No.13,3568-3571中所述的那些,所述文献通过引用并入本文。

[0103] 框架:本文中使用的术语“框架”或“框架序列”是指可变区减去CDR的剩余序列。由于CDR序列的确切定义可通过不同的系统确定,因此框架序列的含义相应地具有不同解释。六个CDR(轻链的CDR-L1、CDR-L2和CDR-L3和重链的CDR-H1、CDR-H2和CDR-H3)也将轻链和重链上的框架区分为每条链上的四个子区域(FR1、FR2、FR3和FR4),其中CDR1位于FR1和FR2之间,CDR2位于FR2和FR3之间,并且CDR3位于FR3和FR4之间。在未将特定子区域指定为FR1、FR2、FR3或FR4的情况下,其他人提到的框架区代表单个天然存在的免疫球蛋白链的可变区内的组合的FR。本文中使用的FR代表四个子区域之一,并且FRs代表构成框架区的四个子区域中的两个或更多个。人重链和轻链接受体序列是本领域已知的。在一个实施方案中,本领域已知的接受体序列可用于本文中公开的抗体中。

[0104] 人抗体:本文中使用的术语“人抗体”旨在包括具有来源于人种系免疫球蛋白序列的可变区和恒定区的抗体。本公开内容的人抗体可包含不是由人种系免疫球蛋白序列编码的氨基酸残基(例如,通过体外随机或位点特异性诱变或通过体内体细胞突变引入的突变),例如在CDR中,并且特别是在CDR3中。然而,本文中使用的术语“人抗体”不旨在包括其中来源于另一哺乳动物物种(例如小鼠)种系的CDR序列已接枝到人框架序列上的抗体。

[0105] 人源化抗体:术语“人源化抗体”是指包含来自非人物种(例如,小鼠)的重链和轻链可变区序列但是其中VH和/或(例如,和)VL序列的至少一部分已被改变为更加“人样”(即,更类似于人种系可变序列)的抗体。一种类型的人源化抗体是CDR接枝抗体,其中人CDR序列被引入非人VH和VL序列中以替代相应的非人CDR序列。在一个实施方案中,提供了人源化抗TfR1抗体和抗原结合部分。这样的抗体可以通过使用传统的杂交瘤技术获得鼠抗TfR1单克隆抗体随后使用体外遗传改造进行人源化来产生,例如在Kasaian et al的PCT公开No.WO 2005/123126 A2中公开的那些。

[0106] 内化细胞表面受体:本文中使用的术语“内化细胞表面受体”是指例如在外部刺激(例如,配体与受体结合)下被细胞内化的细胞表面受体。在一些实施方案中,内化细胞表面受体通过内吞作用内化。在一些实施方案中,内化细胞表面受体通过网格蛋白介导的内吞作用内化。然而,在一些实施方案中,内化细胞表面受体通过不依赖于网格蛋白的途径内化,所述途径例如如吞噬作用、巨胞饮作用、小窝和筏介导的摄取或组成型网格蛋白非依赖性内吞作用。在一些实施方案中,内化细胞表面受体包含胞内结构域、跨膜结构域和/或(例如,和)胞外结构域,其可任选地还包含配体结合结构域。在一些实施方案中,细胞表面受体在配体结合后被细胞内化。在一些实施方案中,配体可以是肌肉靶向剂或肌肉靶向抗体。在一些实施方案中,内化细胞表面受体是转铁蛋白受体。

[0107] 分离的抗体:本文中使用的“分离的抗体”旨在是指代基本上不含具有不同抗原特

异性的其他抗体的抗体(例如,特异性结合转铁蛋白受体的分离的抗体基本上不含特异性结合转铁蛋白受体以外的抗原的抗体)。然而,特异性结合转铁蛋白受体复合物的分离的抗体可能与其他抗原(例如来自其他物种的转铁蛋白受体分子)具有交叉反应性。此外,分离的抗体可基本上不含其他细胞物质和/或(例如,和)化学物质。

[0108] Kabat编号:术语"Kabat编号"、"Kabat定义和"Kabat标记"在本文中可互换使用。在本领域中公认的这些术语是指对抗体或其抗原结合部分的重链和轻链可变区中的比其他氨基酸残基更加可变(即高变)的氨基酸残基进行编号的系统(Kabat et al.(1971)Ann.NY Acad.Sci.190:382-391以及Kabat,E.A.,et al.(1991)Sequences of Proteins of Immunological Interest,Fifth Edition,U.S.Department of Health and Human Services,NIH Publication No.91-3242)。对于重链可变区,CDR1的高变区为第31至35位氨基酸,CDR2的高变区为第50至65位氨基酸,并且CDR3的高变区为第95至102位氨基酸。对于轻链可变区,CDR1的高变区为第24至34位氨基酸,CDR2的高变区为第50至56位氨基酸,并且CDR3的高变区为第89至97位氨基酸。

[0109] 分子载荷:本文中使用的术语"分子载荷"是指发挥调节生物学结局作用的分子或物质。在一些实施方案中,分子载荷与肌肉靶向剂连接或以其他方式缔合。在一些实施方案中,分子载荷是小分子、蛋白质、肽、核酸或寡核苷酸。在一些实施方案中,分子载荷发挥调节DNA序列的转录、调节蛋白质的表达或调节蛋白质的活性的作用。在一些实施方案中,分子载荷是寡核苷酸,其包含具有靶基因的互补区的链。

[0110] 肌肉靶向剂:本文中使用的术语"肌肉靶向剂"是指与肌细胞上表达的抗原特异性结合的分子。肌细胞内或其上的抗原可以是膜蛋白,例如整合膜蛋白或外周膜蛋白。通常来说,肌肉靶向剂与肌细胞上的抗原特异性结合,这有助于将肌肉靶向剂(和任何缔合的分子载荷)内化到肌细胞中。在一些实施方案中,肌肉靶向剂与肌肉上的内化细胞表面受体特异性结合,并且能够通过受体介导的内化而内化到肌细胞中。在一些实施方案中,肌肉靶向剂是小分子、蛋白质、肽、核酸(例如,适配体)、或抗体。在一些实施方案中,肌肉靶向剂与分子载荷连接。

[0111] 肌肉靶向抗体:本文中使用的术语"肌肉靶向抗体"是指作为与肌细胞内或其上存在的抗原特异性结合的抗体的肌肉靶向剂。在一些实施方案中,肌肉靶向抗体与肌细胞上的抗原特异性结合,这有助于将肌肉靶向抗体(和任何缔合的分子载荷)内化到肌细胞中。在一些实施方案中,肌肉靶向抗体与存在于肌细胞上的内化细胞表面受体特异性结合。在一些实施方案中,肌肉靶向抗体是与转铁蛋白受体特异性结合的抗体。

[0112] 寡核苷酸:本文中使用的术语"寡核苷酸"是指长度长至200个核苷酸的寡聚核酸化合物。寡核苷酸的一些实例包括但不限于RNAi寡核苷酸(例如,siRNA、shRNA)、微RNA、间隔聚体、混合聚体、磷酸二酰胺吗啉代、肽核酸、适配体、指导核酸(例如,Cas9指导RNA)等。寡核苷酸可以是单链或双链。在一些实施方案中,寡核苷酸可包含一个或更多个经修饰核苷(例如,2'-O-甲基糖修饰、嘌呤或嘧啶修饰)。在一些实施方案中,寡核苷酸可包含一个或更多个经修饰核苷间键联。在一些实施方案中,寡核苷酸可包含一个或更多个硫代磷酸酯键联,其可为Rp或Sp立体化学构型。

[0113] 重组抗体:本文中使用的术语"重组人抗体"旨在包括通过重组方式制备、表达、产生或分离的所有人抗体,例如使用转染到宿主细胞中的重组表达载体表达的抗体(在本公

开内容中更详细地描述),从重组、组合人抗体文库分离的抗体(Hoogenboom H.R.,(1997) TIB Tech.15:62-70;Azzazy H.,and Highsmith W.E.,(2002)Clin.Biochem.35:425-445; Gavilondo J.V.,and Larrick J.W.(2002)BioTechniques 29:128-145;Hoogenboom H., and Chames P.(2000)Immunology Today 21:371-378),从人免疫球蛋白基因转基因的动物(例如,小鼠)分离的抗体(参见例如Taylor,L.D.,et al.(1992)Nucl.Acids Res.20: 6287-6295;Kellermann S-A.,and Green L.L.(2002)Current Opinion in Biotechnology 13:593-597;Little M.et al(2000)Immunology Today 21:364-370),或通过涉及将人免疫球蛋白基因序列剪接至其他DNA序列的任何其他方式制备、表达、产生或分离的抗体。这样的重组人抗体具有来源于人种系免疫球蛋白序列的可变区和恒定区。然而,在某些实施方案中,对这样的重组人抗体进行体外诱变(或当使用人Ig序列转基因的动物时,进行体内体细胞诱变),并且因此重组抗体的VH和VL区的氨基酸序列是这样的序列,尽管其来源于人种系VH和VL序列并与之相关,但可能不是体内人抗体种系库中天然存在的。本公开内容的一个实施方案提供了能够结合人转铁蛋白受体的完全人抗体,其可使用本领域公知的技术产生,例如但不限于使用人Ig噬菌体文库,例如Jermutus et al.的PCT公开No.WO 2005/007699 A2中公开的那些。

[0114] 互补区:本文中使用的术语“互补区”是指与例如靶核酸的同源核苷酸序列充分互补的例如寡核苷酸的核苷酸序列,使得两个核苷酸序列能够在生理条件下(例如,在细胞中)彼此退火。在一些实施方案中,互补区与靶核酸的同源核苷酸序列完全互补。然而,在一些实施方案中,互补区与靶核酸的同源核苷酸序列部分互补(例如,至少80%、90%、95%或99%互补)。在一些实施方案中,与靶核酸的同源核苷酸序列相比,互补区包含1、2、3或4个错配。

[0115] 特异性结合:本文中使用的术语“特异性结合”是指分子以一定程度的亲和力或亲合力与结合配偶体结合的能力,其使得分子能够用于在结合测定或其他结合环境中将结合配偶体与合适的对照区分开。关于抗体,术语“特异性结合”是指与合适的一种或更多种参考抗原相比,抗体以一定程度的亲和力或亲合力与特异性抗原结合的能力,其使得抗体能够用于将特异性抗原与其他抗原区分开,例如至允许通过与如本文中所述的抗原结合而优先靶向某些细胞(例如,肌细胞)的程度。在一些实施方案中,如果抗体与靶标结合的K_D为至少约10^{-4}M、10^{-5}M、10^{-6}M、10^{-7}M、10^{-8}M、10^{-9}M、10^{-10}M、10^{-11}M、10^{-12}M、10^{-13}M或更小,则抗体与靶标特异性结合。在一些实施方案中,抗体与转铁蛋白受体(例如,转铁蛋白受体的顶端结构域(apical domain)的表位)特异性结合。

[0116] 剪接受体位点:本文中使用的术语“剪接受体位点”或“剪接受体”是指在内含子的3’末端处或者跨越基因或前体mRNA的内含子/外显子连接处的核酸序列基序,其参与将前体mRNA剪接成mRNA(即从前体mRNA中去除内含子),并且可称为剪接特征。剪接受体位点包含在内含子的3’末端处的末端AG序列,所述末端AG序列之前(5’方向)通常是高嘧啶(C/U)区。剪接受体位点的上游是分支点。通过分支点与剪接供体位点之间的酯交换反应形成套索环中间结构,释放5’外显子的3’-OH,其随后与3’外显子的第一核苷酸反应,从而连接外显子并释放内含子套索。已知剪接受体位点中内含子3’末端处的AG序列对于正确的剪接至关重要,因为改变这些核苷酸中的一个会导致抑制剪接。罕有地,替代的剪接受体位点在内含子的3’末端处具有AC而不是更常见的AG。常见的剪接受体位点基序具有[富含Y的区域]-

NCAGG或YxNYAGG的序列或者与其相似的序列,其中Y代表嘧啶,N代表任意核苷酸,并且x是4至20的数字。切割位点在AG之后,AG表示已切除的内含子的3'端核苷酸。

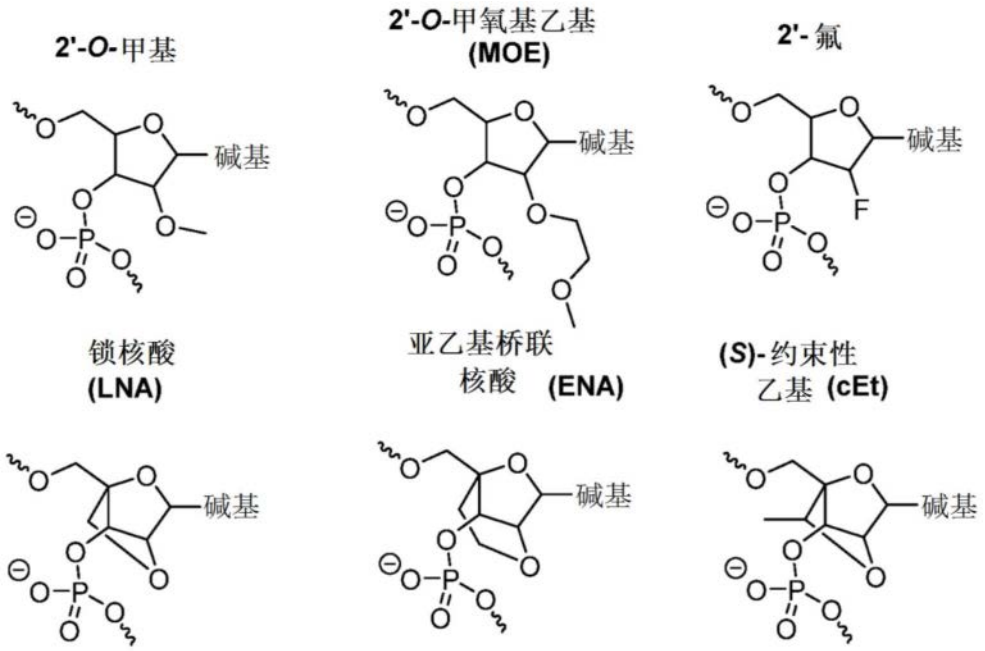
[0117] 剪接供体位点:本文中使用的术语"剪接供体位点"或"剪接供体"是指在内含子的5'末端处或者跨越基因或前体mRNA的外显子/内含子连接处的核酸序列基序,其参与将前体mRNA剪接成mRNA(即从前体mRNA中去除内含子)并且可称为剪接特征。在较大且相当不受限制的序列内,剪接供体位点包含在内含子的5'末端处的末端GU序列。在剪接期间,分支点内核苷酸的2'-OH通过对剪接供体位点内的内含子的5'G进行亲核攻击来引发酯交换反应。从而将G从前体mRNA上切割并转而与分支点核苷酸结合,形成环套索结构。上游外显子的3'核苷酸随后结合剪接受体位点,连接外显子并切除内含子。典型的剪接供体位点具有GGGURAGU或AGGURNG的序列或者与其相似的序列,其中R代表嘌呤并且N代表任意核苷酸。切割位点在第一个GU(即,GG/GURAGU或AG/GURNG)之前,所述第一个GU代表切除的内含子的5'端核苷酸。

[0118] 对象:本文中使用的术语"对象"是指哺乳动物。在一些实施方案中,对象是非人灵长类或啮齿动物。在一些实施方案中,对象是人。在一些实施方案中,对象是患者,例如患有或怀疑患有疾病的人患者。在一些实施方案中,对象是患有或怀疑患有由突变DMD基因序列(例如,DMD基因序列的外显子中的突变)引起的疾病的人患者。在一些实施方案中,对象患有肌养蛋白病,例如迪谢内肌营养不良。在一些实施方案中,对象是具有适合于外显子51跳读的DMD基因的突变的患者。

[0119] 转铁蛋白受体:本文中使用的术语"转铁蛋白受体"(也称为TFRC、CD71、p90或TFR1)是指结合转铁蛋白以促进通过内吞作用摄取铁的内化细胞表面受体。在一些实施方案中,转铁蛋白受体可以是人来源的(NCBI基因ID 7037)、非人灵长类来源的(例如,NCBI基因ID 711568或NCBI基因ID 102136007)或啮齿动物来源的(例如,NCBI基因ID 22042)。另外,已经表征了编码受体的不同同种型的多种人转录物变体(例如,如以以下GenBank RefSeq登录号注释的:NP_001121620.1、NP_003225.2、NP_001300894.1和NP_001300895.1)。

[0120] 2'-经修饰核苷:本文中使用的术语"2'-经修饰核苷"和"2'-经修饰核糖核苷"可互换使用,并且是指在2'位置具有经修饰糖部分的核苷。在一些实施方案中,2'-经修饰核苷是2'-4'双环核苷,其中糖的2'和4'位置是桥联的(例如,通过亚甲基、亚乙基或(S)-约束性乙基桥联)。在一些实施方案中,2'-经修饰核苷是非双环的2'-经修饰核苷,例如,其中糖部分的2'位置被取代。2'-经修饰核苷的一些非限制性实例包括:2'-脱氧、2'-氟(2'-F)、2'-O-甲基(2'-O-Me)、2'-O-甲氧基乙基(2'-MOE)、2'-O-氨基丙基(2'-O-AP)、2'-O-二甲基氨基乙基(2'-O-DMAOE)、2'-O-二甲基氨基丙基(2'-O-DMAP)、2'-O-二甲基氨基乙基氧基乙基(2'-O-DMAEOE)、2'-O-N-甲基乙酰胺基(2'-O-NMA)、锁核酸(locked nucleic acid,LNA,亚甲基桥联核酸)、亚乙基桥联核酸(ethylene-bridged nucleic acid,ENA)和(S)-约束性乙基桥联核酸(cEt)。在一些实施方案中,本文中所述的2'-经修饰核苷是高亲和力的经修饰核苷和寡核苷酸,其包含相对于未经修饰的寡核苷酸而对靶序列具有提高的亲和力的2'-经修饰核苷。下面提供了2'-经修饰核苷之结构的一些实例:

[0121]



[0122] 这些实例显示带有磷酸基团,但是任何核苷间键联都考虑在2'-经修饰核苷之间。

[0123] II.复合物

[0124] 本文中提供了包含与分子载荷共价连接的靶向剂(例如,抗体)的复合物。在一些实施方案中,复合物包含与寡核苷酸共价连接的肌肉靶向抗体。复合物可包含特异性结合单个抗原位点或结合可存在于相同或不同抗原上的至少两个抗原位点的抗体。

[0125] 复合物可用于调节至少一种基因、蛋白质和/或(例如,和)核酸的活性或功能。在一些实施方案中,存在于复合物内的分子载荷负责基因、蛋白质和/或(例如,和)核酸的调节。分子载荷可以是能够调节细胞中基因、蛋白质和/或(例如,和)核酸的活性或功能的小分子、蛋白质、核酸、寡核苷酸或任何分子实体。

[0126] 在一些实施方案中,复合物包含与分子载荷(例如靶向DMD以促进外显子跳读的反义寡核苷酸)共价连接的肌肉靶向剂(例如抗转铁蛋白受体抗体),例如在由突变DMD等位基因编码的转录物中。在一些实施方案中,复合物靶向DMD前体mRNA以促进DMD前体mRNA中外显子51的跳读。

[0127] A.肌肉靶向剂

[0128] 本公开内容的一些方面提供了肌肉靶向剂,例如用于将分子载荷递送至肌细胞。在一些实施方案中,这样的肌肉靶向剂能够例如通过与肌细胞上的抗原特异性结合而与肌细胞结合,并且将缔合的分子载荷递送至肌细胞。在一些实施方案中,分子载荷与肌肉靶向剂结合(例如,共价结合),并且在肌肉靶向剂与肌细胞上的抗原结合之后内化到肌细胞中,例如通过内吞作用。应理解,根据本公开内容可以使用多种类型的肌肉靶向剂。还应理解,任何肌肉靶标(例如,肌肉表面蛋白)都可以被本文中所述的任何类型的肌肉靶向剂靶向。例如,肌肉靶向剂可包含核酸(例如,DNA或RNA)、肽(例如,抗体)、脂质(例如,微泡(microvesicle))或糖部分(例如,多糖),或者由其组成。肌肉靶向剂可包含小分子或由小分子组成。示例性的肌肉靶向剂在本文中进一步详细描述,然而,应理解,本文中提供的示例性肌肉靶向剂并不意味着是限制性的。

[0129] 本公开内容的一些方面提供了与肌肉(例如骨骼肌、平滑肌或心肌)上的抗原特异

性结合的肌肉靶向剂。在一些实施方案中,本文中提供的任何肌肉靶向剂均与骨骼肌细胞、平滑肌细胞和/或(例如,和)心肌细胞上的抗原结合(例如,与之特异性结合)。

[0130] 通过与肌肉特异性细胞表面识别元件(例如,细胞膜蛋白)相互作用,可实现组织定位和选择性摄取到肌细胞中二者。在一些实施方案中,作为肌肉摄取转运体之底物的分子可用于将分子载荷递送到肌肉组织中。与肌肉表面识别元件结合之后是内吞作用,其可允许甚至大分子(例如,抗体)进入肌细胞。作为另一个实例,与转铁蛋白或抗TfR1抗体缀合的分子载荷可通过与转铁蛋白受体结合而被肌细胞摄取,然后可例如通过网格蛋白介导的内吞作用被内吞。

[0131] 肌肉靶向剂的使用可用于将分子载荷(例如,寡核苷酸)集中在肌肉中,同时降低与其他组织中的作用相关的毒性。在一些实施方案中,与对象内的另一种细胞类型相比,肌肉靶向剂将结合的分子载荷集中在肌细胞中。在一些实施方案中,肌肉靶向剂将结合的分子载荷以是非肌细胞(例如,肝、神经元、血液或脂肪细胞)中的量的至少1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90或100倍高的量集中在肌细胞(例如,骨骼肌、平滑肌或心肌细胞)中。在一些实施方案中,当分子载荷在与肌肉靶向剂结合时递送至对象时,其在对象中的毒性降低至少1%、2%、3%、4%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、90%或95%。

[0132] 在一些实施方案中,为了实现肌肉选择性,可能需要肌肉识别元件(例如,肌细胞抗原)。作为一个实例,肌肉靶向剂可以是作为肌肉特异性摄取转运体之底物的小分子。作为另一个实例,肌肉靶向剂可以是通过转运体介导的内吞作用进入肌细胞的抗体。作为另一个实例,肌肉靶向剂可以是与肌细胞上的细胞表面受体结合的配体。应理解,尽管基于转运体的方法为细胞进入提供了直接途径,但是基于受体的靶向可能涉及刺激的内吞作用以达到期望的作用部位。

[0133] i.肌肉靶向抗体

[0134] 在一些实施方案中,肌肉靶向剂是抗体。通常来说,抗体对其靶抗原的高特异性提供了用于选择性靶向肌细胞(例如,骨骼肌、平滑肌和/或(例如,和)心肌细胞)的潜力。这种特异性也可以限制脱靶毒性。已经报道了能够靶向肌细胞表面抗原的抗体的一些实例并且在本公开内容的范围内。例如,靶向肌细胞表面的抗体在以下中有描述:Arahata K.,et al.“Immunostaining of skeletal and cardiac muscle surface membrane with antibody against Duchenne muscular dystrophy peptide”Nature 1988;333:861-3;Song K.S.,et al.“Expression of caveolin-3in skeletal,cardiac,and smooth muscle cells.Caveolin-3is a component of the sarcolemma and co-fractionates with dystrophin and dystrophin-associated glycoproteins”J Biol Chem 1996;271:15160-5;以及Weisbart R.H.et al.,“Cell type specific targeted intracellular delivery into muscle of a monoclonal antibody that binds myosin IIb”Mol Immunol.2003Mar,39(13):78309;其各自的全部内容均通过引用并入本文。

[0135] a.抗转铁蛋白受体(TfR)抗体

[0136] 本公开内容的一些方面是基于这样的认识:与转铁蛋白受体结合的物质(例如,抗转铁蛋白受体抗体)能够靶向肌细胞。转铁蛋白受体是内化细胞表面受体,其通过细胞膜转运转铁蛋白并参与胞内铁水平的调节和稳态。本公开内容的一些方面提供了能够与转铁蛋

白受体结合的转铁蛋白受体结合蛋白。因此,本公开内容的一些方面提供了与转铁蛋白受体结合的结合蛋白(例如,抗体)。在一些实施方案中,与转铁蛋白受体结合的结合蛋白与任何结合的分子载荷一起被内化到肌细胞中。本文中使用的与转铁蛋白受体结合的抗体可以可互换地称为转铁蛋白受体抗体、抗转铁蛋白受体抗体或抗TfR1抗体。与转铁蛋白受体结合(例如,特异性结合)的抗体可在与转铁蛋白受体结合之后例如通过受体介导的内吞作用而被内化到细胞中。

[0137] 应理解,可使用数种已知的方法(例如使用噬菌体展示的文库设计)来产生、合成和/或(例如,和)衍生抗TfR1抗体。示例性方法已经在本领域中表征并且通过引用并入(Díez,P.et al.“High-throughput phage-display screening in array format”,Enzyme and microbial technology,2015,79,34-41.;Christoph M.H.and Stanley,J.R.“Antibody Phage Display:Technique and Applications”J Invest Dermatol.2014,134:2.;Engleman,Edgar(Ed.)“Human Hybridomas and Monoclonal Antibodies.”1985,Springer.)。在另一些实施方案中,抗TfR1抗体先前已被表征或公开。与转铁蛋白受体特异性结合的抗体是本领域中已知的(参见,例如1979年12月4日提交的美国专利No.4,364,934,“Monoclonal antibody to a human early thymocyte antigen and methods for preparing same”;2006年6月14日提交的美国专利No.8,409,573,“Anti-CD71 monoclonal antibodies and uses thereof for treating malignant tumor cells”;2014年5月20日提交的美国专利No.9,708,406,“Anti-transferrin receptor antibodies and methods of use”;2014年12月19日提交的US 9,611,323,“Low affinity blood brain barrier receptor antibodies and uses therefor”;2014年12月24日提交的WO2015/098989,“Novel anti-Transferrin receptor antibody that passes through blood-brain barrier”;Schneider C.et al.“Structural features of the cell surface receptor for transferrin that is recognized by the monoclonal antibody OKT9.”J Biol Chem.1982,257:14,8516-8522.;Lee et al.“Targeting Rat Anti-Mouse Transferrin Receptor Monoclonal Antibodies through Blood-Brain Barrier in Mouse”2000,J Pharmacol.Exp.Ther.,292:1048-1052)。

[0138] 在一些实施方案中,本文中所述的抗TfR1抗体以高特异性和亲和力与转铁蛋白受体结合。在一些实施方案中,本文中所述的抗TfR1抗体与转铁蛋白受体的任何胞外表位或暴露于抗体的表位特异性结合。在一些实施方案中,本文中提供的抗TfR1抗体与来自人、非人灵长类、小鼠、大鼠等的转铁蛋白受体特异性结合。在一些实施方案中,本文中提供的抗TfR1抗体与人转铁蛋白受体结合。在一些实施方案中,本文中所述的抗TfR1抗体与人或非人灵长类转铁蛋白受体(如SEQ ID NO:105至108中提供的)的氨基酸区段结合。在一些实施方案中,本文中所述的抗TfR1抗体与这样的氨基酸区段结合:对应于人转铁蛋白受体(如SEQ ID NO:105中所示)的第90至96位氨基酸,其不在转铁蛋白受体的顶端结构域中。

[0139] 在一些实施方案中,本文中所述的抗TfR1抗体(例如,下表2中的抗TfR克隆8)结合TfR1中的表位,其中所述表位包含SEQ ID NO:105的第214至241位氨基酸和/或第354至381位氨基酸中的残基。在一些实施方案中,本文中所述的抗TfR1抗体结合包含SEQ ID NO:105的第214至241位氨基酸和第354至381位氨基酸中的残基的表位。在一些实施方案中,本文中所述的抗TfR1抗体结合包含如SEQ ID NO:105中所示的人TfR1的一个或更多个残基

Y222、T227、K231、H234、T367、S368、S370、T376和S378的表位。在一些实施方案中,本文中所述的抗TfR1抗体结合包含如SEQ ID NO:105中所示的人TfR1的残基Y222、T227、K231、H234、T367、S368、S370、T376和S378的表位。

[0140] 在一些实施方案中,本文中所述的抗TfR1抗体(例如,下表2中的3M12及其变体)结合TfR1中的表位,其中所述表位包含SEQ ID NO:105的第258至291位氨基酸和/或第358至381位氨基酸中的残基。在一些实施方案中,本文中所述的抗TfR1抗体(例如,下表2中的3M12及其变体)结合包含SEQ ID NO:105的第258至291位氨基酸和第358至381位氨基酸中的残基的表位。在一些实施方案中,本文中所述的抗TfR1抗体(例如,下表2中的3M12及其变体)结合包含如SEQ ID NO:105中所示的人TfR1的一个或更多个残基K261、S273、Y282、T362、S368、S370和K371的表位。在一些实施方案中,本文中所述的抗TfR1抗体(例如,下表2中的3M12及其变体)结合包含如SEQ ID NO:105中所示的人TfR1的残基K261、S273、Y282、T362、S368、S370和K371的表位。

[0141] 对应于NCBI序列NP_003225.2(转铁蛋白受体蛋白1同种型1,智人(homo sapiens))的示例性人转铁蛋白受体氨基酸序列如下:

[0142]
MMDQARSAFSNLFGGEPLSYTRFSLARQVDGDNSHVEMKLAVDEEENADNNTKANVT
KPKRCSGSICYGTIAVIVFFLIGFMIGYLGYCKGVEPKTECERLAGTESPVREEPGEDFPA
ARRLYWDDLKRKLSEKLDSTDFTGTIKLLNENSYVPREAGSQKDENLALYVENQFREF
KLSKVWRDQHFVKIQVKDSAQNSVIIVDKNGRLVYLVENPGGYVAYSKAATVTGKLV
HANFGTKKDFEDLYTPVNGSIVIVRAGKITFAEKVANAESLNAIGVLIYMDQTKFPIVNA
ELSFFGHAHLGTGDPYTPGFPSFNHTQFPPSRSSGLPNIPVQTISRAAAEKLFGNMEGDCP
SDWKTDSTCRMVTSESKNVKLTVSNVLKEIKILNIFGVIKGFVEPDHYVVVGAQRDAW
GPGAAKSGVGTALLLKLAQMFSDMVLKDGFQPSRSIIFASWSAGDFGSVGATEWLEGY
LSSLHLKAFTYINLDKAVLGTSNFKVSASPLLYTLIEKTMQNVKHPVTGQFLYQDSNWA
SKVEKLTLDNAAFPFLAYSGIPAVSFCFCEDTDYPYLGTTMDTYKELIERIPELNKVARA
AAEVAGQFVIKLTHDVELNLDYERYNSQLLSFVRDLNQYRADIKEMGLSLQWLYSARG
DFFRATSRLTTDFGNAEKTDRFVMKKLNDRVMRVEYHFLSPYVSPKESPFRHVFWGSG
SHTLPALLENLKLRKQNNGAFNETLFRNQLALATWTIQGAANALSGDVWDIDNEF
(SEQ ID NO: 105)

[0143] 对应于NCBI序列NP_001244232.1(转铁蛋白受体蛋白1,恒河猴(Macaca mulatta))的示例性非人灵长类转铁蛋白受体氨基酸序列如下:

[0144] MMDQARSAFSNLFGGEPLSYTRFSLARQVDGDNSHVEMKLGVDEEENTDNNTKPNGT
KPKRCGGNICYGTIAVIIFFLIGFMIGYLGYCKGVEPKTECERLAGTESPAREEPEEDFPA
APRLYWDDLKRKLSEKLDTTDFTSTIKLLNENLYVPREAGSQKDENLALYIENQFREFK
LSKVWRDQHFVKIQVKDSAQNSVIIVDKNGGLVYLVENPGGYVAYSKAATVTGKLVH
ANFGTKKDFEDLDSPVNGSIVIVRAGKITFAEKVANAESLNAIGVLIYMDQTKFPIVKAD
LSFFGHAHLGTGDPYTPGFPSFNHTQFPPSQSSGLPNIPVQTISRAAAEKLFGNMEGDCPS
DWKTDSTCKMVTSENKSVKLTVSNVLKETKILNIFGVIKGFVEPDHYVVVGAQRDAW
GPGAAKSSVGTALLLKLAQMFSDMVLKDGFQPSRSIIFASWSAGDFGSVGATEWLEGY
LSSLHLKAFTYINLDKAVLGTSNFKVSASPLLYTLIEKTMQDVKHPVTGRSLYQDSNWA
SKVEKLTLDNAAFPFLAYSGIPAVSFCFCEDTDYPYLGTTMDTYKELVERIPELNKVAR
AAAEVAGQFVIKLTHDTELNLDYERYNSQLLLFLRDLNQYRADVKEMGLSLQWLYSA
RGDFFRATSRLTTDFRNAEKRDKFVMKKLNDRVMRVEYYFLSPYVSPKESPFRHVFWG
SGSHTLSALLESLKLRRQNNSAFNETLFRNQLALATWTIQGAANALSGDVWDIDNEF
(SEQ ID NO: 106)

。

[0145] 对应于NCBI序列XP_005545315.1(转铁蛋白受体蛋白1,食蟹猴(Macaca fascicularis))的示例性非人灵长类转铁蛋白受体氨基酸序列如下：

[0146] MMDQARSAFSNLFGGEPLSYTRFSLARQVDGDNSHVEMKLGVDEEENTDNNTKANGT
KPKRCGGNICYGTIAVIIFFLIGFMIGYLGYCKGVEPKTECERLAGTESPAREEPEEDFPA
APRLYWDDLKRKLSEKLDTTDFTSTIKLLNENLYVPREAGSQKDENLALYIENQFREFK
LSKVWRDQHFVKIQVKDSAQNSVIIVDKNGGLVYLVENPGGYVAYSKAATVTGKLVH
ANFGTKKDFEDLDSPVNGSIVIVRAGKITFAEKVANAESLNAIGVLIYMDQTKFPIVKAD
LSFFGHAHLGTGDPYTPGFPSFNHTQFPPSQSSGLPNIPVQTISRAAAEKLFGNMEGDCPS
DWKTDSTCKMVTSENKSVKLTVSNVLKETKILNIFGVIKGFVEPDHYVVVGAQRDAW
GPGAAKSSVGTALLLKLAQMFSDMVLKDGFQPSRSIIFASWSAGDFGSVGATEWLEGY
LSSLHLKAFTYINLDKAVLGTSNFKVSASPLLYTLIEKTMQDVKHPVTGRSLYQDSNWA
SKVEKLTLDNAAFPFLAYSGIPAVSFCFCEDTDYPYLGTTMDTYKELVERIPELNKVAR
AAAEVAGQFVIKLTHDTELNLDYERYNSQLLLFLRDLNQYRADVKEMGLSLQWLYSA
RGDFFRATSRLTTDFRNAEKRDKFVMKKLNDRVMRVEYYFLSPYVSPKESPFRHVFWG
SGSHTLSALLESLKLRRQNNSAFNETLFRNQLALATWTIQGAANALSGDVWDIDNEF
(SEQ ID NO: 107)

。

[0147] 对应于NCBI序列NP_001344227.1(转铁蛋白受体蛋白1,小家鼠(mus musculus))的示例性小鼠转铁蛋白受体氨基酸序列如下：

[0148] MMDQARSAFSNLFGGEPLSYTRFSLARQVDGDNSHVEMKLAADEEENADNNMKASV
RKPKRFNGRLCFAAIALVIFFLIGFMSGYLGYCKRVEQKEECVKLAETEETDKSETMETE
DVPTSSRLYWADLKTLLSEKLNSIEFADTIKQLSQNTYTPREAGSQKDESLAYYIENQFH
EFKFSKVWRDEHYVKIQVKSSIGQNMVTIVQSNGNLDPVESPEGYVAFSKPTEVSGKLV
HANFGTKKDFEELSYSVNGSLVIVRAGEITFAEKVANAQSFNAIGVLIYMDKNKFPVVE
ADLALFGHAHLGTGDPYTPGFPSFNHTQFPPSQSSGLPNIPVQTISRAAAEKLFGKMEGS
CPARWNIDSSCKLELSQNQNVKLIVKNVLKERRILNIFGVIKGYEEPDRYVVVGAQRDA
LGAGVAAKSSVGTGLLLKLAQVFSDMISKDGFRPSRSIIFASWTAGDFGAVGATEWLEG
YLSSLHLKAFTYINLDKVVLGTSNFKVSASPLLYTLMGKIMQDVKHPVDGKSLYRDSN
WISKVEKLSFDNAAYPFLAYSGIPAVSFCFCEDADYPYLGTRLDTYEALTQKVPQLNQM
VRTAAEVAGQLIIKLTHDVELNLDYEMYNSKLLSFMKDLNQFKTDIRDMGLSLQWLYS
ARGDYFRATSRLTTDFHNAEKTNRFVMREINDRIMKVEYHFLSPYVSPRESPFRHIFWG
SGSHTLSALVENLKLRQKNITAFNETLFRNQLALATWTIQGVANALSGDIWNIDNEF
(SEQ ID NO: 108)。

[0149] 在一些实施方案中,抗TfR1抗体与如下的受体氨基酸区段结合:

[0150] FVKIQVKDSAQNSVIIVDKNGRLVYLVENPGGYVAYSKAATVTGKLVHANFGTKKDFE
DLYTPVNGSIVIVRAGKITFAEKVANAESLNAIGVLIYMDQTKFPIVNAELSFFGHAHLG
TGDPYTPGFPSFNHTQFPPSRSSGLPNIPVQTISRAAAEKLFGNMEGDCPSDWKTDSTCR
MVTSESKNVKLTVSNVLKE (SEQ ID NO: 109)

[0151] 并且不抑制转铁蛋白受体与转铁蛋白和/或(例如,和)人血色素沉着蛋白(human hemochromatosis protein,也称为HFE)之间的结合相互作用。在一些实施方案中,本文中所述的抗TfR1抗体不与SEQ ID NO:109中的表位结合。

[0152] 可使用适当的方法来获得和/或(例如,和)产生抗体、抗体片段或抗原结合剂,例如,通过使用重组DNA方案。在一些实施方案中,也可通过杂交瘤的产生来产生抗体(参见,例如,Kohler,G and Milstein,C.“Continuous cultures of fused cells secreting antibody of predefined specificity”Nature,1975,256:495-497)。目的抗原可以以任何形式或实体(例如,重组或天然存在的形式或实体)用作免疫原。使用标准方法(例如ELISA筛选)筛选杂交瘤,以发现至少一种产生靶向特定抗原之抗体的杂交瘤。也可通过筛选表达抗体的蛋白质表达文库(例如,噬菌体展示文库)来产生抗体。在一些实施方案中,也可使用噬菌体展示文库设计(参见,例如,1991年3月1日提交的美国专利No 5,223,409,“Directed evolution of novel binding proteins”;1992年4月10日提交的WO 1992/18619,“Heterodimeric receptor libraries using phagemids”;1991年5月1日提交的WO 1991/17271,“Recombinant library screening methods”;1992年5月15日提交的WO 1992/20791,“Methods for producing members of specific binding pairs”;1992年2月28日提交的WO 1992/15679,“Improved epitope displaying phage”)。在一些实施方案中,目的抗原可用于对非人动物,例如啮齿动物或山羊进行免疫接种。在一些实施方案中,然后从非人动物获得抗体,并且可任选地使用多种方法(例如,使用重组DNA技术)对其进行修饰。抗体产生的其他实例和方法是本领域已知的(参见,例如Harlow et al.“Antibodies:A Laboratory Manual”,Cold Spring Harbor Laboratory,1988.)。

[0153] 在一些实施方案中,对抗体进行修饰,例如,通过糖基化、磷酸化、SUMO化和/或(例如,和)甲基化进行修饰。在一些实施方案中,抗体是与一个或更多个糖或碳水化合物分子缀合的糖基化抗体。在一些实施方案中,一个或更多个糖或碳水化合物分子通过N-糖基化、O-糖基化、C-糖基化、糖基磷脂酰肌醇化(GPI锚定附着)和/或(例如,和)磷酸糖基化与抗体缀合。在一些实施方案中,一个或更多个糖或碳水化合物分子是单糖、二糖、寡糖或聚糖。在一些实施方案中,一个或更多个糖或碳水化合物分子是支化的寡糖或支化的聚糖。在一些实施方案中,一个或更多个糖或碳水化合物分子包含甘露糖单元、葡萄糖单元、N-乙酰葡糖胺单元、N-乙酰半乳糖胺单元、半乳糖单元、岩藻糖单元或磷脂单元。在一些实施方案中,存在约1至10、约1至5、约5至10、约1至4、约1至3或约2个糖分子。在一些实施方案中,糖基化抗体是完全或部分糖基化的。在一些实施方案中,通过化学反应或通过酶促手段使抗体糖基化。在一些实施方案中,抗体在体外或细胞内被糖基化,其可任选地缺少N-或O-糖基化途径中的酶,例如糖基转移酶。在一些实施方案中,用糖或碳水化合物分子对抗体进行官能化,如2014年5月1日公开的标题为“Modified antibody,antibody-conjugate and process for the preparation thereof”的国际专利申请公开WO2014065661中所述。

[0154] 在一些实施方案中,本公开内容的抗TfR1抗体包含选自表2至7中任一者的任一种抗TfR1抗体的VL结构域和/或(例如,和)VH结构域,并且包含恒定区,所述恒定区包含IgG、IgE、IgM、IgD、IgA或IgY免疫球蛋白分子、免疫球蛋白分子的任何类别(例如,IgG1、IgG2、IgG3、IgG4、IgA1和IgA2)或任何亚类(例如,IgG2a和IgG2b)的恒定区的氨基酸序列。本领域中描述了人恒定区的一些非限制性实例,例如参见上文Kabat E A et al.,(1991)。

[0155] 在一些实施方案中,与转铁蛋白受体结合的物质,例如抗TfR1抗体,能够靶向肌细胞和/或(例如,和)介导物质穿过血脑屏障的转运。转铁蛋白受体是将转铁蛋白转运通过细胞膜并参与胞内铁水平的调节和稳态的内化细胞表面受体。本公开内容的一些方面提供了能够与转铁蛋白受体结合的转铁蛋白受体结合蛋白。与转铁蛋白受体结合(例如,特异性结合)的抗体可在与转铁蛋白受体结合之后被内化至细胞中,例如通过受体介导的内吞作用而被内化至细胞中。

[0156] 在一些方面中,本文中提供了以高特异性和亲和力与转铁蛋白受体结合的人源化抗体。在一些实施方案中,本文中所述的人源化抗TfR1抗体与转铁蛋白受体的任何胞外表位或暴露于抗体的表位特异性结合。在一些实施方案中,本文中提供的人源化抗TfR1抗体与来自人、非人灵长类、小鼠、大鼠等的转铁蛋白受体特异性结合。在一些实施方案中,本文中提供的人源化抗TfR1抗体与人转铁蛋白受体结合。在一些实施方案中,本文中所述的人源化抗TfR1抗体与如SEQ ID NO:105至108中提供的人或非人灵长类转铁蛋白受体的氨基酸区段结合。在一些实施方案中,本文中所述的人源化抗TfR1抗体与这样的氨基酸区段结合:对应于如SEQ ID NO:105中所示的人转铁蛋白受体的第90至96位氨基酸,其不在转铁蛋白受体的顶端结构域中。在一些实施方案中,本文中所述的人源化抗TfR1抗体与TfR1结合但与不TfR2结合。

[0157] 在一些实施方案中,抗TFR1抗体以至少约10^{-4}M、10^{-5}M、10^{-6}M、10^{-7}M、10^{-8}M、10^{-9}M、10^{-10}M、10^{-11}M、10^{-12}M、10^{-13}M或更小的结合亲和力(例如,如Kd所指示)特异性结合TfR1(例如,人或非人灵长类TfR1)。在一些实施方案中,本文中所述的抗TfR1抗体以亚纳摩范围的KD与TfR1结合。在一些实施方案中,本文中所述的抗TfR1抗体选择性地与转铁蛋白受体1

(TfR1)结合但不与转铁蛋白受体2(TfR2)结合。在一些实施方案中,本文中所述的抗TfR1抗体与人TfR1和食蟹猴TfR1结合(例如,Kd为10^{-7}M、10^{-8}M、10^{-9}M、10^{-10}M、10^{-11}M、10^{-12}M、10^{-13}M或更小),但不与小鼠TfR1结合。抗TfR1抗体的亲和力和结合动力学可使用任何合适的方法测试,所述方法包括但不限于生物传感器技术(例如,OCTET或BIACORE)。在一些实施方案中,本文中所述的任一种抗TfR1抗体的结合不竞争或抑制转铁蛋白与TfR1的结合。在一些实施方案中,本文中所述的任一种抗TfR1抗体的结合不竞争或抑制HFE-β-2-微球蛋白与TfR1的结合。

[0158] 表2中提供了抗TfR1抗体的一些非限制性实例。

[0159] 表2.抗TfR1抗体的一些实例

[0160]

Ab	编号系统	IMGT	Kabat	Chothia
3-A4	CDR-H1	GFNIKDDY (SEQ ID NO: 1)	DDYMY (SEQ ID NO: 7)	GFNIKDD (SEQ ID NO: 12)
	CDR-H2	IDPENGDT (SEQ ID NO: 2)	WIDPENGDTEYASKFQD (SEQ ID NO: 8)	ENG (SEQ ID NO: 13)
	CDR-H3	TLWLRRGLDY (SEQ ID NO: 3)	WLRRGLDY (SEQ ID NO: 9)	LRRGLD (SEQ ID NO: 14)
	CDR-L1	KSLLHSNGYTY (SEQ ID NO: 4)	RSSKSLLHSNGYTYLF (SEQ ID NO: 10)	SKSLLHSNGYTY (SEQ ID NO: 15)
	CDR-L2	RMS (SEQ ID NO: 5)	RMSNLAS (SEQ ID NO: 11)	RMS (SEQ ID NO: 5)
	CDR-L3	MQHLEYPFT (SEQ ID NO: 6)	MQHLEYPFT (SEQ ID NO: 6)	HLEYPF (SEQ ID NO: 16)
	VH	EVQLQQSGAELVRPGASVKLSCTASGFNIKDDYMYWVKQRPEQGLEWIGWIDPENGDTEYASKFQDKATVTADTSSNTAYLQLSSLTSEDTAVYYCTLWLRRGLDYWGQGTSVTVSS (SEQ ID NO: 17)		
	VL	DIVMTQAAPSVPVTPGESVSISCRSSKSLLHSNGYTYLFWFLQRPGQSPQLLIYRMSNLASGVPDRFSGSGSGTAFTLRISRVEAEDVGVYYCMQHLEYPFTFGGGTKLEIK (SEQ ID NO: 18)		

[0161]

Ab	编号系统	IMGT	Kabat	Chothia
3-A4 N54T*	CDR-H1	GFNIKDDY (SEQ ID NO: 1)	DDYMY (SEQ ID NO: 7)	GFNIKDD (SEQ ID NO: 12)
	CDR-H2	IDPETGDT (SEQ ID NO: 19)	WIDPETGDTEYASKFQD (SEQ ID NO: 20)	ETG (SEQ ID NO: 21)
	CDR-H3	TLWLRRGLDY (SEQ ID NO: 3)	WLRRGLDY (SEQ ID NO: 9)	LRRGLD (SEQ ID NO: 14)
	CDR-L1	KSLLHSNGYTY (SEQ ID NO: 4)	RSSKSLLHSNGYTYLF (SEQ ID NO: 10)	SKSLLHSNGYTY (SEQ ID NO: 15)
	CDR-L2	RMS (SEQ ID NO: 5)	RMSNLAS (SEQ ID NO: 11)	RMS(SEQ ID NO: 5)
	CDR-L3	MQHLEYPFT (SEQ ID NO: 6)	MQHLEYPFT (SEQ ID NO: 6)	HLEYPF (SEQ ID NO: 16)
	VH	EVQLQQSGAELVRPGASVKLSCTASGFNIKDDYMYWVKQRPEQGLEWIGWIDPETGDTEYASKFQDKATVTADTSSNTAYLQLSSLTSEDTAVYYCTLWLRRGLDYWGQGTSVTVSS (SEQ ID NO: 22)		
	VL	DIVMTQAAPSVPVTPGESVSISCRSSKSLLHSNGYTYLFWFLQRPGQSPQLLIYRMSNLASGVPDRFSGSGSGTAFTLRISRVEAEDVGVYYCMQHLEYPFTFGGGTKLEIK (SEQ ID NO: 18)		
3-A4 N54S*	CDR-H1	GFNIKDDY (SEQ ID NO: 1)	DDYMY (SEQ ID NO: 7)	GFNIKDD (SEQ ID NO: 12)
	CDR-H2	IDPESGDT (SEQ ID NO: 23)	WIDPESGDTEYASKFQD (SEQ ID NO: 24)	ESG (SEQ ID NO: 25)
	CDR-H3	TLWLRRGLDY (SEQ ID NO: 3)	WLRRGLDY (SEQ ID NO: 9)	LRRGLD (SEQ ID NO: 14)
	CDR-L1	KSLLHSNGYTY (SEQ ID NO: 4)	RSSKSLLHSNGYTYLF (SEQ ID NO: 10)	SKSLLHSNGYTY (SEQ ID NO: 15)
	CDR-L2	RMS (SEQ ID NO: 5)	RMSNLAS (SEQ ID NO: 11)	RMS (SEQ ID NO: 5)
	CDR-L3	MQHLEYPFT (SEQ ID NO: 6)	MQHLEYPFT (SEQ ID NO: 6)	HLEYPF (SEQ ID NO: 16)
	VH	EVQLQQSGAELVRPGASVKLSCTASGFNIKDDYMYWVKQRPEQGLEWIGWIDPESGDTEYASKFQDKATVTADTSSNTAYLQLSSLTSEDTAVYYCTLWLRRGLDYWGQGTSVTVSS (SEQ ID NO: 26)		
	VL	DIVMTQAAPSVPVTPGESVSISCRSSKSLLHSNGYTYLFWFLQRPGQSPQLLIYRMSNLASGVPDRFSGSGSGTAFTLRISRVEAEDVGVYYCMQHLEYPFTFGGGTKLEIK (SEQ ID NO: 18)		
3-M12	CDR-H1	GYSITSGYY (SEQ ID NO: 27)	SGYYWN (SEQ ID NO: 33)	GYSITSGY (SEQ ID NO: 38)
	CDR-H2	ITFDGAN (SEQ ID NO: 28)	YITFDGANNYNPSLKN (SEQ ID NO: 34)	FDG (SEQ ID NO: 39)
	CDR-H3	TRSSYDYDVLDY (SEQ ID NO: 29)	SSYDYDVLDY (SEQ ID NO: 35)	SYDYDVLD (SEQ ID NO: 40)
	CDR-L1	QDISNF (SEQ ID NO: 30)	RASQDISNFLN (SEQ ID NO: 36)	SQDISNF (SEQ ID NO: 41)
	CDR-L2	YTS (SEQ ID NO: 31)	YTSRLHS (SEQ ID NO: 37)	YTS (SEQ ID NO: 31)
	CDR-L3	QQGHTLPYT (SEQ ID NO: 32)	QQGHTLPYT (SEQ ID NO: 32)	GHTLPY (SEQ ID NO: 42)
	VH	DVQLQESGPGLVKPSQSLSLTCSVTGYSITSGYYWNWIRQFPGNKLEWMGYITFDGANNYNPSLKNRISITRDTSKNQFFLKLTSVTTEDTATYYCTRSSYDYDVLDYWGQGTTLTVSS (SEQ ID NO: 43)		
	VL	DIQMTQTTSSLSASLGDRVTISCRASQDISNFLNWYQQRPDGTVKLLIYYTSRLHSGVPSRFSGSGSGTDFSLTVSNLEQEDIATYFCQQGHTLPYTFGGGTKLEIK (SEQ ID NO: 44)		
5-H12	CDR-H1	GYSFTDYC (SEQ ID NO: 45)	DYCIN (SEQ ID NO: 51)	GYSFTDY (SEQ ID NO: 56)
	CDR-H2	IYPGSGNT (SEQ ID NO: 46)	WIYPGSGNTRYSERFKG (SEQ ID NO: 52)	GSG (SEQ ID NO: 57)
	CDR-H3	AREDYYPYHGMDY (SEQ ID NO: 47)	EDYYPYHGMDY (SEQ ID NO: 53)	DYYPYHGMD (SEQ ID NO: 58)

[0162]

Ab	编号系统	IMGT	Kabat	Chothia
	CDR-L1	ESVDGYDNSF (SEQ ID NO: 48)	RASESVDGYDNSFMH (SEQ ID NO: 54)	SESVDGYDNSF (SEQ ID NO: 59)
	CDR-L2	RAS (SEQ ID NO: 49)	RASNLES (SEQ ID NO: 55)	RAS (SEQ ID NO: 49)
	CDR-L3	QQSSEDPWT (SEQ ID NO: 50)	QQSSEDPWT (SEQ ID NO: 50)	SSEDPW (SEQ ID NO: 60)
	VH	QIQLQQSGPELVRPGASVKISCKASGYSFTDYCINWVNQRPGQGLEWIGWIYPGSGNTRYSERFKGKATLTVDTSSNTAYMQLSSLTSEDSAVYFCAREDYYPYHGMDYWGQGTSVTVSS (SEQ ID NO: 61)		
	VL	DIVLTQSPTSLAVSLGQRATISCRASESVDGYDNSFMHWYQQKPGQPPKLLIFRASNLESGIPARFSGSGSRTDFTLTINPVEAADVATYYCQQSSEDPWTFGGGTKLEIK (SEQ ID NO: 62)		
5-H12 C33Y*	CDR-H1	GYSFTDYY (SEQ ID NO: 63)	DYYIN (SEQ ID NO: 64)	GYSFTDY (SEQ ID NO: 56)
	CDR-H2	IYPGSGNT (SEQ ID NO: 46)	WIYPGSGNTRYSERFKG (SEQ ID NO: 52)	GSG (SEQ ID NO: 57)
	CDR-H3	AREDYYPYHGMDY (SEQ ID NO: 47)	EDYYPYHGMDY (SEQ ID NO: 53)	DYYPYHGMD (SEQ ID NO: 58)
	CDR-L1	ESVDGYDNSF (SEQ ID NO: 48)	RASESVDGYDNSFMH (SEQ ID NO: 54)	SESVDGYDNSF (SEQ ID NO: 59)
	CDR-L2	RAS (SEQ ID NO: 49)	RASNLES (SEQ ID NO: 55)	RAS (SEQ ID NO: 49)
	CDR-L3	QQSSEDPWT (SEQ ID NO: 50)	QQSSEDPWT (SEQ ID NO: 50)	SSEDPW (SEQ ID NO: 60)
	VH	QIQLQQSGPELVRPGASVKISCKASGYSFTDYYINWVNQRPGQGLEWIGWIYPGSGNTRYSERFKGKATLTVDTSSNTAYMQLSSLTSEDSAVYFCAREDYYPYHGMDYWGQGTSVTVSS (SEQ ID NO: 65)		
	VL	DIVLTQSPTSLAVSLGQRATISCRASESVDGYDNSFMHWYQQKPGQPPKLLIFRASNLESGIPARFSGSGSRTDFTLTINPVEAADVATYYCQQSSEDPWTFGGGTKLEIK (SEQ ID NO: 62)		
5-H12 C33D*	CDR-H1	GYSFTDYD (SEQ ID NO: 66)	DYDIN (SEQ ID NO: 67)	GYSFTDY (SEQ ID NO: 56)
	CDR-H2	IYPGSGNT (SEQ ID NO: 46)	WIYPGSGNTRYSERFKG (SEQ ID NO: 52)	GSG (SEQ ID NO: 57)
	CDR-H3	AREDYYPYHGMDY (SEQ ID NO: 47)	EDYYPYHGMDY (SEQ ID NO: 53)	DYYPYHGMD (SEQ ID NO: 58)
	CDR-L1	ESVDGYDNSF (SEQ ID NO: 48)	RASESVDGYDNSFMH (SEQ ID NO: 54)	SESVDGYDNSF (SEQ ID NO: 59)
	CDR-L2	RAS (SEQ ID NO: 49)	RASNLES (SEQ ID NO: 55)	RAS (SEQ ID NO: 49)
	CDR-L3	QQSSEDPWT (SEQ ID NO: 50)	QQSSEDPWT (SEQ ID NO: 50)	SSEDPW (SEQ ID NO: 60)
	VH	QIQLQQSGPELVRPGASVKISCKASGYSFTDYDINWVNQRPGQGLEWIGWIYPGSGNTRYSERFKGKATLTVDTSSNTAYMQLSSLTSEDSAVYFCAREDYYPYHGMDYWGQGTSVTVSS (SEQ ID NO: 68)		
	VL	DIVLTQSPTSLAVSLGQRATISCRASESVDGYDNSFMHWYQQKPGQPPKLLIFRASNLESGIPARFSGSGSRTDFTLTINPVEAADVATYYCQQSSEDPWTFGGGTKLEIK (SEQ ID NO: 62)		
抗TfR克隆 8	CDR-H1	GYSFTSYW (SEQ ID NO: 138)	SYWIG (SEQ ID NO: 144)	GYSFTSY (SEQ ID NO: 149)
	CDR-H2	IYPGDSDT (SEQ ID NO: 139)	IIYPGDSDTRYSPSFQGQ (SEQ ID NO: 145)	GDS (SEQ ID NO: 150)
	CDR-H3	ARFPYDSSGYYSFDY (SEQ ID NO: 140)	FPYDSSGYYSFDY (SEQ ID NO: 146)	PYDSSGYYSFD (SEQ ID NO: 151)
	CDR-L1	QSISSY (SEQ ID NO: 141)	RASQSISSYLN (SEQ ID NO: 147)	SQSISSY (SEQ ID NO: 152)
	CDR-L2	AAS (SEQ ID NO: 142)	AASSLQS (SEQ ID NO: 148)	AAS (SEQ ID NO: 142)

[0163]

Ab	编号系统	IMGT	Kabat	Chothia
	CDR-L3	QQSYSTPLT (SEQ ID NO: 143)	QQSYSTPLT (SEQ ID NO: 143)	SYSTPL (SEQ ID NO: 153)

[0164] *突变位置根据包含突变的相应VH序列的Kabat编号

[0165] 在一些实施方案中,本公开内容的抗TfR1抗体是表2中提供的任一种抗TfR1抗体

的人源化变体。在一些实施方案中,本公开内容的抗TfR1抗体包含与表2中提供的任一种抗TfR1抗体中的CDR-H1、CDR-H2和CDR-H3相同的CDR-H1、CDR-H2、CDR-H3、CDR-L1、CDR-L2和CDR-L3,并且包含人源化重链可变区和/或(例如,和)人源化轻链可变区。

[0166] 表3中提供了本文中所述的抗TfR1抗体的氨基酸序列的一些实例。

[0167] 表3.抗TfR1抗体的可变区

[0168]

抗体	可变区氨基酸序列 **
3A4 VH3 (N54T*)/Vκ4	V_H: EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDP ETGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLD Y**WGQGTLVTVSS (SEQ ID NO: 69)
	V_L: DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**R MSNLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTK VEIK (SEQ ID NO: 70)
3A4 VH3 (N54S*)/Vκ4	V_H: EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDP ESGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLD Y**WGQGTLVTVSS (SEQ ID NO: 71)
	V_L: DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**R MSNLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTK VEIK (SEQ ID NO: 70)
3A4 VH3 /Vκ4	VH: EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDP ENGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLD Y**WGQGTLVTVSS (SEQ ID NO: 72)
	VL: DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**R MSNLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTK VEIK (SEQ ID NO: 70)
3M12 VH3/Vκ2	VH: QVQLQESGPGLVKPSQTLSLTCSVTGYSIT**SGYYWN**WIRQPPGKGLEWMG**YITF DGANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY** WGQGTTVTVSS (SEQ ID NO: 73)
	VL: DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLH S**GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC**QQGHTLPYT**FGQGTKLEIK (SEQ ID NO: 74)
3M12 VH3/Vκ3	VH: QVQLQESGPGLVKPSQTLSLTCSVTGYSIT**SGYYWN**WIRQPPGKGLEWMG**YITF DGANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY** WGQGTTVTVSS (SEQ ID NO: 73)

[0169]

抗体	可变区氨基酸序列 **
	VL: DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLH**SGVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQGHTLPYT**FGQGTKLEIK (SEQ ID NO: 75)
3M12 VH4/Vκ2	VH: QVQLQESGPGLVKPSQTLSLTCTVTGYSIT**SGYYWN**WIRQPPGKGLEWIG**YITFD GANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WGQGTTVTVSS (SEQ ID NO: 76)
	VL: DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLH**SGVPSRFSGSGSGTDFTLTISSLQPEDFATYFC**QQGHTLPYT**FGQGTKLEIK (SEQ ID NO: 74)
3M12 VH4/Vκ3	VH: QVQLQESGPGLVKPSQTLSLTCTVTGYSIT**SGYYWN**WIRQPPGKGLEWIG**YITFD GANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WGQGTTVTVSS (SEQ ID NO: 76)
	VL: DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLH**SGVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQGHTLPYT**FGQGTKLEIK (SEQ ID NO: 75)
5H12 VH5 (C33Y*)/Vκ3	VH: QVQLVQSGAEVKKPGASVKVSCKASGYSFT**DYYIN**WVRQAPGQGLEWMG**WIY PGSGNTRYSERFKG**RVTITRDTSASTAYMELSSLRSEDTAVYYCAR**EDYYPYH GMDY**WGQGTLVTVSS (SEQ ID NO: 77)
	VL: DIVLTQSPDSLAVSLGERATINC**RASESVDGYDNSFMH**WYQQKPGQPPKLLIF**R ASNLES**GVPDRFSGSGSRTDFTLTISSLQAEDVAVYYC**QQSSEDPWT**FGQGTKL EIK (SEQ ID NO: 78)
5H12 VH5 (C33D*)/Vκ4	VH: QVQLVQSGAEVKKPGASVKVSCKASGYSFT**DYDIN**WVRQAPGQGLEWMG**WIY PGSGNTRYSERFKG**RVTITRDTSASTAYMELSSLRSEDTAVYYCAR**EDYYPYH GMDY**WGQGTLVTVSS (SEQ ID NO: 79)
	VL: DIVMTQSPDSLAVSLGERATINC**RASESVDGYDNSFMH**WYQQKPGQPPKLLIF**R ASNLES**GVPDRFSGSGSGTDFTLTISSLQAEDVAVYYC**QQSSEDPWT**FGQGTKL EIK (SEQ ID NO: 80)
5H12 VH5 (C33Y*)/Vκ4	VH: QVQLVQSGAEVKKPGASVKVSCKASGYSFT**DYYIN**WVRQAPGQGLEWMG**WIY PGSGNTRYSERFKG**RVTITRDTSASTAYMELSSLRSEDTAVYYCAR**EDYYPYH GMDY**WGQGTLVTVSS (SEQ ID NO: 77)
	VL: DIVMTQSPDSLAVSLGERATINC**RASESVDGYDNSFMH**WYQQKPGQPPKLLIF**R ASNLES**GVPDRFSGSGSGTDFTLTISSLQAEDVAVYYC**QQSSEDPWT**FGQGTKL EIK (SEQ ID NO: 80)
抗 TfR 克隆 8	VH: QVQLVQSGAEVKKPGESLKISCKGSGYSFT**SYWIG**WVRQMPGKGLEWMG**IIYP GDSDTRYSPSFQGQ**VTISADKSISTAYLQWSSLKASDTAMYYCAR**FPYDSSGYY SFDY**WGQGTLVTVSS (SEQ ID NO: 154)
	VL: DIQMTQSPSSLSASVGDRVTITC**RASQSISSYLN**WYQQKPGKAPKLLIY**AASSLQ S**GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQSYSTPLT**FGGGTKVEIK (SEQ ID NO: 155)

[0170] *突变位置根据包含突变的相应VH序列的Kabat编号

[0171] **根据Kabat编号系统的CDR被加粗

[0172] 在一些实施方案中,本公开内容的抗TfR1抗体包含VH,所述VH包含表3中提供的任一种抗TfR1抗体的CDR-H1、CDR-H2和CDR-H3,并且与表3中提供的相应VH相比,在框架区中包含一个或更多个(例如,1、2、3、4、5、6、7、8、9、10或更多个)氨基酸变异。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含VL,所述VL包含表3中提供的任一种抗TfR1抗体的CDR-L1、CDR-L2和CDR-L3,并且与表3中提供的相应VL相比,在框架区中包含一个或更多个(例如,1、2、3、4、5、6、7、8、9、10或更多个)氨基酸变异。在一些实施方案中,抗TfR1抗体

的VH是人源化VH,和/或抗TfR1抗体的VL是人源化VL。

[0173] 在一些实施方案中,本公开内容的抗TfR1抗体包含VH,所述VH包含表3中提供的任一种抗TfR1抗体的CDR-H1、CDR-H2和CDR-H3,并且与表3中提供的相应VH相比在框架区中包含具有至少70%(例如,至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少99%)同一性的氨基酸序列。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含VL,所述VL包含表3中提供的任一种抗TfR1抗体的CDR-L1、CDR-L2和CDR-L3,并且与表3中提供的相应VL相比在框架区中包含具有至少70%(例如,至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少99%)同一性的氨基酸序列。在一些实施方案中,抗TfR1抗体的VH是人源化VH,和/或抗TfR1抗体的VL是人源化VL。

[0174] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:69的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL。

[0175] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:71的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL。

[0176] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:72的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL。

[0177] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:73的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL。

[0178] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:73的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL。

[0179] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL。

[0180] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL。

[0181] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:78的氨基酸序列的VL。

[0182] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:79的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL。

[0183] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL。

[0184] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:154的氨基酸序列的VH和含有SEQ ID NO:155的氨基酸序列的VL。

[0185] 在一些实施方案中,本文中所述的抗TfR1抗体是全长IgG,其可包含来自人抗体的重恒定区和轻恒定区。在一些实施方案中,本文中所述的任何抗TfR1抗体的重链可包含重链恒定区(CH)或其一部分(例如,CH1、CH2、CH3或其组合)。重链恒定区可具有任何合适的来源,例如人、小鼠、大鼠或兔。在一个具体实例中,重链恒定区来自人IgG例如IgG1、IgG2或IgG4(γ重链)。以下给出了人IgG1恒定区的一个实例:

[0186] ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS
SGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLG
GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE
QYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLP
PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLT
VDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK (SEQ ID NO: 81)。

[0187] 在一些实施方案中,本文中所述的任何抗TfR1抗体的重链包含突变体人IgG1恒定区。例如,已知在人IgG1的CH2结构域中引入LALA突变(来源于mAb b12的突变体,其已突变以用Ala234和Ala235替换较低的铰链残基Leu234 Leu235)降低了Fc γ 受体结合(Bruhns,P.,et al.(2009)和Xu,D.et al.(2000))。以下提供了突变体人IgG1恒定区(突变加粗并带有下划线):

[0188] ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS
SGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPE**AA**
GGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPRE
EQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTL
PPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLT
VDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK (SEQ ID NO: 82)。

[0189] 在一些实施方案中,本文中所述的任何抗TfR1抗体的轻链还可包含轻链恒定区(CL),其可以是本领域已知的任何CL。在一些实例中,CL是κ轻链。在另一些实例中,CL是λ轻链。在一些实施方案中,CL是κ轻链,以下提供了其序列:

[0190] RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQ
DSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 83)。

[0191] 另外的抗体重链和轻链恒定区是本领域公知的,例如在IMGT数据库(www.imgt.org)或www.vbase2.org/vbstat.php中提供的那些,二者均通过引用并入本文。

[0192] 在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含表3中所列出的任一种VH或其任意变体以及与SEQ ID NO:81或SEQ ID NO:82具有至少80%、至少85%、至少90%、至少95%或至少99%同一性的重链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含表3中所列出的任一种VH或其任意变体以及与SEQ ID NO:81或SEQ ID NO:82相比包含不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)的重链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含表3中所列出的任一种VH或其任意变体以及SEQ ID NO:81中所示的重链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含表3所中列出的任一种VH或其任意变体以及SEQ ID NO:82中所示的重链恒定区。

[0193] 在一些实施方案中,本文中所述的抗TfR1抗体包含轻链,所述轻链包含表3中所列出的任一种VL或其任意变体以及与SEQ ID NO:83具有至少80%、至少85%、至少90%、至少95%或至少99%同一性的轻链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含轻链,所述轻链包含表3中所列出的任一种VL或其任意变体以及与SEQ ID NO:83相比包含

不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)的轻链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含轻链,所述轻链包含表3中所列出的任一种VL或其任意变体以及SEQ ID NO:83中所示的轻链恒定区。

[0194] 下表4中提供了所述抗TfR1抗体的IgG重链氨基酸序列和轻链氨基酸序列的一些实例。

[0195] 表4.抗TfR1 IgG实例的重链序列和轻链序列

[0196]

抗体	IgG 重链/轻链序列 **
3A4 VH3 (N54T*)/Vκ4	重链(具有野生型人IgG1恒定区) EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDPE** **TGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLDY**W GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS CDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ PENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL SLSPGK (SEQ ID NO: 84)
	轻链(具有κ轻链恒定区) DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**RMS** **NLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTKVEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 85)
3A4 VH3 (N54S*)/Vκ4	重链(具有野生型人IgG1恒定区) EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDPE** **SGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLDY**W GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS CDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ PENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL SLSPGK (SEQ ID NO: 86)
	轻链(具有κ轻链恒定区) DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**RMS** **NLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTKVEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 85)
3A4 VH3 /Vκ4	重链(具有野生型人IgG1恒定区) EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDPE** **NGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLDY**W GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS CDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ PENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL SLSPGK (SEQ ID NO: 87)

[0197]

抗体	IgG 重链/轻链序列 **
	轻链（具有κ轻链恒定区） DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**RMS NLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTKVEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 85)
3M12 VH3/Vκ2	重链（具有野生型人 IgG1 恒定区） QVQLQESGPGLVKPSQTLSLTCSVTGYSIT**SGYYWN**WIRQPPGKGLEWMG**YITFD GANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WG QGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALT SGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ PENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL SLSPGK (SEQ ID NO: 88)
	轻链（具有κ轻链恒定区） DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLHS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC**QQGHTLPYT**FGQGTKLEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 89)
3M12 VH3/Vκ3	重链（具有野生型人 IgG1 恒定区） QVQLQESGPGLVKPSQTLSLTCSVTGYSIT**SGYYWN**WIRQPPGKGLEWMG**YITFD GANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WG QGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALT SGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC DKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN WYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ PENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL SLSPGK (SEQ ID NO: 88)
	轻链（具有κ轻链恒定区） DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLHS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQGHTLPYT**FGQGTKLEIKRTVAA PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDS KDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 90)
3M12 VH4/Vκ2	重链（具有野生型人 IgG1 恒定区） QVQLQESGPGLVKPSQTLSLTCTVTGYSIT**SGYYWN**WIRQPPGKGLEWIG**YITFDG ANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WGQ GTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTS GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA PIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP ENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLS LSPGK (SEQ ID NO: 91)
	轻链（具有κ轻链恒定区） DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLHS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC**QQGHTLPYT**FGQGTKLEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 89)

[0198]

抗体	IgG 重链/轻链序列 **
3M12 VH4/Vκ3	重链（具有野生型人 IgG1 恒定区） QVQLQESGPGLVKPSQTLSLTCTVTGYSITSGYYWNWIRQPPGKGLEWIGYITFDG ANNYNPSLKNRVSISRDTSKNQFSLKLSSVTAEDTATYYCTRSSYDYDVLDYWGQ GTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTS GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNW YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPA PIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP ENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLS LSPGK (SEQ ID NO: 91)
	轻链（具有κ轻链恒定区） DIQMTQSPSSLSASVGDRVTITCRASQDISNFLNWYQQKPGQPVKLLIYYTSRLHS GVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQGHTLPYTFGQGTKLEIKRTVAA PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDS KDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 90)
5H12 VH5 (C33Y*)/Vκ3	重链（具有野生型人 IgG1 恒定区） QVQLVQSGAEVKKPGASVKVSCKASGYSFTDYYINWVRQAPGQGLEWMGWIYP GSGNTRYSERFKGRVTITRDTSASTAYMELSSLRSEDTAVYYCAREDYYPYHGM DYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWE SNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT QKSLSLSPGK (SEQ ID NO: 92)
	轻链（具有κ轻链恒定区） DIVLTQSPDSLAVSLGERATINCRASESVDGYDNSFMHWYQQKPGQPPKLLIFRAS NLESGVPDRFSGSGSRTDFTLTISSLQAEDVAVYYCQQSSEDPWTFGQGTKLEIKR TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVT EQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 93)
5H12 VH5 (C33D*)/Vκ4	重链（具有野生型人 IgG1 恒定区） QVQLVQSGAEVKKPGASVKVSCKASGYSFTDYDINWVRQAPGQGLEWMGWIYP GSGNTRYSERFKGRVTITRDTSASTAYMELSSLRSEDTAVYYCAREDYYPYHGM DYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWE SNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT QKSLSLSPGK (SEQ ID NO: 94)
	轻链（具有κ轻链恒定区） DIVMTQSPDSLAVSLGERATINCRASESVDGYDNSFMHWYQQKPGQPPKLLIFRA SNLESGVPDRFSGSGSGTDFTLTISSLQAEDVAVYYCQQSSEDPWTFGQGTKLEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 95)
5H12 VH5 (C33Y*)/Vκ4	重链（具有野生型人 IgG1 恒定区） QVQLVQSGAEVKKPGASVKVSCKASGYSFTDYYINWVRQAPGQGLEWMGWIYP GSGNTRYSERFKGRVTITRDTSASTAYMELSSLRSEDTAVYYCAREDYYPYHGM DYWGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWE SNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT QKSLSLSPGK (SEQ ID NO: 92)

[0199]

抗体	IgG 重链/轻链序列 **
抗 TfR 克隆 8	轻链（具有κ轻链恒定区） DIVMTQSPDSLAVSLGERATINC**RASESVDGYDNSFMH**WYQQKPGQPPKLLIF**RA SNLES**GVPDRFSGSGSGTDFTLTISSLQAEDVAVYYC**QQSSEDPWT**FGQGTKLEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 95)
	VH: QVQLVQSGAEVKKPGESLKISCKGSGYSFT**SYWIG**WVRQMPGKGLEWMG**IIYPG DSDTRYSPSFQGQ**VTISADKSISTAYLQWSSLKASDTAMYYCAR**FPYDSSGYYSF DY**WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSN KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWE SNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYT QKSLSLSPGK (SEQ ID NO: 156)
	VL: DIQMTQSPSSLSASVGDRVTITC**RASQSISSYLN**WYQQKPGKAPKLLIY**AASSLQS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQSYSTPLT**FGGGTKVEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 157)

[0200] *突变位置根据包含突变的相应VH序列的Kabat编号

[0201] **根据Kabat编号系统的CDR被加粗;VH/VL序列带有下划线

[0202] 在一些实施方案中,本公开内容的抗TfR1抗体包含重链,所述重链与SEQ ID NO:84、86、87、88、91、92、94和156中任一者所示的重链相比,包含不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含轻链,所述轻链与SEQ ID NO:85、89、90、93、95和157中任一者所示的轻链相比包含不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)。

[0203] 在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含与SEQ ID NO:84、86、87、88、91、92、94和156中任一者具有至少75%(例如,75%、80%、85%、90%、95%、98%或99%)同一性的氨基酸序列。作为替代或补充(例如,补充),本文中所述的抗TfR1抗体包含轻链,所述轻链包含与SEQ ID NO:85、89、90、93、95和157中任一者具有至少75%(例如,75%、80%、85%、90%、95%、98%或99%)同一性的氨基酸序列。在一些实施方案中,本文中所述的抗TfR1抗体包含含有SEQ ID NO:84、86、87、88、91、92、94和156中任一者的氨基酸序列的重链。作为替代或补充(例如,补充),本文中所述的抗TfR1抗体包含含有SEQ ID NO:85、89、90、93、95和157中任一者的氨基酸序列的轻链。

[0204] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:84的氨基酸序列的重链和含有SEQ ID NO:85的氨基酸序列的轻链。

[0205] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:86的氨基酸序列的重链和含有SEQ ID NO:85的氨基酸序列的轻链。

[0206] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:87的氨基酸序列的重链和含有SEQ ID NO:85的氨基酸序列的轻链。

[0207] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:88的氨基酸序列的重链和含有SEQ ID NO:89的氨基酸序列的轻链。

[0208] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:88的氨基酸序列的重链和含有SEQ ID NO:90的氨基酸序列的轻链。

[0209] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:91的氨基酸序列的重链和含有SEQ ID NO:89的氨基酸序列的轻链。

[0210] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:91的氨基酸序列的重链和含有SEQ ID NO:90的氨基酸序列的轻链。

[0211] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:92的氨基酸序列的重链和含有SEQ ID NO:93的氨基酸序列的轻链。

[0212] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:94的氨基酸序列的重链和含有SEQ ID NO:95的氨基酸序列的轻链。

[0213] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:92的氨基酸序列的重链和含有SEQ ID NO:95的氨基酸序列的轻链。

[0214] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:156的氨基酸序列的重链和含有SEQ ID NO:157的氨基酸序列的轻链。

[0215] 在一些实施方案中,抗TfR1抗体是完整抗体(全长抗体)的Fab片段、Fab'片段或F(ab')$_2$片段。完整抗体(全长抗体)的抗原结合片段可以通过常规方法制备(例如,重组地制备或通过使用酶例如木瓜蛋白酶消化全长IgG的重链恒定区来制备)。例如,F(ab')$_2$片段可通过胃蛋白酶或木瓜蛋白酶消化抗体分子产生,并且Fab片段可通过还原F(ab')$_2$片段的二硫桥产生。在一些实施方案中,本文中所述的抗TfR1抗体的Fab片段中的重链恒定区包含以下氨基酸序列:

[0216] ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS SGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHT (SEQ ID NO: 96)。

[0217] 在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含表3中所列出的任一种VH或其任意变体以及与SEQ ID NO:96具有至少80%、至少85%、至少90%、至少95%或至少99%同一性的重链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含表3中列出的任一种VH或其任意变体以及与SEQ ID NO:96相比包含不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)的重链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含表3中列出的任一种VH或其任意变体以及SEQ ID NO:96中所示的重链恒定区。

[0218] 在一些实施方案中,本文中所述的抗TfR1抗体包含轻链,所述轻链包含表3中所列出的任一种VL或其任意变体以及与SEQ ID NO:83具有至少80%、至少85%、至少90%、至少95%或至少99%同一性的轻链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含轻链,所述轻链包含表3中所列出的任一种VL或其任意变体以及与SEQ ID NO:83相比包含不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)的轻链恒定区。在一些实施方案中,本文中所述的抗TfR1抗体包含轻链,所述轻链包含表3中所列出的任一种VL或其任意变体以及SEQ ID NO:83中所示的轻链恒定区。

[0219] 下表5中提供了所述抗TfR1抗体的Fab重链和轻链氨基酸序列的一些实例。

[0220] 表5.抗TfR1 Fab实例的重链和轻链序列

[0221]

抗体	Fab 重链/轻链序列 **
3A4 VH3 (N54T*)/Vκ4	重链(具有部分人 IgG1 恒定区) EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDPE** **TGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLDY**W GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS CDKTHT (SEQ ID NO: 97)
	轻链(具有κ轻链恒定区) DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**RMS** **NLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTKVEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 85)
3A4 VH3 (N54S*)/Vκ4	重链(具有部分人 IgG1 恒定区) EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDPE** **SGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLDY**W GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS CDKTHT (SEQ ID NO: 98)
	轻链(具有κ轻链恒定区) DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**RMS** **NLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTKVEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 85)
3A4 VH3 /Vκ4	重链(具有部分人 IgG1 恒定区) EVQLVQSGSELKKPGASVKVSCTASGFNIK**DDYMY**WVRQPPGKGLEWIG**WIDPE** **NGDTEYASKFQD**RVTVTADTSTNTAYMELSSLRSEDTAVYYCTL**WLRRGLDY**W GQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGAL TSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKS CDKTHT (SEQ ID NO: 99)
	轻链(具有κ轻链恒定区) DIVMTQSPLSLPVTPGEPASISC**RSSKSLLHSNGYTYLF**WFQQRPGQSPRLLIY**RMS** **NLAS**GVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQHLEYPFT**FGGGTKVEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 85)
3M12 VH3/Vκ2	重链(具有部分人 IgG1 恒定区) QVQLQESGPGLVKPSQTLSLTCSVTGYSIT**SGYYWN**WIRQPPGKGLEWMG**YITFD** **GANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WG QGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALT SGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC DKTHT (SEQ ID NO: 100)
	轻链(具有κ轻链恒定区) DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLHS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC**QQGHTLPYT**FGQGTKLEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 89)
3M12 VH3/Vκ3	重链(具有部分人 IgG1 恒定区) QVQLQESGPGLVKPSQTLSLTCSVTGYSIT**SGYYWN**WIRQPPGKGLEWMG**YITFD** **GANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WG QGTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALT SGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSC DKTHT (SEQ ID NO: 100)
	轻链(具有κ轻链恒定区) DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLHS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQGHTLPYT**FGQGTKLEIKRTVAA PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDS KDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 90)

[0222]

抗体	Fab 重链/轻链序列 **
3M12 VH4/Vκ2	重链(具有部分人 IgG1 恒定区) QVQLQESGPGLVKPSQTLSLTCTVTGYSIT**SGYYWN**WIRQPPGKGLEWIG**YITFDG** **ANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WGQ GTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTS GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHT (SEQ ID NO: 101)
	轻链(具有κ轻链恒定区) DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLHS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC**QQGHTLPYT**FGQGTKLEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 89)
3M12 VH4/Vκ3	重链(具有部分人 IgG1 恒定区) QVQLQESGPGLVKPSQTLSLTCTVTGYSIT**SGYYWN**WIRQPPGKGLEWIG**YITFDG** **ANNYNPSLKN**RVSISRDTSKNQFSLKLSSVTAEDTATYYCTR**SSYDYDVLDY**WGQ GTTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTS GVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCD KTHT (SEQ ID NO: 101)
	轻链(具有κ轻链恒定区) DIQMTQSPSSLSASVGDRVTITC**RASQDISNFLN**WYQQKPGQPVKLLIY**YTSRLHS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQGHTLPYT**FGQGTKLEIKRTVAA PSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDS KDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 90)
5H12 VH5 (C33Y*)/Vκ3	重链(具有部分人 IgG1 恒定区) QVQLVQSGAEVKKPGASVKVSCKASGYSFT**DYYIN**WVRQAPGQGLEWMG**WIYP** **GSGNTRYSERFKG**RVTITRDTSASTAYMELSSLRSEDTAVYYCAR**EDYYPYHGM** **DY**WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHT (SEQ ID NO: 102)
	轻链(具有κ轻链恒定区) DIVLTQSPDSLAVSLGERATINC**RASESVDGYDNSFMH**WYQQKPGQPPKLLIF**RAS** **NLES**GVPDRFSGSGSRTDFTLTISSLQAEDVAVYYC**QQSSEDPWT**FGQGTKLEIKR TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVT EQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 93)
5H12 VH5 (C33D*)/Vκ4	重链(具有部分人 IgG1 恒定区) QVQLVQSGAEVKKPGASVKVSCKASGYSFT**DYDIN**WVRQAPGQGLEWMG**WIYP** **GSGNTRYSERFKG**RVTITRDTSASTAYMELSSLRSEDTAVYYCAR**EDYYPYHGM** **DY**WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHT (SEQ ID NO: 103)
	轻链(具有κ轻链恒定区) DIVMTQSPDSLAVSLGERATINC**RASESVDGYDNSFMH**WYQQKPGQPPKLLIF**RA** **SNLES**GVPDRFSGSGSGTDFTLTISSLQAEDVAVYYC**QQSSEDPWT**FGQGTKLEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 95)
5H12 VH5 (C33Y*)/Vκ4	重链(具有部分人 IgG1 恒定区) QVQLVQSGAEVKKPGASVKVSCKASGYSFT**DYYIN**WVRQAPGQGLEWMG**WIYP** **GSGNTRYSERFKG**RVTITRDTSASTAYMELSSLRSEDTAVYYCAR**EDYYPYHGM** **DY**WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHT (SEQ ID NO: 102)
	轻链(具有κ轻链恒定区) DIVMTQSPDSLAVSLGERATINC**RASESVDGYDNSFMH**WYQQKPGQPPKLLIF**RA** **SNLES**GVPDRFSGSGSGTDFTLTISSLQAEDVAVYYC**QQSSEDPWT**FGQGTKLEIK RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 95)

[0223]

抗体	Fab 重链/轻链序列 **
抗 TfR 克隆 8 形式 1	VH: QVQLVQSGAEVKKPGESLKISCKGSGYSFT**SYWIG**WVRQMPGKGLEWMG**IIYPG DSDTRYSPSFQGQ**VTISADKSISTAYLQWSSLKASDTAMYYCAR**FPYDSSGYYSF DY**WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCP (SEQ ID NO: 158)
	VL: DIQMTQSPSSLSASVGDRVTITC**RASQSISSYLN**WYQQKPGKAPKLLIY**AASSLQS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQSYSTPLT**FGGGTKVEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 157)
抗 TfR 克隆 8 形式 2	VH: QVQLVQSGAEVKKPGESLKISCKGSGYSFT**SYWIG**WVRQMPGKGLEWMG**IIYPG DSDTRYSPSFQGQ**VTISADKSISTAYLQWSSLKASDTAMYYCAR**FPYDSSGYYSF DY**WGQGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWN SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHT (SEQ ID NO: 159)
	VL: DIQMTQSPSSLSASVGDRVTITC**RASQSISSYLN**WYQQKPGKAPKLLIY**AASSLQS** GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC**QQSYSTPLT**FGGGTKVEIKRTVAAP SVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSK DSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 157)

[0224] *突变位置根据包含突变的相应VH序列的Kabat编号

[0225] **根据Kabat编号系统的CDR被加粗;VH/VL序列带有下划线

[0226] 在一些实施方案中,本公开内容的抗TfR1抗体包含重链,所述重链与SEQ ID NO:97至103、158和159中任一者中所示的重链相比包含不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含轻链,所述轻链与SEQ ID NO:85、89、90、93、95和157中任一者中所示的轻链相比不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)。

[0227] 在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含与SEQ ID NO:97至103、158和159中任一者具有至少75%(例如,75%、80%、85%、90%、95%、98%或99%)同一性的氨基酸序列。作为替代或补充(例如,补充),本文中所述的抗TfR1抗体包含轻链,所述轻链包含与SEQ ID NO:85、89、90、93、95和157中任一者具有至少75%(例如,75%、80%、85%、90%、95%、98%或99%)同一性的氨基酸序列。在一些实施方案中,本文中所述的抗TfR1抗体包含重链,所述重链包含SEQ ID NO:97至103、158和159中任一者的氨基酸序列。作为替代或补充(例如,补充),本文中所述的抗TfR1抗体包含轻链,所述轻链包含SEQ ID NO:85、89、90、93、95和157中任一者的氨基酸序列。

[0228] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:97的氨基酸序列的重链和含有SEQ ID NO:85的氨基酸序列的轻链。

[0229] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:98的氨基酸序列的重链和含有SEQ ID NO:85的氨基酸序列的轻链。

[0230] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:99的氨基酸序列的重链和含有SEQ ID NO:85的氨基酸序列的轻链。

[0231] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:100的氨基酸

序列的重链和含有SEQ ID NO:89的氨基酸序列的轻链。

[0232] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:100的氨基酸序列的重链和含有SEQ ID NO:90的氨基酸序列的轻链。

[0233] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:101的氨基酸序列的重链和含有SEQ ID NO:89的氨基酸序列的轻链。

[0234] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:101的氨基酸序列的重链和含有SEQ ID NO:90的氨基酸序列的轻链。

[0235] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:102的氨基酸序列的重链和含有SEQ ID NO:93的氨基酸序列的轻链。

[0236] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:103的氨基酸序列的重链和含有SEQ ID NO:95的氨基酸序列的轻链。

[0237] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:102的氨基酸序列的重链和含有SEQ ID NO:95的氨基酸序列的轻链。

[0238] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:158的氨基酸序列的重链和含有SEQ ID NO:157的氨基酸序列的轻链。

[0239] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:159的氨基酸序列的重链和含有SEQ ID NO:157的氨基酸序列的轻链。

[0240] 其他已知的抗TfR1抗体

[0241] 本领域中已知的任何其他合适的抗TfR1抗体均可用作本文中公开的复合物中的肌肉靶向剂。表6中列出了已知抗TfR1抗体的一些实例,包括相关参考文献和结合表位。在一些实施方案中,抗TfR1抗体包含本文中提供的任何抗TfR1抗体(例如表6中列出的抗TfR1抗体)的互补决定区(CDR-H1、CDR-H2、CDR-H3、CDR-L1、CDR-L2和CDR-L3)。

[0242] 表6-抗TfR1抗体克隆的列表,包括相关参考文献和结合表位信息。

[0243]

抗体克隆名称	参考文献	表位/注释
OKT9	1979 年 12 月 4 日提交的美国专利 No. 4,364,934，标题为“MONOCLONAL ANTIBODY TO A HUMAN EARLY THYMOCYTE ANTIGEN AND METHODS FOR PREPARING SAME” Schneider C. et al. “Structural features of the cell surface receptor for transferrin that is recognized by the monoclonal antibody OKT9.” J Biol Chem. 1982, 257:14, 8516-8522.	TfR1 的顶端结构域（人 TfR1 序列 XM_052730.3 的第 305 至 366 位残基，可在 GenBank 获得）
（来自 JCR） 克隆 M11 克隆 M23 克隆 M27 克隆 B84	• 2014 年 12 月 24 日提交的 WO 2015/098989，“Novel anti-Transferrin receptor antibody that passes through blood-brain barrier” • 2014 年 12 月 24 日提交的美国专利 No. 9,994,641，“Novel anti-Transferrin receptor antibody that passes through blood-brain barrier”	顶端结构域（TfR1 的第 230 至 244 位和第 326 至 347 位残基）和蛋白酶样结构域（第 461 至 473 位残基）

[0244]

抗体克隆名称	参考文献	表位/注释
（来自Genentech） 7A4、8A2、15D2、10D11、7B10、15G11、16G5、13C3、16G4、16F6、7G7、4C2、1B12和13D4	• 2016年5月26日提交的WO 2016/081643，标题为“ANTI-TRANSFERRIN RECEPTOR ANTIBODIES AND METHODS OF USE” • 2014年5月20日提交的美国专利No. 9,708,406，“Anti-transferrin receptor antibodies and methods of use”	顶端结构域和非顶端区
（来自Armagen） 8D3	• Lee et al. “Targeting Rat Anti-Mouse Transferrin Receptor Monoclonal Antibodies through Blood-Brain Barrier in Mouse” 2000, J Pharmacol. Exp. Ther., 292: 1048-1052. • 2008年9月11日提交的美国专利申请2010/077498，标题为“COMPOSITIONS AND METHODS FOR BLOOD-BRAIN BARRIER DELIVERY IN THE MOUSE”	
OX26	• Haobam, B. et al. 2014. Rab17-mediated recycling endosomes contribute to autophagosome formation in response to Group A Streptococcus invasion. Cellular microbiology. 16: 1806-21.	
DF1513	• Ortiz-Zapater E et al. Trafficking of the human transferrin receptor in plant cells: effects of tyrphostin A23 and brefeldin A. Plant J 48:757-70 (2006).	

[0245]

抗体克隆名称	参考文献	表位/注释
1A1B2、66IG10、MEM-189 、JF0956、29806、1A1B2 、TFRC/1818 、1E6 、66Ig10 、TFRC/1059 、Q1/71、23D10、13E4 、TFRC/1149 、ER-MP21 、YTA74.4 、BU54、2B6、RI7 217	• 市售抗转铁蛋白受体抗体。	Novus Biologicals 8100 Southpark Way, A-8 Littleton CO 80120
（来自INSERM） BA120g	• 2005 年 6 月 15 日提交的美国专利申请 2011/0311544A1，标题为“ANTI-CD71 MONOCLONAL ANTIBODIES AND USES THEREOF FOR TREATING MALIGNANT TUMOR CELLS”	不与 OKT9 竞争
LUCA31	• 2004 年 6 月 7 日提交的美国专利 No. 7,572,895，标题为“TRANSFERRIN RECEPTOR ANTIBODIES”	“LUCA31 表位”
（Salk Institute） B3/25 T58/30	• Trowbridge, I.S. et al. “Anti-transferrin receptor monoclonal antibody and toxin–antibody conjugates affect growth of human tumour cells.” Nature, 1981, 第 294 卷, 第 171 至 173 页	
R17 217.1.3、5E9C11、OKT9（BE0023 克隆）	• 市售抗转铁蛋白受体抗体。	BioXcell 10 Technology Dr., Suite 2B West Lebanon, NH 03784-1671 USA

抗体克隆名称	参考文献	表位/注释
BK19.9、B3/25、T56/14 和 T58/1	• Gatter, K.C. et al. “Transferrin receptors in human tissues: their distribution and possible clinical relevance.” J Clin Pathol. 1983 May;36(5):539-45.	

[0246]

抗 TfR1 抗体

CDRH1 (SEQ ID NO: 952)
CDRH2 (SEQ ID NO: 953)
CDRH3 (SEQ ID NO: 954)
CDRL1 (SEQ ID NO: 955)
CDRL2 (SEQ ID NO: 956)
CDRL3 (SEQ ID NO: 957)
VH (SEQ ID NO: 958)
VL (SEQ ID NO: 959)

另外的抗 TfR1 抗体 SEQ ID NO

	VH/VL	CDR1	CDR2	CDR3
VH1	967	960	961	954
VH2	968	960	962	954
VH3	969	960	963	954
VH4	970	960	962	954
VL1	971	955	956	115
VL2	972	955	956	115
VL3	973	955	964	957
VL4	974	965	966	957

[0247] 在一些实施方案中,本公开内容的抗TfR1抗体包含来自选自表6的任一种抗TfR1抗体的一个或更多个CDR-H(例如,CDR-H1、CDR-H2和CDR-H3)氨基酸序列。在一些实施方案中,抗TfR1抗体包含如针对选自表6的任一种抗TfR1抗体所提供的CDR-L1、CDR-L2和CDR-L3。在一些实施方案中,抗TfR1抗体包含如针对选自表6的任一种抗TfR1抗体所提供的CDR-H1、CDR-H2、CDR-H3、CDR-L1、CDR-L2和CDR-L3。

[0248] 在一些实施方案中,本公开内容的抗TfR1抗体包括包含任何抗TfR1抗体(例如选自表6的任一种抗TfR1抗体)的重链可变结构域和/或(例如,和)轻链可变结构域的任何抗体。在一些实施方案中,本公开内容的抗TfR1抗体包括包含任何抗TfR1抗体(例如选自表6的任一种抗TfR1抗体)的重链可变和轻链可变对的任何抗体。

[0249] 本公开内容的一些方面提供了抗TfR1抗体,其具有与本文中所述的那些中的任何同源的重链可变(VH)和/或(例如,和)轻链可变(VL)结构域氨基酸序列。在一些实施方案中,抗TfR1抗体包含与任何抗TfR1抗体(例如选自表6的任一种抗TfR1抗体)的重链可变序列和/或任何轻链可变序列具有至少75%(例如80%、85%、90%、95%、98%或99%)同一性的重链可变序列或轻链可变序列。在一些实施方案中,同源重链可变和/或(例如,和)轻链可变氨基酸序列在本文中提供的任何CDR序列内均不变化。例如,在一些实施方案中,序列变异的程度(例如,75%、80%、85%、90%、95%、98%或99%)可发生在不包括本文中提供的任何CDR序列的重链可变和/或(例如,和)轻链可变序列中。在一些实施方案中,本文中提供的任何抗TfR1抗体包含重链可变序列和轻链可变序列,其包含与任何抗TfR1抗体(例如选自表6的任一种抗TfR1抗体)的框架序列具有至少75%、80%、85%、90%、95%、98%或99%同一性的框架序列。

[0250] 根据本公开内容可使用的转铁蛋白受体抗体的实例描述于国际申请公开WO 2016/081643中,其通过引用并入本文。该抗体的氨基酸序列在表7中提供。

[0251] 表7.已知的抗TfR1抗体实例的重链和轻链CDR

[0252]

序列类型	Kabat	Chothia	Contact
CDR-H1	SYWMH (SEQ ID NO: 110)	GYTFTSY (SEQ ID NO: 116)	TSYWMH (SEQ ID NO: 118)
CDR-H2	EINPTNGRTNYIEKFKS (SEQ ID NO: 111)	NPTNGR (SEQ ID NO: 117)	WIGEINPTNGRTN (SEQ ID NO: 119)
CDR-H3	GTRAYHY (SEQ ID NO: 112)	GTRAYHY (SEQ ID NO: 112)	ARGTRA (SEQ ID NO: 120)
CDR-L1	RASDNLYSNLA (SEQ ID NO: 113)	RASDNLYSNLA (SEQ ID NO: 113)	YSNLAWY (SEQ ID NO: 121)
CDR-L2	DATNLAD (SEQ ID NO: 114)	DATNLAD (SEQ ID NO: 114)	LLVYDATNLA (SEQ ID NO: 122)
CDR-L3	QHFWGTPLT (SEQ ID NO: 115)	QHFWGTPLT (SEQ ID NO: 115)	QHFWGTPL (SEQ ID NO: 123)
鼠 VH	QVQLQQPGAELVKPGASVKLSCKASGYTFTSYWMHWVKQRPGQGLEWIGEINP TNGRTNYIEKFKSKATLTVDKSSSTAYMQLSSLTSEDSAVYYCARGTRAYHYW GQGTSVTVSS (SEQ ID NO: 124)		
鼠 VL	DIQMTQSPASLSVSVGETVTITCRASDNLYSNLAWYQQKQGKSPQLLVYDATNL ADGVPSRFSGSGSGTQYSLKINSLQSEDFGTYYCQHFWGTPLTFGAGTKLELK (SEQ ID NO: 125)		
人源化 VH	EVQLVQSGAEVKKPGASVKVSCKASGYTFTSYWMHWVRQAPGQRLEWIGEIN PTNGRTNYIEKFKSRATLTVDKSASTAYMELSSLRSEDTAVYYCARGTRAYHY WGQGTMVTVSS (SEQ ID NO: 128)		
人源化 VL	DIQMTQSPSSLSASVGDRVTITCRASDNLYSNLAWYQQKPGKSPKLLVYDATNL ADGVPSRFSGSGSGTDYTLTISSLQPEDFATYYCQHFWGTPLTFGQGTKVEIK (SEQ ID NO: 129)		
嵌合全长 IgG1的 HC	QVQLQQPGAELVKPGASVKLSCKASGYTFTSYWMHWVKQRPGQGLEWIGEINP TNGRTNYIEKFKSKATLTVDKSSSTAYMQLSSLTSEDSAVYYCARGTRAYHYW GQGTSVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSG ALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVE PKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPE VKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV SNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV EWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEAL HNHYTQKSLSLSPGK (SEQ ID NO: 132)		
嵌合全长 IgG1的LC	DIQMTQSPASLSVSVGETVTITCRASDNLYSNLAWYQQKQGKSPQLLVYDATNL ADGVPSRFSGSGSGTQYSLKINSLQSEDFGTYYCQHFWGTPLTFGAGTKLELKR TVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQES VTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 133)		

[0253]

序列类型	Kabat	Chothia	Contact
完全人全长 IgG1的 HC	EVQLVQSGAEVKKPGASVKVSCKASGYTFTSYWMHWVRQAPGQRLEWIGEIN PTNGRTNYIEKFKSRATLTVDKSASTAYMELSSLRSEDTAVYYCARGTRAYHY WGQGTMVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNS GALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDP EVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCK VSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA VEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEA LHNHYTQKSLSLSPGK (SEQ ID NO: 134)		
完全人全长 IgG1的 LC	DIQMTQSPSSLSASVGDRVTITCRASDNLYSNLAWYQQKPGKSPKLLVYDATNL ADGVPSRFSGSGSGTDYTLTISSLQPEDFATYYCQHFWGTPLTFGQGTKVEIKRT VAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESV TEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 135)		
嵌合Fab的 HC	QVQLQQPGAELVKPGASVKLSCKASGYTFTSYWMHWVKQRPGQGLEWIGEINP TNGRTNYIEKFKSKATLTVDKSSSTAYMQLSSLTSEDSAVYYCARGTRAYHYW GQGTSVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSG ALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVE PKSCDKTHTCP (SEQ ID NO: 136)		
完全人Fab 的HC	EVQLVQSGAEVKKPGASVKVSCKASGYTFTSYWMHWVRQAPGQRLEWIGEIN PTNGRTNYIEKFKSRATLTVDKSASTAYMELSSLRSEDTAVYYCARGTRAYHY WGQGTMVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNS GALTSGVHTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKV EPKSCDKTHTCP (SEQ ID NO: 137)		

[0254] 在一些实施方案中,本公开内容的抗TfR1抗体包含与表7中所示的CDR-H1、CDR-H2和CDR-H3相同的CDR-H1、CDR-H2和CDR-H3。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含与表7所示的CDR-L1、CDR-L2和CDR-L3相同的CDR-L1、CDR-L2和CDR-L3。

[0255] 在一些实施方案中,本公开内容的抗TfR1抗体包含CDR-L3,其与表7中所示的CDR-L3相比包含不超过3个氨基酸变异(例如,不超过3、2或1个氨基酸变异)。在一些实施方案中,本公开内容的抗TfR1抗体包含CDR-L3,其与表7中所示的CDR-L3相比包含一个氨基酸变异。在一些实施方案中,本公开内容的抗TfR1抗体包含QHFAGTPLT(SEQ ID NO:126)的CDR-L3(根据Kabat和Chothia定义系统)或QHFAGTPL(SEQ ID NO:127)的CDR-L3(根据Contact定义系统)。在一些实施方案中,本公开内容的抗TfR1抗体包含与表7中所示的CDR-H1、CDR-H2和CDR-H3相同的CDR-H1、CDR-H2、CDR-H3、CDR-L1和CDR-L2,并且包含QHFAGTPLT(SEQ ID NO:126)的CDR-L3(根据Kabat和Chothia定义系统)或QHFAGTPL(SEQ ID NO:127)的CDR-L3(根据Contact定义系统)。

[0256] 在一些实施方案中,本公开内容的抗TfR1抗体包含重链CDR,其共同与表7中所示的重链CDR具有至少80%(例如,80%、85%、90%、95%或98%)同一性。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含轻链CDR,其共同与表7所示的轻链CDR具有至少80%(例如,80%、85%、90%、95%或98%)同一性。

[0257] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:124的氨基酸序列的VH。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含含有SEQ ID NO:125的氨基酸序列的VL。

[0258] 在一些实施方案中,本公开内容的抗TfR1抗体包含含有SEQ ID NO:128的氨基酸序列的VH。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含含有SEQ ID NO:129的氨基酸序列的VL。

[0259] 在一些实施方案中,本公开内容的抗TfR1抗体包含VH,所述VH与如SEQ ID NO:128中所示的VH相比包含不超过25个氨基酸变异(例如,不超过25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2或1个氨基酸变异)。作为替代或补充(例如,补充),本公开内容的抗TfR1抗体包含VL,所述VL与如SEQ ID NO:129中所示的VL相比包含不超过15个氨基酸变异(例如,不超过20、19、18、17、16、15、14、13、12、11、9、8、7、6、5、4、3、2或1个氨基酸变异)。

[0260] 在一些实施方案中,本公开内容的抗TfR1抗体是全长IgG1抗体,其可包含来自人抗体的重恒定区和轻恒定区。在一些实施方案中,本文中所述的任何抗TfR1抗体的重链可包含重链恒定区(CH)或其一部分(例如,CH1、CH2、CH3或其组合)。重链恒定区可具有任何合适的来源,例如人、小鼠、大鼠或兔。在一个具体实例中,重链恒定区来自人IgG例如IgG1、IgG2或IgG4(γ重链)。以下给出了人IgG1恒定区的一个实例:

[0261] ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQS
SGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLG
GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE
QYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLP
PSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLT
VDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK (SEQ ID NO: 81)

。

[0262] 在一些实施方案中,本文中所述的任何抗TfR1抗体的轻链还可包含轻链恒定区(CL),其可以是本领域已知的任何CL。在一些实例中,CL是κ轻链。在另一些实例中,CL是λ轻链。在一些实施方案中,CL是κ轻链,以下提供了其序列:

[0263] RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQ DSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC (SEQ ID NO: 83)。

[0264] 在一些实施方案中,本文中所述的抗TfR1抗体是嵌合抗体,其包含含有SEQ ID NO:132的氨基酸序列的重链。作为替代或补充(例如,补充),本文中所述的抗TfR1抗体包含含有SEQ ID NO:133的氨基酸序列的轻链。

[0265] 在一些实施方案中,本文中所述的抗TfR1抗体是包含重链的完全人抗体,所述重链包含SEQ ID NO:134的氨基酸序列。作为替代或补充(例如,补充),本文中所述的抗TfR1抗体包含含有SEQ ID NO:135的氨基酸序列的轻链。

[0266] 在一些实施方案中,抗TfR1抗体是完整抗体(全长抗体)的抗原结合片段(Fab)。在一些实施方案中,本文中所述的抗TfR1 Fab包含含有SEQ ID NO:136的氨基酸序列的重链。作为替代或补充(例如,补充),本文中所述的抗TfR1 Fab包含含有SEQ ID NO:133的氨基酸序列的轻链。在一些实施方案中,本文中所述的抗TfR1 Fab包含含有SEQ ID NO:137的氨基酸序列的重链。作为替代或补充(例如,补充),本文中所述的抗TfR1 Fab包含含有SEQ ID NO:135的氨基酸序列的轻链。

[0267] 本文中所述的抗TfR1抗体可为任何抗体形式,包括但不限于完整(即,全长)抗体、其抗原结合片段(例如Fab、Fab'、F(ab')2、Fv)、单链抗体、双特异性抗体或纳米抗体。在一些实施方案中,本文中所述的抗TfR1抗体是scFv。在一些实施方案中,本文中所述的抗TfR1抗体是scFv-Fab(例如,与恒定区的一部分融合的scFv)。在一些实施方案中,本文中所述的抗TfR1抗体是与恒定区(例如,SEQ ID NO:81中所示的人IgG1恒定区)融合的scFv。

[0268] 在一些实施方案中,可在残基不太可能参与与靶抗原(例如,转铁蛋白受体)相互作用的位置(例如,如基于晶体结构确定的)处将保守突变引入抗体序列(例如,CDR或框架序列)中。在一些实施方案中,将一个、两个或更多个突变(例如,氨基酸替换)引入本文中所述的抗TfR1抗体的Fc区(例如,在CH2结构域(人IgG1的第231至340位残基)和/或(例如,和)CH3结构域(人IgG1的第341至447位残基)和/或(例如,和)铰链区中,根据Kabat编号系统(例如,Kabat中的EU索引)编号)中,以改变抗体的一种或更多种功能特性,例如血清半衰期、补体固定、Fc受体结合和/或(例如,和)抗原依赖性细胞毒性。

[0269] 在一些实施方案中,将一个、两个或更多个突变(例如,氨基酸替换)引入Fc区(CH1结构域)的铰链区中,使得铰链区中的半胱氨酸残基的数目改变(例如,增加或减少),如例如美国专利No.5,677,425中所述。可改变CH1结构域的铰链区中半胱氨酸残基的数目,例如以促进轻链和重链的组装,或改变(例如,提高或降低)抗体的稳定性或促进接头缀合。

[0270] 在一些实施方案中,将一个、两个或更多个突变(例如,氨基酸替换)引入本文中所述的肌肉靶向抗体的Fc区(例如,在CH2结构域(人IgG1的第231至340位残基)和/或(例如,和)CH3结构域(人IgG1的第341至447位残基)和/或(例如,和)铰链区中,根据Kabat编号系统(例如,Kabat中的EU索引)编号)中,以提高或降低抗体对效应细胞表面上Fc受体(例如,激活的Fc受体)的亲和力。降低或提高抗体对Fc受体的亲和力的抗体Fc区中的突变以及用于将这样的突变引入Fc受体或其片段的技术是本领域技术人员已知的。可被进行以改变抗

体对Fc受体的亲和力的抗体Fc受体中的突变的一些实例描述于以下中:例如Smith P et al.,(2012)PNAS109:6181-6186、美国专利No.6,737,056,以及国际公开No.WO 02/060919、WO 98/23289和WO 97/34631,其通过引用并入本文。

[0271] 在一些实施方案中,将一个、两个或更多个氨基酸突变(即,替换、插入或缺失)引入IgG恒定结构域或其FcRn结合片段(优选地,Fc或铰链-Fc结构域片段)中以改变(例如,降低或提高)抗体在体内的半衰期。例如,将改变(例如,降低或提高)抗体在体内的半衰期的突变参见例如国际公开No.WO 02/060919、WO 98/23289和WO 97/34631,以及美国专利No.5,869,046、6,121,022、6,277,375和6,165,745。

[0272] 在一些实施方案中,将一个、两个或更多个氨基酸突变(即,替换、插入或缺失)引入IgG恒定结构域或其FcRn结合片段(优选地,Fc或铰链-Fc结构域片段)中以降低体内抗TfR1抗体的半衰期。在一些实施方案中,将一个、两个或更多个氨基酸突变(即,替换、插入或缺失)引入IgG恒定结构域或其FcRn结合片段(优选地,Fc或铰链-Fc结构域片段)中以提高抗体在体内的半衰期。在一些实施方案中,抗体可在第二恒定(CH2)结构域(人IgG1的第231至340位残基)和/或(例如,和)第三恒定(CH3)结构域(人IgG1的第341至447位残基)(根据Kabat中的EU索引(Kabat E A et al.,(1991)同上)编号)中具有一个或更多个氨基酸突变(例如,替换)。在一些实施方案中,本文中所述的抗体的IgG1的恒定区包含在第252位处的甲硫氨酸(M)至酪氨酸(Y)替换,在第254位处的丝氨酸(S)至苏氨酸(T)替换,以及在第256位处的苏氨酸(T)至谷氨酸(E)替换,所述位置根据Kabat中的EU索引编号。参见美国专利No.7,658,921,其通过引用并入本文。这种类型的突变体IgG(称为“YTE突变体”)已显示出与同一抗体的野生型形式相比半衰期提高4倍(参见Dall'Acqua W F et al.,(2006)J Biol Chem 281:23514-24)。在一些实施方案中,抗体包含IgG恒定结构域,该结构域包含在第251至257、285至290、308至314、385至389和428至436位处的氨基酸残基的一个、两个、三个或更多个氨基酸替换,所述位置根据Kabat中的EU索引编号。

[0273] 在一些实施方案中,将一个、两个或更多个氨基酸替换引入IgG恒定结构域Fc区中,以改变抗TfR1抗体的效应物功能。对其的亲和力被改变的效应配体可以是例如Fc受体或补体的C1组分。该方法在美国专利No.5,624,821和5,648,260中有更详细的描述。在一些实施方案中,恒定区结构域的缺失或失活(通过点突变或其他方式)可降低循环抗体的Fc受体结合,从而提高肿瘤定位。对于使恒定结构域缺失或失活从而提高肿瘤定位的突变的描述,参见例如美国专利No.5,585,097和8,591,886。在一些实施方案中,可将一个或更多个氨基酸替换引入本文中所述的抗体的Fc区中,以去除Fc区上潜在的糖基化位点,这可降低Fc受体结合(参见,例如Shields R L et al.,(2001)J Biol Chem 276:6591-604)。

[0274] 在一些实施方案中,可将本文中所述的抗TfR1抗体的恒定区中的一个或更多个氨基替换为不同的氨基酸残基,使得抗体具有改变的C1q结合和/或(例如,和)降低或消除的补体依赖性细胞毒性(complement dependent cytotoxicity,CDC)。这种方法在美国专利No.6,194,551(Idusogie et al)中有更详细的描述。在一些实施方案中,改变本文中所述的抗体的CH2结构域的N端区域中的一个或更多个氨基酸残基,从而改变抗体的固定补体的能力。这种方法在国际公开No.WO 94/29351中有进一步描述。在一些实施方案中,对本文中所述的抗体的Fc区进行修饰以提高抗体的介导抗体依赖性细胞毒性(antibody dependent cellular cytotoxicity,ADCC)的能力和/或(例如,和)提高抗体对Fcγ受体的亲和力。这

种方法在国际公开No.WO 00/42072中有进一步描述。

[0275] 在一些实施方案中,本文中提供的抗体的重链和/或(例如,和)轻链可变结构域序列可用于产生例如CDR接枝、嵌合、人源化或复合的人抗体或抗原结合片段,如本文中其他地方所述。如本领域普通技术人员所理解的,来源于本文中提供的任何抗体的任何变体(CDR接枝、嵌合、人源化或复合的抗体)可用于本文中所述的组合物和方法中,并将保持特异性结合转铁蛋白受体的能力,从而使得相对于其所来源的原始抗体,变体(CDR接枝、嵌合、人源化或复合的抗体)具有至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或更多的与转铁蛋白受体的结合。

[0276] 在一些实施方案中,本文中提供的抗体包含赋予抗体以期望性质的突变。例如,为了避免归因于已知与天然IgG4 mAb发生的Fab臂交换而引起的潜在并发症,本文中提供的抗体可包含稳定性'Adair'突变(Angal S.,et al.,"Asingle amino acid substitution abolishes the heterogeneity of chimeric mouse/human(IgG4)antibody,"Mol Immunol 30,105-108;1993),其中第228位(EU编号;根据Kabat编号为第241位残基)丝氨酸转化为脯氨酸,从而产生了IgG1样铰链序列。因此,任何抗体都可包含稳定性'Adair'突变。

[0277] 在一些实施方案中,对抗体进行修饰,例如,通过糖基化、磷酸化、SUMO化和/或(例如,和)甲基化进行修饰。在一些实施方案中,抗体是与一个或更多个糖或碳水化合物分子缀合的糖基化抗体。在一些实施方案中,一个或更多个糖或碳水化合物分子通过N-糖基化、O-糖基化、C-糖基化、糖基磷脂酰肌醇化(GPI锚定附着)和/或(例如,和)磷酸糖基化与抗体缀合。在一些实施方案中,一个或更多个糖或碳水化合物分子是单糖、二糖、寡糖或聚糖。在一些实施方案中,一个或更多个糖或碳水化合物分子是支化的寡糖或支化的聚糖。在一些实施方案中,一个或更多个糖或碳水化合物分子包含甘露糖单元、葡萄糖单元、N-乙酰葡糖胺单元、N-乙酰半乳糖胺单元、半乳糖单元、岩藻糖单元或磷脂单元。在一些实施方案中,存在约1至10、约1至5、约5至10、约1至4、约1至3或约2个糖分子。在一些实施方案中,糖基化抗体是完全或部分糖基化的。在一些实施方案中,通过化学反应或通过酶促方式使抗体糖基化。在一些实施方案中,抗体在体外或细胞内被糖基化,其可任选地缺少N-或O-糖基化途径中的酶,例如糖基转移酶。在一些实施方案中,用糖或碳水化合物分子对抗体进行官能化,如2014年5月1日公开的标题为"Modified antibody,antibody-conjugate and process for the preparation thereof"的国际专利申请公开WO2014065661中所述。

[0278] 在一些实施方案中,本文中所述的任一种抗TfR1抗体可在重链序列和/或(例如,和)轻链序列中包含信号肽(例如,N端信号肽)。在一些实施方案中,本文中所述的抗TfR1抗体包含VH和VL序列中的任一种、IgG重链序列和轻链序列中的任一种、或者本文中所述的F(ab')重链序列和轻链序列中的任一种,并且还包含信号肽(例如,N端信号肽)。在一些实施方案中,信号肽包含氨基酸序列MGWSCIILFLVATATGVHS(SEQ ID NO:104)。

[0279] 在一些实施方案中,本文中提供的抗体可具有一种或更多种翻译后修饰。在一些实施方案中,也称为焦谷氨酸形成(pyro-Glu)的N端环化可在产生期间在抗体的N端谷氨酸(Glu)和/或谷氨酰胺(Gln)残基处发生。因此,应理解指定为具有包含N端谷氨酸或谷氨酰胺残基的序列的抗体涵盖已经历由翻译后修饰引起的焦谷氨酸形成的抗体。在一些实施方案中,焦谷氨酸形成发生在重链序列中。在一些实施方案中,焦谷氨酸形成发生在轻链序列中。

[0280] b.其他肌肉靶向抗体

[0281] 在一些实施方案中,肌肉靶向抗体是特异性结合血幼素(hemojuvelin)、小窝蛋白-3、迪谢内肌营养不良肽(Duchenne muscular dystrophy peptide)、肌球蛋白IIb或CD63的抗体。在一些实施方案中,肌肉靶向抗体是特异性结合肌原性前体蛋白的抗体。一些示例性的肌原性前体蛋白包括但不限于ABCG2、M-钙黏着蛋白/钙黏着蛋白-15、小窝蛋白-1、CD34、FoxK1、整联蛋白α7、整联蛋白α7β1、MYF-5、MyoD、肌细胞生成蛋白、NCAM-1/CD56、Pax3、Pax7和Pax9。在一些实施方案中,肌肉靶向抗体是特异性结合骨骼肌蛋白的抗体。一些示例性的骨骼肌蛋白包括但不限于α-肌聚糖蛋白(alpha-Sarcoglycan)、β-肌聚糖蛋白、钙蛋白酶抑制剂、肌酸激酶MM/CKMM、eIF5A、烯醇化酶2/神经元特异性烯醇化酶、ε-肌聚糖蛋白、FABP3/H-FABP、GDF-8/肌生成抑制蛋白、GDF-11/GDF-8、整联蛋白α7、整联蛋白α7β1、整联蛋白β1/CD29、MCAM/CD146、MyoD、肌细胞生成蛋白、肌球蛋白轻链激酶抑制剂、NCAM-1/CD56和肌钙蛋白I。在一些实施方案中,肌肉靶向抗体是特异性结合平滑肌蛋白的抗体。一些示例性的平滑肌蛋白包括但不限于α-平滑肌肌动蛋白、VE-钙黏着蛋白、钙调蛋白结合蛋白/CALD1、钙调理蛋白1、结蛋白(Desmin)、组胺H2 R、胃动素R/GPR38、转凝蛋白/TAGLN、和波形蛋白。然而,应理解,针对另外的靶标的抗体在本公开内容的范围内,并且本文中提供的靶标的示例性列表并不意味着是限制性的。

[0282] c.抗体特征/改变

[0283] 在一些实施方案中,可在残基不太可能参与与靶抗原(例如,转铁蛋白受体)相互作用的位置(例如,如基于晶体结构确定的)处将保守突变引入抗体序列(例如,CDR或框架序列)中。在一些实施方案中,将一个、两个或更多个突变(例如,氨基酸替换)引入本文中所述的肌肉靶向抗体的Fc区(例如,在CH2结构域(人IgG1的第231至340位残基)和/或(例如,和)CH3结构域(人IgG1的第341至447位残基)和/或(例如,和)铰链区中,根据Kabat编号系统(例如,Kabat中的EU索引)编号)中,以改变抗体的一种或更多种功能特性,例如血清半衰期、补体固定、Fc受体结合和/或(例如,和)抗原依赖性细胞毒性。

[0284] 在一些实施方案中,将一个、两个或更多个突变(例如,氨基酸替换)引入Fc区(CH1结构域)的铰链区中,使得铰链区中的半胱氨酸残基的数目改变(例如,增加或减少),如例如美国专利No.5,677,425中所述。可改变CH1结构域的铰链区中半胱氨酸残基的数目,例如以促进轻链和重链的组装,或改变(例如,提高或降低)抗体的稳定性或促进接头缀合。

[0285] 在一些实施方案中,将一个、两个或更多个突变(例如,氨基酸替换)引入本文中所述的肌肉靶向抗体的Fc区(例如,在CH2结构域(人IgG1的第231至340位残基)和/或(例如,和)CH3结构域(人IgG1的第341至447位残基)和/或(例如,和)铰链区中,根据Kabat编号系统(例如,Kabat中的EU索引)编号)中,以提高或降低抗体对效应细胞表面上Fc受体(例如,激活的Fc受体)的亲和力。降低或提高抗体对Fc受体的亲和力的抗体Fc区中的突变以及用于将这样的突变引入Fc受体或其片段中的技术是本领域技术人员已知的。可被进行以改变抗体对Fc受体的亲和力的抗体Fc受体中的突变的一些实例描述于以下中:例如Smith P et al.,(2012)PNAS109:6181-6186、美国专利No.6,737,056,以及国际公开No.WO 02/060919、WO 98/23289、和WO 97/34631,其通过引用并入本文。

[0286] 在一些实施方案中,将一个、两个或更多个氨基酸突变(即,替换、插入或缺失)引入IgG恒定结构域或其FcRn结合片段(优选地,Fc或铰链-Fc结构域片段)中以改变(例如,降

低或提高)抗体在体内的半衰期。例如,将改变(例如,降低或提高)抗体在体内的半衰期的突变参见例如国际公开No.WO 02/060919、WO 98/23289和WO 97/34631,以及美国专利No.5,869,046、6,121,022、6,277,375和6,165,745。

[0287] 在一些实施方案中,将一个、两个或更多个氨基酸突变(即,替换、插入或缺失)引入IgG恒定结构域或其FcRn结合片段(优选地,Fc或铰链-Fc结构域片段)中以降低体内抗转铁蛋白受体抗体的半衰期。在一些实施方案中,将一个、两个或更多个氨基酸突变(即,替换、插入或缺失)引入IgG恒定结构域或其FcRn结合片段(优选地,Fc或铰链-Fc结构域片段)中以提高抗体在体内的半衰期。在一些实施方案中,抗体可在第二恒定(CH2)结构域(人IgG1的第231至340位残基)和/或(例如,和)第三恒定(CH3)结构域(人IgG1的第341至447位残基)(根据Kabat中的EU索引(Kabat E A et al.,(1991)同上)编号)中具有一个或更多个氨基酸突变(例如,替换)。在一些实施方案中,本文中所述的抗体的IgG1的恒定区包含在第252位处的甲硫氨酸(M)至酪氨酸(Y)替换,在第254位处的丝氨酸(S)至苏氨酸(T)替换,以及在第256位处的苏氨酸(T)至谷氨酸(E)替换,所述位置根据Kabat中的EU索引编号。参见美国专利No.7,658,921,其通过引用并入本文。这种类型的突变体IgG(称为"YTE突变体")已显示出与同一抗体的野生型形式相比半衰期提高4倍(参见Dall'Acqua W F et al.,(2006)J Biol Chem 281:23514-24)。在一些实施方案中,抗体包含IgG恒定结构域,该结构域包含在第251至257、285至290、308至314、385至389和428至436位处的氨基酸残基的一个、两个、三个或更多个氨基酸替换,所述位置根据Kabat中的EU索引编号。

[0288] 在一些实施方案中,将一个、两个或更多个氨基酸替换引入IgG恒定结构域Fc区中,以改变抗转铁蛋白受体抗体的效应物功能。对其的亲和力被改变的效应配体可以是例如Fc受体或补体的C1组分。该方法在美国专利No.5,624,821和5,648,260中有更详细的描述。在一些实施方案中,恒定区结构域的缺失或失活(通过点突变或其他方式)可降低循环抗体的Fc受体结合,从而提高肿瘤定位。对于使恒定结构域缺失或失活从而提高肿瘤定位的突变的描述,参见例如美国专利No.5,585,097和8,591,886。在一些实施方案中,可将一个或更多个氨基酸替换引入本文中所述的抗体的Fc区中,以去除Fc区上潜在的糖基化位点,这可降低Fc受体结合(参见,例如Shields R L et al.,(2001)J Biol Chem 276:6591-604)。

[0289] 在一些实施方案中,可将本文中所述的肌肉靶向抗体的恒定区中的一个或更多个氨基替换为不同的氨基酸残基,使得抗体具有改变的C1q结合和/或者(例如,和)降低或消除的补体依赖性细胞毒性(CDC)。这种方法在美国专利No.6,194,551(Idusogie et al)中有更详细的描述。在一些实施方案中,改变本文中所述的抗体的CH2结构域的N端区域中的一个或更多个氨基酸残基,从而改变抗体的固定补体的能力。这种方法在国际公开No.WO 94/29351中有进一步描述。在一些实施方案中,对本文中所述的抗体的Fc区进行修饰以提高抗体的介导抗体依赖性细胞毒性(ADCC)的能力和/或(例如,和)提高抗体对Fcγ受体的亲和力。这种方法在国际公开No.WO 00/42072中有进一步描述。

[0290] 在一些实施方案中,本文中提供的抗体的重链和/或(例如,和)轻链可变结构域序列可用于产生例如CDR接枝、嵌合、人源化或复合的人抗体或抗原结合片段,如本文中其他地方所述。如本领域普通技术人员所理解的,来源于本文中提供的任何抗体的任何变体(CDR接枝、嵌合、人源化或复合的抗体)可用于本文中所述的组合物和方法中,并将保持特

异性结合转铁蛋白受体的能力,从而使得相对于其所来源的原始抗体,变体(CDR接枝、嵌合、人源化或复合的抗体)具有至少50%、至少60%、至少70%、至少80%、至少90%、至少95%或更多的与转铁蛋白受体的结合。

[0291] 在一些实施方案中,本文中提供的抗体包含赋予抗体以期望性质的突变。例如,为了避免归因于已知与天然IgG4 mAb发生的Fab臂交换而引起的潜在并发症,本文中提供的抗体可包含稳定性'Adair'突变(Angal S.,et al.,"Asingle amino acid substitution abolishes the heterogeneity of chimeric mouse/human(IgG4)antibody,"Mol Immunol 30,105-108;1993),其中第228位(EU编号;根据Kabat编号为第241位残基)丝氨酸转化为脯氨酸,从而产生了IgG1样铰链序列。因此,任何抗体都可包含稳定性'Adair'突变。

[0292] 如本文中所提供,本公开内容的抗体可任选地包含恒定区或其一部分。例如,VL结构域可在其C末端连接至轻链恒定结构域,如Cκ或Cλ。类似地,VH结构域或其一部分可连接至如IgA、IgD、IgE、IgG和IgM(以及任何同种型亚类)的重链的全部或一部分。抗体可包括合适的恒定区(参见,例如,Kabat et al.,Sequences of Proteins of Immunological Interest,No.91-3242,National Institutes of Health Publications,Bethesda,Md.(1991))。因此,在本公开内容范围内的抗体可包含与任何合适的恒定区组合的VH和VL结构域或其抗原结合部分。

[0293] ii.肌肉靶向肽

[0294] 本公开内容的一些方面提供了肌肉靶向肽作为肌肉靶向剂。已描述了与特定细胞类型结合的短肽序列(例如,长度为5至20个氨基酸的肽序列)。例如,细胞靶向肽已在以下中进行了描述:Vines e.,et al.,A."Cell-penetrating and cell-targeting peptides in drug delivery"Biochim Biophys Acta 2008,1786:126-38;Jarver P.,et al.,"In vivo biodistribution and efficacy of peptide mediated delivery"Trends Pharmacol Sci 2010;31:528-35;Samoylova T.I.,et al.,"Elucidation of muscle-binding peptides by phage display screening"Muscle Nerve 1999;22:460-6;美国专利No.6,329,501,其于2001年12月11日授权,标题为"METHODS AND COMPOSITIONS FOR TARGETING COMPOUNDS TO MUSCLE";以及Samoylov A.M.,et al.,"Recognition of cell-specific binding of phage display derived peptides using an acoustic wave sensor."Biomol Eng 2002;18:269-72;其各自的全部内容均通过引用并入本文。通过设计肽与特定的细胞表面抗原(例如,受体)相互作用,可实现对所期望组织例如肌肉的选择性。已研究了骨骼肌靶向并且能够递送一系列分子载荷。这些方法可对肌肉组织具有高度选择性,而没有大的抗体或病毒颗粒的许多实际缺点。因此,在一些实施方案中,肌肉靶向剂是长度为4至50个氨基酸的肌肉靶向肽。在一些实施方案中,肌肉靶向肽的长度为4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49或50个氨基酸。可使用数种方法中的任一种(例如噬菌体展示)来产生肌肉靶向肽。

[0295] 在一些实施方案中,肌肉靶向肽可与和某些其他细胞相比在肌细胞中过表达或相对高地表达的内化细胞表面受体,例如转铁蛋白受体结合。在一些实施方案中,肌肉靶向肽可靶向转铁蛋白受体(例如,与之结合)。在一些实施方案中,靶向转铁蛋白受体的肽可包含天然存在配体(例如转铁蛋白)的区段。在一些实施方案中,靶向转铁蛋白受体的肽如2000

年11月30日提交的美国专利No.6,743,893,“RECEPTOR-MEDIATED UPTAKE OF PEPTIDES THAT BIND THE HUMAN TRANSFERRIN RECEPTOR”中所述。在一些实施方案中,靶向转铁蛋白受体的肽如Kawamoto,M.et al,“A novel transferrin receptor-targeted hybrid peptide disintegrates cancer cell membrane to induce rapid killing of cancer cells.”BMC Cancer.2011Aug 18;11:359中所述。在一些实施方案中,靶向转铁蛋白受体的肽如2011年5月20日提交的美国专利No.8,399,653,“TRANSFERRIN/TRANSFERRIN RECEPTOR-MEDIATED SIRNA DELIVERY”中所述。

[0296] 如上所讨论,已报道了肌肉靶向肽的一些实例。例如,使用呈递表面七肽的噬菌体展示文库鉴定了肌肉特异性肽。作为一个实例,具有氨基酸序列ASSLNIA(SEQ ID NO:943)的肽在体外与C2C12鼠肌管结合,并且在体内与小鼠肌肉组织结合。因此,在一些实施方案中,肌肉靶向剂包含氨基酸序列ASSLNIA(SEQ ID NO:943)。该肽在小鼠中进行静脉内注射之后展示出提高的与心脏和骨骼肌组织结合的特异性,以及降低的与肝、肾和脑的结合。使用噬菌体展示已鉴定了另外的肌肉特异性肽。例如,通过噬菌体展示文库鉴定了12个氨基酸的肽用于在迪谢内肌营养不良治疗的情况下进行肌肉靶向。参见Yoshida D.,et al.,“Targeting of salicylate to skin and muscle following topical injections in rats.”Int J Pharm 2002;231:177-84;其全部内容在此通过引用并入。在此,鉴定了具有序列SKTFNTHPQSTP(SEQ ID NO:944)的12个氨基酸的肽,并且该肌肉靶向肽相对于ASSLNIA(SEQ ID NO:943)肽显示出提高的与C2C12细胞的结合。

[0297] 用于鉴定相对于其他细胞类型而对肌肉(例如,骨骼肌)具有选择性的肽的另一方法包括体外选择,这已在Ghosh D.,et al.,“Selection of muscle-binding peptides from context-specific peptide-presenting phage libraries for adenoviral vector targeting”J Virol 2005;79:13667-72中进行了描述;其全部内容通过引用并入本文。通过将随机的12聚体(12-mer)肽噬菌体展示文库与非肌细胞类型的混合物预孵育选择出了非特异性细胞结合物。在数轮选择之后,12个氨基酸的肽TARGEHKEEELI(SEQ ID NO:945)最频繁地出现。因此,在一些实施方案中,肌肉靶向剂包含氨基酸序列TARGEHKEEELI(SEQ ID NO:945)。

[0298] 肌肉靶向剂可以是含氨基酸的分子或肽。肌肉靶向肽可对应于优先与肌细胞中存在的蛋白质受体结合的蛋白质序列。在一些实施方案中,肌肉靶向肽包含高倾向性的疏水性氨基酸,例如缬氨酸,使得该肽优先靶向肌细胞。在一些实施方案中,肌肉靶向肽是先前未表征或公开过的。可使用数种方法中的任一种(例如噬菌体展示肽文库、单珠单化合物肽文库或位置扫描合成肽组合文库)来构思、产生、合成和/或(例如,和)衍生这些肽。示例性方法已在本领域中表征并通过引用并入(Gray,B.P.and Brown,K.C.“Combinatorial Peptide Libraries:Mining for Cell-Binding Peptides”Chem Rev.2014,114:2,1020-1081.;Samoylova,T.I.and Smith,B.F.“Elucidation of muscle-binding peptides by phage display screening.”Muscle Nerve,1999,22:4.460-6.)。在一些实施方案中,先前已公开了肌肉靶向肽(参见,例如Writer M.J.et al.“Targeted gene delivery to human airway epithelial cells with synthetic vectors incorporating novel targeting peptides selected by phage display.”J.Drug Targeting.2004;12:185;Cai,D.“BDNF-mediated enhancement of inflammation and injury in the aging heart.”Physiol

Genomics.2006,24:3,191-7.;Zhang,L.“Molecular profiling of heart endothelial cells.”Circulation,2005,112:11,1601-11.;McGuire,M.J.et al.“In vitro selection of a peptide with high selectivity for cardiomyocytes in vivo.”J Mol Biol.2004,342:1,171-82.)。示例性的肌肉靶向肽包含以下组的氨基酸序列:CQAQGQLVC(SEQ ID NO:946),CSERSMNFC(SEQ ID NO:947),CPKTRRVPC(SEQ ID NO:948),WLSEAGPVVTVRALRGTGSW(SEQ ID NO:949),ASSLNIA(SEQ ID NO:943),CMQHSMRVC(SEQ ID NO:950),和DDTRHWG(SEQ ID NO:951)。在一些实施方案中,肌肉靶向肽可包含约2至25个氨基酸、约2至20个氨基酸、约2至15个氨基酸、约2至10个氨基酸或约2至5个氨基酸。肌肉靶向肽可包含天然存在的氨基酸例如半胱氨酸、丙氨酸或者非天然存在的或经修饰的氨基酸。非天然存在的氨基酸包括β-氨基酸、高氨基酸(homo-amino acid)、脯氨酸衍生物、3-经取代的丙氨酸衍生物、线性核心氨基酸、N-甲基氨基酸和本领域中已知的其他氨基酸。在一些实施方案中,肌肉靶向肽可以是线性的;在另一些实施方案中,肌肉靶向肽可以是环状的,例如双环的(参见,例如Silvana,M.G.et al.Mol.Therapy,2018,26:1,132-147.)。

[0299] iii.肌肉靶向受体配体

[0300] 肌肉靶向剂可以是配体,例如与受体蛋白结合的配体。肌肉靶向配体可以是蛋白质,例如与由肌细胞表达的内化细胞表面受体结合的转铁蛋白。因此,在一些实施方案中,肌肉靶向剂是转铁蛋白或与转铁蛋白受体结合的转铁蛋白衍生物。肌肉靶向配体可替代地是小分子,例如相对于其他细胞类型而优先靶向肌细胞的亲脂性小分子。可靶向肌细胞的一些示例性亲脂性小分子包括包含以下的化合物:胆固醇、胆固醇基、硬脂酸、棕榈酸、油酸、油烯基、亚麻烯(linolene)、亚油酸、肉豆蔻酸、甾醇类、二氢睾酮、睾酮衍生物、甘油、烷基链、三苯甲基类和烷氧基酸。

[0301] iv.肌肉靶向适配体

[0302] 肌肉靶向剂可以是适配体,例如RNA适配体,其相对于其他细胞类型而优先靶向肌细胞。在一些实施方案中,肌肉靶向适配体是先前未表征或公开过的。可使用数种方法中的任一种(例如通过指数富集的配体的系统进化)来构思、产生、合成和/或(例如,和)衍生这些适配体。示例性方法已在本领域中表征并通过引用并入(Yan,A.C.and Levy,M.“Aptamers and aptamer targeted delivery”RNA biology,2009,6:3,316-20.;Germer,K.et al.“RNA aptamers and their therapeutic and diagnostic applications.”Int.J.Biochem.Mol.Biol.2013;4:27-40.)。在一些实施方案中,先前已公开了肌肉靶向适配体(参见,例如Phillippou,S.et al.“Selection and Identification of Skeletal-Muscle-Targeted RNA Aptamers.”Mol Ther Nucleic Acids.2018,10:199-214.;Thiel,W.H.et al.“Smooth Muscle Cell-targeted RNA Aptamer Inhibits Neointimal Formation.”Mol Ther.2016,24:4,779-87.)。示例性的肌肉靶向适配体包括A01B RNA适配体和RNA Apt 14。在一些实施方案中,适配体是基于核酸的适配体、寡核苷酸适配体或肽适配体。在一些实施方案中,适配体可以是约5kDa至15kDa、约5kDa至10kDa、约10kDa至15kDa、约1Da至5Da、约1kDa至3kDa或更小。

[0303] v.其他肌肉靶向剂

[0304] 用于靶向肌细胞(例如,骨骼肌细胞)的一种策略是使用肌转运体蛋白(例如在肌膜上表达的转运体蛋白)的底物。在一些实施方案中,肌肉靶向剂是对肌肉组织具有特异性

的流入转运体的底物。在一些实施方案中,流入转运体对骨骼肌组织具有特异性。两类主要的转运体在骨骼肌肌膜上表达:(1)三磷酸腺苷(ATP)结合盒(ABC)超家族,其促进从骨骼肌组织流出和(2)溶质运载体(SLC)超家族,其可促进底物流入骨骼肌中。在一些实施方案中,肌肉靶向剂是与转运体的ABC超家族或SLC超家族结合的底物。在一些实施方案中,与转运体的ABC或SLC超家族结合的底物是天然存在的底物。在一些实施方案中,与转运体的ABC或SLC超家族结合的底物是非天然存在的底物,例如,与转运体的ABC或SLC超家族结合的其合成衍生物。

[0305] 在一些实施方案中,肌肉靶向剂是本文中所述的靶向转运体的SLC超家族的任何肌肉靶向剂(例如,抗体、核酸、小分子、肽、适配体、脂质、糖部分)。在一些实施方案中,肌肉靶向剂是转运体的SLC超家族的底物。SLC转运体是平衡型的,或者使用跨膜而产生的质子或钠离子梯度来驱动底物的转运。具有高骨骼肌表达的示例性SLC转运体包括但不限于SATT转运体(ASCT1;SLC1A4)、GLUT4转运体(SLC2A4)、GLUT7转运体(GLUT7;SLC2A7)、ATRC2转运体(CAT-2;SLC7A2)、LAT3转运体(KIAA0245;SLC7A6)、PHT1转运体(PTR4;SLC15A4)、OATP-J转运体(OATP5A1;SLC21A15)、OCT3转运体(EMT;SLC22A3)、OCTN2转运体(FLJ46769;SLC22A5)、ENT转运体(ENT1;SLC29A1和ENT2;SLC29A2)、PAT2转运体(SLC36A2)和SAT2转运体(KIAA1382;SLC38A2)。这些转运体可促进底物流入骨骼肌中,为肌肉靶向提供机会。

[0306] 在一些实施方案中,肌肉靶向剂是平衡型核苷转运体2(equilibrative nucleoside transporter 2,ENT2)转运体的底物。相对于其他转运体,ENT2在骨骼肌中具有最高的mRNA表达之一。虽然人ENT2(hENT2)在大多数身体器官例如脑、心脏、胎盘、胸腺、胰腺、前列腺和肾中表达,但其在骨骼肌中特别丰富。人ENT2根据其浓度梯度促进其底物的吸收。ENT2通过转运广泛范围的嘌呤和嘧啶核苷碱基在维持核苷稳态中发挥作用。hENT2转运体对除肌苷之外的所有核苷(腺苷、鸟苷、尿苷、胸苷和胞苷)均具有低的亲和力。因此,在一些实施方案中,肌肉靶向剂是ENT2底物。示例性的ENT2底物包括但不限于肌苷、2',3'-二脱氧肌苷和氯法拉滨(calofarabine)。在一些实施方案中,本文中提供的任何肌肉靶向剂均与分子载荷(例如,寡核苷酸载荷)缔合。在一些实施方案中,肌肉靶向剂与分子载荷共价连接。在一些实施方案中,肌肉靶向剂与分子载荷非共价连接。

[0307] 在一些实施方案中,肌肉靶向剂是有机阳离子/肉碱转运体(OCTN2)的底物,所述有机阳离子/肉碱转运体是钠离子依赖性的高亲和力肉碱转运体。在一些实施方案中,肌肉靶向剂是肉碱、米屈肼(mildronate)、乙酰肉碱或与OCTN2结合的其任意衍生物。在一些实施方案中,肉碱、米屈肼、乙酰肉碱或其衍生物与分子载荷(例如,寡核苷酸载荷)共价连接。

[0308] 肌肉靶向剂可以是蛋白质,该蛋白质是以靶向肌细胞的至少一种可溶性形式存在的蛋白质。在一些实施方案中,肌肉靶向蛋白可以是血幼素(也称为排斥性导向分子C或血色素沉着症2型蛋白),所述血幼素是参与铁超负荷和稳态的蛋白。在一些实施方案中,血幼素可以是全长或片段,或者与功能性血幼素蛋白具有至少75%、至少80%、至少85%、至少90%、至少95%、至少98%或至少99%序列同一性的突变体。在一些实施方案中,血幼素突变体可以是可溶性片段,可缺乏N端信号传导,和/或(例如,和)缺乏C端锚定结构域。在一些实施方案中,血幼素可以以GenBank RefSeq登录号NM_001316767.1、NM_145277.4、NM_202004.3、NM_213652.3或NM_213653.3进行注释。应理解,血幼素可以是人、非人灵长类或啮齿动物来源的。

[0309] B.分子载荷

[0310] 本公开内容的一些方面提供了分子载荷,例如用于调节生物学结局,例如DNA序列的转录、RNA序列的剪接和加工、蛋白质的表达或蛋白质的活性。在一些实施方案中,分子载荷与肌肉靶向剂连接或以其他方式缔合。在一些实施方案中,这样的分子载荷能够靶向肌细胞,例如在通过缔合的肌肉靶向剂递送至肌细胞之后通过与肌细胞中的核酸或蛋白质特异性结合。应理解,根据本公开内容可使用多种类型的分子载荷。例如,分子载荷可包含以下或由以下组成:寡核苷酸(例如,反义寡核苷酸)、肽(例如,结合肌细胞中与疾病相关的核酸或蛋白质的肽)、蛋白质(例如,结合肌细胞中与疾病相关的核酸或蛋白质的蛋白质)或小分子(例如,调节肌细胞中与疾病相关的核酸或蛋白质之功能的小分子)。在一些实施方案中,分子载荷是包含具有突变DMD等位基因的互补区的链的寡核苷酸。本文中进一步详细描述了示例性分子载荷,然而,应理解,本文中提供的示例性分子载荷并不意味着是限制性的。

[0311] i.寡核苷酸

[0312] 本公开内容的一些方面涉及被配置成调节(例如,提高)肌养蛋白(例如,来自DMD等位基因的)表达的寡核苷酸。在一些实施方案中,本文中提供的寡核苷酸被配置成改变DMD前体mRNA的剪接以促进肌养蛋白(例如,功能性截短的肌养蛋白)的表达。在一些实施方案中,本文中提供的寡核苷酸被配置成促进DMD中(例如突变DMD等位基因中)的一个或更多个外显子的跳读,以恢复阅读框。在一些实施方案中,寡核苷酸允许功能性肌养蛋白表达(例如,如Kinali M,Arechevala-Gomeza V,Feng L,et al.Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy:a single-blind,placebo-controlled,dose-escalation,proof-of-concept study.Lancet Neurol.2009;8(10):918-928和Watanabe N,Nagata T,Satou Y,et al.NS-065/NCNP-01:an antisense oligonucleotide for potential treatment of exon53 skipping in Duchenne muscular dystrophy.Mol Ther Nucleic Acids.2018;13:442-449中所述)。在一些实施方案中,提供的寡核苷酸被配置成促进外显子51的跳读以产生肌养蛋白的较短但是功能性的形式(例如,包含框内缺失)。在一些实施方案中,提供了促进外显子51跳读的寡核苷酸(例如,其可与大量患者相关,所述患者包括例如适合于外显子51跳读的患者,例如在DMD外显子3至50、4至50、5至50、6至50、9至50、10至50、11至50、13至50、14至50、15至50、16至50、17至50、19至50、21至50、23至50、24至50、25至50、26至50、27至50、28至50、29至50、30至50、31至50、32至50、33至50、34至50、35至50、36至50、37至50、38至50、39至50、40至50、41至50、42至50、43至50、45至50、47至50、48至50、49至50、50、52、52至58、52至61、52至63、52至64、52至66、52至76或52至77中具有缺失的那些患者)。

[0313] 表8提供了可用于靶向DMD(例如用于外显子跳读和用于DMD内的靶序列)的寡核苷酸序列的一些非限制性实例。在一些实施方案中,寡核苷酸可包含表8中提供的任何反义序列或表8中提供的靶序列的互补序列。

[0314] 表8.用于靶向DMD的寡核苷酸序列。

[0315]

SEQ ID NO	靶序列 † (5'至3')	SEQ ID NO	反义序列 † (5'至3')	SEQ ID NO	反义序列 † (5'至3')	靶位点
160	GUAAGUAUACUGG AUCCCAUUC	384	GAAUGGGAUCCAG UAUACUUAC	608	GAATGGGATCCAG TATACTTAC	内含子50
161	GUAAGUAUACUGG AUCCCAUUCU	385	AGAAUGGGAUCCA GUAUACUUAC	609	AGAATGGGATCCA GTATACTTAC	内含子50
162	GUAAGUAUACUGG AUCCCAUUCUC	386	GAGAAUGGGAUCC AGUAUACUUAC	610	GAGAATGGGATCC AGTATACTTAC	内含子50

[0316]

163	GUAAGUAUACUGG AUCCCAUUCUCU	387	AGAGAAUGGGAUC CAGUAUACUUAC	611	AGAGAATGGGATC CAGTATACTTAC	内含子50
164	UAAGUAUACUGGA UCCCAUUCUC	388	GAGAAUGGGAUCC AGUAUACUUA	612	GAGAATGGGATCC AGTATACTTA	内含子50
165	AAGUAUACUGGAU CCCAUUCUC	389	GAGAAUGGGAUCC AGUAUACUU	613	GAGAATGGGATCC AGTATACTT	内含子50
166	AGUAUACUGGAUC CCAUUCUC	390	GAGAAUGGGAUCC AGUAUACU	614	GAGAATGGGATCC AGTATACT	内含子50
167	GUAUACUGGAUCC CAUUCUCUUUGG	391	CCAAAGAGAAUGG GAUCCAGUAUAC	615	CCAAAGAGAATGG GATCCAGTATAC	内含子50
168	UACUGGAUCCCAU UCUCUUUGGCUC	392	GAGCCAAAGAGAA UGGGAUCCAGUA	616	GAGCCAAAGAGAA TGGGATCCAGTA	内含子50
169	ACUGGAUCCCAUU CUCUUUGGCUC	393	GAGCCAAAGAGAA UGGGAUCCAGU	617	GAGCCAAAGAGAA TGGGATCCAGT	内含子50
170	UGUGGUUACUAAG GAAACUGCCAU	394	AUGGCAGUUUCCU UAGUAACCACA	618	ATGGCAGTTTCCT TAGTAACCACA	外显子51
171	UGUGGUUACUAAG GAAACUGCCAUC	395	GAUGGCAGUUUCC UUAGUAACCACA	619	GATGGCAGTTTCC TTAGTAACCACA	外显子51
172	GUGGUUACUAAGG AAACUGCCAU	396	AUGGCAGUUUCCU UAGUAACCAC	620	ATGGCAGTTTCCT TAGTAACCAC	外显子51
173	GUGGUUACUAAGG AAACUGCCAUC	397	GAUGGCAGUUUCC UUAGUAACCAC	621	GATGGCAGTTTCC TTAGTAACCAC	外显子51
174	GUGGUUACUAAGG AAACUGCCAUCU	398	AGAUGGCAGUUUC CUUAGUAACCAC	622	AGATGGCAGTTTC CTTAGTAACCAC	外显子51
175	UGGUUACUAAGGA AACUGCCAUC	399	GAUGGCAGUUUCC UUAGUAACCA	623	GATGGCAGTTTCC TTAGTAACCA	外显子51
176	GGUUACUAAGGAA ACUGCCAUC	400	GAUGGCAGUUUCC UUAGUAACC	624	GATGGCAGTTTCC TTAGTAACC	外显子51
177	GAAACUGCCAUCU CCAAACUAGAA	401	UUCUAGUUUGGAG AUGGCAGUUUC	625	TTCTAGTTTGGAG ATGGCAGTTTC	外显子51
178	AAACUGCCAUCUC CAAACUAG	402	CUAGUUUGGAGAU GGCAGUUU	626	CTAGTTTGGAGAT GGCAGTTT	外显子51
179	AAACUGCCAUCUC CAAACUAGA	403	UCUAGUUUGGAGA UGGCAGUUU	627	TCTAGTTTGGAGA TGGCAGTTT	外显子51
180	AAACUGCCAUCUC CAAACUAGAA	404	UUCUAGUUUGGAG AUGGCAGUUU	628	TTCTAGTTTGGAG ATGGCAGTTT	外显子51
181	AACUGCCAUCUCC AAACUAG	405	CUAGUUUGGAGAU GGCAGUU	629	CTAGTTTGGAGAT GGCAGTT	外显子51
182	AACUGCCAUCUCC AAACUAGA	406	UCUAGUUUGGAGA UGGCAGUU	630	TCTAGTTTGGAGA TGGCAGTT	外显子51
183	AACUGCCAUCUCC AAACUAGAA	407	UUCUAGUUUGGAG AUGGCAGUU	631	TTCTAGTTTGGAG ATGGCAGTT	外显子51
184	ACUGCCAUCUCCA AACUAGA	408	UCUAGUUUGGAGA UGGCAGU	632	TCTAGTTTGGAGA TGGCAGT	外显子51
185	ACUGCCAUCUCCA AACUAGAA	409	UUCUAGUUUGGAG AUGGCAGU	633	TTCTAGTTTGGAG ATGGCAGT	外显子51
186	UCUCCAAACUAGA AAUGCCAUC	410	GAUGGCAUUUCUA GUUUGGAGA	634	GATGGCATTTCTA GTTTGGAGA	外显子51
187	CUCCAAACUAGAA AUGCCAUC	411	GAUGGCAUUUCUA GUUUGGAG	635	GATGGCATTTCTA GTTTGGAG	外显子51
188	UCCAAACUAGAAA UGCCAUC	412	GAUGGCAUUUCUA GUUUGGA	636	GATGGCATTTCTA GTTTGGA	外显子51
189	GAUUUCAACCGGG CUUGGACAGA	413	UCUGUCCAAGCCC GGUUGAAAUC	637	TCTGTCCAAGCCC GGTTGAAATC	外显子51
190	GAUUUCAACCGGG CUUGGACAGAA	414	UUCUGUCCAAGCC CGGUUGAAAUC	638	TTCTGTCCAAGCC CGGTTGAAATC	外显子51
191	AUUUCAACCGGGC UUGGACAGA	415	UCUGUCCAAGCCC GGUUGAAAU	639	TCTGTCCAAGCCC GGTTGAAAT	外显子51
192	AUUUCAACCGGGC UUGGACAGAACU	416	AGUUCUGUCCAAG CCCGGUUGAAAU	640	AGTTCTGTCCAAG CCCGGTTGAAAT	外显子51
193	UUCAACCGGGCUU GGACAGAACU	417	AGUUCUGUCCAAG CCCGGUUGAA	641	AGTTCTGTCCAAG CCCGGTTGAA	外显子51

[0317]

194	UCAACCGGGCUUGGACAGAA	418	UUCUGUCCAAGCCCGGUUGA	642	TTCTGTCCAAGCCCGGTTGA	外显子51
195	UCAACCGGGCUUGGACAGAACU	419	AGUUCUGUCCAAGCCCGGUUGA	643	AGTTCTGTCCAAGCCCGGTTGA	外显子51
196	UCAACCGGGCUUGGACAGAACUUAC	420	GUAAGUUCUGUCCAAGCCCGGUUGA	644	GTAAGTTCTGTCCAAGCCCGGTTGA	外显子51
197	CAACCGGGCUUGGACAGAACU	421	AGUUCUGUCCAAGCCCGGUUG	645	AGTTCTGTCCAAGCCCGGTTG	外显子51
198	CAACCGGGCUUGGACAGAACUUAC	422	GUAAGUUCUGUCCAAGCCCGGUUG	646	GTAAGTTCTGTCCAAGCCCGGTTG	外显子51
199	CAACCGGGCUUGGACAGAACUUACC	423	GGUAAGUUCUGUCCAAGCCCGGUUG	647	GGTAAGTTCTGTCCAAGCCCGGTTG	外显子51
200	AUGAUCAUCAAGCAGAAGGUA	424	UACCUUCUGCUUGAUGAUCAU	648	TACCTTCTGCTTGATGATCAT	外显子51/内含子51连接
201	AUGAUCAUCAAGCAGAAGGUAUG	425	CAUACCUUCUGCUUGAUGAUCAU	649	CATACCTTCTGCTTGATGATCAT	外显子51/内含子51连接
202	AUGAUCAUCAAGCAGAAGGUAUGA	426	UCAUACCUUCUGCUUGAUGAUCAU	650	TCATACCTTCTGCTTGATGATCAT	外显子51/内含子51连接
203	AUGAUCAUCAAGCAGAAGGUAUGAG	427	CUCAUACCUUCUGCUUGAUGAUCAU	651	CTCATACCTTCTGCTTGATGATCAT	外显子51/内含子51连接
204	UGAUCAUCAAGCAGAAGGUA	428	UACCUUCUGCUUGAUGAUCA	652	TACCTTCTGCTTGATGATCA	外显子51/内含子51连接
205	UGAUCAUCAAGCAGAAGGUAUG	429	CAUACCUUCUGCUUGAUGAUCA	653	CATACCTTCTGCTTGATGATCA	外显子51/内含子51连接
206	UGAUCAUCAAGCAGAAGGUAUGA	430	UCAUACCUUCUGCUUGAUGAUCA	654	TCATACCTTCTGCTTGATGATCA	外显子51/内含子51连接
207	UGAUCAUCAAGCAGAAGGUAUGAG	431	CUCAUACCUUCUGCUUGAUGAUCA	655	CTCATACCTTCTGCTTGATGATCA	外显子51/内含子51连接
208	UGAUCAUCAAGCAGAAGGUAUGAGA	432	UCUCAUACCUUCUGCUUGAUGAUCA	656	TCTCATACCTTCTGCTTGATGATCA	外显子51/内含子51连接
209	GAUCAUCAAGCAGAAGGUAUG	433	CAUACCUUCUGCUUGAUGAUC	657	CATACCTTCTGCTTGATGATC	外显子51/内含子51连接
210	GAUCAUCAAGCAGAAGGUAUGA	434	UCAUACCUUCUGCUUGAUGAUC	658	TCATACCTTCTGCTTGATGATC	外显子51/内含子51连接
211	GAUCAUCAAGCAGAAGGUAUGAG	435	CUCAUACCUUCUGCUUGAUGAUC	659	CTCATACCTTCTGCTTGATGATC	外显子51/内含子51连接
212	GAUCAUCAAGCAGAAGGUAUGAGA	436	UCUCAUACCUUCUGCUUGAUGAUC	660	TCTCATACCTTCTGCTTGATGATC	外显子51/内含子51连接
213	GAUCAUCAAGCAGAAGGUAUGAGAA	437	UUCUCAUACCUUCUGCUUGAUGAUC	661	TTCTCATACCTTCTGCTTGATGATC	外显子51/内含子51连接
214	AUCAUCAAGCAGAAGGUAUG	438	CAUACCUUCUGCUUGAUGAU	662	CATACCTTCTGCTTGATGAT	外显子51/内含子51连接
215	AUCAUCAAGCAGAAGGUAUGA	439	UCAUACCUUCUGCUUGAUGAU	663	TCATACCTTCTGCTTGATGAT	外显子51/内含子51连接
216	AUCAUCAAGCAGAAGGUAUGAG	440	CUCAUACCUUCUGCUUGAUGAU	664	CTCATACCTTCTGCTTGATGAT	外显子51/内含子51连接
217	AUCAUCAAGCAGAAGGUAUGAGA	441	UCUCAUACCUUCUGCUUGAUGAU	665	TCTCATACCTTCTGCTTGATGAT	外显子51/内含子51连接
218	AUCAUCAAGCAGAAGGUAUGAGAA	442	UUCUCAUACCUUCUGCUUGAUGAU	666	TTCTCATACCTTCTGCTTGATGAT	外显子51/内含子51连接
219	AUCAUCAAGCAGAAGGUAUGAGAAA	443	UUUCUCAUACCUUCUGCUUGAUGAU	667	TTTCTCATACCTTCTGCTTGATGAT	外显子51/内含子51连接
220	UCAUCAAGCAGAAGGUAUGA	444	UCAUACCUUCUGCUUGAUGA	668	TCATACCTTCTGCTTGATGA	外显子51/内含子51连接
221	UCAUCAAGCAGAAGGUAUGAG	445	CUCAUACCUUCUGCUUGAUGA	669	CTCATACCTTCTGCTTGATGA	外显子51/内含子51连接
222	UCAUCAAGCAGAAGGUAUGAGA	446	UCUCAUACCUUCUGCUUGAUGA	670	TCTCATACCTTCTGCTTGATGA	外显子51/内含子51连接
223	UCAUCAAGCAGAAGGUAUGAGAA	447	UUCUCAUACCUUCUGCUUGAUGA	671	TTCTCATACCTTCTGCTTGATGA	外显子51/内含子51连接

[0318]

224	UCAUCAAGCAGAAGGUAUGAGAAA	448	UUUCUCAUACCUUCUGCUUGAUGA	672	TTTCTCATACCTTCTGCTTGATGA	外显子51/内含子51连接
225	UCAUCAAGCAGAAGGUAUGAGAAAA	449	UUUUCUCAUACCUUCUGCUUGAUGA	673	TTTTCTCATACCTTCTGCTTGATGA	外显子51/内含子51连接
226	CAUCAAGCAGAAGGUAUGAG	450	CUCAUACCUUCUGCUUGAUG	674	CTCATACCTTCTGCTTGATG	外显子51/内含子51连接
227	CAUCAAGCAGAAGGUAUGAGA	451	UCUCAUACCUUCUGCUUGAUG	675	TCTCATACCTTCTGCTTGATG	外显子51/内含子51连接
228	CAUCAAGCAGAAGGUAUGAGAA	452	UUCUCAUACCUUCUGCUUGAUG	676	TTCTCATACCTTCTGCTTGATG	外显子51/内含子51连接
229	CAUCAAGCAGAAGGUAUGAGAAA	453	UUUCUCAUACCUUCUGCUUGAUG	677	TTTCTCATACCTTCTGCTTGATG	外显子51/内含子51连接
230	CAUCAAGCAGAAGGUAUGAGAAAA	454	UUUUCUCAUACCUUCUGCUUGAUG	678	TTTTCTCATACCTTCTGCTTGATG	外显子51/内含子51连接
231	CAUCAAGCAGAAGGUAUGAGAAAAA	455	UUUUUCUCAUACCUUCUGCUUGAUG	679	TTTTTCTCATACCTTCTGCTTGATG	外显子51/内含子51连接
232	AUCAAGCAGAAGGUAUGAGA	456	UCUCAUACCUUCUGCUUGAU	680	TCTCATACCTTCTGCTTGAT	外显子51/内含子51连接
233	AUCAAGCAGAAGGUAUGAGAA	457	UUCUCAUACCUUCUGCUUGAU	681	TTCTCATACCTTCTGCTTGAT	外显子51/内含子51连接
234	AUCAAGCAGAAGGUAUGAGAAA	458	UUUCUCAUACCUUCUGCUUGAU	682	TTTCTCATACCTTCTGCTTGAT	外显子51/内含子51连接
235	AUCAAGCAGAAGGUAUGAGAAAA	459	UUUUCUCAUACCUUCUGCUUGAU	683	TTTTCTCATACCTTCTGCTTGAT	外显子51/内含子51连接
236	AUCAAGCAGAAGGUAUGAGAAAAA	460	UUUUUCUCAUACCUUCUGCUUGAU	684	TTTTTCTCATACCTTCTGCTTGAT	外显子51/内含子51连接
237	AUCAAGCAGAAGGUAUGAGAAAAAA	461	UUUUUUCUCAUACCUUCUGCUUGAU	685	TTTTTTCTCATACCTTCTGCTTGAT	外显子51/内含子51连接
238	UCAAGCAGAAGGUAUGAGAA	462	UUCUCAUACCUUCUGCUUGA	686	TTCTCATACCTTCTGCTTGA	外显子51/内含子51连接
239	UCAAGCAGAAGGUAUGAGAAA	463	UUUCUCAUACCUUCUGCUUGA	687	TTTCTCATACCTTCTGCTTGA	外显子51/内含子51连接
240	UCAAGCAGAAGGUAUGAGAAAA	464	UUUUCUCAUACCUUCUGCUUGA	688	TTTTCTCATACCTTCTGCTTGA	外显子51/内含子51连接
241	UCAAGCAGAAGGUAUGAGAAAAA	465	UUUUUCUCAUACCUUCUGCUUGA	689	TTTTTCTCATACCTTCTGCTTGA	外显子51/内含子51连接
242	UCAAGCAGAAGGUAUGAGAAAAAA	466	UUUUUUCUCAUACCUUCUGCUUGA	690	TTTTTTCTCATACCTTCTGCTTGA	外显子51/内含子51连接
243	UCAAGCAGAAGGUAUGAGAAAAAAU	467	AUUUUUUCUCAUACCUUCUGCUUGA	691	ATTTTTTCTCATACCTTCTGCTTGA	外显子51/内含子51连接
244	CAAGCAGAAGGUAUGAGAAA	468	UUUCUCAUACCUUCUGCUUG	692	TTTCTCATACCTTCTGCTTG	外显子51/内含子51连接
245	CAAGCAGAAGGUAUGAGAAAA	469	UUUUCUCAUACCUUCUGCUUG	693	TTTTCTCATACCTTCTGCTTG	外显子51/内含子51连接
246	CAAGCAGAAGGUAUGAGAAAAA	470	UUUUUCUCAUACCUUCUGCUUG	694	TTTTTCTCATACCTTCTGCTTG	外显子51/内含子51连接
247	CAAGCAGAAGGUAUGAGAAAAAA	471	UUUUUUCUCAUACCUUCUGCUUG	695	TTTTTTCTCATACCTTCTGCTTG	外显子51/内含子51连接
248	CAAGCAGAAGGUAUGAGAAAAAAU	472	AUUUUUUCUCAUACCUUCUGCUUG	696	ATTTTTTCTCATACCTTCTGCTTG	外显子51/内含子51连接
249	AAGCAGAAGGUAUGAGAAAA	473	UUUUCUCAUACCUUCUGCUU	697	TTTTCTCATACCTTCTGCTT	外显子51/内含子51连接
250	AAGCAGAAGGUAUGAGAAAAA	474	UUUUUCUCAUACCUUCUGCUU	698	TTTTTCTCATACCTTCTGCTT	外显子51/内含子51连接
251	AAGCAGAAGGUAUGAGAAAAAA	475	UUUUUUCUCAUACCUUCUGCUU	699	TTTTTTCTCATACCTTCTGCTT	外显子51/内含子51连接
252	AAGCAGAAGGUAUGAGAAAAAAAU	476	AUUUUUUCUCAUACCUUCUGCUU	700	ATTTTTTCTCATACCTTCTGCTT	外显子51/内含子51连接
253	AGCAGAAGGUAUGAGAAAAA	477	UUUUUCUCAUACCUUCUGCU	701	TTTTTCTCATACCTTCTGCT	外显子51/内含子51连接

[0319]

254	AGCAGAAGGUAUG AGAAAAAA	478	UUUUUUCUCAUAC CUUCUGCU	702	TTTTTTCTCATAC CTTCTGCT	外显子51/内含子51连接
255	AGCAGAAGGUAUG AGAAAAAAU	479	AUUUUUUCUCAUA CCUUCUGCU	703	ATTTTTTCTCATA CCTTCTGCT	外显子51/内含子51连接
256	GCAGAAGGUAUGA GAAAAAA	480	UUUUUUCUCAUAC CUUCUGC	704	TTTTTTCTCATAC CTTCTGC	外显子51/内含子51连接
257	GCAGAAGGUAUGA GAAAAAAU	481	AUUUUUUCUCAUA CCUUCUGC	705	ATTTTTTCTCATA CCTTCTGC	外显子51/内含子51连接
258	CAGAAGGUAUGAG AAAAAAU	482	AUUUUUUCUCAUA CCUUCUG	706	ATTTTTTCTCATA CCTTCTG	外显子51/内含子51连接
259	AAAUGAUAAAAGU UGGCAGAAGU	483	ACUUCUGCCAACU UUUAUCAUUU	707	ACTTCTGCCAACT TTTATCATTT	内含子51
260	UCACUUUACUCUC CUAGACCAU	484	AUGGUCUAGGAGA GUAAAGUGA	708	ATGGTCTAGGAGA GTAAAGTGA	内含子51
261	UCACUUUACUCUC CUAGACCAUU	485	AAUGGUCUAGGAG AGUAAAGUGA	709	AATGGTCTAGGAG AGTAAAGTGA	内含子51
262	UCACUUUACUCUC CUAGACCAUUU	486	AAAUGGUCUAGGA GAGUAAAGUGA	710	AAATGGTCTAGGA GAGTAAAGTGA	内含子51
263	CACUUUACUCUCC UAGACCAUUUCC	487	GGAAAUGGUCUAG GAGAGUAAAGUG	711	GGAAATGGTCTAG GAGAGTAAAGTG	内含子51
264	ACUUUACUCUCCU AGACCAUUUCC	488	GGAAAUGGUCUAG GAGAGUAAAGU	712	GGAAATGGTCTAG GAGAGTAAAGT	内含子51
265	CUUUACUCUCCUA GACCAUUUCCCA	489	UGGGAAAUGGUCU AGGAGAGUAAAG	713	TGGGAAATGGTCT AGGAGAGTAAAG	内含子51
266	UUACUCUCCUAGA CCAUUUCCCA	490	UGGGAAAUGGUCU AGGAGAGUAA	714	TGGGAAATGGTCT AGGAGAGTAA	内含子51
267	UUACUCUCCUAGA CCAUUUCCCAC	491	GUGGGAAAUGGUC UAGGAGAGUAA	715	GTGGGAAATGGTC TAGGAGAGTAA	内含子51
268	UUACUCUCCUAGA CCAUUUCCCACC	492	GGUGGGAAAUGGU CUAGGAGAGUAA	716	GGTGGGAAATGGT CTAGGAGAGTAA	内含子51
269	UACUCUCCUAGAC CAUUUCCCA	493	UGGGAAAUGGUCU AGGAGAGUA	717	TGGGAAATGGTCT AGGAGAGTA	内含子51
270	UACUCUCCUAGAC CAUUUCCCAC	494	GUGGGAAAUGGUC UAGGAGAGUA	718	GTGGGAAATGGTC TAGGAGAGTA	内含子51
271	UACUCUCCUAGAC CAUUUCCCACC	495	GGUGGGAAAUGGU CUAGGAGAGUA	719	GGTGGGAAATGGT CTAGGAGAGTA	内含子51
272	UACUCUCCUAGAC CAUUUCCCACCA	496	UGGUGGGAAAUGG UCUAGGAGAGUA	720	TGGTGGGAAATGG TCTAGGAGAGTA	内含子51
273	ACUCUCCUAGACC AUUUCCCA	497	UGGGAAAUGGUCU AGGAGAGU	721	TGGGAAATGGTCT AGGAGAGT	内含子51
274	ACUCUCCUAGACC AUUUCCCAC	498	GUGGGAAAUGGUC UAGGAGAGU	722	GTGGGAAATGGTC TAGGAGAGT	内含子51
275	ACUCUCCUAGACC AUUUCCCACC	499	GGUGGGAAAUGGU CUAGGAGAGU	723	GGTGGGAAATGGT CTAGGAGAGT	内含子51
276	ACUCUCCUAGACC AUUUCCCACCA	500	UGGUGGGAAAUGG UCUAGGAGAGU	724	TGGTGGGAAATGG TCTAGGAGAGT	内含子51
277	ACUCUCCUAGACC AUUUCCCACCAG	501	CUGGUGGGAAAUG GUCUAGGAGAGU	725	CTGGTGGGAAATG GTCTAGGAGAGT	内含子51
278	CUCUCCUAGACCA UUUCCCA	502	UGGGAAAUGGUCU AGGAGAG	726	TGGGAAATGGTCT AGGAGAG	内含子51
279	CUCUCCUAGACCA UUUCCCAC	503	GUGGGAAAUGGUC UAGGAGAG	727	GTGGGAAATGGTC TAGGAGAG	内含子51
280	CUCUCCUAGACCA UUUCCCACC	504	GGUGGGAAAUGGU CUAGGAGAG	728	GGTGGGAAATGGT CTAGGAGAG	内含子51
281	CUCUCCUAGACCA UUUCCCACCA	505	UGGUGGGAAAUGG UCUAGGAGAG	729	TGGTGGGAAATGG TCTAGGAGAG	内含子51
282	CUCUCCUAGACCA UUUCCCACCAG	506	CUGGUGGGAAAUG GUCUAGGAGAG	730	CTGGTGGGAAATG GTCTAGGAGAG	内含子51
283	UCUCCUAGACCAU UUCCCAC	507	GUGGGAAAUGGUC UAGGAGA	731	GTGGGAAATGGTC TAGGAGA	内含子51
284	UCUCCUAGACCAU UUCCCACC	508	GGUGGGAAAUGGU CUAGGAGA	732	GGTGGGAAATGGT CTAGGAGA	内含子51

285	UCUCCUAGACCAU UUCCCACCA	509	UGGUGGGAAAUGG UCUAGGAGA	733	TGGTGGGAAATGG TCTAGGAGA	内含子51
286	UCUCCUAGACCAU UUCCCACCAG	510	CUGGUGGGAAAUG GUCUAGGAGA	734	CTGGTGGGAAATG GTCTAGGAGA	内含子51
287	UCUCCUAGACCAU UUCCCACCAGUU	511	AACUGGUGGGAAA UGGUCUAGGAGA	735	AACTGGTGGGAAA TGGTCTAGGAGA	内含子51
288	CUCCUAGACCAUU UCCCACC	512	GGUGGGAAAUGGU CUAGGAG	736	GGTGGGAAATGGT CTAGGAG	内含子51
289	CUCCUAGACCAUU UCCCACCA	513	UGGUGGGAAAUGG UCUAGGAG	737	TGGTGGGAAATGG TCTAGGAG	内含子51
290	CUCCUAGACCAUU UCCCACCAG	514	CUGGUGGGAAAUG GUCUAGGAG	738	CTGGTGGGAAATG GTCTAGGAG	内含子51
291	CUCCUAGACCAUU UCCCACCAGUU	515	AACUGGUGGGAAA UGGUCUAGGAG	739	AACTGGTGGGAAA TGGTCTAGGAG	内含子51
292	UCCUAGACCAUUU CCCACCA	516	UGGUGGGAAAUGG UCUAGGA	740	TGGTGGGAAATGG TCTAGGA	内含子51
293	UCCUAGACCAUUU CCCACCAG	517	CUGGUGGGAAAUG GUCUAGGA	741	CTGGTGGGAAATG GTCTAGGA	内含子51
294	UCCUAGACCAUUU CCCACCAGUU	518	AACUGGUGGGAAA UGGUCUAGGA	742	AACTGGTGGGAAA TGGTCTAGGA	内含子51
295	UCCUAGACCAUUU CCCACCAGUUCU	519	AGAACUGGUGGGA AAUGGUCUAGGA	743	AGAACTGGTGGGA AATGGTCTAGGA	内含子51
296	CCUAGACCAUUUC CCACCAG	520	CUGGUGGGAAAUG GUCUAGG	744	CTGGTGGGAAATG GTCTAGG	内含子51
297	CCUAGACCAUUUC CCACCAGUU	521	AACUGGUGGGAAA UGGUCUAGG	745	AACTGGTGGGAAA TGGTCTAGG	内含子51
298	CCUAGACCAUUUC CCACCAGUUCU	522	AGAACUGGUGGGA AAUGGUCUAGG	746	AGAACTGGTGGGA AATGGTCTAGG	内含子51
299	CCUAGACCAUUUC CCACCAGUUCUU	523	AAGAACUGGUGGG AAAUGGUCUAGG	747	AAGAACTGGTGGG AAATGGTCTAGG	内含子51
300	CUAGACCAUUUCC CACCAGUU	524	AACUGGUGGGAAA UGGUCUAG	748	AACTGGTGGGAAA TGGTCTAG	内含子51
301	CUAGACCAUUUCC CACCAGUUCU	525	AGAACUGGUGGGA AAUGGUCUAG	749	AGAACTGGTGGGA AATGGTCTAG	内含子51
302	CUAGACCAUUUCC CACCAGUUCUU	526	AAGAACUGGUGGG AAAUGGUCUAG	750	AAGAACTGGTGGG AAATGGTCTAG	内含子51
303	UAGACCAUUUCCC ACCAGUU	527	AACUGGUGGGAAA UGGUCUA	751	AACTGGTGGGAAA TGGTCTA	内含子51
304	UAGACCAUUUCCC ACCAGUUCU	528	AGAACUGGUGGGA AAUGGUCUA	752	AGAACTGGTGGGA AATGGTCTA	内含子51
305	UAGACCAUUUCCC ACCAGUUCUU	529	AAGAACUGGUGGG AAAUGGUCUA	753	AAGAACTGGTGGG AAATGGTCTA	内含子51
306	UAGACCAUUUCCC ACCAGUUCUUAG	530	CUAAGAACUGGUG GGAAAUGGUCUA	754	CTAAGAACTGGTG GGAAATGGTCTA	内含子51
307	AGACCAUUUCCCA CCAGUUCU	531	AGAACUGGUGGGA AAUGGUCU	755	AGAACTGGTGGGA AATGGTCT	内含子51
308	AGACCAUUUCCCA CCAGUUCUU	532	AAGAACUGGUGGG AAAUGGUCU	756	AAGAACTGGTGGG AAATGGTCT	内含子51
309	AGACCAUUUCCCA CCAGUUCUUAG	533	CUAAGAACUGGUG GGAAAUGGUCU	757	CTAAGAACTGGTG GGAAATGGTCT	内含子51
310	AGACCAUUUCCCA CCAGUUCUUAGG	534	CCUAAGAACUGGU GGGAAAUGGUCU	758	CCTAAGAACTGGT GGGAAATGGTCT	内含子51
311	GACCAUUUCCCAC CAGUUCU	535	AGAACUGGUGGGA AAUGGUC	759	AGAACTGGTGGGA AATGGTC	内含子51
312	GACCAUUUCCCAC CAGUUCUU	536	AAGAACUGGUGGG AAAUGGUC	760	AAGAACTGGTGGG AAATGGTC	内含子51
313	GACCAUUUCCCAC CAGUUCUUAG	537	CUAAGAACUGGUG GGAAAUGGUC	761	CTAAGAACTGGTG GGAAATGGTC	内含子51
314	GACCAUUUCCCAC CAGUUCUUAGG	538	CCUAAGAACUGGU GGGAAAUGGUC	762	CCTAAGAACTGGT GGGAAATGGTC	内含子51
315	GACCAUUUCCCAC CAGUUCUUAGGC	539	GCCUAAGAACUGG UGGGAAAUGGUC	763	GCCTAAGAACTGG TGGGAAATGGTC	内含子51

[0320]

316	ACCAUUUCCCACC AGUUCUUAGGC	540	GCCUAAGAACUGG UGGGAAAUGGU	764	GCCTAAGAACTGG TGGGAAATGGT	内含子51
317	ACCAUUUCCCACC AGUUCUUAGGCA	541	UGCCUAAGAACUG GUGGGAAAUGGU	765	TGCCTAAGAACTG GTGGGAAATGGT	内含子51
318	CCAUUUCCCACCA GUUCUUAGGC	542	GCCUAAGAACUGG UGGGAAAUGG	766	GCCTAAGAACTGG TGGGAAATGG	内含子51
319	CCAUUUCCCACCA GUUCUUAGGCA	543	UGCCUAAGAACUG GUGGGAAAUGG	767	TGCCTAAGAACTG GTGGGAAATGG	内含子51
320	CCAUUUCCCACCA GUUCUUAGGCAA	544	UUGCCUAAGAACU GGUGGGAAAUGG	768	TTGCCTAAGAACT GGTGGGAAATGG	内含子51
321	CAUUUCCCACCAG UUCUUAGGCA	545	UGCCUAAGAACUG GUGGGAAAUG	769	TGCCTAAGAACTG GTGGGAAATG	内含子51
322	CAUUUCCCACCAG UUCUUAGGCAA	546	UUGCCUAAGAACU GGUGGGAAAUG	770	TTGCCTAAGAACT GGTGGGAAATG	内含子51
323	AGUGUUUUGGCUG GUCUCAC	547	GUGAGACCAGCCA AAACACU	771	GTGAGACCAGCCA AAACACT	内含子51
324	AGUGUUUUGGCUG GUCUCACA	548	UGUGAGACCAGCC AAAACACU	772	TGTGAGACCAGCC AAAACACT	内含子51
325	AGUGUUUUGGCUG GUCUCACAA	549	UUGUGAGACCAGC CAAAACACU	773	TTGTGAGACCAGC CAAAACACT	内含子51
326	AGUGUUUUGGCUG GUCUCACAAU	550	AUUGUGAGACCAG CCAAAACACU	774	ATTGTGAGACCAG CCAAAACACT	内含子51
327	GUGUUUUGGCUGG UCUCACAAU	551	AUUGUGAGACCAG CCAAAACAC	775	ATTGTGAGACCAG CCAAAACAC	内含子51
328	GUUUUGGCUGGUC UCACAAUUGUAC	552	GUACAAUUGUGAG ACCAGCCAAAAC	776	GTACAATTGTGAG ACCAGCCAAAAC	内含子51
329	UUUGGCUGGUCUC ACAAUUGUAC	553	GUACAAUUGUGAG ACCAGCCAAA	777	GTACAATTGTGAG ACCAGCCAAA	内含子51
330	UUGGCUGGUCUCA CAAUUGUAC	554	GUACAAUUGUGAG ACCAGCCAA	778	GTACAATTGTGAG ACCAGCCAA	内含子51
331	UUGGCUGGUCUCA CAAUUGUACU	555	AGUACAAUUGUGA GACCAGCCAA	779	AGTACAATTGTGA GACCAGCCAA	内含子51
332	UGGCUGGUCUCAC AAUUGUAC	556	GUACAAUUGUGAG ACCAGCCA	780	GTACAATTGTGAG ACCAGCCA	内含子51
333	UGGCUGGUCUCAC AAUUGUACU	557	AGUACAAUUGUGA GACCAGCCA	781	AGTACAATTGTGA GACCAGCCA	内含子51
334	UGGCUGGUCUCAC AAUUGUACUU	558	AAGUACAAUUGUG AGACCAGCCA	782	AAGTACAATTGTG AGACCAGCCA	内含子51
335	UGGCUGGUCUCAC AAUUGUACUUU	559	AAAGUACAAUUGU GAGACCAGCCA	783	AAAGTACAATTGT GAGACCAGCCA	内含子51
336	GGCUGGUCUCACA AUUGUAC	560	GUACAAUUGUGAG ACCAGCC	784	GTACAATTGTGAG ACCAGCC	内含子51
337	GGCUGGUCUCACA AUUGUACU	561	AGUACAAUUGUGA GACCAGCC	785	AGTACAATTGTGA GACCAGCC	内含子51
338	GGCUGGUCUCACA AUUGUACUU	562	AAGUACAAUUGUG AGACCAGCC	786	AAGTACAATTGTG AGACCAGCC	内含子51
339	GGCUGGUCUCACA AUUGUACUUU	563	AAAGUACAAUUGU GAGACCAGCC	787	AAAGTACAATTGT GAGACCAGCC	内含子51
340	GGCUGGUCUCACA AUUGUACUUUAC	564	GUAAAGUACAAUU GUGAGACCAGCC	788	GTAAAGTACAATT GTGAGACCAGCC	内含子51
341	GCUGGUCUCACAA UUGUACUUUAC	565	GUAAAGUACAAUU GUGAGACCAGC	789	GTAAAGTACAATT GTGAGACCAGC	内含子51
342	GCUGGUCUCACAA UUGUACUUUACU	566	AGUAAAGUACAAU UGUGAGACCAGC	790	AGTAAAGTACAAT TGTGAGACCAGC	内含子51
343	UGUAAAAGGAAUA CACAACGCUGA	567	UCAGCGUUGUGUA UUCCUUUUACA	791	TCAGCGTTGTGTA TTCCTTTTACA	内含子51
344	UGUAAAAGGAAUA CACAACGCUGAA	568	UUCAGCGUUGUGU AUUCCUUUUACA	792	TTCAGCGTTGTGT ATTCCTTTTACA	内含子51
345	GUAAAAGGAAUAC ACAACGCUGA	569	UCAGCGUUGUGUA UUCCUUUUAC	793	TCAGCGTTGTGTA TTCCTTTTAC	内含子51
346	GUAAAAGGAAUAC ACAACGCUGAA	570	UUCAGCGUUGUGU AUUCCUUUUAC	794	TTCAGCGTTGTGT ATTCCTTTTAC	内含子51

[0321]

347	GUAAAAGGAAUAC ACAACGCUGAAG	571	CUUCAGCGUUGUG UAUUCCUUUUAC	795	CTTCAGCGTTGTG TATTCCTTTTAC	内含子51
348	UAAAAGGAAUACA CAACGCUGA	572	UCAGCGUUGUGUA UUCCUUUUA	796	TCAGCGTTGTGTA TTCCTTTTA	内含子51
349	UAAAAGGAAUACA CAACGCUGAA	573	UUCAGCGUUGUGU AUUCCUUUUA	797	TTCAGCGTTGTGT ATTCCTTTTA	内含子51
350	UAAAAGGAAUACA CAACGCUGAAG	574	CUUCAGCGUUGUG UAUUCCUUUUA	798	CTTCAGCGTTGTG TATTCCTTTTA	内含子51
351	UAAAAGGAAUACA CAACGCUGAAGA	575	UCUUCAGCGUUGU GUAUUCCUUUUA	799	TCTTCAGCGTTGT GTATTCCTTTTA	内含子51
352	AAAAGGAAUACAC AACGCUGA	576	UCAGCGUUGUGUA UUCCUUUU	800	TCAGCGTTGTGTA TTCCTTTT	内含子51
353	AAAAGGAAUACAC AACGCUGAA	577	UUCAGCGUUGUGU AUUCCUUUU	801	TTCAGCGTTGTGT ATTCCTTTT	内含子51
354	AAAAGGAAUACAC AACGCUGAAG	578	CUUCAGCGUUGUG UAUUCCUUUU	802	CTTCAGCGTTGTG TATTCCTTTT	内含子51
355	AAAAGGAAUACAC AACGCUGAAGA	579	UCUUCAGCGUUGU GUAUUCCUUUU	803	TCTTCAGCGTTGT GTATTCCTTTT	内含子51
356	AAAAGGAAUACAC AACGCUGAAGAA	580	UUCUUCAGCGUUG UGUAUUCCUUUU	804	TTCTTCAGCGTTG TGTATTCCTTTT	内含子51
357	AAAGGAAUACACA ACGCUGA	581	UCAGCGUUGUGUA UUCCUUU	805	TCAGCGTTGTGTA TTCCTTT	内含子51
358	AAAGGAAUACACA ACGCUGAA	582	UUCAGCGUUGUGU AUUCCUUU	806	TTCAGCGTTGTGT ATTCCTTT	内含子51
359	AAAGGAAUACACA ACGCUGAAG	583	CUUCAGCGUUGUG UAUUCCUUU	807	CTTCAGCGTTGTG TATTCCTTT	内含子51
360	AAAGGAAUACACA ACGCUGAAGA	584	UCUUCAGCGUUGU GUAUUCCUUU	808	TCTTCAGCGTTGT GTATTCCTTT	内含子51
361	AAAGGAAUACACA ACGCUGAAGAA	585	UUCUUCAGCGUUG UGUAUUCCUUU	809	TTCTTCAGCGTTG TGTATTCCTTT	内含子51
362	AAAGGAAUACACA ACGCUGAAGAAC	586	GUUCUUCAGCGUU GUGUAUUCCUUU	810	GTTCTTCAGCGTT GTGTATTCCTTT	内含子51
363	AAGGAAUACACAA CGCUGAA	587	UUCAGCGUUGUGU AUUCCUU	811	TTCAGCGTTGTGT ATTCCTT	内含子51
364	AAGGAAUACACAA CGCUGAAG	588	CUUCAGCGUUGUG UAUUCCUU	812	CTTCAGCGTTGTG TATTCCTT	内含子51
365	AAGGAAUACACAA CGCUGAAGA	589	UCUUCAGCGUUGU GUAUUCCUU	813	TCTTCAGCGTTGT GTATTCCTT	内含子51
366	AAGGAAUACACAA CGCUGAAGAA	590	UUCUUCAGCGUUG UGUAUUCCUU	814	TTCTTCAGCGTTG TGTATTCCTT	内含子51
367	AAGGAAUACACAA CGCUGAAGAAC	591	GUUCUUCAGCGUU GUGUAUUCCUU	815	GTTCTTCAGCGTT GTGTATTCCTT	内含子51
368	AGGAAUACACAAC GCUGAAG	592	CUUCAGCGUUGUG UAUUCCU	816	CTTCAGCGTTGTG TATTCCT	内含子51
369	AGGAAUACACAAC GCUGAAGA	593	UCUUCAGCGUUGU GUAUUCCU	817	TCTTCAGCGTTGT GTATTCCT	内含子51
370	AGGAAUACACAAC GCUGAAGAA	594	UUCUUCAGCGUUG UGUAUUCCU	818	TTCTTCAGCGTTG TGTATTCCT	内含子51
371	AGGAAUACACAAC GCUGAAGAAC	595	GUUCUUCAGCGUU GUGUAUUCCU	819	GTTCTTCAGCGTT GTGTATTCCT	内含子51
372	GGAAUACACAACG CUGAAGA	596	UCUUCAGCGUUGU GUAUUCC	820	TCTTCAGCGTTGT GTATTCC	内含子51
373	GGAAUACACAACG CUGAAGAA	597	UUCUUCAGCGUUG UGUAUUCC	821	TTCTTCAGCGTTG TGTATTCC	内含子51
374	GGAAUACACAACG CUGAAGAAC	598	GUUCUUCAGCGUU GUGUAUUCC	822	GTTCTTCAGCGTT GTGTATTCC	内含子51
375	GGAAUACACAACG CUGAAGAACCC	599	GGGUUCUUCAGCG UUGUGUAUUCC	823	GGGTTCTTCAGCG TTGTGTATTCC	内含子51
376	GGAAUACACAACG CUGAAGAACCCU	600	AGGGUUCUUCAGC GUUGUGUAUUCC	824	AGGGTTCTTCAGC GTTGTGTATTCC	内含子51
377	GAAUACACAACGC UGAAGAAC	601	GUUCUUCAGCGUU GUGUAUUC	825	GTTCTTCAGCGTT GTGTATTC	内含子51

[0322]

378	GAAUACACAACGC UGAAGAACCC	602	GGGUUCUUCAGCG UUGUGUAUUC	826	GGGTTCTTCAGCG TTGTGTATTC	内含子51
379	GAAUACACAACGC UGAAGAACCCU	603	AGGGUUCUUCAGC GUUGUGUAUUC	827	AGGGTTCTTCAGC GTTGTGTATTC	内含子51
380	AAUACACAACGCU GAAGAACCC	604	GGGUUCUUCAGCG UUGUGUAUU	828	GGGTTCTTCAGCG TTGTGTATT	内含子51
381	AUACACAACGCUG AAGAACCC	605	GGGUUCUUCAGCG UUGUGUAU	829	GGGTTCTTCAGCG TTGTGTAT	内含子51
382	UACACAACGCUGA AGAACCC	606	GGGUUCUUCAGCG UUGUGUA	830	GGGTTCTTCAGCG TTGTGTA	内含子51
383	ACACAACGCUGAA GAACCCUGAU	607	AUCAGGGUUCUUC AGCGUUGUGU	831	ATCAGGGTTCTTC AGCGTTGTGT	内含子51

[0323] 表8中提供的任一个寡核苷酸和/或靶序列中的每个胸腺嘧啶碱基(T)可独立且任选地用尿嘧啶碱基(U)替换,和/或每个U可独立且任选地用T替换。表8中列出的靶序列包含U,但是考虑了DMD靶向寡核苷酸与RNA和/或DNA的结合。

[0324] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD序列的区域。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD RNA(例如,SEQ ID NO:130的Dp427m转录物)的区域。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA(例如,SEQ ID NO:130的Dp427m转录物)的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA(例如,SEQ ID NO:131、838或854)的外显子的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA(例如,SEQ ID NO:834或846)的内含子的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD序列(例如,由SEQ ID NO:832、844、833、845、835、836、847至852、837、853、839至843中任一者提供的序列)的一部分的互补区。以下提供了DMD序列的一些实例。以下提供的每个DMD序列均包含胸腺嘧啶核苷酸(T),但是应理解,每个序列可代表其中任何T或所有T将用尿嘧啶核苷酸(U)替换的DNA序列或RNA序列。

[0325] 智人肌养蛋白(DMD),转录物变体Dp427m,mRNA(NCBI参考序列:NM_004006.2)

[0326]
```
TCCTGGCATCAGTTACTGTGTTGACTCACTCAGTGTTGGGATCACTCACTTTCCCCCTACAGGACTCAGATCTGGGA
GGCAATTACCTTCGGAGAAAAACGAATAGGAAAAACTGAAGTGTTACTTTTTTTAAAGCTGCTGAAGTTTGTTGGTT
TCTCATTGTTTTTAAGCCTACTGGAGCAATAAAGTTTGAAGAACTTTTACCAGGTTTTTTTTATCGCTGCCTTGATA
TACACTTTTCAAAATGCTTTGGTGGGAAGAAGTAGAGGACTGTTATGAAAGAGAAGATGTTCAAAAGAAAACATTCA
CAAAATGGGTAAATGCACAATTTTCTAAGTTTGGGAAGCAGCATATTGAGAACCTCTTCAGTGACCTACAGGATGGG
AGGCGCCTCCTAGACCTCCTCGAAGGCCTGACAGGGCAAAAACTGCCAAAAGAAAAAGGATCCACAAGAGTTCATGC
CCTGAACAATGTCAACAAGGCACTGCGGGTTTTGCAGAACAATAATGTTGATTTAGTGAATATTGGAAGTACTGACA
TCGTAGATGGAAATCATAAACTGACTCTTGGTTTGATTTGGAATATAATCCTCCACTGGCAGGTCAAAAATGTAATG
AAAAATATCATGGCTGGATTGCAACAAACCAACAGTGAAAAGATTCTCCTGAGCTGGGTCCGACAATCAACTCGTAA
TTATCCACAGGTTAATGTAATCAACTTCACCACCAGCTGGTCTGATGGCCTGGCTTTGAATGCTCTCATCCATAGTC
ATAGGCCAGACCTATTTGACTGGAATAGTGTGGTTTGCCAGCAGTCAGCCACACAACGACTGGAACATGCATTCAAC
```

[0327]

ATCGCCAGATATCAATTAGGCATAGAGAAACTACTCGATCCTGAAGATGTTGATACCACCTATCCAGATAAGAAGTC
CATCTTAATGTACATCACATCACTCTTCCAAGTTTTGCCTCAACAAGTGAGCATTGAAGCCATCCAGGAAGTGGAAA
TGTTGCCAAGGCCACCTAAAGTGACTAAAGAAGAACATTTTCAGTTACATCATCAAATGCACTATTCTCAACAGATC
ACGGTCAGTCTAGCACAGGGATATGAGAGAACTTCTTCCCCTAAGCCTCGATTCAAGAGCTATGCCTACACACAGGC
TGCTTATGTCACCACCTCTGACCCTACACGGAGCCCATTTCCTTCACAGCATTTGGAAGCTCCTGAAGACAAGTCAT
TTGGCAGTTCATTGATGGAGAGTGAAGTAAACCTGGACCGTTATCAAACAGCTTTAGAAGAAGTATTATCGTGGCTT
CTTTCTGCTGAGGACACATTGCAAGCACAAGGAGAGATTTCTAATGATGTGGAAGTGGTGAAAGACCAGTTTCATAC
TCATGAGGGGTACATGATGGATTTGACAGCCCATCAGGGCCGGGTTGGTAATATTCTACAATTGGGAAGTAAGCTGA
TTGGAACAGGAAAATTATCAGAAGATGAAGAAACTGAAGTACAAGAGCAGATGAATCTCCTAAATTCAAGATGGGAA
TGCCTCAGGGTAGCTAGCATGGAAAAACAAAGCAATTTACATAGAGTTTTAATGGATCTCCAGAATCAGAAACTGAA
AGAGTTGAATGACTGGCTAACAAAAACAGAAGAAAGAACAAGGAAAATGGAGGAAGAGCCTCTTGGACCTGATCTTG
AAGACCTAAAACGCCAAGTACAACAACATAAGGTGCTTCAAGAAGATCTAGAACAAGAACAAGTCAGGGTCAATTCT
CTCACTCACATGGTGGTGGTAGTTGATGAATCTAGTGGAGATCACGCAACTGCTGCTTTGGAAGAACAACTTAAGGT
ATTGGGAGATCGATGGGCAAACATCTGTAGATGGACAGAAGACCGCTGGGTTCTTTTACAAGACATCCTTCTCAAAT
GGCAACGTCTTACTGAAGAACAGTGCCTTTTTAGTGCATGGCTTTCAGAAAAAGAAGATGCAGTGAACAAGATTCAC
ACAACTGGCTTTAAAGATCAAAATGAAATGTTATCAAGTCTTCAAAAACTGGCCGTTTTAAAAGCGGATCTAGAAAA
GAAAAAGCAATCCATGGGCAAACTGTATTCACTCAAACAAGATCTTCTTTCAACACTGAAGAATAAGTCAGTGACCC
AGAAGACGGAAGCATGGCTGGATAACTTTGCCCGGTGTTGGGATAATTTAGTCCAAAAACTTGAAAAGAGTACAGCA
CAGATTTCACAGGCTGTCACCACCACTCAGCCATCACTAACACAGACAACTGTAATGGAAACAGTAACTACGGTGAC
CACAAGGGAACAGATCCTGGTAAAGCATGCTCAAGAGGAACTTCCACCACCACCTCCCCAAAAGAAGAGGCAGATTA
CTGTGGATTCTGAAATTAGGAAAAGGTTGGATGTTGATATAACTGAACTTCACAGCTGGATTACTCGCTCAGAAGCT
GTGTTGCAGAGTCCTGAATTTGCAATCTTTCGGAAGGAAGGCAACTTCTCAGACTTAAAAGAAAAAGTCAATGCCAT
AGAGCGAGAAAAAGCTGAGAAGTTCAGAAAACTGCAAGATGCCAGCAGATCAGCTCAGGCCCTGGTGGAACAGATGG
TGAATGAGGGTGTTAATGCAGATAGCATCAAACAAGCCTCAGAACAACTGAACAGCCGGTGGATCGAATTCTGCCAG
TTGCTAAGTGAGAGACTTAACTGGCTGGAGTATCAGAACAACATCATCGCTTTCTATAATCAGCTACAACAATTGGA
GCAGATGACAACTACTGCTGAAAACTGGTTGAAAATCCAACCCACCACCCCATCAGAGCCAACAGCAATTAAAAGTC
AGTTAAAAATTTGTAAGGATGAAGTCAACCGGCTATCAGGTCTTCAACCTCAAATTGAACGATTAAAAATTCAAAGC
ATAGCCCTGAAAGAGAAAGGACAAGGACCCATGTTCCTGGATGCAGACTTTGTGGCCTTTACAAATCATTTTAAGCA
AGTCTTTTCTGATGTGCAGGCCAGAGAGAAAGAGCTACAGACAATTTTTGACACTTTGCCACCAATGCGCTATCAGG
AGACCATGAGTGCCATCAGGACATGGGTCCAGCAGTCAGAAACCAAACTCTCCATACCTCAACTTAGTGTCACCGAC
TATGAAATCATGGAGCAGAGACTCGGGGAATTGCAGGCTTTACAAAGTTCTCTGCAAGAGCAACAAAGTGGCCTATA
CTATCTCAGCACCACTGTGAAAGAGATGTCGAAGAAAGCGCCCTCTGAAATTAGCCGGAAATATCAATCAGAATTTG
AAGAAATTGAGGGACGCTGGAAGAAGCTCTCCTCCCAGCTGGTTGAGCATTGTCAAAAGCTAGAGGAGCAAATGAAT
AAACTCCGAAAAATTCAGAATCACATACAAACCCTGAAGAAATGGATGGCTGAAGTTGATGTTTTTCTGAAGGAGGA
ATGGCCTGCCCTTGGGGATTCAGAAATTCTAAAAAAGCAGCTGAAACAGTGCAGACTTTTAGTCAGTGATATTCAGA
CAATTCAGCCCAGTCTAAACAGTGTCAATGAAGGTGGGCAGAAGATAAAGAATGAAGCAGAGCCAGAGTTTGCTTCG
AGACTTGAGACAGAACTCAAAGAACTTAACACTCAGTGGGATCACATGTGCCAACAGGTCTATGCCAGAAAGGAGGC
CTTGAAGGGAGGTTTGGAGAAAACTGTAAGCCTCCAGAAAGATCTATCAGAGATGCACGAATGGATGACACAAGCTG
AAGAAGAGTATCTTGAGAGAGATTTTGAATATAAAACTCCAGATGAATTACAGAAAGCAGTTGAAGAGATGAAGAGA
GCTAAAGAAGAGGCCCAACAAAAAGAAGCGAAAGTGAAACTCCTTACTGAGTCTGTAAATAGTGTCATAGCTCAAGC
TCCACCTGTAGCACAAGAGGCCTTAAAAAAGGAACTTGAAACTCTAACCACCAACTACCAGTGGCTCTGCACTAGGC
TGAATGGGAAATGCAAGACTTTGGAAGAAGTTTGGGCATGTTGGCATGAGTTATTGTCATACTTGGAGAAAGCAAAC
AAGTGGCTAAATGAAGTAGAATTTAAACTTAAAACCACTGAAAACATTCCTGGCGGAGCTGAGGAAATCTCTGAGGT
GCTAGATTCACTTGAAAATTTGATGCGACATTCAGAGGATAACCCAAATCAGATTCGCATATTGGCACAGACCCTAA
CAGATGGCGGAGTCATGGATGAGCTAATCAATGAGGAACTTGAGACATTTAATTCTCGTTGGAGGGAACTACATGAA
GAGGCTGTAAGGAGGCAAAAGTTGCTTGAACAGAGCATCCAGTCTGCCCAGGAGACTGAAAAATCCTTACACTTAAT
CCAGGAGTCCCTCACATTCATTGACAAGCAGTTGGCAGCTTATATTGCAGACAAGGTGGACGCAGCTCAAATGCCTC
AGGAAGCCCAGAAAATCCAATCTGATTTGACAAGTCATGAGATCAGTTTAGAAGAAATGAAGAAACATAATCAGGGG
AAGGAGGCTGCCCAAAGAGTCCTGTCTCAGATTGATGTTGCACAGAAAAAATTACAAGATGTCTCCATGAAGTTTCG
ATTATTCCAGAAACCAGCCAATTTTGAGCAGCGTCTACAAGAAAGTAAGATGATTTTAGATGAAGTGAAGATGCACT
TGCCTGCATTGGAAACAAAGAGTGTGGAACAGGAAGTAGTACAGTCACAGCTAAATCATTGTGTGAACTTGTATAAA
AGTCTGAGTGAAGTGAAGTCTGAAGTGGAAATGGTGATAAAGACTGGACGTCAGATTGTACAGAAAAAGCAGACGGA
AAATCCCAAAGAACTTGATGAAAGAGTAACAGCTTTGAAATTGCATTATAATGAGCTGGGAGCAAAGGTAACAGAAA
GAAAGCAACAGTTGGAGAAATGCTTGAAATTGTCCCGTAAGATGCGAAAGGAAATGAATGTCTTGACAGAATGGCTG
GCAGCTACAGATATGGAATTGACAAAGAGATCAGCAGTTGAAGGAATGCCTAGTAATTTGGATTCTGAAGTTGCCTG
GGGAAAGGCTACTCAAAAAGAGATTGAGAAACAGAAGGTGCACCTGAAGAGTATCACAGAGGTAGGAGAGGCCTTGA
AAACAGTTTTGGGCAAGAAGGAGACGTTGGTGGAAGATAAACTCAGTCTTCTGAATAGTAACTGGATAGCTGTCACC
TCCCGAGCAGAAGAGTGGTTAAATCTTTTGTTGGAATACCAGAAACACATGGAAACTTTTGACCAGAATGTGGACCA
CATCACAAAGTGGATCATTCAGGCTGACACACTTTTGGATGAATCAGAGAAAAAGAAACCCCAGCAAAAAGAAGACG
TGCTTAAGCGTTTAAAGGCAGAACTGAATGACATACGCCCAAAGGTGGACTCTACACGTGACCAAGCAGCAAACTTG
ATGGCAAACCGCGGTGACCACTGCAGGAAATTAGTAGAGCCCCAAATCTCAGAGCTCAACCATCGATTTGCAGCCAT
TTCACACAGAATTAAGACTGGAAAGGCCTCCATTCCTTTGAAGGAATTGGAGCAGTTTAACTCAGATATACAAAAAT
TGCTTGAACCACTGGAGGCTGAAATTCAGCAGGGGGTGAATCTGAAAGAGGAAGACTTCAATAAAGATATGAATGAA
GACAATGAGGGTACTGTAAAAGAATTGTTGCAAAGAGGAGACAACTTACAACAAAGAATCACAGATGAGAGAAAGCG

[0328]

AGAGGAAATAAAGATAAAACAGCAGCTGTTACAGACAAAACATAATGCTCTCAAGGATTTGAGGTCTCAAAGAAGAA
AAAAGGCTCTAGAAATTTCTCATCAGTGGTATCAGTACAAGAGGCAGGCTGATGATCTCCTGAAATGCTTGGATGAC
ATTGAAAAAAAATTAGCCAGCCTACCTGAGCCCAGAGATGAAAGGAAAATAAAGGAAATTGATCGGGAATTGCAGAA
GAAGAAAGAGGAGCTGAATGCAGTGCGTAGGCAAGCTGAGGGCTTGTCTGAGGATGGGGCCGCAATGGCAGTGGAGC
CAACTCAGATCCAGCTCAGCAAGCGCTGGCGGGAAATTGAGAGCAAATTTGCTCAGTTTCGAAGACTCAACTTTGCA
CAAATTCACACTGTCCGTGAAGAAACGATGATGGTGATGACTGAAGACATGCCTTTGGAAATTTCTTATGTGCCTTC
TACTTATTTGACTGAAATCACTCATGTCTCACAAGCCCTATTAGAAGTGGAACAACTTCTCAATGCTCCTGACCTCT
GTGCTAAGGACTTTGAAGATCTCTTTAAGCAAGAGGAGTCTCTGAAGAATATAAAAGATAGTCTACAACAAAGCTCA
GGTCGGATTGACATTATTCATAGCAAGAAGACAGCAGCATTGCAAAGTGCAACGCCTGTGGAAAGGGTGAAGCTACA
GGAAGCTCTCTCCCAGCTTGATTTCCAATGGGAAAAAGTTAACAAAATGTACAAGGACCGACAAGGGCGATTTGACA
GATCTGTTGAGAAATGGCGGCGTTTTCATTATGATATAAAGATATTTAATCAGTGGCTAACAGAAGCTGAACAGTTT
CTCAGAAAGACACAAATTCCTGAGAATTGGGAACATGCTAAATACAAATGGTATCTTAAGGAACTCCAGGATGGCAT
TGGGCAGCGGCAAACTGTTGTCAGAACATTGAATGCAACTGGGGAAGAAATAATTCAGCAATCCTCAAAAACAGATG
CCAGTATTCTACAGGAAAAATTGGGAAGCCTGAATCTGCGGTGGCAGGAGGTCTGCAAACAGCTGTCAGACAGAAAA
AAGAGGCTAGAAGAACAAAAGAATATCTTGTCAGAATTTCAAAGAGATTTAAATGAATTTGTTTTATGGTTGGAGGA
AGCAGATAACATTGCTAGTATCCCACTTGAACCTGGAAAAGAGCAGCAACTAAAAGAAAAGCTTGAGCAAGTCAAGT
TACTGGTGGAAGAGTTGCCCCTGCGCCAGGGAATTCTCAAACAATTAAATGAAACTGGAGGACCCGTGCTTGTAAGT
GCTCCCATAAGCCCAGAAGAGCAAGATAAACTTGAAAATAAGCTCAAGCAGACAAATCTCCAGTGGATAAAGGTTTC
CAGAGCTTTACCTGAGAAACAAGGAGAAATTGAAGCTCAAATAAAAGACCTTGGGCAGCTTGAAAAAAAGCTTGAAG
ACCTTGAAGAGCAGTTAAATCATCTGCTGCTGTGGTTATCTCCTATTAGGAATCAGTTGGAAATTTATAACCAACCA
AACCAAGAAGGACCATTTGACGTTCAGGAAACTGAAATAGCAGTTCAAGCTAAACAACCGGATGTGGAAGAGATTTT
GTCTAAAGGGCAGCATTTGTACAAGGAAAAACCAGCCACTCAGCCAGTGAAGAGGAAGTTAGAAGATCTGAGCTCTG
AGTGGAAGGCGGTAAACCGTTTACTTCAAGAGCTGAGGGCAAAGCAGCCTGACCTAGCTCCTGGACTGACCACTATT
GGAGCCTCTCCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGA
AATGCCATCTTCCTTGATGTTGGAGGTACCTGCTCTGGCAGATTTCAACCGGGCTTGGACAGAACTTACCGACTGGC
TTTCTCTGCTTGATCAAGTTATAAAATCACAGAGGGTGATGGTGGGTGACCTTGAGGATATCAACGAGATGATCATC
AAGCAGAAGGCAACAATGCAGGATTTGGAACAGAGGCGTCCCCAGTTGGAAGAACTCATTACCGCTGCCCAAAATTT
GAAAAACAAGACCAGCAATCAAGAGGCTAGAACAATCATTACGGATCGAATTGAAAGAATTCAGAATCAGTGGGATG
AAGTACAAGAACACCTTCAGAACCGGAGGCAACAGTTGAATGAAATGTTAAAGGATTCAACACAATGGCTGGAAGCT
AAGGAAGAAGCTGAGCAGGTCTTAGGACAGGCCAGAGCCAAGCTTGAGTCATGGAAGGAGGGTCCCTATACAGTAGA
TGCAATCCAAAAGAAAATCACAGAAACCAAGCAGTTGGCCAAAGACCTCCGCCAGTGGCAGACAAATGTAGATGTGG
CAAATGACTTGGCCCTGAAACTTCTCCGGGATTATTCTGCAGATGATACCAGAAAAGTCCACATGATAACAGAGAAT
ATCAATGCCTCTTGGAGAAGCATTCATAAAAGGGTGAGTGAGCGAGAGGCTGCTTTGGAAGAAACTCATAGATTACT
GCAACAGTTCCCCCTGGACCTGGAAAAGTTTCTTGCCTGGCTTACAGAAGCTGAAACAACTGCCAATGTCCTACAGG
ATGCTACCCGTAAGGAAAGGCTCCTAGAAGACTCCAAGGGAGTAAAAGAGCTGATGAAACAATGGCAAGACCTCCAA
GGTGAAATTGAAGCTCACACAGATGTTTATCACAACCTGGATGAAAACAGCCAAAAAATCCTGAGATCCCTGGAAGG
TTCCGATGATGCAGTCCTGTTACAAAGACGTTTGGATAACATGAACTTCAAGTGGAGTGAACTTCGGAAAAAGTCTC
TCAACATTAGGTCCCATTTGGAAGCCAGTTCTGACCAGTGGAAGCGTCTGCACCTTTCTCTGCAGGAACTTCTGGTG
TGGCTACAGCTGAAAGATGATGAATTAAGCCGGCAGGCACCTATTGGAGGCGACTTTCCAGCAGTTCAGAAGCAGAA
CGATGTACATAGGGCCTTCAAGAGGGAATTGAAAACTAAAGAACCTGTAATCATGAGTACTCTTGAGACTGTACGAA
TATTTCTGACAGAGCAGCCTTTGGAAGGACTAGAGAAACTCTACCAGGAGCCCAGAGAGCTGCCTCCTGAGGAGAGA
GCCCAGAATGTCACTCGGCTTCTACGAAAGCAGGCTGAGGAGGTCAATACTGAGTGGGAAAAATTGAACCTGCACTC
CGCTGACTGGCAGAGAAAAATAGATGAGACCCTTGAAAGACTCCAGGAACTTCAAGAGGCCACGGATGAGCTGGACC
TCAAGCTGCGCCAAGCTGAGGTGATCAAGGGATCCTGGCAGCCCGTGGGCGATCTCCTCATTGACTCTCTCCAAGAT
CACCTCGAGAAAGTCAAGGCACTTCGAGGAGAAATTGCGCCTCTGAAAGAGAACGTGAGCCACGTCAATGACCTTGC
TCGCCAGCTTACCACTTTGGGCATTCAGCTCTCACCGTATAACCTCAGCACTCTGGAAGACCTGAACACCAGATGGA
AGCTTCTGCAGGTGGCCGTCGAGGACCGAGTCAGGCAGCTGCATGAAGCCCACAGGGACTTTGGTCCAGCATCTCAG
CACTTTCTTTCCACGTCTGTCCAGGGTCCCTGGGAGAGAGCCATCTCGCCAAACAAAGTGCCCTACTATATCAACCA
CGAGACTCAAACAACTTGCTGGGACCATCCCAAAATGACAGAGCTCTACCAGTCTTTAGCTGACCTGAATAATGTCA
GATTCTCAGCTTATAGGACTGCCATGAAACTCCGAAGACTGCAGAAGGCCCTTTGCTTGGATCTCTTGAGCCTGTCA
GCTGCATGTGATGCCTTGGACCAGCACAACCTCAAGCAAAATGACCAGCCCATGGATATCCTGCAGATTATTAATTG
TTTGACCACTATTTATGACCGCCTGGAGCAAGAGCACAACAATTTGGTCAACGTCCCTCTCTGCGTGGATATGTGTC
TGAACTGGCTGCTGAATGTTTATGATACGGGACGAACAGGGAGGATCCGTGTCCTGTCTTTTAAAACTGGCATCATT
TCCCTGTGTAAAGCACATTTGGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATTTTGTGA
CCAGCGCAGGCTGGGCCTCCTTCTGCATGATTCTATCCAAATTCCAAGACAGTTGGGTGAAGTTGCATCCTTTGGGG
GCAGTAACATTGAGCCAAGTGTCCGGAGCTGCTTCCAATTTGCTAATAATAAGCCAGAGATCGAAGCGGCCCTCTTC
CTAGACTGGATGAGACTGGAACCCCAGTCCATGGTGTGGCTGCCCGTCCTGCACAGAGTGGCTGCTGCAGAAACTGC
CAAGCATCAGGCCAAATGTAACATCTGCAAAGAGTGTCCAATCATTGGATTCAGGTACAGGAGTCTAAAGCACTTTA
ATTATGACATCTGCCAAAGCTGCTTTTTTTCTGGTCGAGTTGCAAAAGGCCATAAAATGCACTATCCCATGGTGGAA
TATTGCACTCCGACTACATCAGGAGAAGATGTTCGAGACTTTGCCAAGGTACTAAAAAACAAATTTCGAACCAAAAG
GTATTTTGCGAAGCATCCCCGAATGGGCTACCTGCCAGTGCAGACTGTCTTAGAGGGGGACAACATGGAAACTCCCG
TTACTCTGATCAACTTCTGGCCAGTAGATTCTGCGCCTGCCTCGTCCCCTCAGCTTTCACACGATGATACTCATTCA
CGCATTGAACATTATGCTAGCAGGCTAGCAGAAATGGAAAACAGCAATGGATCTTATCTAAATGATAGCATCTCTCC
TAATGAGAGCATAGATGATGAACATTTGTTAATCCAGCATTACTGCCAAAGTTTGAACCAGGACTCCCCCCTGAGCC

[0329] AGCCTCGTAGTCCTGCCCAGATCTTGATTTCCTTAGAGAGTGAGGAAAGAGGGGAGCTAGAGAGAATCCTAGCAGAT
CTTGAGGAAGAAAACAGGAATCTGCAAGCAGAATATGACCGTCTAAAGCAGCAGCACGAACATAAAGGCCTGTCCCC
ACTGCCGTCCCCTCCTGAAATGATGCCCACCTCTCCCCAGAGTCCCCGGGATGCTGAGCTCATTGCTGAGGCCAAGC
TACTGCGTCAACACAAAGGCCGCCTGGAAGCCAGGATGCAAATCCTGGAAGACCACAATAAACAGCTGGAGTCACAG
TTACACAGGCTAAGGCAGCTGCTGGAGCAACCCCAGGCAGAGGCCAAAGTGAATGGCACAACGGTGTCCTCTCCTTC
TACCTCTCTACAGAGGTCCGACAGCAGTCAGCCTATGCTGCTCCGAGTGGTTGGCAGTCAAACTTCGGACTCCATGG
GTGAGGAAGATCTTCTCAGTCCTCCCCAGGACACAAGCACAGGGTTAGAGGAGGTGATGGAGCAACTCAACAACTCC
TTCCCTAGTTCAAGAGGAAGAAATACCCCTGGAAAGCCAATGAGAGAGGACACAATGTAGGAAGTCTTTTCCACATG
GCAGATGATTTGGGCAGAGCGATGGAGTCCTTAGTATCAGTCATGACAGATGAAGAAGGAGCAGAATAAATGTTTTA
CAACTCCTGATTCCCGCATGGTTTTTATAATATTCATACAACAAAGAGGATTAGACAGTAAGAGTTTACAAGAAATA
AATCTATATTTTTGTGAAGGGTAGTGGTATTATACTGTAGATTTCAGTAGTTTCTAAGTCTGTTATTGTTTTGTTAA
CAATGGCAGGTTTTACACGTCTATGCAATTGTACAAAAAAGTTATAAGAAAACTACATGTAAAATCTTGATAGCTAA
ATAACTTGCCATTTCTTTATATGGAACGCATTTTGGGTTGTTTAAAAATTTATAACAGTTATAAAGAAAGATTGTAA
ACTAAAGTGTGCTTTATAAAAAAAAGTTGTTTATAAAAACCCCTAAAAACAAAACAAACACACACACACACATAC
ACACACACACACAAAACTTTGAGGCAGCGCATTGTTTTGCATCCTTTTGGCGTGATATCCATATGAAATTCATGGCT
TTTTCTTTTTTTGCATATTAAAGATAAGACTTCCTCTACCACCACACCAAATGACTACTACACACTGCTCATTTGAG
AACTGTCAGCTGAGTGGGGCAGGCTTGAGTTTTCATTTCATATATCTATATGTCTATAAGTATATAAATACTATAGT
TATATAGATAAAGAGATACGAATTTCTATAGACTGACTTTTTCCATTTTTTAAATGTTCATGTCACATCCTAATAGA
AAGAAATTACTTCTAGTCAGTCATCCAGGCTTACCTGCTTGGTCTAGAATGGATTTTTCCCGGAGCCGGAAGCCAGG
AGGAAACTACACCACACTAAAACATTGTCTACAGCTCCAGATGTTTCTCATTTTAAACAACTTTCCACTGACAACGA
AAGTAAAGTAAAGTATTGGATTTTTTTAAAGGGAACATGTGAATGAATACACAGGACTTATTATATCAGAGTGAGTA
ATCGGTTGGTTGGTTGATTGATTGATTGATTGATACATTCAGCTTCCTGCTGCTAGCAATGCCACGATTTAGATTTA
ATGATGCTTCAGTGGAAATCAATCAGAAGGTATTCTGACCTTGTGAACATCAGAAGGTATTTTTTAACTCCCAAGCA
GTAGCAGGACGATGATAGGGCTGGAGGGCTATGGATTCCCAGCCCATCCCTGTGAAGGAGTAGGCCACTCTTTAAGT
GAAGGATTGGATGATTGTTCATAATACATAAAGTTCTCTGTAATTACAACTAAATTATTATGCCCTCTTCTCACAGT
CAAAAGGAACTGGGTGGTTTGGTTTTTGTTGCTTTTTTAGATTTATTGTCCCATGTGGGATGAGTTTTTAAATGCCA
CAAGACATAATTTAAAATAAATAAACTTTGGGAAAAGGTGTAAAACAGTAGCCCCATCACATTTGTGATACTGACAG
GTATCAACCCAGAAGCCCATGAACTGTGTTTCCATCCTTTGCATTTCTCTGCGAGTAGTTCCACACAGGTTTGTAAG
TAAGTAAGAAAGAAGGCAAATTGATTCAAATGTTACAAAAAAACCCTTCTTGGTGGATTAGACAGGTTAAATATATA
AACAAACAAACAAAAATTGCTCAAAAAAGAGGAGAAAAGCTCAAGAGGAAAAGCTAAGGACTGGTAGGAAAAAGCTT
TACTCTTTCATGCCATTTTATTTCTTTTTGATTTTTAAATCATTCATTCAATAGATACCACCGTGTGACCTATAATT
TTGCAAATCTGTTACCTCTGACATCAAGTGTAATTAGCTTTTGGAGAGTGGGCTGACATCAAGTGTAATTAGCTTTT
GGAGAGTGGGTTTTGTCCATTATTAATAATTAATTAATTAACATCAAACACGGCTTCTCATGCTATTTCTACCTCAC
TTTGGTTTTGGGGTGTTCCTGATAATTGTGCACACCTGAGTTCACAGCTTCACCACTTGTCCATTGCGTTATTTTCT
TTTTCCTTTATAATTCTTTCTTTTTCCTTCATAATTTTCAAAAGAAAACCCAAAGCTCTAAGGTAACAAATTACCAA
ATTACATGAAGATTTGGTTTTTGTCTTGCATTTTTTTCCTTTATGTGACGCTGGACCTTTTCTTTACCCAAGGATTT
TTAAAACTCAGATTTAAAACAAGGGGTTACTTTACATCCTACTAAGAAGTTTAAGTAAGTAAGTTTCATTCTAAAAT
CAGAGGTAAATAGAGTGCATAAATAATTTTGTTTTAATCTTTTTGTTTTTCTTTTAGACACATTAGCTCTGGAGTGA
GTCTGTCATAATATTTGAACAAAAATTGAGAGCTTTATTGCTGCATTTTAAGCATAATTAATTTGGACATTATTTCG
TGTTGTGTTCTTTATAACCACCAAGTATTAAACTGTAAATCATAATGTAACTGAAGCATAAACATCACATGGCATGT
TTTGTCATTGTTTTCAGGTACTGAGTTCTTACTTGAGTATCATAATATATTGTGTTTTAACACCAACACTGTAACAT
TTACGAATTATTTTTTTAAACTTCAGTTTTACTGCATTTTCACAACATATCAGACTTCACCAAATATATGCCTTACT
ATTGTATTATAGTACTGCTTTACTGTGTATCTCAATAAAGCACGCAGTTATGTTAC (SEQ ID NO: 130)

[0330] 智人肌养蛋白(DMD),转录物变体Dp427m,外显子50(NCBI参考序列:NM_004006.2的第7445至7553位核苷酸,NCBI参考序列:NG_012232.1的第1524527至1524635位核苷酸)

[0331] AGGAAGTTAGAAGATCTGAGCTCTGAGTGGAAGGCGGTAAACCGTTTACTTCAAGAGCTGAGGGCAAAGCAGCCTGA
CCTAGCTCCTGGACTGACCACTATTGGAGCCT (SEQ ID NO: 131)

[0332] 智人肌养蛋白(DMD),外显子50/内含子50连接(NCBI参考序列:NG_012232.1的第1524606至1524665位核苷酸)

[0333] TAGCTCCTGGACTGACCACTATTGGAGCCTGTAAGTATACTGGATCCCATTCTCTTTGGC (SEQ ID NO: 832)

[0334] 智人肌养蛋白(DMD),外显子50/内含子50连接靶序列1(NCBI参考序列:NG_012232.1的第1524626至1524677位核苷酸)

[0335] ATTGGAGCCTGTAAGTATACTGGATCCCATTCTCTTTGGCTCTAGCTATTTG (SEQ ID NO: 833)

[0336] 智人肌养蛋白(DMD),内含子50(NCBI参考序列:NG_012232.1的第1524636至1570417位核苷酸)

[0337]

GTAAGTATACTGGATCCCATTCTCTTTGGCTCTAGCTATTTGTTCAAAAGTGCAACTATGAAGTGATGACTGGGTGA
GAGAGAAAATTTGTTTCAATTCTAAAGATAGAGATAAACCTTTGTGTTATTGACTGTGCAAAAAGTCTTAGAGTACA
TTCCTTGGAAATTGACTCTGATTCAAAGTGTTGCATGACAACGGGATATGGGGAGTGTTCTCTGGAGATACACCCAC
AAGGAAGAGAAGAGCACAAGGGAGATTGTGGGAGAGTCTGAAATGTGATTTGTCTGCAGCAGAGGCCTAAGCCAGTC
TCGCAGGAGCCCTACATCTGGGCTGGCTGTGCAGAGCTGTCCTGAATTGCAGGCAGTGGGCCTGGCCCTTGTATTCC
TGATCCAGCCAGCCATTGGCCAGGGGCTGGCTGCTGCCTGAGAGTGGAAGGACAACTTGGACAAGTTTTCTGAGGCC
GAAGGCAATTCTTAGTAAGGAACACCATTAACAACCAATATTCCTAGCATCCAGGGATGTGTGCATTGTTCCTGAAG
AGGGACAAGTATGTCTACAAAAATCACAGAAACCACAGAAACACACACAGTCCTACTAGCACCTCTCCCTGTCCCAT
TTGCAAACAATTTAAGAGCTCTCCCATTTTTAGTTCAAGAAAAAGAAAAATGGATTGGGAGGACCACAAGCTGACTT
GGGGGAGGAATATTTCCTCATTTAGCTGTAGTTTTAACTTTTGTTTTCACTGCATATTTTCAGTCTATTTTATTTTC
TTTCCTCTTCAGTTGTTGATAGAAGGTATTCATAAATTCTCATGGCAATGTTAATGCTGGCTTTGACTCTCAGGGGA
AAGAGGCCAGAAAACTTCTTTGCTGTACCATTCCATAATTAGGCAGAACTAAAAACATCTTTGGGTGTTGTTTTTTG
TTTTTGTTTTTTTTTTTTGCCTTGTCTGCTTTTCAAAGATCAAATGATTGAAGCATTAAAGCATGGTGACTGGTTCTT
CAGGTAAAGTTGATTTTTATTTTATGTCAAGTAGAAAAATACTGAACTGGAAGAATCACAGCTGGGGTAGCACAATC
ATAATTCATTAGAAGGCATAAATAGTGCTTGGATTAAAAGAAGCCCTACAATCTGGGGACAGTGCATCTCATGTGCC
CTCTGGGATTACTCGGCAGTCATCAGAGTTAGATTTAACGACTTTGGAGACTTAAGCATTATGGTTTTTTTTTTTTG
TCAATCTGGGACACTGAAATTGCTGTATCAGGGTTATACTCAACTGTGTCAGGTTTATTTGTTTTTATGAGCTGTAA
TTTTTGGTTCCCTCAGCGCATATGCATAGTTTGTTCCTATGTTATCATTTATTGGTGTCTGTTTTCTGGCTGTCTCT
GGTAGGTTCAGCCTCAGACTCTGTAACTCCATGAAGAGATTATGTTCCAATGATGTTTTATAAGTTTGTTAAACTCT
GAACTCATGAGTTTATGTCCCATATAAGCCACGTTACACATGGTAGGAAGGCTCCAAAACCAGGGCGCCGAAATCCA
TTTAACGTGTAACTTACCTAAATGTAACAATGTTTATAAGAAAAATACATTGGAAGTTCCAGTTTTGACTTCCAGCA
ACATATATAATTCATCCACTTTATTTATTAACTTCCATGTGTTGAGCATCATACTGGTGCTGCGAGTACAGCATAGA
ATAAAAGTCTCTCCTTTCATAAAACATATATTGTAATTGAAAGAGAAAGACAATAAACTAATGAAGAAAATATATAC
TGTCTCAATAATTATAAGTGCTGTAGAGTGTAGTCTACAGATTGATGTCAATGGGTATTTGTTAGACACACAGAATC
TGAGGCCCCTATCCTAGACCCACTGAATCACAATCTGCATTTTAATAGAATCCCCAGGTGAATCCTGTGCACACTGA
CATTTGAGAAACACCATTACAGAGAACAACTAAGCAGGGAGATGGGATGAGGGTATTATTGTCTATAGTGTGGTCAG
GGAAGCTTGTCTGTTAAGAGAACATCAGAAAACTGATGTAAGTGAGGAAGTGAGCCTGGTGTATTTCTGGGAAAATT
ATTCCAGGCAGGGAGAGAAAAGACTGAGCAACGATACTGAAGTAGGAACAAGATGACGGAATATTAAGGAGATCAGT
GAGACTAGAGGAGTGGGTCAGGGGAAGTGTGATGGAAGCCATGAGAGATACTCATCTTTCATAGCACTGCCCTACTT
CCTTCTCCCCAACATGAGGGTCTCATCACCCCCCACCACTCTTGTCTTCTCCTATGTCCTCCACATTGCTGCCAGTA
TGGAGAGTCTGGGAATGCCCTCAGCTCAAAGCTGTTTGGTGATAGCTGGCAGAGTTGTGGTAGTAGCTAAAAAAGAA
TTAAGGGAAAGAGGAATTTTCTCAAAAGCAGGTGCTTTTCATCCTCTTTAGCAAACCGAAACAGATCTGAGCATTAA
GTCAAGATGTTAAATACACAAATGTTGAATGAAAAAAAAAACAAAAGGTAGTCATTTAAATTCAGAGCTGCTTTTAT
TAAAATAAGATTTTCTTTTTTCTTTACTGTGGTAGTTCAAATATCAGAATAAAGAATTGTTTCTATTCCCGACTTCC
TGACTTGCAGGAAGTTAATCAGAAATAAATGCAATATAAAAAAAGAAAATCTAATTTGTATTATGCTTCTTGTATAT
GTTTATTATTTCATGTACTGTATTACAATGTAATAGAATTTATAATTCATTATAGCAGATTGTTTCCATTGCATTCC
TACTATTAAATATGTAGAAGCTACACATATACTTGTAGCTTTAACATATATGTCTTTATCCTCAAAATAACTGCAAA
GAACATATAGATAATTTTTAAAGATTAAGGAGCCTGAGGTTTAGAGGGGAGATAGCTAGATTAAGGCCACACAGCTA
GAAAGCAAGCAAGCAAGGGTTTCAGTCCACATGTCAAGCTCCACAGCCTGTGTTTTGTTTTGTGGCTGTGCTTTACA
CTATCTCTCTGTCCAAGAACCTAATGGAAAATTACAGATACAGATGCAGCTGGCCAGCAGTTAATATAATTTAACTC
AATCTTAAATTTATCTGGAGTAAAAGTGATACAAGTTTCCGTGTTTTTTCTTTTCTTTCTTTCTTTTTTCTTTGTGT
GTGTGTGTGTGTGTGTGTGTGTGACAGAATCTTGCTCTGTTGCCCAGGCTGGAGTGCAGTGGTGCGATCTCGGTTTA
CTACAACCTCTGCCTTTCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATACAGGTGCGCACCA
CCATGCCCAGCTAATTTTTGTAGTTTTAGTAAAGACGGGGTTGCTCCATGTTGATCAGGCTGGTCTTAAACTCCTGA
CCTCAAGTGATCCGCCTGCCTCAGCCTCTCAAAGTGCTGGGATTACAGGCATGAGCCACCTTGACTGGCCTGTTTCT
GTGTTTTTTCTACTTAAGAAGTAGAAAAAATTGGTTCACTCTATTTGAATTTCTTAGAACCATAGAAATCCAAACTT
GGAATAATTATTGCAATTATTATCTAGTTGAGTATTCTTCTTTTATAGATGTAGAAGCTGAGGCCTAAAGTTGCTTG
TTTTGTTTTATGAACACTAAACCAAACTACCTTGTGGCTGCTTACTTAAATTATAAATTATAATGGGGTGGCCTTGC
CTGAGCTGTATAAATTGTTTTAATTATCAGGACAAATCAACATACTGGAAAAAAAAAGCAAAACTTGCAATTGTTGT
TGCTTAGACACCTGTCCATATCAGTTTCTATTGTTGCATAGAAGCACCCCAAAACTTAACAGTTCAAAACGAGTGAT
TTATTGCTCATAATTCTGTGTGTCTGAAGTTTGCTCTGTGATTAGCTGGATAGTTATTCTGATGGTCTTGCTTGGTG
TTACTTATGTGGATGCAGTTATCTAGCAAGTGAACAGGAGCTAGATGGTCTAAAATGCACTCTCTCAGATATCTGAC
AGCTGGCTGTTGGTTGCAGAACCTTGATTCTTCATAAGGCCACTCATCATCCTGTAGGCTAGATTAGGCTTCATTAC
ATGGTATTCTCAGGGCAGTTTTCCAAGAGAGGGCGGGTGGAAGCTACAAGACCTTCTGATGCCTAGGCTTTGAAACG
TGTAGGTTACTTCTGCTAAGTTATATTGGTCAAAGCACCTCAAAAGATCAGCCCAGATTCAAGAGATGAGGAAATTA
CTTCATCTCGTGAAAGGAGGAGATGCCACATCGCATATCAAAGGGGTATGCATATTGGGGATAGAAGGTTTTATTGT
CACCGTATTTATACACAAATCACTACATTGGGGAAAAGGAGGGAAACTGGAGATCAAAAGTGTTGGTCTTAATCCAA
CTTAATTCTCAAAAATTACCATGCGTTAGAACCACCCAGTTCTTCGCAAAGTATAGATTACTGGGCTTCAACCCCAG
AGTTTCTGATTTACTAGCTCTGGGGTGAGACCTGCATCTCTCTCTCTCTTTTTTTTTGAGACTGAGTCTTGCTCT

[0338]

GTCACCCAGGCTGGAGTGCAGTGGTGCAATCTCGGCTCACAGCAACCTCTGCCTCCTGGGTTGAAGCGGTTCTCCTG
CCTCAGCCTCCTGAGTAGCTAGGATTATAGGCACCCGCCACCACGCCTGGCTAATTTTTGTATTTTTAGTAGAGACA
GGGTTTCACCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCTGCCTCAGCCTCCCATAGTGC
TGGGATTACAGGCGTGAGCCACTGTGCCCAGCCAAGACTTGCATCTCTGAAAAGTTCCTAGGTTATGCTGATGCTGG
CCTATGCTTTGAGAACTACTACCACAGACATACAGTGAGTGGGGAAGAATAAATTCATCCCTTCTGCTGTGTGCAGC
AAGGAGTGGGATTCCAATGAGATCCAGTGCTGTGAATGCTAAAGGGAAATCCATCTTATTTTAGCACCTCTACTCCC
CATCTCCCCACCCCGAGGATGTTATAGCTTAGAAGTTCAAGGAGATGGACAACACACTAAACCAGGCAGTATTTGCC
CTGCAGAGCTGTTCAGTGTTCCTGGATGAGACCTCTGAGAAGAAAAGCCATAAGTTCCTCTAGAGACTTTCACAATC
ATTTAGGTAGACAGGACTTTGCATGGGTCTGAAGGCTTGCATGGCAGATGGAGGCAAAGAGCCAGCAAATCTGGTTG
TAAATGTCAATGTGAATCCTTTCTTATCCACAAGCTGCTGGGCCTGAGAACATTAATGTTCTACAATACCCGATTTA
GCATTTTTGAAAGAAATTGCATATAGACATGCTTAATGTGAAGACTCCAAATCAGGATATTTGATTCAAATGTCTCT
TGGTAATAACTATGGAATGAATAACCCATTGTATATGGACATATAGAAGAGCCAGTTAACAGAGTTTTCTTTTTTTT
TTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCG
GCTCACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCCCGTAGCTGGGACTACAGGCGC
CCGCCACCACGCCCGGCTAATTTTTTTGTGTTTTTTAGTAGAGACGGGGTTTCACTGTGTTAGCCAGGATGGTCTCG
ATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCC
CAGAGTTTTCTAGTTGATTAAACACTGGTAAAATCATCTCTTCCTGTAATTAAGTTTGGAGAGAGCAAGTCTCAAGT
GTAATTAGGAAAGCACAGTATTGGAGGTCAAGAAGCCTAAATTCTGATATCTCAGTTGTGCTACCAACTATATAAGG
GATCAACAGACTGTTTCTGCAAAGGACCGATACTAGATATTTTCTGCTTTATGGGTCATGTGGTCAACTCTGTCATT
GTATACTAAAAGCCACAGACAATACCTAAATAAAGGGACATGGCTGTGCTCCAGTGACACTTTATTTACAGAAACGA
GGCAGGTAAGATAGATTTGACCCTTAGTGGGCTACCCATTTCCCCTTAGCACTCAGCCTTCCTATAAACCATGAAGC
GTCTAACTACAAGTACCTGTGAGTCTTCCTGAGTCCTTTCTCTGCCTGTAGGAGTAGCTCTCCCACTTGCAGAGCAG
GCTGGAAGCTGGGGAGGAGATTATCTCTGGAATAGCACTTAACCAATGGACAAAAGCTGGGGGATAAAGGTGTTGGT
CTTCATAGTTTTCATACTGCTATGAAGAAATACCCGAGACTGGATAATTTATACAGGAAAAAGAAGTTTAATGGACT
CACAGTTCCACATGGCTGGAGAGGCCTCACAATCATGGCAGAAGGCGAAGGAAGAGCAGAGGCACTTCTTACATGAC
AGCAGGCAAGAGAGCATGTGCAGGGAAACTGCTCTTTATAAAACCATCAGATCTCATGAGACTTATTCACTATCAGG
AGAACAGCACAGGAAAACCCCACCCCCATGATTCAGTTACCTCCCACATGGTCCCTCCCACAACACGTGGGGATTAT
TGGAGTTACAATTCAAGATGAGATTTGGGTGAGAACACAGCCAAACCATATCAATAAATATCCCACCTTCTTCTCCC
CTTGGGTAGGACACCTCCAATGCATGTTCCACACTAAAAATCTCCAGTGAATTGGGCAGTTGACCACAATGATAAGC
ATACTTATTAGCATGCCTCGTATAGGCTTCCTTCCGATGTGTAGGCTTTCTTCCCATTTTGGCAGAATCTGGACCCA
ACCCCACATCTCCTGAGTTTGATTTCAGTCCTCTTCAGTATTCTCATATATCTCAACTTTGTTCCCTGCTGAGATAC
AACTGACAGATATATATGTATCTTCAGCTCTGCAGCTCTATCTTCTCTTCTACCATGGCTTTCTGGGATCAGCTCCC
AATTGAATTATTTGTATTTAAATCCTTGTCTCAGGGTCTACTTCTGAGCAAACCTAAATTAATACAAGTCTTAGAAT
TAGTGCCACAGTTATACCTTTTGAGTAAGTTATGTTGCAATTCCCATTTTATAGATGAGATAACCAAGGGTCAGAAA
GGTTAAATAATTTGTCCAAGGTTACAGAGCTAGTCAGTGACAGAGTGGGTATTCAAAAATAAGTCTCTATGATTCCA
GAGCTTATGCTATTAATCAGTGCACCATGTATCAAGCAGTGATTTAGTTATCTTCATTTGAATTTTTTGAAGTCTCT
ATTTAAATTGAATATCAAATCCTCACATATAAGGTTATTTATACCTTTTTATTTGTTTTATTTATTTATTTATTTTT
GCTTAACTTTTTTTTATTATACTTTAAGTTTTAGGGTACATGTGCACAATGTGCAGGTCTGTTACATATATATACAT
GTGCCATGTTGGTGTGCTGCACCCATTAACTCGTCATTGAACATTAGGTATATCTCCTAATGCTATCCCTCCCCCCT
CCCCCCACCCCACAGCAGGCCCCAGTGTGTGATGTTCCCCTTCCTGTGACCATGTGTTCTCATTGTTCAATTCCCAC
CTGTGAGTGAGAATATGCGGTGTTTGGTTTTTTGTCCTTGCGATAGTTTGCTGAGAATGATGGTTTACAGCTTCATC
CATATCCCTACAAAGGACATGAACTCATCATTTTTTATGGCTGCATAGTATTCCATGGTGATTTATACATTTTAATT
TGAACTTCACTCACCTCTACAAAGTTAAATAAAGCATTCCATTTCCATTTCAATAATACTTTAAATATCTAGGCAGT
TAAAGCATGAGGTTACTTGGCACTTAAATGTACTCTTGTCATACTACTTTTGTTGGCTTAAATTGCAAAAAAAAAAA
ATGGTTAATTTATTGTGACAATACCCTACATTTATCTGCAGTGATCATTTTTTTTGTGAAAATGGCCTTCGTTTATC
TGCAGCAAAGGAAAAAGAGGATGGCAATTAGTTCTTGCATTCTTATTCCTCTCTTGGGTCCTGATCCTTCTCATTAA
TAGAAACATGGCAGGGGAGGGGTATATAACCCACACCCTTTCCTGTTGTGGTTATGTTTCCACTGTTGATTCTGCTT
CAGGTGAACCTTTAGGATTAGGCAAATAAATTTCCGTGAGGCCAAATCTTTTTCTTCCTCATTAACAGATGATTTCT
CTGCTAAAAACACTTACGACATGGCTATACTATTGCCGGTTTTATAGTTACAGGCTCTAAACCTTGAAAACTTCCTC
AAAGTCTAATACGTCAGGAGCAAGCTTTTGTACAAAAAATGTGAAGACCCTTAATCAGTTCCAATAACAAAATAAAT
CCATTTTAAACCCTATCCCAAGATACTGCAAGGCCTTGGAGCAGCTGGAGAGACTCCTTAACTCTTGACATTAATTA
ATTAATTTAAAAATTCATATTTGTATGTATCAGTGAGGTAAGAGTGCTTGAAATATAATGAGATGTGTCACACTGTA
GAAAGGGGAGTGACAACAGAAAGCCCTGGCTGGTGAGGCCCCAGCACTTCTCACACTCATCAGAAGGAAGTCTTTCC
ATGAAGGCAGTAGGGGGTCCCTGGTCCCAGGCCAGGTCCTTCAACAATGCTCACTGGTTAACAGGAAAGGCACTACA
GTGGCATCATTGTTAACATCCAAGACAGTGATAGAATGTGAAACTTCTACTGTTTAGTATTTAGTATTCAGCATTTA
GATGTTAATTATCCATTTTGTGATGAAGTTCCCTTTCTTCTCCCTCTCTTAACCTTTTGGTAGTTTTATTGCATGGT
TACCATTTCCAGTTAGGGTTGTGCTTTGGGGTCTGAACTGATGAAGGAAGAGAAACTCTAGTTCATCATTTCTAGGA
AAAAGAGAAGGCTAACATCAATTCTGATGATTAGAGATTTTTTGATTACCTATGTCCTGCATTTTAACAGAAATAAA
ATGTTAATTACAGTAAACTTACTTTAACCTTCTTATGCTATTTAACACTCTTCAAGAAAGGACTTTTGCTTGAAATC
ACTACAGAGTATACCATAATCTTGACTGTCATACTCTGACCAAGAAACCAGAACTATTTAGGCATATTTATGAGAGT
AAGTACCCTACCGCTCAAAATGGAAGGCTTCAACATGTATTCCTAATATTTCCTAAAACTTCACTGAATAGTTTTAA
AATTGAATATAACTTTACCTTCAGAGAGAAACAAGATTTCGAAAGGAAATTGTAGCAGTTTTGTACTTCAATTGTCG
ATTTTAAGATTGTGGTCTGAAAAGTATTTGAACAGTCATTCTTTCTTTTCCACTATCTCAGGAAAGGCTCTCATTCT
ATTAGAAGCAATCTTAGAAGCGTAACTGCCAACTCTCTTCTTAAAAAGTGAAGAGCAGATGGAGTTCTACCAATCTG

[0339]

CTATACTTGACAGTTTTTCAACTACCTAACCTAAGTTCCAGTTACTGTCTAAATTATTTTTATTCGAATAGGAAAAA
TACATTTATGCTTACTGTACATTTCGACAGTGTTCATTCTTTGAGGTCTGTTATTTAATCTCCTTCCCACTTTTCTT
TTCTATCTTAAATCAGATGAGTCCCCATGGTCTGGCAACAAAACAGCTGTTTCTCAAATTTTTAACGTCTCTGTGGC
CCTCCTCCTACCCTTCCAAGTGTTTTCCATACTATTTTTCTGTTGTGGGCTTCAGAATTGGATGCTGTTCCATAATT
ACTGGTCTTCCCATGACTCTGTCACATCATAAAGTCTATATATACCCTTTTCCTACGTTTACATGCCAAATACCAGG
TGTGTCTAGCTTTAATTCAATGACTAGCATATGGTAGGTATTGCACTAATATTTTTTGGTCAAGTGAACATATATTT
TACAGATAGATCAAATACATCAACCTTCCCATGTAATAATAATAATACCAAGAAGATTAAATGATCGCGTGCTTGAC
CAGCTGTTTAGTGGCAGAATCTGGACCCAACCCCACATCTCCTGAGTTTAATTTCAGTGCTCTTCAGCATTCTTATA
TATCTCAACTTTGTTCCCTGCTGAGACACAGCTGACAGATATATATATATATCTTCACAATCTCAGTAGTAAGCCCT
ATTTTTTTAAGATTACAGTGCTGATCAAAGAGGGATATTCTATGGCATTGATGCAAATCTTCCAGGCCAGATATAAC
CCCTATCTCCTTAAACCCTCATTCTTGACTGTGTCTTGAAATCAAATTTGAGAACCTCTGCTTTACATGCTACCTTC
CTCTTAAAATTCATAAGGTTTATTCCTTCTCTCATCTTCAGTATTTTTATGAAAATATGTTTTACTCTTATTTCCAA
AGTGCTTGCCAGCCCATGCTGAGTTCATTAGACAAACAATCTAGGTATCCTAATATTTAGGTAAAATTGCTAGCAGC
ACTAACTATACCAGTACCCACTTAAATTGCAATATAAACCTACAATAGGAAAAAAAAAAGTCAAAATTATACTACTTT
CAATTCCTACTTCTGGAATAATTATCACACCTTCAAAAAAACTCATAATTGTTCTCCAATATTAAAAACCAGGAACT
AAATTACATCACTATATATATATAGCTATATAGACATATAACTGTATATAGTTTATATATATATTACATGTCACTAT
ATATGGTTATATATACAGATAGCTATATCTCTATATAGTTTTTATATAGGAAACTATATATATAAACTCATATATAT
TTATATATAAACTCATATATATATAAACTCATATATATTTATATATAAACTCATATATATTTATATATAAACTCATA
TATATATAAGAGATTTTATATGTATATATGAGAGACAATATTGAAATAAAACAAACTAGCCAATCTCCAATGTCCCT
TCATTTTCCCAGGACTCTTCTTTATTCTCAAAGAAATGTATGAATACACAATATAAATAAATGTTCTATAATGTTCA
TTATGGATTATAATAGTTTTCTTGTCATACTGAAGTGAATGGGAATTGTTTCTTCGATTATTAGCATACTTTTCATA
ACTATGTGATTTGGTGATAGGCACTTTTGCTTACTAAGTTCAGCATGATCAAAACAGAACTCCTATTATTTTTATTG
TTGATTCTTATTTCTTTAAGGTAAAATATGCCCTTTTCCTTTGAACCTGCACACTGTTTAACAAATAAGCACATAAC
TCAGGATTGAATTGTACACTTCGATTTGAGCTTTTTTTTCAAGGTCAGGACCCAGTTTCACAAGAAGTTTTATTTTT
TCCAATACAACTGACATCCACTCCCACCAGCTGAAAGACAAGAAAAACTTGTCTAATAAAGCTTTCAGATTCAATTT
GCTGCCTGCATACAGCTTGAGGAATCTCTGGAGGTCACTCACAGCATGTGTTGCAACCCCAACAGGGAGAAGTAATG
AAAAGATTCTAGTTAAAAAGCTGACACTGCCCCTTCCAACCTCTTTGAATGTGAATATAATAAGCCAGTTTACAGAC
GCAAATCTCTATGATTCTGGGGATTTCCATCTTGATCTCTGACTCCAAGGAACATTTGAATGCATGGATTTGTATCC
ATTATCTGGGTGAATAAATGCTTCATATTGAAAAAAGGGGTGCTTTAACAACATAAGTCTGATGTAAATCAGGCAAA
ACAACATTGTCACTTCATGTTTAACTCTCCTGGAGGGTCTCTAAGGTCTCACAGTTTGGTTCTATTCCAGTAATATA
TAGGCCTATCATAGCCATTTTCAAAAATAATACCTGCTTTCATTTCGATTATTCCCCCTAGCTTTTGCATTGACCCG
AACATACCTAATATTTATCTTAGGGCTAACACGCATTAATGCCTTGCTCTGTACCAGGCATTTTGCCAAGTATTCTT
TGTGCATTTTTCTGTTTAATCTTACAGCAGCCTTATGAAATAGGTACTACATTATTATTATTTTTCACCATGAGAGG
AAATGAAAGCCTAGAGAGAATGGTTATCCAAAAACACCCAGCTACTAAGTGGGCACAGCATGGCCTTGAACCTGAGT
CTTTATAAAGTTCATGCCTGTCTTTTACCTTTATGTTAAACATACTGAATCTTGGTCATGCAGTCTATGAATGAAGA
CTCCATATACTCTAGGACCAATTCTACCATATTGTGCATGCTTTTGTACATATTTTCCAAATGAAATTAAACAACAA
TACTCTCTCTTCCCCTTTCTTTTCCGTTTTGCATGACATCTAAATCTTTTATTAAATCTCCGTGGGTGGAATTGGGC
CTTAAGTAGTAGTACCTTTGAAGCTTAATACTATAACCTCAGAGTTACGGAAGTGGTTTCAATATGAAGAAATATAT
ACGTTCTTTTCTTTTCTTTTCTTTTCTTTTTTTGGCTCCTCTAGAATAGAAGGCAATCAGGAGAGAAAAAGACATTA
AGATGATAGGCTTGATTCTCCCACAGTGTTACTACTTAGCTTTATTTATTCTCCCACATCTGTAACTGTCAATTCAA
GGTCAAAAAGCCAGTACAGCAGCAGAATTACAAAAAGTAGATCTGGAAAATCTATAGGGTACCATAGTCCAGCACCC
TGCTGCAAGTCAAAATCAATTAAATAAATATGTTTAAACAGCACTTCATTTAATCTTAAGGGCACCACGACTCCCTT
GGAAAATTCTTTTTTGTGTTCAGACCTGATTCTCTGAAAGTATTTTCTAGTGTTTTCTTTGTTCGGTGGCTGTGGTG
AATAAGTGGACCCAATTGCCTACAGCATGAAATCGTAGAAATGAATATGGGCTGGAACTTCAATCAATCACACAAAC
CAGAAACATAGAGTTCATCACCTCTTCTGGAAAGCGTCCCCTTCGGCCGTCAAAAGGGTAGAGTTTCCTCCAACATA
TGTTCATACTTCCCTTTGCCTCCCATCTGCAATTTCAACCAGTCTGTTGCTCTCGCCCTTTTTAGGCTTCACATACC
GGATTTCTTCTTAGGAGCTCACTTGAAAAAAGGGTTGTATGCTTTAAAAATACTAAAAGTCACTGGACTAGATGATA
ATAAAATTTCTGTAAAATAAAAAGGGAGTTAACAGAGGTGTTACCACTTATCATAATGGATATTGTTGACCTTTCCT
CTAGCCAGTGTCCCACTTAAGCATGCTTCTTGGCTGGAATTTATTTTTCTCATCTGGATAAATCTGAAGCATTATAG
TAAAACCCAGTGTGAGCAAGGCCATGGATGCAAGTGGATTGAGATGAATGAGTGGATAGACCTGCGTGGGCAGAAAT
AATGAGGTCAGCAATAAGCCATACCAAGGGATGCACACTCAAGGAATAAATTGGAGGTAGCAGAATAACTGAACAAA
TGATACATTTGATCAGACTGTCACATGAATATCAACTGAAACAAGTTATGACTGTAGTTAGTGAACCCTAGAGTGAA
GAAAACAAAATATTGAAATCCATCTGATCAAAAAATAATAAAATTGTTTACATGTATATATTATCTAAAATTTCCCT
TAAGAATGAAGATTATAAATTCCCGCAGTTGAGGAAACTGAATAAGAGGAGAATCAGGACAGAATCTTTTCTCTTCG
TATTCCTAGTCCATTTTTCTTCTAACTACAGTAGTGAGAACAAGAATTTATTCACCAAATCATAATACATTGCAATT
AGGGGATGCCATTTGAATCTTGGAAGACACGATTTGAGGCAAATAAAAAATCCATATTTATAAAGTCAATTATTGGT
TTATATAATTTACTGTTTCAGCGGGAGGTATAGGTTAGTAATGTAAGTCACTTCAAGAAGGTTTTGGTCGAAGTTTT
GCCAATTTTTTCTGTAAAAAGCCAGATAGTAAATATTTTGGGTTTTACAAGCCAGAAGATCTTTGTTGAAGTTACTC
AACTCTGCTAATGAGGTGCAAAGCAGCCATAGGCAGTTTGTAAATGCATAAGTGTGGCTGTGTTCCAATAAAAGTGT
GTTTGCAAGAACAGGCAGTTGGTTGGATTTGGCCCACAAGCCGTAGTTCACTGACCCCTGTTTTAGATAAATACTAG
TCAATAGATTCCTTGTAGCTAATAAAGGGAATATAGATTGTTTAATGTGTTTCCCTAATCTTCTTAAGGCTTCCTCC
AGGGGAAAAAATATTTCTTATCATCAGCAGTGTTATTAATGCTTACTCAGAACAAAGAATACCACTTACTTAGCATA
GAATGGACATTTATAAAGCCACTCATGGAAAGAATAAAGAGATGTGAAATACCTGGTACAACTCTTAAGAGTGTTCC
ACTATCATTTGAAAAAATGATGAGTAGTAGTTATATAAAAGGACTCGAGCCTAAAACATTTGTTCTCACTTTGGATA

[0340]

ACGTTTTCCTATCTCAGCATCAAAAAATACAAAGAAAGGGGATAACTAGGTCTTAGATTTTCTAAATTCTCATGACT
ATAAGTCCCATATATTTAGATATTGAAGTACCCTTAAATATACCTAAGTAGAAGTAATAATTTCGATATTATGGGAA
ATCTAATTACTGCTCATAGTTCTCGAGATTAAAATGTTATACACCACATCTCAGCTTTCTTTGAAACTTGAAATGGC
ATACTTGTACAAAGGTAGAATGGTCTCCCAAAGCAGCGTCATAGATTAGCAGAAAAGAGCTATGATAATGCCACTTC
AACTGACTATCATTTAAATAATGAATTAGTTAGGAAGAGTTAAATTATATAGTTACAGGCAGCAGGAGATGTCACCA
TCAGGGGGATAAGAAGCCATTTTCCAGTACACTATTTGACTTAACTGGTGCAGTTCCTTCTTTAACTGTAGGTTACT
CATTTGTATCCTTAATTTATCAATATACTAATCTGACAAATAATTTATTAAGGATTACTGTGTGCCCTCCTGTGTGT
TAGATTCTAGGGATACAATATGGATCAAGATAAGCGTGATTCCTAGACTCCTAGAAAGTAAGAAAACTAAATAACTG
TACGTGACAAATACTAGGATGTAGAATACTAGGGAAGAATCAGGAAGAATCACCTCACCCAGTCTGGAGTATTGCAA
CAGAGCGATGTCAGAGGAAAAATTCTGGAAAAAAAAAAAAAAAAAGGAAATTTGAAGATGAGACTTGAAGGATAAGTA
GGTGTACCTAGGCAAACAAGTAGGGGAGAGTGTTCCAGGCAAGAGAAACAAGTAACAGTTGGAGGCAAAATGGAGCC
TAGCACGTTTTAAAAATATCAGTATGGCTGAAGCACTGAGTTGGGGTTGGGAATGGGTGAGGAGCGATGTGTCAGGG
AGGACATTCCAGATGATGCTGGAAATGCATGCATGTTCGCAAACTGCATTTTAAGACATGTTAGAGAAATCAAACAC
ATTTTATTTGAAGGGTAATTGAAAGCCACTGAAGCATTGTAGAGAGACTAGTGACATGAGCATATTTGGGTTTTAGA
GGGATTACTCTGGCTGAAGAATAGGTAACTGATTGGATAAGAGCAACTTTGGAGGCTATTCCAGTGGTTCAGGTGAT
AGATAATAATGACCTAAACTAAAGAAGTAATAAAGGGAATGGAGACAAAGTAGACAGATTCAAGTGATACTTGGCAC
GTGGAAACAACAGGCCCTGGTAACTGATTGGACATAAAATGGAGAGATACAGAGAACGCGAATAGAACACCCAGGCA
TTTGGCTTGATAAACCGAGTAGACCGTGGTAGAAATTACTAAACTATGGAGTATTTTAGGGTTCAGGGGAACTGTGA
GTTCAATTTTAGACATCTAGTTTTGCAAGGCCACTGAGGTAACTGAGCATATGAGTCAGGAGAGAAGTTTAGGCTAA
AAATATGGATGCCAGAGTTAACCAGCGTTATGATAAAACAAAGCTGAAGTCTCGAGGGTAAGAACACCTAAACATCT
AGGCTGGAGAAGGCAGCAGGTAAAAAGTAAAAGGAGAGGCCAGGCACGGTGGCTCACGCCTGTAATCCCAACACTTT
GGGGGGCCGAGGTGGGTGGATCACCTGAGGTCGGGAGTTTGAGACCAGCCTGACCAACATGGAGAAACCGTGTCTCT
ACTAAAAATACAAAATCAGCCAGGCATGGTGGCGCATGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAA
TCGCTTGAACCCAGGAGGTGGAGGTTGCGGTGAGCCAAGATTGTGCCATTGCGCTCCAGCCTGGGCAACAAGAGTGA
AACTCCGTCTCAAAAAATAAAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAAACAGAGAAAGAGAGAGAGGAAGACA
AAGAAAGGAAGGAAGGGAGGGAGGGAGGGAGGGAGGGAAAGAAAAGAAAAGGAGAAAAATTCAGGAGAATGGT
TACTTCCAGGGAGATGGAGGCGATTGTGCTGTGGGGAACACAAGGGTGGGGTCAAGGTATTAGCAGTATTCTATTTC
TTGATTGGGGTTGTATTTACATAAAGTGTTGCTTTATAATTATTCTTCACACTTTATGTGTACGTTCTATGTAATCA
TCTATAGATAAGACAGATTTCACTGTAAAAGAAAATAAAAGCTTCCAAAAGATTATCATCACAATTGTAACAGATTC
CCCTGGTGCCTGGAGTCACACGCCATTTTCCTGCACTGCAGTTGCAGCTGCAGTGGACAGCCCTGTGTGAGTTCAGA
CTTGCCTTTAGCTGACAGCATCCCATGTCAAGGGAATGGCTCCCATTTTTCTACTTTCTATCTAAGGGACTTCTCTG
ACATCCCAGGAGCCCACAGATTTTGTGAGCTTTCTCACCCTTGAAGTTTTAGTGAGTGAGCAACCTTCAACCAATGG
AGATGGGAGCCCATGGATATATTTTTAACCACTATTCCTTCCGGGGGCAAGGGGAATTCTCTGTGATTCTCAGGAAC
ATACAAAAGTTCTGTCAAAATACAGTCCCCATGGTCCATAAGCATTACCTTGATGATAATACATTTGATTGGCATTT
CCTCCCCCTCTGTCTCACTCTTTTGGTTTTTCATTCTTGCTTCCTAGGGATCAGCTTCCAAATGAGCTATCTGTACC
CAAGTCCTCATCCAGGCCCTGTTTTCAGGGGACCCAAAGACAACAATAAGAAAAATGGAACTGAAAGAGGAGAAGAC
TTTAAGGCATGAAAAAAGTTCCTCTGTATTCCATACTGCATATTTAACTGCCTACTCAACAGTTCCACTTAGATGTC
TCAAAAATAATCTCATGATCTTTTACTGCTATGAAATGCATGACCTTCCCCTGATATATTCCTTTCTTCAGGTTTTG
TAGCACCACTTAGCTATCCAGTAACAAAATCTTGGGGGTCATTCTTAAAACCTTCCCACCTCACCCCTGCTGGACAT
CCACCACTAAGTTCAGTTGATTTTTTTGCTCCTAAATATTTCTTGGTTCAGCTTTCATTTTTAGATTTATCCTGCAT
ACCCCTGTACAATCTCATCATCTCTTTCCTGGACTGTTACAGTAGTCTTATATTAGGATGATCATAGTCTTGATTTG
CCTAGAGAAAACATGCTTTATGTCATTGCCTCAGAGTGACTAATAGTCCCACCTTTCACTTGTAACACTATGCAGGT
TATTGGGTTAAATAATACTGTCATCCAAACCTAATCTCTCCTCATTCATTTCTTCTCTCCTATGCATCTTAGCTGTG
TGAAAATTCAAAATATAAATTTGATGACAGACAGAACTCTGTTTAAAATGCTCCAGTACCTTACCATTTATCTCAAA
ATAAAAATTTCAAAAAAAGAAATGATATCCCACAGGATCTTGTATAGTGTGACCCTTTCCATTTCATCATCCTTATA
CCAGATACTGTGAACTACTTACTCTCTTATCTCCACATTGACCTCCCTCCTAGTTTTGTTTTGCTTATGAAAGTGCT
TATTTTCTGTTTTTTAAAGCCATTGCACATATTGGTCTCTCTGATTGAGACACTATCCTTTTGAATTTTTGACCTCA
TACCTACTCACTTTTCAGGTCTCAGCTCAAATGTTATACTCAGGAAAGACGCTCCTTACCTCCCAAACTAGGTTAGT
GTAAATGGCACTATTTATACTCCTCTTCCAGAGCACACACCAAATCTTATATTTATTTGTGGGACAATTTGAATTCA
TGTTGTTCCCCCTACTCAATTGTGAGCTTCTTGAGGACATACTCTGCCTCCTCCTTACCTAGCTTTATTGAAGTATA
ATTGAGAAATAAAAACTGTATGTATTCAAGGTATACAACATGATGATTTTATATGACTATATTGTGAGATGATTACC
ACAATCAAATTAATTAATACATCTAGCACAAGAAATAGTTACTATTGTGTGTGTGTTGGGGGGGGGGGATGAGGACA
CTTAAGATCCAGTCTTGTAGCAAATTTCAAGTAAACAGTAAAGTATTATTAACTATAGTAACCATAATGTACATTAG
ATCCCCAGACATCTTATAACTGAAAGTTTGTGCCCTTTGACCAATGATATGGTTTGGCTGTGTCCCCACCCAAATCT
CACCTTGAATTGTAATCCCTGTAATTCCCATGTGTTGTGGGAGGGACCCAGGGGGAAGTAATTGAATCATGGGTTTG
TTTCCCCCCATGCTGTTGTCGTGATAGTGAATGAGTTCTCATGAGATCTGATGGTTTTATAAGCATCTGGCATTTCC
CTTGCTGGCACTCATTCTCTCTCCTATCACCCTGTGAAAAGGTGCCTTCCTCCATGATTGTAAGTTTCCGGAGGCCT
CTGAAGCCATGCGGAACTGTGAGTCTATTAAACCTCTTTTCTTTATAAATTACCCGGTCTTGGGTATTTCTTCATAG
CAGCATGAGAACGGACTAATACAGTAAATTGGTACTGCAGAGAGTGGGGTATTGCTGTAAATATACCTGAAAATGTG
GAAGTGACTTGGAACTGGGTAAGAGGCACAGGTTGGAACAGTTTGGAGGGCTAGAAGATGACAGGAAAATGTGGGAA
AATTTGAAACTTCCTAGAGACTTGTTGAATGGTTTTGACCAAAATGCCGATGGTGATGTGGACGATGAAGTCCAGGC
TGAGGTGGTCTCAGATGGAGATTAGGAACTTCTTGGGAACTGGAGCAAAGGACACTGTTGCTAAGCTTTAGCAAAGA
GACTGGCAGCATTTTCCCCTGGCCTAGAGATCTGTGGAAATTTGAACTTGAGAGAGATGATCTGAAATTGGAACTTT
GTTTTAAAGGGAAGCAGAGCATCAAAGTTTGGAAAATCTGCAGCCTAACAATGTGATAGAAAAGAAAAACCCATTTT

[0341]

CTGAGAAATACAAGCTGGCTGCAGAACTTTGCTTAAGTAAAGGAGCCAAATGTTAAGCGCCAAGACAATGGGGAAGA
TGTCTCCAGGGCATGTCAGAGGTCTTAATGGCAGCCCCTCCCATCACAAGCCCAGAGGCCTAAAAGGAAAACATGGT
TTCACGGGCCAGGCACAGGGCCGTGCTGCTTTGTGGAGTCTCAGGACTTCGTGCCCTGCATACCAGCTGTGGCTCAA
AGAAGCCAAAGTACAGCTCAGGCTGTTGCTTCAGAGGGTGCAGGCCTTAGTGGCTTACATGTGGTGTTGGGCCTGGG
GTTGGACAGAAGTCAAGAATTGAGGTTTGCAACCTCTGCCTAGAATTCAGAGGATGTATGGAAAAGCCTAGGTGTCC
AGGCAGAAGTTTGCTGCAGGGGCAGGGCCCTCATGGAAAACCTCTCTGCTGGGACAGTGCAGAAGGAAAATGTGGGG
TCGGAGCTCCACACAGAGTCCCCACTGGGGCACTGCCTAGTGGAGCTGTGAGAAGAGGACCACCATCATCCAAACCC
CAGAATAGTCAGAAATGCTGATAGTTTGAACCATGCATCTAGAAAAGCTGCAGATACTCAGTGCTAGCCATGATAGC
AGCTGGGAGGGGGCTGTACCCTGCAAAGCCACAAGGGCAGAGCTGCCCAAGGCCATGGGAGCCCACCTCTTGCATCA
GCATGACCTGGATGTGAGACATGTAGTCAAAGGAGATCATTTGGGCAGTTCAAAGTGTAATGACTGCCCTATTGGAT
TTAAGACTTGCATGTGGCCTGTAACCCCTTTATTTTGGCCAATTTCTCCCATTCGGAACAGGTGTATTTACCCAATG
CCTGTACCCCCCATTGTATCCTGGAAGTAACTAACTTGTTTTGGATTTTCAGGCTCATAGGCGGATGGGACTTGCTT
TGCCTTAGATGAAACTTTGGACTTGGACTTTTGGGTTAATGTTGGAATGAGTTAAGACTTTGGGTGACTGTTGGGAA
GGCATTATTGTGTTTTGAAATGTGATGACATGAGATTTGGGAGGGGCCAGGAGCAGAATGATATGTTTTGGCTGTGT
CCCCACCCAAATCTCACCTTGAATTGTAATCTCCAAAATCCCCAGGTGTCATGGGAGGGACCCAGTGAGAGGTAATT
GAATCATGGAGGCAGTTTCCCCCATGCTATTCTCATGATAGTGGGTGATTTCACATGAGATCTGATGGTTTTATAAG
TGTCTGGCGTTTCCCCTGCTGGCACTCATTCTCCCTCCTGCCACCCTGTGAAGAGGTGCCTTCTGCCATGATTGTAA
GTTTCCTGAGGCCTCCCCAGCCATGCAGAATGGTGAATCAATTAAAACTGTTTTCTTCATAAATTACCCAGTCTCGG
GTATTTCTTCATAGCAGCATGAGAACAGACTAATACAACCAACATCCCCTCTTGTCGATTTTCATTGAATCCCCATT
ACCTGGGAATGAATGAGGAAGGCAGAAATAGAAATCAGCGTATTGTTCAATATAGAAATCCTGAAGTTACACAAGAT
AATTACGAGCAAGACTCAGTACAGAGAAGGAACCCTTGAGCTGGGGATCAGAGTCTTCTGGAAACAAGGAGTACCAA
AGGGGAGTGAAAAAGAGAAGTGGAGAATGAGTAGTAGCAGTAAAAGAGGAAGTATGTAAAGAAATGGAAATAAAGAT
GAAATAAACTGTCATCTTGTTCCAGAGACAAAGCTAGCTAGGGGAGCTAAGGAAATTTAGAGGGATAAAAAGTCTCA
TCCAAACCTTCATAATGAGCTAAGGGCTAGTTTTTCATAGACCCTAGCCCAACTTGCAACCCTCTCACTCTCTTACT
TTTTTCCTGGCTCTGTTTTGCTTTACTGCCTTTCATATAATCTAAAATGATGTTAAATGTATATTTCCTGTCTCTTT
TCTATATCACAATATAAATTCCATGAGGCCATTGCCTGTGCCACATTTACCACTGTACTCCCAGTGCCTAGAATGGT
GCCCAGATCAAGTAGACTCCCAGAGTGTCTGTTGTCAAATGATTCTCAGTCCAGTGATGTTTGCACTACATGGTGCT
GCCACTTTTTTTTTTTTTTTTACTTCATTCCATTTGGAATTTGGATCATTTTGCCAACTTTAGGGTTTATTCCTCTT
TATTTCCCCTGTGTTTGCTGAATATCTGACTATCTATTTTTCTCAGCTCTCTTCTCCTCACACTTTGCCTCCAACTT
CCTCTCTGCTACAAATTATACTTAGCACAAGACTGGAGCCACTTTTCATTGATTTTGATGTCAGTGCTTGAGTGACA
ACTCCATGTGGAGAACTTTCAACCTCCTGCTGTCTCGTTGTTTGCTTTGAAAAAAAAAAATACTCATCTTCTTCCTG
AACCCACTGGGTAAATTTATGCTCACAAAAAAGACAAAACTTCTAGAACAAAAACTGTATGAGCTTCAAAAGTATTA
GTTTTCCCTTGATTTTGAGACTCAAAAGTACCATAGAAATTTTAAAAGGGGGATGAAGTAATTTTGTCATTCAACCC
GATATATTTTACAATTTGAATGTCAGAACCATGAAATGTCTTTTCAGTGTACTATCATCCTCACTATTTCCACCTAT
AAAAGAAAAGCAAAAATATTGTTTGAGAACAAAAAAATATTGCAGTAGCGGAAAAATTCACTTTAAGTCTTTTTTTG
GAGGTTGCCACATCACGCCTGTCTTCAGCTCCTCTCACTGAAGATTCTTTGTTTTCTTGAAACCAGTTGGTCACTTC
CAATTCTTACCACCTATAAATCTTCCTGGTTCTTTATATTTCACTTTATATGTCCCTTATCTCCTCTGATTTTTGTT
CCAGTCTTTCATTTTTTACATTCTTTGTCATCCTATATGTCAGAAAGAATCTGCCACTCCAAATTCCTTAATATATT
TATCTACCAGTGTTTATTTGCCTGTCACATGTACTCTTCAGTAATCCATTGACTTGAAAACTTTCTACTCTGTCTTT
CTTACTATAGTTTGCTTCTTGAGTTCAAATTATTAATATATTAGGTCAAGTAATTCTATTTTTATGCACATTTTACC
TAACATTCTTGGTAAATATAATTGTTTTTTAAATTGCATCACATGGTGAGTATTAATTTAAGACGTAATTAGTATTA
ATCTACTGGACAAGATTTTATTTATTAACTGTTTTGCCCTTTGCTAGGGTTCTGTAATCTCACCCTCAGTGCCTTTA
AATAGTCTTACCCAAGGATCACCTATCCCTCTTCAATTTAAATTCTCTATTGTGTTAGTTTCTGAGATTGCACATTC
TCTGGTTTGGTTTGGTTTTTCAGTGATCTCTGTGGTCACTTCCAACCATTTCTACTTAGTTTACTTTTTTCCCTGTT
AGCCCTCTTCAGTGTTAATATCACCTGGAAGTTAACCCATTCCTAGTCTGCATTGCTCTATGGCCAGCATCTTGGCT
GATTGGCCATCACCTATCCTGTACTTATTGTCAATTTTTTTAAATTTTTTATTGTGGCAAACAGCACATAACGGAAA
ATTTACCATCTGTCTTTGTTCATGTATGTTGCTATAAAGGAATGCCAGAGACTGGGCAATTTATAAAGAATGGAAGG
TTTATTTGGCTCATGGTTCTGCAGGCTGTACAAAAAAGCATGGCACCACTACCAACTTCTCATGAGGGCCTCAGGCT
TCTTCCATTCATGGTGAAGGGCAAAGGGGAGCTGGTGTGTAGAGATCACATGGTTAGAGAGCATAAACAAGAGAGAG
AGGGGAAGTGCCAGGCTGTTATTTGGCAACTAGCTCTTGCAGAACCTAACAGAATGAGCAGTCACTCAACCTGCCCC
CAGGAAGGGCATTAAGCCATTCAAGAAGGATCCACCCCCATGACCCCACACTTCCCGTTAAGCCCCACCTCCAATAT
TGAGGATCAAATTTCAACATGAGATTTGGAGGGGACAACATCTGAACTATAGCATTATCTTAATTCAAATAATTTTA
CTACCAATTCCATTTATTAGAGAATCTAAGATTAAGACTACTGAAACTATGCCAATGATTATCTCTAGTTCAGGCTT
CTCTCTTCAATTCCAAACTCATATTTTCACCTGCTGACAAGCTACCACGTAGATGTTCCACAGGTATTTCTAACTCA
GGATATGTAAAATTGATGTTATCATTATTTTCTGAAAATCTATTCCTCTTACTATATTCCCTATTTTTGTGAATGAC
ACAATCTAATGGACTGGCCATGTTAAAAATCAGGAAGTTTTTCTGGCACTTTCTTTATTCCTTAATATTAAGCCCCA
CGATAAATTACCAATCCTTTCGATCTTGTTAAGACAAATAAGCTTGAAATTTCTTATCTCTTATTTAGCACTGCAAT
TTTGTTTTCTCATTCATTCAGATAGCATCTTCTACAAAACCTCTTCACTAGTTTTTCTGTCTCTCAGCTCATCTTGT
ACACTGCAAATCCATTAAAACTACACTTTTTAACATGCCAATCATAATGTATTAGTAAAGGTTGTATAATAGTTCTC
TTTCTCTTCTAGGATAAAAGCTAAATTTTTCAGCTTGGTACACAAGGCCGTTCATGGATTTGGTACCTACTAGTCTA
TCTAGGGTATTTCCTACATCTTCCCTCCTCTCTCTTGAAAAATCAAGTAATACTAAACTATTTGTAACTCCCTAATC
ACCAATTGCTGCCTTTATACTTTAACATATATTCTTCCTTTTGCTAGAAATACATTTGTCATAATAGGCTCTTTGGT
AAACTTACCATTACATTCTCAGTTTATAATATTATTTGTGTAAGTACCTTTTTTCCTACTTAGTCCTCCAGGGGAAT
GCATGTGCTTCTCTCTCCATATTATCATTTTAAAATTTACTCCTATTTCAGCAAGTATCATAATATTTTACTGAGAT

[0342]

TAATTGTTGGCATATTTATGTTACCCTGGTGACATATTTAATAGTTAAGATGCTTTGGGCTGCAAGAAAGAGAATGC
ACTTTCAAAGTGGCTTCAAAAATAAGAATAATTCAACTCACATAACTGGAAGTTCAGAGGTTAAAGCTGGCTCCAGA
TGAAGTACAAGCAGAGCTCTGTCACTCTTTCTCTATGACTTTCTTTGTCCTGATTTTTTTCTCTGCGTCAGCTTTGA
CCTCCTCACTAGTTGCCCTCCAGGTTCCAAGATGACTGCCAGCAACAAATGGGTAACATAATTTCTTGTTACAGTGA
GAGACATAAACACTTACCTCACAGATTCAGAGATTTCACCCTCAGATAAGAACAATCAAAATTTTAATTGTGTATCT
ACTGTATCATTAGTATATGTATTAGGTATTCATATATATCTACTATTGTACTAAAGATGATTCAGAATGGTTTCATA
AGCTGGTTGATTTAAACTTTTCTTCCTGAACCAAGAGGATGAGATTATCTTGACTGGCTTAGAATCGAAGATTTAGG
TGAATCTAAATCCATCTCCAGAACTGAGGATGGATTCCGTAGAAACTTGAAGACAATTGGAATTCTCCTAGAAAGCA
GAATTTCAGAGTGGGTGTAGGGATATCTTTGAATGGAAGTTGTTTAAGCCACCAACAATATTCACTCCATTGGGGTG
ACACTGGGCCCAAAGGTACCCAGTTAATATTTGTTCAACGAACCTATAGGAAGACAATATTATCTGCAAGCTATTTA
TAAACTAAATGACAACGAGAAGTATTTTTAGAAAAGTTAATCATATACAATTTAGTCATTCACATTGGGAGGCTCTT
CCTAAATTTCGTTTCTGATCCTTATCTTGATCTTTGACGCAATCATATCTTGGCCTCCATCTAAAAGCCAGTAGAAG
CCTCTGGCCATTTTTTAAACAATTGGTTTAACAACATTAATTGGAATCTATATAGCTCAACTTAAAAAAAAAAAAAA
GAGGAACTAGCTTATCATAATACAATATTGTCAAGGCAAGTGATACTTCCATGCAGCTATTCAGCATTTGGTGTATT
CTTGGTTTCTGCAGACCTAATAAAAAGAGATTATCTCAGCGTCCTCTAGGTTTAATTTTAAAAAACAAGAAACCATG
AACAATGAAAATGATTTCAAAACATCTTCACTCACATCTAAAATATGAAATAGCTAGATCAGAGGTTCTTAATAATT
TAATCGGTGAAGCATAACTGAAGGATTTAGAAAAAGTGCCTAAACTATTAAAAGATTCTTCAGCTGCTGCTTAATGT
TGATATCAAGTTATTTTTTACACTGGCAAACAGCTAAGCCATTCCCTGGTCACATCCACTTTTCAGCATTTAGCGCT
CCTCTCAATCATCCCGATCACAGCCCCAGGAATGTTTCATGGCATCTCCGCAATAATAGTATATTACTATTGGGTCC
TGAATTTTGAGGTAGTTTATCTTTCAGAAAGGATATGAATAGAGCAATAAGTCCAAGTAGAGTGGGCTTTATTAATG
ATGCCCAAGATTTTAGACGTATACTATAATCTTCAATCAATTTTAACACCACTTCCCTTACCCTAGTTTTAGTCTAA
ACCAGTCCAACTGTGCTGTAGTCTCAGATAATACAATGTTAGATTTTTTTTTTCAAGATATAGAATTGTCGGGAAAA
CTTGTATGTAACCTATACCAATCTAAATTTCTTAGCATTTAACTCTAAAGAGTAAAGCTTTTAGCACTTCCTGTGTA
ATACAAGGGCAAGTGCTGATTTCTCTGAATATAATTTTCCTTTGTTAGCACATATGTTCACTGTTTAATAAAAATAA
GAATGCTTCAATTATCTCCTTTTGTCAAGGGCCTGAGAAATAAAGAAATATACAGAGGGTCCCAATCTAAAAATGTT
TTGATTTACAATGGTTCGGCTTAGGATATTTTTATTTTACAATGGTGTGAAAGTAATATGCACTCAGTAGAAATCAT
ACTTCAAGTACTCATACAACTGTTCTGCTTTTCATGTTCAATACAGTATTCAATAAATTACATGAGATGTTCAACAC
CTTTAATAAAATAGGCTTTGTGTTAGATGATTTTGCCCAACTGTAGGCTAATGTAAGTGTTCTGACCATGTTGAAGG
TAGGCTAGGCTAAGCAATGATGTTCAGTAGGTTAGGTTTATTAAATGCATTTTTCAATTTATGATGTTTTCAACTTA
TGATGGGTTTTTCAATTTATGATGTTTTCAACGTATGATGGGTTTATCAGGAAATAATCCCATTGTACGTTGAGGAG
CACCTGTATCAATATAGGCATTTACACAACTCTCGTACTGATAGCGGCAGGAGGCAGAGAAGCTCTAGGCAGAAAAG
GGATGGTCCCCAGCGAAAACCCCACCCTCAAGCCAAAAAGCCTGAAACCGCAGCTCAAAGTGGGAACTTATATCCCA
GTTTTCCTGCTCGAATGTTGCCTTTTTCTAAACCACCCATGGCCCCACCCCACCCCATCCTGTGCCTATAAAAACCC
CAGACTCAGCTGGTAGACAGGACTACAGCTGGACATCAGAGAGAAGCAGCTTGACTTCAGAGGGACAACTTGATGGC
ATAACTTTAGAGAAGAATCCGGCTGGACTTCAGGGGAAGATTACTTGCCACCCCCATCCCCTTTTCAGCTCCCCTTC
CCACTGAGAGCCACTTTCATCGGCAGTACAATCCCTCACATTTACAATCCTTCAATTTGTTCATGTGACCTCATTTT
CCCTAGATGCTGGACAAGAGCTCAGGAGCCACAAGTGTGAATACAAAAGGCCCTTTGCCCTTGCTGGTGGAGGGCAG
CTGCCTCCTGTGAAAAGACAAATGGCCCACTGAGCTGTTAACGCCTAAGCTGTCCGTGGATGGCAGAGCTAACAGAG
CACTGTAACACACCCTCTGGGGCTTCAGGGGTCGCAGACGCCTCCACCTAGATGCTGCTCCAGTGCTCATGCACTCC
AGTTCCCACCTCGTTTGCTTGCACACTCCCTCCAGTGAGGAGTTGAGAGCAGTGGGCTAAGTAAATAAGGCACCCCT
GTTGCGAGTTCCACAAAGGGGTCAGGGAAATATCCTGCTTCGTTACTAAGTAGGTATAAACCCTTTACTAAGTTAGT
AATTATTCAGTAATAATACTTAAATGAAAGCTATTCTGTTAGACTAAATTCAGATTATCAAAGACACGACAGAAAAC
AGCTTTCTCATTCATGAACCTATTATTTCCTTTTGTAAAATATGTTAAATGAAAATACAGCCGTCATGCTAATTTCT
AATGGTAGAACATATTTTGAAAACTCCTTTATGTTTGGAAGATTTTGCTTTAGTGCAGATAATCAGAATGATGTGAT
ATACTAAGAAATAATTTTTAAAATGAGATGTGACATTTCTCATAATCTAAATAAGAAATGGCAAAACATTGTCCTAA
GCTAAATAACTCATATGAAGTGATAAAATATTGCTTTCTAAAGGTCCATGATATGTAGTGATTTCTAATGTGTTTAA
TAGCATTGATCCTATTGGGTAATTGGGTGGTTCTAACATGTAAGGAAGGCCTCCAGTACTAATTTCATGTACTGGGA
AACTACTGGGCAGGGATGAATCCCTTAACTCCTAAGTAAGACTATCAGATCATATAAATCTGCTTTTTGATTTGCAA
AGACTCTTAGGCATACCTCTTGAGAATATTAAACATCTACTAAATTATATGTAAAGCATTTCAGTCTAAGATTTACA
ATGCTCAAAGGAGAAAGATTTTAAAGTTCAGCATTGGAATTTCCATAATTTCCTTCCAATTGTAGAATTTTACAATT
GAGATAGCAAATAATAATATAAATAAATATGTAAGACTGAAATCAACATAGGAGTTTGGAAAGAAGACTGTATGGGT
TAACTAGAGTTGTCAACAAAGACTTCACAGAAGATTGTGATCGTAGACATTAGTAAGGAATAAAATGGGTGCAAAAA
AAAGTCTAAGAAACAACAAGGATAAATTTATTTTGTAAAGAGTGAAAGCATATTGGGTATTACATAAAGAGACTGGC
TAGTTTGACACAATGTGTACTTTTGGGGTAAGCAGTGGGAAATCACTTTCTGTAAGTAAAGTGGAAGAGAATTCAGG
TATTGGATATGGTCATATAGGATTTTTCAACCTTTGTTTTGCCCCCATCCCTATCCTACCAGTATCATCACAGATGT
TGGCAATATTCACTCACTTGCCTGAAGATAGATATTCATGTCAGGGGCTACAGATATGTAAAGGTGAATGAAGTTAA
ATAGCAAAATTCAAGCACACTAGTTCTGACTGTGCACCCTCTCTCACTCAGAAAAGCACTTTGGAATGCGAGAGTAC
TCTTATATTGACAACCTTTAGTTGGTTCTAATTTATAAAAAAATATAAAAATATTTAGACAATTTGGATATTTTATA
TCTATGCCACAACCCATTGTCAAATGATTGAGATATCATTGTTTGGCAACAATGAAACTGAAATCAACTCTAAGAGT
AAAGAGTACGACTTTTAAAGATATTTGGTTAAGAGTAATGTTGCAAGGCAGAGACCAACATTATAGAAGTTCTCAAC
CTGTGTGCTGGATATAGCTAAAATTTCCCAGGACCATTGAAAGAAATTCTCAAATCCACAACCACCTCAAGTGTTGT
CTGTGAATCAAAGGAGTGATGCATTACCCACAAACATGTAGTAGACATTTTTTAAGTGATGTAGATGCTACTGCAAT
TAATAAATGCAAAATTATATTACAATATTACTGAATTCAGAATAAGTTGATTTTGTCACTATTTTTCTCAATCTACG
CATGCTAGATATAAAGTTATTATATAGAGACGTAACATTTTTAATATTTCAGAAGCATGCTCATGATATTGTAAACC

[0343]

AGGCACTGTTAAAAGATCACTGGATTGGGAATTAATAAACCTGGGATCTAGTCCTCATTTGATCTCTTGTTGGCTAT
GGTTTGGGGGTTTTGGGCAAACTATTCAACCTTTTCTGCCTTGGTTTAATTGATAAATGATGGTGTTTTCCCAGATA
AATCCCTTCCCAACTTAATATTACCAGATGTAGCATCTATGGTTTAGAATGTACAGTATAAATTAACCTTCCTGAAG
ATCTTCACAAGTTACTATAACCTATATTTTCATGGCACTGAAAACTAAGTTTTTGATAGCTTACGTTTTTATAAATA
ATTTACTTACTCATTTTTATCATAATAATAAATCTGATTCACTAATCACCAAAATATCATTTTTGAAAATAGATGCA
TGAAAGGATCCGAACTTGTTATGGTTTTATCTGTTCAAGTCACCTAATTTTGGTAGCCACAGGCCCCCATTGTCAAT
AGGGGAAGATTATCCATTTTAGCAATACAGTCTCTGATAACTTCAGCCCATATGCCCTTCAATTCCACTTTTACGTC
AGATTGAATGGGAATGTGAGGCCCCCAGTAGTAAACTAACCCTACTTTCTCTTTGGAAATTGGTCTTTCTACCTAGC
TCTTTGCCTCTGTTCACGTTCTTGTATTCCACAGAAAATATATACATTAGGTGTTAAAATCACAATGATTAACAATT
TTTAAGTAGAAATAGTTATTAAGTATAGCATAATCATGCCTTTGAATTAGTACAAAAGTAGGAAACAGAGCTTTAGT
GACTTTTTTCATTCTTTCCACCATTTACAGGGCAAAAATGAAGAATTTTACCAATTCAAAACTATGCACATGTATAG
TTTCCACCAGTATTTAGTAGTTATGTTTCTCAAGATGTATAATTCCTTTCCTTTCTGTTTTCTGTAGTTTGATAAAC
CCTAGATAGGAGTTAATGTTGTTTCAGTTAGGTTTATTATTTCTTTTATGTGGTTTAATTTCATGCAATAAGCTAGA
GATTTTGTAACATAACTTGATAAAAATTTTCTCCCTTCGTATCTTTTTGTTTTAAAATAATGGATTAATATAGATTG
TAATTTTACAGTAGAGGAAATACATTTTACTTTTAGTTCTTATCTAGATATCTTAGGAAAAAAGAAACATCATTTTT
AAGGATTATTATTTTTCTACTAATGAAAAAAATAGCATGATTTTCCATTCTGGACTTTGTAATTAACTTTACCTTGG
AATAATTGAATCTTAAATATAACTTCATTGAAAATTTTATTTTCAAGTAACATTTTAAAATGCAAAATTTGTGTGAT
CTCTACTAAAAAGAACTCTCACATCCCAATGTGCTTATAGCTACAATAATTGGTTAAAAGAGACAAATTATAAAGAA
GATATAGGTTGTGACATTGAAATTACAAAATGTGGATTAGGGAAGTAAGCGTGTAGAGTTTTGTATGCAATCAAAGT
TAAGTTATTACCAGCTTAAAAAACCCTTACGTAAGATGTAAAGAAAAGCAAAAACCTATAGTAGCAAAAGCAAAAAC
CTATAGTAGATATACAAAAGGTAAAAAGTAAGGAATCAAAGTATACTACTGAGCAAAACAGTCAAACCATAAAAGAA
GACAACAAAAAAAGGCATATAGAACAAAGGATCTACAAAACAATTAGAAAACAACTTTTTAAATGGCAGTAGTAAAT
TCCTACCTATCAATATTTACTTTGAATGTAAATGGATTAAATTCACCAGTCTAAAGATAGAGTGGCCAAATGGATTA
AAACAACAAGACTCAACTACATGCTGCCCATAAGAGACTCACGTCATCTTTTAGGACACATGTTGACTGAAATAGAA
GAGATAGAAAAAGATATTTTGGTTTCCATGCAAATGGAAACCAAAAGAGAATGGGGATAGCCATACTTCTATTAGAC
AAAATAGGCTTTAAAAATCAAAAACTAAAAAGACACAAAGAAGGTCATTAAATAATGATAAAAGGATCAATTCATCA
AGAAGGTATAACAATTGTAAATATATATGCACCCAACATTGGAGTACCTAAATATATAAAGCAAATATAAAGTGATA
AAAAGAAAGAGACAAACTACAGTACAATAATAGTAGGGGACTTCTACCCCAAATTCAACAATAGACAGAAAATCCAC
ATTAAAAAATCAATAAGAATGCATTGGACTTTATGCTTTAGATCAAATAGACCTAGCAGACATATACAGTACATCTC
ATCTAACAGCAGGAGAATATACATTCTTATCAAGTGCACAAGAAACAATTCTTCAGGGTAGATCATATGTTAGGCCA
CAAAATGAGTCCTAACAAATTTAACAAGATTGAAAGCATGTATTTTATGTAAAACAACGCCATCATGGAAAAAGTAC
TAGTGTAAACAAGGTTTAAATATGATTTACTGATTGTTTAAAAAGGAATTATCTTAGCCCTGATCTGATGGGATTTC
CCCTTTGTAAGCAGCAAAAATAAGTTCATAATGAAGCAACTGTAATAATACAGCTTCACAGATCTTTCTGAATAAAC
AGAGTTGGATATGTTTCTACTTCAGAAACCATTTACTGTGGGCTCACAGCTTTTCCATACACTCTTTACACTCTTAA
TTTTAAACCCATTCATCAAAAGGATTAAGACAATGAGATTCAAGTCCAAGACAATAGGAAGTATGTGCATCAAAACT
GTCATGCTAATGCTCTGAGGAACATTGTTATTTCAATAGCATAATTTAAAACCACTGAAACCATGTTTTATTTATGA
TTACCTTTCATACGTTCAAAAAGAATTTGAGATGGTTTGGCAGGGATGTCTTAAAGAAACAAACACCAAATTTTTAT
TTGTTGTGCTTCTAACAAGCAATTTTTCCTACGTAAGTGTTACCTGTTTTCTCCCCTTGATTTTGATCTCTTTTGTA
TGGTTGGTAGTTGTCGCACTTCTGGGCATAGTATGGTTTTCTACAAGTTAAATAAAAACATAAATTGATTTTGATGT
AGCTAGTTCCACTATTTCAGTTAGTGTGTTTTCCATACCTGCAATTCATAACATGTTTACTGACCCAGAACGTTAAA
ACTCAAGTTTAGTTCAATACAGTCACTTCAAACCAACTTAAGATAGCAGTTACATTTCCAAGGTCATTATGAAGACG
GAGTTCAAGACTTCAGCAGTTGAATTTGTCAAGCTCATGGGTCTTTTAGTTACAGGAAGTGTGCATATTTCACATAA
GAAACAGCAACTCTGATGATACATTGAAACTCAAATATACCCAAGGAGTGTAAAACTACTTTATAAGCCCTTAAACA
ATAAATATGCCAACAATCCTCTGCATACTTTTTGTCATTTTTTAGAGCATTCAATTGAATTATATAACATGTGATAC
CAATAAATAATTAACTTTTTATTTATTTATTTAGAGAATGATTCTTGCTTTGTCTCCCAGGCTGGAGTGCAATGGCA
TGATCTCAGCTCACTGCAACCTCCACCTCTTAGTTTCAAGTGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATT
ACAGCCATGCACCACCACGCCTGGCTAATTTTGTACTTTTAGTAGAGACAGGGTTTCACCATAATGATCAGGCTGGC
CTTGAACTCCGGACCTCAGGTGTTCTACCACCTCAGCCTCCCAAAGTGCTGGTATTACAGGCATGAGCCACAGCGCC
TGGCTTATAATTAACTTTAAAAAATATTCTACTATCGAATGCCTGAAAAAATCATATTACTTCTATGTATTAAAAAC
AAACTATTACGTAAATAGACCAGAGATACGGTGTCAGAGATGAGATTCCTTGGCAACAGTCTCTGAGACAAAAAGTT
GCACACAGAAAGTATTTTGAGAAGTATTGTTGCTATTTATAAGGAAGTGAAGGAGGCAGTATTGGGCAGAGAGAAAA
GCTCATCCACACTGCAGTTGCTACTGAGGCTTCAGCCCATGTGACAGGGACACTGGAGCTGGGATGGTCTTTCAGAG
TTCTACCAAATTGAGGCAAATGGGCGAGACTTTTGTATCTTTGCACTAGCCAAGAGCAGGCACCAGGGAGAAATGCA
GCTGTGATCCTTTTGTAGATATTATTATTCTGGAGTCAATGCAACCACACCACAAATACTAGGAATAATATTGGTAG
TGTGGGTGCATCGACTCCAGAAGAGGATCTTGATGAAGCATTACAGTATCCACTACAGAGAGGCACTGGCATATTCT
GCTTACCCCACAAATATTGCAATGAGGATGCAAGTAGGAGTAGTGAGGACATAAAATAAATCGATTATTTTCCACTG
GGCCTTAATATATCAGAACCATTGGAATTTACAGGATGATATTTATTATAGTACTCAAAAAAATCTCTTTTAAATCT
CCTTAACTCAGAAGGGAATTTTAAAAAGTCCACAATTCACCTGGTCTATGAATTCCCTTTTAAAATAAAATGTGGTT
TAATCAAATTTACAAGAAAAAAACAAACAACCCCATCAAAAAGTGGGGGAAGGATATAAACAGACTCTTCTCAAAAG
AAGACATTTATGTGGCCAAAAGACACATGAAAAAAAGCTCATCATCACTGGTCATTAGAGAAATGCAAATCAAAACC
GCAATGAGATACCATCTCACATCAGTTAGAAAGGCGATCATTAAAAAGTCAGGAAACAACAGATGCTGGAGAGGATG
TGGAGAAATAGGAACGCTTTTACACTGTTGGTGGGAGTGTGAATTACTTCAACCATTGTGGAAGACCGTGTGGCGAT
TCCTCAAGGATCTACAACCAGAAATACCATTTAACTCAGCCATCCCATTACTGGGTATATACCCAAAGGATTATAAA
TCATTCTACTGTAAAGACACATGCACACGTACGTTTGTTGCAGCACTGTTTACAATAGCAAAGACTTGGAACCAACC

[0344]

CAAATGTCCATCAATAATAGACTGGATACAGAAAATGTGGCACATATACACCATAGAATACTATGCAGCCATAAAAA
AGGATGCGTTCGTGTCCTTTGTAAGGACATGGATGAAGCTGGAAACCATCATTCTCAGCAAACTAACACAGGAACAG
AAAACCAAACACCGCATGTTCTCACTCATAAGTGGGAGTTGAACGATGAGAACACATGGAGATGGGGGAGGGGAACA
TCACACACCAGTTGGGGGATTGGGGGGCAGAGGGAGGGATAACGTTAGGAGAAATACCTAATGTAGATGAAGGGTTG
ATGGGTACAGCAAACCACCATGGCACGTGTATACCTACGTAGCAAACCTGTACATTCTGCACATGTATCCCAGAACT
TAAAGTATAATTTAAAAAATGTGGTTGAAAAACAAAACCCACATAATACAAACTTTGCCAGCTTAACTGTTTTATGT
GTACAAACCAGTAGTGTTAACTATACATACATTGTTATTCAACAGATCTCTAGAATGTTTTCATCTCTCAAATCCGA
AACTCCAAACCCACTGAAGAGCTCCCATTGCTCCCTGTACCCCAGCACTGGCAATATGACCACTCTACTTTCTGTCT
CTAAAGAGTTTAACTACTTTAGATACATCATATAAAAGGAAGCATGCAGTATTTGTCATTTTATGACTGGCTTATTT
CACTTAGCATAATGTCTTCAAGGTACACCCATGTTGTAGCATAGGAAAGGATTTCCTTCTTTTTTTGTGGCCGAGTA
ATAATATTCCGTTGTGTCCTATACCACATTTTTTAATCCGTTTATCAATCAATGGACATTTTAGTTACTTCAATTTT
TGGCTATTGTGAATTATGCCGCAGTTAATATGAGTGTGCAAATATCTCTTTAAGATCCTGTTTTTAATTCTTTTGGA
TATACAGATGCTCATTGATTTACACTGGGATTACATCCCAATCAGCCCATCATAAGTTGAAAATATCATGTCAAAAA
TTAATTTAATATGCCTAAGCTACCAAACATCATTGCTTAGTCTAGCATACTTTAAATGTTATCAGAACATGTAGATT
ACAGTACATCTGGGCAAAATCATCTTGCAACACAGTACATAGTAGAGTATCAATTATTTATCCTCATGATAAATATC
ATGCTGCAACCCAATATGTCAGAAGAGAGTATCATACTGCATATTTACTAGCCCTAGAAAAGATCAAAATTCAAAAT
TTGAAGTATGGTTTCTTAATGAATGCATATCGCTTTCACACCATTGTAAAGTCCAAAAATTACAAGTTAGACCTTAG
TAAGTTGGGGACTAACCGTACACCCAGAAATGGGATTGCTGGATTATATGATAATTCTATTTTTATTTTTTGAGGAC
TGGATTATATGATAATTCTATTTTTATGTTTTGAGGAACTTCCATACGGTTTTCCATAGTGGCTACACCTTTTCACA
TTCCCACCAACAATGCAGAAGTGTTTCAGTTTCTCCACATCCTTGCCAGCACATGTTATTTTCTGTTTATGATCGTG
GTCATCCTCATGGTTGTGAGGTAATATCTCATTGTGGTTTTCGTTCTTCATTTGCATTTCCCCGATGATTGACGATG
TTGAGTATGTTTTCAGATGCTTGTTGGCTGTGTATATATTTTCTTTGGAGAAATGTCTATTTAAATCCTTGCCCATT
TTAAAATCAGGTTAATTGTTTTTGGTGAATTCCTTATTAATTACTCACCTCAGAGCCTTTCTTCTTAAAATACAGAT
TTCTCAGAACCTTTCATCTAATTTACTAAATGTTGATCTTGAAGAAGTATTTTAATTTATCTGAATTTAGTTTCCCG
TTCTATAAATTGATAGTAATCATGTTTTCCCTATCCACACCAGAGTGGTATGATGAGGAACCAATGCAGAAAATGAC
TAAAAAGGCATTTTTATGCTGTCCAGATCTTGTGCAAATATATTATAATTGATTTGAACCAAAAGAATCTACATTTT
AAGACTATTTAATATTGTCGCATTTACATTGCACTAATGGTTCCTCTTTTTTCTCACTGATGAAGTTCTAGATTGAA
ACTCTGAGGAAACTGTAAATCACAGTTACATAACTCTGTTATATTAATTTATATAACTTCTCTGTCTCTCTCTAAAT
ATATATATTTAGACTATATATACATATGCATATATATATTTAGAGAGAGAACGAGATTACACTTAAGGACTCTGCTT
AGTACATTGCAACTTGCACCTCATTCTAATTGTGAAAACAACAACAAATTTTGATTGAGACCCTGTTATTTTTCAGA
CCTTTGGTCTTGTGAGGAAGGAAAAGATGATTCTGATACTTATCTCAAATAGCTTATCATCTCCTCTGGTGGATGAG
GCTTGTGAGGGCCTGACACATGGTAGAGGATCTTATTCATTATATAACAACTTTCAAAACACATCACTAATTATTTT
TCTGCTTAAGGTGAAAACGTAGCTCTCCCAAACAAATAAGAGATTTAGACTGAATTCTGTAGAAAACACTGCTTACC
TCATTTTGCTTAGTTTTCACTGCAGAGTCTTGTCTCTAACATTCAGGTAGGAGGAACTCTATTTTAAAAAATTGAAT
AAACAAACCCAAATCAGAATAAATTATTTAATTCAGGAGTTTTTTAAGGTGCTCAGAATTTTGGTAGAATAGCAGTG
TATTTTCTGTAATAAGAACTACACAATGTTCTCTAGCTGGTTATACTCAAATGGTTATATGGATTTTATTCTGATAT
CAGTTTTGCAATTGATGACCACCCTTCATCAATAGCGTGGTCTTTATTATGGTTTCATCTGGCTTTTTCTTCCCTTG
TCTTCTGCACTGGGCTGCCTACTTCTTGGCTCTTATTATTACATACTCTACTTTTTTCATATACAAAGGTCTTAGTG
CATTAATAGTGAGTTTCTGCTAAAAGTGTTACTTTCTTCAAATATCCTGTCAAATGCTGGCTGCCTGATTTATTGAC
CTCATAGAGTGACTATGTGAATCTTCCATTCTCTGTGGAATTCCATTCCACATTTTAACTTAAAGACTCCGTCTTTT
CCAGCTGTGTTCTCTCTGCTGGATAAGCATTTTTCTTTTAATTTCTGTGGTAATGGGAAGGGAATTTTAAATCCTGT
TTTGCGTACAAGTGTTTAACAGGTAAAAATGGATGCTGTGTCAGCTATCTATTGCTGTGTAACAAGCTACCCCAAAT
GTAGAGACTTAAAATAAGAATCATTTTTTTTTTTTAGCTCATGATTCTTGATTCTGCTTGGGGCTTGGTTCATCGAG
GCGATTCTTTTGCTAGACTCAGCTGTGCTCATTCATGCACTCGTGGTCAGCTGGTAGTTGATGACCACATGTGGCAT
TCATATGTCTGGTATAACTTTTTGGTTATAAGTTGAGACGGTGTGGGTGACAGGATCACTTGCCTGTAATTATTCAT
CAGGCTAGCATAGTTCCATAGCGGTGGAAGAGTTCCAAGCATAGCCCAACACCCAAGCATTGTTAAGCCTCTGCTTA
TGTCATGTTTGCTGATACCTCATTAGCCGAAACAAGTGACTTGACCAATCCATATTCAAGGAGTGGAGAAACAGACT
ATGGCTCTTAATGGGAGAATCTGGAATATCTCTTGACAAAAGTGTAGATGCAAGAGGAGAATCTGTGGACATTTACT
CTCCAGCAGAGATGTTGTCATTTTTGTCACTAAGACAGTTCCTGTGTAATATTATTTATTATTCCTTGGCAGCCATA
AGTTTGTGGTTTACTTATGGATACTTCAGTGCCTGCCCAATATCATGTTGGAAAGAGAAGCCCTTAACAATTTCTCT
TAATTTCTACTTCTGAGGAATGAGAAAAAAAAAAATTGGATAGGCATAAATGATTTCCAGCACAAATTCTAAGCAGCT
ATTTGAGAGGTGGGTGGGTAGGGAGAGATGGAAAATCCTTTTAATTGAGACATGCTCAACCATAAGCAATTTTTCTT
TCCTCAGGAGACATTTGGTAATGTCAGGAGACATTTTTGGTTGTACATCTAGGAGGGAGCTCTGGCCAGAGATACTG
ACAAGCAATTCTGATCCTACAATGCACAGGACAGCCACTCACGACAAAAAATTATCTGGGACAAAAATGTCAATAGT
GCTGAGATTGAGAAACCCTGCTTTTCACTGAGCGCATGTAGATGAATTCATATAATGTTAGTTATAGAAGGGCACTT
AGATGATGCCAACTAAAACAGGAAAAGGCAATGATTATTTCTCTATTCAAGTTAGTGAAGGAAGGATATGTTTAAAG
TCAGGTTGGAGATATTTCTCAAGGGTAATATTTAAGTTGTGAGTGGATGCCAGATGCGCTTGATTGTTACAATCCAT
ATTTTTATAGCTATATCAAATGATCTTCTCTCCCAAATTAAAATGGATATTAGATATATTTCATGGATACCATATAT
TATAAGCTATAGCTCTTAGAAGTATTTGAGTGATTAATAATTCTTGCATATTGAATGAGACATGTAAGAAAGATACC
TGATGTCCTATGATTCTTAGAAATATGTGCAGTGGTTAAGCACATTGTCTTGGAAACATCACTTGTTAATGTGAGAT
CCTGAACAAGTTTCGTTACCTTGGTAAAAGGGGGATGATAACATCTACCCTGAAGATGGTTTGGTAAAGGTGAAATA
AAAGAGGGTTTATGAAGAACCTTGCCTAGCACATAAGAATCACTGAGTAAATGATGGGGGTGATGGCAGTGTTGGTG
GCAATTATCTTACTTGTAAGTAGTAGCAAATATTTATTTGGATTATATTTTTAATAGCAAGAAAGAATCTATTAAAA
TGTAAGCAAATAGCACCATTAAGCTTTAAAAATCCATAGCTCCCAAACTTAACATTTTTTTTTCTTAAATCCAGAAA

[0345]

TACAAATGGTTATCATGAATCCTTGAGTTCCAGTACTCCGGCCTTGCCTATGTAGAGTACAAGCATTTATTTTTCTA
AAATGAGAATTATTGATGTTAGGTGTGTCATATATTTTATTTACTCTTTAGAGTGACAGAAAAGAAAGCAGACAAAA
ATAAGTGTTTATGTGTTCATAGTGTTTGTTTCTAATTTTCCAACTTGTATGTGCAGACTGTAAATATTTAGAGGGAA
GACATGAAATTCTTCCTTACCTACACAATAGATTGTCTTTCTTAGAATCTAGACATGATTTCAAATTTTCTTCAGAA
GTTCTCTTGAATTCTTACGGAATTCATTAATCCAAATGCCATCCTTTCAAATATTCAAAAAGAAAATATTCTCCAAT
GTTGCTTACATGGAGGAGAGTTATTTTATCTACTAGTGAGATGGAAGTCGCTAACAATCTTTCTTATACCCCATAGT
ATTAAATAGTAATGCAATTTGATCACATATAACATTGCTCTCCTGTTGACATGGAATGATAAGGAGGAAGGGAATAT
CTAACAGCACTTTTTAGATATTCTCTCATAGAATTGTTAATGTGAGCCCAATAACTTATTTTGTAAGTTGAGTGATT
TTTCAAATATAAAAGTGAAGCAAACAAAAAAAGTTACCCACATTCATTCAACAAAGCTAGAAATAGAGTATGTGCTT
CTTGCTTTCTTATCAACCAAAGGAGACCTGTGAATTACAGGAAAACCATGTGTGAATACTTCAAAGGAAATATGCCA
AACACATTGTGTGTGTGTTTTAAAAGATTTAGTAGCTCCTTTAAAGCTCTAAGTAAAGCTGTTAGTCTGACTAATCA
TGTACCCCTGAAACCAATTAAATTTTCAGACTAGAGAAGTGTTCTTCAAACACTTGAAAAAAATAAGAGTTCCTTCC
TAAAGCCAACACCTTTTAATTAACAAAATACTCAGAGCAAAGCTGTGGTTTGCAATAGTATATTACACCTACTGTTA
TTCATGCTACCATTCTGGTGTGCCAAACTCAATTCATTTCAACTATTTTGAGCTTGATTTTCTGTCTGCTTAACTAT
GGAGGGAAAACCACAATTTTAGTAGAGTCATTTTAATTTTCTAATGAATGCATGAGAAAGGTATTATGCTACATAGT
ATGTCCTCAGACTATTCAGATGCCCTCTCTCTCTCTTGCTCTGTCTCTGTCTGTCTCTCTTTACACACACACACACA
CACACACACACACACACAATGTGTGTATCTATCTGTCTACCTAGATATACAGTTTAATTTTGTACATTGTAATACAT
CTGGTTCTGGCATAAGGTGATTGAGCAAAAAACAAACAAACAAAACCCACCCAAATCTGTTACTGGAGGAACCTTTA
ATAGCATAACATCAGAAAGACTATTTTCAGTAAGGACCCTAGGCAGGAAAAGACAATTCCTCCAATGGACTTGAGAA
GAATACCCTTAGTCTCTGAGGTGGTTAGAGAGGCTGAGATCTGAACCAGCTAATACCAGAGCCAGCTTCTTGTAACC
CAAGATAGGGGCATAAGAAACGGGGTGAATGCCCAGGCATGAACAGACTTGGGGTTAAGATTAAATAGGGTACACGG
GAAAGGATTTCCGCCAGCTGGAGGGGAAGCCATTTCCCATGCTCCTTAGGGATGACAACAGGAGACCCAGCCCAGAA
CATCATTGAAAATTCTATGTAATAATTTATATATCCTAACTCTTTACCCGTTCGTTCTAGATTATGGATTCAGGCAC
TCTTAGCAAAACAAATCCTCTTTCAGTCTTTTAAAATATTGACTTAACAGGAAATATCCTTGGCACAACAATACGAC
CCTGCCTGACCCTTAGAACTGTGAACAGGTTGCTCAGACTGTACAAAACCACTAACAAATCTGACCAGCCTACACTT
CACTGAAAAAGGTGATTTGCATATATTTCTAGTACTATGACATGGTGGAAAAGAAAGAACCGCATATGTGTTAGATA
AAGGTAAACGCATCTGAAAGTGCAGTAAGAATCAATGTGATATATACAATAAAATAAGATAAATAGGCCGGGCATGG
TGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCACGAGGTCAGGAGATCGAGACCATCCT
GGCCAACATGGTGAAACCCCGTCTCCACTAAAGTACAAAAAATTAGTTGGGCATGGTGGCAAATGCCTGTAGTCCCA
GATACTCAGGAGGCTGAGGCAGGGGAATCGCTTGAACCCGCGAGGCAGAGATTGCAGTGAGCTGAGATCACGCCACT
GCACTCCAGCCTGGCAACAGAGCAACACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAATTTCAGGACTTCAAAAC
TTCTGTCCTTTGAAACACTTTGATTAAAAAATGAAAAAAATATATACATGCCTGAATCCAGAGGCAAGTCTTTTTAG
AAAACATAAAATATTTTAAGTAATGTTTTTCTGAACAGAAATGGCTCTAATGTGAGAGGTCGATGAAGATTAGCCTC
TTGCTTCTTTATGTTTTAAAGTATATTCCTTGAGACCCAAAGTTTCTGACTTTAGAATCCAACTCTCAACTACTATC
AGCTATCCTCTAGACACTTTCAAAATCCTCCAAAACCTAATAGTGGTGCAAGATTTTGTCCTTCCTAGTGATAACTA
GATTTAAGTAAGAAAACGTTAAAGACTAAAAAGTGTTTAGGACCCACAAGAAAGGTGTAATTCTGTTCCTCATGGAG
ATATTTTAGTTACACACTTCTTATTCTGGTCCTCCCCAAAACACAGAAATCTAGAGTTGACATTTAAAGAGAAAATG
AGATTTTTGTCTGGAAACAGGAGAGAAAAAGTTTTCAAGAAGTATAGAAAGCGTATGAAAGGACACAGGTTTGAAAT
GGCCTGTGCTCATGAGAGAATGCACAATAGACATGGGTGTGGAGGAGCCTGGAAAGAATGGAGAGAGGTAAGTCTGG
AGGGTCAAATCTGGTTCAGAGTACAGAGCATAAAATCCCATACGAATGTAGAACATGAGGAGCCAATAACCATTGTG
TGAGAAATGCAGGAAAAAGAGGAAGAAAAAAATCCACATAGTGTACTTACAGTCAGAACAATCCTAGGAGGTGACTT
ACATGTATGATGTCATAAAATCCGTAACACAATTCCATTAGATTGGTACTGTTTTCATAGTTACACAGGTGAAAACT
TTTGAGGGTCAAGATGTTCATTGTAGCAGGCATAGTATAATTCCTTACCTCAAGATGCCTCCAATGTCAGTGAACAT
TACTTTTCCAAGATTGTCTATTTAGGAAAAAGTAAGGCCAGTTGTCTATGAAGGAAGTACTATTATAATCGCCACTT
TACAGAGGAGTAAATGGAGATTCAGAGAGGTTAAGTCATCTGACTATAGTTACACAGCTAAAAAAAAAAAAAATTAT
AGAGCAAGGATTCAAATACAACAATCTGGTCCTAGAAAGCCTAAATATTCTACACCCTGCTTTCTGATTCAGAGCTA
TTTCCTGCTTTTCTCCCTTCGATAGGAAAAAAAATGTATGTGTGTGGGGTTGGGGGCAAATAAATTGACTGGCTGTA
TTTTCTCTCTTTAATATTTGATCTTTCATTGTCTGCTCCAAGCAGTGGGTCTTCCTCCATTATCTTCTCGTGAAAAT
AGCTAAAATCACCTAACTAGTAAAAACCTAGGCATGTTTTGTGTTCTAGCCTCCATTCTTCCATTCTAATGGGTGGC
TGTTAAATTTTAGTACTCATTGTATGTAGAAGACCTCTAAGTTCACTAAAGAGCTTACTGCACATTGACTTTTCTTT
AGATTCAAAGGGAATCAATTGAGCACCTCCTATACTCCAGGACATGTGCAAGATGCAAGAGATAGAGCAGTGAGAAA
GTCTGATAAGGTTTCCCCTCCCTTTCCCAGCAGTGTACAGATAGATGTTCTTCAGGTGGTGCATGCTGGAGAAGAGA
AGACCAGAAATACTGGCCTGACGGGTCTATAGATGAGATTATCTGCCCATGACTGGCGCAGGGATGGCCGCATTGCT
AGACCTCTCCATCATCCTCTGAATATGCTTCAGCTAATTTATCATCCCAATGGCATATTTAGTCATCATGAACCATT
CTCTCTTTTTGAGTCTCAGGCCCTGGCCTTGCTTTTCAATAGACTTCCAATAGATTCCTTCCTCCTTCCTTCCCTTA
CTGCCTGGCTTCCTGACTTCCTTCCTTCCTGCCTTCCTGCCTTGTCTTGCCTTGCTTTTCAATAGATTCCCTCCCTC
CCTCCCTCCCTCCTTCCTCCCTCCCTCCCTCCCTGACTTTCTTCTTTCCTTCTTCCTTTCTGTCTTTTTCCAAACGT
GCTTTTCAGGAAACAGTGGTCTGCTTGTTGAAGTCTGATAATTCTCTAGTTCCTCATCCTTCACTTTATTGAAATTT
AGTGTGACATGATCATTTCCACCTCATTAGGTACTTCTTTCATGGTCAGAAGTAAGTCTGGAGAAAAAAAAAAAAATC
TCTTGTCTCTGCTCTTATTTGAGGGTCAAATTTTCAATAAGTCACTTTTTAAAAAAGCATTTTCTGACACTTCCCAT
AAGCTTGAAATCCTCCTAAATTGCTGTATATTGCCTCAGTATACCTTGTGTATCTATTTTAGGAACACTCCATCCAC
ATTTGCCAGTCAGCCTGGTGTTCTGCAGTTAGTTCCTACAGTGATATTTTAATTTAGCTCTCTTTTCATCCTCACAC
ATGCATCCTCTCTGGATATTTAGCTCCTTTCCTGGAATCCCTTTTAAGATTTCTAGATCTTTTTGCTTCCATGATTT
CTTCTCCTGGACTGTGACACAAAATGCTATTTCTTCTTTACATTACATTTAATTCTTTCTAGAAAGAGCCTCAGAGT

[0346] AGTCAGATAGCTTTGAGAAACAAAACTTTTTCTTTATTGCCTCACTGTTACTGCCTTTCAATCATTGTTTCGTGACA
CAAATTTTTTTATTCTCTCTGACAATTAAAACACTATTTTTTTCTGTCTGCATTGATCAAAATTAGTTCCTTCATTC
ATAGAAAACTCTTGGTGTCCCTGAGAAGCTTGAGAGACAAGAAACATTCTTCCATTCTACTCATCTTCTTCTCTAAT
GAGGAGACAACCTTAAAAGCACAGTTACATAGCCATAAAAATTAATGATTGGCTACCTCAGAATGAAAATTCAATGT
CTCATTTTTTTTTTAATATTCTTAGAATCGTTCACTGGTTGTCCAGTGTGAGTCTCCTGTTGAGATGTCTTTTGCAGC
TTTCCTTGAAACCTTTCATTCCAAACTACATAGTCCAATAATTTTGCCACCAATCTTCTGGTTATATTATGCTCTTG
AGTCTGTTGTCTATAAACTTGATTAGGCATTCCTTCCCCTCACCACTCACCTCTGATAACCCAGCTGTGTGTTGGTA
TTTAGTATCAATTCACACCAGCAAGTTCAGCCCTCTTCAATCAATATAGGGCCACACACGGACTTTTGACTGACTAC
TCCCCAAGTATTTCACATTTTGGGGCCTTATCTCCAGTTTCTCACCACAGTTGTTCATCACTGTGTTTCTTACTAGC
CAGGCGTTTATAAAAACAGTAATACCTAACACTATTGATCACCTACTATAGTGTCAGGCGCTGTAATAATATTATTG
TGATGATGATGATTATGCTGCTCTTTCTGGCATTGTCATACGTGTATTGCTTGTACTACTCACTGAATCTACACAAC
TGCCCTTATGACATTTACCCTGTTATTATTCCTCTTTTAAGGTAAATACATGAAAAATGCTTCCCACTTTGCCTTGC
TTACTGCTTATTGCTAGTACTGAACAAATGTTAGAACTGAAACTTAGAGAGGTTATGTGGCTTTACCAAGGTCCCAG
AGTTCCTAGGGCAGAGAACAGGATTGTCTACCAGACATTTTAATTCTAGTACTATGCATCTTAACCATTACCATAGG
CTGACTTACTCTACAGTGTCCAACACTATTCATATTAAGATTTATTTAATGACTTTGAAACAGTATTTCATGTCTAA
ATAGAAAAACTACTAACTCGCATTTTTAAGAAAATATTGTATCTTGGTTTTTCTTCACTGCTGGCCAGTTTACTAAC
AATCTGAAATAAAAAGAAAAAAATATGATAAACTGCTCCCAGTATAAAATACAGAGCTAAGACAAGAACGTTTCATT
GGCTTTGATTTCCCTAGGGTCCAGCTTCAAATTAATTTACTTCCTATTCAAGGGAATTTTAAATCAGAAAGAAGATC
TTATCCCATCTTGTTTTGCCTTTGTTTTTTCTTGAATAAAAAAAAAAATAAGTAAAATTTATTTCCCTGGCAAGGTCT
GAAAACTTTTGTTTTCTTTACCACTTCCACAATGTATATGATTGTTACTGAGAAGGCTTATTTAACTTAAGTTACTT
GTCCAGGCATGAGAATGAGCAAAATCGTTTTTTAAAAAATTGTTAAATGTATATTAATGAAAAGGTTGAATCTTTTC
ATTTTCTACCATGTATTGCTAAACAAAGTATCCACATTGTTAGAAAAAGATATATAATGTCATGAATAAGAGTTTGG
CTCAAATTGTTACTCTTCAATTAAATTTGACTTATTGTTATTGAAATTGGCTCTTTAGCTTGTGTTTCTAATTTTTC
TTTTTCTTCTTTTTTCCTTTTTGCAAAAACCCAAAATATTTTAG (SEQ ID NO: 834)

[0347] 智人肌养蛋白(DMD),内含子50靶序列1(NCBI参考序列:NG_012232.1的第1524636至1524685位核苷酸)

[0348] GTAAGTATACTGGATCCCATTCTCTTTGGCTCTAGCTATTTGTTCAAAAG (SEQ ID NO: 835)

[0349] 智人肌养蛋白(DMD),内含子50靶序列2(NCBI参考序列:NG_012232.1的第1570168至1570417位核苷酸)

[0350] CGTTTTTTAAAAAATTGTTAAATGTATATTAATGAAAAGGTTGAATCTTTTCATTTTCTACCATGTATTGCTAAACA
AAGTATCCACATTGTTAGAAAAAGATATATAATGTCATGAATAAGAGTTTGGCTCAAATTGTTACTCTTCAATTAAA
TTTGACTTATTGTTATTGAAATTGGCTCTTTAGCTTGTGTTTCTAATTTTTCTTTTTCTTCTTTTTTCCTTTTTGCA
AAAACCCAAAATATTTTAG (SEQ ID NO: 836)

[0351] 智人肌养蛋白(DMD),内含子50/外显子51连接(NCBI参考序列:NG_012232.1的第1570388至1570447位核苷酸)

[0352] TCCTTTTTGCAAAAACCCAAAATATTTTAGCTCCTACTCAGACTGTTACTCTGGTGACAC (SEQ ID NO: 837)

[0353] 智人肌养蛋白(DMD),转录物变体Dp427m,外显子51(NCBI参考序列:NM_004006.2的第7554至7786位核苷酸,NCBI参考序列:NG_012232.1的第1570418至1570650位核苷酸)

[0354] CTCCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCA
TCTTCCTTGATGTTGGAGGTACCTGCTCTGGCAGATTTCAACCGGGCTTGGACAGAACTTACCGACTGGCTTTCTCT
GCTTGATCAAGTTATAAAATCACAGAGGGTGATGGTGGGTGACCTTGAGGATATCAACGAGATGATCATCAAGCAGA
AG (SEQ ID NO: 838)

[0355] 智人肌养蛋白(DMD),外显子51靶序列1(NCBI参考序列:NG_012232.1的第1570442至1570487位核苷酸)

[0356] TGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGA (SEQ ID NO: 839)

[0357] 智人肌养蛋白(DMD),外显子51靶序列2(NCBI参考序列:NG_012232.1的第1570455至1570498位核苷酸)

[0358] GGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTT (SEQ ID NO: 840)

[0359] 智人肌养蛋白(DMD),外显子51靶序列3(NCBI参考序列:NG_012232.1的第1570465至1570506位核苷酸)

[0360] GAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATG (SEQ ID NO: 841)。

[0361] 智人肌养蛋白(DMD),外显子51靶序列4(NCBI参考序列:NG_012232.1的第1570442至1570506位核苷酸)

[0362] TGACACAACCTGTGGTTACTAAGGAAACTGCCATCTCCAAACTAGAAATGCCATCTTCCTTGATG (SEQ ID NO: 842)。

[0363] 智人肌养蛋白(DMD),外显子51靶序列5(NCBI参考序列:NG_012232.1的第1570518至1570567位核苷酸)

[0364] TGCTCTGGCAGATTTCAACCGGGCTTGGACAGAACTTACCGACTGGCTTT (SEQ ID NO: 843)。

[0365] 智人肌养蛋白(DMD),外显子51/内含子51连接(NCBI参考序列:NG_012232.1的第1570621至1570680位核苷酸)

[0366] GATATCAACGAGATGATCATCAAGCAGAAGGTATGAGAAAAAATGATAAAAGTTGGCAGA (SEQ ID NO: 844)。

[0367] 智人肌养蛋白(DMD),外显子51/内含子51连接靶标1(NCBI参考序列:NG_012232.1的第1570623至1570674位核苷酸)

[0368] TATCAACGAGATGATCATCAAGCAGAAGGTATGAGAAAAAATGATAAAAGTT (SEQ ID NO: 845)。

[0369] 智人肌养蛋白(DMD),内含子51(NCBI参考序列:NG_012232.1的第1570651至1614861位核苷酸)

[0370]
GTATGAGAAAAAATGATAAAAGTTGGCAGAAGTTTTTCTTTAAAATGAAGATTTTCCACCAATCACTTTA
CTCTCCTAGACCATTTCCCACCAGTTCTTAGGCAACTGTTTCTCTCTCAGCAAACACATTACTCTCACTA
TTCAGCCTAAGTATAATCAAGGATATAAATTAATGCAAATAACAAAAGTAGCCATACATTAAAAAGGAAA
TATACAAAAAAAAAAAAAAAAAAAAAGCAGAAACCTTACAAGAATAGTTGTCTCAGTTAAATTTACTAAAC
AACCTGGTATTTTAAAAATCTATTTTATACCAAATAAGTCACTCAACTGAGCTATTTACATTTAAACTGT
TTGTTTTGGCACTACGCAGCCCAACATATTGCAGAATCAAATATAATAGTCTGGGAATTGATTATTATCC
ACTCTTCTAAGTTGTCTGTGCCAATTTGCCTTCTCCAATGATAAGGATAATTGAAAGAGAGCTATAACTT
AAAAAGAGAAAAGTAACAAAACATAAGATATTTAAAATTACCCTAGATCTTAAAGTTGGCATTTATGCAA
TGCCATGTTCAAATGAACATGTTTTTAATACAAATAGTGCATTTTTCAGCCTCAGTGTAATCCATTTGGT
AAAATTATGACATCAACTAGAAACATTAGAATACATTGATGTAAATATGGTTTACCTAGCTAGATCAAAT
ATACTATATATCTTTTATATTTGTGAATGATTAAGAAAAATAATGTTGGAATTGTTATACATTAAAGTTT
TTTCACTTGTAACAGCTTTCAAGCCTTTCTAAAGAAATACAAAGTTGTGCTGAAGGTATTTAGGTATTAA
AGTACTACCTTTTGAAAAAACAAGAAGTGAGGCAGACAGAGTAAGGGGAATTTCTTTGTAAAATAAACTT
CACCAATTCCATAGGAATAAAAGTAATTTGATAGTAAACAACCTGCATTTAAAGGCCTTGAGCTTGAATA
CAGAAGACCTGAATTCAGTGCCATTTGCAAATGATGATTGTGGTCAAGCCATCTCTGGATCTTCGTTTCC
TATTCTGAGTACAGAGCATACAGAGTACACATTCACATTCACAATATAGTTATGGATATGGATGTATATA
AATATATGTAAATACTACATATATGTACCTAAAATTTGTTTTACTTCTGCTTTAAAAAAAGTAATTATAG
CCACATTTTTCAGAAAAAGTAACTGAGGCTCATAGATGTCAAATTCCCAGTAAGTAGCAGAACAAGGATT
CAAATCCAAGTCCATTTGATTCCTAAGCTTGTGTTATTACTTGCTACTGCAGAGAGTATACGTAGCAAGT
AATATATGTACTGCAAGCAATACATACTATTGCTGCGGTAATAACTGTAACTGCAGTTACTATTTAGTGA
TTTGTATGTAGATGTAGATGTAGTCTATGTCAGACACTATGCTGAGCATTTTATGGTTGCTATGTACTGA

[0371]

TACATACAGAAACAAGAGGTACGTTCTTTTACAATACCATATTGAGTTATATAATACTCCCAGGACTTTT
ATTTACCAAAGGAAACAATATTTTATAATGTTTAAAGCCCAGGTTTTGAAGTTACATTGTCTGGGTTCAA
AGCTTGGCTCCCAAGCTGTGTGACCTTGAGTAAGTTATTCTGCCTACCTGAGCCCAAGTTTATCTAGCTA
TAAAATGGGGATAGTTGTACTATCTGCCTTGCAGTTTGTCATCAGGATTAAGTTGGTTGGTACATGAAAA
ATGCTTCCCACTTTGCCTTGCTTACTGCTTACTGCTAGTATTGAACAAATGTTAGTAATTATATTTGGTT
CCACCACGAACTCTAGAAATCTAACCAATGATGGCATTTGTATTATGCAAACTGTATATCACATCATAAT
ATTATATGGAAATGAGAGCTTGTTTCCGCTTCTGTAGCCTAGTCTACCATTGACATAGCTTCCTGCAGAA
GTTACCAGATAATAGATTGGGAGAGAAAGTCCACACTTCCTTGTGACGGGTTTGTGAGTCCAGCATTTAG
GGAAGCCATTGATGTGCTCAGTAGTCTCCAGAGTTCTCTAAATAAATGTGTCCTTTTCAGAAAGGACTAC
TGATTTGATGCCCCCTCACAGAGATCGTCTTTAAATATAGGTCAAAAACTAATGTAGAGGGCCAGGTGCA
ATGTTTCACGCCTGCACTCCCAGCGCTTTGGGAGGCTGAGGCAGGTGGATCACTTGAGGTCAGGAGTTTG
AGACCAGCCTGGTCAACATGGCCAAACCCCATCTCTACTGAAAATAAAAAAATTAGCTGGTGTGGTGGCC
CATGCCTATAATCCCAGCTACTAGGGAGGCTGAGGCAGGAGAATCACTTGAATCCTGGACCAGAGGTTCC
ATTGAGCTGAGATCACACCATTGCACTCCAGACTGAGTGACAGAGTGAGACTCCATCTCAGAAAAAAAAA
AATTTAGGGGGAAAAATCAAAAGCCATTTCTGAGACACAAAAATACAGGATTTATAAATTATATATGGTA
TATATAAAAATATTTTTAAAATAGTATATATAGCATATTATATATAATGATATGTAATGTTCATATATTA
CATATTTATAAAAAAATCTAATCTCCCTTCTCTTGCTTGCTGAATAGGGGGATGCTTTGCCTGCCTCTTC
CTCTTATATTAAAAAATAATTCTTAAAGACATTGTCAGTTCTTGGCTTTTATAGCCTCAATCACCAAATT
GTCGGTAAAATGGCCCTAAATAATCATTAAACAAATGTGTGTGAGAGGGGAAATAAGAAGGATAAGTAAG
TATGGGGAGGATTTTGTTATAATTTCAGGAAATCAATATCAATTTTATGTAAAGTTTTAAATAAAGCAAT
CCCAACTTTAATGTTTGATGTGTGAAAAATTAGGCAAAATTCCAAAAGGGCTTTATAAACTGAAAAAAAC
TTTACTAACACCTATCCATTTTTATTATTTTAACCAACTTCTATTGAGCTGCCACTAAGTACCTGGGAAA
CATAAAGTTGTACAACATAGAATGTGCAGGTAAAAGAGGTTGAAGGAAGAAAATAATAACACTATGATAG
AGATAAATTTTAGGATAATAGCTAACACATATGATATGCCAGTCATTGATCTAAGTACTTCACGTGAATT
CTTTAATGCTTACAACATTACTGTGAGGTAGATAGAGAGGCACAGTAAGGATAATAACCTGCCTGAGATC
GAGGAAGAAAGACAATGATGAGATGTGAACTCAGGCAGTTTGGTTCCAGAGTCCTCTCCCTTAAACCTCA
TAGTTTTCAACTTCTCTGATATTGTGTGGGTGATGCTGTTGGGGCTTTCTTCAGGGAAAACTAAGCCAGG
AGAGAGAATGGATGCTAGTGAGATATTCCTGAAGAAGGAAAAACTTAAGCCAGGCATTAAAGAATGAGTT
GGAATTACCTAGCTAGATAAAACGAGAAGGGCAATCCAGGCAGAGGGAACAGACTGTGCTTTTCACTGAG
GTGGAAAAAAAACAGAGTATATCAGAGGAATTGTGATTCCATATGGCTGAAGTTAAGGGTATATGATGAG
GAAGAAATTGATGAGGTTGAATAGAGAGGACTGGGGCTAAATAATGGGAATCCTTTGTTGCCAGACTGAG
GAATTTTGATGATGGCCTACAGGCAGTGGCAACTCTGAAAGGATTGTAAACAGGAAAATAAAATCATCAC
ATATAGTTTAGTTGCCTATCAATTAGAGCTCTCTGGATGCAAGCAACAGAAATCATTCTCTGATTAAATC
AGGCAGAAAGTAAATGTGCTGTAATTAGCACAAAGGCATTGGAACAAAACTTACAAAAGGAAAAAGAATC
TGAGCATGCCTTTCTGGGCATGTGGCTAGCAAGAAGTATTCCAGTCTGTTTGTGATACTCTCTTTTCTCC
ATCCTGTGTGTAACTCTGTTCAAATTTTAAAGTCTTAAAAGAGAGTCCAGTTCACCTTGTTTGGGTCACA
TGTTAATACATGAGCTAGAAGGGAGCAGAAAACTTTGATTTAAATCCCTCTCCTCCCAAAGTCTCAAAAT
TAGGGAAAGGCAATTCTCCTGAATAGAAACTGGGTTCTATTGACAATAGAAGAAGGAAATGATTCTGACC
AACCACTAAACAATAATTGTCCACTGAACTCAGTCAAGAACATGTAGAATAAGTTGGAGGATAGAGCAAA
TAAAGGAGATTTGTAGGAGGTAATTATTATGATCTAAAGCAAGCTTGTTCAACTCATGGCCTGTGAGCCA
CATGCGTCCCAGGATGGCTTTGAATGTGGCCCAACACAAATTTGTAAACTTTCTTAAAACATGAGATACT
TTTTGTGACTTTTTTTTGCTCATCAGCTATCATTAGTGTTAGTGTATTTTATGTGTGGCCCAGGACAATT
CTTCTTCCAGTGTGGCCCAGGGAAGCCAAAAGATTGGATACAGCTGATCTAAAGCAACAGGTTCATCTAC
TCAACTTCACAACGTGTAGACCTGAAATAAAGACCATTCATATACCAATACCTGAAATATAAATTTGTTT
GACCATGACACGTACAGTAATTGGTTCTCAATAAATGTGGATAGCTTGATGGATAATGTGAATGCAATGT
GATAAGGAAACTTCATATTCAACAAAGACTGGAATGTGAGGATTATAATTCCAAAGCACCAGAAGATAGA
TAAGATAATGCAATGAGACATTTTATGACTCAAGGCAAAGTTAGTTATGAGATTCAGACCAAACCTTAGA
CGTGCAGTAATTGAAATATTTGCCACAGAAGGGGTATAAGGACATGACATTCAAGTAAGCTAACCTTTCA
CTAGCTTTAGACTTTGAACTCAGAAAACATATTTGGTGAAAAGCTTATGGTCCCCTTTAGTATGTATTGC
TTGATTAAAGTATTATTTTAGAAAATGGTGAGCTGCTTCCATTTTGAAATAAAAATAATTTTTACTAAGT
GAATTATATTCAGTGAAAAAAATGGAAGCTACAATTACAACTTTAATTTTTTTAAGTTTTAAGAATACAG
CCATTTAAAAAAATTAAGCAAATCTGCTTCATTTTAGACAGTAGAAAATATACCATTATCTTTTAGAAGA
ATAGAGATGTGAAATATGCAAATTAAGCCTTTAGAAGTAAAGCACACATGAAGTTCAAAGTTTAATTTCT
AGAATTGTGAATCAATAGCAGTGGATGATTTGTACTTTATAGCTTAGTGTCGGAGAAATCTGATTAAAAA
ATGCTTTTTCTGTTTCATCACATAAACATAAGTAAAATTGCTCTGAAACAACAATATTTGACAAGAATTA
GCAGTTTTCTTTTTTGACATAATCTATCAAATGAAGGGAAAAATATGTCCTGGGTTTTGCTTTGAGAGTG
ATTACTAAATCTGACCCTTAAGGAAAGGAAGGAGAGAACAAAGAAGGGAGGAAAGAAAGGGAAGGAAGGA
GGAGGAAAGGGAAAAAAAGAAGGAAGAGAGGAAGGAAAGCAGGAAAAAGGGGAAGGAGAGAGAAAAGACA

[0372]

AAAGAAAGGTAGCAAGGAAAGAAAAAAAGACAAGAAAGGAATATTAAAGAGGACAAAAGAGGAGTGAGGA
AAGGAGGAAATGGAAGAGGGATGGTGGGAGACAGGAGGGAGAAAGGTGGAGGGGGAAATATGAAGAGAGG
TTCCCAGCAGTGGAGACTAGTGTTGCTATCAACAAATAGAATTTAGATGGCCATATGATATTATTTTTCA
TAATACTGGTGTCTGATTGCCTGTGCTGAGTTAATTGTAGTCTTTTTTTTCAATTCCGTTTGGCCAGGTG
TTCAGGATAATTCACCACAAAATCTCAACCACTGCACTTGTATTGAATAAAGAATTGAGTTGGCAAAGGC
ATTTTATCCTCCAGTAAGACCTTTCCAGATTGGGGTTGAGACAAATTGGCCAATCTGGACAAGATGATAA
TAGCATTGTTCAAGATTAATTTTTAACCACACATTGCACTGTTACCTGGGAGATTTCATTATCTAAAAAT
TGAATGAGCAGTTTTAGTGGGTATAGTGTATATTTAAATGGGACATAATTACTTGAATGAGTTTAATTTT
TGTTGTTGTTGTTAAGGTCAAAGTACTTAAAAATTATGATTTTTTAAAACTCTGTCTATACACAAAAAGC
ATTTGAATTAGCTACAGAATAATTCTGATTATAACTTTTGGTGAATAGATTCAGTCAAAATCTGATTACT
AAACAACTTGTGTAGTATAGCCCTGGAAGAATTGATGGGACAATGTGTGGGTAAAGTGGCATTGGCTATT
TAAACTAAAAGCAATACAAAACAGAATGTTTCTTGGTTTTATTCTGTTGTCACAAACCCAGCAGAAAGTG
GCTATTACAATAGTTTCCCTTATTCAACAAATGAGAGAAGTTATAGACAATTTAGTTAATTGATCTAAAG
TCACTTAGTAAATGTAATTGTCCTAACATAAACCCAGACCCCCAGACCTCTTGGGAATAGATAATGTTTC
TTACTTCTTTTCTATTTCCTCAGCCACCCCCCTCAACTTCTTACACATCTCATTTCTCCATCCAAATTAT
AACAAAACAAAGCAAACATGGTTTATTTCCATGGGCATCAAATGGATTTCACGAGGTTGGGTGACAGTCA
TCTTAGGGTGAGGAGATTGATTATTCTGTTTTTCTCTTTCATCGATCAACAATCCAGCCCTTCTCATCTC
ATCATTTCATTTCTGCACAAACTTGTTTAAGAAATACCAATTAAGAAATTAATTAAGAAATTAATGTTGT
AATCTGTTTGGCTGAAGATATTTACAAATTTTGTGCTTTAATTATCTTCCAACAAATGTACATGTCTCTG
GTAGACAGCTTGCGACCATCTGGATGACTGATCCATATTTATATAATTTTCTTTCTTTACCTAATGAGAC
CAAATCCACTATTATCTTCAACGAAGGATGTAAAGATATGTCAGTGTCAGTAATGTGACTTATTTTATAT
TCTCTGGTCATAACAAAAATAAACCGCCCCTTAAATAAAAAGGTCATAGAGTTGCAAACACACACACACA
CACACACACACACACAAAATCATATTTTCTAAGTCTCCTAATTACCTTTTTATGGAAAATGATACCATAT
GCTTTTTTCTTAAAGAAACTACATAAACTTATAAACTATACTAAACTACACATTTCAAAGTCTATGAATG
GAAATGTGTATCTTATTATATTTTAATTCAATTCACTGTAAACTTTTCTGTCAAAATCTTATCAAGCAAA
ACTGATCCAGGATATTTACATGAATTCTGATGGAAGTCACTGTACTGTGTTTTCCATAAAATACCAGTGG
GATTCTGATAAGGAAGTTTATGTTTGCCATTGTGTTTAAATAGAGAATTCTGGGCCGGGCATGGTAGCTC
ACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGTGTATCACCTGAGGTCAGGAGTTTGAGACCAG
CCTGACCAACATGGAGAAACCCCGTCTCTACTAAAAATACAAAATTAGCCGGGCGTGGTGGCGCATGCCT
GTAGCCCCAGCTACTCCTGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGATGTTGCAGTGAG
CCGAGATCACACCATTGCACTCCAGCCTGGGCAACAAGAGCGAAACTCCGTCTCAAATAAATAAATAATT
AGAGAATTCTATTTACAAATTTCCTTTCTTGGATCTAGTTAGGGCTCCTTTATATGGAGTGATTTTTATT
GTTTTCATAGAAATACGTAGAATCTGGGTCTTCTCTAACTTTCTTACAGGAAAGCAATGTAATAAGGTTT
TTTTTTAATTTTCTGAAAGTTATATAATGTTATTGTTCCCTAAAGTTTAGGACCTGCCTTTTAGGCTTTC
CATTTCACCATAACTTTTGGTCCTTAAAGTCTGTAATTGAAGTTACAGTGTGTTATGATGTAAATTTTTC
TTATTATTACCTTTAATGTTAGGGTAATGTTAACTAATGTTAATGTTAGGTATATGGTTGTTTTTTTCAT
TCCTTCGTTCAACAAATTGTCTTTGAAACCCATGTTACAAAGCACTCTAGAGTCAGGTTGAAGACTATTA
AGAAAGGAGAATAGAAAGAGACACTAGAGTAATAATTTGGATTTAAATTTGATTTCCTTGTGTATGATAG
TGAATAAGTGTGAATAAGATGAGGCAGTGATCACACATCACTATTAGTAAAAGTGTTTCTGTACCTGTAT
CCACACTTTTATGTATATGGTTACTTATGTTAAAGTGATACATATTATATAAAATTAACGTATACATTAA
GTAGATATTTTAATAGTCTGTAATTAAATACTACTAGTATTTTCTTTCCTCCTTCAAGTGCTTACTTTTG
ATACCTCGAGTTACAGTGTCATAAAGATTCTTTAGAAATATATTGACTGTCTTTTAAGAGCTTTTGATAC
AATACTGAGTTTACATTCATCTGTTATTTATTGAACACTTGCTGGTGAAAGGCATCAGTGTTATCTGCTC
TTAGGGAACAAAAATTAAAAAGGGATAGGCCCTAATTTTAGAGTGTATCCTCTATAAGAAAAACATAAAA
GATAGGGCAGTCATGGCCACAAAAGAAAAAAGTGTTATGGTGGTTTCAATCATATATGTATTAGAATGAA
TAAATCAACTGATCAATTGTGATTTCTTATTCTAAATATGTGCCTGCCTTTTTCATATAGATGAAAATTA
AGCTATGTTTATCTTTCCAGGGATCTTGTTGATTTTTATTCAATAACTTGGGAGTGAAAGTTGATTTTTG
CATATGTTTTAATGTTTTTAAATTTCATAAATGAATTGATCAGTAATTTCCAAGGTAGTAATGGCTGCAT
TGTTTTTGAAAAAAAAAAAGCAACAGGATTTGATTGTGCTTTTATGATTTTTAAAGAATTCATTAAAAAT
AATGCCACGGTTTCTAAAATGATTTGAGTCAATTTCTTATTCGATTTATAAAAATAACTTTGAATACAAT
TTTAGTAATTCACAAATGCTTTCAGTTCCCTTACCTTTATATTTTATATTCTGTGTAAACAAGTGACATA
ATATTTAAGAATTATATATCTCCTATGATTTATTCAAGAAAAGAATATATACTGTATTATTTATTTCAAG
AACAGAAATGCTTTGATTTAACTGTCATCTTCTCTCTTCAATTATGGAAGCAAAATAAACTGTAATGACC
AATGTAACCCCTCCCCCATATCAAGTTAATCTATGTTCAACTCCAGAATTATTTTTGAACACTCAAACTA
GAAATTAAAAAAAATTAAATCCATGAAGACGATTTTTGCCAAAAGCATATAGATAAATTGAGTTGATTCT
ATACTTAAGAAAGTGGAGAGGAGAGAGTAATTTGGAGAGAGTAATTTACTCTTAATCCCATATTTTTTCC
CTAAATGTGAAAGAAGTAGATTGTAGTGAGAGGGAAAATAACCTGTAGCAACTTCATTGAGGCTAAGCTT
TCTGTCATGTTATATTATACGAAAGTAATGAAATGCTTCCACAGATAGAATCAGAAGTCCCCTCTGAGAA

[0373]

ATTCTACATAAAAATTAGCCTGCCACTTTACCACACTTACTCAAGTTTGATTTTTTTAAGTTATGTAATA
GATGTTAGGCACTAGAAGAGGACATTTACTGGGGGCAAAGATCAGTAGTTGGAAAGAATGCAAGCAGGCA
AGAAGCTATATATAATGAGATTTTACAGTACAATTGTTTTCTAAATGAAAATGAGGACGGGTCCAGACAC
AATGGCTCACACTTGTAATATCAGTGCCAGGATGGAGGATCCCTTGAGGCCAGAAGTTCAAGAGCAACCT
GGGCAACAGAGTGAGACTTCATCTCTACAAAAAAATGAAATAAAAAGTTAGCTGGCTGTGATGGTGTGGG
CTTGTAGCCTTAGCTACTCAGGAGGCTGAGGTGACATGATCTCTTGGGCCCAGGAGTTCGAGGCTGCAGT
GAGCTATGATAGCGCCACTGGATTCCAACCTGGGCAATGGAACAAAACTCCATTTCTAAAAAAGAATAAA
ATATAAAACTAAAATAATAAATAAATAAAAATGAGGATATATTTTATTTTAACATTTGGAAACTTTGTAG
GTGAGGACCATGCAAACATTCAAGGTGTGAGTTCTGACCAAATCCAATTATTAACCATACCAATGACTTA
AGGTTTCTTCACACTCCTTAAAGTTGATTAATATAATGATTATATAGTTGACTGGTATGTCACAGCTTGA
AGCCTTTGAGATTTATTCCTGCCTTTTCTGTAAAGGTTGTTTTGTTAATTCCAGTATGTACTGGTCGTTT
TTGTTTTGTTTTGTTTTTGTTTTTGTTTTGTTTTTTTGAGATGGAGTCTCGCACTGTTGCCCGGGCTGGA
GTGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCCACCTAGGTTCAAGCGATTCTCCTGCTTCAGCC
TCCTGAGTAGCTGGGATTACAGGCACTCACCACCACACCCGGCTAATTTTTTTTATTTTTAGTAGAGATG
GGGTTTCACTATGGTGGCCAGGCTGGTCTCAAATACCTGACCTCATGATCCACCTGCCTTGGCCTCCCAA
AGTGCGGGGATTACAGGCGTGAGCCACCGTGCCCGGCTGCCAATATGTATTGGTCTTTTTCATCAATGAT
TCAGTCCAAAATCATTTTGTCCTTTAACTATATATTTTCTTGTAAAGCTGCTTCTGTTGTCTTGAACTTT
TCTTTTCAAATGTATGTTGTCATTTGACTTTTTAGATTGTTATTTTCTGGTCCTCGAAATAAATTTAAAT
TTCCTGTAAAGGAAGGTGTAATATTCTATTTGACATAGCCGCTAAAGATGTACTAGGTGCTTTATAAATA
TTGTTGATTTACTTTATCTTCACAGATTACTAGTTTTACTTAGTATTTGGAATATGACAACATTTTATAG
AGCTATATTCATATATATGTTTATCTTAACTGTTAAATGCAATATGATTCATGTCTTGTTTTGGTCAATG
ATGAATGAAAGTCTCCTGAGAATTAAATTTACTGCATCGATGCAAAAACAATCATAATTTTAGACACTCT
AAGAATTTTAGAAATTAAAGGATTTTTTTTTTCCAGTTTACTCTGTTAAGATTGTGTTTAGCTATGCGTG
ACAGCATTCTCACTACAGTGGCTTATCCAGATAGTTTCTTTTTCTCATAGAGCAAGACTTCCAGAATTAT
GTGTTCCAGGGTCAGTGCAGCACCTCCAAAACCGTATGTCCCAACTTTTTCCTCCAACCCCAGTCATCTC
CAACATGAGACTTTCTTTTTGTTTTGTTTTGTTGTTTTTGTTTTTGTTTTTGTTTTGAGATGGAGTCTCT
GTCGCCAGGCTGGAGTGCAGTGGCGCGATCTCGACTCACTGCAACCTCTGACTCCCTGGTTCAAGGGATT
CTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACGGGAACGCACCACCACGCCCAGCTAATTTTTGTATT
TTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCCTGACCTCATGATCTGCCTGC
CTCAGCCTCCCAAAGTGGTGGGATTACAGGCGTGAGCCACCGTGCCCAGCAAGACTTTTTTTCCTGTGGT
CTCAACATGGCTACTCTGCCTCCAGGCACTATGTCTGTACTTTAAAATGGAAGAAGGGAAAATGGGGAAA
GTAAAAGCATATTCCAGCTGTGTCAGCTCCTGTTTGTAAGGAAAACCAGTGCTTTTCTGGCAGCCCCACA
CAGAAGAGTTTCTACTTGAACAGTGCATTAACCAGAAATGTGTCACGTGACCATTCCTAACTTCTAAGGA
TCTTGGGAGGATTGAGTGTTTTAACTGAATAGGTGTGTTTCTTTCTTCATAATTCAAAATGTGAAAATTG
GTAACTTAGTTATAAAACCTTGCTAGTCTGAACAGAATTTGGATTTTTTTAGCTAAGAAGGAAAGAAAGG
GTATCGGATAGGCAGCTGGCTATGCTAGCCAAGATACTCTTAATAATGCACATTTTTCTTCTTTGGACAT
AAGCAGTTTTAACTTAGCTAAATATGATGTGATTGTTTTCCTGTCTTCTTAGTTCTGTTTAAATTTGTTT
CAGAAATCAAGGAAATAAAATGGAGAAAAAACTCTATTATTCATGTCTATCTTTCTGCCTCTGAATATTT
TTATGTTGGAGAAAGAGAAAGCAGTAACTTTCATAATAGCTTACATAGTCTGACAAATTCTAAACATGTC
CGTTAGCATCAATATACGTGGTATTAGGTCATCAGTTTTTATATTCTGAATTATTAAGACCCAAATAAAC
CCACTGAGTTCAAGAGAAAGTATACACTGAGCAATAAAAACATTACCAGTTCTAGCAATGATAATCAAAC
AAGAAAACAAGTAATATAGTCTGTTTAGAATAACATGTTTTAAAGATCAAGTTTTTCTTCCTTACCAATG
TTGCCTTTCTTGTAACACTTTTTTTTCCTTCTTGAGATAGGCTTTCCTATCTTTTGTCACAAACCCCAAT
ATTTACATGGCCATTCGTAGTCTATTCATAGCAGCACCACCCCATGGCCCAAACTTGTAGATATTGCCCT
CCTTCTATGGTTGTTCTAATAAGAATAACCACTCTTGTCTCTCATAATCTCAGCTGTTTTGTGCCGTTAA
AATGGAAAATAATGAGTATTAAGATACTAACTAGGGGCCAGGCGCGGTGGCTCACGCCTGTAATCCCAGC
ACTTTGGGAGGATGAGGCGGGCGGATCACGAGGTCAGAAGATTGAGACCATCCTGGCTAACACGGTGAAA
CCACGTCTCTACTAAAAATACAAAAAATTAGCCAGGCGTGGTGGTGGGTGCCTGTGGTCCCAGCTACTCG
GGAGGCTACTGCACTCCAGCCTGGGCGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAAAAAATAC
TAACTAGGTTTCAGTCATATGAGATGAATAAGTCCTACAAATCTGTACAGCCTAGGGCTTGTAGTTAACA
ATGTTTTATATTTAAAACTTTGCTAAGAGAGTAAATCTTTTTTTACCTGTTATTATCGTACATGCAAAAA
TAATAGTAAATAAGGAAGGTAAGAGAAAAATTTTGGAGGTGATGCATAGGTTTATGGCATAGATTCTGGA
CATGATTTCACAGGGGTATACAGTCATGCATTGCTCAACAACAGATATACATTCGGAGAAATGCATGGTT
AGCTGATTGATCTTGTTGTGCAAACATCATACAGTGTACTTACACAAACCTAGATGGTACATCCTACTAT
ACATCTAGGTTGTATGGTATAGCCTATTGCTCCTAGGCTATAAACCAGTACAGCATGTGACTGTACTGAA
TACTTCAGGCAATTGGAACACGATGGCAAGTATTTGTGTATCTAATCACATCTAAACATAGAAAAGTTAC
AGAAAAATATTATAATCTTATAGGACCACTGTCCTACATGTATACAGACTGTGCACATGATAGTGGTCCA
TTGATCAAAATGTCCTTATGCAGCACATAACTGTAATTATCTCGAAACTCATCAAGCTGTGTTCATTAAA

[0374]

TGTGTATAGCTTTTTATGTCAATAAAGTGGTTAAGAAATCAATAAAGTGGTTAAAAAATATTTTGACTAG
GAAATATACTATCATTTCTAGTTGATAAAAGATCTCAACATTTCCAAAATTGTCCTACAGAAAACCAGGT
TCATCAGGTGTTCATACATGATCCTCATGAAAAGGTCAAATAAGCTGAAAAACATGCATAGACGTTGCCT
ATCCTAGCAATCTATGATGTACATCTCCATAGTAAGGTCACTGAAAAGTCTTTTAGGAATGTTAGTATTG
TTAGCTCAGTATTTCTCAGTGATTTCTCCATGGAAACCATTTGTGTAGAGCATCTTGAGGAGCACAGCTG
AGAGAACATTATCTTAGTTGGATGTGTATGTCCCTCTGAGTCACTTGATTTCTCTACATATGCTTTTACC
AAATTAATCTTTTAGAAATCTTTTCTCTTCGCACTATGTCTATAATTTGTGAAGTTGTTACCAGGATAAC
ATTTGTGCCTCTCACCATGATGTACCTACCAGGGTCCAAGCAGCCATTCCTTCTCTAGAGCAACTGTCTG
AGGGAAAGAATTTAACACAGCATTCTACGAAATCATTTTATTTATAAAAATAGATTACTGCTTTACATAT
AGTAATTTATATTTAGAATATTGATTAATTATTAAAATCTGCATGAGAGCTTTAAAGAGTAGTACATAAT
ATATAGCAGTTTGTACTCAAACTGTCTTCTAAAAAGGATTCACTTTTTGTTTGTATTCTATTGTCCTATT
CGTTGATAGTGTTACGTAAGTAATTATAAAACTTAAAATCTGGAAAGAGAATGTGGACTCAGAATGCCAT
CTCTTTTGTTATTTCAAATGGATTAGAAATGAACATACATATTCATTTTCTTTCATTACACATCCAGAGA
AATAGAATGGATTTTATAAATATGTAAAAGCAAGGATTTTGATCACTGATAAAAAGGGAAGGTTTGGTCA
CTACCTTATTTCATTCCTTTTTTCTTATCCTTTTTTTTTTTTTTGTCAATTATTTGATGACATCTCTGAA
CATCACCTTTTATTCATGACAAGAATTGGGTATCATGGTAAAGAACACTGTTAATATAATTCAGTTACTT
CACCCCCTCCTGAAATATAGAGAAGCTTTAAGACTATGTGAATATTTTTTTCTGGTTTTCTTGTATTTGT
AGAAATAGCATGAGCTTTGTTTAAAGTCAGGCATCTAAAACCTTGCCCTGTATGTTATTGACAACCTGCA
CAAATTTTAGGATCTATTCTATTACAGTTTGTTCAACTGTAAAACTAGGATAGCAAACTCTATGTCATAT
TTTCGTTATCAGAATTTAAAAAGCATGTTTTAAGATCTTAGTAAATAATAAATCTCTACTCTGTAGTTGA
ATTTGTTCTATATTCTTTAAGAAATTCCCTTTGATGGTTATGCCAACCTCTGTATTACTTTTCTTCACAC
TTTAACTTTGCGCTGAAATCATAGTAGTATTTTACGTTATCAGTCAAAATAACAGTCATCCTTAAAACAA
ATATGAATTTTAGATGATTAAATAGATTTGTATGGAGGTTCTTCTTGCTAATCATAGCAGTTATCCTTGG
TGAAAAATGATAGACACTTGAAAAAACCAATTAATCATGATGGCTATTTTTGCATCATAAATAAAGCTTT
CAAATTTGAGAGGGAATCAAAAGGGCAATGGTAGTATAGTGTCTCAAAGCCCCTTTCCAATTGATGGTAC
AAATTTAAAAAGAGAGAGAGAGAGAGAAACATGTTTCACTGTAATTGTTTTCTAAGAGCTTCCAAAAAAG
CGTATTTTCTTAATAGATTCAAATTTTTCAGTTGGATTGAAAGGGAAGTCTTGGAGTGTAGTGAGGAGGG
CACCTTCTGTTGAGAGGTGTTCAGACGACAGAGTGTGCCCAAGGCCAAAGATGAGATGGTTTTGCGAAAG
TCAGTGGCCACAAACAGGTGTGTTTGACCCCTGAGAGATATGCAGGAAGTCTACCCCACTTTAATTCTTC
CAAATATTCTTTACCTTAATTCCCAAGTACTTGATAAAGGAGCAATGGGGAGAAAATATGCACACTATTA
TGGAAAAGTTTTGACCTACACTTTGGAGAGTTTTAGATTAAGAGCATTCTAGAAATCAGTCCCAAATGCC
TAGGGTTTACTTACTTAAAGATAATATCATAGTTTGGGTGACTGGGAAGCATACCCTGAGATTGAGGTGA
GCATGCAGTATGTCTATTTAGGAGTGTTCTTGGGGTCAACGTGTAGGGGCAGAGGGAGAAGTTGAGCTCT
GACGCAGTCTTAGTAAGGGCCTCAGCTGACCGTTCAGGGAGTTCTTAAGCTGGAATGACCCTTCAGAAGT
GCTAGGAAACGAAGAAAGGGGACTGGATCTTTATAACCCCGTGTCAAGTCATGCACTGGATGTGGGCTAC
TCCAGGAAGGCAACGAACTTTAGCAAGATGATTCTCTTTAGCCACGGGAATTTCCATAAGGGGGCTGCTA
TGGTCTGAATGTTTTTGTCCCTCCAAAATGTGTATGTTGAAACCTAACACTCAAGGTGATGGTATTAGAA
GGTGGGGGTTTGGGGGGGTGATTAGGTCATGCGGGCTCTGCCTTCAGAAACAGGATCAGTGCCCTTATAA
AAGCGGCTCCAGAAAGCTTCCTTGCCCTCCCACCATGTAAGGACACACCGAAGATGCCATTTAACAGGAG
TGGGCCCTCACCAGACAATGAATCTGCTGATGTCTTGATCTTGGAATTCCCAGCCTCCAGAACTATAAGC
AATAAATTCTGTTGTTTATAAATTACCCAGTCTAAGGTATTTAGCTATAGCGGCCCAGACTAAGACAAGG
GCTGACAGCTGAAGGCTGTCTACCAGCAGCACTCCTAGCAGCTGGGGAACTAAGTCCTTCATTTCCAAAG
GGGAATCTAGGCAGCATATTTACAGCTTTTCACTACAGATAAGCTCATTATTTCAAATAGGGACTAGCAG
GAAAAAATTAAATTGCCCAAAATTTAGTGGGATGCTGAAATAGATTGTGGTGTGTAAATTGGAGTATAGT
GAGGAGAGCACCTTCAAACCAGTATGTACTACATGATATTGTTTTTGTTGCAATATTTATTATATACCCA
AACACACATATATTACTTTTAGAAACACACACCACATATATATCTATGAATATTTTATATACACATAGGG
AAGGATTGTTGATGTTATTTATGCTATTTTAAAGATCGATGTTTTCATATAATTATGTATTGGTTATATA
TTATTTCTTGATATAAGGTAAAAAAAAAAAGCAAAACAAACTTTAAGTGATCACTATGAAAAGAATCCCA
ATGCTGCACATTTAGGTTTATCCAACTCTTCCCATTAAAATATTAAATAGTAGAAATAATTGTGAATAAG
AAAGAGCAGATTTTGAAAAATGGAAAGAAATGCTTAAAGACATAGCATTGTTGCCCAACCATCATTATTT
AAACATACAGTGTTTGGCTTTGACCAAATTGCCTTCAAACACTTCCTTTTGGCCCAAAATGTTAGGTCAT
ATATACTACCATAAAATTCATGATGCTTACCATGCATTAATTTCTAGTATATACCAGGCATTGTGCTATG
CATATCATATTCAATATTTCTAATCCTCTCAAAAGTGGTACAAGCTAACTGGCGTTTTTCTTGTTTTGAA
AGGGAGAAACTCAGAGAGGTTAAGTGACTTGCCCAAGGCAATGCCATTGATAAGTGCCAGATTCTATCAC
AGGTTTATTGGCAACAAACCATATGTGCGCGTGCATGCGCGTGTGTGTGTGTGTGTGTGTGTGTACAC
ATACACGAATAACATATATGGTATAAATACGTGGAAACATAATAAACTGCATTGAGCTGCGTTTATAATT
AGTATTTAGGACATGTTTGGCAAATAAAAACAGTGGAGATTGAAATGGATTTGCTTAGGAAAAATGATAC
ATTAAAATAGGCTTTATTATGAGTCTTCAACTATTCTGTGAAAATAGATACCCAGGGAAGAAATAATAGA

[0375]

GAATATGAATCTTGAGCAGGCAACTGAGAACTTGTCGAAGAGCCAAGATAAAAATGTCAGAGAGGAGAAT
ATTTTGGCAGCTCAGATGAGCCCCCAGAGGGTGGGAGGCAATGATCTCACCGCAGTCTCGTATCTGAACC
CCAGGTTTTTGCATCTCCATAAAGTAATTTCTTACACCCCTCAATAATGATCGGGCTTACTCTCAATCTC
TCGCTCTCTCTCTGTGTCTCTCTCACGCACACAAACATGCAGAACATTTCTTGCACATGCATAACTCATA
AGACGATTATGTAAATACCAGCCTTTTTATTTCATAACTAAATTACAAGGCCTGGTTATTGTTTGGACTG
TGAAAAAATAATTATGTGAATAGGTGCCTCAAGATGAAAGACAAGGCAAGATTGTGAAATTATTCATATG
ATAGTAATAGTATGCAAAAAATAACACAATCTTTAAAGATCTTTAACGACCTAGTTATAAAACCATGCTT
TATAACAAATATAACCATGAGGAAATAAAAAGAAAAATGTAATAATATACTCCAAGAATAAAGTCAAATG
TATTGTTGAATGTAAGGAGTTGGTTACACTTCCTTATAGTGGAGGTTATTTTAAAATTTGTGGCTTACGT
GGTGTTATGAATTGCCCTAGATCAACACTATTATGCAAGGCCAACTATTAGGTTATTTTTGGTAGATAAC
CACAGCAAAACTTTAGTATAATAGGTAAAGGTTAGCTACACTCCCATACCCTCACTCTCAGGTGTTGTCA
TACTCCGTATAAAAGGTTCAATCAAGGGAGACATGAGAATATTCCAGAATCTAGAGGCAGGATGCAGTTA
ACCTTAGAGAAGGCATCAGACAACTAGAATCTTCGGATTCAATGTGGAAACAAAGCATAGTTTAGGCATT
AAATCTTGGGCACCATTCCAAAGAATACAGGTTCCATAACTTACTATATTTTTATACCTAGCAAGCTAGA
GATGAGGAATTGCTCTCAAATATTTTAACCAAAGCATGTATCTTAAGTAACACTAATCTCATAAGTGAAA
ACTCATTTCTAATATTCATTTTGCTCATTAGCAAGGCCTCTAGTGTTGACTGTGATAAAAAATAGTTCAA
ATGCTGGTAGAACCCACCCCAGGAGACTGGCCTTTCTGATTAAATTCTAACTCTATCCCCACGTGAATTC
CTGACTTAAGTAACTGAGTTCCTGCACATCAGAATATAAGTATATTATAGATATAAAAACATATGTAATT
AATAAATATTTTAAGTGAGACACTTCTTTCATCTTTATGGCTTAACTATATCAGACATTTGATTATTTTT
AGCGGTCTAACTACAAAACAAAACACAAAGCCCACAACTAAAAATTTCTTTGTATATATTGCAAAGAGGC
AACCATTTGGTGTCAATTCAATCATGAGTGAAATGCTATTATACGAGTACATCTCCCTGGCTTGTATGGG
GGTAATAGGGCATGGAATTTACAGATTCACAATAACTGAGATATTCACAATAACAAAGATATCAATATGT
AGCTTTTCCCATAACTTTGTGTAATGAAATCCTCAGTTTGTGCTGTGTAAAAAGCTTATTGTTTACTTCT
CATGAAAATCATCTTAGTTTTTATCTTTATTTAATAGTCTGTAATTTGGGGGTAATACATTCGTTTTGTT
GATACTATGTGAAGTGGCAAGCAGAAAATTCTAACAGGAATAGATAAGCAAGTATCCTATAAATCAGAGT
CAGTGTCTCTCTCTCTCTCTTTTAATGAGTCAGTCTGTCTCTCTCTCTTTTTCCCTGCCTGGCTATCTAT
CTGTATTTTTCAGTTTTGCTTTGCAAATAAGAGAATTGTGTGTTGTAAACCAACCAACTTACCATTAATT
TTCTCTGAATTCAAAAGCAATTACAAGCGGACTCTTGAGTTTGTGCTGCCTGGTTGTCTGCATATAGGCC
AGATGTCTAGAATAGGATCTTTATTTACTATTTTTACCCTCCTAATTTCATGGTAACTCCAAGGTAGATG
ATATTTGTAATCGTACACTACTTGTCAGAAATCTTTCTAATAACACTGCTATTTTATAAAAATAAACATT
AATTCAGTATAAAATTTTATTTTAAATTGTTAATTCAAGCAAATCAGTGAGGTAACTTTTACACTGCCGA
GCGTACGTGTGTGTGGATTAGTACAGCCATGCCATAGACTTCACTTGTAATCTTTTCTTTATATTTTTTA
TACACCTGAAATGTTCATCATTGTGCTGTAGAAAACAATCTCATTGTGTTTTTAAAAGCTAGAGTGGGTA
TTGAGAAGGGGAAGAGGATCATAGAAAAAGTTGGTTAACATGCTACTTAACACTTCAAATCTTTACTCGA
TGTCATCATCAGCAACATTTTAAATTTATGCTTCTACTAGTTTGCAGTTCTTTTCCTTTGATTATTCTTA
TGATACAAGCCTTTCCACACAAAATTTATGTACAGGAATTGTGTAGAATTTTTCTTTGGAAAATATGGTG
ATTTATTACAATTTGGGCAACATCATCATTTTAAAAATTCAGAATTTGATTTTTCTCAGAATCATCAGAA
AGAATAAAGCATATATATGGTTCATGTCAGGAGAATTAGAACATGAGAATTAATATATCTCTGATCTTTT
AAAATATTTTCATGTTTGTGAATCAGCAGATTTTTCCTAGTTTGAGATTTAAAAAATCTAGATATAATTA
AAATCTCACTGATGTTTCACCATCAGATGATTTTATATTTGTATTTTCTTCCACTTCATAACTTGTATAG
AGAAGAATAGAAGAAAGAAAAAGGGAGGATTGATAATCTTTCTCTCTCAGTTCTTATAGCACTTCATTTT
TTAAACTTATTACTTCCTTCTGCCTGCTTTGTTTGTCTACATGTTTGTATTTCATGATTTCTTAGAAATC
CATCTACTGCCATTCTGAAGGTCATTTACCTGAAAATGATAGAAAGCAGCATATATTCAAACAACTGCAG
AGTAATTGTCTATATCAGTTATCATTGTTCATTACTTTTCTGTTTTAGGATTGAGGGGCTGCCTCGCCAC
CTCCCTCACACCCCCAGCATATTATCACAAAGCCTACTGATTCATTCACATCCCTGGGCTGAATTTGCCA
CCCACTGTGTGTTCCTGTTGTTTTGTGTATGGAAGTGAAAAGATTTAATTTGATGTTGTTGAAAAGACAC
AGAGGCTAACTTTCAATTTTCATATGTAGTTCTTCCCTCTCCCTCTGCACCACCTCCTTTACTTGTTGAG
AAAATTGCCCTCTCCATGGTAACAATAGAAGAAGCTTTCAGATTTTAGTAGTAGTTGTTGCAGAGAAAAG
AATTCAAAAAGTAGATGAAGTTTAAAAATGAAAAAGAGAGAGGAAGACAGCTGGGAAGAAGGCTTAATGT
TTATGAGTGGGTGTGGAGGGGAAGAACTAAGTTGAATGAACAAAGCTGAGCTAAGGGGAAGATGGTTTTT
CTGCATCCCAGAAGGCAATACCCTAGCCTTTCCTGCAGCCTTCACTCCCCAAAAGATAAGAGCTTTATCT
GAAATTCTTATAGGATTCATTCCTGAAGAGCAGCTTGTCACCAAACAGAAACACTGTGATTTCCTCAGGG
AGTCACAGTTTATTATTATTTTTTTAATGTAACGCTTTTGTGAACTCCAGTTTCCACCTCAATTCAAATG
GTCTTTTGGTTACAGGGTGAAAGAGACCCAACAATACACCTTTCCCACTTCCGGAGGCCTTTGGTTAAAC
CATGTCTGCCACAAGGACACAGGAGCCTGGTATGACTGGTTGTTTTTTGTTTGCTTTTTTGCCTCCTGTG
CTTTCTAGATTGTGAGATACTGTAACTCTTGTCGATGACACATAGTACCGAACCCACCCGAAGAAGTATG
TCAGTATGTCACATTGTGACAAACAGCTTCTCATGCTAAGTAAATGCAGAACCATTGTGAAAGGTTTAAT
AATGCCCACTCCTCCCCCGCCAAAGATGTCCATATCCTAATCCCAGGAACCTGTGAATATGTTACCTTAC

[0376]

ATGGCAAAAGGCTTTGTATTAACAGATGTGGTTAAGTTAAAAATCTTGACACGGAGAGATAGCCTGGGTT
ACCCCAGTGCGCCCAATGTAATCACAAGAGTCCTCCTAAGAGAGAAGGAGGTGATGATACAAGCAGAGTA
AAAGAGAGATTGGAAGATGCTACACTACTGGCATTGAAGATGAAGGACAGGGCCAAGAGCCAAGAAATGC
AGGCAGGCTCTAAAAGCTGGAAAAGGCATGGAAAAGAATCCTCCCCTACATCCCTTAGAGGGAATGCAAG
CTCTGCCAACACATTGTTTCTAGCTTGTGAGACCCATTTTTTGGACTTTGGACCTCCAAAATTGTAAGAT
AATAAATTTGGGTTGTTTTAAGCCATTAAGTCTGTAATCATTTGTTACAACAGCCACAGGCAGCTAATAC
AGCCATGAACATTTAGTAATGACTAACTTTGCACAATTTTAATACAAGCTTCTTATTAAGGTTTATTTTT
TCTTAATTACAAGGAATAAAAGTGGGGTCTGGGGGCAATGTCATGGTCCACTCCGTTTTAGCCATATGAA
TTTGTATTTCCAGCATTAGAACAAAAGGTGACAAATCTGAATGTATTTGTGTGAAATAATAATAAAGCAG
AACAAAAAGGGAAAAGTGTCCAGCTGGAAATGAAGTTAGAGAAAGATGAGGAGAAGCAAGCCAATTGTGT
AGTTTTCCCTTCTGCTTTTTAAAATCATGATTTGTTTAACCCACTGAATTCTATTTTAGAAACAGGACTG
CAAGGAAGTGTTGATGGATTTGGTGGCATGAGAACCAGAGTCACAGAGGCAGGAAAGTAAGGAATAAGTG
TTAGAATAGGAAGCAGAGTTGCTTGGGAAGAGACCTTATGACATGTGGACAGGGCTAGACTTAGGAGTCA
GAAAGACCTGAGTTCAAATGCTATCCTTTAGTATAGTTTGAAGTCAGGTAGCGTGATGCCTCCAGCTTTG
TTCTTTTGGCTTAGGATTGACTTGGCGATGCGGGCTCTTTTTTGGTTCCATATGAACTTTAAAGTAGTTT
TTTCCAATTCTGTGAAGAAAGTCATTGGTAGCTTGATGGGGATGGCATTGAATCTGTAAATTACCTTGGG
CAGTATGGCCATTTACACGATATTGATTCTTCCTACCCATGAGCACGGAATGTTCTTCCATTTGTTTGTG
TCCTCTTTTATTTCCTTGAGCAGTGGTTTGTAGTTCTCCTTGAAGAGGTCCTTCACATCGCTTGTAAGTT
GGATTCCTAGGTATTTTATTCTCTTTGAAGCAATTGTGAATGGGAGTTCACTCATGATTTGGCTCTCTGT
TTGTCTGTCGTTGGTGTATAAGAATGCTTGTGATTTTTGTACATTGATTTTGTATCCTGAGACTTTGCTG
AAGTTGCTTATCAGCTTAAGGAGATTTTGGGCTGAGACAATGGGGTTTTCTAGATATACAATCATGTCGT
CTGCAAACAGGGACAATTTGACTTCCTCTTCTCCTAATTGAATACCCTTTATTTCCTTCTCCTGCCTGAT
TGCCCTGGCCAGAACTTCCAACACTATGTTGAATAGGAGTGGTGAGAGAGGGCATCCCTGTCTTGTGCCA
GTTTTCAAAGGGAATGCTTCTATAGTACAAGGCTACAGTAACCAAAACAGCATGGTACTGGTACCAAAAC
AGACATATAGATCAATGGAACAGAACAGAGCCCTCAGAAGTAACGCCGCATATCTACCACTATCTGATCT
TTGACAAACCTGAGAAAAACAAGCAATGGGGAAAGGATTCCCTATTTAATAAATGGTGCTGGGAAAACTG
GCTAGCCATATGTAGAAAGCTGAAACTGGATCCCTTCCTTACACCTTATACAAAAATCAATTCAAGATGG
ATTAAAGACTTAAACGTTAGACCTAAAACCATAAAAACCCTAGAAGAAAACCTAGGCATTACCATTCAGG
ACATAGGCATGGGCAAGGACTTCATGTCTAAAACACCAAAAGCAAGGGCAACAAAAGCCAAAATTGACAA
ATGGGATCTAACTAAACTAAAGAGCTTCTGCACAGCAAAAGAAACTACCATCAGAGTGAACAGGCAACCT
ACAACATGGGAGAAAATTTTCGCAACCTGCTTATCTGACAAAGAGCTAATATCCAGAATCTACAATGAAC
TCCAACAAATTTACAAGAAAAAAACAAACAACCCCATCCAAAAGTGGGCGAAGGACATGAACAGACACTT
CTCAAAAGAAGACATTTATGCAGCCAAAAGACACATGAAAAAATGCTCACCATCACTGGCCATCAGAGAA
ATGCAAATCAAAACCACAATGAGATACCATCTCACACCAGTTAGAATGGCAATCATTAAAAAGTCAGGAA
ACAACAGGTGCTGGAGAGGATGTGGAGAAATAGGAACACTCTTACACTGTTGGTGGGACTGTAAACTAGT
TCAACCATTGTGGAAGTCAGTGTGGCGATTCCTCAGGGATCTAGAACTAGAAATACCATTTGACCCAGCC
ATCCCATTACTGGGTATATACCCAAAGGACTATAAATCATGCTGCTATAAAGACACATGCACATGTATGT
TTATTGAGGCACTATTTACAATAGCAAAGACTTGGAACCAACCCAAATGTCCAACAATGATAGACTGGAT
TAAGAAAATGTGGCACATATACACCATGGAATACTATGCAGCCATAAAAAAGGATGAGTTCATGTCCTTT
GTAGGGACATGGATGAAATTGGAAATCATCATTCTCAGTAAACTATTGCAAGAACAAAAAACCAAACACC
GCATATTCTCATTCATAGGTGGGAATTGAACAATGAGAACACATGGACACAGGAAGGGGAACATCACACT
CTGGGGACTGTTGTGGGGTGGGGGGAGGGGGGAGGGATGGCATTGGGAGATATACCTAATGCTAGATGAC
GGGTTAGTGGGTGCAGCGTGCCAGCATGGCACATGTATACATATGTAACTAACCTGCACATTGTGCACAT
GTACCCTAAAACTTAAAGTATAATAATAATAATAATAATAAAATCTCAAAATAATTAAAAAAAGAAACAA
ACAAATGCTATCCTGATCCTAACTGGCTGGCTGTCTTTGGGGAAGTTGGTAATCTTTTCTGTGCTTATTT
CCTCATGTGTAAAAAAATGAATATAGTACCCAGCTAGGTAGAGTTGTTGTTGGGATTAAATGATGACTAT
AAAGCATCTAGCCCAGCTTCGGCTACATTATAGCTGCTTACGAAATTGTAGTTACGATGTAAAAGAGAAA
AACACTGGAAAAGGAGGATATGGGCCATTTTATTCCACCTTCACCACCTTTTAGCTTGGTGACCTTGGGC
AAATTATGCTTCATTCCGTGCTTCATTTTCCTTGTCTATAAAAGGGTGTAAGTACAGAACCATTGAGGGG
TGGTCATTATTAACCTACCTCAAATGGTGTCTGTAAGTTAATATATATTGTGCTTTTCCTATGTACAATA
TCTAGCACATAATTACAAATCAAATCCATCCCATGTGCAATATCTAGCACATAGGAAAAGCACAATAACT
AGTTATTACTCTTGTTGTAGTAATTGCTACGCTGTAGGAGTTTGAATTGTAAGGCAGTGGAGAGTCACTG
ACCTTTACGAGAAAGTGTAGCAGAACATTTGAGTAGATAGTAATGGGGAATATTACATAAATGGATAGAT
ATTAGGGGCAGATATTACTATTAAAATATTACAGCATGGATATTTATTAAGGCCAAACTGGTTAATTAGT
TGCATCTCTCAGGTTCCTAATGTTGCTTAATTTTTTAACCTCCCATTTTGTGCTGCCCTTTGTACGAATA
TTTAATGCTCCCAACACCTCTTCAGTAGCACATGTACTGTGAGTTTGTTTTGTTATTACTTGTGTGTATT
AGCATTCCTTTGTGAACCAAAAGCATGGAATTAGCTGTTGCCTCTAGGCTACCTAGTTTTGTAGTTTGGA
TTGAAGCCTTCACCTCAGTAACACCTATTCTGTCTACTATCTTACAGAAAACTTGTAAAATTAAGACAGA

[0377]

TCATTAATATAGCAGAAAGAGACAAAGGGCAGAGAACATTGAGATACTGGATATTGGAACCACCCAATAG
TGTTGATTTATTTATGATTATCAGTTTTTGTCTCTGCCTAGCCTCATGCCACTAAAGTCTCTGAGGCAAC
AAAGAATAAGCAATTTTGCTCACCTTATACAAATAAAACACAGAAAAAGGAATCACTAGAGAAATGGTAC
TGCAGCCTTTCTGCAGGGATTACTGCTTATTTTTAAATTACTTAAAAGGTATTGAAATTATTGTTCATAA
TGAGAAACCTGCCTAATAAAACAGAAAATTAAACTTAACACTTCCCTATAATGTAAACAGCTCGGTTAGG
AACACAACATTACAGAAACCACTTAAGAATTGATTGTACTTGTTCTTGGAGCAGAACTAGAAGCTCACCG
TTTAGAAGCTGTGCACATTTCCCTATCAAACAGTACATAAAGTTTCCATATTCCTCAGAATCGGCTTCAT
TTGTGCCATGTGTTTGCTTGGAACTATGCCACAGAAAGCAGTTCTCCCCCTCAAGCTGGGCTCCTTTCAT
GCCGCAGTGCAAGTGTGTGATATACTGGCACCATGTGCTAATGTAGACCCATTTTTATATGATAAGAATT
AGTACGGCCTAGGGAATAGACAAGTATGTCTAAAATCCTCCCCATAGAATATGTCCCTTCCTTTAAAAGC
TGTCATACTGTAAGTTCCAGCTGAGTTAAAGGCCACTGTGCTCCTATAGGGAAATATATTCTATTGTAAT
TTTTACGTTCTCCAATAACAGTCTGTTCTTTGTTTACTGAAGAGAGCTTTCATGTCATAAAATGGTGTTT
TTTGACAGAGAAGCAGAATCATTGTTTTATTATAGAAATTTGCTCTTACAACAGCAAAAATAAATAGCTC
ATCTCTTAAGCTCCTGATCAATGTCTAACACCTCCTACCCCCAGCAACACTTCACTGCAAGTATATTAAC
ACTCTATAATAGCAATTCCACTCACCTACCAAGAAATGATCTTCACAAATGATTTACAGCTAAACCAGAG
CTTAAACACATAGCACCCAATCAAGGGCAGATTTTTATCTTTTTCCCAGTCATATAAGTTCTGAGAAGAA
ATAGATTAATGTTGATCTCCCAGACAACTGCTGAGAAAATGTACAAAGGATGTTGTTTATTTTGAAGAAT
GAGACCTAGTTGTTAAGCACTTTTTCCCCTTATATGTACGTCCAAAGGTAACCATTACACCATTTTGATG
CAAATTTAGGATATATATTTATTCATACCTCTCTTCTCCATTCGGATGTTGTCTGTGTGAGTGCTCACAG
ACACATGCACACATACACACATGCACTCCTGTTTCACACTTATTTGTAAAACTCACAAGGATTTCCAAGC
CATTAATATAGCATTGTTTAAGGTGAACACATGGTTGTTCACCATCCATATGTATCTTCACTTTGTAGCA
CTCAGAATTTGGCAAAATCAGAAGGCTGAAACCTCATGGATTAAATATATTCTATATAACATATGTCTTA
ATTGCTGTTACTGTAAAGAAACCTGGACTAGCCATATTTGACTAATTTCTACCTAAGGTATTTGAATTCT
TATAAATAGATTCATTGCTTTAATCACACAAGAGTGGTTTATATGAATGTAATTATCTCCACTTTATAGC
TGAATAAACTGAGCCTGATTCTATCCCTATATGGGAAACATGAATTGAACAGCTGTGCCAATTATTTTGA
TAATTCAAATTTCACATCTACCATGTGAAGACAGCAGAAGAGGGTTAGGGGGCTTGAATTATTCTGATTA
ACTGTGTTCATGAGTGTAATCGCCTCTAGATAATCACTCATTTCTTCACTTCACTTCCCATCTACAGGTA
GTATCAGCGAATGGTAACATCTCTTGCTTTGCCTCAGTTTATTGATGGCTGCCGTTAGAAATAAAAAGCA
TGGTTTTGTTTCAGTACTTAAATGAATAATATCTTCAAAATGTTTTTAAAACGTGAAAAGGTTGAATGCA
TTTTTAACAAATGTTTTGTTAGCTTTGACTTTTATTTTTGAACAGATGAGCACAATACCCCAGTACTCCT
TTCCTAGAAATAGGAGTACTACCTGAAGACTTATTTCCCAAAGAAAAATATCAGGTCTAGTGCAGCAATA
CGTATTAAGGGCATTGAAAAGTTATATTCACAAAATGTGACATCATAACATATGTTAATACTTCTTATCA
CTGATAATAATCCTTGAAGTTGTATTTCCAGAGAGATCTCAATTTCTTCTCACACTCTGAAAGTCTCTGT
TTATCCTTTAGAGTAGGAATGTAAGAATTTAACAAAACATTCTGAATGTTTACCTTTTTTCTAAACTGAA
ACTACAATCCCTTTTTACCCCTATATAGTAAAATATAATATTACAAGTAGAATCAGCAAATTTGTTTAAT
AATTCTTTGGGACAGTTTTTACAAGCAATGGGGTTGAATTTATGTTCCTTGTGTGCAGTGGAGTTATTAT
ATTCTTCTTAAAAAGATGCATGAAGGTAAATTAGAAATGTTTTACATGTTTTCATGACAGGACATTTAAT
CAAAGAGGAGATACAAGAGGCTTTTCTTGGGTTACTGCAATTTAATTTTCCATTTCTTTCTTGGAAGGGA
CATTGGATGCAGTTGTACGAGGTTAATATTTCTAACATGCCACTTTTATTGTGGCATTCTTCTGCTCTCA
AACCTCTGATGATTTCCCATGGCCTTCAACATGATGTCCTTATTTGGGCATCTAGTGTTCTGAGCCCTCA
CATTCTGGCCCCAGCTTCCCTTCTCAACTTGATCACAGCCATCGTAATTCTGCTAACTAATTACACAGCT
GCCCTTCCCACTCTTTCTCAACCACTCTGCTGTATTCTGACTTCTACACTTTAGTTTTAAAGCTACTCCT
TGTCCTGAAATTCCTTCTCCCATTAGGTCATTATTAATTGAGTTCATCCTCTAAGTTACAGTTCATGTTG
CTGTTCCTCTATGGCACCTTCCCTGAACCTGGTCCCATATAAAATCCCATCTCTCAGAATCCCATTAGGG
TAGGATATGTGGTCTGTAGACTATTACGTTTGTCATTTACATATTGTCTTCTATTATTGGGTAACTGCGT
GTGCGAGCATGCATGGGCTGGCCTGAAGGTCACCTCCCCAACTGCATTGAAAGTTCATCACAAGTTCAAT
TATTTCACTAGAGAGTCTCTTTCATGTCATCCATGAGGCACATTCCCTACTGTGATGTGTTATGCATACA
ATTATTTCAATAAATATTTTCTAGCTCTTTGATTGACCAAAGCTTAATTACCTGTCAACTCTAGCCTCTT
GTATCTGGAATTTCTACAGTCTTTGGATAGTATCTTTAGGATGCAAAATTAGGAGGAGTATGTACCAGGC
AAACTATTACAAATAATGCCCTCAAATAGTTACATTTCACTATTCATGTCTTGTAATTTATCTTCTGCTT
TGGTATTTTAGTACACTTATGATTCAATTTGCTGTATAGATTCCTCTGAATAGGGACAAGAGAATTCGTC
TTGATAAGTGGAAGTTCGAAGGATTCCAAAATGATGTTATTCAAGGTAGAACAAGAAATTAATACTGAAA
AAATTGAGGAGTAATAATCCCCAAATATGTACATGCGTATCTCGTTTTATTGGGTTTCACTTTATTGCAC
TTTGCAGATATTTCACTTTTTTGTAAATTGAAAGTTTGTGGCAAGGCTGCATTGAGCAAGTCCGTCGGGCA
CAATTTTTCCAACAGCATGTGCTCGCTTTACGTCTCTGTGTCACGTTTTGGTAATTTGCTCAATATTTCA
AACATTATTATTATTATTATATTTGTTATGATCTGTGATCAGTGACCTTTGATGTTACTATTGTAATTGT
TTTGAGATGTCATGAACTGCACTCATATGAGATGGCAAACTTTGTGGGGTGCAGTGGCTCACACCTGTAA
TCCTAACACTTTGGGAGGCCAAGGCAGGAGGATCGTGTTAGCCCAGAAGTTTGAGACCAGTCTGGGAAAC

AAAGTGAGACCCTGTCTTTAAAATATATATGTAGAAAAATTAACTGGGCATGGTGGCACATGGCTGTAAG
GAGCCCTGCCAGCTGCATGGGAGGCTGACACAGGAGGATCACTTGAGCCCAGGAGGTCAAGGCGGCAGTA
AGCCATGTTCACTCCAGTGCCCTCCAGCCAGAATGACAGAGCAAGACCCTGTGTGGAAAAAAAAAAAAAG
ACAAACATTTTTTCAACTTTGATTTTAGATTCAGGGAGTACATGTGTAGGTTTATTACCTTGATATATTA
CATGATGCCGAGGTTTGGAGTACAAATGATACTGTCACCCAGGTACTGAGCATAGTAACCGATAGTTAGT
TTTTCAACCCTTGTTTCCCTCCCTCCCCACTCTAGTAGTCCTCCGTTTCTATTGTTGCCATCATTATGTC
CATGACAACGCACTGTTTAGCTCCACATGAGAACACGTGGTATTTGGTATTGTGTTTCTGCATTAATTCA
CTTAGAATAATGGCCTCCAGCTGCATCCATGTTGCTGCAAAGGACATGATTTTGTTCTTTTTTATGGCTG
CATAGTATTCCATGGTGTATATGCACCGTATTTTCTTTCTCCAGTCTGCCACTGATGGGCACCTAGGCTG
ACTTCATACCTTTGCTATTGTGAATAGTGCTGAAATGAGGATGAGAATACATGTGGTTTTTTAGTAAAGC
AATTTGTTTTATTTGGGCTATATGCCCAGTAATGGGATCACTAGGTTGAACGATAGTTGTGTTTTAAGTC
CTTTGAGAAATCTTCAAACTGTTTCACTGTGGCTGAACTAATTTGCATTCCCAGCAACAGTGTATCAGAG
TTCCCTCTTCTCTACAGCGTCAGCAGCATCTGTCATTTTTTTGACTTTTTAATAATAGCCAGTATGAATG
GTGTGAGACGGTATCTCGTTGTGGTTTTGATTTGCATTTCTCTGATGTTGAGTGATGTGGAGCATTTTTT
CATGTTTGTTGGCCACTTGTATGTCTTCTTTTGAGTAGTGTCTGTTTATGTCTTTTGCCCATTTTTTTTG
ATGGGGTTATTTGTTTTTGACTTGTTGAATTGTTTAAGTTCCCTATAGATTCTGAATATTAGACCTTTGT
CAGATGCATAGTTTGCAAATATATTCTCTTATTCTGTAGACTGTCTGTTTACTCTGTTGATAAATTCTTT
CACTGTGAAGAGCTCTTTAGTTTAATTAAGTTCCACTTGTCAATTTTTGGTTTTGTTATAATTGCTTTTG
AGGACTTAGTTATAAATTCTTTCCCAAGTCTGATGTCCAGAGTGGTGTTTCCTAGGCTTTCTTATAGGAT
TCTTATAATTTGAGATCTAATGTTTAAACCTTTATTCCATCTTGAGTTAATTTTTGTATATGGTGTAAGG
AGGGGGTCCAGTTGCATTCTTCTGCATATGGCTAACCAGCCATCCCAGCATCATGTCTTAAATACAGAGT
CCTTTTCCCATTACTTATTTTCATAAGATGGCAAACTTAATCAATCAATGTTTTGTGTGTTCTGGCTACT
CCACTGATCAGCCATTCCCTCATCTCTCTTCCTCTCCTTGGGCCTCCCTATTCCCTGAGACACAACAATA
TTGAAATTATGCCAGTCAGTAACCCTACAATGTCCTCTAAGTGTTCATGGGAAAAAAAAGAGTCACATGT
TTGTCACTTTAAATCAAAAGTCAGAAATGATTAAGATTGGTGAGGAAGGCATGTCAAAAGCCAAGACAGG
CTGAAAGCCAGACCTCTTGTGCCAGTTGGCCAAGTTGTGAATGCAAAGGAAACGTTCTTGAAGAAAATTA
AAAGTGCTACTCTGTTGAACACAGGAATAAGAAAGTGGAACGGCCTTATTTTTAATATGGAGAATGTCTT
AGTGGTCTGGATAGAGGATCAAAACAGCCACACCATTATCTTAAGCAAAAGGCTAATCTAAAGCAAAGGA
CTAATTCTCTGCAATTCTGTGAATGCCGAGAGAGGTGAGGAAGCTGCAGCAGAAAAGTTGGAAGCTAGCA
[0378] GAAGTATGTTCATGAGCCTTAATGAATAAGCCCTCTCTATAACATAAAAGTGCAAGGCAGAGCAACAAGT
GCTGATGGAGAAGCTGCAGCAAACTATCCAGAAGATCAAACTAACATCTAAGTTAATAAAGGTGGCAATA
CTAAACAACACATTTTCAATATAGACGAAACAGCCTTCTATTGGAAAAGGATGCCATCTAGGACTTTCAT
AGCTAGAGAGAAGTCAATGCCTGGCTTCAAAAGTTCAAAGGACAGGCTGACTGTCTTGTGAGGGGCTAAT
GCAACTGGTGACTTTCAGTTGAAGCCAGTGCTCGTTTACCATTCCAAAAATCCTAGGGCCCTTCAGGTTA
TGCTAGATCTAGTCTGCCTATGCTCTGTAAATCAAACAACAAAGACTAGATGACTGCACATCTCTTTACA
GCATGGTTTGATGAACATTTTGAGCCCTCTGTTGAGACCTACTTCTCAAAAAAATGTGTCTTTCAAAATA
TTACTGCTCTTTGACAATGTCCCTGGTCACCCAAGAGCCCTGAGGAATATGTCCAAAGAGAATGATTTTA
TTTTCATACCTGCTAACACAACATCCATTCTGCAGTCCTTGGATCAAGAAGTCACTTCAACTTTCAAGTC
TTATTACTTAAATCATACATTTCATAAAGGTATAGCTGCTATATTGTGGTGAGTCCACTGATGGATCTGG
ATAAAGTAAACTGAAAACGGAAAGTCCTCACCATTCCAGATGCCATGAAGAAAATTCATGATTGAGGGGA
GGAGGTCAAAATATCAACATTATCAGGAGTTTGGAAGTAGTTAATTCCAACACTCATCAATGACTTTGAG
GGTCCAAGACTTCAGTAGAAGAAGTCACTGCAGATGTGGTAGAAATAGCATGAGAACCAGAATGAGAAGT
GGAGTCAGAAGGTGTGACTGAATAGCTGTAATCTCAAGGTAAAACTTCAGCGGATCCAGAGTTGCTGCTT
ATGGATGAGCAAAGATTGTGGTTTCATGAGATGCACTCTACTCTTGGTGATGATGCTGTGAACATTGTTG
AAATGACAACAAAGGATTTAGAATATTCCATAAACTTAGTTGATAAATTAGCATCAGGGTTTGAGAAGAT
TGACTCAAATTTTGAAAGAAGTTCTACTGAGTAAAATGCTATCAAACAGCATCACATGCTACAAAGAAAT
CATTGGCTATAAAGGAGAGCCAATCGATGCAGCAAACTTCGTTGTTGTCTTATTTTAAGAAATTGCTACA
GCCATCCCGACCTTCAGCAACCGCCACCCTGATCAATCAGCAGCCATCCTCACTGAGGCAAGAGCCTGTA
CCAGCAAAAAGATTATGACTCCCTGCAGGCTCAGATGATTGTTACCATTTTTTTTAGGAGTAATGTATTT
TCAAATTAAGGTATGTACATTTTTTAGATACAATGCTATTGCACACTTGACAGATGATAGTAGAGTGTAA
ACATAACATTTAATGCACTGGGGAACCAAAAATATTCGTGTGCCTCATTTTATTGTGATATTTACTTTAT
TGCAGCAGTCTGGAACCAAACACACAATATTCCAAGGCAGTCCTATATATTGTTTCATGTCTCTCTTTCT
TCTAAAATGTGTGTAGGAGAAAAAATTATATCACTTATGTTACTGGGCCATAATATCAAAAGCCTGCGAA
TCTGATGCACATGATAAAAACTGGTCTTAAACCTGTCTTGATTATCCTTTGTTAATATGCCAAATTTATA
GAACAATAGAGTTTCAAGAAATGTGAACAATGTAGAATAACTAAAAGATCTAACGTTGAATAGCTAAAAT
TATCAACGCCCTTTATATCACTTTAGAAATGCGTTTGCAAATCATTATCAACAAATTGTAGCATAAAATT
CTTTTTTTTTCTTGACTTTGAAATTGTATTTCAAGATCCGCAATTTACACCACATTATTTACTTACTGGC
TTGTGAAAGTGAAAGGCATTTTCATTTTTGGATGTTAAGGGTTTTAGTAAATGACAAGTAAATCAATCCT

[0379]

TTAAGTTCCGTGTGGTATAATAGCTTGGAAAGGACCATCTTAATCTTTTTTTCAACACAGGGAGATAATT
TATTTTAAAAATAACACACTAATAGTAGATTAATATTATTACCATTTCAAAGAAGTCACCAAATTTGGTA
GGCTTAAGAGGTTTTTTTTTTTTTTTTGAGATTACTATGCTCTTTTTTTATTTTATTTATTTTATATTTA
TTTATTTATTTTTTAGTTTTTTAATTTTACTTTAAGTTCTGGGATACATGTGCAGAACGTGCAGGTTTGT
TACCTAGGTATACATGTATCATGGTGGTTTGCTATACTCATCAACCCAACATCCAGGTTTTAAGCCCCCA
ATGCATTAGGTATTTGTCCTAATGCTCTCTGTCCCCTTGCCCCTCACCCCTTGACAGGCCCCGGTGTGTG
ATGTTCCCCTCCCTGTGTCCATGTGTTCTCATTGTTCAATTCCCACTTACGAGTGAGAACATGCAGTGCT
AGGCTTAAGAGTTTTTTTAACTCCCCAAAATATCAACAAATTGAAACATTACTACAAAGAAATAGAGAAA
TAAATTTCACTGACTGTCTTATTGTTTTTTAAGTTTGAATAGCTAATAATGATTTCTTAAACAGCTATCA
TATTTTTTATTTTTAAAGCTAGCCAAATGATCAGTGATTTTTATAATACTGATAAATACTGCTTAGAAAA
GGAACATGTGTTCTAGCAATTTCCACACATTTCTGATTCTAATTACTTGTTTCTTTTTGTTTTATTTTTA
TATTTTTAAGTTTTGTGAGTACCTAGCAGGTGTATATATTTATTTGGTACTTGAGATGTTTTGATACAGG
CATGCAATGTGTAGTAAGCACATCATGTAAAATGGGGTATTCAACCCCTCAAGCATTTATTCTTTATGTT
CCAAACAATCCAATTATACTCTTTTAATTAATGTAAAATGTACAATTAAATCATTATTGGTTATAGTCCC
CCTGTTGTGCTATCAAATAGTAGGTCTTACTCATTCTTTCTAACTAATTTTTTGTACCCAGTAACTATAC
CTACACCACCCCCACCTCCCCACGACCCTTCCCATCCTCAGGTAACCATCCTTCTACTCTCTATGTCCAT
GAGTTCAATTGTTTTGAATTTTAAATCCCACAAATAAGTGAGAACATGCAATGTTTGTCTTTCTGTGTCT
GATTTATTTCACCTAGCATACTGACCTCCATTTCCAACAATGTTGTTGCAGATGACAGAATCTCATTCTT
TTTTGTGGCTGAGTAGTATTCCATTGTGCATAGGTACCACATGTTCTTTGTCCATTCATCTACTGATGGA
CACTTAGGTTGCTTCCAAATCTTGGCTATTGTGAATAGCGCTGCAATAAACATGGGAGTGCAGATATCTC
TTCAATATACTGATTTATTTTCTTTTGGGTATATACCCAGCAGCGGGAATGCTATATTATATGGTAGCTC
TATTTTTAGTTTTTTGAGGAACCTCCAAACGGTTCTCCATAGTGGTTGTGCTAATTTACATTTCCACCAA
CAGTATACAAGGGTTCCTTTTTCTTCACATACTTGTCAACATTGGTTATTGCCTGTCTTTTGGATATAAG
CCATTTTAACGGGAGTGAGGTAATATGTCATTGTAATTTTGATTTTCATTTCTCTGGTTATCAATGATTC
AGTGATGTTGAGCACCTTTTCATATGCTTGTTTGCCATTTGTATGTATTCTTGATCTGCCACAAGTCTCT
AATTACTTGTTTCTTGTACCACTGTTTCCCTTCATTGTAGTCAGTGATTTGGTCAATCAACACTTTTTAC
GTATGGGTATCTGTTAACTGTATTTCTAGGAATGGACCAAGAATTGAGAAATTTATCCCAACCAGAAAGA
ACAAAATCTATGGAGGCACAGGATTACTGAAAGCTTTGCTCCTATAGCCTCAGTTTTTTTCTGCAAGTTC
GCTGCTTCCCAGTTGTCCTTGTGATAAAATTCAAAACCTCAAACTGATTTTTAAAAAGCCAACTATATAC
TGCTTCTACTATGTCTACCTAAACTTATTTGCCATTATTTGTAAAGAATTCCAGGCTCTAACCAGCCCAT
TTGAGGCTTTATTTTACGTTCATGCCTTTGTTCATGTCTGTCTTTCCCTTATAATGCTCTTCTCATTCTC
TAGCCAAATCTGCCAAAGTTCAGCTGTAGTGCCATACTTGATATGAAGTCTTTGCTGAGAATTCCCATAT
GTGGTCATTTCCTTCACTGATTTGTTTCAATAGTTGTTGCCATTATTATTATTTTTTTTATACTGTTGCT
GGATACTTTCGCAGATCCAATCTCTAACCTGTTGTTCTACCCTAGGAGGTTGACCAATATGGCAGTAGCA
TTGGGCTTGCTTGTCTATTGGCTTGGTCCTTGTGTTGGGAGGCATCAGCAGATCAATGGATGAGAGGAGA
GTGAGTCGAGGGATGGCTAGGTTCCTTTACTCACAGCCCCAGCTCCTGTCAAGAGGTCTTGTCCCTGCAG
CCACTTCTCAGATTCTACTAACTATACCCTCCCCTTAGCCTTTCAGGCCTTGAGGAGGAAAGCCTCCTCA
TCCCACTTTGAATGAACTATTTATTGCAAGAACCATGACTGATTCAGAATATAATCTGTCCTTGAATGAA
ATAGTTAATTTCCCCTTTTGTATTCTCGTTTCATATCCCTGTCCCTTTATCCATTCCTATATTATATCAT
ACTATCATATAGCTACCTTATATTATACTATGACATGAGTACCTTGTAAGTATATCATGCTATATTATAG
TTCTTAAGTGTGTGGGCTTTGGGGCCAACTTACCTATGTTAGAAGTTCTGCCACTTATTACCATGTGATC
CTAGACAAATTGTTTAACTTCTTTTTCATTCATTTTCTTTACTAGTAAAATAGGGGCAATATTAGTACCC
ACCTCACAGAGTAATATGATAATTACATTGACTGATAGTTTTTATTTAAAACTTTTATTGTTCTTACTAT
GAAATGGGAAATGTTCTAAACACCTTACAAATATTAACTCTATTAATTCTCATAATGAATCTGAGAGGTA
GAGACTCAGTATCTCCATTTCACAGTGCTGGAAATGATACAGACAGCATTTAAGTAACATACTGAAAATA
GTAAACCGAATAGGTGGCAGCCAGTATGCAAATTGGGAAAGCCTGACTCCAGAGTTCATTCTTTTAACTA
GTATGCTGTGATAGACTCTTTTGGCATATGGTATGCACTCAATAAACAACTATTGTTGTTGTCGTTAGCA
TTACTTCCCATTGTATTATTACACAAAAAATAATTGTTTATGAGTATTTGTTTGCTAGATTATAAGTTCC
TTGAAAAAAGAAGACAAGTTTTATTTGTCTATGTGTACTAGCTGAGTGCTTGGCAAATAGTTGCAGTTTA
GTAAATGTTTCTAAAACAAATTATTAGTTGTTTCTTATGTATTTCCCAAGTCTATCCTAGCCTTGGAAAC
AGCTAACACTTAGCTAAACCTAGAAATGTCATTTGAGATTTCAGCAGCCATCATTGTTGCTGAAGCCACA
GGTTCTCCTTTTTATCTCCAATTTCTTCCTCAGATGATTCACCACTTTCTTTGTCGTTTCCCTACTTCAT
TTTTCCTATAGTGACTTTTGTTCTCCAATAACCATTGATAGTGATGACAGTCCCACCTTGGAACCACTTC
TAACTGTCCTGTTCAGCTTTTCCTGAGGCCAGATTTCCCCCAAAATACATTTTTTAGCTTCATAACTGCT
CCTTCCAGAAATTTGAAACGCATTCTCAGGAATTAAGCAAGAGGCATTTGATTGACGTCATTTAATTTGT
GTTTAAATGTGTTCTTCCCATCAACGATTAACCAGTGCTTCAGCCAAGATACATAACTTTTATTCTCCAC
TAACCTTAGGGTTGAGGTCACAATCAGGTACGTAATCTGTAGAGCCCAGTGAAAAATGAAAACGCAGGGC
CCTTGGTTAAAAAAAAATTAAGAATTTCAAGATGGCAACAGTAGAGCATGAAACCAAGTATGAATCCCTT

CTAATGCAGATCCTTGTGTAACTAACTGCACAAGTTATACGTTCAAGAAGCTGGTCCTGGTTGAGGTCTT
TGTCTACCTAGGGCACTTCTCACTCAGAGAGGTGGAATAATCTTAACTTTCTGTCTTCCGTCATCTGAAA
ACTTGCCCCAAAGTTCCTTTTGCCATCTCTTGCAGCAAACCTAGTGGTTGTTATCTGTTAAGGATTCCTG
GTGGCTTGCTTTAGTTTCCTACAGTTCTTGAGCTGCATGTGTTGAATGGAACTCCACCATTACCATGTAA
CACATCTGAATACTCTTATTTTCCTCCAGTTAAATTCTGCGCTCCTTGGAAATAATGGTCATGCACCCTA
CCTTAGTGATTTCTAAATAGAGGTAGTATGTAAAAATACATATTTACTTGAATTGCATCTTAAGAGGCCA
GCCCATGGGAGGCATGAAGAAACCTATGTTACTAATTAATATTCAATAATTGACAGTTACACATTCAGGT
AGTAGTGTCATGGAACCATCACGTTACATGGAACTGAAGAGCTGATGAGAGGTTTAGGGTCTAGGAGGAC
AAAAGTGAGTGTATTAGTCAGGGCCTGAGGTTATATGTACCCAGACAGGATAAAATGGGACAATATTTTA
TTAGTGGAAATATCTATGTGAAAGAAAAGAAGGAGGAAGCCAAGGAAGGTGCAAGAGATGTTAGATCAAG
ATGCAATTCTGACTCTGCAAGGAAGAGAGAGGGAGAGAAGGCTAGGCAGAAGCATCCCAGTGTGCGGTCT
AGTGGAAGGAAATTTTGGCAAAGCTGTTGGGAAGTCATTGAGGCAGAGCCAGGCAAAGAAGTCCCATGTC
TCCCAAGGAAGGCTCTCTGCCTTAGTATTCCCACCACACCCAATCATTGGGTGAGGAGAAGTCTGTGAGA
AGCTTGGCTTTGGTGCAGTGCAATCATGGATTTCAAAATGCAGTAACAGGAGTCCTCAGTCAGTTAAGAC
CCAATAATAGAAGGCCTGCATATTCTCATGGTGGCCACTTGGGTATGAGGAGCAGTGGTGTTTAAAACTG
TTTTGTATAATTTAAATAAAGAAAAGCTGAGGACTACTTAAGCTTGATTCCTTCAGAAGACAGTCTTTGG
CCTTTATATTTCATGTGATATCTTTGCCATATGCTCTGATAACTCTCCATGTTTCCCCTACCATAATACC
TCAGGTTGTTGCAATTGCTTTTTTATTGCCTGTCATCACTACTCGTTTGTTTTCTCTTTGAGAACAGGAA
TTATTTTCTCCTTCACTGCTACAGCCCCTGCACCTAGCTCAATGAGTGACACAACAGAAGCACTTAAAAA
ATTGCTATAACTCAAATTTGGGGGATATTCTACAAAATACCTGGCCTCTGTTGGTCAAAACTGTAAGTGT
TATAAAGTTCAAAGAAAGGCAAATAACTATTTCAACTTAAAAAAGACTAAGGAGATACGACAACTAAAAG
CAATGCGTGAGTCTGGACCAGACCAGGAAATAAAAATATAGCTATAAAGTTCATTAATGGGGTAATTGTC
ATACTCTGAATATAGAGTATGGATAAGATTATGGTACTAAAGCCACGCTAAATTTCCTGACTTTGAAACT
GAGCTGAGGTTATTTAAGAGAATGTCTTTGCTATTAAGAAATAAAACCTAAAGTATTTAAGGGTAAAGGG
GCATGATATCTGCAAATTATTCTCAAATGATTCAGAAAAAAAATATATATATATAAAACATTACATATAC
AGTTATATATCTATACACACAGAGAGAATGATAAAATGATAAATAATTAAAAATAATGTGGCAAAATGTT
AACAAATGGTGAATCCGGGCAAGGGTTACCTGGGAGTTCTTTATCTTGCAAATATTCTGCAAGCTTGAAA
TTATATAAAAATAAAATGCATCTTATAAAGTATTTATTCAGTGAATAAAGAACAAAAAAGGCTAACTGCA
GTTGGAAGATATTTATGAAGTTGGTTATGAAGATTCTTAAGAAGTTACCGATGGAGGTGCTAGTAGAACA
[0380] TTCAAAAAAGAAGTAGGGAAGGTTGACCAAGTAGGAAAATATCAATATCCAAGCCAGTATAAAATGGAAA
TGAGAAGTGAGGAAGTAAGTGGAAGACAATGAGGATGGTTTTCGATTTGGCCACTCTATTTGGAGAGGCA
GCCGAATGCAAGAATAGAAGCCAGAAGAAGATGCTCAGTGAGATGGTTGAAAGCTAGATAGATTACAGCA
TCCTCACCAGTAAAACCCTTTCGGTAACTAGAAAGGCTACAATTTAGTACCTTCCTGACTTCTATGCTTA
TTTTCTTCAATACATAAAATGGTTCCGTAAACTCTTTTACCTTCTGAATTCTTTATATTAATTTTTTGAA
GTTGTAAATAAAATAGCATCAGTTCTACATTGTTACATTTCAGCTTAATTCATATTCATTTACTGAAAAT
GGGAACATTTGAAAAATCATCATGGGCATTTATGCTATGTAGATTGTTGATTTTTATAGAAAAATATAAA
AATATGACCAGTTTGATTTTCAAAGTCTTTTCTTAGACATGTAAATACTAAGCATTCAACTCAACATATA
GAGTTTTTATTTGAGTATTATTTAGGTGGAATTCTATTTTAATGAATACAATAAAAAATTGTAATTTTGT
CTAAAAGCCTAAAATGCCCTAGTTATAATATGTATGATTTCACTGTTTAACTTCCTATTTCATAGGGTTG
CTATTTATAACCACTTCACTCAACTCTGGGGGGACTTAGTGAGATTAAAGACTTCTGATTCACTTTGTAT
TTGAAGAATTTTTTTTCCTCCATCTTTGCTCAGCTAGTGGAATCCATGATGAATTCTCATCTCCAAGGGG
TAAGCAGTTTTTAGTAAAGCCCAGTAGCTGACTTATGACTCCTTAGAAATAGCATTGATTCCTTCCTTCT
CCTGTGTTTTGTTTCCTCTAGAATGATAGAATCCATGTAGACACGATCCATTATCATGCTTAGGTACTGG
TAAGCATGTAATGATTTTAGTTTTGTTCGCTTTAAGTTATTTGTGTCACAAATATCTGGGATCATATCAG
AGAAATAAATAAGCACAATTAGCATTCTACTTGTTTGTTATGACTAAAGCTAGGTTGAGGAAACAGAAAA
GGACCAGAGGTCATATGAGGATGAAGATAATACTAGGAACAGCATGTTTGGGAGAGTAACATCTGGTAGG
GGTAGCAGATTGGGGGCAGAGAACAGAATTTTATAGATGGATATTTTGGAGGCAAGTAGTTTGAGTAATG
ATTAGATCTAAGGTGTTTTCTCATCTGTGGGTGGCTCGAAGGAATAGAGGTGAAGGTCAGTTTATTTGAG
AAGTTCTGGAATTATAAAACTAAGTTGAAGTCAAAGAAAGTATAGTAGCAAATAAATAGAATACCCTTAA
AAGGAAACCAAATGAAAAATAATCGTTACTCTCACCATATGCTTGTGTTCTTATTAGCAAGAAATTCTTT
TAACCACTGTTTTTATAATATCTTAATGAAAAAATACTGAAGCGTATGCCATATTAAATCCCTCTCTTTA
TTTCTAGAAAGGGAATCAAAGGAGAAAATTCCCATTCTGCTATACTAAAAGACCACTAAGTAAAGAGCCT
ATTAGTGTATGATAAATCCCATAGCAATATACATTATCATTTTACAGCTTCTTTGTTGAAATGAATGTTT
GTATGTGTTGACCATAGAGTGGGATAAAAAGTTGAAATTTTGTTTTGAAATATTTTAGAAATGCATAGTT
GTACTGCAGTTGTGAACCTCCTTAGATTTTTAAGGAGGCTGCTTCAAAGGATCTCATTAATAATCTTCTC
AGGTGCTTACAAAGCATGTGTCTGTCAGCAGAATTAGAGAATCACCCAACTAGAGAACAGGTTTCACAAT
ACCCTGAGACCTATTTTGTTCATTAGAGAGGAAAATGGCTTGTTTTGAGTCTAAGTTGACATGCTTGCTA
ATTTCAGCAATAAAAGCTGTTCATTGTGGTCAGGTTTAATTTAGAGCCTGGTAAGGTTCAGATTAAAGTT

[0381] GATCAACTTACTTTTACAACATACTTCTTAAATGAACTTTGAAATCTTAAAAGAAGGAAAAAAGTATAGC
AAACAGTGAATAATGTATCTAAAACTGAGAAGCAAAAAAAATCTGGTTATGTGAGAGTGAATTAAAAGAA
GAACAACCCAATAAAGATAATCTTTGTTATATAAAAATTTCCAAGTATCGAAAAGCACGATTTTTCATGT
GAGTTACACACTATACCGAATATATTTGTCCACTGCCACATGCATAGTCCCTAGAATAGTGCCTAGTGCT
GAAAAATATTTATTAAAATGAATGGATGAGTAAATGAATTTATGTATTTTGCCAGCCCTGTGTATTTAAA
GTTCTCTGTTAACTTTGAGGTGAAAATTTGACTTCATCTGAGGTTTCTGGGTAAGTCCGTTTTAAAAATT
CTATTGAACATATTCAAACATTTTAGGGTAGGCAATTCCAAAGCAACCTTTCAGCTTCCCATGTCACAGA
TGACCAGAGTTTCACATTCTAACACTGGAAAACATCTTATTTCATAAAATCTACCTGCTACTATATGGTT
CCTACTTTAAAATTTGTTCAGTACTCTCCAGCTGACTTATGCCACTTACTTCAATAGCTGTCTTTGGCAA
TTTGTTCCATATTTCAAACACTCTGTTGTGTAAAGAAACCATAAAAGTTAGAAACCTGAAAATTGGATTT
TTTTTCTGAGCAATCACAGCTTACAAATGTGGAAAATTTGTTAAAAGTTAGCCCCTCCAATTTTTCAATA
CAGAGAGGAGAAAGTGCCTAAAGTAGATGTACAATGTTTGGAAAAGTTTTTTGCATTATTTTACTATTAC
CAAAAGCAATTGAGTTTAAATCACAAAGCCTGTCTCCCTACCTCTTCACAGAAGAAACACTACAGAATGA
TCAAAATTTGGCCTTTCCAAAACCAAAATTCGTTAGAAAATCAGCAGGAGTCAAAGTACAGAGTAAATAA
CTAAGTTTCATATAAGTTTCAGATACATTACTATCTACCACTTTCATCTCTTCATCTTCATTGGCCTCAT
GTGGTAGAACATCATATTTAAAATTATACAAACTTGCTGGCTTGTTTACTAGTGTGGTTATTATAAGAAA
AAAATGAGAAAATATAGATAAAACATCTGTCACATATTGCTTATAAACTAACAGTAAATATTACTTGTAT
TTTCCCCAATTAAAATAAAATTTACTGAGTTTTAGAACCAGAGCTAGTCAGATGCCTTTTTTTCATAAAT
TTCTTCATAAATACCTCTGAGATTGTGGTCCTTAAATCTAGAGAGACTAAGATGACAGAGAAAATAGACA
CTGAAGAAAGGGAAGAATATCTTAATGATTTACATTACTACCTATAAATTAAAAATTGTTAACTTTATTA
TATTTGTATTTTTATTTAAAATAGTGCTATATTAAAGTCATTTATAATACAGGGGAATAGGAATACTAAC
CTGTAATCTGATGCTCTCCAAACTTGCCTAAATCATAAAAGTTAATTAGATAATTTATTTAAAATGCAAG
ATTTGCAGCCCTTTCCACTACATATTCATTAGTTCTGGAGGGAGGCAAAAAGGTTTGGTGATTCTTACGG
ACAGGCAGCTTAGGAGAAAAGCTGATTTAGCTCGTCTACTTCACCTTTTCATTTGACAGGTGAGAAATCT
CAGGGGTACAATGAAGTTAAATAAGTAATATCTCTTAAATCGGTTCTGTGCTTTTTCTGTTTTTAAAATA
AATATACCTTAATTTTGACGTCACACAGAATGATATTATAAGTATAAATAGTTATCTATCTTTTAAATAC
ATTGTCGTAATTCAGAATAACATTTCTTACTCAAGGCATTCAGACAGTGGTTTAAGTAATCCGAGGTACT
CCGGAATGTCTCCATTTGAGCCTTTAAATGAAGAAAATCTATAGTCAAGATTTTCATTTGAAATATTTTT
GATATCTAAGAATGAAACATATTTCCTGTTAAATTGTTTTCTATAAACCCTTATACAGTAACATCTTTTT
TATTTCTAAAAGTGTTTTGGCTGGTCTCACAATTGTACTTTACTTTGTATTATGTAAAAGGAATACACAA
CGCTGAAGAACCCTGATACTAAGGGATATTTGTTCTTACAG (SEQ ID NO: 846)

[0382] 智人肌养蛋白(DMD),内含子51靶序列1(NCBI参考序列:NG_012232.1的第1570651至1570700位核苷酸)

[0383] GTATGAGAAAAAATGATAAAAGTTGGCAGAAGTTTTTCTTTAAAATGAAG (SEQ ID NO: 847)

[0384] 智人肌养蛋白(DMD),内含子51靶序列2(NCBI参考序列:NG_012232.1的第1570651至1570693位核苷酸)

[0385] GTATGAGAAAAAATGATAAAAGTTGGCAGAAGTTTTTCTTTAA (SEQ ID NO: 848)

[0386] 智人肌养蛋白(DMD),内含子51靶序列3(NCBI参考序列:NG_012232.1的第1570703至1570765位核苷酸)

[0387] TTTCCACCAATCACTTTACTCTCCTAGACCATTTCCCACCAGTTCTTAGGCAACTGTTTCTCT (SEQ ID NO: 849)

[0388] 智人肌养蛋白(DMD),内含子51靶序列4(NCBI参考序列:NG_012232.1的第1614751至1614804位核苷酸)

[0389] TATTTCTAAAAGTGTTTTGGCTGGTCTCACAATTGTACTTTACTTTGTATTATG (SEQ ID NO: 850)

[0390] 智人肌养蛋白(DMD),内含子51靶序列5(NCBI参考序列:NG_012232.1的第1614793至1614847位核苷酸)

[0391] CTTTGTATTATGTAAAAGGAATACACAACGCTGAAGAACCCTGATACTAAGGGAT (SEQ ID NO: 851)

[0392] 智人肌养蛋白(DMD),内含子51靶序列6(NCBI参考序列:NG_012232.1的第1614612至1614861位核苷酸)

[0393] CGGAATGTCTCCATTTGAGCCTTTAAATGAAGAAAATCTATAGTCAAGATTTTCATTTGAAATATTTTTGATATCTA AGAATGAAACATATTTCCTGTTAAATTGTTTTCTATAAACCCTTATACAGTAACATCTTTTTTATTTCTAAAAGTGT TTTGGCTGGTCTCACAATTGTACTTTACTTTGTATTATGTAAAAGGAATACACAACGCTGAAGAACCCTGATACTAA GGGATATTTGTTCTTACAG (SEQ ID NO: 852)。

[0394] 智人肌养蛋白(DMD),内含子51/外显子52连接(NCBI参考序列:NG_012232.1的第1614832至1614891位核苷酸)

[0395] CCTGATACTAAGGGATATTTGTTCTTACAGGCAACAATGCAGGATTTGGAACAGAGGCGT (SEQ ID NO: 853)。

[0396] 智人肌养蛋白(DMD),转录物变体Dp427m,外显子52(NCBI参考序列:NM_004006.2的第7787至7904位核苷酸,NCBI参考序列:NG_012232.1的第1614862至1614979位核苷酸)

[0397] GCAACAATGCAGGATTTGGAACAGAGGCGTCCCCAGTTGGAAGAACTCATTACCGCTGCCCAAAATTTGAAAAACAA GACCAGCAATCAAGAGGCTAGAACAATCATTACGGATCGAA (SEQ ID NO: 854)。

[0398] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD序列(例如,DMD前体mRNA)中的剪接特征。在一些实施方案中,DMD序列中的剪接特征是DMD序列中的外显子剪接增强子(ESE)、分支点、剪接供体位点或剪接受体位点。在一些实施方案中,ESE在DMD序列(例如,DMD前体mRNA)的外显子51中。在一些实施方案中,分支点在DMD序列(例如,DMD前体mRNA)的内含子50或内含子51中。在一些实施方案中,剪接供体位点跨越DMD序列(例如,DMD前体mRNA)的外显子50与内含子50的连接处、在其内含子50中、跨越其外显子51与内含子51的连接处、或者在其内含子51中。在一些实施方案中,剪接受体位点在DMD序列(例如,DMD前体mRNA)的内含子50中、跨越其内含子50与外显子51的连接处、在其内含子51中、或跨越其内含子51与外显子52的连接处。在一些实施方案中,可用于靶向DMD的寡核苷酸促进外显子51的跳读,例如通过靶向DMD序列(例如,DMD前体mRNA)中的剪接特征(例如,ESE、分支点、剪接供体位点或剪接受体位点)。ESE、分支点、剪接供体位点和剪接受体位点的一些实例在表9中提供。

[0399] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD序列中的外显子剪接增强子(ESE)。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD外显子51中的ESE(例如,表9中列出的ESE)。

[0400] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸包含靶序列的互补区,所述靶序列包含DMD转录物的一个或更多个完整或部分ESE(例如,表9中列出的一个或更多个完整或部分ESE)。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含DMD外显子51的一个或更多个完整或部分ESE。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:860至894中所示的一个或更多个完整或部分ESE。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:860至894中任一者所示的ESE的至少4个(例如,4、5、6、7或8个)连续核苷酸。在一些实施方案中,寡核苷酸包含如SEQ ID NO:904至938中任一者所示的ESE反义序列的至少4个(例如,4、5、6、7或8个)连续核苷酸。

[0401] 在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含DMD外显子51的一个或更多个ESE(例如,2、3、4个或更多个相邻ESE)的至少6个(例如,6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个)核苷酸。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:860至894中所示的一个或更多个ESE(例如,2、3、4个或更多个相邻ESE)的至少6个(例如,6、7、8、9、10、11、12、13、14、15、16、17、18、19、20

个或更多个)核苷酸。在一些实施方案中,寡核苷酸包含如SEQ ID NO:904至938中所示的一个或更多个ESE反义序列(例如,2、3、4个或更多个相邻ESE的反义序列)的至少6个(例如,6、7、8、9、10、11、12、13、14、15、16、17、18、19、20个或更多个)核苷酸。

[0402] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为18至35个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:860至894中任一者所示的ESE的至少4个(例如,4、5、6、7或8个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为20至30(例如,20、25、30)个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:860至894中任一者所示的ESE的至少4个(例如,4、5、6、7或8个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为20至25(即20、21、22、23、24或25)个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:860至894中任一者所示的ESE的至少4个(例如,4、5、6、7或8个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸的长度为30个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:860至894中任一者所示的ESE的至少4个(例如,4、5、6、7或8个)连续核苷酸。

[0403] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD序列中的分支点。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD内含子50或内含子51中的分支点(例如,表9中列出的分支点)。

[0404] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸包含靶序列的互补区,所述靶序列包含DMD转录物的完整或部分分支点(例如,表9中列出的完整或部分分支点)。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含DMD内含子50或内含子51的完整或部分分支点。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:856至858、896和897中任一者所示的完整或部分分支点。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:856至858、896和897中任一者所示的分支点的至少4个(例如,4、5、6或7个)连续核苷酸。在一些实施方案中,寡核苷酸包含如SEQ ID NO:900至902、940和941中任一者所示的分支点反义序列的至少4个(例如,4、5、6或7个)连续核苷酸。

[0405] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为18至35个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:856至858、896和897中任一者所示的分支点的至少4个(例如,4、5、6或7个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为20至30(例如,20、25、30)个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:856至858、896和897中任一者所示的分支点的至少4个(例如,4、5、6或7个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为20至25(即20、21、22、23、24或25)个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:856至858、896和897中任一者所示的分支点的至少4个(例如,4、5、6或7个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸的长度为30个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:856至858、896和897中任一者所示的分支点的至少4个(例如,4、5、6

或7个)连续核苷酸。

[0406] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD序列中的剪接供体位点。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向跨越外显子50与内含子50的连接处、在内含子50中、跨越外显子51与内含子51的连接处、或在内含子51中的剪接供体位点(例如,表9中列出的剪接供体位点)。

[0407] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸包含靶序列的互补区,所述靶序列包含DMD转录物的完整或部分剪接供体位点(例如,表9中列出的完整或部分剪接供体位点)。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含跨越DMD的外显子50与内含子50的连接处、在其内含子50中、跨越其外显子51与内含子51的连接处、或在其内含子51中的完整或部分剪接供体位点。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:855或895所示的完整或部分剪接供体位点。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:855或895所示的剪接供体位点的至少4个(例如,4、5、6或7个)连续核苷酸。在一些实施方案中,寡核苷酸包含如SEQ ID NO:899或939所示的剪接供体位点反义序列的至少4个(例如,4、5、6或7个)连续核苷酸。

[0408] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为18至35个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:855或895所示的剪接供体位点的至少4个(例如,4、5、6或7个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为20至30(例如,20、25、30)个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:855或895所示的剪接供体位点的至少4个(例如,4、5、6或7个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为20至25(即20、21、22、23、24或25)个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:855或895所示的剪接供体位点的至少4个(例如,4、5、6或7个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸的长度为30个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:855或895所示的剪接供体位点的至少4个(例如,4、5、6或7个)连续核苷酸。

[0409] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向DMD序列中的剪接受体位点。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸靶向在内含子50中、跨越内含子50与外显子51的连接处、在内含子51中、或跨越内含子51与外显子52的连接处的剪接受体位点(例如,表9中列出的剪接受体位点)。

[0410] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸包含靶序列的互补区,所述靶序列包含DMD转录物的完整或部分剪接受体位点(例如,表9中列出的完整或部分剪接受体位点)。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含在DMD的内含子50中、跨越其内含子50与外显子51的连接处、在其内含子51中、或跨越其内含子51与外显子52的连接处的完整或部分剪接受体位点。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:859或898所示的完整或部分剪接受体位点。在一些实施方案中,寡核苷酸包含靶序列的互补区,所述靶序列包含如SEQ ID NO:859或898所示的剪接受体位点的至少4个(例如,4、5、6、7、8或9

个)连续核苷酸。在一些实施方案中,寡核苷酸包含如SEQ ID NO:903或942所示的剪接受体位点反义序列的至少4个(例如,4、5、6、7、8或9个)连续核苷酸。

[0411] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为18至35个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:859或898所示的剪接受体位点的至少4个(例如,4、5、6、7、8或9个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为20至30(例如,20、25、30)个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:859或898所示的剪接受体位点的至少4个(例如,4、5、6、7、8或9个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读,例如用于跳读外显子51)的寡核苷酸的长度为20至25(即20、21、22、23、24或25)个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:859或898所示的剪接受体位点的至少4个(例如,4、5、6、7或8个)连续核苷酸。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸的长度为30个核苷酸,并且包含靶序列的互补区,所述靶序列包含如SEQ ID NO:859或898所示的剪接受体位点的至少4个(例如,4、5、6、7、8或9个)连续核苷酸。

[0412] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA的外显子与内含子的连接(例如,由SEQ ID NO:832、833、837、844、845和853提供的外显子/内含子连接中任一者)的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA的外显子与内含子的连接(例如,由SEQ ID NO:832、833、837、844、845和853提供的外显子/内含子连接中任一者)的至少10个(例如,10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25个或更多个)连续核苷的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸与SEQ ID NO:832、833、837、844、845和853中任一者互补。

[0413] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA的靶序列(例如,由SEQ ID NO:833、835至837、845、847至853和839至843中任一者提供的靶序列)的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA的靶序列(例如,由SEQ ID NO:833、835至837、845、847至853和839至843中任一者提供的靶序列)的至少10个(例如,10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25个或更多个)连续核苷的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸与SEQ ID NO:833、835至837、845、847至853和839至843中任一者互补。

[0414] 表9.靶序列基序实例

[0415]

在DMD中的位置	类型	SEQ ID NO:	基序序列 †	SEQ ID NO:	基序反义序列 †
跨越外显子50/内含子50连接处	剪接供体	855	CTGTAAG	899	CTTACAG
内含子50	分支点	856	TATTAAT	900	ATTAATA
内含子50	分支点	857	TTGAC	901	GTCAA
内含子50	分支点	858	TTCTAAT	902	ATTAGAA
跨越内含子50/外显子51连接处	剪接受体	859	TATTTTAGC	903	GCTAAAATA
外显子51	ESE	860	CTACTCA	904	TGAGTAG
外显子51	ESE	861	CAGACTG	905	CAGTCTG
外显子51	ESE	862	GACTGTTA	906	TAACAGTC
外显子51	ESE	863	TTACTCT	907	AGAGTAA
外显子51	ESE	864	ACTCTGG	908	CCAGAGT
外显子51	ESE	865	CTCTGGT	909	ACCAGAG
外显子51	ESE	866	TGACACA	910	TGTGTCA
外显子51	ESE	867	GACACAA	911	TTGTGTC
外显子51	ESE	868	ACACAAC	912	GTTGTGT
外显子51	ESE	869	AACCTGTG	913	CACAGGTT
外显子51	ESE	870	CCTGTGG	914	CCACAGG
外显子51	ESE	871	CTGTGGT	915	ACCACAG
外显子51	ESE	872	GGTTACTA	916	TAGTAACC
外显子51	ESE	873	TTACTAA	917	TTAGTAA
外显子51	ESE	874	CTAAGGA	918	TCCTTAG
外显子51	ESE	875	CTGCCAT	919	ATGGCAG
外显子51	ESE	876	CAAACTA	920	TAGTTTG
外显子51	ESE	877	TGGAGGT	921	ACCTCCA
外显子51	ESE	878	GGTACCTG	922	CAGGTACC

在DMD中的位置	类型	SEQ ID NO:	基序序列 †	SEQ ID NO:	基序反义序列 †
外显子51	ESE	879	GATTTCAA	923	TTGAAATC
外显子51	ESE	880	TTTCAAC	924	GTTGAAA
外显子51	ESE	881	TCAACCG	925	CGGTTGA
外显子51	ESE	882	CAACCGG	926	CCGGTTG
外显子51	ESE	883	GGACAGAA	927	TTCTGTCC
外显子51	ESE	884	CCGACTG	928	CAGTCGG
外显子51	ESE	885	CGACTGG	929	CCAGTCG
外显子51	ESE	886	TTTCTCTG	930	CAGAGAAA
外显子51	ESE	887	TCTCTGC	931	GCAGAGA
外显子51	ESE	888	TCACAGA	932	TCTGTGA
外显子51	ESE	889	CACAGA	933	TCTGTG
外显子51	ESE	890	ACAGAGG	934	CCTCTGT
外显子51	ESE	891	CAGAGGG	935	CCCTCTG
外显子51	ESE	892	CTTGAGG	936	CCTCAAG
外显子51	ESE	893	TGAGGA	937	TCCTCA
外显子51	ESE	894	GAGATGA	938	TCATCTC
跨越外显子51/内含子51连接处	剪接供体	895	AGGTATG	939	CATACCT
内含子51	分支点	896	TTTTGAT	940	ATCAAAA
内含子51	分支点	897	CCCTGAT	941	ATCAGGG
跨越内含子51/外显子52连接处	剪接受体	898	TCTTACAGG	942	CCTGTAAGA
跨越外显子50/内含子50连接处	剪接供体	855	CTGTAAG	899	CTTACAG

[0416] †表9中提供的任一个序列中的每个胸腺嘧啶碱基(T)可独立且任选地用尿嘧啶碱基(U)替换。表9中列出的基序序列和反义序列包含T,但是考虑了RNA和/或DNA中的基序序列的结合。

[0417] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的任一种寡核苷酸是磷酸二酰胺吗啉代寡聚物(PMO)。

[0418] 在一些实施方案中,寡核苷酸可具有突变体DMD等位基因的互补区,所述突变体DMD等位基因例如在人DMD的外显子1至79中的任一者中具有至少一个突变的DMD等位基因,所述突变导致移码和不正确的RNA剪接/加工。

[0419] 在一些实施方案中,寡核苷酸中的任一种可以是盐形式,例如,作为钠盐、钾盐或镁盐。

[0420] 在一些实施方案中,本文中所述的任一种寡核苷酸的5'或3'核苷(例如,末端核苷)与胺基缀合,任选地通过间隔基(spacer)。在一些实施方案中,间隔基包含脂族部分。在一些实施方案中,间隔基包含聚乙二醇部分。在一些实施方案中,磷酸二酯键联存在于间隔基与寡核苷酸的5'或3'核苷之间。在一些实施方案中,本文中所述的任何寡核苷酸的5'或3'核苷(例如,末端核苷)与间隔基缀合,所述间隔基是经取代或未经取代的脂族、经取代或未经取代的杂脂族、经取代或未经取代的亚碳环基、经取代或未经取代的亚杂环基、经取代或未经取代的亚芳基、经取代或未经取代的亚杂芳基、-O-,-N(R^A)-,-S-,-C(=O)-,-C(=O)O-,-C(=O)NR^A-,-NR^AC(=O)-,-NR^AC(=O)R^A-,-C(=O)R^A-,-NR^AC(=O)O-,-NR^AC(=O)N(R^A)-,-OC(=O)-,-OC(=O)O-,-OC(=O)N(R^A)-,-S(O)$_2$$NR^A$-,-$NR^A$S(O)$_2$-,或其组合;每个$R^A$独立地是氢或者经取代或未经取代的烷基。在某些实施方案中,间隔基是经取代或未经取代的亚烷基、经取代或未经取代的亚杂环基、经取代或未经取代的亚杂芳基、-O-、-N(R^A)-或-C(=O)N(R^A)$_2$,或其组合。

[0421] 在一些实施方案中,本文中所述的任一种寡核苷酸的5'或3'核苷与式-NH_2-$(CH_2)_n$-的化合物缀合,其中n为1至12的整数。在一些实施方案中,n是6、7、8、9、10、11或12。在一些实施方案中,磷酸二酯键联存在于在式-NH_2-$(CH_2)_n$-的化合物与寡核苷酸的5'或3'核苷之间。在一些实施方案中,式-NH_2-$(CH_2)_6$-的化合物通过6-氨基-1-己醇(NH_2-$(CH_2)_6$-OH)与寡核苷酸的5'磷酸之间的反应与寡核苷酸缀合。

[0422] 在一些实施方案中,寡核苷酸例如通过胺基与靶向剂,例如肌肉靶向剂(例如抗TfR1抗体)缀合。

[0423] a.寡核苷酸大小/序列

[0424] 寡核苷酸可具有多种不同的长度,例如,取决于格式。在一些实施方案中,寡核苷酸的长度为7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、75个或更多个核苷酸。在一些实施方案中,寡核苷酸的长度为8至50个核苷酸、长度为8至40个核苷酸、长度为8至30个核苷酸、长度为10至15个核苷酸、长度为10至20个核苷酸、长度为15至25个核苷酸、长度为21至23个核苷酸、长度为20至25个核苷酸等。

[0425] 在一些实施方案中,当用于本公开内容目的的寡核苷酸的核酸序列可与靶核酸特异性杂交时,其与靶核酸"互补"。在一些实施方案中,与靶核酸(例如,mRNA或前体mRNA分子)杂交的寡核苷酸导致对靶标的活性或表达的调节(例如,mRNA翻译降低、前体mRNA剪接改变、外显子跳读、靶mRNA降解等)。在一些实施方案中,寡核苷酸的核酸序列与其靶核酸具有足够程度的互补性,使得其在期望避免非特异性结合的条件下,例如在生理条件下,不与非靶序列杂交。因此,在一些实施方案中,寡核苷酸可与靶核酸的连续核苷酸至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少

97％、至少98％、至少99％或100％互补。在一些实施方案中,互补核苷酸序列不需要与其靶标100％互补来与靶核酸可特异性地杂交或对靶核酸具有特异性。在某些实施方案中,寡核苷酸包含一个或更多个相对于靶核酸错配的核碱基。在某些实施方案中,与靶标相关的活性因这样的错配而降低,但与非靶标相关的活性以更大的量降低(即,对靶核酸的选择性提高并且脱靶效应降低)。

[0426] 在一些实施方案中,寡核苷酸包含靶核酸的互补区,所述互补区的长度在8至15、8至30、8至40、或10至50、或5至50、15至20、20至25、或5至40个核苷酸的范围内。在一些实施方案中,寡核苷酸与靶核酸的互补区的长度为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49或50个核苷酸。在一些实施方案中,互补区与靶核酸的至少8个连续核苷酸互补。在一些实施方案中,与靶核酸的连续核苷酸部分相比,寡核苷酸可包含1、2或3个碱基错配。在一些实施方案中,寡核苷酸可在15个碱基上具有多至3个错配,或在10个碱基上具有多至2个错配。

[0427] 在一些实施方案中,寡核苷酸与本文中所述的任一种寡核苷酸(例如,表8中列出的寡核苷酸)的靶序列互补(例如,至少85％、至少90％、至少95％或100％互补)。在一些实施方案中,寡核苷酸与由SEQ ID NO:384至831提供的任一种寡核苷酸的靶序列互补(例如,至少85％、至少90％、至少95％或100％互补)。在一些实施方案中,这样的靶序列与表8中列出的寡核苷酸100％互补。在一些实施方案中,这样的靶序列与由SEQ ID NO:384至831提供的寡核苷酸100％互补。在一些实施方案中,寡核苷酸与本文中提供的靶序列(例如,表8中列出的靶序列)互补(例如,至少85％、至少90％、至少95％或100％互补)。在一些实施方案中,寡核苷酸与SEQ ID NO:160至383中任一者互补(例如,至少85％、至少90％、至少95％或100％互补)。

[0428] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA的靶序列(例如,由SEQ ID NO:160至383中任一者提供的靶序列)的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含DMD RNA的靶序列(例如,由SEQ ID NO:160至383中任一者提供的靶序列)的至少8个(例如,8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25个或更多个)连续核苷的互补区。在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸与SEQ ID NO:160至383中任一者互补。

[0429] 在一些实施方案中,可用于靶向DMD(例如,用于外显子跳读)的寡核苷酸包含含有本文中提供的DMD靶向序列(例如,表8中列出的反义序列)的至少8个(例如,8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25个或更多个)连续核碱基的序列。在一些实施方案中,寡核苷酸包含含有SEQ ID NO:384至831中任一者的至少8个(例如,8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25个或更多个)连续核碱基的序列。在一些实施方案中,寡核苷酸包含SEQ ID NO:384至831中任一者的序列。

[0430] 在一些实施方案中,应理解在C5位置处的核碱基尿嘧啶的甲基化形成胸腺嘧啶。因此,在一些实施方案中,具有C5甲基化尿嘧啶(或5-甲基-尿嘧啶)的核苷酸或核苷可以等同地鉴定为胸腺嘧啶核苷酸或核苷。

[0431] 在一些实施方案中,本文中提供的任一种寡核苷酸(例如,表8中列出的寡核苷酸)中的任一个或更多个胸腺嘧啶碱基(T)可独立且任选地是尿嘧啶碱基(U),和/或本文中提

供的寡核苷酸中的任一个或更多个U可独立且任选地是T。在一些实施方案中,由SEQ ID NO:608至831提供的任一种寡核苷酸中或与SEQ ID NO:160至383中任一者互补的寡核苷酸中的任一个或更多个胸腺嘧啶碱基(T)可任选地是尿嘧啶碱基(U),和/或寡核苷酸中的任一个或更多个U可任选地是T。在一些实施方案中,由SEQ ID NO:384至607提供的任一种寡核苷酸中或与SEQ ID NO:160至383中任一者互补的寡核苷酸中的任一个或更多个尿嘧啶碱基(U)可任选地是胸腺嘧啶碱基(T),和/或寡核苷酸中的任一个或更多个T可任选地是U。

[0432] b.寡核苷酸修饰:

[0433] 本文中所述的寡核苷酸可被修饰,例如,包含经修饰糖部分、经修饰核苷间键联、经修饰核苷酸或核苷和/或者(例如,和)其组合。另外,在一些实施方案中,寡核苷酸可表现出以下特性中的一种或更多种:不介导选择性剪接;不是免疫刺激性的;具有核酸酶抗性;与未经修饰寡核苷酸相比具有提高的细胞摄取;对细胞或哺乳动物无毒;提高了在细胞中内部排出内体;使TLR刺激最小化;或避免模式识别受体。本文中所述的寡核苷酸的任何经修饰的化学组成(chemistry)或形式可彼此组合。例如,同一寡核苷酸内可包含一种、两种、三种、四种、五种或更多种不同类型的修饰。

[0434] 在一些实施方案中,可使用某些核苷酸或核苷修饰,所述核苷酸或核苷修饰使并入它们的寡核苷酸比天然寡脱氧核苷酸或寡核糖核苷酸分子对核酸酶消化更具抗性;这些经修饰寡核苷酸比未经修饰寡核苷酸完整存活更长的时间。经修饰寡核苷酸的一些具体实例包括含有经修饰的主链(backbone)的那些,例如经修饰核苷间键联,例如硫代磷酸酯键联、磷酸三酯键联、甲基膦酸酯键联、短链烷基键联或环烷基糖间键联或短链杂原子键联或杂环糖间键联。因此,本公开内容的寡核苷酸可例如通过并入修饰例如核苷酸或核苷修饰而稳定化以抵抗溶核降解。

[0435] 在一些实施方案中,寡核苷酸的长度可以是多至50个或多至100个核苷酸,其中寡核苷酸的2至10、2至15、2至16、2至17、2至18、2至19、2至20、2至25、2至30、2至40、2至45或更多个核苷酸或核苷是经修饰核苷酸/核苷。寡核苷酸的长度可以是8至30个核苷酸,其中寡核苷酸的2至10、2至15、2至16、2至17、2至18、2至19、2至20、2至25、2至30个核苷酸或核苷是经修饰核苷酸/核苷。寡核苷酸的长度可以是8至15个核苷酸,其中寡核苷酸的2至4、2至5、2至6、2至7、2至8、2至9、2至10、2至11、2至12、2至13、2至14个核苷酸或核苷是经修饰核苷酸/核苷。任选地,寡核苷酸可具有除1、2、3、4、5、6、7、8、9或10个经修饰核苷酸/核苷之外的每个核苷酸或核苷。寡核苷酸修饰在本文中进一步描述。

[0436] c.经修饰核苷

[0437] 在一些实施方案中,本文中所述的寡核苷酸包含在糖的2'位置处经修饰的至少一个核苷。在一些实施方案中,寡核苷酸包含至少一个2'-经修饰核苷。在一些实施方案中,寡核苷酸中的所有核苷均是2'-经修饰核苷。

[0438] 在一些实施方案中,本文中所述的寡核苷酸包含一个或更多个非双环2'-经修饰核苷,例如2'-脱氧、2'-氟(2'-F)、2'-O-甲基(2'-O-Me)、2'-O-甲氧基乙基(2'-MOE)、2'-O-氨基丙基(2'-O-AP)、2'-O-二甲基氨基乙基(2'-O-DMAOE)、2'-O-二甲基氨基丙基(2'-O-DMAP)、2'-O-二甲基氨基乙基氧基乙基(2'-O-DMAEOE)或2'-O-N-甲基乙酰胺基(2'-O-NMA)修饰的核苷。

[0439] 在一些实施方案中,本文中所述的寡核苷酸包含一个或更多个2'-4'双环核苷,其

中核糖环包含连接环中的两个原子的桥联部分,例如,通过亚甲基(LNA)桥联、亚乙基(ENA)桥联或(S)-约束性乙基(cEt)桥联将2'-O原子与4'-C原子连接。LNA的一些实例描述于国际专利申请公开WO/2008/043753中,其于2008年4月17日公开,并且标题为"RNA Antagonist Compounds For The Modulation Of PCSK9",其内容通过引用整体并入本文。ENA的一些实例在以下中提供:国际专利公开No.WO2005/042777,其于2005年5月12日公开,并且标题为"APP/ENA Antisense";Morita et al.,Nucleic Acid Res.,增刊1:241-242,2001;Surono et al.,Hum.Gene Ther.,15:749-757,2004;Koizumi,Curr.Opin.Mol.Ther.,8:144-149,2006以及Horie et al.,Nucleic Acids Symp.Ser(Oxf),49:171-172,2005;其公开内容通过引用整体并入本文。cEt的一些实例在以下中提供:美国专利7,101,993、7,399,845和7,569,686,其各自通过引用整体并入本文。

[0440] 在一些实施方案中,寡核苷酸包含在以下美国专利或专利申请公开之一中公开的经修饰核苷:美国专利7,399,845,其于2008年7月15日授权,并且标题为"6-Modified Bicyclic Nucleic Acid Analogs";美国专利7,741,457,其于2010年6月22日授权,并且标题为"6-Modified Bicyclic Nucleic Acid Analogs";美国专利8,022,193,其于2011年9月20日授权,并且标题为"6-Modified Bicyclic Nucleic Acid Analogs";美国专利7,569,686,其于2009年8月4日授权,并且标题为"Compounds And Methods For Synthesis Of Bicyclic Nucleic Acid Analogs";美国专利7,335,765,其于2008年2月26日授权,并且标题为"Novel Nucleoside And Oligonucleotide Analogues";美国专利7,314,923,其于2008年1月1日授权,并且标题为"Novel Nucleoside And Oligonucleotide Analogues";美国专利7,816,333,其于2010年10月19日授权,并且标题为"Oligonucleotide Analogues And Methods Utilizing The Same"和美国公开号2011/0009471,现在为美国专利8,957,201,其于2015年2月17日授权,并且标题为"Oligonucleotide Analogues And Methods Utilizing The Same",其各自的全部内容出于所有目的通过引用并入本文。

[0441] 在一些实施方案中,寡核苷酸包含至少一个经修饰核苷,其导致所述寡核苷酸与不具有至少一个经修饰核苷的寡核苷酸相比,Tm提高1℃、2℃、3℃、4℃或5℃。寡核苷酸可具有多个经修饰核苷,其导致所述寡核苷酸与不具有经修饰核苷的寡核苷酸相比,Tm总体提高2℃、3℃、4℃、5℃、6℃、7℃、8℃、9℃、10℃、15℃、20℃、25℃、30℃、35℃、40℃、45℃或更高。

[0442] 寡核苷酸可包含不同种类的核苷的混合物。例如,寡核苷酸可包含2'-脱氧核糖核苷或核糖核苷和2'-氟修饰核苷的混合物。寡核苷酸可包含脱氧核糖核苷或核糖核苷和2'-O-Me修饰核苷的混合物。寡核苷酸可包含2'-氟修饰核苷和2'-O-Me修饰核苷的混合物。寡核苷酸可包含2'-4'双环核苷和2'-MOE、2'-氟或2'-O-Me修饰核苷的混合物。寡核苷酸可包含非双环2'-经修饰核苷(例如,2'-MOE、2'-氟或2'-O-Me)和2'-4'双环核苷(例如,LNA、ENA、cEt)的混合物。

[0443] 寡核苷酸可包含不同种类的替代核苷。例如,寡核苷酸可包含替代的2'-脱氧核糖核苷或核糖核苷和2'-氟修饰核苷。寡核苷酸可包含替代的脱氧核糖核苷或核糖核苷和2'-O-Me修饰核苷。寡核苷酸可包含替代的2'-氟修饰核苷和2'-O-Me修饰核苷。寡核苷酸可包含替代的2'-4'双环核苷和2'-MOE、2'-氟或2'-O-Me修饰核苷。寡核苷酸可包含替代的非双

环2'-经修饰核苷(例如,2'-MOE、2'-氟或2'-O-Me)和2'-4'双环核苷(例如,LNA、ENA、cEt)。

[0444] 在一些实施方案中,本文中所述的寡核苷酸包含5'-乙烯基膦酸酯修饰、一个或更多个无碱基残基和/或一个或更多个倒置无碱基残基。

[0445] d.核苷间键联/主链

[0446] 在一些实施方案中,寡核苷酸可包含硫代磷酸酯键联或其他经修饰核苷间键联。在一些实施方案中,寡核苷酸包含硫代磷酸酯核苷间键联。在一些实施方案中,寡核苷酸在至少两个核苷之间包含硫代磷酸酯核苷间键联。在一些实施方案中,寡核苷酸在所有核苷之间均包含硫代磷酸酯核苷间键联。例如,在一些实施方案中,寡核苷酸在核苷酸序列的5'或3'末端的第一、第二和/或(例如,和)第三核苷间键联处包含经修饰核苷间键联。

[0447] 可使用的含磷的键联包括但不限于: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、甲基膦酸酯和包含3'亚烷基膦酸酯的其他烷基膦酸酯、以及手性膦酸酯、次膦酸酯、包含3'-氨基磷酰胺酯和氨基烷基磷酰胺酯的磷酰胺酯、硫羰基磷酰胺酯、硫羰基烷基膦酸酯、硫羰基烷基磷酸三酯和具有正常3'-5'键联的硼烷磷酸酯、这些的2'-5'连接类似物、以及其中相邻的核苷单元对连接在3'-5'到5'-3'或2'-5'到5'-2'的具有相反极性的那些;参见美国专利no:3,687,808;4,469,863;4,476,301;5,023,243;5,177,196;5,188,897;5,264,423;5,276,019;5,278,302;5,286,717;5,321,131;5,399,676;5,405,939;5,453,496;5,455,233;5,466,677;5,476,925;5,519,126;5,536,821;5,541,306;5,550,111;5,563,253;5,571,799;5,587,361;和5,625,050。

[0448] 在一些实施方案中,寡核苷酸可具有杂原子主链,例如亚甲基(甲基亚氨基)或MMI主链;酰胺主链(参见De Mesmaeker et al.Ace.Chem.Res.1995,28:366-374);吗啉代主链(参见Summerton and Weller,美国专利No.5,034,506);或肽核酸(peptide nucleic acid,PNA)主链(其中寡核苷酸的磷酸二酯主链被聚酰胺主链替换,核苷酸与聚酰胺主链的氮杂氮原子直接或间接结合,参见Nielsen et al.,Science 1991,254,1497)。

[0449] e.立体特异性寡核苷酸

[0450] 在一些实施方案中,寡核苷酸的核苷酸间磷原子是手性的,并且基于手性磷原子的构型调节寡核苷酸的特性。在一些实施方案中,可使用适当方法来以立体控制的方式合成P-手性寡核苷酸类似物(例如,如Oka N,Wada T,Stereocontrolled synthesis of oligonucleotide analogs containing chiral internucleotidic phosphorus atoms.Chem Soc Rev.2011Dec;40(12):5829-43中所述)。在一些实施方案中,提供了含硫代磷酸酯的寡核苷酸,其包含通过基本上所有的Sp或基本上所有的Rp硫代磷酸酯糖间键联连接在一起的核苷单元。在一些实施方案中,具有基本上手性纯糖间键联的这样的硫代磷酸酯寡核苷酸通过酶或化学合成制备,如例如于1996年12月12日授权的美国专利5,587,261中所述,其内容通过引用整体并入本文。在一些实施方案中,手性控制的寡核苷酸提供了靶核酸的选择性切割模式。例如,在一些实施方案中,手性控制的寡核苷酸在核酸的互补序列内提供了单个位点切割,如例如美国专利申请公开20170037399A1中所述,其于2017年2月2日公开,标题为"CHIRAL DESIGN",其内容通过引用整体并入本文。

[0451] f.吗啉代

[0452] 在一些实施方案中,寡核苷酸可以是基于吗啉代的化合物。基于吗啉代的寡聚化合物描述于Dwaine A.Braasch and David R.Corey,Biochemistry,2002,41(14),4503-

4510);Genesis,volume 30,issue 3,2001;Heasman,J.,Dev.Biol.,2002,243,209-214;Nasevicius et al.,Nat.Genet.,2000,26,216-220;Lacerra et al.,Proc.Natl.Acad.Sci.,2000,97,9591-9596;和1991年7月23日授权的美国专利No.5,034,506。在一些实施方案中,基于吗啉代的寡聚化合物是磷酸二酰胺吗啉代寡聚物(PMO)(例如,如Iverson,Curr.Opin.Mol.Ther.,3:235-238,2001;和Wang et al.,J.Gene Med.,12:354-364,2010中所述;其公开内容通过引用整体并入本文)。

[0453] g.肽核酸(PNA)

[0454] 在一些实施方案中,寡核苷酸的核苷酸单元的糖和核苷间键联(主链)二者均被新的基团替换。在一些实施方案中,碱基单元被维持以用于与适当的核酸靶化合物杂交。一种这样的寡聚化合物(已显示具有优异的杂交特性的寡核苷酸模拟物)被称为肽核酸(PNA)。在PNA化合物中,寡核苷酸的糖-主链被含酰胺的主链(例如,氨基乙基甘氨酸主链)替换。核碱基被保留并且与主链酰胺部分的氮杂氮原子直接或间接结合。报道制备PNA化合物的代表性出版物包括但不限于美国专利no.5,539,082;5,714,331;和5,719,262,其各自通过引用并入本文。PNA化合物的进一步教导可见于Nielsen et al.,Science,1991,254,1497-1500中。

[0455] h.混合聚体

[0456] 在一些实施方案中,本文中所述的寡核苷酸可以是混合聚体或者包含混合聚体序列模式。通常来说,混合聚体是包含天然和非天然存在的核苷二者的寡核苷酸或者通常以替代模式包含两种不同类型的非天然存在的核苷的寡核苷酸。混合聚体通常比未经修饰寡核苷酸具有更高的结合亲和力,并且可用于特异性结合靶分子,例如,来阻断靶分子上的结合位点。通常来说,混合聚体不向靶分子募集RNA酶并且因此不促进靶分子的切割。已描述了不能募集RNA酶H的这样的寡核苷酸,例如,参见WO2007/112754或WO2007/112753。

[0457] 在一些实施方案中,混合聚体包含以下或者由以下组成:重复模式的核苷类似物和天然存在的核苷、或者一种类型的核苷类似物和第二种类型的核苷类似物。然而,混合聚体不需要包含重复模式,并且可替代地包含经修饰核苷和天然存在的核苷的任何排列,或者一种类型的经修饰核苷和第二种类型的经修饰核苷的任何排列。重复模式可以是例如每第二个或每第三个核苷是经修饰核苷(例如LNA),并且剩余核苷是天然存在的核苷,例如DNA,或是2'经取代的核苷类似物,例如2'MOE或2'氟类似物,或者本文中所述的任何其他经修饰核苷。公认的是,经修饰核苷的重复模式,例如LNA单元,可在固定位置例如在5'或3'末端与经修饰核苷组合。

[0458] 在一些实施方案中,混合聚体不包含多于5个、多于4个、多于3个或多于2个连续的天然存在的核苷(例如DNA核苷)的区域。在一些实施方案中,混合聚体包含至少一个由至少两个连续的经修饰核苷,例如至少两个连续的LNA组成的区域。在一些实施方案中,混合聚体包含至少一个由至少三个连续的经修饰核苷单元,例如至少三个连续的LNA组成的区域。

[0459] 在一些实施方案中,混合聚体不包含多于7个、多于6个、多于5个、多于4个、多于3个或多于2个连续的核苷类似物例如LNA的区域。在一些实施方案中,LNA单元可被其他核苷类似物例如本文中提及的那些替换。

[0460] 混合聚体可被设计为包含亲和力增强的经修饰核苷(例如在非限制性实例中LNA核苷和2'-O-Me核苷)的混合物。在一些实施方案中,混合聚体在至少两个、至少三个、至少

四个、至少五个或更多个核苷之间包含经修饰核苷间键联(例如,硫代磷酸酯核苷间键联或其他键联)。

[0461] 可使用任何合适的方法产生混合聚体。教导制备混合聚体的代表性美国专利、美国专利出版物和PCT出版物包括美国专利公开No.US20060128646、US20090209748、US20090298916、US20110077288和US20120322851以及美国专利No.7687617。

[0462] 在一些实施方案中,混合聚体包含一个或更多个吗啉代核苷。例如,在一些实施方案中,混合聚体可包含与一个或更多个其他核苷(例如,DNA、RNA核苷)或经修饰核苷(例如,LNA、2'-O-Me核苷)混合(例如,以交替方式混合)的吗啉代核苷。

[0463] 在一些实施方案中,混合聚体可用于剪接校正或外显子跳读,例如,如以下中报道的:Touznik A.,et al.,LNA/DNA mixmer-based antisense oligonucleotides correct alternative splicing of the SMN2 gene and restore SMN protein expression in type 1SMA fibroblasts Scientific Reports,volume 7,Article number:3672(2017),Chen S.et al.,Synthesis of a Morpholino Nucleic Acid(MNA)-Uridine Phosphoramidite,and Exon Skipping Using MNA/2′-O-Methyl Mixmer Antisense Oligonucleotide,Molecules 2016,21,1582,其各自的内容通过引用并入本文。

[0464] i.多聚体

[0465] 在一些实施方案中,分子载荷可包含通过接头连接的2个或更多个寡核苷酸的多聚体(例如,串联体)。在一些实施方案中,以这种方式,复合物的寡核苷酸负载可提高到超过靶向剂上的可用连接位点(例如,抗体上的可用硫醇位点),或者以其他方式调节以实现特定的载荷负载量。多聚体中的寡核苷酸可相同或不同(例如,靶向不同基因或者同一基因或其产物上的不同位点)。

[0466] 在一些实施方案中,多聚体包含通过可切割接头连接在一起的2个或更多个寡核苷酸。然而,在一些实施方案中,多聚体包含通过不可切割接头连接在一起的2个或更多个寡核苷酸。在一些实施方案中,多聚体包含2、3、4、5、6、7、8、9、10个或更多个连接在一起的寡核苷酸。在一些实施方案中,多聚体包含2至5、2至10或4至20个连接在一起的寡核苷酸。

[0467] 在一些实施方案中,多聚体包含2个或更多个首尾相连(以线性排列)的寡核苷酸。在一些实施方案中,多聚体包含2个或更多个通过基于寡核苷酸的接头(例如,聚-dT接头,一种无碱基接头)首尾相连的寡核苷酸。在一些实施方案中,多聚体包含一个寡核苷酸的5'末端,其与另一个寡核苷酸的3'末端连接。在一些实施方案中,多聚体包含一个寡核苷酸的3'末端,其与另一个寡核苷酸的3'末端连接。在一些实施方案中,多聚体包含一个寡核苷酸的5'末端,其与另一个寡核苷酸的5'末端连接。尽管如此,在一些实施方案中,多聚体可包含分支结构,所述分支结构包含通过分支接头连接在一起的多个寡核苷酸。

[0468] 可用于本文中提供的复合物中的多聚体的另一些实例在以下中公开:例如,美国专利申请号2015/0315588A1,标题为Methods of delivering multiple targeting oligonucleotides to a cell using cleavable linkers,其于2015年11月5日公开;美国专利申请号2015/0247141A1,标题为Multimeric Oligonucleotide Compounds,其于2015年9月3日公开;美国专利申请号US2011/0158937A1,标题为Immunostimulatory Oligonucleotide Multimers,其于2011年6月30日公开;和美国专利号5,693,773,标题为Triplex-Forming Antisense Oligonucleotides Having Abasic Linkers Targeting

Nucleic Acids Comprising Mixed Sequences Of Purines And Pyrimidines,其于1997年12月2日授权,其各自的内容通过引用整体并入本文。

[0469] C.接头

[0470] 本文中所述的复合物通常包含共价连接本文中所述的任一种抗TfR1抗体与分子载荷的接头。接头包含至少一个共价键。在一些实施方案中,接头可以是共价连接抗TfR1抗体与分子载荷的单键,例如二硫键或二硫桥。然而,在一些实施方案中,接头可通过多个共价键共价连接本文中所述的任一种抗TfR1抗体与分子载荷。在一些实施方案中,接头可以是可切割接头。然而,在一些实施方案中,接头可以是不可切割接头。接头通常在体外和体内是稳定的,并且在某些细胞环境中可以是稳定的。另外,通常来说接头不负面影响抗TfR1抗体或分子载荷的功能特性。接头合成的一些实例和方法是本领域中已知的(参见,例如Kline,T.et al.“Methods to Make Homogenous Antibody Drug Conjugates.”Pharmaceutical Research,2015,32:11,3480-3493.;Jain,N.et al.“Current ADC Linker Chemistry”Pharm Res.2015,32:11,3526-3540.;McCombs,J.R.and Owen,S.C.“Antibody Drug Conjugates:Design and Selection of Linker,Payload and Conjugation Chemistry”AAPS J.2015,17:2,339-351.)。

[0471] 接头通常将包含允许与抗TfR1抗体和分子载荷二者连接的两种不同的反应性物质。在一些实施方案中,两种不同的反应性物质可以是亲核体和/或亲电体。在一些实施方案中,接头包含两种不同的亲电体或亲核体,其对两种不同的亲核体或亲电体具有特异性。在一些实施方案中,接头通过与抗TfR1抗体的赖氨酸残基或半胱氨酸残基缀合而与抗TfR1抗体共价连接。在一些实施方案中,接头通过含马来酰亚胺的接头与抗TfR1抗体的半胱氨酸残基共价连接,其中任选地,含马来酰亚胺的接头包含马来酰亚胺基己酰基或马来酰亚胺基甲基环己烷-1-羧酸根基团。在一些实施方案中,接头通过3-芳基丙腈官能团与抗TfR1抗体的半胱氨酸残基或巯基官能化的分子载荷共价连接。在一些实施方案中,接头与抗TfR1抗体的赖氨酸残基共价连接。在一些实施方案中,接头通过酰胺键、氨基甲酸酯键、酰肼、三唑、硫醚和/或二硫键与抗TfR1抗体和/或(例如,和)分子载荷独立地共价连接。

[0472] i.可切割接头

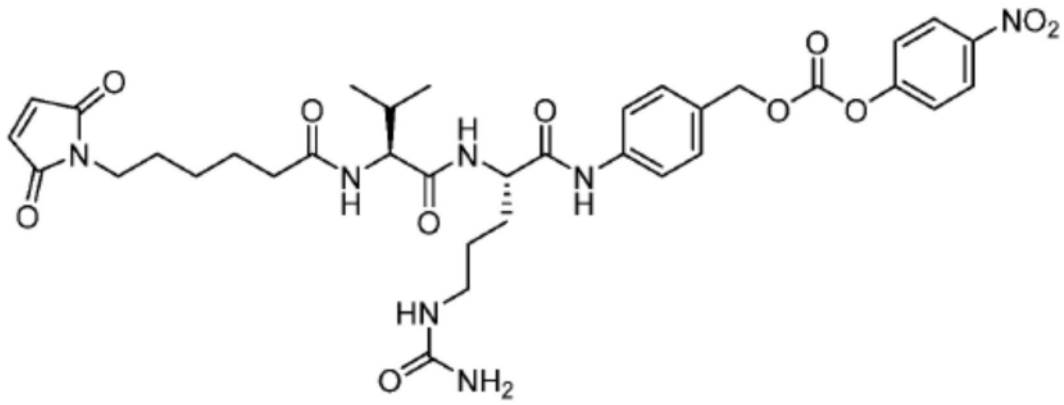
[0473] 可切割接头可以是蛋白酶敏感性接头、pH敏感性接头或谷胱甘肽敏感性接头。这些接头通常仅在胞内可切割,并且优选在胞外环境,例如肌细胞胞外中稳定。

[0474] 蛋白酶敏感性接头可通过蛋白酶酶活性切割。这些接头通常包含肽序列,并且长度可为2至10个氨基酸、约2至5个氨基酸、约5至10个氨基酸、约10个氨基酸、约5个氨基酸、约3个氨基酸或约2个氨基酸。在一些实施方案中,肽序列可包含天然存在的氨基酸(例如半胱氨酸、丙氨酸)或者非天然存在的或经修饰的氨基酸。非天然存在的氨基酸包括β-氨基酸、高氨基酸、脯氨酸衍生物、3-经取代的丙氨酸衍生物、线性核心氨基酸、N-甲基氨基酸和本领域已知的其他氨基酸。在一些实施方案中,蛋白酶敏感性接头包含缬氨酸-瓜氨酸或丙氨酸-瓜氨酸序列。在一些实施方案中,蛋白酶敏感性接头可被溶酶体蛋白酶(例如组织蛋白酶B(cathepsin B))和/或(例如,和)内体蛋白酶切割。

[0475] pH敏感性接头是在高或低pH环境中容易降解的共价键联。在一些实施方案中,pH敏感性接头可在4至6的pH下被切割。在一些实施方案中,pH敏感性接头包含腙或环缩醛。在一些实施方案中,pH敏感性接头在内体或溶酶体内被切割。

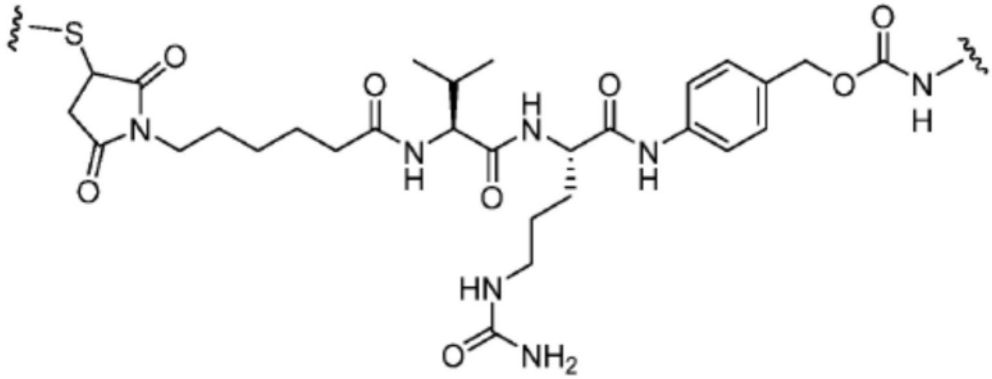
[0476] 在一些实施方案中,谷胱甘肽敏感性接头包含二硫化物部分。在一些实施方案中,谷胱甘肽敏感性接头通过与细胞内的谷胱甘肽物质进行二硫化物交换反应来切割。在一些实施方案中,二硫化物部分还包含至少一种氨基酸,例如半胱氨酸残基。

[0477] 在一些实施方案中,接头包含缬氨酸-瓜氨酸序列(例如,如美国专利6,214,345中所述,其通过引用并入本文)。在一些实施方案中,在缀合之前,接头包含以下结构:

[0478] 

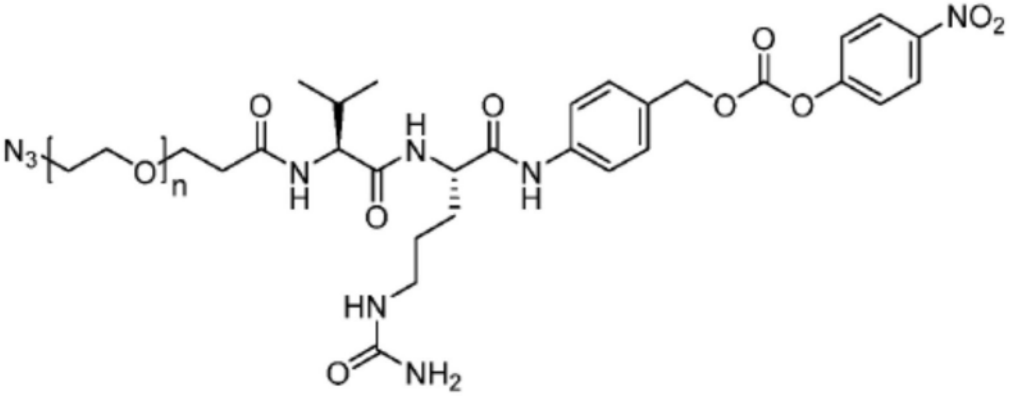
。

[0479] 在一些实施方案中,在缀合之后,接头包含以下结构:

[0480] 

。

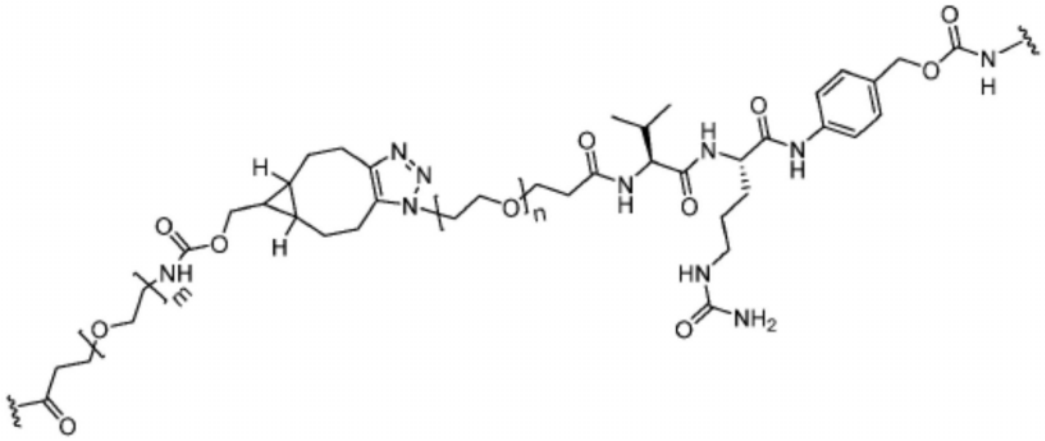
[0481] 在一些实施方案中,在缀合之前,接头包含以下结构:

[0482] 

(A),

[0483] 其中n是0至10的任意数字。在一些实施方案中,n是3。

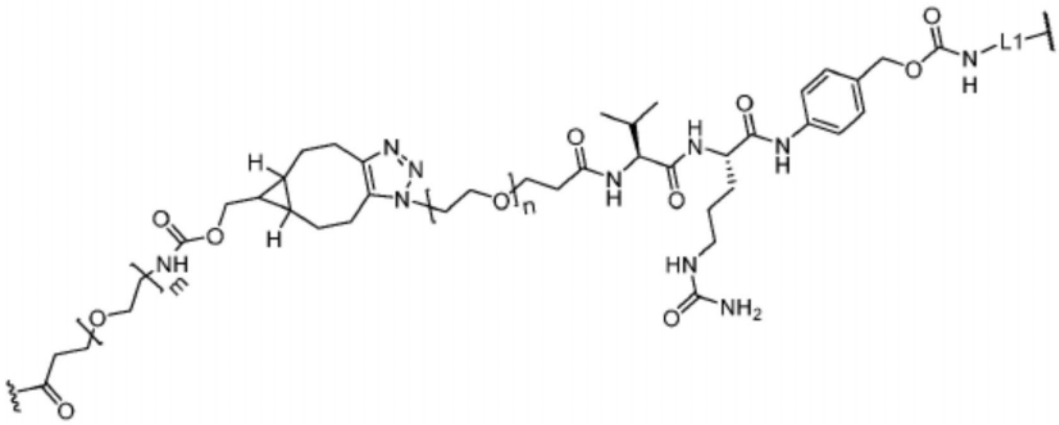
[0484] 在一些实施方案中,接头包含以下结构:

[0485] 

(H),

[0486] 其中n是0至10的任意数字,其中m是0至10的任意数字。在一些实施方案中,n是3和/或(例如,和)m是4。

[0487] 在一些实施方案中,接头包含以下结构:

[0488] 

(I),

[0489] 其中n是0至10的任意数字,其中m是0至10的任意数字。在一些实施方案中,n是3和/或(例如,和)m是4。

[0490] ii.不可切割接头

[0491] 在一些实施方案中,可使用不可切割接头。通常来说,不可切割接头在细胞或生理环境中不容易被降解。在一些实施方案中,不可切割接头包含任选经取代的烷基,其中所述取代可包括卤素、羟基、氧物质和其他常见的取代。在一些实施方案中,接头可包含任选经取代的烷基、任选经取代的亚烷基、任选经取代的亚芳基、亚杂芳基、包含至少一种非天然氨基酸的肽序列、截短的聚糖、不能被酶降解的一种或更多种糖、叠氮化物、炔-叠氮化物、包含LPXT序列的肽序列、硫醚、生物素、联苯、聚乙二醇或等同化合物的重复单元、酸性酯、酰胺、磺酰胺和/或烷氧基-胺接头。在一些实施方案中,分选酶介导的连接可用于共价连接包含LPXT序列的抗TfR1抗体与包含$(G)_n$序列的分子载荷(参见,例如Proft T.Sortase-mediated protein ligation:an emerging biotechnology tool for protein modification and immobilization.Biotechnol Lett.2010,32(1):1-10.)。

[0492] 在一些实施方案中,接头可包含经取代的亚烷基、任选经取代的亚烯基、任选经取代的亚炔基、任选经取代的亚环烷基、任选经取代的亚环烯基、任选经取代的亚芳基、还包含至少一个选自N、O和S的杂原子的任选经取代的亚杂芳基;还包含至少一个选自N、O和S的

杂原子的任选经取代的亚杂环基;亚氨基、任选经取代的氮物质、任选经取代的氧物质O、任选经取代的硫物质或聚(环氧烷烃),例如聚环氧乙烷或聚环氧丙烷。在一些实施方案中,接头可以是不可切割的N-γ-马来酰亚胺基丁酰基-氧基琥珀酰亚胺酯(GMBS)接头。

[0493] iii.接头缀合

[0494] 在一些实施方案中,接头通过磷酸酯、硫醚、醚、碳-碳、氨基甲酸酯、或酰胺键与抗TfR1抗体和/或(例如,和)分子载荷共价连接。在一些实施方案中,接头通过磷酸酯或硫代磷酸酯基团例如寡核苷酸主链的末端磷酸酯与寡核苷酸共价连接。在一些实施方案中,接头通过抗TfR1抗体上存在的赖氨酸或半胱氨酸残基与抗TfR1抗体共价连接。

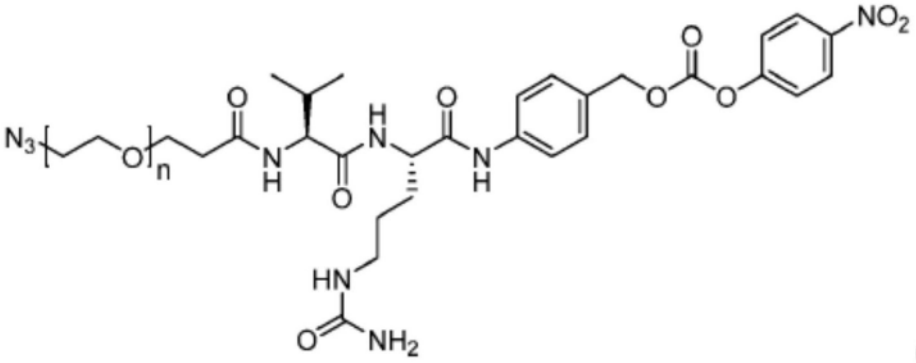
[0495] 在一些实施方案中,接头或其一部分通过叠氮化物和炔烃之间的环加成反应与抗TfR1抗体和/或(例如,和)分子载荷共价连接以形成三唑,其中叠氮化物或炔烃可位于抗TfR1抗体、分子载荷或接头上。在一些实施方案中,炔烃可以是环炔烃,例如环辛炔。在一些实施方案中,炔烃可以是双环壬炔(也称为双环[6.1.0]壬炔或BCN)或经取代的双环壬炔。在一些实施方案中,环辛炔如国际专利申请公开WO2011136645中所述,其于2011年11月3日公开,标题为"Fused Cyclooctyne Compounds And Their Use In Metal-free Click Reactions"。在一些实施方案中,叠氮化物可以是包含叠氮化物的糖或碳水化合物分子。在一些实施方案中,叠氮化物可以是6-叠氮基-6-脱氧半乳糖或6-叠氮基-N-乙酰半乳糖胺。在一些实施方案中,包含叠氮化物的糖或碳水化合物分子如国际专利申请公开WO2016170186中所述,其于2016年10月27日公开,标题为"Process For The Modification Of A Glycoprotein Using A Glycosyltransferase That Is Or Is Derived From Aβ(1,4)-N-Acetylgalactosaminyltransferase"。在一些实施方案中,在叠氮化物和炔烃之间进行环加成反应以形成三唑,其中叠氮化物或炔烃可位于抗TfR1抗体、分子载荷或接头上,如以下中所述:国际专利申请公开WO2014065661,其于2014年5月1日公开,标题为"Modified antibody,antibody-conjugate and process for the preparation thereof";或国际专利申请公开WO2016170186,其于2016年10月27日公开,标题为"Process For The Modification Of A Glycoprotein Using A Glycosyltransferase That Is Or Is Derived From Aβ(1,4)-N-Acetylgalactosaminyltransferase"。

[0496] 在一些实施方案中,接头包含间隔基,例如聚乙二醇间隔基或酰基/胺甲酰基磺酰胺间隔基,例如HydraSpace™间隔基。在一些实施方案中,间隔基如Verkade,J.M.M.et al.,"A Polar Sulfamide Spacer Significantly Enhances the Manufacturability,Stability,and Therapeutic Index of Antibody-Drug Conjugates",Antibodies,2018,7,12中所述。

[0497] 在一些实施方案中,接头通过亲二烯体和二烯/杂二烯之间的第尔斯-阿尔德反应(Diels-Alder reaction)与抗TfR1抗体和/或(例如,和)分子载荷共价连接,其中亲二烯体或二烯/杂二烯可位于抗TfR1抗体、分子载荷或接头上。在一些实施方案中,接头通过其他周环反应(pericyclic reaction),例如烯反应与抗TfR1抗体和/或(例如,和)分子载荷共价连接。在一些实施方案中,接头通过酰胺、硫代酰胺或磺酰胺键合反应与抗TfR1抗体和/或(例如,和)分子载荷共价连接。在一些实施方案中,接头通过缩合反应与抗TfR1抗体和/或(例如,和)分子载荷共价连接,以形成存在于接头和抗TfR1抗体和/或(例如,和)分子载荷之间的肟、腙或氨基脲基团。

[0498] 在一些实施方案中,接头通过亲核体(例如胺或羟基)和亲电体(例如羧酸、碳酸或醛)之间的共轭加成反应与抗TfR1抗体和/或(例如,和)分子载荷共价连接。在一些实施方案中,在进行接头和抗TfR1抗体或分子载荷之间的反应之前,亲核体可存在于接头上并且亲电体可存在于抗TfR1抗体或分子载荷上。在一些实施方案中,在进行接头和抗TfR1抗体或分子载荷之间的反应之前,亲电体可存在于接头上并且亲核体可存在于抗TfR1抗体或分子载荷上。在一些实施方案中,亲电体可以是叠氮化物、五氟苯基、硅中心、羰基、羧酸、酸酐、异氰酸酯、硫代异氰酸酯、琥珀酰亚胺酯、磺基琥珀酰亚胺酯、马来酰亚胺、烷基卤化物、烷基假卤化物、环氧化物、环硫化物、氮丙啶、芳基、活化的磷中心和/或活化的硫中心。在一些实施方案中,亲核体可以是任选经取代的烯烃、任选经取代的炔烃、任选经取代的芳基、任选经取代的杂环基、羟基、氨基、烷基氨基、苯胺基和/或硫醇基。

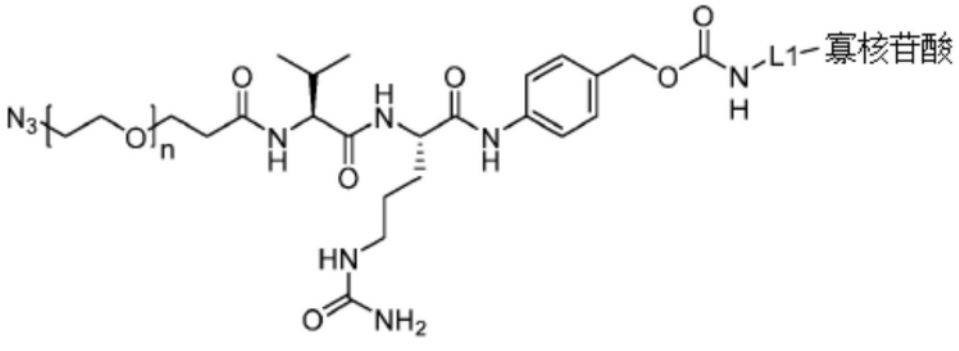
[0499] 在一些实施方案中,接头包含与反应性化学部分(例如,用于点击化学的叠氮化物部分或BCN部分)共价连接的缬氨酸-瓜氨酸序列。在一些实施方案中,包含与反应性化学部分(例如,用于点击化学的叠氮化物部分)共价连接的缬氨酸-瓜氨酸序列的接头包含以下结构:

[0500] 

(A),

[0501] 其中n是0至10的任意数字。在一些实施方案中,n是3。

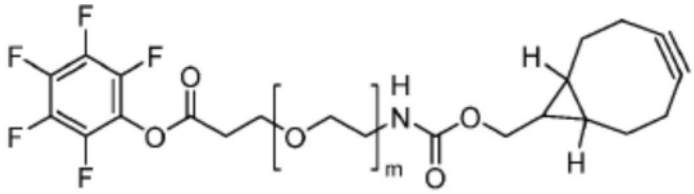
[0502] 在一些实施方案中,包含式(A)结构的接头与分子载荷(例如,寡核苷酸)共价连接(例如,任选地经由另外的化学部分)。在一些实施方案中,包含式(A)结构的接头与寡核苷酸共价连接(例如,通过与胺-L1-寡核苷酸的亲核取代形成氨基甲酸酯键),产生包含以下结构的化合物:

[0503] 

(B),

[0504] 其中n是0至10的任意数字。在一些实施方案中,n是3。

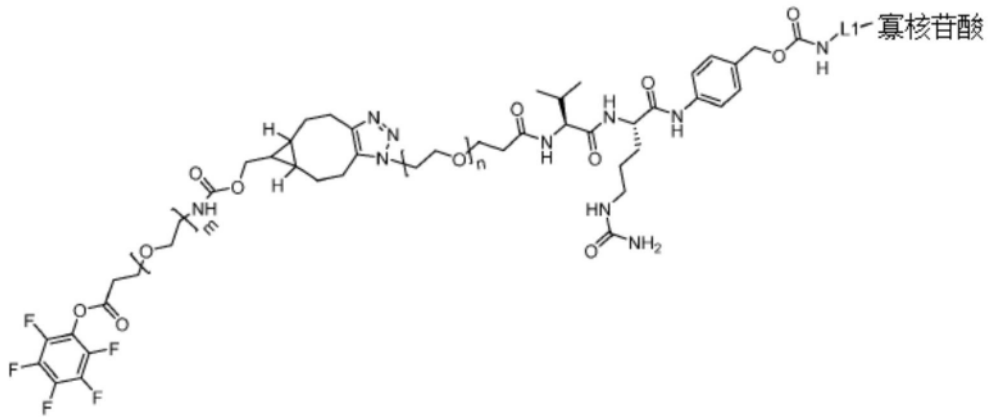
[0505] 在一些实施方案中,式(B)化合物通过三唑与另外的部分进一步共价连接,其中三唑通过式(A)或式(B)的叠氮化物与双环壬炔上提供的炔烃之间的点击反应形成。在一些实施方案中,包含双环壬炔的化合物包含以下结构:

[0506] 

(C)，

[0507] 其中m是0至10的任意数字。在一些实施方案中，m是4。

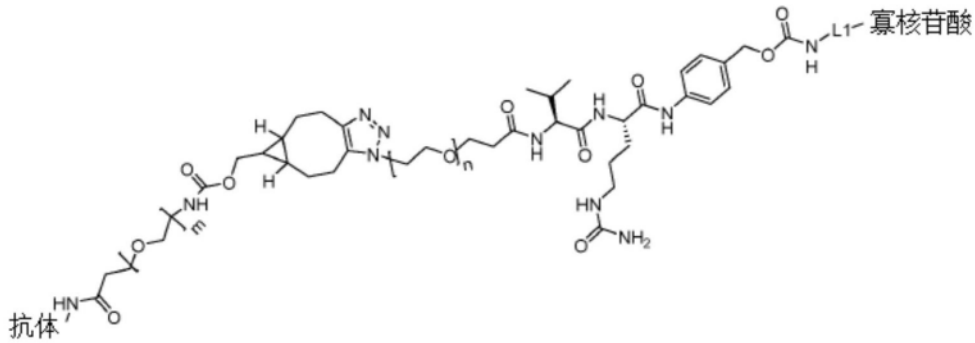
[0508] 在一些实施方案中，结构(B)化合物的叠氮化物通过与结构(C)化合物的炔烃的点击反应形成三唑，形成包含以下结构的化合物：

[0509] 

(D)，

[0510] 其中n是0至10的任意数字，并且其中m是0至10的任意数字。在一些实施方案中，n是3并且m是4。

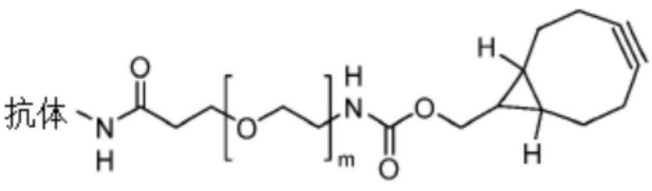
[0511] 在一些实施方案中，结构(D)化合物还与抗TfR1抗体的赖氨酸共价连接，形成包含以下结构的复合物：

[0512] 

(E)，

[0513] 其中n是0至10的任意数字，其中m是0至10的任意数字。在一些实施方案中，n是3和/或(例如，和)m是4。应理解，在式(E)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

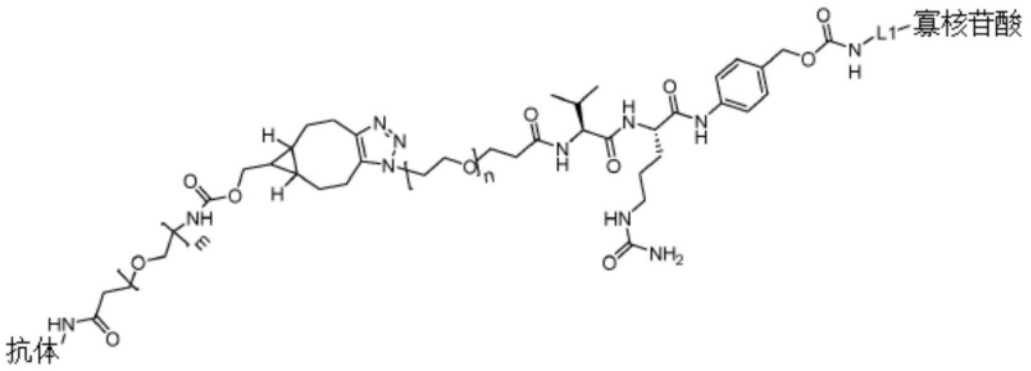
[0514] 在一些实施方案中，式(C)化合物还与抗TfR1抗体的赖氨酸共价连接，形成包含以下结构的化合物：

[0515] 

(F)，

[0516] 其中m是0至15(例如4)。应理解,在式(F)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

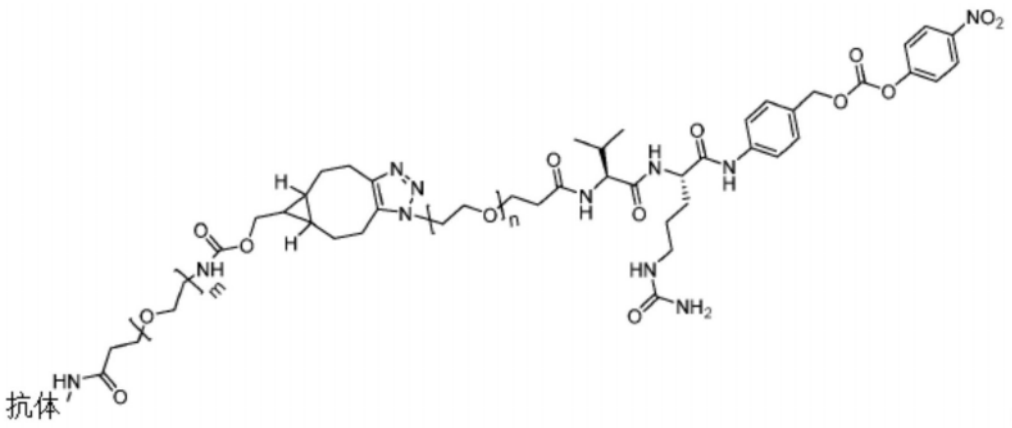
[0517] 在一些实施方案中,结构(B)化合物的叠氮化物通过与结构(F)化合物的炔烃的点击反应形成三唑,形成包含以下结构的复合物:

[0518] 

(E),

[0519] 其中n是0至10的任意数字,其中m是0至10的任意数字。在一些实施方案中,n是3和/或(例如,和)m是4。应理解,在式(E)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

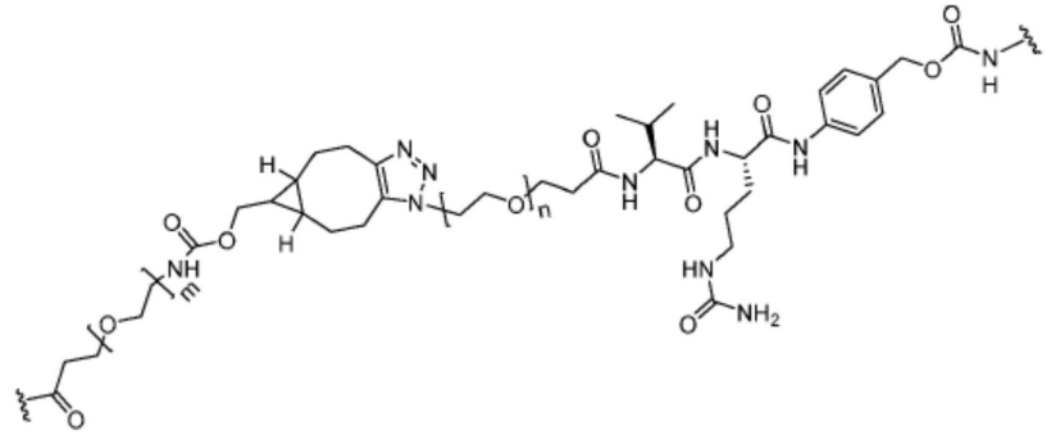
[0520] 在一些实施方案中,结构(A)化合物的叠氮化物通过与结构(F)化合物的炔烃的点击反应形成三唑,形成包含以下结构的化合物:

[0521] 

(G),

[0522] 其中n是0至10的任意数字,其中m是0至10的任意数字。在一些实施方案中,n是3和/或(例如,和)m是4。在一些实施方案中,寡核苷酸与包含式(G)结构的化合物共价连接,从而形成包含式(E)结构的复合物。应理解,在式(G)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

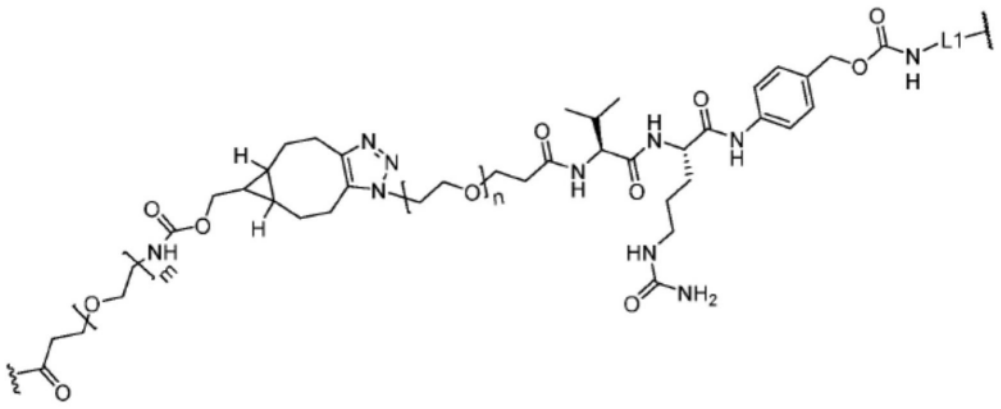
[0523] 在一些实施方案中,在本文中所述的任一复合物中,抗TfR1抗体经由包含以下结构的接头通过抗TfR1抗体的赖氨酸与分子载荷(例如,寡核苷酸)共价连接:

[0524] 

(H),

[0525] 其中n是0至10的任意数字,其中m是0至10的任意数字。在一些实施方案中,n是3和/或(例如,和)m是4。

[0526] 在一些实施方案中,在本文所述的任一复合物中,抗TfR1抗体经由包含以下结构的接头通过抗TfR1抗体的赖氨酸与分子载荷(例如,寡核苷酸)共价连接:

[0527] 

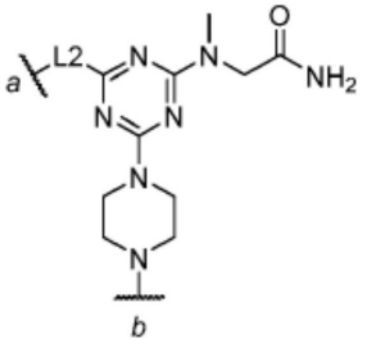
(I),

[0528] 其中n是0至10的任意数字,其中m是0至10的任意数字。在一些实施方案中,n是3和/或(例如,和)m是4。

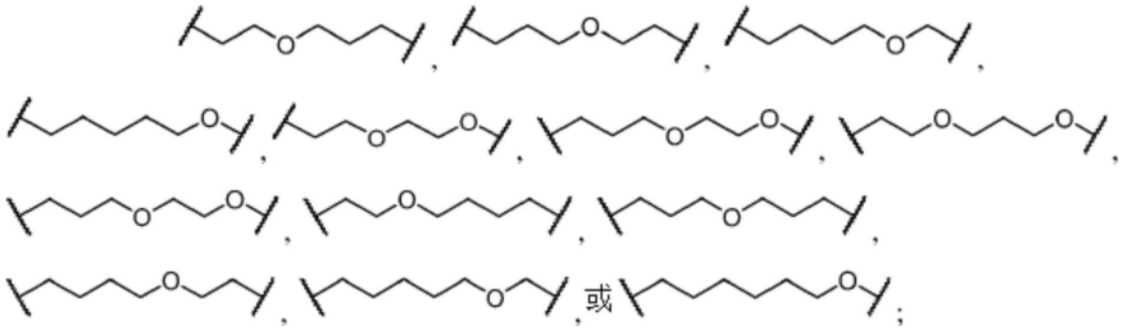
[0529] 在一些实施方案中,在式(B)、(D)、(E)和(I)中,L1是间隔基,所述间隔基是经取代或未经取代的脂族、经取代或未经取代的杂脂族、经取代或未经取代的亚碳环基(carbocycylene)、经取代或未经取代的亚杂环基、经取代或未经取代的亚芳基、经取代或未经取代的亚杂芳基、

[0530] -O-,-N(R^A)-,-S-,-C(=O)-,-C(=O)O-,-C(=O)NR^A-,-NR^AC(=O)-,-NR^AC(=O)R^A-,-C(=O)R^A-,-NR^AC(=O)O-,-NR^AC(=O)N(R^A)-,-OC(=O)-,-OC(=O)O-,-OC(=O)N(R^A)-,-$S(O)_2NR^A$-,-$NR^AS(O)_2$-,

[0531] 或其组合,其中每个R^A独立地是氢或者经取代或未经取代的烷基。在一些实施方案中,L1是

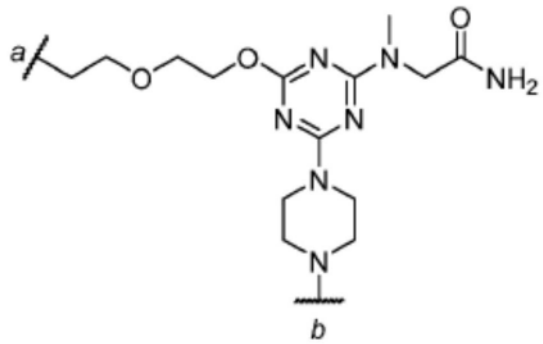
[0532] 

[0533] 其中L2是

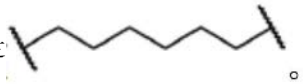
[0534] 

[0535] 其中a标记与式(B)、(D)、(E)和(I)的氨基甲酸酯部分直接连接的位点;以及b标记与寡核苷酸共价连接(直接或通过另外的化学部分进行连接)的位点。

[0536] 在一些实施方案中,L1是:

[0537] 

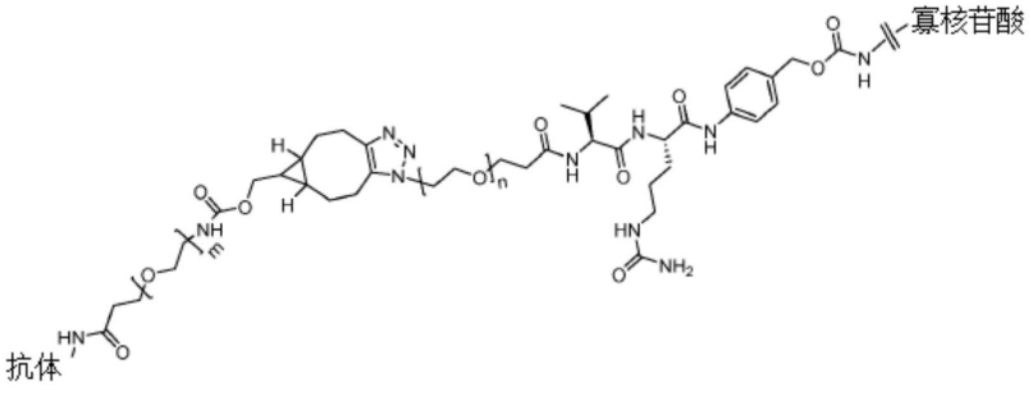
[0538] 其中a标记与式(B)、(D)、(E)和(I)的氨基甲酸酯部分直接连接的位点;以及b标记与寡核苷酸共价连接(直接或通过另外的化学部分进行连接)的位点。

[0539] 在一些实施方案中,L1是  。

[0540] 在一些实施方案中,L1与寡核苷酸的5'磷酸连接。在一些实施方案中,磷酸是磷酸二酯。在一些实施方案中,L1与寡核苷酸的5'硫代磷酸酯连接。在一些实施方案中,L1与寡核苷酸的5'膦酰胺酯(phosphonoamidate)连接。在一些实施方案中,L1通过磷酸二酰胺键联与寡核苷酸的5'末端连接。

[0541] 在一些实施方案中,L1是任选的(例如,不是必要存在的)。

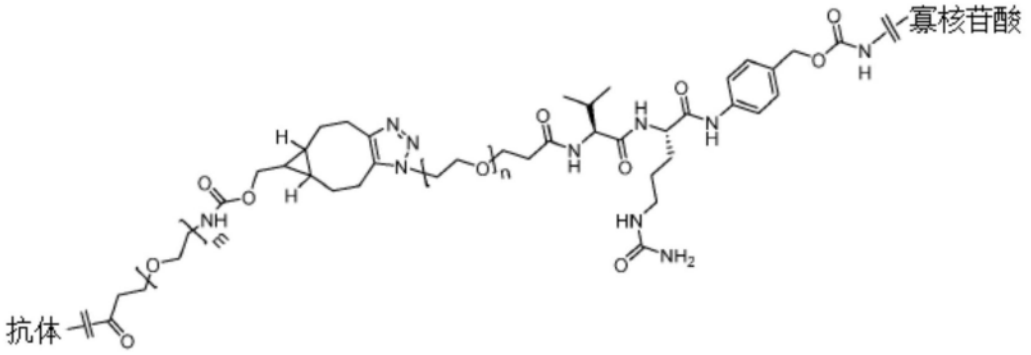
[0542] 在一些实施方案中,本文中所述的任一种复合物具有以下结构:

[0543] 

(J)，

[0544] 其中n是0至15(例如3)并且m是0至15(例如4)。应理解,在式(J)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

[0545] 在一些实施方案中,本文中所述的任一种复合物具有以下结构:

[0546] 

(K)，

[0547] 其中n是0至15(例如3)并且m是0至15(例如4)。

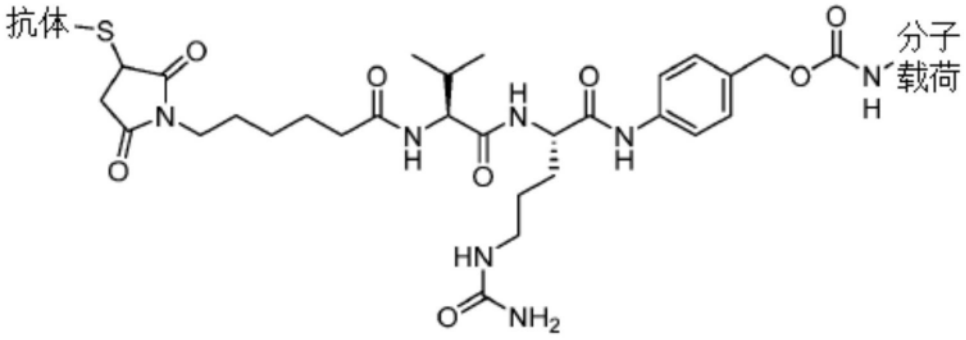
[0548] 在一些实施方案中,在与化合物(例如,式(A)或式(G)化合物)连接之前,寡核苷酸经修饰以在5'末端、3'末端处或在内部包含胺基(例如,作为胺官能化的核碱基)。

[0549] 尽管在抗TfR1抗体和寡核苷酸分子载荷的背景下描述了接头缀合,但应理解,预期在其他肌肉靶向剂(例如,其他肌肉靶向抗体)和/或其他分子载荷上使用这样的接头缀合。

[0550] D.抗体-分子载荷复合物的一些实例

[0551] 本文中还提供了包含与本文中所述的任何分子载荷(例如,寡核苷酸)共价连接的任一种本文中所述的抗TfR1抗体的复合物的一些非限制性实例。在一些实施方案中,抗TfR1抗体(例如,表2至表7中提供的任一种抗TfR1抗体)通过接头与分子载荷(例如,寡核苷酸,例如表8中提供的寡核苷酸)共价连接。可使用本文中所述的任何接头。在一些实施方案中,如果分子载荷是寡核苷酸,则接头与寡核苷酸的5'末端、寡核苷酸的3'末端、或寡核苷酸的内部位点连接。在一些实施方案中,接头通过硫醇反应性键联(例如,通过抗TfR1抗体中的半胱氨酸)与抗TfR1抗体连接。在一些实施方案中,接头(例如,包含缬氨酸-瓜氨酸序列的接头)通过胺基(例如,通过抗体中的赖氨酸)与抗体(例如,本文中所述的抗TfR1抗体)连接。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

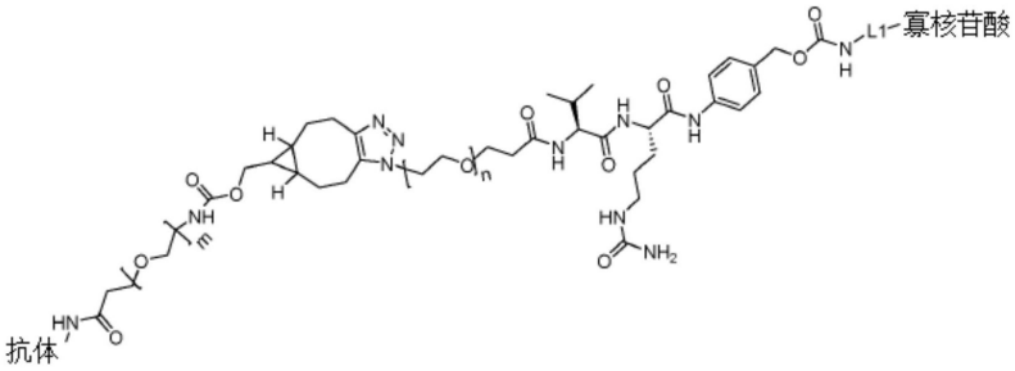
[0552] 以下提供了复合物的结构的一个实例,所述复合物包含通过接头与分子载荷共价连接的抗TfR1抗体:

[0553] 

,

[0554] 其中接头通过硫醇反应性键联(例如,通过抗体中的半胱氨酸)与抗体连接。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0555] 以下提供了复合物的结构的另一个实例,所述复合物包含通过接头与分子载荷共价连接的抗TfR1抗体:

[0556] 

(E) ,

[0557] 其中n是0至10的数字,其中m是0至10的数字,其中接头通过胺基(例如,赖氨酸残基上的)与抗体连接,和/或(例如,和)其中接头与寡核苷酸(例如,在5'末端、3'末端或内部)连接。在一些实施方案中,接头通过赖氨酸与抗体连接,接头在5'末端与寡核苷酸连接,n是3并且m是4。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。应理解,在式(E)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

[0558] 应理解,抗体可与具有不同化学计量的分子载荷连接,该特性可被称为药物抗体比(drug to antibody ratio,DAR),其中"药物"是分子载荷。在一些实施方案中,一个分子载荷与一个抗体连接(DAR=1)。在一些实施方案中,两个分子载荷与一个抗体连接(DAR=2)。在一些实施方案中,三个分子载荷与一个抗体连接(DAR=3)。在一些实施方案中,四个分子载荷与一个抗体连接(DAR=4)。在一些实施方案中,提供了不同复合物的混合物,每种复合物具有不同的DAR。在一些实施方案中,这样的混合物中的复合物的平均DAR可在1至3、1至4、1至5或更大的范围内。混合物中复合物的平均DAR不必要是整数值。可通过将分子载荷缀合至抗体上的不同位点和/或(例如,和)通过将多聚体缀合至抗体上的一个或更多个位点来提高DAR。例如,可通过将单个分子载荷缀合至抗体上的两个不同位点或通过将二聚体分子载荷缀合至抗体的单个位点来实现DAR为2。

[0559] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的本文中所述

的抗TfR1抗体(例如,表2至7中提供的抗体)。在一些实施方案中,本文中所述的复合物包含通过接头(例如,包含缬氨酸-瓜氨酸序列的接头)与分子载荷共价连接的本文中所述的抗TfR1抗体(例如,表2至7中提供的抗体)。在一些实施方案中,接头(例如,包含缬氨酸-瓜氨酸序列的接头)通过硫醇反应性键联(例如,通过抗体中的半胱氨酸)与抗体(例如,本文中所述的抗TfR1抗体)连接。在一些实施方案中,接头(例如,包含缬氨酸-瓜氨酸序列的接头)通过胺基(例如,通过抗体中的赖氨酸)与抗体(例如,本文中所述的抗TfR1抗体)连接。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0560] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含表2中列出的任一种抗体的CDR-H1、CDR-H2、CDR-H3、CDR-L1、CDR-L2和CDR-L3。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0561] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:69、SEQ ID NO:71或SEQ ID NO:72的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0562] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:73或SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0563] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:73或SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0564] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:78的氨基酸序列的VL。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0565] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:77或SEQ ID NO:79的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0566] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗

体,其中所述抗TfR1抗体包含含有SEQ ID NO:154的氨基酸序列的VH和含有SEQ ID NO:155的氨基酸序列的VL。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0567] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:84、SEQ ID NO:86或SEQ ID NO:87的氨基酸序列的重链和含有SEQ ID NO:85的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0568] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:88或SEQ ID NO:91的氨基酸序列的重链和含有SEQ ID NO:89的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0569] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:88或SEQ ID NO:91的氨基酸序列的重链和含有SEQ ID NO:90的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0570] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:92或SEQ ID NO:94的氨基酸序列的重链和含有SEQ ID NO:95的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0571] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:92的氨基酸序列的重链和含有SEQ ID NO:93的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0572] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:156的氨基酸序列的重链和含有SEQ ID NO:157的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0573] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:97、SEQ ID NO:98或SEQ ID NO:99的氨基酸序列的重链和含有SEQ ID NO:85的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0574] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:100或SEQ ID NO:101的氨基酸序列的重链和含有SEQ ID NO:89的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

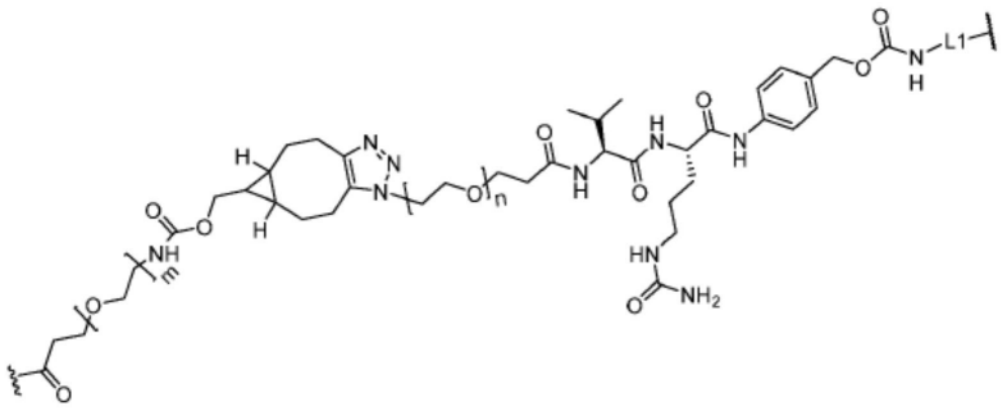
[0575] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:100或SEQ ID NO:101的氨基酸序列的重链和含有SEQ ID NO:90的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0576] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:102的氨基酸序列的重链和含有SEQ ID NO:93的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0577] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:102或SEQ ID NO:103的氨基酸序列的重链和含有SEQ ID NO:95的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0578] 在一些实施方案中,本文中所述的复合物包含与分子载荷共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含含有SEQ ID NO:158或SEQ ID NO:159的氨基酸序列的重链和含有SEQ ID NO:157的氨基酸序列的轻链。在一些实施方案中,分子载荷是DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)。

[0579] 在本文中所述的任何示例性复合物中,在一些实施方案中,抗TfR1抗体通过包含以下结构的接头与分子载荷共价连接:

[0580] 

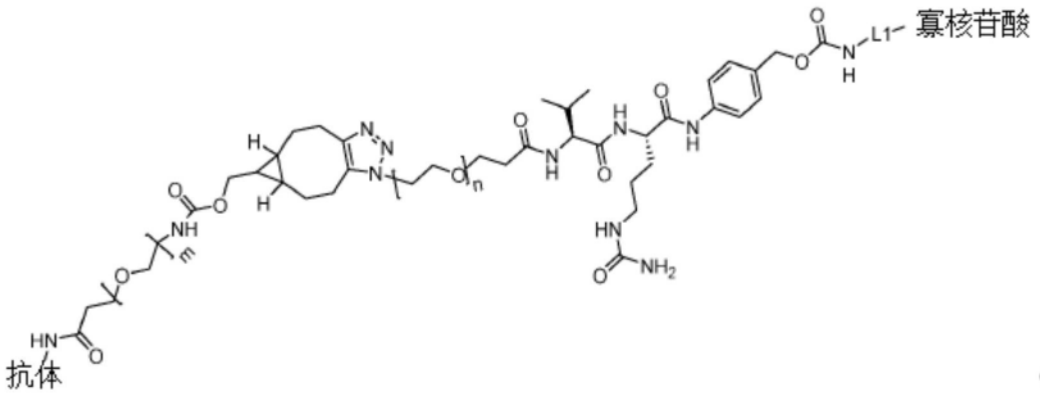
(I)

,

[0581] 其中n是3,m是4。

[0582] 在一些实施方案中,本文中所述的复合物包含通过抗TfR1抗体中的赖氨酸与DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID

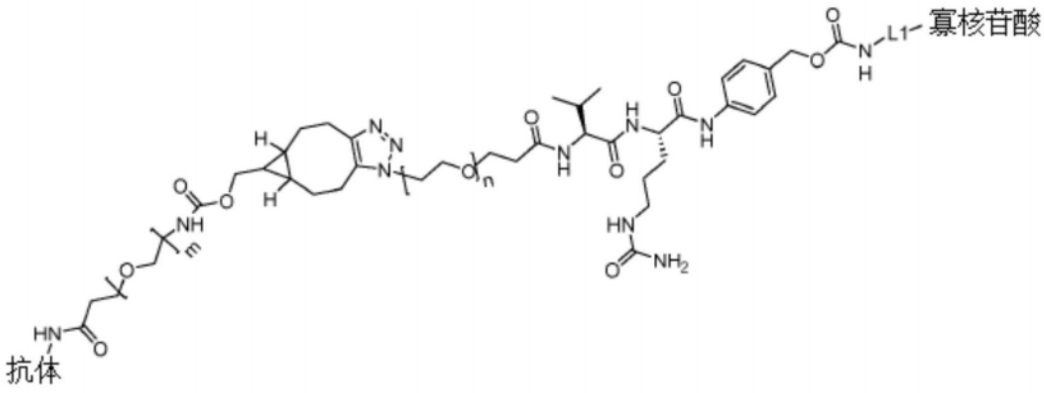
NO:160至383中任一者互补的DMD靶向寡核苷酸)的5'末端共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含表2中列出的任一种抗体的CDR-H1、CDR-H2、CDR-H3、CDR-L1、CDR-L2和CDR-L3,其中所述复合物具有以下结构:

[0583] 

(E),

[0584] 其中n是3并且m是4。应理解,在式(E)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

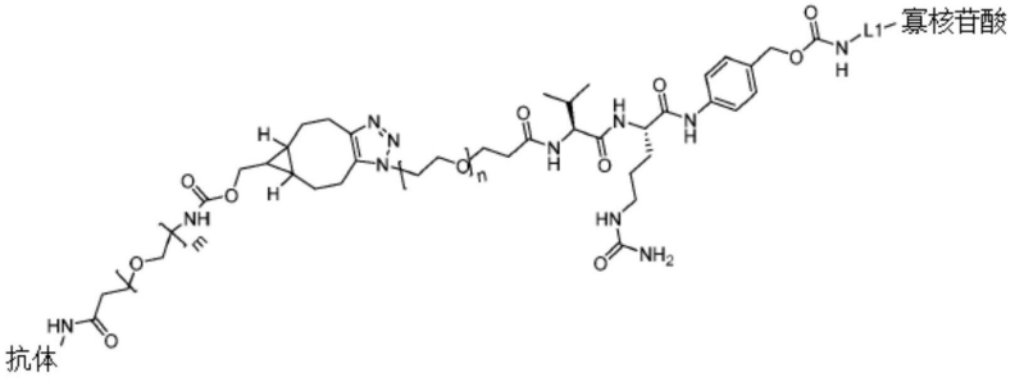
[0585] 在一些实施方案中,本文中所述的复合物包含通过抗TfR1抗体中的赖氨酸与DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)的5'末端共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含表3中列出的任一种抗体的VH和VL,其中所述复合物具有以下结构:

[0586] 

(E),

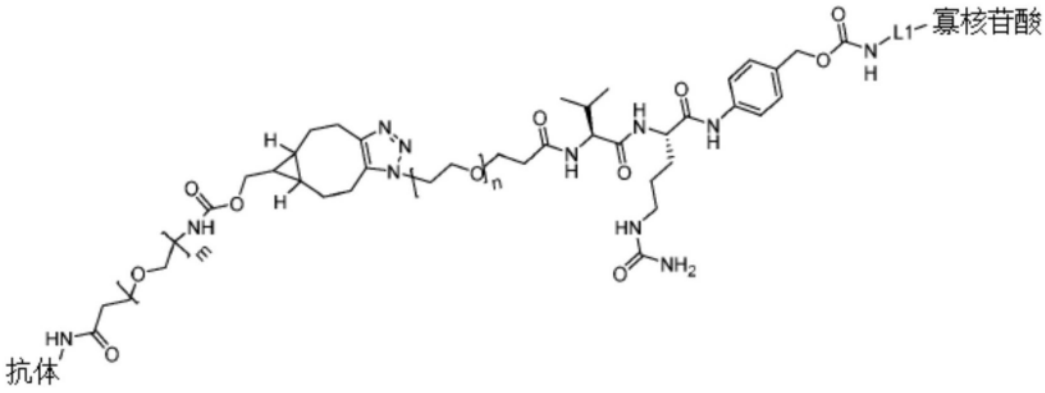
[0587] 其中n是3并且m是4。应理解,在式(E)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

[0588] 在一些实施方案中,本文中所述的复合物包含通过抗TfR1抗体中的赖氨酸与DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)的5'末端共价连接的抗TfR1抗体,其中所述抗TfR1抗体包含表4中列出的任一种抗体的重链和轻链,其中所述复合物具有以下结构:

[0589] 

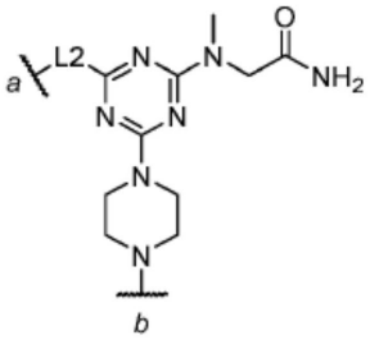
[0590] 其中n是3并且m是4。应理解,在式(E)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

[0591] 在一些实施方案中,本文中所述的复合物包含通过抗TfR1抗体中的赖氨酸与DMD靶向寡核苷酸(例如,表8中列出的、由SEQ ID NO:384至831中任一者提供的、或与SEQ ID NO:160至383中任一者互补的DMD靶向寡核苷酸)的5'末端共价连接的抗TfR1 Fab,其中所述抗TfR1 Fab包含表5中列出的任一种抗体的重链和轻链,其中所述复合物具有以下结构:

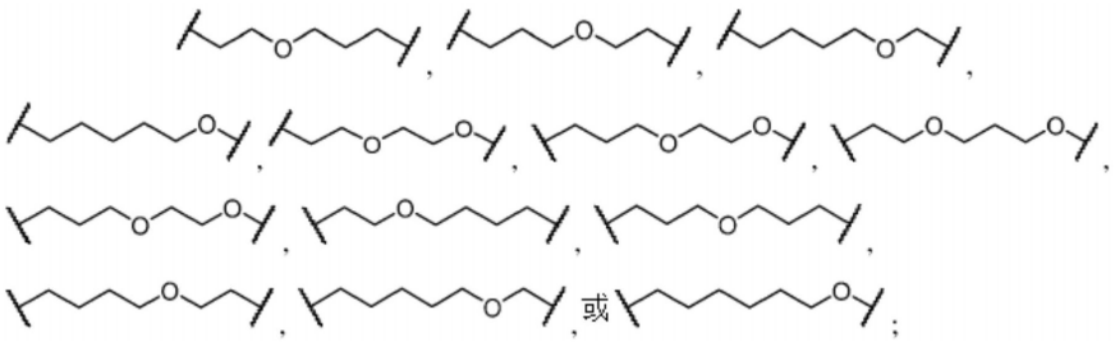
[0592] 

[0593] 其中n是3并且m是4。应理解,在式(E)中示出的与抗TfR1抗体相邻的酰胺是由与抗TfR1抗体的胺(例如赖氨酸ε胺)的反应产生的。

[0594] 在一些实施方案中,在本文中所述的复合物的任一个实例中,L1是:

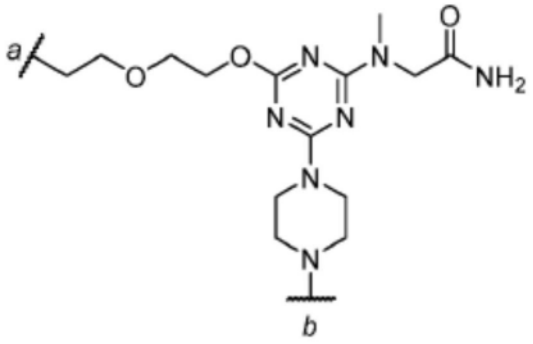
[0595] 

[0596] 其中L2是

[0597] 

[0598] 其中a标记与式(B)、(D)、(E)和(I)的氨基甲酸酯部分直接连接的位点;以及b标记与寡核苷酸共价连接(直接或通过另外的化学部分进行连接)的位点。

[0599] 在一些实施方案中,L1是:

[0600] 

[0601] 其中a标记与式(B)、(D)、(E)和(I)的氨基甲酸酯部分直接连接的位点;以及b标记与寡核苷酸共价连接(直接或通过另外的化学部分进行连接)的位点。

[0602] 在一些实施方案中,L1与寡核苷酸的5’磷酸连接。在一些实施方案中,磷酸是磷酸二酯。在一些实施方案中,L1与寡核苷酸的5’硫代磷酸酯连接。在一些实施方案中,L1与寡核苷酸的5’膦酰胺酯(phosphonoamidate)连接。在一些实施方案中,L1通过磷酸二酰胺键联与寡核苷酸的5’末端连接。

[0603] 在一些实施方案中,L1是任选的(例如,不是必要存在的)。

[0604] III.制剂

[0605] 本文中提供的复合物可以以任何合适的方式配制。通常来说,本文中提供的复合物以适用于药物用途的方式配制。例如,可使用使降解最小化、促进递送和/或(例如,和)摄取或者为制剂中的复合物提供另外的有益特性的制剂将复合物递送至对象。在一些实施方案中,本文中提供了包含复合物和可药用载体的组合物。这样的组合物可适当地配制,使得当施用于对象时,无论是施用至靶细胞的直接环境中或全身性施用,足够量的复合物都能进入靶肌细胞。在一些实施方案中,复合物被配制在缓冲溶液例如磷酸缓冲盐水溶液、脂质体、胶束结构和衣壳中。

[0606] 应理解,在一些实施方案中,组合物可分别包含本文中提供的复合物的一种或更多种组分(例如,肌肉靶向剂、接头、分子载荷或它们中任一者的前体分子)。

[0607] 在一些实施方案中,复合物被配制在水或水溶液(例如,用pH调节的水)中。在一些实施方案中,复合物被配制在碱性缓冲水溶液(例如,PBS)中。在一些实施方案中,本文中公开的制剂包含赋形剂。在一些实施方案中,赋形剂赋予组合物以改善的稳定性、改善的吸收、改善的溶解性和/或(例如,和)活性成分的治疗性增强。在一些实施方案中,赋形剂是缓

冲剂(例如,柠檬酸钠、磷酸钠、tris碱或氢氧化钠)或载剂(例如,缓冲溶液、矿脂(petrolatum)、二甲基亚砜或矿物油)。

[0608] 在一些实施方案中,将复合物或其组分(例如,寡核苷酸或抗体)冻干用于延长其保质期,并随后在使用(例如,施用于对象)之前制成溶液。因此,包含本文中所述的复合物或其组分的组合物中的赋形剂可以是冻干保护剂(例如,甘露醇、乳糖、聚乙二醇或聚乙烯吡咯烷酮)或崩解温度调节剂(例如,右旋糖酐、ficoll或明胶)。

[0609] 在一些实施方案中,药物组合物被配制成与其预期施用途径相容。施用途径的一些实例包括肠胃外施用,例如静脉内、皮内、皮下施用。通常来说,施用途径是静脉内或皮下。

[0610] 适用于可注射使用的药物组合物包含无菌水溶液(在具有水溶性时)或分散体,以及用于即时制备无菌可注射溶液或分散体的无菌粉末。载体可以是溶剂或分散介质,其包含例如水、乙醇、多元醇(例如,甘油、丙二醇和液体聚乙二醇等),及其合适的混合物。在一些实施方案中,组合物中的制剂包含等张剂,例如糖、多元醇例如甘露醇、山梨醇以及氯化钠。无菌可注射溶液可通过将所需量的复合物与以上列举的成分之一或组合(根据需要)一起并入选定的溶剂中,然后过滤灭菌来制备。

[0611] 在一些实施方案中,组合物可包含至少约0.1%的复合物或其组分,或者更多,尽管活性成分的百分比可为总组合物的重量或体积的约1%至约80%或更多。本领域技术人员将考虑制备这样的药物制剂的因素例如溶解性、生物利用度、生物学半衰期、施用途径、产品保质期以及其他药理学考虑因素,并且因此多种剂量和治疗方案可以是期望的。

[0612] IV.使用/治疗方法

[0613] 如本文中所述的包含与分子载荷共价连接的肌肉靶向剂的复合物在治疗患有肌养蛋白病(例如,迪谢内肌营养不良)的对象中是有效的。在一些实施方案中,复合物包含分子载荷,所述分子载荷是寡核苷酸,例如促进由突变DMD等位基因表达的前体mRNA的外显子跳读的反义寡核苷酸。

[0614] 在一些实施方案中,对象可以是人对象、非人灵长类对象、啮齿动物对象或任何合适的哺乳动物对象。在一些实施方案中,对象可患有迪谢内肌营养不良或其他肌养蛋白病。在一些实施方案中,对象具有突变DMD等位基因,其可任选地在DMD外显子中包含至少一个突变,该突变导致移码突变并导致不正确的RNA剪接/加工。在一些实施方案中,对象患有严重肌养蛋白病的症状,例如肌肉萎缩或肌肉损失。在一些实施方案中,对象具有肌酸磷酸激酶(CK)的血清浓度的无症状提高和/或(例如,和)伴随有肌红蛋白尿的肌肉痉挛。在一些实施方案中,对象患有进行性肌肉疾病,例如迪谢内肌营养不良或贝克肌营养不良或DMD相关扩张型心肌病(DCM)。在一些实施方案中,对象不患有肌养蛋白病的症状。

[0615] 在一些实施方案中,对象具有适合于外显子51跳读的DMD基因中的突变。在一些实施方案中,如本文中所述的包含与分子载荷共价连接的肌肉靶向剂的复合物在治疗在DMD基因中具有适合于外显子51跳读的突变的对象方面是有效的。在一些实施方案中,复合物包含分子载荷,所述分子载荷是寡核苷酸,例如,促进前体mRNA的外显子51的跳读的反义寡核苷酸,例如在由突变DMD基因(例如,适合于外显子51跳读的突变DMD基因)编码的前体mRNA中。

[0616] 本公开内容的一个方面包括涉及向对象施用有效量的本文中所述的复合物的方

法。在一些实施方案中,可向有治疗需要的对象施用有效量的包含含有与分子载荷共价连接的肌肉靶向剂的复合物的药物组合物。在一些实施方案中,可通过合适的途径施用包含如本文中所述的复合物的药物组合物,所述途径可包括静脉内施用,例如作为推注(bolus)或通过在一段时间内的连续输注。在一些实施方案中,可通过肌内、腹膜内、脑脊髓内、皮下、关节内、滑膜内或鞘内途径进行施用。在一些实施方案中,药物组合物可以是固体形式、水性形式或液体形式。在一些实施方案中,可将水性或液体形式雾化或冻干。在一些实施方案中,雾化或冻干形式可用水溶液或液体溶液重构。

[0617] 用于静脉内施用的组合物可包含多种载体,例如植物油、二甲基乙酰胺、二甲基甲酰胺、乳酸乙酯、碳酸乙酯、豆蔻酸异丙酯、乙醇以及多元醇(甘油、丙二醇、液体聚乙二醇等)。对于静脉内注射,水溶性抗体可通过滴注方法施用,通过该方法输注包含抗体和生理学上可接受的赋形剂的药物制剂。生理上可接受的赋形剂可包括,例如5%右旋糖、0.9%盐水、林格液(Ringer's solution)或其他合适的赋形剂。可将肌内制剂,例如抗体的合适可溶性盐形式的无菌制剂溶解在药用赋形剂例如注射用水、0.9%盐水或5%葡萄糖溶液中并施用。

[0618] 在一些实施方案中,包含含有与分子载荷共价连接的肌肉靶向剂的复合物的药物组合物通过位点特异性或局部递送技术施用。这些技术的一些实例包括复合物的可植入储库源、局部递送导管、位点特异性载体、直接注射或直接应用。

[0619] 在一些实施方案中,包含含有与分子载荷共价连接的肌肉靶向剂的复合物的药物组合物以赋予对象治疗作用的有效浓度施用。如本领域技术人员所公认的,有效量根据疾病的严重程度、所治疗对象的独特特征(例如年龄、身体状况、健康或体重)、治疗的持续时间、任何同时治疗的性质、施用途径和相关因素而变化。这些相关因素是本领域技术人员已知的,并且仅通过常规实验即可解决。在一些实施方案中,有效浓度是被认为对患者安全的最大剂量。在一些实施方案中,有效浓度将是提供最大效力的最低的可能浓度。

[0620] 经验考虑因素(例如复合物在对象中的半衰期)通常将有助于确定用于治疗的药物组合物的浓度。施用频率可凭经验确定和调整以使治疗效力最大化。

[0621] 可使用任何合适的方法评估治疗的效力。在一些实施方案中,治疗的效力可如下进行评估:通过对与肌养蛋白病相关的症状(例如肌肉萎缩或肌无力)的观察结果进行的评价,通过对对象的自我报告结局(例如移动性、自我护理、日常活动、疼痛/不适和焦虑/抑郁)的测量,或通过生活质量指标(例如寿命)。

[0622] 在一些实施方案中,将包含含有与分子载荷共价连接的肌肉靶向剂的本文中所述的复合物的药物组合物以相对于对照(例如治疗之前基因表达的基线水平)足以调节至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%或至少95%的靶基因活性或表达的有效浓度施用于对象。

[0623] 另外的实施方案

[0624] 1.复合物,其包含与分子载荷共价连接的抗转铁蛋白受体1(TfR1)抗体,所述分子载荷被配置成用于诱导DMD前体mRNA中外显子51的跳读,其中所述抗TfR1抗体是表2至7中的任一者中标识的抗体。

[0625] 2.实施方案1所述的复合物,其中所述抗TfR1抗体包含:

[0626] (i)SEQ ID NO:33的重链互补决定区1(CDR-H1)、SEQ ID NO:34的重链互补决定区

2(CDR-H2)、SEQ ID NO:35的重链互补决定区3(CDR-H3)、SEQ ID NO:36的轻链互补决定区1(CDR-L1)、SEQ ID NO:37的轻链互补决定区2(CDR-L2)和SEQ ID NO:32的轻链互补决定区3(CDR-L3)；

[0627] (ii)SEQ ID NO:7的CDR-H1、SEQ ID NO:8的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3；

[0628] (iii)SEQ ID NO:7的CDR-H1、SEQ ID NO:20的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3；

[0629] (iv)SEQ ID NO:7的CDR-H1、SEQ ID NO:24的CDR-H2、SEQ ID NO:9的CDR-H3、SEQ ID NO:10的CDR-L1、SEQ ID NO:11的CDR-L2和SEQ ID NO:6的CDR-L3；

[0630] (v)SEQ ID NO:51的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3；

[0631] (vi)SEQ ID NO:64的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3；或者

[0632] (vii)SEQ ID NO:67的CDR-H1、SEQ ID NO:52的CDR-H2、SEQ ID NO:53的CDR-H3、SEQ ID NO:54的CDR-L1、SEQ ID NO:55的CDR-L2和SEQ ID NO:50的CDR-L3。

[0633] 3.实施方案1或实施方案2所述的复合物,其中所述抗TfR1抗体包含：

[0634] (i)含有与SEQ ID NO:76具有至少85％同一性的氨基酸序列的重链可变区(VH)；和/或含有与SEQ ID NO:75具有至少85％同一性的氨基酸序列的轻链可变区(VL)；

[0635] (ii)含有与SEQ ID NO:69具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85％同一性的氨基酸序列的VL；

[0636] (iii)含有与SEQ ID NO:71具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85％同一性的氨基酸序列的VL；

[0637] (iv)含有与SEQ ID NO:72具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:70具有至少85％同一性的氨基酸序列的VL；

[0638] (v)含有与SEQ ID NO:73具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:74具有至少85％同一性的氨基酸序列的VL；

[0639] (vi)含有与SEQ ID NO:73具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:75具有至少85％同一性的氨基酸序列的VL；

[0640] (vii)含有与SEQ ID NO:76具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:74具有至少85％同一性的氨基酸序列的VL；

[0641] (viii)含有与SEQ ID NO:77具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:78具有至少85％同一性的氨基酸序列的VL；

[0642] (ix)含有与SEQ ID NO:79具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:80具有至少85％同一性的氨基酸序列的VL;或者

[0643] (x)含有与SEQ ID NO:77具有至少85％同一性的氨基酸序列的VH;和/或含有与SEQ ID NO:80具有至少85％同一性的氨基酸序列的VL。

[0644] 4.实施方案1至3中任一项所述的复合物,其中所述抗TfR1抗体包含：

[0645] (i)含有SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL；

[0646] (ii)含有SEQ ID NO:69的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

[0647] (iii)含有SEQ ID NO:71的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

[0648] (iv)含有SEQ ID NO:72的氨基酸序列的VH和含有SEQ ID NO:70的氨基酸序列的VL;

[0649] (v)含有SEQ ID NO:73的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL;

[0650] (vi)含有SEQ ID NO:73的氨基酸序列的VH和含有SEQ ID NO:75的氨基酸序列的VL;

[0651] (vii)含有SEQ ID NO:76的氨基酸序列的VH和含有SEQ ID NO:74的氨基酸序列的VL;

[0652] (viii)含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:78的氨基酸序列的VL;

[0653] (ix)含有SEQ ID NO:79的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL;或者

[0654] (x)含有SEQ ID NO:77的氨基酸序列的VH和含有SEQ ID NO:80的氨基酸序列的VL。

[0655] 5.实施方案1至4中任一项所述的复合物,其中所述抗TfR1抗体是Fab片段、Fab'片段、F(ab')2片段、scFv、Fv或全长IgG。

[0656] 6.实施方案5所述的复合物,其中所述抗TfR1抗体是Fab片段。

[0657] 7.实施方案6所述的复合物,其中所述抗TfR1抗体包含:

[0658] (i)含有与SEQ ID NO:101具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:90具有至少85%同一性的氨基酸序列的轻链;

[0659] (ii)含有与SEQ ID NO:97具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:85具有至少85%同一性的氨基酸序列的轻链;

[0660] (iii)含有与SEQ ID NO:98具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:85具有至少85%同一性的氨基酸序列的轻链;

[0661] (iv)含有与SEQ ID NO:99具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:85具有至少85%同一性的氨基酸序列的轻链;

[0662] (v)含有与SEQ ID NO:100具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:89具有至少85%同一性的氨基酸序列的轻链;

[0663] (vi)含有与SEQ ID NO:100具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:90具有至少85%同一性的氨基酸序列的轻链;

[0664] (vii)含有与SEQ ID NO:101具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:89具有至少85%同一性的氨基酸序列的轻链;

[0665] (viii)含有与SEQ ID NO:102具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:93具有至少85%同一性的氨基酸序列的轻链;

[0666] (ix)含有与SEQ ID NO:103具有至少85%同一性的氨基酸序列的重链;和/或含有

与SEQ ID NO:95具有至少85%同一性的氨基酸序列的轻链;或者

[0667] (x)含有与SEQ ID NO:102具有至少85%同一性的氨基酸序列的重链;和/或含有与SEQ ID NO:95具有至少85%同一性的氨基酸序列的轻链。

[0668] 8.实施方案6或实施方案7所述的复合物,其中所述抗TfR1抗体包含:

[0669] (i)含有SEQ ID NO:101的氨基酸序列的重链;和含有SEQ ID NO:90的氨基酸序列的轻链;

[0670] (ii)含有SEQ ID NO:97的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

[0671] (iii)含有SEQ ID NO:98的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

[0672] (iv)含有SEQ ID NO:99的氨基酸序列的重链;和含有SEQ ID NO:85的氨基酸序列的轻链;

[0673] (v)含有SEQ ID NO:100的氨基酸序列的重链;和含有SEQ ID NO:89的氨基酸序列的轻链;

[0674] (vi)含有SEQ ID NO:100的氨基酸序列的重链;和含有SEQ ID NO:90的氨基酸序列的轻链;

[0675] (vii)含有SEQ ID NO:101的氨基酸序列的重链;和含有SEQ ID NO:89的氨基酸序列的轻链;

[0676] (viii)含有SEQ ID NO:102的氨基酸序列的重链;和含有SEQ ID NO:93的氨基酸序列的轻链;

[0677] (ix)含有SEQ ID NO:103的氨基酸序列的重链;和含有SEQ ID NO:95的氨基酸序列的轻链;或者

[0678] (x)含有SEQ ID NO:102的氨基酸序列的重链;和含有SEQ ID NO:95的氨基酸序列的轻链。

[0679] 9.实施方案1至8中任一项所述的复合物,其中所述抗TfR1抗体不与转铁蛋白受体1的转铁蛋白结合位点特异性结合,和/或其中所述抗TfR1抗体不抑制转铁蛋白与转铁蛋白受体1的结合。

[0680] 10.实施方案1至9中任一项所述的复合物,其中所述分子载荷包含寡核苷酸。

[0681] 11.实施方案10所述的复合物,其中所述寡核苷酸促进所述DMD前体RNA中反义介导的外显子跳读。

[0682] 12.实施方案10或11所述的复合物,其中所述寡核苷酸包含所述DMD前体mRNA的剪接特征的互补区。

[0683] 13.实施方案12所述的复合物,其中所述剪接特征是所述DMD前体mRNA的外显子剪接增强子(ESE)。

[0684] 14.实施方案13所述的复合物,其中所述剪接特征在所述DMD前体mRNA的外显子51中,任选地其中所述ESE包含SEQ ID NO:860至894中任一者的序列。

[0685] 15.实施方案12所述的复合物,其中所述剪接特征是分支点、剪接供体位点或剪接受体位点。

[0686] 16.实施方案15所述的复合物,其中所述剪接特征跨越所述DMD前体mRNA的外显子

50与内含子50的连接处、在其内含子50中、跨越其内含子50与外显子51的连接处、跨越其外显子51与内含子51的连接处、在其内含子51中、或跨越其内含子51与外显子52的连接处,任选地其中所述剪接特征包含SEQ ID NO:855至8595和895至898中任一者的序列。

[0687] 17.实施方案12至16中任一项所述的复合物,其中所述互补区包含与所述剪接特征互补的至少4个连续核苷。

[0688] 18.实施方案1至9中任一项所述的复合物,其中所述分子载荷包含寡核苷酸,所述寡核苷酸包含SEQ ID NO:160至383中任一者的互补序列或包含SEQ ID NO:384至831中任一者的序列,其中每个胸腺嘧啶碱基(T)可独立且任选地用尿嘧啶碱基(U)替换,并且每个U可独立且任选地用T替换。

[0689] 19.实施方案10至18中任一项所述的复合物,其中所述寡核苷酸包含至少一个经修饰核苷间键联。

[0690] 20.实施方案19所述的复合物,其中所述至少一个经修饰核苷间键联是硫代磷酸酯键联。

[0691] 21.实施方案10至20中任一项所述的复合物,其中所述寡核苷酸包含一个或更多个经修饰核苷。

[0692] 22.实施方案21所述的复合物,其中所述一个或更多个经修饰核苷是2'-经修饰核苷。

[0693] 23.实施方案10至18中任一项所述的复合物,其中所述寡核苷酸包含一个或更多个磷酸二酰胺吗啉代,任选地其中所述寡核苷酸是磷酸二酰胺吗啉代寡聚物(PMO)。

[0694] 24.实施方案1至23中任一项所述的复合物,其中所述抗TfR1抗体通过可切割接头与所述分子载荷共价连接。

[0695] 25.实施方案24所述的复合物,其中所述可切割接头包含缬氨酸-瓜氨酸序列。

[0696] 26.实施方案1至25中任一项所述的复合物,其中所述抗TfR1抗体通过与所述抗体的赖氨酸残基或半胱氨酸残基的缀合而与所述分子载荷共价连接。

[0697] 27.复合物,其包含与寡核苷酸共价连接的抗TfR1抗体,所述寡核苷酸被配置成用于诱导DMD前体mRNA中外显子51的跳读,其中所述寡核苷酸包含SEQ ID NO:160至383中任一者的互补区。

[0698] 28.实施方案27所述的复合物,其中所述抗TfR1抗体是表2至7中的任一者中标识的抗体。

[0699] 29.复合物,其包含与寡核苷酸共价连接的抗TfR1抗体,所述寡核苷酸被配置成用于诱导DMD前体mRNA中外显子51的跳读,其中所述寡核苷酸包含所述DMD前体mRNA的剪接特征的互补区。

[0700] 30.靶向DMD的寡核苷酸,其中所述寡核苷酸包含SEQ ID NO:160至383中任一者的互补区。

[0701] 31.实施方案30所述的寡核苷酸,其中所述互补区包含与SEQ ID NO:160至383中任一者互补的至少15个连续核苷。

[0702] 32.实施方案30或31所述的寡核苷酸,其中所述寡核苷酸包含SEQ ID NO:384至831中任一者的至少15个连续核苷,任选地其中所述寡核苷酸包含SEQ ID NO:384至831中任一者的序列,其中每个胸腺嘧啶碱基(T)可独立且任选地用尿嘧啶碱基(U)替换,并且每

个U可独立且任选地用T替换。

[0703] 33.将分子载荷递送至细胞的方法,所述方法包括使所述细胞与实施方案1至26中任一项所述的复合物接触。

[0704] 34.将寡核苷酸递送至细胞的方法,所述方法包括使所述细胞与实施方案27至29中任一项所述的复合物接触。

[0705] 35.促进细胞中肌养蛋白的表达或活性的方法,所述方法包括使所述细胞与有效促进分子载荷内化至所述细胞的量的实施方案1至26中任一项所述的复合物接触,任选地其中所述细胞是肌细胞。

[0706] 36.促进细胞中肌养蛋白的表达或活性的方法,所述方法包括使所述细胞与有效促进寡核苷酸内化至所述细胞的量的实施方案27至29中任一项所述的复合物接触,任选地其中所述细胞是肌细胞。

[0707] 37.实施方案35或36所述的方法,其中所述细胞是体外的。

[0708] 38.实施方案35或36所述的方法,其中所述细胞在对象中。

[0709] 39.实施方案38所述的方法,其中所述对象是人。

[0710] 40.实施方案39所述的方法,其中所述对象具有适合于外显子51跳读的DMD基因。

[0711] 41.实施方案35至40中任一项所述的方法,其中所述肌养蛋白是截短的肌养蛋白。

[0712] 42.治疗具有与肌养蛋白病相关的突变DMD等位基因的对象的方法,所述方法包括向所述对象施用有效量的实施方案1至29中任一项所述的复合物。

[0713] 43.促进细胞中DMD前体mRNA转录物的外显子51跳读的方法,所述方法包括使所述细胞与有效量的实施方案1至29中任一项所述的复合物接触。

[0714] 44.治疗具有与肌养蛋白病相关的突变DMD等位基因的对象的方法,所述方法包括向所述对象施用有效量的实施方案1至29中任一项所述的复合物。

[0715] 实施例

[0716] 实施例1.迪谢内肌营养不良患者肌管中抗TfR1抗体缀合物的外显子跳读活性

[0717] 在本研究中,评价了包含与DMD外显子51跳读反义寡核苷酸(antisense oligonucleotide,ASO)共价连接的抗TfR1 Fab(3M12 VH4/Vκ3)的抗TfR1抗体缀合物的外显子跳读活性。DMD外显子51跳读ASO是长度为30个核苷酸的磷酸二酰胺吗啉代寡聚物(PMO),并且靶向具有序列TGGAGGT(SEQ ID NO:877)的DMD外显子51中的ESE。将在DMD基因中携带外显子52缺失的永生化人成肌细胞解冻并以1e6个细胞/瓶的密度接种在Promocell骨骼细胞生长培养基(含5% FBS和1×Pen-Strep)中,并允许其生长至汇合。一旦汇合,就对细胞进行胰蛋白酶化并通过离心进行沉淀并重悬在新鲜Promocell骨骼细胞生长培养基中。对细胞数进行计数并将细胞以50,000个细胞/孔的密度接种到经基质胶包被的96孔板中。允许细胞恢复24小时。通过抽吸生长培养基并用不含血清的分化培养基替代来诱导细胞分化成肌管。然后将细胞用10μM ASO的DMD外显子51跳读寡核苷酸(未与抗体共价连接——“裸”)或10μM ASO等同物的与DMD外显子51跳读寡核苷酸共价连接的抗TfR1 Fab(3M12VH4/Vκ3)处理。将细胞与测试品一起孵育十天,然后从96孔板收获总RNA。对75ng总RNA进行cDNA合成,并进行突变特异性PCR以评价细胞中外显子51跳读的程度。将突变特异性PCR产物在4%琼脂糖凝胶上运行,并使用SYBR gold使其显现。使用密度测定法来计算跳读和未跳读的扩增子的相对量,并将外显子跳读确定为外显子51跳读的扩增子除以存在的

扩增子总量的比：

[0718] $$外显子跳读\% = \frac{跳读扩增子}{(跳读扩增子 + 未跳读扩增子)} * 100$$

。

[0719] 结果表明,在患者肌管中,与裸DMD外显子51跳读寡核苷酸相比,缀合物导致外显子跳读增强(图1)。这表明抗TfR1 Fab 3M12 VH4/Vκ3使得缀合物能够细胞内化至肌细胞中,导致肌细胞中外显子51跳读寡核苷酸的活性。类似地,抗TfR1抗体(例如,抗TfR1 Fab 3M12 VH4/Vκ3)可使得包含与其他外显子跳读寡核苷酸(例如,本文中提供的外显子跳读寡核苷酸,例如外显子51跳读寡核苷酸)共价连接的抗TfR1抗体的缀合物能够内化至肌细胞中,并促进肌细胞中外显子跳读寡核苷酸的活性。

[0720] 实施例2.在食蟹猴中在体内抗TfR1 Fab-ASO缀合物的外显子跳读活性

[0721] 将抗TfR1 Fab 3M12 VH4/Vκ3与在实施例1中使用的DMD外显子51跳读反义寡核苷酸(ASO)共价连接。在健康的非人灵长类中体内测试缀合物的外显子跳读活性。在第1天和第8天,通过静脉内输注向原初雄性食蟹猴(每组n＝4至5只)施用两个剂量的载剂、30mg/kg裸ASO(即未与抗体共价连接)、或122mg/kg与DMD外显子51跳读寡核苷酸共价连接的抗TfR1 Fab(3M12 VH4/Vκ3)(30mg/kg ASO等同物)。在施用第一个剂量之后2周或4周处死动物并收获组织。使用Promega Maxwell® RSC仪器从组织样品收集总RNA,并使用qScript cDNA SuperMix进行cDNA合成。使用终点PCR进行外显子51跳读的评估。

[0722] PCR产物的毛细管电泳用于评估外显子跳读,并且使用下式计算外显子51跳读%：

[0723] $$外显子跳读\% = \frac{跳读带的摩尔浓度}{跳读带的摩尔浓度+未跳读带的摩尔浓度} * 100$$

。

[0724] 表10中示出了计算的外显子51跳读结果。

[0725] 表10.食蟹猴中DMD mRNA的外显子51跳读

时间	2 周			4 周	
组	载剂	裸 ASO[a]	缀合物	裸 ASO[a]	缀合物
缀合物剂量[b]	0	n/a	122	n/a	122
ASO剂量[c]	0	30	30	30	30

[0726]

四头肌[d]	0.00 (0.00)	1.216 (1.083)	4.906 (3.131)	0.840 (1.169)	1.708 (1.395)
膈肌[d]	0.00 (0.00)	1.891 (2.911)	7.315 (1.532)	0.717 (1.315)	9.225 (4.696)
心脏[d]	0.00 (0.00)	0.043 (0.096)	3.42 (1.192)	0.00 (0.00)	4.525 (1.400)
二头肌[d]	0.00 (0.00)	0.607 (0.615)	3.129 (0.912)	1.214 (1.441)	4.863 (3.881)
胫骨前肌[d]	0.00 (0.00)	0.699 (0.997)	1.042 (0.685)	0.384 (0.615)	0.816 (0.915)
腓肠肌[d]	0.00 (0.00)	0.388 (0.573)	2.424 (2.329)	0.00 (0.00)	5.393 (2.695)

[0727] [a]ASO＝反义寡核苷酸。

[0728] [b]缀合物剂量作为抗TfR1 Fab 3M12 VH4/Vκ3-ASO缀合物的mg/kg列出。

[0729] [c]ASO剂量作为mg/kg的ASO或抗TfR1 Fab 3M12 VH4/Vκ3-ASO剂量的ASO等同物列出。

[0730] [d]外显子跳读值是平均外显子51跳读%,其中括号中为标准偏差(n=5)。

[0731] 还使用杂交ELISA和与ASO序列互补的探针对组织ASO累积进行定量。生成标准曲线并从标准曲线的线性回归导出ASO水平(以ng/g计)。与施用裸ASO相比,在施用抗TfR1 Fab VH4/Vκ3-ASO缀合物之后,ASO以更高水平分布到所有评价的组织。静脉内施用裸ASO导致在施用第一个剂量之后2周和4周在评价的所有组织中的ASO水平接近背景水平。抗TfR1 Fab VH4/Vκ3-ASO缀合物的施用导致在第一次给药之后2周,ASO以以下排列顺序分布在所评价的组织中:心脏>膈肌>二头肌>四头肌>腓肠肌>胫骨前肌。还评估了组织浓度的持续时间。在评价的时间段(2至4周)内,四头肌、二头肌和膈肌中的ASO浓度下降小于50%,而心脏、胫骨前肌和腓肠肌中的ASO水平几乎保持不变(表11)。这表明抗TfR1 Fab 3M12 VH4/Vκ3使得缀合物能够在体内细胞内化至肌细胞中,导致肌细胞中外显子跳读寡核苷酸的活性。类似地,在体内抗TfR1抗体(例如,抗TfR1 Fab 3M12 VH4/Vκ3)可使包含与其他外显子跳读寡核苷酸(例如,本文中提供的外显子跳读寡核苷酸,例如外显子51跳读寡核苷酸)共价连接的抗TfR1抗体的缀合物能够内化至肌细胞中,并促进肌细胞中外显子跳读寡核苷酸的活性。

[0732] 表11.食蟹猴中DMD外显子51跳读ASO的组织分布

时间	2 周			4 周	
组	载剂	裸 ASO[a]	缀合物	裸 ASO[a]	缀合物
缀合物剂量[b]	0	n/a	122	n/a	122
ASO 剂量[c]	0	30	30	30	30
四头肌[d]	0 (59.05)	696.8 (868.15)	2436 (954.0)	197 (134)	682 (281)

[0733]

膈肌[d]	0± (144.3)	580.02 (360.11)	6750 (2256)	60 (120)	3131 (1618)
心脏[d]	0 (396.03)	1449 (1337)	27138 (6315)	943 (1803)	30410 (9247)
二头肌[d]	0 (69.58)	615.63 (335.17)	2840 (980.31)	130 (80)	1326 (623)
胫骨前肌[d]	0 (76.31)	564.71 (327.88)	1591 (253.50)	169 (110)	1087 (514)
腓肠肌[d]	0 (41.15)	705.47 (863.75)	2096 (474.04)	170 (69)	1265 (272)

[0734] [a]ASO=反义寡核苷酸。

[0735] [b]缀合物剂量作为抗TfR1 Fab 3M12 VH4/Vκ3ASO缀合物的mg/kg列出。

[0736] [c]ASO剂量作为mg/kg的ASO或抗TfR1 Fab 3M12 VH4/Vκ3-ASO缀合物剂量的ASO等同物列出。

[0737] [d]ASO值是组织中ASO的平均浓度(以ng/g计),其中括号中为标准偏差(n=5)。

[0738] 等同方案和术语

[0739] 本文中举例说明性地描述的公开内容可在不存在本文中未具体公开的任何一个或更多个要素、一个或更多个限制的情况下适当地实践。因此,例如,在本文中的每种情况下,术语“包含/包括”、“基本上由.....组成”和“由.....组成”中的任一个可用其他两个术语中的任一个替换。已采用的术语和表达作为描述而非限制的术语使用,并且使用这样的术语和表达不旨在排除所示出和所描述的特征的任何等同形式或其一部分,而是应认识到,在所公开内容的范围内可进行多种修改。因此,应理解,尽管已通过一些优选实施方案、任选的特征具体公开了本公开内容,但是本领域技术人员可获取本文中所公开概念的修改和变化,并且这样的修改和变化被认为是在本公开内容的范围内。

[0740] 另外,在根据马库什组(Markush group)或其他替代组描述本公开内容的特征或方面的情况下,本领域技术人员将认识到,本公开内容也因此以马库什组或其他组的任何个体成员或成员亚组的方式描述。

[0741] 应理解,在一些实施方案中,在描述寡核苷酸或其他核酸的结构时可参考序列表中所示的序列。在这样的实施方案中,实际的寡核苷酸或其他核酸与指定序列相比可具有一个或更多个替代核苷酸或核苷(例如,DNA核苷的RNA对应物或RNA核苷的DNA对应物)和/或者(例如,和)一个或更多个经修饰核苷酸/核苷和/或者(例如,和)一个或更多个经修饰核苷间键联和/或者(例如,和)一个或更多个其他修饰,同时保留与指定序列基本相同或相似的互补特性。

[0742] 除非在本文中另外指明或与上下文明显矛盾,否则在描述本发明的上下文中(尤其是在所附权利要求书的上下文中)使用没有数量词修饰的名词将被解释为一个/种或更多个/种。除非另有说明,否则术语“包含”、“具有”、“包括”和“含有”将被解释为开放式术语(即,意指“包括但不限于”)。除非本文中另外指明,否则本文中值范围的记载仅旨在用作分别指代落入该范围内的每个单独值的速记方法,并且每个单独值均被并入说明书中,如同其在本文中被单独记载一样。除非在本文中另外指明或在其他情况下与上下文明显矛盾,否则本文中所述的所有方法均可以以任何合适的顺序进行。除非另有说明,否则本文中提供的任何和所有实例或示例性语言(如“例如”)的使用仅仅旨在更好地说明本发明,并且不对本发明的范围构成限制。说明书中的语言均不应被解释为表示对本发明的实践必要的任何未要求保护的要素。

[0743] 本文中描述了本发明的一些实施方案。在阅读前述说明之后,那些实施方案的变化方案对于本领域普通技术人员可变得明显。

[0744] 本发明人预期技术人员在适当时采用这样的变化方案,并且本发明人希望以除本文中具体描述的之外的方式实践本发明。因此,如适用法律所允许的,本发明包括在此所附权利要求书中记载的主题的所有修改方案和等同方案。此外,除非在本文中另外指明或在其他情况下与上下文明显矛盾,否则本发明涵盖其所有可能变化方案中的上述要素的任何组合。本领域技术人员将认识到或仅使用常规实验就能够确定本文中所述的本发明的具体实施方案的许多等同方案。这样的等同方案旨在由所附权利要求书涵盖。

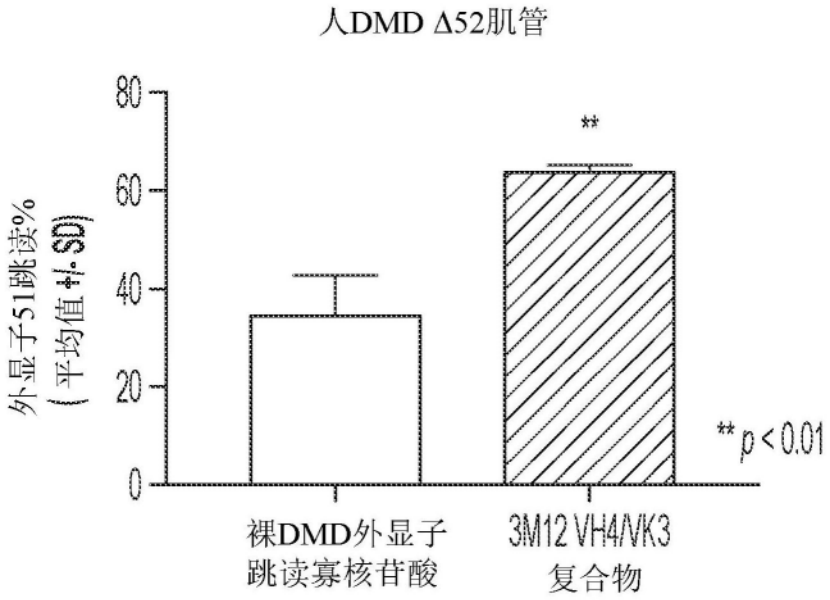

人DMD Δ52肌管
外显子51跳读%
(平均值 +/- SD)
80
60
40
20
0
**
**p < 0.01
裸DMD外显子
跳读寡核苷酸
3M12 VH4/VK3
复合物

图1