I872078

(19)中華民國智慧財產局

(12)發明說明書公告本 **(11)證書號數：TW I872078 B**

(45)公告日： 中華民國 114 (2025) 年 02 月 11 日

(21)申請案號：109118224 (22)申請日： 中華民國 109 (2020) 年 05 月 29 日

(51)Int. Cl.： ***C12N15/79 (2006.01)*** ***A61P25/00 (2006.01)***

(30)優先權：2019/05/29 美國 62/854,238

2019/06/05 美國 62/857,727

2020/04/10 美國 63/008,569

(71)申請人：美商編碼製藥公司 (美國) ENCODED THERAPEUTICS, INC. (US)
美國

(72)發明人：德萊特拉 史蒂芬妮 TAGLIATELA, STEPHANIE (US)；坦涅浩斯 安妮 TANENHAUS, ANNE (US)；瑞瑪穆斯 卡提克 RAMAMOORTHI, KARTIK (US)；楊 安卓 YOUNG, ANDREW (US)；歐博克夫勒 大衛 OBERKOFLER, DAVID (US)

(74)代理人：陳長文

(56)參考文獻：

JP 2018-088888A WO 2017/180915A2

WO 2018/049079A1

審查人員：張茜毓

申請專利範圍項數：31 項 圖式數：13 共 505 頁

(54)名稱

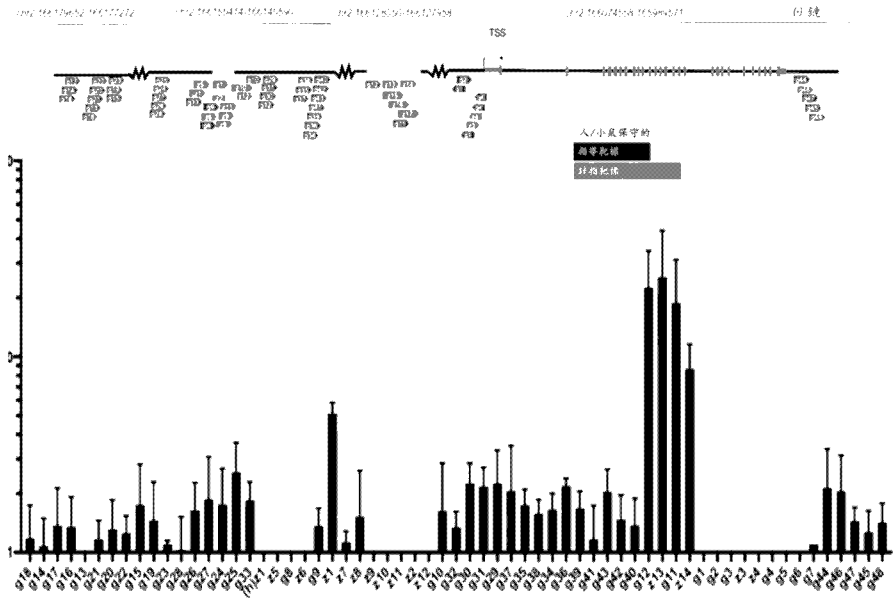
用於選擇性基因調節的組合物和方法

(57)摘要

本文提供了用於選擇性上調 SCN1a 的工程化轉錄因子及其在治療疾病和病症如 Dravet 綜合症中的用途。還提供了微小 RNA 結合位點及其用於在小白蛋白神經元中選擇性表達的用途。

Provided herein are engineered transcription factors for selective upregulation of SCN1a and uses thereof for treating diseases and disorders, such as, Dravet syndrome. Also provided are microRNA binding sites and uses thereof for selective expression in parvalbumin neurons.

指定代表圖：



不同靶位點

【圖1】

公告本

I872078

【發明摘要】

【中文發明名稱】

用於選擇性基因調節的組合物和方法

【英文發明名稱】

COMPOSITIONS AND METHODS FOR SELECTIVE GENE REGULATION

【中文】

本文提供了用於選擇性上調SCN1a的工程化轉錄因子及其在治療疾病和病症如Dravet綜合症中的用途。還提供了微小RNA結合位點及其用於在小白蛋白神經元中選擇性表達的用途。

【英文】

Provided herein are engineered transcription factors for selective upregulation of SCN1a and uses thereof for treating diseases and disorders, such as, Dravet syndrome. Also provided are microRNA binding sites and uses thereof for selective expression in parvalbumin neurons.

【指定代表圖】

圖1

【代表圖之符號簡單說明】

無

【發明說明書】

【中文發明名稱】

用於選擇性基因調節的組合物和方法

【英文發明名稱】

COMPOSITIONS AND METHODS FOR SELECTIVE GENE REGULATION

【技術領域】

【先前技術】

【0001】 眾多的人類疾病與基因的異常表達相關。在一些情況下，基因中的遺傳突變會導致其調節異常、下調或根本不表達，從而導致單倍性不足。在一些情況下，基因中的遺傳突變會導致其上調，從而導致基因的過表達。在治療遺傳性病症或疾病中存在許多挑戰。一種方法是基因療法，它涉及將核酸治療性地遞送到患者細胞中。然而，與基因療法相關的各種挑戰仍未得到解決，如基因療法引發的不希望的免疫應答、脫靶效應、對基因治療媒介物(例如病毒)克隆能力的限制、在較長時段內維持治療效果等。中樞神經系統(CNS)對開發解決基因和/或蛋白質表達潛在損害的療法提出了許多獨特的挑戰。雖然有一些藥物可以幫助控制CNS疾病/病症的症狀，但許多CNS疾病/病症，例如Dravet綜合症，缺乏特定的治療或治癒方法。因此，需要能夠調節任何內源基因的表達以幫助逆轉疾病或病症的影響的新組合物和方法，特別是具有降低的免疫原性、降低的脫靶效應、增加的靶基因特異性和/或增加的治療功效的療法。

【發明內容】

【0002】 在一個方面，本申請提供了一種表達盒，其包含編碼增加細胞

中SCN1A基因表達的非天然存在的轉錄因子的序列，其中所述非天然存在的轉錄因子包含以如下方式與至少兩個轉錄啟動域(TAD)可操作地連接的DNA結合域(DBD)：TAD1-TAD2-DBD、DBD-TAD3-TAD4或TAD1-TAD2-DBD-TAD3-TAD4。在某些實施方案中，TAD1、TAD2、TAD3和TAD4獨立地選自以下：VP16、VP64、Viper、CITED2、CITED4、CREB3或其功能片段。在某些實施方案中，TAD1和TAD2是相同的TAD。在某些實施方案中，TAD1和TAD2是CITED2或其功能片段。在某些實施方案中，TAD1和TAD2是CITED4或其功能片段。在某些實施方案中，TAD3和TAD4是相同的TAD。在某些實施方案中，TAD3和TAD4是CITED2或其功能片段。在某些實施方案中，TAD3和TAD4是CITED4或其功能片段。在某些實施方案中，TAD1、TAD2、TAD3和TAD4是相同的TAD。在某些實施方案中，TAD1、TAD2、TAD3和TAD4是CITED2或其功能片段。在某些實施方案中，TAD1、TAD2、TAD3和TAD4是CITED4或其功能片段。

【0003】 在某些實施方案中，在所述至少兩個TAD結構域之間沒有連接體。

【0004】 在某些實施方案中，在所述至少兩個TAD結構域之間存在連接體。在某些實施方案中，所述連接體包含GGSGGGSG(SEQ ID NO: 177)或GGSGGGSGGGSGGGSG(SEQ ID NO: 178)或由其組成。

【0005】 在某些實施方案中，所述DBD與具有18-27個核苷酸的基因組區域結合。

【0006】 在某些實施方案中，所述DBD與其最接近的人類對應物具有至少80%的序列同一性。在某些實施方案中，所述DBD與其最接近的人類對應物具有至少90%的序列同一性。在某些實施方案中，所述DBD和所

述至少兩個TAD各自與其最接近的人類對應物具有至少80%的序列同一性。在某些實施方案中，所述DBD和所述至少兩個TAD各自與其最接近的人類對應物具有至少90%的序列同一性。

【0007】 在某些實施方案中，所述DBD包含指導RNA和核酸酶滅活的Cas蛋白。在某些實施方案中，所述核酸酶滅活的Cas蛋白是核酸酶滅活的Cas9。

【0008】 在某些實施方案中，所述DBD包含鋅指結構域。在某些實施方案中，所述DBD包含6至9個鋅指結構域。在某些實施方案中，所述DBD包含6個鋅指。在某些實施方案中，所述DBD與具有18個核苷酸的基因組區域結合。在某些實施方案中，所述DBD包含9個鋅指。在某些實施方案中，所述DBD與具有27個核苷酸的基因組區域結合。

【0009】 在某些實施方案中，所述DBD包含與SEQ ID NO: 148-151中的任一個具有至少95%序列同一性的序列。在某些實施方案中，所述DBD包含具有SEQ ID NO: 148-151中的任一個的序列。

【0010】 在某些實施方案中，所述DBD衍生自人EGR1或人EGR3。

【0011】 在某些實施方案中，所述DBD包含與SEQ ID NO: 77-98中的任一個具有至少90%同一性的序列。在某些實施方案中，所述DBD包含SEQ ID NO: 77-98。

【0012】 在某些實施方案中，所述DBD包含與SEQ ID NO: 92具有至少90%同一性的序列。在某些實施方案中，所述DBD包含SEQ ID NO: 92。

【0013】 在某些實施方案中，所述非天然存在的轉錄因子包含與SEQ ID NO: 130或131具有至少90%同一性的序列。在某些實施方案中，

所述非天然存在的轉錄因子包含SEQ ID NO: 130或131。

【0014】 在某些實施方案中，所述表達盒包含與SEQ ID NO: 72或73中的任一個具有至少90%同一性的核苷酸序列。在某些實施方案中，所述表達盒包含SEQ ID NO: 72或73中的任一個的核苷酸序列。

【0015】 在某些實施方案中，所述表達盒進一步包含調節元件，該調節元件在PV神經元中以比在其他細胞類型中更高的水準驅動所述轉錄因子的表達。在某些實施方案中，所述調節元件包含SEQ ID NO: 1-4中的任一個。在某些實施方案中，所述調節元件包含SEQ ID NO: 2或3。

【0016】 在某些實施方案中，所述表達盒進一步包含PV選擇性微小RNA結合位點。在某些實施方案中，所述PV選擇性微小RNA結合位點與SEQ ID NO: 7、14或15中的任一個具有至少90%的同一性。在某些實施方案中，所述PV選擇性微小RNA結合位點包含SEQ ID NO: 7、14或15中的任一個。

【0017】 在某些實施方案中，所述表達盒是病毒載體的一部分。在某些實施方案中，所述病毒載體是AAV病毒。在某些實施方案中，所述AAV病毒是AAV9病毒或scAAV9病毒。在某些實施方案中，所述病毒載體是慢病毒。

【0018】 在另一方面，本申請提供了一種表達盒，其包含編碼增加細胞中SCN1A基因表達的非天然存在的轉錄因子的序列，其中所述非天然存在的轉錄因子包含與轉錄啟動域可操作地連接的DNA結合域，其中所述DNA結合域是包含序列LEPGEKP – [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH – TGKKTS(SEQ ID NO: 147)的鋅指蛋白，並且其中所述DNA結合域與所述轉錄啟動域之間沒有HA標籤

(SEQ ID NO: 303)。在某些實施方案中，所述轉錄啟動域包含VP16、VPR或VP64序列，或其功能片段。在某些實施方案中，所述轉錄啟動域包含VP64。

【0019】 在某些實施方案中，所述DNA結合域與具有18-27個核苷酸的基因組區域結合。在某些實施方案中，所述DNA結合域是包含SEQ ID NO: 147的鋅指結構域，其中n = 6至9。在某些實施方案中，所述DNA結合域是包含SEQ ID NO: 147的鋅指結構域，其中n = 6。在某些實施方案中，所述DNA結合域與具有18個核苷酸的基因組區域結合。在某些實施方案中，所述DNA結合域是包含SEQ ID NO: 147的鋅指結構域，其中n = 9。在某些實施方案中，所述DNA結合域與具有27個核苷酸的基因組區域結合。

【0020】 在某些實施方案中，所述DNA結合域包含與SEQ ID NO: 148-151中的任一個具有至少95%序列同一性的序列。在某些實施方案中，所述DNA結合域包含具有SEQ ID NO: 148-151中的任一個的序列。

【0021】 在某些實施方案中，所述DNA結合域包含與SEQ ID NO: 77-91中的任一個具有至少90%同一性的序列。在某些實施方案中，所述DNA結合域包含SEQ ID NO: 77-91中的任一個。

【0022】 在某些實施方案中，所述表達盒進一步包含調節元件，該調節元件在PV神經元中以比在其他細胞類型中更高的水準驅動所述轉錄因子的表達。在某些實施方案中，所述調節元件包含SEQ ID NO: 1-4中的任一個。在某些實施方案中，所述調節元件包含SEQ ID NO: 2或3。

【0023】 在某些實施方案中，所述非天然存在的轉錄因子包含與SEQ ID NO: 127具有至少90%同一性的序列。在某些實施方案中，所述

非天然存在的轉錄因子包含SEQ ID NO: 127。

【0024】在某些實施方案中，所述表達盒包含與SEQ ID NO: 93或71中的任一個具有至少90%同一性的核苷酸序列。在某些實施方案中，所述表達盒包含SEQ ID NO: 93或71中的任一個的核苷酸序列。

【0025】在某些實施方案中，所述表達盒進一步包含PV選擇性微小RNA結合位點。在某些實施方案中，所述PV選擇性微小RNA結合位點與SEQ ID NO: 7、14或15中的任一個具有至少90%的同一性。在某些實施方案中，所述PV選擇性微小RNA結合位點包含SEQ ID NO: 7、14或15中的任一個。

【0026】在某些實施方案中，所述表達盒是病毒載體的一部分。在某些實施方案中，所述病毒載體是AAV病毒。在某些實施方案中，所述AAV病毒是AAV9病毒或scAAV9病毒。在某些實施方案中，所述病毒載體是慢病毒。

在另一方面，本申請提供了一種包含PV選擇性微小RNA結合位點的多核苷酸，該PV選擇性微小RNA結合位點包含與SEQ ID NO: 14或15具有至少80%序列同一性的序列，其中所述微小RNA結合位點降低轉基因在興奮性神經元中的表達。在某些實施方案中，所述PV選擇性微小RNA結合位點包含SEQ ID NO: 14。在某些實施方案中，所述PV選擇性微小RNA結合位點包含SEQ ID NO: 15。在另一方面，本申請提供了一種包含PV選擇性微小RNA結合位點以及啟動子和/或增強子的表達盒。在某些實施方案中，所述啟動子和/或增強子是PV選擇性調節元件，該PV選擇性調節元件在小白蛋白(PV)神經元中以比在其他細胞類型中更高的水準驅動轉基因的表達。在某些實施方案中，所述PV選擇性調節元件與轉基因可操

作地連接。

【0027】在另一方面，本申請提供了一種表達盒，其包含與轉基因可操作地連接的調節元件和至少一個微小RNA結合位點，其中所述調節元件在小白蛋白(PV)神經元中以比在其他細胞類型中更高的水準驅動所述轉基因的表達，並且其中所述微小RNA結合位點降低所述轉基因在興奮性神經元中的表達。在某些實施方案中，所述表達盒不包含SEQ ID NO: 67。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR128的結合位點(SEQ ID NO: 9)。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR221的結合位點(SEQ ID NO: 11)。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR222的結合位點(SEQ ID NO: 13)。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR128的結合位點(SEQ ID NO: 9)和至少一個針對MIR221的結合位點(SEQ ID NO: 11)。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR128的結合位點(SEQ ID NO: 9)、至少一個針對MIR221的結合位點(SEQ ID NO: 11)和至少一個針對MIR222的結合位點(SEQ ID NO: 13)。在某些實施方案中，所述微小RNA結合位點包含與SEQ ID NO: 7、14或15中的任一個具有至少90%同一性的序列。在某些實施方案中，所述微小RNA結合位點包含SEQ ID NO: 7、14或15。

【0028】在某些實施方案中，所述轉基因編碼一種多肽，該多肽包含增加細胞中SCN1A基因表達的非天然存在的轉錄因子。在某些實施方案中，所述轉錄因子與具有18-27個核苷酸的基因組區域結合。在某些實施方案中，所述轉錄因子包含DNA結合域。在某些實施方案中，所述轉

錄因子包含DNA結合域和轉錄啟動域。

【0029】 在某些實施方案中，所述DNA結合域與其最接近的人類對應物具有至少80%的序列同一性。在某些實施方案中，所述DNA結合域與其最接近的人類對應物具有至少90%的序列同一性。在某些實施方案中，所述DNA結合域和轉錄啟動域均與其最接近的人類對應物具有至少80%的序列同一性。在某些實施方案中，所述DNA結合域和轉錄啟動域均與其最接近的人類對應物具有至少90%的序列同一性。

【0030】 在某些實施方案中，所述DNA結合域包含指導RNA和核酸酶滅活的Cas蛋白。在某些實施方案中，所述核酸酶滅活的Cas蛋白是核酸酶滅活的Cas9。

【0031】 在某些實施方案中，所述DNA結合域包含鋅指結構域。在某些實施方案中，所述DNA結合域包含6至9個鋅指結構域。在某些實施方案中，所述DNA結合域包含6個鋅指。在某些實施方案中，所述DNA結合域與具有18個核苷酸的基因組區域結合。在某些實施方案中，所述DNA結合域包含9個鋅指。在某些實施方案中，所述DNA結合域與具有27個核苷酸的基因組區域結合。

【0032】 在某些實施方案中，所述DNA結合域包含與SEQ ID NO: 148-151中的任一個具有至少95%序列同一性的序列。在某些實施方案中，所述DNA結合域包含具有SEQ ID NO: 148-151中的任一個的序列。在某些實施方案中，所述DNA結合域包含與SEQ ID NO: 92-98中的任一個具有至少90%同一性的序列。在某些實施方案中，所述DNA結合域包含SEQ ID NO: 92-98中的任一個。

【0033】 在某些實施方案中，所述DNA結合域是包含序列

LEPGEKP – [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH – TGKKTS(SEQ ID NO: 147)的鋅指蛋白。

【0034】在某些實施方案中，所述DNA結合域包含與SEQ ID NO: 77-91中的任一個具有至少90%同一性的序列。在某些實施方案中，所述DNA結合域包含SEQ ID NO: 77-91中的任一個。

【0035】在某些實施方案中，所述DNA結合域衍生自人EGR1或人EGR3。

【0036】在某些實施方案中，所述轉錄啟動域包含VP16、VPR、VP64、CITED2、CITED4或CREB3序列，或其功能片段。在某些實施方案中，所述轉錄啟動域包含人CITED2、CITED4或CREB3序列，或其功能片段。

【0037】在某些實施方案中，所述調節元件包含具有SEQ ID NO: 1-4中的任一個的序列。在某些實施方案中，所述調節元件包含具有SEQ ID NO: 2或3的序列。

【0038】在某些實施方案中，所述非天然存在的轉錄因子包含與SEQ ID NO: 105、106和127-129中的任一個具有至少90%同一性的序列。在某些實施方案中，所述非天然存在的轉錄因子包含SEQ ID NO: 105、106和127-129中的任一個。

【0039】在某些實施方案中，所述轉基因包含與SEQ ID NO: 71、74、75、76或184中的任一個具有至少90%同一性的核苷酸序列。在某些實施方案中，所述轉基因包含SEQ ID NO: 71、74、75、76或184中的任一個。

【0040】在某些實施方案中，所述表達盒是病毒載體的一部分。在

某些實施方案中，所述病毒載體是AAV病毒。在某些實施方案中，AAV病毒是AAV9病毒或scAAV9病毒。在某些實施方案中，所述病毒載體是慢病毒。

【0041】 在另一方面，本申請提供了一種在靈長類動物的小白蛋白(PV)神經元中選擇性表達轉基因的方法，其包括向靈長類動物施用包含轉基因和至少一個微小RNA結合位元元點的病毒載體，其中所述微小RNA結合位點降低所述轉基因在興奮性神經元中的表達。

【0042】 在某些實施方案中，所述病毒載體進一步包含與所述轉基因可操作地連接的調節元件，其中所述調節元件在小白蛋白(PV)神經元中以比在其他細胞類型中更高的水準驅動所述轉基因的表達。

【0043】 在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR128的結合位點(SEQ ID NO: 9)。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR221的結合位點(SEQ ID NO: 11)。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR222的結合位點(SEQ ID NO: 13)。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR128的結合位點(SEQ ID NO: 9)和至少一個針對MIR221的結合位點(SEQ ID NO: 11)。在某些實施方案中，所述微小RNA結合位點包含至少一個針對MIR128的結合位點(SEQ ID NO: 9)、至少一個針對MIR221的結合位點(SEQ ID NO: 11)和至少一個針對MIR222的結合位點(SEQ ID NO: 13)。在某些實施方案中，所述微小RNA結合位點包含與SEQ ID NO: 7、14或15中的任一個具有至少90%同一性的序列。在某些實施方案中，所述微小RNA結合位點包含SEQ ID NO: 7、14或15。

【0044】在某些實施方案中，所述轉基因包含編碼非天然存在的轉錄因子的序列，該非天然存在的轉錄因子增加細胞中SCN1A基因的表達。

【0045】在某些實施方案中，所述轉錄因子與具有18-27個核苷酸的基因組區域結合。

【0046】在某些實施方案中，所述轉錄因子包含DNA結合域。

【0047】在某些實施方案中，所述轉錄因子包含DNA結合域和轉錄啟動域。

【0048】在某些實施方案中，所述DNA結合域與其最接近的人類對應物具有至少80%的序列同一性。在某些實施方案中，所述DNA結合域與其最接近的人類對應物具有至少90%的序列同一性。在某些實施方案中，所述DNA結合域和轉錄啟動域均與其最接近的人類對應物具有至少80%的序列同一性。在某些實施方案中，所述DNA結合域和轉錄啟動域均與其最接近的人類對應物具有至少90%的序列同一性。

【0049】在某些實施方案中，所述DNA結合域包含指導RNA和核酸酶滅活的Cas蛋白。在某些實施方案中，所述核酸酶滅活的Cas蛋白是核酸酶滅活的Cas9。

【0050】在某些實施方案中，所述DNA結合域包含鋅指結構域。在某些實施方案中，所述DNA結合域包含6至9個鋅指結構域。在某些實施方案中，所述DNA結合域包含6個鋅指。在某些實施方案中，所述DNA結合域與具有18個核苷酸的基因組區域結合。在某些實施方案中，所述DNA結合域包含9個鋅指。在某些實施方案中，所述DNA結合域與具有27個核苷酸的基因組區域結合。

【0051】在某些實施方案中，所述DNA結合域包含與SEQ ID NO: 148-151中的任一個具有至少95%序列同一性的序列。在某些實施方案中，所述DNA結合域包含具有SEQ ID NO: 148-151中的任一個的序列。

【0052】在某些實施方案中，所述DNA結合域衍生自人EGR1或人EGR3。

【0053】在某些實施方案中，所述DNA結合域包含與SEQ ID NO: 92-98中的任一個具有至少90%同一性的序列。在某些實施方案中，所述DNA結合域包含SEQ ID NO: 92-98中的任一個。

【0054】在某些實施方案中，所述DNA結合域是包含序列LEPGEKP – [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH – TGKKTS(SEQ ID NO: 147)的鋅指蛋白。

【0055】在某些實施方案中，所述DNA結合域包含與SEQ ID NO: 77-91中的任一個具有至少90%同一性的序列。在某些實施方案中，所述DNA結合域包含SEQ ID NO: 77-91中的任一個。

【0056】在某些實施方案中，所述轉錄啟動域包含VP16、VPR、VP64、CITED2、CITED4或CREB3序列，或其功能片段。在某些實施方案中，所述轉錄啟動域包含人CITED2、CITED4或CREB3序列，或其功能片段。

【0057】在某些實施方案中，所述調節元件包含具有SEQ ID NO: 1-4中的任一個的序列。在某些實施方案中，所述調節元件包含具有SEQ ID NO: 2或3的序列。

【0058】在某些實施方案中，所述非天然存在的轉錄因子包含與SEQ ID NO: 105、106和127-129中的任一個具有至少90%同一性的序

列。在某些實施方案中，所述非天然存在的轉錄因子包含SEQ ID NO: 105、106和127-129中的任一個。

【0059】 在某些實施方案中，所述轉基因包含與SEQ ID NO: 71、74、75、76或184中的任一個具有至少90%同一性的核苷酸序列。在某些實施方案中，所述轉基因包含SEQ ID NO: 71、74、75、76或184中的任一個。

【0060】 在某些實施方案中，所述病毒載體是AAV病毒。在某些實施方案中，所述AAV病毒是AAV9病毒或scAAV9病毒。在某些實施方案中，所述病毒載體是慢病毒。

【0061】 在某些實施方案中，所述靈長類動物是人。在某些實施方案中，所述靈長類動物為非人靈長類動物。在某些實施方案中，所述非人靈長類動物是舊大陸猴、猩猩、大猩猩、黑猩猩、狨猴、食蟹猴、恒河猴或豬尾猴。

【0062】 在另一方面，本申請提供了一種表達盒，其包含編碼增加細胞中SCN1A基因表達的非天然存在的轉錄因子的序列，其中所述非天然存在的轉錄因子包含與SEQ ID NO: 128或129具有至少90%同一性的序列。在某些實施方案中，所述非天然存在的轉錄因子包含SEQ ID NO: 128或129。

【0063】 在另一方面，本申請提供了一種通過施用本文提供的任何表達盒來增加細胞中SCN1A的表達的方法。在某些實施方案中，所述細胞是神經元細胞。在某些實施方案中，所述神經元細胞選自單極、雙極、多極或假單極神經元。在某些實施方案中，所述細胞是GABA能神經元。在某些實施方案中，所述細胞是PV神經元。在某些實施方案中，所述細

胞為非神經元細胞。在某些實施方案中，所述細胞是神經膠質細胞。在某些實施方案中，所述神經膠質細胞選自星形膠質細胞、少突膠質細胞、室管膜細胞、許旺細胞和衛星細胞。在某些實施方案中，所述細胞在受試者內。在某些實施方案中，所述受試者是哺乳動物。在某些實施方案中，所述受試者是人。在某些實施方案中，增加SCN1A的表達治療疾病、病症或症狀。在某些實施方案中，所述病症是中樞神經系統病症。在某些實施方案中，所述病症是與SCN1A單倍性不足相關的癲癇。在某些實施方案中，所述單倍性不足是受試者對於SCN1A基因的功能喪失突變而言為雜合的結果。在某些實施方案中，所述病症是與SCN1A基因中的插入、缺失或置換相關的癲癇。在某些實施方案中，所述病症是與SCN1A基因中的點突變相關的癲癇。在某些實施方案中，所述病症是Dravet綜合症。在某些實施方案中，所述中樞神經系統病症的症狀是神經元活動過度。在某些實施方案中，治療所述中樞神經系統病症包括減輕神經元活動過度。在某些實施方案中，所述中樞神經系統病症的症狀是癲癇發作。在某些實施方案中，治療所述中樞神經系統病症包括降低癲癇發作的頻率。在某些實施方案中，治療所述中樞神經系統病症包括降低癲癇發作的嚴重程度。

【0064】在另一方面，本申請提供了一種通過施用本文提供的任一種表達盒來增加CNS中SCN1A的表達的方法。在某些實施方案中，所述表達盒通過單側腦室內(ICV)給藥來施用。在某些實施方案中，所述表達盒通過雙側腦室內(ICV)給藥來施用。在某些實施方案中，所述增加的SCN1A表達在腦中發生。在某些實施方案中，所述增加的SCN1A表達在額葉皮質、頂葉皮質、顳葉皮質、海馬體、髓質和/或枕葉皮質中發生。在某些實施方案中，所述增加的SCN1A表達在脊柱中發生。在某些實施

方案中，所述增加的SCN1A表達在脊髓和/或背根神經節中發生。

援引併入

【0065】 本說明書中所提及的所有出版物、專利和專利申請均通過引用併入本文，其程度猶如具體地且單獨地指出每個單獨的出版物、專利或專利申請均通過引用而併入。

【圖式簡單說明】

【0066】 本發明的新穎特徵在所附申請專利範圍中具體闡述。通過參考以下對利用本發明原理的說明性情況加以闡述的詳細描述以及附圖，將會對本發明的特徵和優點獲得更好的理解，在這些附圖中：

【0067】 圖1示出了使用與染色體2上的各個區域(參照GRCh38.p12)結合的工程化轉錄因子，內源SCN1A的上調。資料以SCN1A表達相對於對照(EGFP-KASH)條件的變化倍數呈現。

【0068】 圖2A、圖2B和圖2C示出了使用SCN1A特異性轉錄啟動物(見表1)，HEK293細胞中內源SCN1A的相對表達。資料以相對於對照條件的變化倍數呈現，並在Log_{10}標度上顯示。

【0069】 圖3A示出了使用SCN1A特異性轉錄啟動物(構建體30)，GABA神經元中內源SCN1A的相對表達。資料以相對於對照條件(CBA-EGFP)的變化倍數呈現。

【0070】 圖3B示出了使用SCN1A特異性轉錄啟動物(構建體25和16)，GABA神經元中內源SCN1A的相對表達。資料以在Log_{10}上相對於對照條件(CBA-EGFP)的變化倍數呈現。

【0071】 圖4示出了由SCN1A特異性轉錄因子(構建體30)驅動的內源SCN1A和40個最近鄰基因的相對表達。資料以在Log_{10}上相對於對照條

件(CBA-EGFP-KASH)的變化倍數呈現。

【0072】圖5A和圖5B示出了與表達eGFP的對照表達盒相比，SCN1A特異性轉錄啟動物在體內的表達。圖5A示出了在注射對照eGFP或包含SCN1A轉錄啟動物的構建體4的小鼠中，SCN1A基因的相對表達。圖5B以平均eGFP百分比示出了SCN1A表達的變化。這些實驗表明，由構建體4引起的轉錄啟動導致SCN1A表達上調約20-30%。

【0073】圖6A、圖6B、圖6C、圖6D、圖6E、圖6F和圖6G示出了與對照相比，使用各種SCN1A特異性轉錄因子，在Dravet綜合症的Scn1a^{tm1Kea}敲除小鼠模型中對高熱癲癇發作的影響。向P1 Scn1a +/-小鼠(雜合；HET)輸注AAV9-EGFP或表達SCN1A特異性轉錄因子(構建體31-34、42和43之一)的AAV9載體。在P26-P28，對輸注的小鼠進行高熱誘發的癲癇發作試驗，並記錄其經歷強直-陣攣性癲癇發作的內部溫度。圖6D顯示了在包含位於DBD與TAD之間的HA標籤的構建體32和不包含HA標籤的構建體34之間的直接比較。圖6E顯示了在包含位於編碼區與聚A尾之間的m1微小RNA結合位點的構建體31和不包含該m1微小RNA結合位點的構建體32之間的直接比較。圖6H示出了使用構建體31(與注射PBS的對照相比)，在Dravet綜合症的Scn1a^{RX}突變小鼠模型中對高熱癲癇發作的影響。

【0074】圖7A、圖7B、圖7C和圖7D示出了在各種條件下在Dravet綜合症的Scn1a^{tm1Kea}敲除小鼠模型中的存活率。圖7A示出了在存活試驗中野生型(PBS WT)與Scn1a +/-小鼠(PBS HET)之間的比較。向P1 Scn1a +/-(N=53)和Scn1a +/+(N=54)小鼠輸注PBS。每天在其飼養籠中觀察小鼠，如有任何死亡，則記錄日期。Scn1a +/-與Scn1a +/+動物之間的存活

率存在顯著差異(P<0.0001)。圖7B-D示出了與對照相比，對於用各種SCN1A特異性轉錄因子治療的小鼠，在Dravet綜合症小鼠模型中對存活率的影響。向P1 Scn1a +/-小鼠輸注PBS或表達SCN1A特異性轉錄因子(構建體31或33)的AAV9載體。每天在其飼養籠中觀察小鼠，如有任何死亡，則記錄日期。圖7D顯示了在包含位於編碼區與聚A尾之間的m1微小RNA結合位點的構建體31和不包含該m1微小RNA結合位點的構建體33之間的直接比較。圖7E示出了使用構建體31(與注射PBS的對照相比)，在Dravet綜合症的Scn1a^{RX}突變小鼠模型中的存活率。

【0075】 圖8示出了在實質內遞送編碼SCN1A特異性轉錄因子(構建體33)的AAV9載體後，在不同腦組織中的相對Scn1A mRNA表達，所述載體以1.2x10^{12} gc/動物施用至兩隻食蟹猴，並相對於兩隻未經治療的對照動物進行歸一化。注射後28天處死所有動物，並通過Taqman PCR對組織樣品中的Scn1A mRNA進行定量。資料被報告為目標mRNA在來自腦的不同組織切片中的歸一化表達。用不同的一組Scnla基因衍生的引物/探針也記錄到相似的結果。

【0076】 圖9A-F顯示了用AAV9載體處理後，絨猴海馬體齒狀回區域中EGFP的表達模式，所述AAV9載體包含在EF1a啟動子控制下的EGFP轉基因，在RE 2啟動子(SEQ ID NO: 2)控制下的EGFP轉基因，或在RE 2啟動子(SEQ ID NO: 2)控制下的EGFP轉基因與位於EGFP編碼區與聚A位元點之間的m1微小RNA結合位點(SEQ ID NO: 7)。對於每種載體處理，顯示了海馬體齒狀回區域的代表性區域。頂行顯示了用DAPI染色的細胞核，而底行顯示了用抗GFP抗體染色的GFP陽性區域。在圖9A(EF1a處理)中，海馬體CA4回(hylus)區域以黃色勾勒出輪廓，而箭頭指向齒狀細胞顆

粒細胞體層(DG)。圖9B和圖9C以相同的區域為中心。突出顯示了CA4區——它是興奮性和抑制性中間神經元的混合體，因為它是在RE 2 + m1條件下唯一顯著表達的區域。通過EF1a和RE 2驅動的轉基因表達，GFP表達更加廣泛，並且包括海馬體的其他區域。認為DG細胞層主要包含興奮性神經元。由EF1a和RE 2驅動的GFP表達在DG細胞層中可見(圖9D和圖9E)，但在RE 2 + m1處理的動物中卻不存在(圖9F)(白色箭頭)。

【0077】 圖10A-L顯示，用AAV9載體(包含在EF1a啟動子控制下的EGFP轉基因，在RE 2啟動子(SEQ ID NO: 2)控制下的EGFP轉基因，或在RE 2啟動子(SEQ ID NO: 2)控制下的EGFP轉基因與位於EGFP編碼區與聚A位元點之間的m1微小RNA結合位點(SEQ ID NO: 7))處理後，絨猴海馬體齒狀回區域中EGFP的表達模式主要位元於RE 2和RE 2 + m1處理的動物的小白蛋白(PV)陽性細胞中。對於每種載體處理，顯示了海馬體齒狀回區域的代表性區域。頂行顯示了GFP陽性區域，而下一行顯示了用針對PV的抑制性中間神經元標誌物染色的相同區域。圖10A-F中框選出的區域在圖10G-L中以更高的放大倍數示出。由RE 2和RE 2 + m1驅動的GFP表達主要與抑制性中間神經元標誌物PV共定位(圖10H和10K，10I和10L，白色箭頭)，而在EG-EF1a中，GFP表達不那麼容易定位於PV陽性細胞(圖10G和10J，白色箭頭)。另外，與EF1a處理的動物中較不明顯的細胞體形態學(圖10G，黃色箭頭)相比，在RE 2和RE 2 + m1處理的動物中，GFP陽性細胞具有高度分支細胞的明顯中間神經元形態學，所述細胞具有錐體細胞體(圖10H和10I，黃色箭頭)。

【0078】 圖11顯示了對於通過單側腦室內(ICV)施用以4.8E+13或8E+13 vg/動物施用AAV9-RE^{GABA}-eTF^{SCN1A}而處理的動物，額葉皮質

(FC)、喙頂葉皮質(Rostral PC)、顳葉皮質(TC)、尾頂葉皮質(Caudal PC)、海馬體(Hip)、髓質(Med)和枕葉皮質(OC)組織樣品中的VG/二倍體基因組(實施例10和實施例11)。每個數據點代表組織樣品的VG/二倍體基因組，而水準條代表每只動物的所有組織樣品的平均VG/二倍體基因組。

【0079】 圖12顯示了對於通過單側腦室內(ICV)施用以4.8E+13或8E+13 vg/動物施用AAV9-RE^{GABA}-eTF^{SCN1A}而處理的動物，額葉皮質(FC)、喙頂葉皮質(Rostral PC)、顳葉皮質(TC)、尾頂葉皮質(Caudal PC)、海馬體(Hip)、髓質(Med)和枕葉皮質(OC)組織樣品中的轉錄物/μg RNA(實施例10和實施例11)。每個數據點代表組織樣品的VG/二倍體基因組，而水準條代表每只動物的所有組織樣品的平均VG/二倍體基因組。*ARFGAP2*的平均轉錄物為1.85E+6/μg RNA，並由上邊界虛線表示。檢測極限由下邊界虛線表示。

【0080】 圖13顯示了腦外周圍組織樣品中的載體生物分佈(VG/二倍體基因組)和轉基因表達(轉錄物/μg RNA)。所示的周圍組織樣品是脊髓C2/L4(SC C2/L4)、背根神經節C2/L4(DRG C2/L4)、肝臟、脾臟、心臟、腎臟、肺、胰腺和睾丸/卵巢。靈長類動物腦中的平均VCN(載體生物分佈)和轉錄物(轉基因表達)由虛線表示。

【實施方式】

交叉引用

【0081】 本申請要求2019年5月29日提交的第62/854,238號美國臨時專利申請、2019年6月5日提交的第62/857,727號美國臨時專利申請和2020年4月10日提交的第63/008,569號美國臨時專利申請的權益，所述臨時申請中的每一個均通過引用整體併入本文。

序列表

【0082】本申請含有序列表，該序列表已經以ASCII格式電子提交，並且通過引用整體併入本文。創建於2020年5月28日的所述ASCII副本被命名為46482-724_601_SL.txt，大小為418,483個位元組。

【0083】本文提供了工程化的轉錄因子或eTF，它們是非天然存在的，並且已被設計成與基因組靶位元點結合並調節目的內源基因的表達。可以將這樣的eTF設計為上調或下調目的基因的表達(RNA和/或蛋白質表達)。本文還提供了可以併入病毒載體中的微小RNA結合位點，並且提供了轉基因在小白蛋白(PV)神經元中的選擇性表達。

【0084】在一方面，本申請提供了能夠上調鈉電壓門控通道α亞單位1(SCN1A)基因的表達並增加其相應蛋白質產物Nav1.1的表達的eTF，以及用其治療與Nav1.1缺乏相關的疾病或病症例如Dravet綜合症的方法。

【0085】在另一方面，本申請提供了降低包含微小RNA結合位點的mRNA在興奮性神經元中的表達從而導致該基因在GABA能或小白蛋白(PV)神經元中選擇性表達的微小RNA結合位點，及其在PV神經元中選擇性表達目的基因的使用方法。

定義

【0086】除非上下文另有明確規定，否則如本文所用的單數形式“一個”、“一種”和“該”也意欲包括複數形式。此外，就具體實施方式和/或申請專利範圍中使用的術語“包括”、“包含”、“具有”或其變化形式而言，這些術語旨在以與術語“包含”類似的方式為包含性的。

【0087】術語“約”或“大約”意指在本領域普通技術人員測定的特定值的可接受誤差範圍內，該可接受誤差範圍將部分取決於該值如何測量或

測定，即，測量系統的局限性。例如，根據本領域中的實踐，“約”可指在1個或大於1個標準差內。或者，“約”可指給定值的至多20%、至多15%、至多10%、至多5%或至多1%的範圍。

【0088】 術語“確定”、“測量”、“評價”、“評估”、“測定”、“分析”及其語法等同語在本文中可互換使用，以指代任何測量形式，並且包括確定元素是否存在(例如，檢測)。這些術語可包括定量和/或定性測定。評估可以是相對的或絕對的。

【0089】 術語“表達”是指核酸序列或多核苷酸從DNA範本轉錄(如轉錄成mRNA或其他RNA轉錄物)的過程，和/或所轉錄的mRNA隨後被翻譯成肽、多肽或蛋白質的過程。轉錄物和編碼的多肽可統稱為“基因產物”。如果多核苷酸衍生自基因組DNA，則表達可以包括真核細胞中mRNA的剪接。

【0090】 如本文所用的，“可操作地連接的”、“可操作的連接”、“有效連接的”或其語法等同語是指例如啟動子、增強子、多腺苷酸化序列等遺傳元件並列，其中這些元件之間的關係允許它們以預期的方式發揮作用。例如，如果可包含啟動子和/或增強子序列的調節元件說明啟動編碼序列的轉錄，則該調節元件與編碼區可操作地連接。在調節元件與編碼區之間可以有間插殘基，只要能夠維持這種功能關係即可。

【0091】 如本文所用的“載體”是指包含多核苷酸或與多核苷酸締合，且可用於介導多核苷酸向細胞的遞送的大分子或大分子締合物。載體的實例包括質粒、病毒載體、脂質體和其他基因遞送媒介物。載體通常包含與基因可操作地連接的遺傳元件，例如調節元件，以促進該基因在靶標中的表達。

【0092】 如本文所用的，“表達盒”和“核酸盒”可互換使用，是指一起表達或可操作地連接以供表達的核酸序列或元件的組合。在一些情況下，表達盒是指調節元件和與其可操作地連接以供表達的一個或多個基因的組合。

【0093】 術語“AAV”是腺伴隨病毒的縮寫，可用來指代該病毒本身或其衍生物。該術語涵蓋所有血清型、亞型以及天然存在形式和重組形式，除非另有要求。縮寫“rAAV”是指重組腺伴隨病毒，也稱為重組AAV載體(或“rAAV載體”)。術語“AAV”包括AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、rh10及其雜合體，禽AAV、牛AAV、犬AAV、馬AAV、靈長類動物AAV、非靈長類動物AAV和羊AAV。AAV的各種血清型的基因組序列，以及天然末端重複序列(TR)、Rep蛋白和衣殼亞單位的序列是本領域已知的。這類序列可見於文獻中或公共資料庫如GenBank中。如本文所用的“rAAV載體”是指包含非AAV來源的多核苷酸序列(即，與AAV相比異源的多核苷酸)的AAV載體，該多核苷酸序列一般是用於細胞的遺傳轉化的目的序列。通常，異源多核苷酸的側翼為至少一個、通常兩個AAV反向末端重複序列(ITR)。rAAV載體可以是單鏈的(ssAAV)或自互補的(scAAV)。“AAV病毒”或“AAV病毒顆粒”是指由至少一個AAV衣殼蛋白和包殼的多核苷酸rAAV載體組成的病毒顆粒。如果該顆粒包含異源多核苷酸(即野生型AAV基因組以外的多核苷酸，如待遞送至哺乳動物細胞的轉基因)，則其一般被稱為“rAAV病毒顆粒”或簡稱為“rAAV顆粒”。因此，rAAV顆粒的產生必然包括rAAV載體的產生，因為這樣的載體包含在rAAV顆粒內。

【0094】 如本文所用的，術語“治療”、“處理”、“療法”等是指減輕、延遲或減慢疾病或病症的進展，預防、減弱、減少影響或症狀，防止其發作，抑制或減輕其發作。本公開的方法可用於任何哺乳動物。示例性的哺乳動物包括但不限於大鼠、貓、狗、馬、牛、綿羊、豬，更優選人。治療性益處包括正在被治療的潛在疾病的根除或改善。另外，治療性益處也可以如下實現：一種或多種與該基礎病症相關的生理學症狀得到根除或改善，使得在受試者中觀察到起色，雖然該受試者可能仍受基礎病症的折磨。在一些情況下，對於預防性益處，可以將治療劑施用於具有發展成特定疾病的風險的受試者或報告了疾病的一個或多個生理學症狀的受試者，即使可能尚未作出該疾病的診斷。本公開的方法可用於任何哺乳動物。在一些情況下，治療可以導致症狀的減少或停止(例如，癲癇發作頻率、持續時間和/或嚴重程度的降低)。預防性效果包括延緩或消除疾病或病況的出現，延緩或消除疾病或病況的症狀發作，減慢、終止或逆轉疾病或病況的進展，或其任意組合。

【0095】 術語“有效量”或“治療有效量”是指足以實現預期應用的本文所述組合物的量，該預期應用包括但不限於如下文定義的疾病治療。治療有效量可根據預期的治療應用(體內)，或所治療的受試者和疾病狀況，例如，受試者的體重和年齡、疾病狀況的嚴重程度、給藥方式等而變化，這些可以容易地由本領域普通技術人員確定。該術語還適用於將在靶細胞中誘導特定應答的劑量。具體劑量將根據所選擇的特定組合物、所遵循的給藥方案、是否與其他化合物聯合施用、給藥時機、所施用的組織、攜帶該組合物的物理遞送系統而變化。

【0096】 核苷酸或肽序列的“片段”是指短於參考或“全長”序列的序

列。

【0097】 分子的“變體”是指此類序列的等位基因變異，即，在結構和生物活性上與整個分子或其片段基本相似的序列。

【0098】 DNA或蛋白質序列的“功能片段”是指保留與全長DNA或蛋白質序列的生物活性基本相似的生物活性(在功能或結構上)的片段。DNA序列的生物活性可以是其以已知歸因於全長序列的方式影響表達的能力。

【0099】 術語“受試者”和“個體”在本文中可互換使用，是指脊椎動物，優選哺乳動物，更優選人。本文所述的方法可用於人類治療、獸醫應用和/或在疾病或病況的動物模型中的臨床前研究。

【0100】 術語“體內”是指在受試者的身體中發生的事件。

【0101】 術語“體外”是指在受試者的身體外發生的事件。例如，體外試驗包括在受試者體外進行的任何試驗。體外試驗包括其中採用活細胞或死細胞的基於細胞的試驗。體外試驗還包括其中不採用完整細胞的無細胞試驗。

【0102】 通常，可以互換使用的“序列同一性”或“序列同源性”是指兩個多核苷酸或多肽序列分別的確切核苷酸-核苷酸對應關係或氨基酸-氨基酸對應關係。一般而言，用於確定序列同一性的技術包括比較兩個核苷酸或氨基酸序列，並確定其同一性百分比。例如為了評估同一性，序列比較可以通過任何合適的比對演算法來進行，包括但不限於Needleman-Wunsch演算法(參見，例如，可以在www.ebi.ac.uk/Tools/psa/emboss_needle/獲得的EMBOSS Needle比對器(aligner)，任選地採用預設設置)、BLAST演算法(參見，例如，可以在blast.ncbi.nlm.nih.gov/Blast.cgi獲得的BLAST比對工具，任選地採用預

設設置)和Smith-Waterman演算法(參見，例如，可以在www.ebi.ac.uk/Tools/psa/emboss_water/獲得的EMBOSS Water比對器，任選地採用默認設置)。可以使用所選演算法的任何合適的參數(包括默認參數)來評估最佳比對。兩個序列之間的“同一性百分比”，也稱為“同源性百分比”，可被計算為兩個最佳比對的序列之間的精確匹配數除以參考序列的長度並乘以100。例如，也可以通過使用可從國立衛生研究院(National Institutes of Health)得到的包括2.2.9版本在內的高級BLAST電腦程式比較序列資訊來確定同一性百分比。BLAST程式基於Karlin和Altschul, Proc. Natl. Acad. Sci. USA 87:2264-2268 (1990)的比對方法，並如以下文獻所討論：Altschul等人, J. Mol. Biol. 215:403-410 (1990)；Karlin和Altschul, Proc. Natl. Acad. Sci. USA 90:5873-5877 (1993)；和Altschul等人, Nucleic Acids Res. 25:3389-3402 (1997)。簡言之，BLAST程式將同一性定義為相同的比對標誌(即核苷酸或氨基酸)的數目除以兩個序列中較短者中的標誌的總數。可以使用該程式來確定在所比較的序列的總長度上的同一性百分比。提供默認參數來優化採用短查詢序列的搜索，例如採用blastp程式。該程式還允許使用SEG篩檢程式來遮罩通過Wootton和Federhen, Computers and Chemistry 17: 149-163 (1993)的SEG程式確定的查詢序列區段。高序列同一性通常包括大約80%至100%的序列同一性範圍，及其之間的整數值。

【0103】 如本文所用的，關於蛋白質的“工程化的”是指非天然存在的蛋白質，包括但不限於衍生自天然存在的蛋白質的蛋白質，或其中天然存在的蛋白質已被修飾或重新程式設計以具有某種性質的蛋白質。

【0104】 如本文所用的，“合成的”和“人工的”可互換使用，是指與

天然存在的人類蛋白質具有低序列同一性(例如，小於50%的序列同一性)的蛋白質或其結構域。例如，VPR和VP64結構域是合成的反式啟動域。

【0105】 如本文所用的，“工程化轉錄因子”或“eTF”是指經修飾或重新程式設計以結合特定靶結合位點和/或包括經修飾或替換的轉錄效應物結構域的非天然存在的DNA結合蛋白或非天然存在的轉錄調節物。

【0106】 如本文所用的，“DNA結合域”可以用來指代一個或多個DNA結合基序，如鋅指或鹼性螺旋-環-螺旋(bHLH)基序，其單獨地或共同地作為DNA結合蛋白的一部分。

【0107】 術語“轉錄啟動域”、“轉錄啟動結構域”、“反式啟動域”、“反式啟動結構域”和“TAD”在本文中可互換使用，並且是指這樣的蛋白質結構域，它與DNA結合域一起，可以通過直接地或通過被稱為共啟動物的其他蛋白質接觸轉錄機制(例如，通用轉錄因子和/或RNA聚合酶)而啟動從啟動子的轉錄。

【0108】 術語“轉錄阻抑域”、“轉錄阻抑結構域”和“TRD”在本文中可互換使用，並且是指這樣的蛋白質結構域，它與DNA結合域一起，可以通過直接地或通過被稱為共阻抑物的其他蛋白質接觸轉錄機制(例如，通用轉錄因子和/或RNA聚合酶)而阻抑從啟動子的轉錄。

【0109】 術語“GRCh38.p12”是指Genome Reference Consortium Human Build 38補丁版本12(GRCh38.p12)，其GenBank Assembly登錄號為GCA_000001405.27，日期為2017/12/21。

【0110】 除非另有說明，否則本文使用的所有術語均具有對於本領域技術人員而言相同的含義，並且本發明的實施將採用分子生物學、微生物學和重組DNA技術中的常規技術，這些技術均在本領域技術人員的知

識範疇內。

上調SCN1A的工程化轉錄因子(eTF)

【0111】在一方面，本申請提供了能夠上調鈉電壓門控通道α亞單位1(SCN1A)基因的表達並增加其相應蛋白質產物Nav1.1的表達的eTF。SCN1A基因屬於編碼用於裝配鈉通道的亞單位的基因家族。這些通道將帶正電荷的鈉離子轉運到細胞內，在細胞產生並傳輸電信號的能力方面起關鍵作用。SCN1A基因編碼被稱為Nav1.1的鈉通道的一部分(α亞單位)。這些通道主要存在於腦中，在腦中它們控制鈉離子向細胞內的流動。Nav1.1通道參與將信號從一個神經細胞(或神經元)傳輸到另一個神經細胞(或神經元)。已經發現SCN1A基因中的數個突變會導致遺傳性癲癇伴熱性癲癇發作+(GEFS+)，這是嚴重程度不同的一系列癲癇發作病症。這些病況包括單純的熱性(發熱相關的)癲癇發作，其開始於嬰兒期，通常到5歲時停止，以及熱性癲癇發作+(FS+)。FS+涉及熱性和其他類型的癲癇發作，包括與發熱無關的癲癇發作(無發熱癲癇發作)，其持續到超過兒童期。GEFS+譜還包括其他病況，如Dravet綜合症(也稱為嬰兒重度肌陣攣性癲癇或SMEI)，其導致更嚴重的癲癇發作，持續時間更長，並且可能難以控制。這些反復的癲癇發作(癲癇)會隨著時間的推移而惡化，並且常常伴有腦功能下降。其他許多突變與家族性偏癱性偏頭痛有關，家族性偏癱性偏頭痛是一種在家族中普遍存在的偏頭痛形式，並且至少一種突變與某些抗癲癇藥物的有效性相關。因此，本文提供的增加SCN1A表達的eTF可用來治療與Nav1.1通道中的突變相關的多種疾病或病症。

【0112】轉錄因子(TF)是結合基因組中的特定序列並控制基因表達的蛋白質。本文提供的上調SCN1A的工程化轉錄因子或eTF是非天然存在

的蛋白質，其包含DNA結合域(DBD)和至少一個作為轉錄調節物的結構域，例如轉錄啟動域(TAD)或轉錄阻抑域(TRD)。在一個實施方案中，上調SCN1A的eTF可包含DBD和TAD(例如，TAD-DBD或DBD-TAD)，其中該DBD和TAD可衍生自相同的蛋白質或不同的蛋白質。在另一個實施方案中，上調SCN1A的eTF可包含DBD和兩個TAD，其中該DBD和TAD衍生自相同的蛋白質，該DBD衍生自第一蛋白質且兩個TAD均衍生自第二蛋白質，該DBD和一個TAD衍生自第一蛋白質且第二個TAD衍生自第二蛋白質，或者該DBD衍生自第一蛋白質、一個TAD衍生自第二蛋白質且第二個TAD衍生自第三蛋白質(例如，TAD1-DBD-TAD1、TAD1-DBD-TAD2、TAD1-TAD1-DBD、TAD1-TAD2-DBD、DBD-TAD1-TAD1或DBD-TAD1-TAD2)。在另一個實施方案中，上調SCN1A的eTF可包含DBD和三個TAD，其中該DBD和TAD衍生自相同的蛋白質，該DBD衍生自第一蛋白質且TAD衍生自一種或多種不同的蛋白質，或者其中該DBD和所有TAD均衍生自不同的蛋白質，例如，TAD_X-TAD_X-TAD_X-DBD、TAD_X-TAD_X-DBD-TAD_X、TAD_X-DBD-TAD_X-TAD_X或DBD-TAD_X-TAD_X-TAD_X，其中每個X獨立地選擇並且可以與一個或全部其他TAD相同或不同。實例包括，例如，TAD1-TAD1-DBD-TAD1、TAD1-TAD1-DBD-TAD2、TAD1-TAD2-DBD-TAD1、TAD1-TAD2-DBD-TAD2、TAD1-TAD2-DBD-TAD3、TAD1-DBD-TAD1-TAD1、TAD1- DBD-TAD2-TAD2、TAD1-DBD-TAD1-TAD2、TAD2-DBD-TAD1-TAD2、TAD1-DBD-TAD2-TAD3、TAD1-TAD1-TAD1-DBD、TAD1-TAD2-TAD2-DBD、TAD1-TAD2-TAD2-DBD、TAD1-TAD2-TAD3-DBD、DBD-TAD1-TAD1-TAD1、DBD-TAD1-TAD1-TAD2、

DBD-TAD1-TAD2-TAD2或DBD-TAD1-TAD2-TAD3等。在另一個實施方案中，上調SCN1A的eTF可包含DBD和四個TAD，其中該DBD和TAD衍生自相同的蛋白質，該DBD衍生自第一蛋白質且TAD衍生自一種或多種不同的蛋白質，或者其中該DBD和所有TAD均衍生自不同的蛋白質，例如，TAD_X-TAD_X-TAD_X-TAD_X-DBD、TAD_X-TAD_X-TAD_X-DBD-TAD_X、TAD_X-TAD_X-DBD-TAD_X-TAD_X、TAD_X- DBD-TAD_X-TAD_X-TAD_X或DBD-TAD_X-TAD_X-TAD_X-TAD_X，其中每個X獨立地選擇並且可以與一個或全部其他TAD相同或不同。實例包括，例如，TAD1-TAD1-DBD-TAD1-TAD1、TAD1-TAD1-DBD-TAD2-TAD2、TAD1-TAD2-DBD-TAD1-TAD2、TAD1-TAD2-DBD-TAD2-TAD1、TAD1-TAD2-DBD-TAD1-TAD3、TAD1-TAD3-DBD-TAD1-TAD2、TAD1-TAD2-DBD-TAD3-TAD4、TAD1-TAD1-TAD1-DBD-TAD2、TAD1-TAD2-TAD3-DBD-TAD4、TAD1-DBD-TAD1-TAD1-TAD2、TAD1-DBD-TAD2-TAD3-TAD4、TAD1-DBD-TAD1-TAD2-TAD3、TAD2-DBD-TAD1-TAD2-TAD3、TAD1-DBD-TAD2-TAD3-TAD4、TAD1-TAD1-TAD1-TAD1-DBD、TAD1-TAD2-TAD2-TAD3-DBD、TAD1-TAD2-TAD3-TAD4-DBD、DBD-TAD1-TAD1-TAD1-TAD1、DBD-TAD1-TAD1-TAD2-TAD2、DBD-TAD1-TAD2-TAD3-TAD4 或 DBD-TAD1-TAD2-TAD3-TAD3等。在一個實施方案中，上調SCN1A的eTF包含DBD和兩個位於該DBD的相同末端(例如，N-末端或C-末端)的TAD，其中該DBD衍生自第一蛋白質且兩個TAD均衍生自第二蛋白質，或者該DBD衍生自第一蛋白質、一個TAD衍生自第二蛋白質且第二個TAD衍生自第三蛋白質(例如，TAD1-TAD1-DBD、TAD1-TAD2-DBD、DBD-TAD1-

TAD1或DBD-TAD1-TAD2)。在某些實施方案中，該DBD可以是包含來自多種蛋白質的結構域的合成構建體。

【0113】在某些實施方案中，DBD和TAD和/或兩個TAD可以直接綴合，例如，沒有間插氨基酸序列，DBD和TAD和/或兩個TAD可以使用肽連接體綴合，或其組合。在某些實施方案中，DBD與TAD綴合，和/或一個TAD經由具有1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、60、70、75、80、90或100個氨基酸或者1-5、1-10、1-20、1-30、1-40、1-50、1-75、1-100、5-10、5-20、5-30、5-40、5-50、5-75、5-100、10-20、10-30、10-40、10-50、10-75、10-100、20-30、20-40、20-50、20-75或20-100個氨基酸的連接體與第二TAD綴合。在一些情況下，DBD和TAD和/或兩個TAD經由在衍生出所述結構域的天然存在的蛋白質中發現的天然存在的間插殘基綴合。在其他實施方案中，DBD和TAD和/或兩個TAD經由合成的或外源的連接體序列綴合。合適的連接體可以是柔性的、可切割的、不可切割的、親水的和/或疏水的。在某些實施方案中，DBD和TAD和/或兩個TAD可經由包含多個甘氨酸和/或絲氨酸殘基的連接體融合在一起。甘氨酸/絲氨酸肽連接體的實例包括[GS]n、[GGGS]n(SEQ ID NO: 179)、[GGGGS]n(SEQ ID NO: 180)、[GGSG]n(SEQ ID NO: 181)，其中n為等於或大於1的整數。在某些實施方案中，可用於綴合DBD和TAD和/或兩個TAD的連接體是GGSGGGSG(SEQ ID NO: 177)。在某些實施方案中，可用於綴合DBD和TAD和/或兩個TAD的連接體是GGSGGGSGGGSGGGSG(SEQ ID NO: 178)。在某些實施方案中，當DBD綴合至兩個TAD時，第一TAD和第二TAD可以通過相同或不同的連

接體綴合至DBD，或者一個TAD可以通過連接體綴合至DBD，而另一個TAD直接綴合至DBD(例如，沒有間插連接體序列)，或者兩個TAD可以直接綴合至DBD(例如，沒有間插連接體序列)。在某些實施方案中，當DBD在同一末端(例如，N-末端或C-末端)綴合至兩個TAD時，連接這兩個TAD的連接體可以與將TAD連接至DBD的連接體相同或不同，或者TAD可以通過連接體彼此綴合，但是TAD直接與DBD綴合(例如，沒有間插連接體序列)，或者TAD可以彼此直接綴合(例如，沒有間插連接體序列)，但是TAD通過連接體與DBD綴合。在某些實施方案中，本文提供的上調SCN1A的eTF不包含位於DBD與一個或多個TAD之間的一個或多個HA標籤(例如，SEQ ID NO: 171)。

【0114】 本文提供的上調SCN1A的eTF具有與天然存在的轉錄因子不同的性質。在某些實施方案中，上調SCN1A的eTF包含衍生自天然存在的蛋白質的DBD，該DBD已被修飾，使得與衍生出該DBD的天然存在的蛋白質相比，該DBD與不同的靶位點結合，並且與衍生出該DBD的天然存在的蛋白質(例如，除SCN1A以外的基因)相比，包含這種修飾的DBD的eTF調節從不同基因(例如，SCN1A)的表達。在其他實施方案中，本文提供的上調SCN1A的eTF包含衍生自天然存在的蛋白質的TAD，該TAD已被修飾，使得與衍生出該TAD的天然存在的蛋白質(例如，除SCN1A以外的基因)相比，包含這種修飾的TAD的eTF調節從不同基因(例如，SCN1A)的表達，並且/或者與衍生出該TAD的天然存在的蛋白質相比，包含這種修飾的TAD的eTF不同地調節SCN1A的表達(例如，上調與下調)。在某些實施方案中，本文提供的上調SCN1A的eTF包含衍生自天然存在的蛋白質的DBD和衍生自天然存在的蛋白質(相同或不同的蛋白質)的TAD，

其中該DBD和TAD均已被修飾。在這樣的實施方案中，與衍生自出DBD的天然存在的蛋白質相比，所述DBD可以與不同的靶位點結合，與衍生出結構域的天然存在的蛋白質(例如，除SCN1A以外的基因)相比，包含這種修飾的DBD和TAD的eTF調節從不同基因(例如，SCN1A)的表達，並且/或者與衍生出DBD和TAD結構域的天然存在的蛋白質相比，包含這種修飾的DBD和TAD的eTF不同地調節SCN1A的表達(例如，上調與下調)。

DNA結合域(DBD)

【0115】 本文提供的上調SCN1A的eTF可包含與目的靶位元元點(例如，當被本文提供的eTF結合時導致SCN1A上調的靶位點)結合的任何合適的DBD。在某些實施方案中，該DBD可以是合成設計的DBD。在其他實施方案中，該DBD可以衍生自天然存在的蛋白質。DBD家族包括鹼性螺旋-環-螺旋(bHLH)(例如，c-Myc)、鹼性亮氨酸拉鍊(例如，C/EBP)、螺旋-轉角-螺旋(例如，Oct-1)和鋅指(例如，EGR1或EGR3)。這些家族表現出廣泛的DNA結合特異性和基因靶標。如本文所預期的，任何一種已知的人轉錄因子蛋白質都可以充當蛋白質平臺，用於工程化和/或重新程式設計DBD，以識別特異性靶位點，從而導致內源SCN1A基因表達的調節。在示例性實施方案中，本文提供的DBD包含鋅指結構域、TALEN結合域或gRNA/Cas複合物。

【0116】 本文提供的DBD可以被設計為識別任何導致SCN1A上調的靶位點。在示例性實施方案中，DBD被設計成識別基因組位置並當被eTF結合時上調內源SCN1A基因的表達。當被本文提供的eTF結合時能夠調節內源SCN1A基因表達的結合位點可以位於基因組中導致調節SCN1A的基因表達的任何位置。在各個實施方案中，該結合位點可以位於與

SCN1A不同的染色體上、與SCN1A相同的染色體上、在SCN1A基因的轉錄起始位點(TSS)的上游、在SCN1A基因的TSS的下游、在SCN1A基因的TSS的近端、在SCN1A基因的遠端、在SCN1A基因的編碼區內、在SCN1A基因的內含子內、在SCN1A基因的聚A尾的下游、在調節SCN1A基因的啟動子序列內或在調節SCN1A基因的增強子序列內。

【0117】 DBD可以被設計成與任何長度的靶結合位點結合，只要它提供DBD對靶結合位點序列的特異性識別，例如具有最小的或沒有脫靶結合。在某些實施方案中，與其他所有基因相比，當被eTF結合時，靶結合位點可以以至少2倍、5倍、10倍、20倍、50倍、75倍、100倍、150倍、200倍、250倍、500倍或更高倍數的水準調節SCN1A的表達。在某些實施方案中，與40個最近鄰基因(例如，位於染色體上最靠近SCN1A的編碼序列上游或下游的40個基因)相比，當被eTF結合時，靶結合位點可以以至少2倍、5倍、10倍、20倍、50倍、75倍、100倍、150倍、200倍、250倍、500倍或更高倍數的水準調節SCN1A的表達。在某些實施方案中，靶結合位點可以是至少5 bp、10 bp、15 bp、20 bp、25 bp、30 bp、35 bp、40 bp、45 bp或50 bp或更大。結合位點的具體長度將取決於eTF中的DBD類型。通常，結合位點的長度越長，對基因表達的結合和調節的特異性就越高(例如，較長的結合位點具有較低的脫靶作用)。在某些實施方案中，相對於使用具有與較短靶位點結合的DBD的eTF所觀察到的脫靶效應，具有識別較長靶結合位點的DBD的eTF具有較低的與非特異性結合相關的脫靶效應(例如，調節脫靶基因或SCN1A以外的其他基因的表達)。在一些情況下，脫靶結合的減少是與具有識別較短靶結合位點的DBD的相當eTF相比低至少1.2、1.3、1.4、1.5、2、3、4、5、6、7、8、

9或10倍。

【0118】 在某些實施方案中，本文提供的DBD可以被修飾，以具有增加的結合親和力，使其與靶結合位點結合更長時間，使得與該DBD綴合的TAD能夠募集更多轉錄因子和/或募集這類轉錄因子更長時間，以對內源SCN1A基因的表達水準產生更大的影響。在某些實施方案中，可以修飾DBD以增加其與所需靶位點的特異性結合(或靶上結合)，和/或修飾DBD以減少其非特異性或脫靶結合。

【0119】 在各個實施方案中，可以使用各種方法確定DBD或eTF與靶結合位點之間的結合。在某些實施方案中，可以使用遷移率變化測定、DNA酶保護測定或本領域已知用於測定蛋白質-DNA結合的其他任何體外方法來確定DBD或eTF與靶結合位點之間的特異性結合。在其他實施方案中，可以使用功能測定，例如，通過測量當靶結合位點被eTF結合時基因(例如，SCN1A)的表達(RNA或蛋白質)來確定eTF與靶結合位點之間的特異性結合。例如，靶結合位點可以位於細胞中包含的或整合到細胞基因組中的載體上的報導基因(例如，eGFP)或SCN1A基因的上游，其中該細胞表達eTF。或者，可以將表達eTF的載體引入天然包含SCN1A基因的細胞類型中。與對照(例如，沒有eTF或識別不同靶位點的eTF)相比，在eTF的存在下報導基因(或SCN1A)的表達水準更高，表明該eTF的DBD與靶位點結合。合適的體外(例如，非基於細胞的)轉錄和翻譯系統也可以以類似的方式使用。在某些實施方案中，與對照(例如，無eTF或識別不同靶位點的eTF)相比，與靶位點結合的eTF可具有至少2倍、3倍、5倍、10倍、15倍、20倍、30倍、50倍、75倍、100倍、150倍或更高的報導基因或SCN1A的表達。

【0120】 在某些實施方案中，本文公開的上調SCN1A的eTF識別以下大小的靶結合位點：至少9bp、12bp、15bp、18bp、21bp、24bp、27bp、30bp、33bp或36bp；超過9bp、12bp、15bp、18bp、21bp、24bp、27bp或30bp；或者9-33bp、9-30bp、9-27bp、9-24bp、9-21bp、9-18bp、9-15bp、9-12bp、12-33bp、12-30bp、12-27bp、12-24bp、12-21bp、12-18bp、12-15bp、15-33bp、15-30bp、15-27bp、15-24bp、15-21bp、15-18bp、18-33bp、18-30bp、18-27bp、18-24bp、18-21bp、21-33bp、21-30bp、21-27bp、21-24bp、24-33bp、24-30bp、24-27bp、27-33bp、27-30bp或30-33bp。在示例性實施方案中，本文公開的上調SCN1A的eTF識別18-27bp、18bp或27bp的靶結合位點。

【0121】 在某些實施方案中，本文公開的上調SCN1A的eTF識別位於染色體2上的靶結合位點。在某些實施方案中，本文公開的上調SCN1A的eTF識別位於染色體2上在SCN1A的TSS上游或下游110 kb、100 kb、90 kb、80 kb、70 kb、60 kb、50 kb、40 kb、30 kb、20 kb、10 kb、5 kb、4 kb、3 kb、2 kb或1 kb內的靶結合位點。在某些實施方案中，本文公開的上調SCN1A的eTF識別位於染色體2上在SCN1A的TSS上游110 kb內的靶結合位點。在某些實施方案中，本文公開的上調SCN1A的eTF識別位於染色體2上在SCN1A的TSS下游110 kb內的靶結合位點。在示例性實施方案中，這類靶結合位點是18-27bp、18bp或27bp。

【0122】 在某些實施方案中，本文公開的上調SCN1A的eTF識別位於染色體2上在位置166179652-165989571內、位置166128050-166127958內、位置166155414-166140590內、位置166179652-1661777272內或位置1659990246-165989592內(均參照GRCh38.p12)的

靶結合位點。在示例性實施方案中，這類靶結合位點是18-27bp、18bp或27bp。

【0123】在某些實施方案中，本文公開的上調SCN1A的eTF識別這樣的靶結合位點，其(i)為18-27bp、18bp或27 bp，(ii)與染色體2上選自166178880、166177369、166177362、166177299、166177299、166155393、166155264、166149373、166149176、166149165、166149118、166148953、166148565、166142396、166142391、166142344、166142239、166141162、166140928、166140590、165990076、165989684、165989571、166155255、166155099、166148843、166148361、166142219、166141090、165990246、165990193、166149168、166127991、166128002、166128037或166128025(均參照GRCh38.p12)的位置重疊，並且(iii)當被本文公開的eTF結合時能夠產生至少1.2倍的SCN1A表達增加。

【0124】在某些實施方案中，本文公開的上調SCN1A的eTF(i)與包含SEQ ID NO: 18、25、30、31或35-66中的任一個或由其組成的靶位點結合，並且(ii)當被本文公開的eTF結合時能夠產生至少1.2倍的SCN1A表達增加。

【0125】在某些實施方案中，本文公開的上調SCN1A的eTF識別這樣的靶結合位點，其(i)為18-27bp、18bp或27 bp，(ii)與染色體2上選自166155255、166155099、166148843、166148361、166142219、166141090、165990246、165990193、166149168、166127991、166128002、166128037或166128025(均參照GRCh38.p12)的位置重疊，並且(iii)當被本文公開的eTF結合時能夠產生至少2倍的SCN1A表達增

加。

【0126】 在某些實施方案中，本文公開的上調SCN1A的eTF(i)與包含SEQ ID NO: 18、30、31、37、38、45、47、48、49、55、61、62或64中的任一個或由其組成的靶位點結合，並且(ii)當被本文公開的eTF結合時能夠產生至少2倍的SCN1A表達增加。

【0127】 在某些實施方案中，本文公開的上調SCN1A的eTF識別這樣的靶結合位點，其(i)為18-27bp、18bp或27 bp，(ii)與染色體2上選自166149168、166127991、166128002、166128037或166128025(均參照GRCh38.p12)的位置重疊，並且(iii)當被本文公開的eTF結合時能夠產生至少5倍的SCN1A表達增加。

【0128】 在某些實施方案中，本文公開的上調SCN1A的eTF(i)與包含SEQ ID NO: 18、30、31、37或38中的任一個或由其組成的靶位點結合，並且(ii)當被本文公開的eTF結合時能夠產生至少5倍的SCN1A表達增加。

【0129】 在某些實施方案中，本文公開的上調SCN1A的eTF識別這樣的靶結合位點，其(i)為18-27bp、18bp或27 bp，(ii)與染色體2上選自166128002、166128037或166128025(均參照GRCh38.p12)的位置重疊，並且(iii)當被本文公開的eTF結合時能夠產生至少15倍的SCN1A表達增加。

【0130】 在某些實施方案中，本文公開的上調SCN1A的eTF(i)與包含SEQ ID NO: 30、37或38中的任一個或由其組成的靶位點結合，並且(ii)當被本文公開的eTF結合時能夠產生至少15倍的SCN1A表達增加。

【0131】 在某些實施方案中，本文公開的上調SCN1A的eTF識別這

樣的靶結合位點，其(i)為18-27bp、18bp或27 bp，(ii)與染色體2上選自166128037或166128025(均參照GRCh38.p12)的位置重疊，並且(iii)當被本文公開的eTF結合時能夠產生至少20倍的SCN1A表達增加。

【0132】 在某些實施方案中，本文公開的上調SCN1A的eTF(i)與包含SEQ ID NO: 30或38中的任一個或由其組成的靶位點結合，並且(ii)當被本文公開的eTF結合時能夠產生至少20倍的SCN1A表達增加。

【0133】 在某些實施方案中，本文公開的上調SCN1A的eTF識別這樣的靶結合位點，其(i)為18-27bp、18bp或27 bp，(ii)與染色體2上位置166128025處的位置重疊，並且(iii)當被本文公開的eTF結合時能夠產生至少25倍的SCN1A表達增加。

【0134】 在某些實施方案中，本文公開的上調SCN1A的eTF(i)與包含SEQ ID NO: 30或由其組成的靶位點結合，並且(ii)當被本文公開的eTF結合時能夠產生至少25倍的SCN1A表達增加。

【0135】 在某些實施方案中，本文公開的上調SCN1A的eTF識別這樣的靶結合位點，其(i)為18-27bp、18bp或27 bp，並且(ii)與基因組區域結合，該基因組區域在具有SEQ ID NO: 18、25、30、31或35-66中任一個序列的基因組位置的至少1 kb、750 bp、500 bp、400 bp、300 bp、200 bp、100 bp或50 bp內。在某些實施方案中，當被本文公開的eTF結合時，靶結合位點能夠產生至少1.2倍、2倍、5倍、15倍、20倍或25倍的SCN1A表達增加。

【0136】 在某些實施方案中，本文公開的上調SCN1A的eTF識別這樣的靶結合位點，其(i)為18-27bp、18bp或27 bp，並且(ii)與基因組區域結合，該基因組區域與具有SEQ ID NO: 18、25、30、31或35-66中任一

個序列的基因組位置至少部分重疊。在某些實施方案中，當被本文公開的eTF結合時，靶結合位點能夠產生至少1.2倍、2倍、5倍、15倍、20倍或25倍的SCN1A表達增加。

【0137】 在某些實施方案中，本文公開的上調SCN1A的eTF識別具有任一個以下序列的靶結合位點：SEQ ID NO: 18、25、30、31或35-66。在某些實施方案中，當被本文公開的eTF結合時，靶結合位點能夠產生至少1.2倍、2倍、5倍、15倍、20倍或25倍的SCN1A表達增加。

【0138】 在某些實施方案中，本文公開的上調SCN1A的eTF導致與對照(例如，沒有eTF或者不識別靶位點的eTF)相比，在細胞中或在體內至少1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍、25倍、50倍、100倍或更高，或至少50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高的SCN1A表達(SCN1A RNA和/或Nav1.1蛋白)上調。在各個實施方案中，可以使用PCR方法、Western印跡法或免疫測定來檢測SCN1A表達的上調。

【0139】 在某些實施方案中，本文公開的上調SCN1A的eTF與靶位點結合，在轉錄啟動試驗中，該靶位點相對於對照能夠將SCN1A表達增加至少1.2倍、1.3倍、1.4倍、1.5倍、1.6倍、1.7倍、1.8倍、1.9倍、2倍、3倍、4倍、5倍、8倍、10倍、12倍、15倍、18倍、20倍、25倍、30倍、40倍、50倍、75倍、100倍或更高，或至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。本文在實施例3中提供了示例性的SCN1A轉錄啟動試驗。簡言之，用攜帶eTF或對照eGFP報導基因構建體

的質粒轉染HEK293。轉染後48h，收集細胞，分離RNA，進行逆轉錄，並通過qPCR(例如，使用具有SEQ ID NO: 185和186的引物)分析所得的cDNA樣品，以對內源SCN1A轉錄物的水準進行定量。GAPDH可用作參考基因，以確定SCN1A表達的相對水準。

【0140】 在某些實施方案中，本文公開的上調SCN1A的eTF具有最小的脫靶效應，例如，與非特異性結合相關的脫靶效應，例如，脫靶基因或除SCN1A以外的基因的表達的調節。在一個實施方案中，本文公開的上調SCN1A的eTF與對照相比特異性上調SCN1A，是該eTF與對照相比對於一個或多個脫靶基因產生的表達的至少5倍、10倍、15倍、20倍、25倍、30倍、40倍或50倍。在示例性實施方案中，本文公開的上調SCN1A的eTF與對照相比特異性上調從SCN1A基因的轉錄，是該eTF相對於對照產生的40個最近鄰基因(例如，位於染色體2上SCN1A編碼序列的40個最接近的基因)的轉錄的至少15倍，這些基因例如是PLA2R1、ITGB6、RBMS1、TANK、PSMD14、TBR1、SLC4A10、DPP4、FAP、IFIH1、GCA、FIGN、GRB14、COBLL1、SLC38A11、SCN3A、SCN2A、CSRNP3、GALNT3、TTC21B、SCN9A、SCN7A、B3GALT1、STK39、CERS6、NOSTRIN、SPC25、ABCB11、DHRS9、BBS5、KLHL41、FASTKD1、PPIG、CCDC173、PHOSPHO2、KLHL23、SSB、METTL5、UBR3和MYO3B(參見表14)。在各個實施方案中，可以使用PCR方法檢測從SCN1A基因的轉錄的上調。

【0141】 在某些實施方案中，本文公開的上調SCN1A的eTF能夠在Dravet綜合症的Scn1a^{tm1kea}小鼠模型中的高熱癲癇發作(HTS)試驗中降低癲癇發作的頻率。在某些實施方案中，與對照(例如，PBS處理的或用包

含編碼eGFP的序列的AAV載體治療)相比，本文公開的eTF能夠在HTS試驗中將在42.6°C時的癲癇發作頻率降低至少1.2倍、1.3倍、1.4倍、1.5倍、1.6倍、1.7倍、1.8倍、1.9倍、2.0倍或更多，或至少20%、30%、40%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、100%或更多。在某些實施方案中，本文公開的eTF能夠在HTS試驗中降低在42.6°C時的癲癇發作頻率，使得該實驗中分析的至少60%、62%、65%、70%、75%、76%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的小鼠在42.6°C時沒有癲癇發作。本文在實施例6中描述了示例性的HTS試驗。簡言之，可以在P1經雙側ICV，向雄性Scn1a +/-小鼠與雌性C57Bl/6J小鼠雜交產生的幼崽給予如本文提供的編碼上調SCN1A的eTF的AAV9載體或編碼eGFP的對照載體。可以以約1.0E10-5.0E12 gc/小鼠向小鼠給藥。通過每2分鐘將小鼠的體溫(在受控條件下並通過體溫監測)提高約0.5°C，直到出現第一次強直-陣攣性癲癇發作伴姿勢喪失或者直到體溫達到43°C，在混合129Stac X C57BL/6背景中的P26-P28 SCN1A雜合小鼠和SCN1A野生型小鼠中進行HTS試驗。如果在整個實驗過程中沒有檢測到癲癇發作伴姿勢喪失，則認為小鼠沒有癲癇發作。

【0142】 在某些實施方案中，本文公開的上調SCN1A的eTF能夠提高SCN1A雜合小鼠，例如Scn1a^{tm1kea}小鼠系的存活率。在某些實施方案中，與對照(例如，PBS處理的或用包含編碼eGFP的序列的AAV載體治療)相比，本文公開的eTF能夠將P100時SCN1A雜合小鼠的存活率提高至少1.2倍、1.3倍、1.4倍、1.5倍、1.6倍、1.7倍、1.8倍、1.9倍、2.0倍或更多，或至少20%、30%、40%、50%、55%、60%、65%、70%、

75%、80%、85%、90%、95%、100%或更多。在某些實施方案中，本文公開的eTF能夠提高P100時SCN1A雜合小鼠的存活率，使得該試驗中分析的至少65%、70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的小鼠在P100時仍然存活。本文在實施例7中描述了示例性的存活試驗。簡言之，可以在P1經雙側ICV，向雄性Scn1a +/-小鼠與雌性C57Bl/6J小鼠雜交產生的幼崽給予AAV9載體。可以以約1.0E10-5.0E12 gc/小鼠向小鼠給藥。確定了存活到P100的小鼠的數目。

【0143】 在某些實施方案中，本文提供的上調SCN1A的eTF可以包含DBD，該DBD來自鋅指蛋白，衍生自鋅指蛋白，或者是核酸酶滅活的鋅指蛋白。鋅指是一種小的蛋白質結構基序，其特徵在於一個或多個鋅離子(Zn^{2+})的配位，以便穩定折疊。鋅指(Znf)結構域是相對較小的蛋白質基序，其含有多個導致與DNA靶位點串聯接觸的手指狀突起。鋅指基序的模組性質允許大量的DNA序列組合以高度親和力和特異性結合，因此理想地適合於工程改造可被靶向並結合特定DNA序列的蛋白質。許多工程化的鋅指陣列基於鼠轉錄因子Zif268的鋅指結構域。Zif268具有三個單獨的鋅指基序，它們共同地以高親和力結合9 bp序列。如本文進一步描述的，已經基於結構鑒定了多種鋅指蛋白並將其表徵為不同的類型。任何此類鋅指蛋白都可與本文所述的DBD關聯使用。

【0144】 多種設計鋅指蛋白的方法是可用的。例如，描述了設計鋅指蛋白以結合目的靶DNA序列的方法，參見，例如，Liu Q等人, Design of polydactyl zinc-finger proteins for unique addressing within complex genomes, Proc Natl Acad Sci USA. 94 (11): 5525–30 (1997)；Wright

DA等人, Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly, Nat Protoc . Nat Protoc. 2006;1(3):1637-52；以及CA Gersbach和T Gaj, Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies, Am Chem Soc 47: 2309−2318 (2014)。另外，用於設計鋅指蛋白以與目的DNA靶序列結合的各種基於Web的工具是可公開獲得的，參見，例如，可從萬維網地址omictools.com/zfns-category處找到的OmicX的Zinc Finger Nuclease Design Software Tools and Genome Engineering Data Analysis網址；和可從萬維網地址scripps.edu/barbas/zfdesign/zfdesignhome.php處找到的Scripps的Zinc Finger Tools設計網址。另外，用於設計鋅指蛋白以與目的DNA靶序列結合的各種市售服務是可獲得的，參見，例如，Creative Biolabs提供的市售服務或試劑盒(萬維網地址creative-biolabs.com/Design-and-Synthesis-of-Artificial-Zinc-Finger-Proteins.html)，可從Addgene獲得的Zinc Finger Consortium Modular Assembly Kit(萬維網地址addgene.org/kits/zfc-modular-assembly/)，或來自Sigma Aldrich的CompoZr Custom ZFN Service(萬維網地址sigmaaldrich.com/life-science/zinc-finger-nuclease-technology/custom-zfn.html)。

【0145】在某些實施方案中，本文提供的上調SCN1A的eTF包含含有一個或多個鋅指的DBD，或衍生自鋅指蛋白的DBD。在一些情況下，該DBD包含多個鋅指，其中每個鋅指在其N-末端或C-末端或均通過氨基酸連接體與另一個鋅指或另一個結構域連接。在一些情況下，本文提供的DBD包含來自表9中描述的一種或多種鋅指類型的一個或多個鋅指。在一

些情況下，本文提供的DBD包含多個鋅指結構或基序，或多個具有SEQ ID NO: 152-167中的一個或多個或其任何組合的鋅指。在某些實施方案中，DBD包含X—[ZF—X]n和/或[X—ZF]n—X，其中ZF是具有表9中列出的任一基序(例如，SEQ ID NO: 136-146中的任一個)的鋅指結構域，X是包含1-50個氨基酸的氨基酸連接體，並且n是1-15的整數，例如1、2、3、4、5、6、7、8、9、10、11、12、13、14或15，其中每個ZF可以獨立地具有與該DBD中的其他ZF序列相同的序列或不同的序列，並且其中每個連接體X可以獨立地具有與該DBD中的其他X序列相同的序列或不同的序列。每個鋅指可以在其C-末端、N-末端或兩端與另一個序列、鋅指或結構域連接。在DBD中，每個連接體X可以在序列、長度和/或性質(例如，柔性或電荷)上相同，或者在序列、長度和/或性質上不同。在一些情況下，兩個或更多個連接體可以相同，而其他連接體則不同。在示例性實施方案中，該連接體可以獲自或衍生自連接表9中提供的一種或多種天然存在的鋅指蛋白中發現的鋅指的序列。在其他實施方案中，合適的連接體序列包括例如長度為5個或更多個氨基酸的連接體。關於長度為6個或更多個氨基酸的示例性連接體序列，還參見美國專利6,479,626、6,903,185和7,153,949，其中每一篇均整體併入本文。本文提供的DBD蛋白可以包括該蛋白質的各個鋅指之間的合適連接體的任何組合。本文描述的DBD蛋白可以包括該蛋白質的各個鋅指之間的合適連接體的任何組合。

【0146】 在某些實施方案中，本文提供的上調SCN1A的eTF包含含有一個或多個經典鋅指的DBD。經典的C2H2鋅指在一條鏈中具有兩個半胱氨酸，在另一條鏈中具有兩個組氨酸殘基，它們由鋅離子配位。經典的鋅指結構域具有兩個β-折疊和一個α-螺旋，其中α-螺旋與DNA分子相互作

用並形成DBD與靶位點結合的基礎，並且可以被稱為“識別螺旋”。在示例性實施方案中，鋅指的識別螺旋在位置-1、2、3或6處包含至少一個氨基酸置換，從而改變鋅指結構域的結合特異性。在其他實施方案中，本文提供的DBD包含一個或多個非經典鋅指，例如，C2–H2、C2–CH和C2–C2。

【0147】 在另一個實施方案中，本文提供的上調SCN1A的eTF包含DBD，該DBD包含具有以下結構的鋅指基序：LEPGEKP – [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH – TGKKTS(SEQ ID NO: 147)，其中n為1-15的整數，例如1、2、3、4、5、6、7、8、9、10、11、12、13、14或15，並且每個X獨立地是能夠與靶序列的3 bp結合的識別序列(例如識別螺旋)。在示例性實施方案中，n為3、6或9。在特別優選的實施方案中，n為6。在各個實施方案中，與DBD中的其他X序列相比，每個X可以獨立地具有相同的氨基酸序列或不同的氨基酸序列。在示例性實施方案中，每個X是包含7個氨基酸的序列，該序列已經使用萬維網位址scripps.edu/barbas/ zfdesign/zfdesignhome.php處的Scripps的Zinger Finger Design Tool設計為與目的靶結合位元元點的3 bp相互作用。

【0148】 由於DBD內的每個鋅指都識別3 bp，因此DBD中包含的鋅指的數目會告知被DBD識別的結合位點的長度，例如，具有1個鋅指的DBD將識別具有3 bp的靶結合位點，具有2個鋅指的DBD將識別具有6 bp的靶結合位點，具有3個鋅指的DBD將識別具有9 bp的靶結合位點，具有4個鋅指的DBD將識別具有12 bp的靶結合位，具有5個鋅指的DBD將識別具有15 bp的靶結合位點，具有6個鋅指的DBD將識別具有18 bp的靶結合

位點，具有9個鋅指的DBD將識別具有27 bp的靶結合位，依此類推。通常，識別更長靶結合位點的DBD將表現出更高的結合特異性(例如，更少的脫靶結合或非特異性結合)。

【0149】 在其他實施方案中，本文提供的上調SCN1A的eTF包含DBD，該DBD通過在鋅指結構域的一個或多個識別螺旋中進行一個或多個氨基酸置換而衍生自天然存在的鋅指蛋白，以便改變DBD的結合特異性(例如，改變被DBD識別的靶位點)。本文提供的DBD可以衍生自任何天然存在的鋅指蛋白。在各個實施方案中，這樣的DBD可以衍生自諸如小鼠、大鼠、人等任何物種的鋅指蛋白。在示例性實施方案中，本文提供的DBD衍生自人鋅指蛋白。在某些實施方案中，本文提供的DBD衍生自表9中列出的天然存在的蛋白質。在示例性實施方案中，本文提供的DBD蛋白衍生自人EGR鋅指蛋白，例如，EGR1、EGR2、EGR3或EGR4。

【0150】 在某些實施方案中，本文提供的上調SCN1A的eTF包含DBD，該DBD通過修飾DBD而衍生自天然存在的蛋白質，該修飾通過在天然存在的蛋白質的DBD內重複一個或多個鋅指而增加DBD蛋白中的鋅指結構域的數目。在某些實施方案中，這樣的修飾包括在天然存在的蛋白質的DBD內鋅指的雙重複製、三重複製、四重複製或進一步的倍增。在一些情況下，來自人類蛋白質的DBD的一個鋅指被複製，例如，2、3、4、5、6、7、8、9、10、11、12個或更多個拷貝的相同鋅指基序在eTF的DBD中被重複。在一些情況下，來自天然存在的蛋白質的DBD的一組鋅指被複製，例如，來自天然存在的蛋白質的DBD的一組3個鋅指被雙重複製，以產生包含具有6個鋅指的DBD的eTF，被三重複製，以產生具有9個鋅指的eTF的DBD，或被四重複製，以產生具有12個鋅指的eTF的DBD，

依此類推。在一些情況下，來自天然存在的蛋白質的DBD的一組鋅指被部分複製，以形成具有更大數目的鋅指的eTF的DBD，例如，eTF的DBD包含四個鋅指，其中所述鋅指代表來自天然存在的蛋白質的第一鋅指的一個拷貝、第二鋅指的一個拷貝和第三鋅指的兩個拷貝，在eTF的DBD中共有四個鋅指。然後通過在鋅指結構域的一個或多個識別螺旋中進行一個或多個氨基酸置換而進一步修飾這樣的DBD，以便改變DBD的結合特異性(例如，改變被DBD識別的靶位點)。在示例性實施方案中，該DBD衍生自天然存在的人類蛋白質，如人EGR鋅指蛋白，例如，EGR1、EGR2、EGR3或EGR4。

【0151】 人EGR1和EGR3的特徵在於三指C2H2鋅指DBD。鋅指的通用結合規則規定，所有三個指與其具有相似幾何形狀的同源DNA序列相互作用，使用每個鋅指的α螺旋中的相同氨基酸來確定靶結合位點序列的特異性或識別。這樣的結合規則允許人們修飾EGR1或EGR3的DBD，以工程改造識別所需靶結合位點的DBD。在一些情況下，根據已發佈的鋅指設計規則，對EGR1或EGR3的鋅指基序中的7個氨基酸的DNA識別螺旋進行修飾。在某些實施方案中，例如通過改變一個或多個識別螺旋的序列和/或通過增加DBD中鋅指的數目來修飾EGR1或EGR3的三指DBD中的每個鋅指。在某些實施方案中，將EGR1或EGR3重新程式設計，以識別在所需靶位點處至少9、12、15、18、21、24、27、30、33、36個或更多個堿基對的靶結合位點。在某些實施方案中，這類衍生自ERG1或EGR3的DBD包含至少3、4、5、6、7、8、9、10、11、12個或更多個鋅指。在示例性實施方案中，該DBD中的一個或多個鋅指在識別螺旋的位置-1、2、3或6處包含至少一個氨基酸置換。

【0152】在各個實施方案中，包含衍生自EGR1或EGR3的DBD的上調SCN1A的eTF具有與天然存在的EGR1或EGR3的結合特異性不同的DNA結合特異性，例如，該DBD識別的靶結合位點具有與被未修飾的EGR1或EGR3識別的結合位點的以下序列不同的序列：(GCG(T/G)GGGCG)(SEQ ID NO: 182)。

【0153】在其他實施方案中，本文提供的上調SCN1A的eTF包含作為gRNA/Cas複合物的DBD。CRISPR(成簇的規律間隔的短回文重複序列)/Cas9是一種基因組編輯工具，其允許進行位點特異性基因組靶向。II型CRISPR/Cas系統是一種原核適應性免疫應答系統，其使用非編碼RNA指導Cas9核酸酶誘導位點特異性DNA切割。已經利用CRISPR/Cas9系統創建了一種簡單的、RNA可程式設計的方法來介導哺乳動物細胞中的基因組編輯。可以產生單個指導RNA(sgRNA)，以將Cas9核酸酶引導至特定的基因組位置，然後被gRNA/Cas9複合物結合。可以使用多種方法和工具將gRNA設計為與目的靶位元元點結合。例如，用於設計gRNA以結合目的靶DNA序列的方法描述於Aach等人. Flexible algorithm for identifying specific Cas9 targets in genomes. BioRxiv, Cold Spring Harbor Labs. doi: http://dx.doi.org/10.1101/005074 (2014)；Bae等人, Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics. 30(10):1473–1475 (2014)；Doench, J. G.等人, Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotech 34, 184–191 (2016)；Gratz等人, Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila.

Genetics. 196(4):961–971 (2014)；Heigwer等人, E-CRISP: fast CRISPR target site identification. Nat Methods. 11(2):122–123 (2014)；Ma等人, A guide RNA sequence design platform for the CRISPR/Cas9 system for model organism genomes. Biomed Res Int. doi:http://doi.org/10.1155/2013/270805 (2013)；Montague等人, CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. 42(W1):W401–W407 (2014)；Liu等人, CRISPR-ERA: a comprehensive design tool for CRISPR-mediated gene editing, repression and activation. Bioinformatics. 31(22):3676–3678 (2015)；Ran等人, *In vivo* genome editing using Staphylococcus aureus Cas9. Nature. 520(7546):186–191 (2015)；Wu等人, Target specificity of the CRISPR-Cas9 system. Quant Biol. 2(2):59–70 (2015)；Xiao等人, CasOT: a genome-wide Cas9/gRNA off-target searching tool. Bioinformatics. 30(8):1180–1182 (2014)；Zetsche等人, Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas System. Cell. 163(3):759–771 (2015)。另外，用於設計gRNA以結合目的DNA靶序列的各種基於Web的工具是可公開獲得的，參見，例如，可從萬維網地址atum.bio/eCommerce/cas9/input?multipleContacts=false處找到的AUTM的CRISPR gRNA Design工具；可從萬維網地址portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design-crisprai處找到的Broad Institute的CRISPRa/i gRNA設計工具；可從萬維網地址e-crisp.org/E-CRISP/處找到的DKFZ German Cancer Research Center的E-CRISP設計工具；以及可從萬維網地址design.synthego.com/#/處找到的

Synthego的Knockout Guide Design工具。另外，用於設計gRNA以結合目的DNA靶序列的各種可商購服務是可以獲得的，參見，例如，IDT(萬維網地址idtdna.com/site/order/designtool/index/ CRISPR_SEQUENCE)、ThermoFisher(萬維網地址thermofisher.com/ order/custom-oligo/crispr)和GenScript(萬維網地址genscript.com/gRNA-design-tool.html)提供的可商購服務。

【0154】在示例性實施方案中，作為gRNA/Cas複合物的DBD包含核酸酶滅活的Cas蛋白或dCas，例如dCas9，如核酸酶滅活的金黃色葡萄球菌Cas9(dSaCas9)或核酸酶滅活的釀膿鏈球菌Cas9(dSpCas9)。提供gRNA作為包含靶向區域和支架區域的序列，該靶向區域將gRNA/Cas複合物靶向所需的靶位點，而該支架區域促進與Cas蛋白的相互作用。任何合適的gRNA支架可以與本文提供的gRNA一起使用。在示例性實施方案中，該gRNA是單個gRNA或sgRNA，並包含以下支架序列：5’-GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA-3’(SEQ ID NO: 183)。指導RNA的靶向區域連接到支架序列的5'端，以形成完整的sgRNA。在某些實施方案中，gRNA和dCas蛋白可以從同一表達盒表達。在某些實施方案中，U6啟動子用來表達gRNA。在其他實施方案中，例如，通過將dCas穩定地整合到基因組中或在染色體外穩定維持的質粒上，可以在經工程改造以穩定表達dCas-TAD蛋白的細胞中表達gRNA。

【0155】在其他實施方案中，本文提供的上調SCN1A的eTF可以包含DBD，該DBD來自TALEN，衍生自TALEN，或者是核酸酶滅活的TALEN。轉錄啟動物樣效應物核酸酶(TALEN)是包含DBD和可被工程改

造以切割特定DNA序列的核酸酶結構域的限制酶。通過將TAL效應物DNA結合域與DNA切割結構域(例如核酸酶)綴合來創建TALEN。轉錄啟動物樣效應物(TALE)可以被工程改造為與所需的靶DNA序列結合，從而將核酸酶結構域引導至特定位置。

【0156】 TAL效應物是來自黃單胞菌屬(*Xanthomonas*)細菌的細菌蛋白質。DNA結合域包含重複的高度保守的33-34個氨基酸的序列，具有不同的第12個和第13個氨基酸。這兩個位置，被稱為重複可變二殘基(RVD)，是高度可變的，並且顯示出與特定核苷酸識別的強相關性。氨基酸序列與DNA識別之間的這種直接關係允許通過選擇含有適當RVD的重複區段的組合來工程改造特異性靶向所需序列的DBD。

【0157】 多種設計TALE的方法是可用的。例如，用於設計TALE以與目的靶DNA序列結合的方法描述於T. Cermak等人, *Nucleic Acids Research*. **39** (12): e82 (2011)；F. Zhang F等人, *Nature Biotechnology*. **29** (2): 149–53 (2011)；R. Morbitzer等人, *Nucleic Acids Research*. **39** (13): 5790–9 (2011)；T. Li等人, *Nucleic Acids Research*. **39** (14): 6315–25 (2011)；R. Geissler等人, *PLOS One*. **6**(5): e19509 (2011)；和E. Weber等人, *PLOS One*. **6** (5): e19722 (2011)。另外，用於設計TALE以與目的DNA靶序列結合的各種基於Web的工具是可公開獲得的，參見，例如，可從萬維網地址e-talen.org/E-TALEN/TAL處找到的E-Talen，以及可從萬維網地址tale-nt.cac.cornell.edu/node/add/single-tale處找到的Effector Nucleotide Targeter 2.0工具。另外，用於設計TALE以與目的DNA靶序列結合的各種市售服務是可獲得的，參見，例如，OmicX(萬維網地址omictools.com/)、Addgene(萬維網

地址addgene.org/talen/guide/)或ThermoFisher(萬維網地址thermofisher.com/us/en/home/life-science/genome-editing/geneart-tals/tal-design-tool.html)提供的市售服務。另外，可以使用可公開獲得的軟體程式(DNAWorks)設計適合於裝配TALE的寡核苷酸，參見，例如，D. Hoover D *Methods in Molecular Biology*. **852**: 215–23 (2012)。

轉錄調節域

【0158】本文提供的上調SCN1A的eTF可包含能夠募集一種或多種蛋白質因子的任何合適的結構域，當eTF通過DBD(例如，鋅指DBD、gRNA/Cas DBD或TALE DBD)與靶位點結合時，該蛋白質因子可以調節從目的基因的轉錄(例如，RNA聚合酶II、CBP/p300、CREB或KRAB)或基因表達水準。在某些實施方案中，這樣的結構域募集增加目的基因的轉錄或基因表達水準的蛋白質因子，並且是轉錄啟動域(TAD)。在其他實施方案中，這樣的結構域募集降低目的基因的轉錄或基因表達水準的蛋白質因子，並且是轉錄阻抑域(TRD)。在某些實施方案中，該轉錄調節域(TAD或TRD)可以是合成設計的結構域。在其他實施方案中，該轉錄調節域(TAD或TRD)可以衍生自天然存在的蛋白質，例如，轉錄因子、轉錄共啟動物、轉錄共阻抑物或沉默子蛋白質。在各個實施方案中，轉錄調節域(TAD或TRD)可以衍生自諸如小鼠、大鼠、猴、病毒或人等任何物種的蛋白質。

【0159】在一個示例性實施方案中，適於在本文提供的上調SCN1A的eTF中使用的TAD衍生自病毒蛋白質。衍生自病毒蛋白質的示例性TAD包括，例如，VP64(SEQ ID NO: 133)、VPR(SEQ ID NO: 132)、VP16、VP128、p65、p300的TAD結構域，或其任何功能片段或變體，

或與之具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

【0160】在另一個示例性實施方案中，適於在本文提供的上調SCN1A的eTF中使用的TAD衍生自人類蛋白質。衍生自人類蛋白質的示例性TAD包括，例如，CBP/p300相互作用反式啟動物2(CITED2)(SEQ ID NO: 134)、CBP/p300相互作用反式啟動物4(CITED4)(SEQ ID NO: 135)、EGR1(SEQ ID NO: 176)、CREB3(SEQ ID NO: 224)或EGR3(SEQ ID NO: 175)的TAD結構域，或其任何功能片段或變體，或與之具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列。

【0161】在某些實施方案中，上調SCN1A的eTF包含與轉錄啟動域或TAD綴合的鋅指DBD。在各個實施方案中，鋅指DBD可以與來自病毒蛋白質如VP64或VPR的TAD或來自人類蛋白質如CITED2、CITED4或CREB3的TAD綴合。在某些實施方案中，衍生自人類蛋白質例如EGR1或EGR3的鋅指DBD與衍生自人類蛋白質例如CITED2、CITED4或CREB3的TAD綴合。在某些實施方案中，衍生自人類蛋白質例如EGR1或EGR3的鋅指DBD與VP64或VPR TAD綴合。在某些實施方案中，與人類蛋白質例如EGR1或EGR3具有小於75%序列同一性的合成鋅指DBD或鋅指DBD與衍生自人類蛋白質例如CITED2、CITED4或CREB3的TAD綴合。在某些實施方案中，與人類蛋白質例如EGR1或EGR3具有小於75%序列同一性的合成鋅指DBD或鋅指DBD與VP64或VPR TAD綴合。

【0162】在某些實施方案中，dCas蛋白與TAD綴合。在各個實施方案中，dCas可以與來自病毒蛋白質如VP64或VPR的TAD或來自人類蛋白

質如CITED2、CITED4或CREB3的TAD綴合。在示例性實施方案中，dCas9與VP64或VPR TAD綴合。

【0163】在某些實施方案中，TALE蛋白與TAD綴合。在各個實施方案中，該TALE可以與來自病毒蛋白質如VP64或VPR的TAD或來自人類蛋白質如CITED2、CITED4或CREB3的TAD綴合。在示例性實施方案中，TALE與VP64或VPR TAD綴合。

上調SCN1A並且與人類蛋白質高度同源的eTF

【0164】在某些實施方案中，本文公開的上調SCN1A的eTF與一種或多種人類蛋白質具有高同一性百分比(如下文進一步描述)。在某些實施方案中，這類eTF與一種或多種人類蛋白質具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整體序列同一性。在某些實施方案中，相比於與一種或多種人類蛋白質具有較低整體序列同一性百分比的eTF，這類eTF表現出降低的免疫原性。在各個實施方案中，免疫原性的降低可以使用elispot測定、免疫測定或電腦類比方法來測量。在某些實施方案中，這類eTF可包含衍生自人EGR1或EGR3的DBD和衍生自人EGR1、EGR3、CITED2、CITED4或CREB3的TAD。相比於與人類蛋白質序列具有較低同一性百分比的eTF，這類eTF當施用於受試者時沒有或幾乎沒有免疫原性，或具有降低的免疫原性。

【0165】在某些實施方案中，本文提供的上調SNC1A的eTF與一種或多種人類蛋白質具有至少75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性。當本

文提供的上調SNC1A的eTF包含衍生自同一蛋白質的DBD和TAD時，與人類蛋白質的同一性百分比可以通過計算eTF中與其衍生自的人類蛋白質(例如EGR1或EGR3)相匹配的氨基酸殘基的總數除以eTF中的氨基酸殘基總數來確定。當提供的上調SNC1A的eTF包含來自一種人類蛋白質的DBD和衍生自不同人類蛋白質的TAD時，與人類對應物的同一性百分比可以通過分別計算每個結構域與人類對應物的同一性百分比並將二者相加來確定，例如，(i)計算DBD中與其衍生自的人類蛋白質(例如，EGR1或EGR3)相匹配的氨基酸殘基總數，除以eTF中的氨基酸殘基總數；(ii)計算TAD中與其衍生自的人類蛋白質(例如，CITED2、CITED4或CREB3)相匹配的氨基酸殘基總數，除以eTF中的氨基酸殘基總數；以及(iii)將(i)和(ii)的總和相加。在這樣的實施方案中，結構域如下劃分：第一結構域從eTF的N-末端延伸到第二結構域的編碼序列的起點，第二結構域從第二結構域的編碼序列的起點延伸到eTF的C-末端(例如，對於具有構型NLS-DBD-連接體-NLS-TAD的eTF，第一結構域為NLS-DBD-連接體，而第二結構域為NLS-TAD)。當本文提供的上調SNC1A的eTF包含來自一種人類蛋白質的DBD和衍生自一種或多種不同人類蛋白質的兩個TAD時，與人類對應物的同一性百分比可以通過分別計算每個結構域與人類對應物的同一性百分比並將全部三者相加來確定，例如，(i)計算DBD中與其衍生自的人類蛋白質(例如，EGR1或EGR3)相匹配的氨基酸殘基總數，除以eTF中的氨基酸殘基總數；(ii)計算第一TAD中與其衍生自的人類蛋白質(例如，CITED2、CITED4或CREB3)相匹配的氨基酸殘基總數，除以eTF中的氨基酸殘基總數；(iii)計算第二TAD中與其衍生自的人類蛋白質(例如，CITED2、CITED4或CREB3)相匹配的氨基酸殘基總數，除以eTF中

的氨基酸殘基總數；以及(iv)將(i)、(ii)和(iii)的總和相加。在這樣的實施方案中，結構域如下劃分：第一結構域從eTF的N-末端延伸到第二結構域的編碼序列的起點，第二結構域從第二結構域的編碼序列的起點延伸到第三結構域的編碼序列的起點，第三結構域從第三結構域的編碼序列的起點延伸到eTF的C-末端(例如，對於具有構型NLS-TAD1-連接體-NLS-DBD-連接體-NLS-TAD2的eTF，第一結構域為NLS-TAD1-連接體，第二結構域為NLS-DBD-連接體，而第三結構域為NLS-TAD2)。如本節所述，與一種或多種人類蛋白質的同一性百分比可以使用同一性百分比輸出來確定，該同一性百分比輸出使用可從NCBI獲得的標準蛋白質BLAST工具(例如，blastp套件比對工具，使用具有默認參數的blastp(蛋白質->蛋白質)演算法)來獲得，該工具可以從萬維網上的NCBI網址blast.ncbi.nlm.nih.gov/獲得。

【0166】 在某些實施方案中，由於與天然存在的人類蛋白質的高度序列同一性，本文提供的上調SNC1A的eTF具有在人類受試者中引發極低、最小或不引發不良免疫應答的益處。在某些實施方案中，與使用與一種或多種人類蛋白質具有較低同一性百分比的eTF(例如，與一種或多種人類蛋白質具有低於50%、55%、65%或70%序列同一性的eTF)所觀察到的免疫原性相比，本文提供的上調SNC1A的eTF引發降低的免疫原性，例如至少0.5、1、1.5、2、2.5、3、3.5、4、4.5、5、10、15、20、25、30、35、40、45或50倍或更多倍的免疫原性降低。在一些情況下，免疫原性的降低可以使用elispot測定、免疫測定或電腦類比方法來測量。具有低或最小免疫原性的基因療法具有幾個優點，包括改善的患者耐受性、降低的達到治療效果所需的劑量、一次給藥後延長的治療效果、根據需要多次施

用或以多劑量施用的能力、每次給藥在較長時段上持續的治療功效、基因療法的提高的安全性和/或提高的有效性。

【0167】 在某些實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含衍生自一種或多種天然存在的人類蛋白質的DBD和TAD。在某些實施方案中，這類eTF可以包含衍生自包含DBD的任何天然存在的人類蛋白質的DBD。在示例性實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含衍生自天然存在的鋅指蛋白的DBD，例如表1中列出的構建體5-27、36-41或44-53中的任一種。在某些實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含衍生自人類EGR蛋白如EGR1、EGR2、EGR3或EGR4的DBD。在示例性實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含衍生自人EGR1或EGR3的DBD。在各個實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含衍生自人鋅指蛋白質的DBD，其中已經在DBD的一個或多個鋅指結構域中進行了最小的氨基酸改變(例如1、2、3、4、5、6、7個，或1-2、1-3、1-4、1-5、1-6、1-7、2-3、2-4、2-5、2-6、2-7、3-4、3-5、36或3-7個氨基酸改變)，以改變DBD識別目的靶結合位元元點的結合特異性。這類序列修飾優選地在DBD的鋅指結構域的識別螺旋中進行，而其餘的人鋅指DBD或蛋白質(包括TAD)保持未修飾，從而與天然存在的人類蛋白質保持盡可能高的序列同一性。

【0168】 在某些實施方案中，本文提供的上調SNC1A並且與一種或

多種人類蛋白質具有高序列同一性百分比的eTF包含與衍生自人類蛋白質的DBD綴合的衍生自人類蛋白質的一個或多個轉錄調節域(例如，TAD)。在各個實施方案中，轉錄調節域可以衍生自任何天然存在的人類蛋白質，該人類蛋白質具有當eTF通過DBD與靶位點結合時能夠募集可以調節從目的基因的轉錄(例如，RNA聚合酶II、共啟動蛋白或共阻抑蛋白)或基因表達水準的一種或多種蛋白質因子的結構域。在示例性實施方案中，TAD衍生自人EGR蛋白，例如人EGR1、EGR2、EGR3或EGR4，或引用的人類蛋白質，例如人CITED2或CITED4蛋白。在示例性實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含來自人EGR1或EGR3蛋白的TAD。在另一個示例性實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含來自人CITED2或CITED4蛋白的TAD。

【0169】 在一個實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF可以包含人DBD(hDBD)和人TAD(hTAD)(例如，hTAD-hDBD或hDBD-hTAD)，其中該hDBD和hTAD可以衍生自相同的人類蛋白質或不同的人類蛋白質。在另一個實施方案中，本文提供的與一種或多種人類蛋白質具有高序列同一性百分比的eTF可以包含hDBD和兩個hTAD，其中該hDBD和hTAD衍生自相同的人類蛋白質，該hDBD衍生自第一人類蛋白質並且兩個hTAD均衍生自第二人類蛋白質，該hDBD和一個hTAD衍生自第一人類蛋白質並且第二hTAD衍生自第二人類蛋白質，或者該hDBD衍生自第一人類蛋白質、一個hTAD衍生自第二人類蛋白質並且第二hTAD衍生自第三人類蛋白質(例如，hTAD1-hDBD-hTAD1、hTAD1-hDBD-hTAD2、hTAD1-hTAD1-

hDBD、hTAD1-hTAD2-hDBD、hDBD-hTAD1-hTAD1或hDBD-hTAD1-hTAD2)。

【0170】在示例性實施方案中，本文提供的與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含以下構型中的任一種：(i)均衍生自人EGR1的hDBD和hTAD；(ii)均衍生自人EGR3的hDBD和hTAD；(iii)衍生自人EGR1的hDBD和衍生自CITED2的hTAD(例如，hEGR1 DBD-hCITED2 TAD或hCITED2 TAD-hEGR1 DBD)；(iv)衍生自人EGR1的hDBD和衍生自人CITED4的hTAD(例如，hEGR1 DBD-hCITED4 TAD或hCITED4 TAD-hEGR1 DBD)；(v)衍生自人EGR3的hDBD和衍生自CITED2的hTAD(例如，hEGR3 DBD-hCITED2 TAD或hCITED2 TAD-hEGR3 DBD)；(vi)衍生自人EGR3的hDBD和衍生自人CITED4的hTAD(例如，hEGR3 DBD-hCITED4 TAD或hCITED4 TAD-hEGR3 DBD)；(vii)衍生自人EGR1的hDBD和兩個衍生自CITED2的hTAD(例如，hCITED2 TAD-hEGR1 DBD-hCITED2 TAD、hCITED2 TAD-hCITED2 TAD-hEGR1 DBD或hEGR1 DBD-hCITED2 TAD-hCITED2 TAD)；(viii)衍生自人EGR1的hDBD和兩個衍生自人CITED4的hTAD(例如，hCITED4 TAD-hEGR1 DBD-hCITED4 TAD、hCITED4 TAD-hCITED4 TAD-hEGR1 DBD或hEGR1 DBD-hCITED4 TAD-hCITED4 TAD)；(ix)衍生自人EGR3的hDBD和兩個衍生自人CITED2的hTAD(例如，hCITED2 TAD-hEGR3 DBD-hCITED2 TAD、hCITED2 TAD-hCITED2 TAD-hEGR3 DBD或hEGR3 DBD-hCITED2 TAD-hCITED2 TAD)；(x)衍生自人EGR3的hDBD和兩個衍生自人CITED4的hTAD(例如，hCITED4 TAD-hEGR3 DBD-hCITED4 TAD、hCITED4 TAD-

hCITED4 TAD-hEGR3 DBD或hEGR3 DBD-hCITED4 TAD-hCITED4 TAD)；(xi)衍生自人EGR1的hDBD，衍生自人CITED2的第一hTAD，衍生自人CITED4的第二hTAD(例如，hCITED2 TAD-hEGR1 DBD-hCITED4 TAD、hCITED4 TAD-hEGR1 DBD-hCITED2 TAD、hCITED2 TAD-hCITED4 TAD-hEGR1 DBD、hCITED4 TAD-hCITED2 TAD-hEGR1 DBD、hEGR1 DBD-hCITED4 TAD- hCITED2 TAD或hEGR1 DBD-hCITED2 TAD- hCITED4 TAD)；或(xii)衍生自人EGR3的hDBD，衍生自人CITED2的第一hTAD，衍生自人CITED4的第二hTAD(例如，hCITED2 TAD-hEGR3 DBD-hCITED4 TAD、hCITED4 TAD-hEGR3 DBD-hCITED2 TAD、hCITED2 TAD-hCITED4 TAD-hEGR3 DBD、hCITED4 TAD-hCITED2 TAD-hEGR3 DBD、hEGR3 DBD-hCITED4 TAD-hCITED2 TAD或hEGR3 DBD-hCITED2 TAD-hCITED4 TAD)。

【0171】在某些實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF包含以下任一種：(i)包含SEQ ID NO: 103-124、128-131、205、207、209、213、217、219、221或223中的任一個的序列；(ii)包含SEQ ID NO: 92-98中的任一個的序列；(iii)與(i)或(ii)的任一個序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列，或(iv)(i)、(ii)或(iii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、

90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性實施方案中，這類eTF與一種或多種人類蛋白質具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整體序列同一性。在某些實施方案中，相比於與一種或多種人類蛋白質具有較低整體序列同一性百分比的eTF，這類eTF表現出降低的免疫原性。在各個實施方案中，免疫原性的降低可以使用elispot測定、免疫測定或電腦類比方法來測量。

【0172】 在某些實施方案中，本文提供的上調SNC1A並且與一種或多種人類蛋白質具有高序列同一性百分比的eTF除DBD和TAD結構域外還可以另外包含一個或多個氨基酸序列或結構域，如核定位元信號或連接體等。另外，編碼本文提供的與一種或多種人類蛋白質具有高序列同一性百分比的eTF的多核苷酸，除eTF的編碼序列外，還可以另外包含一個或多個核酸序列，如啟動子、增強子、聚A尾等。在這樣的實施方案中，一個或多個另外的氨基酸序列和/或核酸序列優選地是人序列，衍生自人序列，或與人類蛋白質具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的序列同一性。

示例性SCN1A eTF

【0173】 在某些實施方案中，本文公開的上調SCN1A的eTF包含具有一個或多個鋅指結構域的DNA結合域，所述鋅指結構域包含含有SEQ ID NO: 152-167中的任一個的識別螺旋。在某些實施方案中，本文公開的上調SCN1A的eTF包含具有至少一個、兩個、三個、四個、五個、六個、七個、八個、九個、十個、十一個或十二個鋅指結構域的DNA結合域，

其中每個鋅指結構域獨立地包含含有SEQ ID NO: 152-167中的任一個的識別螺旋。在某些實施方案中，本文公開的上調SCN1A的eTF包含具有六個鋅指結構域的DNA結合域，其中每個鋅指結構域獨立地包含含有SEQ ID NO: 152-167中的任一個的識別螺旋。在某些實施方案中，本文公開的上調SCN1A的eTF包含具有九個鋅指結構域的DNA結合域，其中每個鋅指結構域獨立地包含含有SEQ ID NO: 152-167中的任一個的識別螺旋。在示例性實施方案中，此類eTF包含具有SEQ ID NO: 147的DNA結合域，其中每個X獨立選自SEQ ID NO: 152-167中的任一個，並且n為6或9。

【0174】在某些實施方案中，本文公開的上調SCN1A的eTF包含具有以下任一序列的DNA結合域：(i)包含RSDNLVR x REDNLHT x RSDELVR x QSGNLTE x TSGHLVR x QNSTLTE(SEQ ID NO: 148)的序列，其中x可以是1-50個氨基酸的連接體，(ii)與SEQ ID NO: 148具有至少80%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列，或(ii)(i)或(ii)的功能片段。在某些實施方案中，這樣的eTF進一步包含一個或多個選自VP64、VPR、CITED2、CITED4或CREB3的TAD。在一個實施方案中，這樣的eTF包含與DBD的C-末端綴合的VPR TAD結構域。在某些實施方案中，這樣的eTF包含與DBD的N-末端、C-末端或N-末端和C-末端綴合的CITED2 TAD。在某些實施方案中，這樣的eTF包含與DBD的N-末端、C-末端或N-末端和C-末端綴合的CITED4 TAD。在某些實施方案中，這樣的eTF包含與DBD的N-末端或C-末端綴合的兩個CITED4 TAD。在某些實施方案中，這樣的eTF能夠與具有SEQ ID NO: 18的靶位點結合，並且將SCN1A的表達上調與對照相比至少2倍、5倍、

10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

【0175】在某些實施方案中，本文公開的上調SCN1A的eTF包含具有以下任一序列的DNA結合域：(i)包含RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT(SEQ ID NO: 149)的序列，其中x可以是1-50個氨基酸的連接體，(ii)與SEQ ID NO: 149具有至少80%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列，或(ii)(i)或(ii)的功能片段。在某些實施方案中，這樣的eTF進一步包含一個或多個選自VP64、VPR、CITED2、CITED4或CREB3的TAD。在一個實施方案中，這樣的eTF包含與DBD的C-末端綴合的VPR TAD結構域。在某些實施方案中，這樣的eTF包含與DBD的N-末端、C-末端或N-末端和C-末端綴合的CITED2 TAD。在某些實施方案中，這樣的eTF包含與DBD的N-末端、C-末端或N-末端和C-末端綴合的CITED4 TAD。在某些實施方案中，這樣的eTF包含與DBD的N-末端或C-末端綴合的兩個CITED4 TAD。在某些實施方案中，這樣的eTF能夠與具有SEQ ID NO: 30的靶位點結合，並且將SCN1A的表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

【0176】在某些實施方案中，本文公開的上調SCN1A的eTF包含具有以下任一序列的DNA結合域：(i)包含RRDELNV x RSDHLTN x RSDDLVR x RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x

QSSSLVR x REDNLHT(SEQ ID NO: 151)的序列，(ii)與SEQ ID NO: 151具有至少80%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列，或(ii)(i)或(ii)的功能片段。在某些實施方案中，這樣的eTF進一步包含一個或多個選自VP64、VPR、CITED2、CITED4或CREB3的TAD。在一個實施方案中，這樣的eTF包含與DBD的C-末端綴合的VPR TAD結構域。在某些實施方案中，這樣的eTF包含與DBD的N-末端、C-末端或N-末端和C-末端綴合的CITED2 TAD。在某些實施方案中，這樣的eTF包含與DBD的N-末端、C-末端或N-末端和C-末端綴合的CITED4 TAD。在某些實施方案中，這樣的eTF包含與DBD的N-末端或C-末端綴合的兩個CITED4 TAD。在某些實施方案中，這樣的eTF能夠與具有SEQ ID NO: 32的靶位點結合，並且將SCN1A的表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

【0177】在某些實施方案中，本文公開的上調SCN1A的eTF包含具有以下任一序列的DNA結合域：(i)包含DPGALVR x RSDNLVR x QSGDLRR x THLDLIR x TSGNLVR x RSDNLVR(SEQ ID NO: 150)的序列，(ii)與SEQ ID NO: 150具有至少89%、85%、90%、95%、96%、97%、98%或99%序列同一性的序列，或(ii)(i)或(ii)的功能片段。在某些實施方案中，這樣的eTF進一步包含一個或多個選自VP64、VPR、CITED2、CITED4或CREB3的TAD。在一個實施方案中，這樣的eTF包含與DBD的C-末端綴合的VPR TAD結構域。在某些實施方案中，這樣的

eTF包含與DBD的N-末端、C-末端或N-末端和C-末端綴合的CITED2 TAD。在某些實施方案中，這樣的eTF包含與DBD的N-末端、C-末端或N-末端和C-末端綴合的CITED4 TAD。在某些實施方案中，這樣的eTF包含與DBD的N-末端或C-末端綴合的兩個CITED4 TAD。在某些實施方案中，這樣的eTF能夠與具有SEQ ID NO: 31的靶位點結合，並且將SCN1A的表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

【0178】在某些實施方案中，本文公開的上調SCN1A的eTF包含以下任一序列：(i)包含SEQ ID NO: 99-131、205、207、209、213、217、219、221或223中的任一個的序列；(ii)包含SEQ ID NO: 77-98中的任一個的序列；(iii)與(i)或(ii)的任一個序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列，或(iv)(i)、(ii)或(iii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

【0179】在某些實施方案中，本文公開的上調SCN1A的eTF包含以下任一序列：(i)包含SEQ ID NO: 99-102或125-127中的任一個的序列；(ii)包含SEQ ID NO: 77-91中的任一個的序列；(iii)與(i)或(ii)的任一個序

列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列，或(iv)(i)、(ii)或(iii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

【0180】在某些實施方案中，本文公開的上調SCN1A的eTF包含以下任一序列：(i)包含SEQ ID NO: 103-124、128-131、205、207、209、213、217、219、221或223中的任一個的序列；(ii)包含SEQ ID NO: 92-98中的任一個的序列；(iii)與(i)或(ii)的任一個序列具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列，或(iv)(i)、(ii)或(iii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性實施方案中，這類eTF與一種或多種人類蛋白質具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整體序列同一性。在某些實施方案中，相比於與一種或多種人類蛋白質具有較低整體序列同一性百分比的eTF，這類eTF表現出降低的免疫原性。在各個實施方案中，免疫原性的降低可以使用elispot測定、免疫測定或電腦類比方法來測

量。

【0181】在某些實施方案中，本文公開的上調SCN1A的eTF包含以下任一序列：(i)包含SEQ ID NO: 127的序列；(ii)與SEQ ID NO: 127具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列，或(iii)(i)或(ii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF包含SEQ ID NO: 77並且與具有SEQ ID NO: 18的靶位點結合。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。

【0182】在某些實施方案中，本文公開的上調SCN1A的eTF包含以下任一序列：(i)包含SEQ ID NO: 128的序列；(ii)與SEQ ID NO: 128具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列，或(iii)(i)或(ii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF包含SEQ ID NO: 92並且與具有SEQ ID NO: 18的靶位點結合。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性實施方案中，這類eTF與一種或多種人類蛋白質具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、

96%、97%、98%或99%的整體序列同一性。在某些實施方案中，相比於與一種或多種人類蛋白質具有較低整體序列同一性百分比的eTF，這類eTF表現出降低的免疫原性。在各個實施方案中，免疫原性的降低可以使用elispot測定、免疫測定或電腦類比方法來測量。

【0183】 在某些實施方案中，本文公開的上調SCN1A的eTF包含以下任一序列：(i)包含SEQ ID NO: 129的序列；(ii)與SEQ ID NO: 129具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列，或(iii)(i)或(ii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF包含SEQ ID NO: 92並且與具有SEQ ID NO: 18的靶位點結合。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性實施方案中，這類eTF與一種或多種人類蛋白質具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整體序列同一性。在某些實施方案中，相比於與一種或多種人類蛋白質具有較低整體序列同一性百分比的eTF，這類eTF表現出降低的免疫原性。在各個實施方案中，免疫原性的降低可以使用elispot測定、免疫測定或電腦類比方法來測量。

【0184】 在某些實施方案中，本文公開的上調SCN1A的eTF包含以下任一序列：(i)包含SEQ ID NO: 130的序列；(ii)與SEQ ID NO: 130具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、

94%、95%、96%、97%、98%或99%序列同一性的序列，或(iii)(i)或(ii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF包含SEQ ID NO: 92並且與具有SEQ ID NO: 18的靶位點結合。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性實施方案中，這類eTF與一種或多種人類蛋白質具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整體序列同一性。在某些實施方案中，相比於與一種或多種人類蛋白質具有較低整體序列同一性百分比的eTF，這類eTF表現出降低的免疫原性。在各個實施方案中，免疫原性的降低可以使用elispot測定、免疫測定或電腦類比方法來測量。

【0185】 在某些實施方案中，本文公開的上調SCN1A的eTF包含以下任一序列：(i)包含SEQ ID NO: 131的序列；(ii)與SEQ ID NO: 131具有至少80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的序列，或(iii)(i)或(ii)中任一序列的功能片段或變體。在示例性實施方案中，這類eTF包含SEQ ID NO: 92並且與具有SEQ ID NO: 18的靶位點結合。在示例性實施方案中，這類eTF能夠將SCN1A表達上調與對照相比至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高，或者與對照相比至少20%、30%、40%、50%、60%、70%、75%、80%、90%、100%、125%、150%、200%、250%、300%、400%或500%或更高。在示例性實

施方案中，這類eTF與一種或多種人類蛋白質具有至少75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%的整體序列同一性。在某些實施方案中，相比於與一種或多種人類蛋白質具有較低整體序列同一性百分比的eTF，這類eTF表現出降低的免疫原性。在各個實施方案中，免疫原性的降低可以使用elispot測定、免疫測定或電腦類比方法來測量。

【0186】在某些實施方案中，本文公開的上調SCN1A的eTF包含含有gRNA/Cas複合物的DBD，其中所述gRNA包含含有SEQ ID NO: 35-66中的任一個的靶向序列。gRNA的靶序列連接到具有以下序列的支架序列的5’端：5’-GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA-3’(SEQ ID NO: 183)。在示例性實施方案中，Cas蛋白是核酸酶滅活的Cas9蛋白。在某些實施方案中，這樣的eTF進一步包含與Cas蛋白綴合的一個或多個TAD，其中該TAD選自VP64、VPR、CITED2、CITED4或CREB3。在一個實施方案中，這樣的eTF包含與Cas蛋白的C-末端綴合的VPR TAD結構域。在某些實施方案中，這樣的eTF包含與Cas蛋白的N-末端、C-末端或N-末端和C-末端綴合的CITED2 TAD。在某些實施方案中，這樣的eTF包含與Cas蛋白的N-末端、C-末端或N-末端和C-末端綴合的CITED4 TAD。在示例性實施方案中，與對照相比，這類eTF能夠將SCN1A表達上調至少2倍、5倍、10倍、15倍、20倍、25倍、30倍、40倍、50倍或更高。

多核苷酸

【0187】在另一方面，本申請提供了編碼本文公開的任何上調SCN1A的eTF的多核苷酸。在另一方面，本申請提供了包含PV選擇性微

小RNA結合位點的多核苷酸。在某些實施方案中，本申請提供了多核苷酸，其包含與轉基因可操作地連接的PV選擇性調節元件和PV選擇性微小RNA結合位點。在某些實施方案中，本申請提供了多核苷酸，其包含編碼本文公開的上調SCN1A的eTF的序列和PV選擇性微小RNA結合位點。在某些實施方案中，本申請提供了與編碼上調SCN1A的eTF的轉基因可操作地連接的PV選擇性調節元件和PV選擇性調節元件。

編碼上調SCN1A的eTF的多核苷酸

【0188】 在某些實施方案中，本申請提供了包含以下任一序列的多核苷酸：(i)編碼上調SCN1A的eTF或其變體或功能片段的核酸序列，該eTF包含SEQ ID NO: 77-131、205、207、209、213、217、219、221或223中的任一個；(ii)編碼上調SCN1A的eTF的功能片段的核酸，該eTF具有SEQ ID NO: 77-131、205、207、209、213、217、219、221或223中的任一個；或(iii)編碼上調SCN1A的eTF或其變體或功能片段的核酸，該eTF與具有SEQ ID NO: 77-131、205、207、209、213、217、219、221或223中任一個的上調SCN1A的eTF具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性。

【0189】 在某些實施方案中，本申請提供了包含以下任一序列的多核苷酸：(i)編碼DBD或其變體或功能片段的核酸序列，該DBD包含SEQ ID NO: 92-98中的任一個；(ii)編碼DBD的功能片段的核酸，該DBD具有SEQ ID NO: 92-98中的任一個；或(iii)編碼DBD或其變體或功能片段的核酸，該DBD與具有SEQ ID NO: 92-98中任一個的DBD具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、

94%、95%、96%、97%、98%、99%或更高的序列同一性，其中該DBD能夠與被SEQ ID NO: 92-98中的任一個結合的靶位點結合。

【0190】 在某些實施方案中，本申請提供了編碼上調內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)編碼包含SEQ ID NO: 103-124、128-131、205、207、209、213、217、219、221或223中任一個的eTF的核酸序列；(ii)編碼具有SEQ ID NO: 103-124、128-131、205、207、209、213、217、219、221或223中任一個的eTF的功能片段的核酸；或(iii)編碼eTF的核酸，該eTF與具有SEQ ID NO: 103-124、128-131、205、207、209、213、217、219、221或223中任一個的eTF具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性，其中該eTF能夠上調SCN1A。

【0191】 在某些實施方案中，本申請提供了編碼DBD的多核苷酸，該DBD與被本文公開的eTF結合時能夠上調內源SCN1A的基因組靶位點結合，其中該多核苷酸包含以下任一個序列：(i)編碼包含SEQ ID NO: 77-98中任一個的DBD的核酸序列；(ii)編碼具有SEQ ID NO: 77-98中任一個的DBD的功能片段的核酸；或(iii)編碼eTF的核酸，該eTF與具有SEQ ID NO: 77-98中任一個的DBD具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性，其中該DBD能夠與被SEQ ID NO: 77-98中的任一個結合的靶位點結合。

【0192】 在某些實施方案中，本申請提供了編碼DBD的多核苷酸，該DBD與被本文公開的eTF結合時能夠上調內源SCN1A的基因組靶位點

結合，其中該多核苷酸包含以下任一個序列：(i)編碼包含SEQ ID NO: 148-151中任一個的DBD的核酸序列；(ii)編碼具有SEQ ID NO: 148-151中任一個的DBD的功能片段的核酸；或(iii)編碼eTF的核酸，該eTF與具有SEQ ID NO: 148-151中任一個的DBD具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性，其中該DBD能夠與被SEQ ID NO: 92-98中的任一個結合的靶位點結合。

【0193】在某些實施方案中，本申請提供了編碼能夠調節內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 70-76或184中的任一個的核酸序列；(ii)具有(i)的任一序列的功能片段的核酸；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸，其中該多核苷酸編碼能夠上調SCN1A的eTF。

【0194】在某些實施方案中，本申請提供了編碼能夠調節內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 70的核酸序列；(ii)具有SEQ ID NO: 70的功能片段的核酸；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性實施方案中，這類多核苷酸編碼具有SEQ ID NO: 127的eTF，或其能夠上調SCN1A的功能片段或變體。

【0195】在某些實施方案中，本申請提供了編碼能夠調節內源

SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 71的核酸序列；(ii)具有SEQ ID NO: 71的功能片段的核酸；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性實施方案中，這類多核苷酸編碼具有SEQ ID NO: 127的eTF，或其能夠上調SCN1A的功能片段或變體。

【0196】在某些實施方案中，本申請提供了編碼能夠調節內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 72的核酸序列；(ii)具有SEQ ID NO: 72的功能片段的核酸；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性實施方案中，這類多核苷酸編碼具有SEQ ID NO: 130的eTF，或其能夠上調SCN1A的功能片段或變體。

【0197】在某些實施方案中，本申請提供了編碼能夠調節內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 73的核酸序列；(ii)具有SEQ ID NO: 73的功能片段的核酸序列；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性實施方案中，這類多核苷酸編碼具有SEQ ID NO: 131的eTF，或其能夠上調SCN1A的功能片段或變體。

【0198】在某些實施方案中，本申請提供了編碼能夠調節內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 74的核酸序列；(ii)具有SEQ ID NO: 74的功能片段的核酸序列；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性實施方案中，這類多核苷酸編碼具有SEQ ID NO: 127的eTF，或其能夠上調SCN1A的功能片段或變體。

【0199】在某些實施方案中，本申請提供了編碼能夠調節內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 75的核酸序列；(ii)具有SEQ ID NO: 75的功能片段的核酸序列；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性實施方案中，這類多核苷酸編碼具有SEQ ID NO: 127的eTF，或其能夠上調SCN1A的功能片段或變體。

【0200】在某些實施方案中，本申請提供了編碼能夠調節內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 76的核酸序列；(ii)具有SEQ ID NO: 76的功能片段的核酸序列；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性實施方案中，這類多核苷酸編碼具有SEQ ID NO: 106的eTF，或其能夠上調SCN1A的功能

片段或變體。

【0201】在某些實施方案中，本申請提供了編碼能夠調節內源SCN1A的eTF的多核苷酸，其中該多核苷酸包含以下任一個序列：(i)具有SEQ ID NO: 184的核酸序列；(ii)具有SEQ ID NO: 184的功能片段的核酸序列；或(iii)與(i)或(ii)的任一序列具有至少70%、75%、80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核酸。在示例性實施方案中，這類多核苷酸編碼具有SEQ ID NO: 106的eTF，或其能夠上調SCN1A的功能片段或變體。

用於在PV神經元中選擇性表達的包含微小RNA結合位點的多核苷酸

【0202】在另一方面，本申請提供了包含微小RNA結合位點的多核苷酸，該微小RNA結合位元元點導致目的基因在小白蛋白(PV)神經元中選擇性表達。微小RNA或miRNA是小的非編碼RNA(約20個核苷酸)，其通過與mRNA分子內的互補識別位點雜交而在轉錄後調節基因表達，並且通過促進mRNA轉錄物的降解或通過抑制被mRNA編碼的蛋白質的翻譯而導致基因表達的抑制。本文提供的微小RNA結合位元元點抑制目的基因在興奮性神經元中的表達，從而促進目的基因在PV神經元中的選擇性表達(例如，PV選擇性微小RNA結合位點)。在某些實施方案中，興奮性神經元是表達STAC、Slc17a7、Car12、Syt17、ITPKA、Col6a1、CamKII、Sv2b、INHBA和/或DKK3中的一種或多種的神經元。在示例性實施方案中，興奮性神經元是表達CamKII的神經元。

【0203】在某些實施方案中，本申請提供了包含針對一種或多種微小RNA的一個或多個微小RNA結合位點的多核苷酸，所述微小RNA促進

PV選擇性表達，例如，促進包含該微小RNA結合位點的mRNA在興奮性神經元中的降解。促進PV選擇性表達的示例性微小RNA包括，例如，miR-128、miR-221和miR-222。在某些實施方案中，本申請提供了包含至少1、2、3、4、5、6、7、8、9、10個或更多個PV選擇性微小RNA結合位點的多核苷酸。在一個實施方案中，本申請提供了包含至少1、2、3、4、5、6、7、8、9、10個或更多個miR-128結合位點(SEQ ID NO: 9)的多核苷酸。在一個實施方案中，本申請提供了包含至少1、2、3、4、5、6、7、8、9、10個或更多個miR-221結合位點(SEQ ID NO: 11)的多核苷酸。在一個實施方案中，本申請提供了包含至少1、2、3、4、5、6、7、8、9、10個或更多個miR-222結合位點(SEQ ID NO: 13)的多核苷酸。在一個實施方案中，本申請提供了包含至少1個miR-128結合位點、至少1個miR-221結合位點和至少1個miR-222結合位點的多核苷酸。在一個實施方案中，本申請提供了包含至少1個miR-128結合位點和至少1個miR-222結合位點的多核苷酸。在一個實施方案中，本申請提供了包含至少1個miR-221結合位點和至少1個miR-222結合位點的多核苷酸。在示例性實施方案中，本申請提供了包含至少1個miR-128結合位點(SEQ ID NO: 9)和至少1個miR-221結合位點(SEQ ID NO: 11)的多核苷酸。在一個實施方案中，本申請提供了包含至少2個miR-128結合位點(SEQ ID NO: 9)和至少2個miR-221結合位點(SEQ ID NO: 11)的多核苷酸。在一個實施方案中，本申請提供了包含至少3個miR-128結合位點(SEQ ID NO: 9)和至少3個miR-221結合位點(SEQ ID NO: 11)的多核苷酸。在一個實施方案中，本申請提供了包含至少4個miR-128結合位點(SEQ ID NO: 9)和至少4個miR-221結合位點(SEQ ID NO: 11)的多核苷酸。在一個實施方案中，

本申請提供了包含至少5個miR-128結合位點(SEQ ID NO: 9)和至少5個miR-221結合位點(SEQ ID NO: 11)的多核苷酸。在這樣的實施方案中，所述結合位元點可以以任何順序排列。例如，對於包含2個miR-128結合位點和2個miR-221結合位點的構建體，所述結合位點可以按以下任何構型排列：miR-128 – miR-128 – miR-221 – miR221、miR-128 – miR-221 – miR-128 – miR-221、miR-128 – miR221 – miR221 – miR-128、miR-221 – miR128 – miR221 – miR128、miR-221 – miR128 – miR128 – miR221或miR221 – miR221- miR128 – miR128。在示例性實施方案中，本文提供的多核苷酸包含具有4個miR-128結合位點(SEQ ID NO: 9)及隨後的4個miR-221結合位點(SEQ ID NO: 11)的序列，例如，miR-128 – miR-128 – miR128 – miR-128 – miR221 – miR221 – miR-221 – miR221。在另一個示例性實施方案中，本文提供的多核苷酸包含具有1個miR-221序列(SEQ ID NO: 11)、1個miR-222序列(SEQ ID NO: 13)和1個miR-128結合位點(SEQ ID NO: 9)的序列，例如，miR-221 – miR222 – miR128。在另一個示例性實施方案中，本文提供的多核苷酸包含具有按以下順序排列的2個miR-221序列(SEQ ID NO: 11)、2個miR-222序列(SEQ ID NO: 13)和2個miR-128結合位點(SEQ ID NO: 9)的序列：miR-221 – miR222 – miR128 - miR-221 – miR222 – miR128。

【0204】在具有超過一個微小RNA結合位點的多核苷酸中，所述結合位點可以在多核苷酸序列中彼此直接相鄰(例如，在結合位點之間沒有連接體或間插序列)，或者可以彼此相隔至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20個或更多個核苷酸，或1-20、1-15、1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3或1-2個

核苷酸。在示例性實施方案中，微小RNA結合位點相隔約5個核苷酸或相隔5個核苷酸。在示例性實施方案中，隔開微小RNA結合位點的序列(以及在微小RNA結合位點之間形成的連接點，或在微小RNA結合位點與隔開微小RNA結合位點的序列之間形成的連接點)與其他任何微小RNA或其他任何神經元微小RNA都不互補。

【0205】 在某些實施方案中，本文提供的多核苷酸包含與SEQ ID NO: 7具有至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的微小RNA結合位點。在示例性實施方案中，本文提供的多核苷酸包含含有SEQ ID NO: 7的微小RNA結合位點。

【0206】 在某些實施方案中，本文提供的多核苷酸包含與SEQ ID NO: 14具有至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的微小RNA結合位點。在示例性實施方案中，本文提供的多核苷酸包含含有SEQ ID NO: 14的微小RNA結合位點。

【0207】 在某些實施方案中，本文提供的多核苷酸包含與SEQ ID NO: 15具有至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的微小RNA結合位點。在示例性實施方案中，本文提供的多核苷酸包含含有SEQ ID NO: 15的微小RNA結合位點。

【0208】 在某些實施方案中，本文提供的微小RNA結合位點位於mRNA轉錄物的3’非翻譯區內，例如，在翻譯終止密碼子(即，TAA、TGA或TAG)之後並且在聚A尾之前。微小RNA結合位點可以直接定位成

與翻譯終止密碼子相鄰，或者可以與翻譯終止密碼子相隔至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20個或更多個核苷酸，或1-20、1-15、1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3或1-2個核苷酸，並且/或者可以與聚A尾相鄰，或者可以與聚A尾相隔至少1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20個或更多個核苷酸，或1-20、1-15、1-10、1-9、1-8、1-7、1-6、1-5、1-4、1-3或1-2個核苷酸。

【0209】 在某些實施方案中，本文提供的微小RNA結合位點導致與脫靶細胞類型相比，在PV細胞中的選擇性基因表達。在一些情況下，脫靶細胞類型包括但不限於興奮性神經元、非PV CNS細胞類型和非神經元CNS細胞類型。在某些實施方案中，PV選擇性微小RNA結合位點導致在PV神經元中相對於在至少一種、兩種、三種、四種、五種或更多種非PV CNS細胞類型中的選擇性基因表達。在一些情況下，非PV CNS細胞是興奮性神經元、多巴胺能神經元、星形膠質細胞、小神經膠質細胞、運動神經元、血管細胞或非GABA能神經元(例如，不表達GAD2、GAD1、NKX2.1、DLX1、DLX5、SST和VIP中的一種或多種的細胞)、非PV神經元(例如，不表達小白蛋白的GABA能神經元)或其他CNS細胞(例如，從未表示任何PV、GAD2、GAD1、NKX2.1、DLX1、DLX5、SST和VIP的CNS細胞類型)。在示例性實施方案中，本文提供的PV選擇性微小RNA結合位點通過降低在興奮性神經元中的表達而導致與興奮性神經元(例如，表達STAC、Slc17a7、Car12、Syt17、ITPKA、Col6a1、CamKII、Sv2b、INHBA和/或DKK3中的一種或多種的神經元)相比，在PV神經元中的基因表達選擇性提高。在一些情況下，通過具有不同的細胞標誌物、

形態學、表型、基因型、功能和/或用於細胞類型分類的其他任何手段來區分細胞類型。

【0210】由PV選擇性微小RNA結合位點驅動的表達的選擇性可以用多種方式來測量。在一個實施方案中，可以通過將從包含PV選擇性微小RNA結合位點的基因表達可檢測水準的轉錄物的PV細胞的數目與表達該基因的細胞總數進行比較(例如，表達該基因的PV與總細胞(PV+非PV細胞)之比)，來測量PV細胞中相對於非PV細胞中基因表達的選擇性。例如，可以使用基於免疫組織化學的共定位測定來確定對PV神經元的選擇性，該共定位測定使用包含編碼螢光蛋白(例如，eGFP)的基因和PV選擇性微小RNA結合位點的表達盒來測量基因表達，並使用與第二螢光標記物(例如，紅色螢光蛋白)連接的鑒定PV細胞的抗體(例如，與PV神經元特異性相互作用的抗PV抗體)。PV細胞中的表達的選擇性可以如下計算：將表達PV和eGFP兩者的細胞(例如，PV細胞)數除以表達eGFP的細胞(例如，PV細胞和非PV細胞)總數，然後乘以100以轉換成百分比。在另一個實例中，可以使用基於免疫組織化學的共定位測定來確定對PV神經元的選擇性，該共定位測定使用包含編碼螢光蛋白(例如，eGFP)的基因和PV選擇性微小RNA結合位點的表達盒來測量基因表達，並使用與第二螢光標記物(例如，紅色螢光蛋白)連接的鑒定PV細胞的第一抗體(例如，與PV神經元特異性相互作用的抗PV抗體)，及鑒定興奮性細胞的第二抗體(例如，與興奮性神經元特異性相互作用的抗CamKII抗體)。PV細胞中的表達的選擇性可以如下計算：將表達PV和eGFP兩者的細胞(例如，PV細胞)數除以表達eGFP + PV和eGFP + CamKII的細胞(例如，PV細胞和興奮性細胞)數，然後乘以100以轉換成百分比。表達轉基因的PV細胞的百分比

越高，微小RNA結合位點對PV細胞的選擇性就越高。在某些實施方案中，本文提供的PV選擇性微小RNA結合位點對於在PV細胞中的表達可能是高度選擇性的。例如，本文提供的PV選擇性微小RNA結合位點可表現出約50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或高於約99%的對PV神經元的選擇性(例如，PV神經元/總細胞x100或PV神經元/PV+興奮性神經元x100)。

【0211】 在一些情況下，本文提供的PV選擇性微小RNA結合位點是短的。在一些情況下，PV選擇性微小RNA結合位點的大小與載體(例如，病毒載體或rAAV)的克隆容量相容，使得轉基因、啟動子(和可選的增強子)和微小RNA結合位元點的組合大小不超過載體的克隆容量。在一些情況下，PV選擇性微小RNA結合位點的長度可達約500bp、400bp、300bp、250 bp、225 bp、215 bp、210 bp、200bp、150 bp、140 bp、135 bp、130 bp、125 bp、120 bp、115 bp、110 bp、100bp、90 bp、80 bp、75 bp、70 bp、65 bp、60 bp或50 bp。在一些情況下，PV選擇性微小RNA結合位點為約50-500 bp、50-400 bp、50-300 bp、50-250 bp、50-200 bp、50-100 bp、50-75 bp、50-70 bp、100-500 bp、100-400 bp、100-300 bp、100-250 bp、100-200 bp、100-150 bp、100-140 bp、100-135 bp、200-500 bp、200-400 bp、200-300 bp或200-250 bp。

【0212】 在示例性實施方案中，本文提供的包含一個或多個PV選擇性微小RNA結合位點的多核苷酸不包含SEQ ID NO: 67。

表達盒

【0213】 在另一方面，本申請提供了包含本文提供的多核苷酸(例

如，包含編碼上調SCN1A和/或含有PV選擇性微小RNA結合位點的eTF的序列的多核苷酸)和一個或多個調節元件的表達盒。在某些實施方案中，本申請提供了包含本文提供的多核苷酸(例如，包含編碼上調SCN1A和/或含有PV選擇性微小RNA結合位點的eTF的序列的多核苷酸)和PV選擇性啟動子的表達盒。

【0214】 在某些實施方案中，本文提供的多核苷酸(例如，包含編碼上調SCN1A和/或含有PV選擇性微小RNA結合位點的eTF的序列的多核苷酸)是除編碼eTF的序列外還包含一個或多個調節元件的表達盒的一部分。在示例性實施方案中，本文提供的多核苷酸(例如，包含編碼上調SCN1A和/或含有PV選擇性微小RNA結合位點的eTF的序列的多核苷酸)是包含啟動子的表達盒的一部分，該啟動子位於轉基因序列的上游，從而能夠驅動該轉基因(例如，選擇性上調SCN1A的eTF)在細胞中的表達。

【0215】 在某些實施方案中，本文公開的表達盒包含本文提供的多核苷酸(例如，包含編碼上調SCN1A和/或含有PV選擇性微小RNA結合位點的eTF的序列的多核苷酸)和組成型啟動子，該組成型啟動子位於編碼轉基因的序列的上游，從而能夠驅動該轉基因(例如，選擇性上調SCN1A的eTF)在細胞中的表達。組成型啟動子的實例包括GAD2啟動子、人突觸蛋白啟動子、CBA啟動子、CMV啟動子、minCMV啟動子、TATA盒、超核心啟動子或EF1α啟動子，或其組合。

【0216】 在某些實施方案中，本文公開的表達盒包含本文提供的多核苷酸(例如，包含編碼上調SCN1A和/或含有PV選擇性微小RNA結合位點的eTF的序列的多核苷酸)和能夠驅動轉基因(例如，選擇性上調SCN1A的eTF)在細胞中的表達的短啟動子。在某些實施方案中，適合根據本文描

述的核酸分子使用的短啟動子包含小於500 bp、450 bp、400 bp、350 bp、300 bp、250 bp、225 bp、200 bp、175 bp、150 bp、145 bp、140 bp、135 bp、130 bp、125 bp、120 bp、115 bp、110 bp、105 bp、100 bp、95 bp、90 bp、85 bp、80 bp或75 bp，或約80-300 bp、80-275 bp、80-250 bp、80-200 bp、80-150 bp、80-125 bp、80-120 bp、80-115 bp、80-110 bp、80-105 bp、80-100 bp、85-300 bp、85-275 bp、85-250 bp、85-200 bp、85-150 bp、85-125 bp、85-120 bp、85-115 bp、85-110 bp、85-105 bp、85-100 bp、90-300 bp、90-275 bp、90-250 bp、90-200 bp、90-150 bp、90-125 bp、90-120 bp、90-115 bp、90-110 bp、90-105 bp、90-100 bp、95-300 bp、95-275 bp、95-250 bp、95-200 bp、95-150 bp、95-125 bp、95-120 bp、95-115 bp、95-110 bp、95-105 bp、95-100 bp、100-300 bp、100-275 bp、100-250 bp、100-200 bp、100-150 bp、100-125 bp、100-120 bp、100-115 bp、100-110 bp或100-105 bp。在示例性實施方案中，適合根據本文描述的表達盒使用的短啟動子包含約100-120 bp、約117 bp或約100 bp。

【0217】 在某些實施方案中，本文公開的表達盒包含短啟動子，該啟動子包含以下任一序列或由以下任一序列組成：(i)SEQ ID NO: 1；(ii)其變體或功能片段；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，該核酸序列與編碼本文公開的選擇性上調SCN1A的任何一種eTF的多核苷酸可操作地連接，並且該eTF任選地包含本文公開的微小RNA結合位點。短啟動子序列的其他實例可見於PCT公開WO 2018/213786。

【0218】在某些實施方案中，本文公開的表達盒包含本文提供的多核苷酸(例如，包含編碼上調SCN1A和/或含有PV選擇性微小RNA結合位點的eTF的序列的多核苷酸)和細胞類型選擇性啟動子，該啟動子位於編碼轉基因(例如，選擇性上調SCN1A的eTF)的序列的上游，從而能夠驅動該轉基因在目的細胞中的選擇性表達。在某些實施方案中，細胞類型選擇性啟動子可以對任何目的細胞類型(例如，心臟細胞、肝細胞、肌肉細胞、骨細胞、神經元或其亞群)具有選擇性(例如，選擇性地驅動在其中的表達)。在示例性實施方案中，本文公開的表達盒包含編碼選擇性上調SCN1A的eTF的多核苷酸和位於編碼eTF的序列上游的PV選擇性調節元件(例如，啟動子、增強子和/或啟動子和增強子)，以便能夠驅動eTF在PV細胞中的選擇性表達，並且任選地包含PV選擇性微小RNA結合位點。PV選擇性調節元件是指特異性調節在PV神經元中的基因表達的調節元件。在某些實施方案中，相對於一種或多種其他CNS細胞類型，PV選擇性調節元件增強在PV神經元中的表達。在某些實施方案中，PV選擇性調節元件抑制脫靶細胞類型中的轉錄和/或翻譯過程。

【0219】在某些實施方案中，與脫靶細胞類型相比，本文提供的PV選擇性調節元件導致在PV細胞中的選擇性基因表達。在一些情況下，脫靶細胞類型包括但不限於興奮性神經元、非PV CNS細胞類型和非神經元CNS細胞類型。在某些實施方案中，PV選擇性調節元件導致在PV神經元中相對於在至少一種、兩種、三種、四種、五種或更多種非PV CNS細胞類型中的選擇性基因表達。在一些情況下，非PV CNS細胞是興奮性神經元、多巴胺能神經元、星形膠質細胞、小神經膠質細胞、運動神經元、血管細胞或非GABA能神經元(例如，不表達GAD2、GAD1、NKX2.1、

DLX1、DLX5、SST和VIP中的一種或多種的細胞)、非PV神經元(例如，不表達小白蛋白的GABA能神經元)或其他CNS細胞(例如，從未表示任何PV、GAD2、GAD1、NKX2.1、DLX1、DLX5、SST和VIP的CNS細胞類型)。在一些情況下，本文提供的PV選擇性調節元件導致在PV神經元中與在非PV GABA能細胞中相比，基因表達的選擇性增加。在一些情況下，通過具有不同的細胞標誌物、形態學、表型、基因型、功能和/或用於細胞類型分類的其他任何手段來區分細胞類型。

【0220】 由PV選擇性調節元件驅動的表達的選擇性可以用多種方式來測量。在一個實施方案中，可以通過將從與PV選擇性調節元件可操作地連接的基因表達可檢測水準的轉錄物的PV細胞的數目與表達該基因的細胞總數進行比較(例如，PV與表達該基因的總細胞(PV+非PV細胞)之比)，來測量PV細胞中相對於非PV細胞中基因表達的選擇性。例如，可以使用基於免疫組織化學的共定位測定來確定對PV神經元的選擇性，該共定位測定使用與PV選擇性調節元件可操作地連接的編碼螢光蛋白(例如，eGFP)的基因來測量基因表達，並使用與第二螢光標記物(例如，紅色螢光蛋白)連接的鑒定PV細胞的抗體(例如，與PV神經元特異性相互作用的抗PV抗體)。PV細胞中的表達的選擇性可以如下計算：將表達PV和eGFP兩者的細胞(例如，PV細胞)數除以表達eGFP的細胞(例如，PV細胞和非PV細胞)總數，然後乘以100以轉換成百分比。表達轉基因的PV細胞的百分比越高，調節元件對PV細胞的選擇性就越高。在某些實施方案中，本文提供的PV選擇性調節元件對於在PV細胞中的表達可能是高度選擇性的。例如，本文提供的PV選擇性調節元件可表現出約50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、

96%、97%、98%、99%或高於約99%的對PV神經元的選擇性(例如，PV神經元/總細胞x 100)。

【0221】 在一些情況下，本文提供的PV選擇性調節元件是短的。在一些情況下，PV選擇性調節元件的大小與載體(例如，病毒載體或rAAV)的克隆容量相容，使得轉基因和一個或多個PV選擇性調節元件的組合大小不超過載體的克隆容量。在一些情況下，PV選擇性調節元件的長度可達約2050bp、2000bp、1900bp、1800bp、1700bp、1600bp、1500bp、1400bp、1300bp、1200bp、1100bp、1000bp、900bp、800bp、700bp、600bp、500bp、400bp、300bp、200bp或100bp。在一些情況下，PV選擇性調節元件為約500-600bp、500-700bp、500-800bp、500-900bp、500-1000bp、500-1500bp、500-2000bp或500-2050bp。

【0222】 在某些實施方案中，本文提供的PV選擇性調節元件包含以下任一序列或由以下任一序列組成：(i)SEQ ID NO: 2-4；(ii)其變體、功能片段或組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在一些情況下，調節元件包含SEQ ID NO: 2-4中的任一個。PV選擇性調節元件的其他實例可見於PCT公開WO 2018/187363。

【0223】 在示例性實施方案中，本申請提供了表達盒，其包含在PV選擇性調節元件的控制下編碼選擇性上調SCN1A的eTF的核酸序列。在某些實施方案中，本申請提供了表達盒，其包含在具有SEQ ID NO: 2-4中任一個的PV選擇性調節元件控制下編碼選擇性上調SCN1A的eTF的核酸序列，該eTF包含具有以下任一序列的DBD：SEQ ID NO: 77-98。在某些實

施方案中，本申請提供了表達盒，其包含在具有SEQ ID NO: 2-4中任一個的PV選擇性調節元件控制下編碼選擇性上調SCN1A的eTF的核酸序列，該eTF包含以下任一序列：SEQ ID NO: 99-131、205、207、209、213、217、219、221或223。在某些實施方案中，本申請提供了表達盒，其包含在具有SEQ ID NO: 2-4中任一個的PV選擇性調節元件控制下的含有以下任一序列的核酸序列：SEQ ID NO: 67-73。在某些實施方案中，本申請提供了表達盒，其包含在具有SEQ ID NO: 2中任一個的PV選擇性調節元件控制下編碼選擇性上調SCN1A的eTF的核酸序列，該eTF包含具有以下任一序列的DBD：SEQ ID NO: 148-151。在某些實施方案中，本申請提供了表達盒，其包含在具有SEQ ID NO: 2中任一個的PV選擇性調節元件控制下編碼選擇性上調SCN1A的eTF的核酸序列，該eTF包含以下任一序列：SEQ ID NO: 99-131、205、207、209、213、217、219、221或223。在某些實施方案中，本申請提供了表達盒，其包含在具有SEQ ID NO: 2中任一個的PV選擇性調節元件控制下的含有以下任一序列的核酸序列：SEQ ID NO: 67-76或316。

【0224】在某些實施方案中，本申請提供了包含PV選擇性微小RNA結合位點以及啟動子和/或增強子序列的表達盒。該表達盒中可包括本文所述的任何啟動子。在示例性實施方案中，本文提供的表達盒包含PV選擇性微小RNA結合位點和PV選擇性調節元件。在某些實施方案中，本文提供的表達盒包含PV選擇性微小RNA結合位點和PV選擇性調節元件，該PV選擇性調節元件包含(i)SEQ ID NO: 2-4中的任一個；(ii)其變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、

92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在某些實施方案中，本文提供的表達盒包含(1)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 7、14或15中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(2)PV選擇性調節元件，其包含(i)SEQ ID NO: 2-4中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在某些實施方案中，本文提供的表達盒包含(1)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 7；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(2)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在一個實施方案中，本文提供的表達盒包含含有SEQ ID NO: 7的微小RNA結合位點和含有SEQ ID NO: 2的PV選擇性調節元件。在某些實施方案中，本文提供的多核苷酸包含(1)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 14；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有

至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(2)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在一個實施方案中，本文提供的表達盒包含含有SEQ ID NO: 15的微小RNA結合位點和含有SEQ ID NO: 2的PV選擇性調節元件。在某些實施方案中，本文提供的表達盒包含(1)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 15；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(2)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在一個實施方案中，本文提供的表達盒包含含有SEQ ID NO: 15的微小RNA結合位點和含有SEQ ID NO: 2的PV選擇性調節元件。

【0225】 在某些實施方案中，本文提供的表達盒包含PV選擇性微小RNA結合位元點和編碼本文提供的上調SCN1A表達的eTF的序列。在示例性實施方案中，本申請提供了一種包含PV選擇性調節元件、如本文提供的上調SCN1A表達的eTF和PV選擇性微小RNA結合位點的表達盒。在

示例性實施方案中，本申請提供了一種表達盒，其包含(1)PV選擇性調節元件，其包含(i)SEQ ID NO: 2-4中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，(2)編碼上調SCN1A的eTF的序列，其包含(i)SEQ ID NO: 77-131、205、207、209、213、217、219、221或223中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(3)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 7、14或15中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在另一個示例性實施方案中，本申請提供了一種表達盒，其包含(1)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，(2)編碼上調SCN1A的eTF的序列，其包含(i)SEQ ID NO: 77或127中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的

核酸序列，以及(3)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 7；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在另一個示例性實施方案中，本申請提供了一種表達盒，其包含(1)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，(2)編碼上調SCN1A的eTF的序列，其包含(i)SEQ ID NO: 77或127中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(3)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 14；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在另一個示例性實施方案中，本申請提供了一種表達盒，其包含(1)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，(2)編碼上調SCN1A的eTF的序列，其包含(i)SEQ ID NO: 77或127中

的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(3)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 15；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在另一個示例性實施方案中，本申請提供了一種表達盒，其包含(1)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，(2)編碼上調SCN1A的eTF的序列，其包含(i)SEQ ID NO: 92或106中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(3)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 7；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在另一個示例性實施方案中，本申請提供了一種表達盒，其包含(1)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中

的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，(2)編碼上調SCN1A的eTF的序列，其包含(i)SEQ ID NO: 92或106中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(3)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 14；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列。在另一個示例性實施方案中，本申請提供了一種表達盒，其包含(1)PV選擇性調節元件，其包含(i)SEQ ID NO: 2；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，(2)編碼上調SCN1A的eTF的序列，其包含(i)SEQ ID NO: 92或106中的任一個；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列同一性的核酸序列，以及(3)PV選擇性微小RNA結合位點，其包含(i)SEQ ID NO: 15；(ii)(i)的變體、功能片段或其組合；或(iii)與(i)或(ii)中的任一個具有至少80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、

94%、95%、96%、97%、98%或99%序列同一性的核酸序列。

【0226】在某些實施方案中，本文提供的包含PV選擇性調節元件和PV選擇性微小RNA結合位點的表達盒在大小上小於5 kb、4.9 kb、4.8 kb、4.7 kb、4.6 kb、4.5 kb、4.4 kb、4.3 kb、4.2 kb、4.1 kb、4.0 kb、3.9 kb、3.8 kb、3.7 kb、3.6 kb、3.5 kb、3.4 kb、3.3 kb、3.2 kb、3.1 kb、3.0 kb、2.9 kb、2.8 kb、2.7 kb、2.6 kb、2.5 kb、2.4 kb、2.3 kb、2.2 kb、2.1 kb、2.0 kb、1.9 kb、1.8 kb、1.7 kb、1.6 kb或1.5 kb或更小，或者約為1.5-5 kb、1.5-4.7 kb、1.5-4.5 kb、1.5-4.0 kb、1.5-3.5 kb、1.5-3.0 kb、1.5-2.5 kb、1.5-2.0 kb。

【0227】在某些實施方案中，本文提供的表達盒除了啟動子外還可包含一個或多個額外的調節元件，例如與轉錄起始或終止相關的序列、增強子序列和有效的RNA加工信號。示例性的調節元件包括，例如，內含子、增強子、UTR、穩定性元件、WPRE序列、Kozak共有序列、翻譯後響應元件、微小RNA結合位點或多腺苷酸化(聚A)序列或其組合。調節元件可以用來在基因表達的轉錄階段、轉錄後階段或翻譯階段調節基因表達。在RNA水準上，調節可以在翻譯(例如，穩定mRNA以供翻譯的穩定性元件)、RNA切割、RNA剪接和/或轉錄終止的水準上發生。在各個實施方案中，調節元件可以將轉錄因子募集到編碼區，該轉錄因子增加在目的細胞類型中的基因表達選擇性，提高產生RNA轉錄物的速率，提高產生的RNA的穩定性，並且/或者提高從RNA轉錄物合成蛋白質的速率。在示例性實施方案中，本文提供的表達盒包含至少一個如本文提供的PV選擇性微小RNA結合位點。

【0228】在某些實施方案中，本文所述的表達盒進一步包含聚A序

列。合適的聚A序列包括，例如，長度約為75 bp的人工聚A(PA75)(參見，例如，WO 2018/126116)、牛生長激素聚A、SV40早期聚A信號、SV40晚期聚A信號、兔β珠蛋白聚A、HSV胸苷激酶聚A、魚精蛋白基因聚A、腺病毒5 EIb聚A、生長激素聚A或PBGD聚A。在示例性實施方案中，適用於本文提供的表達盒的聚A序列是hGH聚A(SEQ ID NO: 17)或合成聚A(SEQ ID NO: 16)。通常，聚A序列在本文所述的表達盒中位於編碼eTF的多核苷酸的下游。

【0229】 在某些實施方案中，本文提供的表達盒進一步包含編碼一種或多種核定位元信號(NLS)的一種或多種核酸序列。可以使用任何有助於將與其附接的蛋白質輸入到細胞核中的NLS肽。NLS的實例包括，例如，SV40大T抗原NLS、核質蛋白NLS、EGL-13 NLS、c-Myc NLS和TUS蛋白NLS。參見，例如，C. Dingwall等人, J. Cell Biol. 107: 841-9 (1988)；J.P. Makkerh等人, Curr Biol. 6: 1025-7 (1996)；和M. Ray等人, Bioconjug. Chem. 26: 1004-7 (2015)。NLS可以位於eTF蛋白序列上的任何位置，但是在優選的實施方案中，其與eTF的N-末端或eTF的結構域綴合。在示例性實施方案中，本文提供的核酸盒編碼具有與eTF的N-末端融合的NLS的eTF。在其他實施方案中，本文提供的核酸盒編碼具有與eTF的N-末端融合的第一NLS和位於eTF的DBD與TAD結構域之間的第二NLS的eTF。

表達載體

【0230】 在某些實施方案中，本文所述的表達盒可以併入表達載體中。表達載體可以用來通過轉染或轉導將表達盒遞送至靶細胞。載體可以是整合載體或非整合載體，是指載體將表達盒或轉基因整合到宿主細胞基

因組中的能力。表達載體的實例包括但不限於：(a)非病毒載體，如核酸載體，包括線性寡核苷酸和環狀質粒；人工染色體，如人類人工染色體(HAC)、酵母人工染色體(YAC)和細菌人工染色體(BAC或PAC)；附加型載體；轉座子(例如，PiggyBac)；和(b)病毒載體，如逆轉錄病毒載體、慢病毒載體、腺病毒載體和腺伴隨病毒載體。

【0231】 表達載體可以是線性寡核苷酸或環狀質粒，並且可以通過各種轉染方法(包括物理和化學方法)遞送至細胞。物理方法通常是指利用物理力對抗細胞膜屏障以促進遺傳物質的細胞內遞送的遞送方法。物理方法的實例包括使用針、彈道DNA、電穿孔、聲孔效應、光穿孔、磁轉染和氫穿孔(hydroporation)。化學方法通常是指其中化學運載體將核酸分子遞送至細胞的方法，並且可以包括無機顆粒、基於脂質的載體、基於聚合物的載體和基於肽的載體。

【0232】 在一些實施方案中，使用陽離子脂質(例如，陽離子脂質體)將表達載體施用於靶細胞。已經研究了用於基因遞送的各種類型的脂質，例如脂質納米乳液(例如，其為一種不混溶液體在另一種不混溶液體中的分散體，被乳化劑穩定)或固體脂質納米顆粒。

【0233】 在一些實施方案中，使用基於肽的遞送媒介物將表達載體施用於靶細胞。基於肽的遞送媒介物可具有保護待遞送的遺傳物質、靶向特定細胞受體、破壞內體膜和將遺傳物質遞送至細胞核中的優點。在一些實施方案中，使用基於聚合物的遞送媒介物將表達載體施用於靶細胞。基於聚合物的遞送媒介物可包含天然蛋白質、肽和/或多糖或合成聚合物。在一個實施方案中，基於聚合物的遞送媒介物包含聚乙烯亞胺(PEI)。PEI可以將DNA濃縮成帶正電荷的顆粒，該顆粒與陰離子細胞表面殘基結

合，並通過內吞作用被帶入細胞。在其他實施方案中，基於聚合物的遞送媒介物可以包含聚-L-賴氨酸(PLL)、聚(DL-乳酸)(PLA)、聚(DL-丙交酯-共-糖苷)(PLGA)、聚鳥氨酸、聚精氨酸、組蛋白、魚精蛋白、樹狀聚體、殼聚糖、葡聚糖的合成氨基衍生物和/或陽離子丙烯酸聚合物。在某些實施方案中，基於聚合物的遞送媒介物可包含諸如PEG和PLL等聚合物的混合物。

【0234】 在某些實施方案中，表達載體可以是適合基因療法的病毒載體。病毒基因療法載體或基因遞送載體的優選特徵可包括可再現地且穩定地增殖並純化至高滴度的能力；介導靶向遞送(例如，將轉基因特異性遞送至目的組織或器官而沒有在他處的廣泛載體散播)的能力；以及介導基因遞送和轉基因表達而不會誘發有害副作用的能力。

【0235】 通過利用病毒感染途徑但避免隨後會導致複製和毒性的病毒基因表達，已經為了基因療法的目的工程改造了幾種類型的病毒，例如被稱為腺伴隨病毒的非致病性細小病毒。可以通過從病毒基因組中刪除全部或部分編碼區，但完整保留對於例如將載體基因組包裝到病毒衣殼中或將載體核酸(例如DNA)整合至宿主染色質中等功能而言可能必需的序列(例如，末端重複序列)，從而獲得這樣的病毒載體。

【0236】 在各個實施方案中，合適的病毒載體包括逆轉錄病毒(例如，A型、B型、C型和D型病毒)、腺病毒、細小病毒(例如，腺伴隨病毒或AAV)、冠狀病毒、負鏈RNA病毒如正粘病毒(例如，流感病毒)、彈狀病毒(例如狂犬病和水皰性口炎病毒)、副粘病毒(例如麻疹和仙台病毒)、正鏈RNA病毒如小RNA病毒和甲病毒，以及雙鏈DNA病毒，包括腺病毒、皰疹病毒(例如單純皰疹病毒1型和2型、EB病毒、巨細胞病毒)和痘

病毒(例如牛痘、禽痘和金絲雀痘)。逆轉錄病毒的實例包括禽造白細胞組織增生-肉瘤病毒、人嗜T-淋巴細胞病毒1型(HTLV-1)、牛白血病病毒(BLV)、慢病毒和泡沫病毒。例如，其他病毒包括Norwalk病毒、披膜病毒、黃病毒、呼腸孤病毒、乳多空病毒、嗜肝DNA病毒和肝炎病毒。根據整合到宿主基因組中的能力，病毒載體可被分類為兩組——整合型和非整合型。致癌逆轉錄病毒和慢病毒可以整合到宿主細胞的染色質中，而腺病毒、腺伴隨病毒和皰疹病毒則主要以染色體外附加體的形式存在於細胞核中。

【0237】 在某些實施方案中，合適的病毒載體是逆轉錄病毒載體。逆轉錄病毒是指逆轉錄病毒科的病毒。逆轉錄病毒的實例包括諸如鼠白血病病毒(MLV)等致癌逆轉錄病毒，以及諸如人類免疫缺陷病毒1(HIV-1)等慢病毒。逆轉錄病毒基因組是單鏈(ss)RNA，並且包含可以以順式或反式提供的各種基因。例如，逆轉錄病毒基因組可以包含順式作用序列，如兩個長末端重複序列(LTR)，以及用於基因表達、逆轉錄和整合到宿主染色體中的元件。其他組分包括用於將特定RNA包裝到新形成的病毒體中的包裝信號(psi或ψ)和聚嘌呤束(PPT)，後者是逆轉錄過程中正鏈DNA合成的起始位點。另外，逆轉錄病毒基因組可包含gag、pol和env基因。gag基因編碼結構蛋白質，pol基因編碼與ssRNA伴隨並將病毒RNA逆轉錄為DNA的酶，而env基因編碼病毒包膜。通常，gag、pol和env以反式提供，以用於病毒複製和包裝。

【0238】 在某些實施方案中，本文提供的逆轉錄病毒載體可以是慢病毒載體。識別出慢病毒的至少五種血清組或血清型。不同血清型的病毒可以差異地感染某些細胞類型和/或宿主。例如，慢病毒包括靈長類動物

逆轉錄病毒和非靈長類動物逆轉錄病毒。靈長類動物逆轉錄病毒包括HIV和猿猴免疫缺陷病毒(SIV)。非靈長類動物逆轉錄病毒包括貓免疫缺陷病毒(FIV)、牛免疫缺陷病毒(BIV)、山羊關節炎-腦炎病毒(CAEV)、馬感染性貧血病毒(EIAV)和維斯納病毒(visnavirus)。慢病毒或慢病毒載體可以能夠轉導休眠細胞。與致癌逆轉錄病毒載體一樣，慢病毒載體的設計可以基於順式作用序列與反式作用序列的分離。

【0239】 在某些實施方案中，本申請提供了已被設計用於通過優化的治療性逆轉錄病毒載體進行遞送的表達載體。該逆轉錄病毒載體可以是慢病毒，其包含左(5')LTR；有助於病毒的包裝和/或核輸入的序列；啟動子；可選的一種或多種額外的調節元件(例如，增強子或聚A序列)；可選的慢病毒反向回應元件(RRE)；包含與編碼eTF的序列可操作地連接的PV選擇性調節元件的構建體；可選的絕緣子；和右(3')逆轉錄病毒LTR。

【0240】 在示例性實施方案中，本文提供的病毒載體是腺伴隨病毒(AAV)。AAV是一種小型、複製缺陷型、無包膜的動物病毒，其可感染人類和其他一些靈長類動物物種。尚不知道AAV是否會引起人類疾病，但會誘導輕度的免疫應答。AAV載體還可以感染分裂中的細胞和休眠細胞，而無需整合到宿主細胞基因組中。

【0241】 AAV基因組由長度約為4.7 kb的線性單鏈DNA組成。該基因組由側翼是長度約為145 bp的反向末端重複(ITR)序列的兩個開放閱讀框(ORF)組成。該ITR由5'端的核苷酸序列(5' ITR)和位於3'端的核苷酸序列(3' ITR)組成，它們包含回文序列。該ITR通過折疊經由互補域基配對形成T形髮夾結構而順式起作用，該結構在第二鏈合成的DNA複製起始過程中起引物的作用。這兩個開放閱讀框編碼參與病毒體複製和包裝的rep

和cap基因。在示例性實施方案中，本文提供的AAV載體不含rep或cap基因。這樣的基因可以反式提供，以供產生如下文進一步描述的病毒體。

【0242】在某些實施方案中，AAV載體可以包括填充核酸。在一些實施方案中，該填充核酸可以編碼綠色螢光蛋白或抗生素抗性基因，如卡那黴素或氨苄青黴素抗性基因。在某些實施方案中，該填充核酸可以位於ITR序列之外(例如，與位於5’ ITR序列與3’ ITR序列之間的eTF轉基因序列和調節序列相比)。

【0243】存在各種AAV血清型，包括AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12和AAV13。這些血清型的向性或所感染的細胞類型不同。AAV可包含來自多種血清型(例如，假型)的基因組和衣殼。例如，AAV可以包含包裝在來自血清型5或血清型9的衣殼中的血清型2的基因組(例如，ITR)。假型可以提高轉導效率並改變向性。

【0244】在一些情況下，可以穿過血腦屏障或感染CNS的細胞的AAV血清型是優選的。在一些情況下，AAV9或其變體用來遞送本公開的表達盒，該表達盒包含與編碼選擇性上調SCN1A的eTF的轉基因可操作地連接的PV選擇性調節元件。在一些情況下，AAV9或其變體用來遞送包含PV選擇性微小RNA結合位點的本公開的表達盒。在一些情況下，AAV9或其變體用來遞送本公開的表達盒，該表達盒包含與編碼選擇性上調SCN1A的eTF的轉基因可操作地連接的PV選擇性調節元件，和PV選擇性微小RNA結合位點。

【0245】在示例性實施方案中，本申請提供了已被設計用於通過AAV進行遞送的表達載體。該AAV可以是任何血清型，例如AAV1、

AAV2、AAV3、AAV3b、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、AAV-DJ，或嵌合、雜合或變異AAV。該AAV也可以是自互補AAV(scAAV)。在某些實施方案中，設計用於通過AAV進行遞送的表達載體包含5’ ITR和3’ ITR。在某些實施方案中，設計用於通過AAV進行遞送的表達載體包含5’ ITR、啟動子、編碼eTF的轉基因和3’ ITR。在某些實施方案中，設計用於通過AAV進行遞送的表達載體包含5’ ITR、增強子、啟動子、編碼eTF的轉基因、聚A序列和3’ ITR。

宿主細胞

【0246】 在另一方面，本發明涉及包含本文公開的表達盒或表達載體的宿主細胞。宿主細胞可以是細菌細胞、酵母細胞、昆蟲細胞或哺乳動物細胞。在示例性實施方案中，宿主細胞是指易被目的病毒感染並適於體外培養的任何細胞系。

【0247】 在某些實施方案中，本文提供的宿主細胞可用於離體基因療法的目的。在這樣的實施方案中，用包含PV選擇性微小RNA結合位元點和/或編碼本文公開的選擇性上調SCN1A的eTF的序列的核酸分子或表達載體轉染細胞，隨後將細胞移植到患者或受試者中。移植的細胞可以具有自體的、同種異體的或異源的來源。對於臨床用途，通常在良好生產規範(GMP)條件下進行細胞分離。移植前，一般要檢查細胞質量以及是否存在微生物或其他污染物，並可以進行預處理，例如採用放射和/或免疫抑制治療。此外，可以將宿主細胞與生長因子一起移植，以刺激細胞增殖和/或分化。

【0248】 在某些實施方案中，宿主細胞可用於離體基因療法。優選

地，所述細胞是真核細胞，如哺乳動物細胞，該哺乳動物包括但不限於人，非人靈長類動物如猿、黑猩猩、猴和猩猩，馴養動物，包括狗和貓，以及牲畜，如馬、牛、豬、綿羊和山羊，或其他哺乳動物物種，包括但不限於小鼠、大鼠、豚鼠、兔、倉鼠等。本領域技術人員將根據待移植的患者或受試者選擇更合適的細胞。

【0249】 在某些實施方案中，本文提供的宿主細胞可以是具有自我更新和多能性性質的細胞，如幹細胞或誘導多能性幹細胞。幹細胞優選為間充質幹細胞。間充質幹細胞(MSC)能夠分化為成骨細胞、軟骨細胞、脂肪細胞或肌細胞中的至少一種，並且可以從任何類型的組織中分離。通常，MSC將從骨髓、脂肪組織、臍帶或外周血中分離。其獲得方法是本領域技術人員公知的。誘導多能幹細胞(也稱為iPS細胞或iPSC)是一種類型的多能幹細胞，其可以直接從成體細胞產生。Yamanaka等人通過將Oct3/4、Sox2、Klf4和c-Myc基因轉移到小鼠和人成纖維細胞中並迫使所述細胞表達該基因，而誘導了iPS細胞(WO 2007/069666)。Thomson等人隨後使用Nanog和Lin28代替Klf4和c-Myc產生了人iPS細胞(WO 2008/118820)。

【0250】 在示例性實施方案中，本文提供的宿主細胞是包裝細胞。所述細胞可以是貼壁細胞或懸浮細胞。包裝細胞和輔助載體或病毒或DNA構建體以反式一起提供了病毒載體的完全複製和包裝所需的所有缺失功能。

【0251】 優選地，所述包裝細胞是真核細胞，如哺乳動物細胞，包括猿猴、人、狗和齧齒動物細胞。人類細胞的實例是PER.C6細胞(WO01/38362)、MRC-5(ATCC CCL-171)、WI-38(ATCC CCL-75)、

HEK-293細胞(ATCC CRL-1573)、HeLa細胞(ATCC CCL2)和胚胎恒河猴肺細胞(ATCC CL-160)。非人靈長類動物細胞的實例是Vero細胞(ATCC CCL81)、COS-1細胞(ATCC CRL-1650)或COS-7細胞(ATCC CRL-1651)。狗細胞的實例是MDCK細胞(ATCC CCL-34)。齧齒動物細胞的實例是倉鼠細胞，如BHK21-F、HKCC細胞或CHO細胞。

【0252】作為哺乳動物來源的替代，用於本發明的細胞系可以來源於禽類來源，如雞、鴨、鵝、鵪鶉或雉雞。禽細胞系的實例包括禽胚胎幹細胞(WO01/85938和WO03/076601)、無限增殖化的鴨視網膜細胞(WO2005/042728)和禽胚胎幹細胞衍生的細胞，包括雞細胞(WO2006/108846)或鴨細胞，如EB66細胞系(WO2008/129058和WO2008/142124)。

【0253】在另一個實施方案中，所述宿主細胞是昆蟲細胞，如SF9細胞(ATCC CRL-1711)、Sf21細胞(IPLB-Sf21)、MG1細胞(BTI-TN-MG1)或High Five™細胞(BTI-TN-5B1-4)。

【0254】在某些實施方案中，本文提供的包含本發明重組AAV載體/基因組(例如，包含PV選擇性微小RNA結合位元點和/或編碼選擇性上調SCN1A的eTF的序列)的宿主細胞可進一步包含一種或多種額外的核酸構建體，例如，(i)編碼rep和cap基因但不攜帶ITR序列的核酸構建體(例如，AAV輔助質粒)，和/或(ii)提供AAV複製所必需的腺病毒功能的核酸構建體(例如質粒)。在示例性實施方案中，本文提供的宿主細胞包含：i)包含PV選擇性微小RNA結合位元點和/或編碼本文提供的選擇性上調SCN1A的eTF的序列(即，重組AAV基因組)的表達載體；ii)編碼不攜帶ITR序列的AAV rep和cap基因的核酸構建體；以及iii)包含腺病毒輔助基因的核酸構

建體(如下文進一步描述)。

【0255】 在某些實施方案中，rep基因、cap基因和腺病毒輔助基因可以組合在一個質粒上(Blouin V等人. J Gene Med. 2004; 6(suppl): S223-S228；Grimm D.等人. Hum. Gene Ther. 2003; 7: 839-850)。因此，在另一個示例性實施方案中，本文提供的宿主細胞包含：i)包含PV選擇性微小RNA結合位元點和/或編碼本文公開的選擇性上調SCN1A的eTF的序列(即，重組AAV基因組)的表達載體；以及ii)編碼不攜帶ITR序列的AAV rep和cap基因且進一步包含腺病毒輔助基因的質粒。

【0256】 在另一個實施方案中，本文提供的宿主細胞包含：a)包含PV選擇性微小RNA結合位元點和/或編碼本文公開的選擇性上調SCN1A的eTF的序列(即，重組AAV基因組)的表達載體；b)編碼不攜帶ITR序列的AAV rep和cap基因的質粒；以及c)包含腺病毒輔助基因E2a、E4和VA RNA的質粒；其中在組成型地表達並反式補充腺病毒E1基因的細胞，優選哺乳動物細胞，如HEK-293細胞(ATCC CRL-1573)中進行共轉染。

【0257】 在某些實施方案中，適合於大規模生產AAV載體的宿主細胞是可以用重組桿狀病毒的組合來感染的昆蟲細胞(Urabe等人, Hum. Gene Ther. 2002;13: 1935-1943)。例如，SF9細胞可以用分別表達AAV rep、AAV cap和待包裝的AAV載體的三種桿狀病毒載體來共感染。重組桿狀病毒載體將提供病毒複製和/或包裝所需的病毒輔助基因功能。

【0258】 關於根據本發明用於基因療法的病毒體的構建和生產的進一步指導可見於：Viral Vectors for Gene Therapy, Methods and Protocols. Series: Methods in Molecular Biology, Vol. 737. Merten和Al-

Rubeai (編), 2011 Humana Press (Springer)；Gene Therapy. M. Giacca. 2010 Springer-Verlag；Heilbronn R.和Weger S. Viral Vectors for Gene Transfer: Current Status of Gene Therapeutics. 於：Drug Delivery, Handbook of Experimental Pharmacology 197；M. Schafer-Korting (編). 2010 Springer-Verlag；pp. 143-170；Adeno-Associated Virus: Methods and Protocols. R. O. Snyder和P. Moulllier (編). 2011 Humana Press (Springer)；Bunning H.等人, Recent developments in adeno-associated virus technology. J. Gene Med. 2008; 10:717-733；以及Adenovirus: Methods and Protocols. M. Chillon和A. Bosch (編); 第三版. 2014 Humana Press (Springer)。

病毒體和生產病毒體的方法

【0259】在某些實施方案中，本申請提供了包含病毒載體的病毒顆粒，該病毒載體包含PV選擇性微小RNA結合位元點和/或編碼本文公開的選擇性上調SCN1A的eTF的序列。術語“病毒顆粒”和“病毒體”在本文中可互換使用，並且涉及感染性的且通常是複製缺陷性的病毒顆粒，其包含包裝在衣殼中的病毒基因組(例如，病毒表達載體)，並且視情況而定，例如，對於逆轉錄病毒，其包含包圍衣殼的脂質包膜。“衣殼”是指其中包裝有病毒基因組的結構。衣殼由若干個寡聚體結構亞單位組成，該亞單位由蛋白質構成。例如，AAV具有由以下三種衣殼蛋白的相互作用形成的二十面體衣殼：VP1、VP2和VP3。在一個實施方案中，本文提供的病毒體是通過包裝包含PV選擇性調節元件和PV選擇性微小RNA結合位點的AAV載體而獲得的重組AAV病毒體或rAAV病毒體。在另一個實施方案中，本文提供的病毒體是通過將包含與編碼如本文所述選擇性上調SCN1A的eTF

的序列可操作地連接的PV選擇性調節元件的AAV載體包裝在蛋白質殼中而獲得的重組AAV病毒體或rAAV病毒體。在另一個實施方案中，本文提供的病毒體是通過將包含與編碼如本文所述選擇性上調SCN1A的eTF的序列可操作地連接的PV選擇性調節元件和PV選擇性微小RNA結合位點的AAV載體包裝在蛋白質殼中而獲得的重組AAV病毒體或rAAV病毒體。

【0260】 在某些實施方案中，可以通過將來源於特定AAV血清型的AAV基因組用衣殼包裝在由天然Cap蛋白(對應於相同特定血清型的AAV)形成的病毒顆粒中來製備本文提供的重組AAV病毒體。在其他實施方案中，本文提供的AAV病毒顆粒包含以下病毒載體，其包含包裝到來自不同血清型的蛋白質中的給定AAV血清型的ITR。參見，例如，Bunning H等人, J Gene Med 2008; 10: 717-733。例如，具有來自給定AAV血清型的ITR的病毒載體可以包裝到：a)由來源於相同或不同AAV血清型的衣殼蛋白(例如，AAV2 ITR和AAV9衣殼蛋白；AAV2 ITR和AAV8衣殼蛋白；等等)構成的病毒顆粒；b)由來自不同AAV血清型或突變體的衣殼蛋白混合物(例如，AAV2 ITR與AAV1和AAV9衣殼蛋白)構成的鑲嵌型病毒顆粒；c)由已通過在不同AAV血清型或變體之間進行域交換而被截短的衣殼蛋白(例如，AAV2 ITR與具有AAV9域的AAV8衣殼蛋白)構成的嵌合病毒顆粒；或d)經工程改造而展示出選擇性結合域，從而能夠與靶細胞特異性受體發生嚴格相互作用的靶向病毒顆粒(例如，AAV5 ITR與通過插入肽配體而被遺傳截短的AAV9衣殼蛋白；或通過將肽配體偶聯至衣殼表面而非遺傳修飾的AAV9衣殼蛋白)。

【0261】 技術人員將會理解，本文提供的AAV病毒體可包含任何AAV血清型的衣殼蛋白。在一個實施方案中，病毒顆粒包含來自選自

AAV1、AAV2、AAV5、AAV8和AAV9的AAV血清型的衣殼蛋白，它們更適於遞送至CNS(M. Hocquemiller等人, Hum Gene Ther 27(7): 478-496 (2016))。在特定實施方案中，病毒顆粒包含本發明的表達盒，其中該表達盒的5’ ITR和3’ ITR序列是AAV2血清型的而衣殼蛋白是AAV9血清型的。

【0262】 用於生產rAAV病毒體的許多方法是本領域已知的，包括轉染、穩定細胞系生產以及包括腺病毒-AAV雜合體、皰疹病毒-AAV雜合體(Conway, J E等人 (1997) J. Virology 71(11):8780-8789)和桿狀病毒-AAV雜合體在內的感染性雜合病毒生產系統。用於生產rAAV病毒顆粒的rAAV生產培養物都需要：1)合適的宿主細胞，例如在桿狀病毒生產系統的情況下，包括人源細胞系，如HeLa、A549或293細胞，或昆蟲來源的細胞系，如SF-9；2)合適的輔助病毒功能，其由野生型或突變型腺病毒(如溫度敏感性腺病毒)、皰疹病毒、桿狀病毒或提供協助工具的質粒構建體提供；3)AAV rep和cap基因及基因產物；4)側翼為AAV ITR序列的轉基因(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的eTF的序列，和/或PV選擇性啟動子)；以及5)支持rAAV生產的合適的培養基和培養基組分。

【0263】 在各個實施方案中，本文所述的宿主細胞包含以下三種組分：(1)rep基因和cap基因，(2)提供協助工具的基因，以及(3)側翼為ITR的轉基因(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的eTF的序列，和/或PV選擇性啟動子)。AAV rep基因、AAV cap基因和提供協助工具的基因可通過將所述基因併入載體如質粒中並將所述載體引入宿主細胞中而引入細胞中。

rep、cap和協助工具基因可以併入到相同質粒或不同質粒中。在優選實施方案中，將AAV rep和cap基因併入一個質粒中，並將提供協助工具的基因併入另一質粒中。可以通過使用本領域公知的任何合適的方法將用於創建產生病毒體的宿主細胞的各種質粒(例如，包含AAV rep和cap基因、協助工具或轉基因)引入細胞中。轉染方法的實例包括但不限於磷酸鈣共沉澱、DEAE-葡聚糖、聚凝胺、電穿孔、顯微注射、脂質體介導的融合、脂質轉染、逆轉錄病毒感染和生物射彈轉染。在某些實施方案中，可將提供rep和cap基因、協助工具和側翼為ITR的轉基因的質粒同時引入細胞中。在另一個實施方案中，提供rep和cap基因以及協助工具的質粒可以在引入包含轉基因的質粒之前或之後被引入細胞中。在示例性實施方案中，用以下三種質粒同時轉染細胞(例如，三重轉染法)：(1)包含側翼為ITR的轉基因(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的eTF的序列，和/或PV選擇性啟動子)的質粒，(2)包含AAV rep和cap基因的質粒，以及(3)包含提供協助工具的基因的質粒。示例性宿主細胞可以是293、A549或HeLa細胞。

【0264】 在其他實施方案中，(1)AAV rep和cap基因、(2)提供協助工具的基因以及(3)側翼為ITR的轉基因(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的eTF的序列，和/或PV選擇性啟動子)中的一個或多個可以由包裝細胞附加型地攜帶和/或整合到包裝細胞的基因組中。在一個實施方案中，宿主細胞可以是包裝細胞，其中AAV rep和cap基因以及協助工具在宿主細胞中穩定地保留，並且用含有側翼為ITR的轉基因(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的

eTF的序列，和/或PV選擇性啟動子)的質粒暫態轉染宿主細胞。在另一個實施方案中，宿主細胞是包裝細胞，其中AAV rep和cap基因在宿主細胞中穩定地保留，並且用含有側翼為ITR的轉基因(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的eTF的序列，和/或PV選擇性啟動子)的質粒和含有協助工具的質粒暫態轉染宿主細胞。在另一個實施方案中，宿主細胞可以是包裝細胞，其中協助工具在宿主細胞中穩定地保留，並且用含有側翼為ITR的轉基因(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的eTF的序列，和/或PV選擇性啟動子)的質粒和含有rep和cap基因的質粒暫態轉染宿主細胞。在另一個實施方案中，宿主細胞可以是用rep和cap基因、協助工具和側翼為ITR的轉基因序列(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的eTF的序列，和/或PV選擇性啟動子)穩定轉染的生產細胞系。示例性包裝和生產細胞可以來源於293、A549或HeLa細胞。

【0265】 在另一個實施方案中，生產細胞系是用提供Rep和Cap蛋白的桿狀病毒表達載體感染的昆蟲細胞系(通常是Sf9細胞)。該系統不需要腺病毒輔助基因(Ayuso E等人, Curr. Gene Ther. 2010, 10:423-436)。

【0266】 如本文所用的，術語“cap蛋白”是指具有天然AAV Cap蛋白(例如，VP1、VP2、VP3)的至少一種功能活性的多肽。cap蛋白的功能活性的實例包括誘導衣殼形成、促進單鏈DNA積累、促進AAV DNA包裝到衣殼中(即衣殼化)、結合細胞受體以及促進病毒體進入宿主細胞中的能力。原則上，任何Cap蛋白均可用於本發明。

【0267】 已經報導，Cap蛋白對宿主向性，細胞、組織或器官特異

性，受體使用、感染效率和AAV病毒的免疫原性具有影響。因此，可以考慮例如受試者的物種(例如人類或非人類)、受試者的免疫學狀態、受試者對於長期或短期治療的適合性或具體的治療應用(例如，具體疾病或病症的治療，或向特定細胞、組織或器官的遞送)來選擇用於rAAV的AAV cap。在某些實施方案中，cap蛋白來源於選自AAV1、AAV2、AAV5、AAV8和AAV9血清型的AAV。在示例性實施方案中，cap蛋白來源於AAV9。

【0268】 在一些實施方案中，用於本發明方法的AAV Cap可以通過一種上述AAV cap或其編碼核酸的誘變(即通過插入、缺失或置換)來產生。在一些實施方案中，AAV cap與一種或多種上述AAV cap至少70%、75%、80%、85%、90%、95%、98%或99%或更高水準相似。

【0269】 在一些實施方案中，所述AAV cap是嵌合的，包含來自兩種、三種、四種或更多種上述AAV cap的結構域。在一些實施方案中，該AAV cap是源自兩種或三種不同AAV或重組AAV的VP1、VP2和VP3單體的嵌合體。在一些實施方案中，rAAV組合物包含超過一種上述cap。

【0270】 在一些實施方案中，用於rAAV病毒體的AAV cap被工程改造為包含異源序列或其他修飾。例如，可以將賦予選擇性靶向或免疫逃避的肽或蛋白序列工程改造到cap蛋白中。備選地或另外地，可以對cap進行化學修飾，使得rAAV的表面被聚乙二醇化(即peg化)，這可以促進免疫逃避。還可以對cap蛋白進行誘變(例如，以去除其天然受體結合或掩蓋免疫原性表位)。

【0271】 如本文所用的，術語“rep蛋白”是指具有天然AAV rep蛋白(例如，rep 40、52、68、78)的至少一種功能活性的多肽。rep蛋白的功能

活性的實例包括與該蛋白質的生理功能相關的任何活性，包括通過識別來促進DNA複製、結合並切刻DNA複製的AAV起點以及DNA解旋酶活性。其他功能包括調節從AAV(或其他異源)啟動子的轉錄以及AAV DNA向宿主染色體中的位點特異性整合。在特定實施方案中，AAV rep基因可以來自血清型AAV1、AAV2、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10或AAVrh10；更優選地，來自選自AAV1、AAV2、AAV5、AAV8和AAV9的AAV血清型。

【0272】 在一些實施方案中，用於本發明方法的AAV rep蛋白可以通過一種上述AAV rep或其編碼核酸的誘變(即通過插入、缺失或置換)來產生。在一些實施方案中，該AAV rep與一種或多種上述AAV rep至少70%、75%、80%、85%、90%、95%、98%或99%或更高水準相似。

【0273】 如本文所用的，表述“協助工具”或“輔助基因”是指AAV複製所依賴的病毒蛋白質。協助工具包括AAV複製所需的那些蛋白質，包括但不限於參與AAV基因轉錄啟動、階段特異性AAV mRNA剪接、AAV DNA複製、cap表達產物的合成和AAV衣殼裝配的那些蛋白質。基於病毒的協助工具可以來源於任何已知的輔助病毒，如腺病毒、皰疹病毒(除單純皰疹病毒1型以外)和痘苗病毒。協助工具包括但不限於腺病毒E1、E2a、VA和E4或皰疹病毒UL5、ULB、UL52和UL29以及皰疹病毒聚合酶。在優選的實施方案中，AAV複製所依賴的蛋白質來源於腺病毒。

【0274】 在一些實施方案中，用於本發明方法的AAV複製所依賴的病毒蛋白質可以通過一種上述病毒蛋白質或其編碼核酸的誘變(即通過插入、缺失或置換)來產生。在一些實施方案中，該病毒蛋白質與一種或多種上述病毒蛋白質至少70%、75%、80%、85%、90%、95%、98%或

99%或更高水準相似。

【0275】 測定AAV複製所依賴的cap蛋白、rep蛋白和病毒蛋白質的功能的方法是本領域公知的。

【0276】 用於表達目的轉基因(例如，包含以下一個或多個：PV選擇性微小RNA結合位元點，編碼如本文所述選擇性上調SCN1A的eTF的序列，和/或PV選擇性啟動子)的宿主細胞可以在足以裝配AAV病毒體的條件下生長。在某些實施方案中，宿主細胞生長一段合適的時間，以促進AAV病毒體的裝配和病毒體向培養基中的釋放。通常，細胞可以生長約24小時、約36小時、約48小時、約72小時、約4天、約5天、約6天、約7天、約8天、約9天或至多約10天。在約10天後(或更早，取決於培養條件和所用的具體宿主細胞)，生產水準通常顯著下降。通常，從病毒產生的點開始測量培養時間。例如，在AAV的情況下，病毒產生通常在如本文所述的合適的宿主細胞中提供輔助病毒功能時開始。通常，在輔助病毒感染後(或在病毒產生開始後)約48至約100、優選約48至約96、優選約72至約96、優選約68至約72小時收穫細胞。

【0277】 rAAV生產培養物可以在適合於所使用的具體宿主細胞的多種條件(在寬溫度範圍內、持續不同時間長度等)下生長。rAAV生產培養物包括附著依賴性培養物，其可以在合適的附著依賴性容器，例如滾瓶、中空纖維篩檢程式、微載體以及填充床或流化床生物反應器中培養。rAAV載體生產培養物還可以包含適應懸浮的宿主細胞，如HeLa、293和SF-9細胞，其可以通過多種方式培養，包括，例如旋轉燒瓶、攪拌罐生物反應器和一次性系統，如Wave袋系統。

【0278】 本領域已知的合適的培養基可以用於產生rAAV病毒體。

這些培養基包括但不限於Hyclone Laboratories和JRH生產的培養基，包括改良Eagle培養基(MEM)、Dulbecco改良Eagle培養基(DMEM)，其各自通過引用整體併入本文。在某些實施方案中，rAAV生產培養基可以補充有0.5%-20%(v/v或w/v)水準的血清或血清衍生重組蛋白。或者，可以在無血清條件下產生rAAV載體，該無血清條件也可以被稱為沒有動物來源的產物的培養基。

【0279】 在培養宿主細胞以允許產生AAV病毒體後，可以收穫並純化所得的病毒體。在某些實施方案中，(1)可以通過裂解宿主細胞從生產培養物的宿主細胞獲得AAV病毒體，和/或(2)可以在轉染一段時間(優選72小時)後從所述細胞的培養基中獲得AAV病毒體。只要在導致rAAV病毒體從完整細胞釋放到培養基中的條件下培養細胞，就可以從生產培養物中用過的培養基收穫rAAV病毒體(參見，例如，美國專利6,566,118)。裂解細胞的合適方法在本領域中也是已知的，並且包括例如多個冷凍/融化迴圈、聲處理、微流化和用化學物質如去汙劑和/或蛋白酶處理。

【0280】 收穫後，可以純化rAAV病毒體。如本文所用的術語“純化的”包括不含至少一些其他組分的rAAV病毒體的製品，這些其他組分可以存在於天然出現rAAV病毒體之處或最初製備rAAV病毒體之處。因此，例如，可以使用分離技術從來源混合物如培養裂解物或生產培養上清液中富集rAAV病毒體，來製備純化的rAAV病毒體。富集可以通過多種方式來測量，例如通過溶液中存在的DNA酶抗性顆粒(DRP)或基因組拷貝(gc)的比例，或通過感染性，或者可以相對於來源混合物中存在的第二種潛在干擾性物質(例如污染物，包括生產培養污染物或過程中污染物，包括輔助病毒、培養基組分等)來測量。

【0281】在某些實施方案中，可以將rAAV生產培養收穫物進行澄清，以去除宿主細胞碎片。在一些實施方案中，可以使用多種標準技術，如離心或通過孔徑為0.2 μm或更大的篩檢程式(例如，乙酸纖維素篩檢程式或一系列深層篩檢程式)過濾，將生產培養收穫物進行澄清。

【0282】在某些實施方案中，將rAAV生產培養收穫物進一步用Benzonase™處理，以消化生產培養物中存在的任何高分子量DNA。在一些實施方案中，Benzonase™消化在標準條件下進行，例如，Benzonase™的終濃度為1-2.5單位/ml，溫度範圍從環境溫度至37°C，時間為30分鐘至數小時。

【0283】在某些實施方案中，可以使用以下一種或多種純化步驟來分離或純化rAAV病毒體：平衡離心；流通式陰離子交換過濾；用於濃縮rAAV顆粒的切向流過濾(TFF)；通過磷灰石色譜法捕獲rAAV；輔助病毒熱滅活；通過疏水相互作用色譜法捕獲rAAV；通過大小排阻色譜法(SEC)進行緩衝液交換；納米過濾；以及通過陰離子交換色譜法、陽離子交換色譜法或親和色譜法捕獲rAAV。這些步驟可以單獨使用，以各種組合使用，或以不同的順序使用。純化rAAV顆粒的方法可見於，例如，Xiao等人, (1998) Journal of Virology 72:2224-2232；美國專利6,989,264和8,137,948；以及WO 2010/148143。

【0284】在某些實施方案中，純化的AAV病毒體可以用PBS透析、過濾並儲存在-80ºC。病毒基因組的滴度可以使用線性化質粒DNA作為標準曲線，通過定量PCR來確定(參見，例如，Lock M等人, Hum. Gene Ther. 2010; 21:1273-1285)。

藥物組合物

【0285】在某些實施方案中，本申請提供了包含PV選擇性微小RNA結合位元點和/或編碼選擇性上調SCN1A的eTF的序列以及藥學上可接受的載體的組合物。在其他實施方案中，本申請提供了包含PV選擇性微小RNA結合位元點和/或編碼選擇性上調SCN1A的eTF的序列以及藥學上可接受的載體的病毒體。在示例性實施方案中，這樣的組合物適用於基因療法應用。藥物組合物優選地在製備和儲存條件下是無菌的和穩定的。無菌溶液可以通過例如通過無菌濾膜過濾來實現。

【0286】藥物組合物中可接受的載體和賦形劑優選地在所採用的劑量和濃度下對接受者無毒。可接受的載體和賦形劑可包括緩衝液，如磷酸鹽、檸檬酸鹽、HEPES和TAE，抗氧化劑，如抗壞血酸和甲硫氨酸，防腐劑，如氯化六甲雙銨、十八烷基二甲基苄基氯化銨、間苯二酚和苯紮氯銨，蛋白質，如人血清白蛋白、明膠、葡聚糖和免疫球蛋白，親水聚合物，如聚乙烯吡咯烷酮，氨基酸，如甘氨酸、穀氨醯胺、組氨酸和賴氨酸，以及碳水化合物，如葡萄糖、甘露糖、蔗糖和山梨醇。本公開的藥物組合物可以以可注射製劑的形式腸胃外施用。可以使用無菌溶液或任何藥學上可接受的液體作為媒介物來配製用於注射的藥物組合物。藥學上可接受的媒介物包括但不限於無菌水和生理鹽水。

【0287】本公開的藥物組合物可以在微膠囊如羥甲基纖維素或明膠-微膠囊和聚甲基丙烯酸甲酯微膠囊中製備。本公開的藥物組合物還可以在其他藥物遞送系統如脂質體、白蛋白微球、微乳劑、納米顆粒和納米膠囊中製備。用於基因療法的藥物組合物可以在可接受的稀釋劑中，或者可以包含其中嵌有基因遞送媒介物的緩釋基質。

【0288】本文提供的藥物組合物可以被配製用於腸胃外給藥、皮下

給藥、靜脈內給藥、肌肉內給藥、動脈內給藥、實質內給藥、鞘內給藥、小腦延髓池內給藥、腦室內給藥或腹膜內給藥。藥物組合物還可以被配製用於或通過鼻、噴霧、口、氣霧劑、直腸或陰道給藥。在一個實施方案中，本文提供的藥物組合物被施用於CNS或腦脊液(CSF)，即通過實質內注射、鞘內注射、小腦延髓池內注射或腦室內注射施用。組織靶標可以是特異性的，例如CNS特異性的，或者它可以是若干種組織例如肌肉和CNS組織的組合。示例性的組織或其他靶標可以包括肝、骨骼肌、心肌、脂肪沉積物、腎、肺、血管內皮、上皮、造血細胞、CNS和/或CSF。在優選的實施方案中，本文提供的包含PV選擇性微小RNA結合位點和/或選擇性上調SCN1A的eTF的藥物組合物被施用於CNS或CSF，即通過實質內注射、鞘內注射、小腦延髓池內注射或腦室內注射施用。這些方法中的一種或多種可以用來施用本公開的藥物組合物。

【0289】 在某些實施方案中，本文提供的藥物組合物包含“有效量”或“治療有效量”。如本文所用的，這樣的量是指在達到所需治療結果所必需的劑量和時間段下有效的量，該治療結果例如是增加SCN1A表達的水準和/或降低癲癇發作的頻率和/或持續時間。

【0290】 本公開的藥物組合物的劑量取決於包括給藥途徑、待治療的疾病和受試者的身體特性(例如年齡、體重、一般健康)在內的因素。可以調整劑量以提供最佳治療反應。一般而言，劑量可以是有效治療疾病而不誘發顯著毒性的量。在一個實施方案中，本文提供的AAV載體可以以$5x10^{11}$至$1x10^{14}$ gc/kg(基因組拷貝/千克患者體重(gc/kg))範圍內的量或劑量施用於患者，以治療SCN1A缺乏(包括，例如，Dravet綜合症)。在更具體的實施方案中，該AAV載體以約$5x10^{11}$ gc/kg至約$3x10^{13}$ gc/kg，或約

$1x10^{12}$至約$1x10^{14}$ gc/kg，或約$1x10^{12}$至約$1x10^{13}$ gc/kg範圍內包含的量，或約$5x10^{11}$ gc/kg、$1x10^{12}$ gc/kg、$1.5x10^{12}$ gc/kg、$2.0x10^{12}$ gc/kg、$2.5x10^{12}$ gc/kg、$3x10^{12}$ gc/kg、$3.5x10^{12}$ gc/kg、$4x10^{12}$ gc/kg、$4.5x10^{12}$ gc/kg、$5x10^{12}$ gc/kg、$5.5x10^{12}$ gc/kg、$6x10^{12}$ gc/kg、$6.5x10^{12}$ gc/kg、$7x10^{12}$ gc/kg、$7.5x10^{12}$ gc/kg、$8x10^{12}$ gc/kg、$8.5x10^{12}$ gc/kg、$9x10^{12}$ gc/kg或$9.5x10^{12}$ gc/kg施用。例如，可以通過qPCR或數位小液滴PCR(ddPCR)來確定gc/kg(參見，例如，M. Lock等人, Hum Gene Ther Methods. 2014年4月;25(2): 115-25)。在另一個實施方案中，本文提供的AAV載體可以以$1x10^{9}$至$1x10^{11}$ iu/kg(載體的感染單位(iu)/受試者或患者體重(kg))範圍內的量或劑量施用於患者，以治療SCN1A缺乏(包括，例如，Dravet綜合症)。在某些實施方案中，所述藥物組合物可以根據需要以單位元劑量形成。這樣的單劑量單位可包含約$1x10^{9}$ gc至約$1x10^{15}$ gc。

【0291】本公開的藥物組合物可以例如每天、每週、每月、每半年、每年一次或多次(例如，1-10次或更多次)或根據醫學需要施用於有需要的受試者。在示例性實施方案中，單次給藥是足夠的。在一個實施方案中，包含編碼PV選擇性微小RNA結合位點和/或選擇性上調SCN1A的eTF的表達盒的藥物組合物適用於人類受試者，並且通過實質內注射、鞘內注射、小腦延髓池內注射或腦室內注射施用。在一個實施方案中，該藥物組合物通過團注經由外周靜脈遞送。在其他實施方案中，該藥物組合物通過經約10分鐘(±5分鐘)、經約20分鐘(±5分鐘)、經約30分鐘(±5分鐘)、經約60分鐘(±5分鐘)或經約90分鐘(±10分鐘)的輸注而經由外周靜脈遞送。

【0292】在另一方面，本申請進一步提供了一種試劑盒，其包含在一個或多個容器中的本文所述的核酸分子、載體、宿主細胞、病毒體或藥

物組合物。試劑盒可包括說明書或包裝材料，其說明如何將試劑盒內包含的核酸分子、載體、宿主細胞或病毒體施用於患者。試劑盒的容器可以是任何合適的材料，例如玻璃、塑膠、金屬等，並且可以是任何合適的大小、形狀或構造。在某些實施方案中，試劑盒可包括一個或多個安瓿或注射器，其中包含合適的液體或溶液形式的核酸分子、載體、宿主細胞、病毒體或藥物組合物。

治療方法

【0293】在一個方面，本申請提供了使用本文公開的選擇性上調SCN1A的eTF的方法。在某些實施方案中，本申請提供了施用表達盒、表達載體或病毒顆粒以上調細胞中SCN1A的表達的方法，該表達盒、表達載體或病毒顆粒包含編碼本文公開的選擇性上調SCN1A的eTF的多核苷酸。在各個實施方案中，本文公開的選擇性上調SCN1A的eTF可用來在體外、體內或離體調節細胞中SCN1A的表達。

【0294】在某些實施方案中，本申請提供了通過向有需要的受試者施用表達盒、表達載體或病毒顆粒來治療與SCN1A相關的疾病或病症的方法，該表達盒、表達載體或病毒顆粒包含編碼本文公開的選擇性上調SCN1A的eTF的多核苷酸。在某些實施方案中，該病症是中樞神經系統病症。在示例性實施方案中，該疾病或病症與SCN1A的單倍性不足相關。在某些實施方案中，該病症是與SCN1A單倍性不足相關的癲癇。在某些實施方案中，該單倍性不足是受試者對於SCN1A基因的功能喪失突變而言為雜合的結果。在某些實施方案中，該病症是與SCN1A基因中的插入、缺失或置換相關的癲癇。在某些實施方案中，該病症是與SCN1A基因中的點突變相關的癲癇。在某些實施方案中，治療疾病或病症的方法包

括施用表達盒、表達載體或病毒顆粒，該表達盒、表達載體或病毒顆粒包含編碼本文公開的選擇性上調SCN1A的eTF的多核苷酸，從而糾正SCN1A的表達不足，使其處於健康個體的水準內，或處於醫護標準所定義的正常範圍內。在某些實施方案中，本文公開的方法用來治療與內源SCN1A相關的疾病或病症，該內源SCN1A包含導致SCN1A的異常表達的一個或多個突變。

【0295】 在某些實施方案中，本申請提供了通過向有需要的受試者施用表達盒、表達載體或病毒顆粒來改善與疾病或病症相關的症狀的方法，該表達盒、表達載體或病毒顆粒包含編碼本文公開的選擇性上調SCN1A的eTF的多核苷酸。

【0296】 在示例性實施方案中，本申請提供了通過向有需要的受試者施用表達盒、表達載體或病毒顆粒來治療與SCN1A突變(例如，點突變、置換、缺失、倒位等)、Nav1.1缺乏和/或Nav1.1活性降低相關的疾病、病症或症狀的方法，該表達盒、表達載體或病毒顆粒包含編碼選擇性上調SCN1A基因或其蛋白質產物Nav1.1的表達的eTF的多核苷酸。電壓門控鈉離子通道對於橫紋肌和神經元組織中動作電位的產生和傳播至關重要。電壓門控鈉離子通道是由大中央孔形成糖基化α亞單位和2個較小的輔助β亞單位組成的異聚複合物。由SCN1A基因編碼的大α亞單位Nav1.1亞單位與多種疾病或病症如Dravet綜合症有關。Nav1.1在神經元中表達，並且可以由包括由SCN1B基因表達的Navβ1在內的各種β亞單位裝配。

【0297】 在某些實施方案中，本申請提供了使用選擇性上調內源SCN1A基因表達的eTF治療與SCN1A突變(例如，缺失、插入、倒位、點突變(例如，無義突變、錯義突變)等)或Nav1.1活性降低相關的疾病的方

法。與SCN1A突變相關的疾病和病症包括但不限於：Dravet綜合症、大田原綜合症、癲癇、早期嬰兒型癲癇性腦病6(EIEE6)、家族性熱性癲癇發作3A(FEB3A)、難治性兒童癲癇伴全身性強直-陣攣性癲癇發作(ICEGTC)、家族性偏癱性偏頭痛3(FHM3)、Panayiotopoulos綜合症、家族性心房顫動13(ATFB13)、全身性癲癇伴熱性癲癇發作+1型(gefs+ 1型)、Brugada綜合症、非特異性心臟傳導缺陷、全身性癲癇伴熱性癲癇發作+、良性家族性嬰兒癲癇發作、早期嬰兒型癲癇性腦病11(EIEE11)、良性家族性嬰兒癲癇、神經變性、tau蛋白病變和阿爾茨海默病。在一些情況下，該神經系統病況是Dravet綜合症。SCN1A的突變或異常還與癲癇病症、癲癇、自閉症、家族性偏癱性偏頭痛3型(FHM3)、遺傳性癲癇伴熱性癲癇發作+(GEFS+)和某些抗癲癇發作藥物的有效性相關。例如，SCN1A中的ICS5N+5G>A突變與抗癲癇藥苯妥英和卡馬西平的最大安全量(劑量)相關。

【0298】在某些實施方案中，本申請提供了通過施用包含編碼選擇性上調SCN1A的eTF的多核苷酸的表達盒、表達載體或病毒顆粒來治療患有Dravet綜合症或具有患Dravet綜合症風險的受試者的方法。Dravet綜合症的特徵在於在兒童期第一年內延長的發熱和非發熱癲癇發作。這種疾病會進展為其他癲癇發作類型，例如肌陣攣性癲癇發作和部分癲癇發作、精神運動遲緩和共濟失調。其特徵在於認知損害、行為失常和運動缺陷。行為缺陷通常包括多動和衝動，在極少數情況下，包括孤獨樣行為。Dravet綜合症還與包括嗜睡和失眠在內的睡眠障礙相關。在許多患者中，Dravet綜合症是由導致產生無功能蛋白質的基因突變引起的。在治療與遺傳原因相關的病症中存在許多挑戰。因此，大多數現有的治療已被用於癲癇發作

和其他症狀的預防性醫療處理。

【0299】 在70-90%的患者中，Dravet綜合症是由SCN1A基因中的無義突變引起的，該突變導致提前終止密碼子，因而導致無功能性蛋白質。通常，鈉通道孔的S5或S6區段中的錯義突變導致通道功能的喪失和Dravet綜合症的發展。SCN1A突變(例如，無義突變、錯義突變、缺失、插入、倒位等)的雜合遺傳是形成有缺陷的鈉通道所必需的；Dravet綜合症患者仍具有該基因的一個正常拷貝。因此，該疾病被表徵為單倍性不足的疾病，因此增加SCN1A的功能性拷貝的表達可以恢復Nav1.1的正常產生水準。

【0300】 與Dravet綜合症相關的症狀包括癲癇發作、記憶缺陷、發育遲緩、肌張力差和/或認知問題。用本文所述的表達盒、表達載體或病毒顆粒治療可導致一個或多個症狀的改善，如癲癇發作的次數、持續時間和/或強度的減小。向具有發生Dravet綜合症的風險的受試者施用本文所述的基因療法可以防止Dravet綜合症的一個或多個症狀的發展或減緩其進展。

【0301】 在某些實施方案中，與未治療的對照相比或與治療前水準相比，用本文所述包含編碼選擇性上調SCN1A的eTF的多核苷酸的表達盒、表達載體或病毒顆粒進行治療將癲癇發作持續時間和/或頻率，例如與Dravet綜合症相關的癲癇發作降低至少1%、2%、3%、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%、45%、50%、55%、60%、65%、70%、75%、80%、85%、

90%、95%或更多。

【0302】在一些阿爾茨海默病患者中，澱粉樣蛋白β(Aβ)的產生涉及許多可以影響神經元的興奮性的肽和蛋白酶，導致癲癇發作和PV神經元中Nav1.1鈉通道的下調。在另一個實施方案中，本申請提供了通過施用包含編碼選擇性上調SCN1A的eTF的多核苷酸的本文所述表達盒、表達載體或病毒顆粒來治療罹患阿爾茨海默病的受試者的方法。與阿爾茨海默病相關的症狀包括短期記憶喪失、認知困難、癲癇發作，以及語言、執行功能、感知(失認)和動作執行(失用)存在困難。用包含編碼選擇性上調SCN1A的eTF的多核苷酸的表達盒、表達載體或病毒顆粒進行治療可導致一個或多個阿爾茨海默病症狀的改善，如記憶喪失的進展減慢，或防止一種或多種症狀。在一些情況下，該治療可以導致高γ功率腦活動的校正。該治療可以導致癲癇發作頻率和/或癲癇發作嚴重程度降低，或與未治療相比將高γ功率活動減少10%、20%、30%、40%、50%、60%、70%或更多。在一些情況下，該治療可以導致認知功能改善。與未治療或用本文公開的編碼選擇性上調SCN1A的eTF的多核苷酸治療前相比，學習和/或記憶可以改善至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%或超過100%。

【0303】在一些情況下，與未治療的對照相比或與治療前水準相比，用包含編碼選擇性上調SCN1A的eTF的多核苷酸的表達盒、表達載體或病毒顆粒進行治療將高γ功率活動(例如，與阿爾茨海默病相關的高γ功率活動)降低至少1%、2%、3%、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%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%或95%。

【0304】帕金森症是指在帕金森病(PD)中發現的體徵和症狀的集合，包括緩慢(運動徐緩)、僵硬(強直)、震顫和失衡(姿勢不穩)。在一些情況下，向具有發生或罹患帕金森病的風險的受試者施用本文所述包含編碼選擇性上調SCN1A的eTF的多核苷酸的表達盒、表達載體或病毒顆粒可以預防其一種或多種症狀的發展，或者與未治療相比，將帕金森病的進展減慢至少10%、至少20%、至少30%、至少40%、至少50%、至少60%、至少70%、至少80%或至少90%。

【0305】在某些實施方案中，本申請提供了可用於治療具有患病風險的受試者的方法。可能已知受試者易患疾病，例如神經系統疾病或與癲癇、癲癇發作和/或腦病相關的疾病。由於遺傳事件，或由於已知的危險因素，受試者可能易患疾病。例如，受試者可能在SCN1A中攜帶與癲癇(例如，Dravet綜合症)相關的突變。SCN1A基因中任何降低其活性(通過降低表達水準、削弱蛋白質功能或二者的組合)的突變都可以使受試者易患疾病，包括SCN1A基因中的插入、缺失、倒位、易位或置換中的任何一種或多種(例如，點突變，包括無義突變和/或錯義突變)。在一些情況下，由於受試者的年齡，受試者可能易患諸如阿爾茨海默病的疾病。在一些情況下，受試者可能具有不足量的SCN1A蛋白，並且治療與SCN1A相關的疾病包括施用本文所述包含編碼選擇性上調內源SCN1A的eTF的多核苷酸的表達盒、表達載體或病毒顆粒。

【0306】在某些實施方案中，使用本文提供的包含編碼選擇性上調內源SCN1A的eTF的多核苷酸的表達盒、表達載體或病毒顆粒進行治療可

導致與Dravet或其他SCN1A相關疾病或病症(例如，與SCN1A單倍性不足相關的癲癇)相關的症狀的減輕或停止。例如，治療可以改善學習、記憶、認知功能和/或運動功能；降低癲癇發作的頻率和/或持續時間；並且/或者降低溫度敏感性(或提高觸發癲癇發作的溫度閾值)。

【0307】 在另一方面，本申請提供了通過施用包含至少一個PV選擇性微小RNA結合位點的表達盒、表達載體或病毒顆粒而在PV神經元中選擇性表達轉基因的方法。在某些實施方案中，本申請提供了通過施用包含轉基因和至少一個PV選擇性微小RNA結合位點的表達盒、表達載體或病毒顆粒而在靈長類動物的PV神經元中選擇性表達轉基因的方法。在某些實施方案中，本申請提供了通過施用包含與轉基因可操作地連接的PV選擇性調節元件和至少一個PV選擇性微小RNA結合位點的表達盒、表達載體或病毒顆粒而在PV神經元中選擇性表達轉基因的方法。在示例性實施方案中，該轉基因包含編碼如本文所述選擇性上調SCN1A的任何eTF的序列。

【0308】 在某些實施方案中，本申請提供了一種基因治療方法，其包括向受試者施用包含轉基因和至少一個PV選擇性微小RNA結合位點的表達盒、表達載體或病毒顆粒。在某些實施方案中，本申請提供了基因治療方法，其包括向受試者施用包含與轉基因可操作地連接的PV選擇性調節元件和至少一個PV選擇性微小RNA結合位點的表達盒、表達載體或病毒顆粒。在示例性實施方案中，該轉基因包含編碼如本文所述選擇性上調SCN1A的任何eTF的序列。

【0309】 在某些實施方案中，本申請提供了一種治療疾病或病症的方法，其包括施用包含轉基因和至少一個PV選擇性微小RNA結合位點的

表達盒、表達載體或病毒顆粒。在某些實施方案中，本申請提供了治療疾病或病症的方法，其包括向受試者施用包含與轉基因可操作地連接的PV選擇性調節元件和至少一個PV選擇性微小RNA結合位點的表達盒、表達載體或病毒顆粒。在示例性實施方案中，該轉基因包含編碼如本文所述選擇性上調SCN1A的任何eTF的序列。在某些實施方案中，包含轉基因和PV選擇性微小RNA結合位點以及可選的PV選擇性調節元件的表達盒、表達載體或病毒顆粒可以用來治療涉及PV神經元的疾病或病症。在某些實施方案中，包含轉基因和PV選擇性微小RNA結合位點以及可選的PV選擇性調節元件的表達盒、表達載體或病毒顆粒用來治療神經元狀況。適合治療的神經元疾病或病症包括但不限於Dravet綜合症、阿爾茨海默病、帕金森病、亨廷頓病、肌萎縮側索硬化(ALS)、脊髓性肌萎縮(SMA)、癲癇、神經變性疾病、運動障礙、活動疾病、情緒障礙、運動神經元疾病、進行性肌萎縮(PMA)、進行性延髓麻痹、假性延髓麻痹、原發性側索硬化、AIDS的神經系統後果、發育障礙、多發性硬化、神經遺傳性疾病、中風、脊髓損傷、外傷性腦損傷、tau蛋白病變、神經元興奮不足和/或癲癇發作。在一些實施方案中，包含轉基因和PV選擇性微小RNA結合位點以及可選的PV選擇性調節元件的病毒載體、病毒顆粒或藥物組合物用來治療精神疾病(例如，精神分裂症、強迫性障礙、成癮、抑鬱、焦慮、精神病)；自閉症譜系障礙(例如，脆性X綜合症、Rett綜合症)；癲癇(例如，Dravet綜合症、慢性外傷性腦病、全身性癲癇伴熱性癲癇發作+(GEFS+)、癲癇性腦病、顳葉癲癇、局灶性癲癇、結節性硬化症、與SCN1A單倍性不足相關的癲癇)；和/或神經變性(例如，阿爾茨海默病、帕金森病)。與功能失調的PV神經元相關的疾病，例如由於SCN1A或

Nav1.1中的功能喪失突變而引起的疾病，包括：Dravet綜合症、大田原綜合症、癲癇、早期嬰兒型癲癇性腦病6(EIEE6)、家族性熱性癲癇發作3A(FEB3A)、難治性兒童癲癇伴全身性強直-陣攣性癲癇發作(ICEGTC)、家族性偏癱性偏頭痛3(FHM3)、Panayiotopoulos綜合症、家族性心房顫動13(ATFB13)、全身性癲癇伴熱性癲癇發作+1型(gefs+ 1型)、Brugada綜合症、非特異性心臟傳導缺陷、全身性癲癇伴熱性癲癇發作+、良性家族性嬰兒癲癇發作、早期嬰兒型癲癇性腦病11(EIEE11)、良性家族性嬰兒癲癇、神經變性、tau蛋白病變和阿爾茨海默病。

【0310】 在某些實施方案中，採用本文所述包含轉基因和PV選擇性微小RNA結合位點以及可選的PV選擇性調節元件的表達盒、表達載體或病毒顆粒進行的治療導致與神經元疾病或病症相關的症狀得到改善。例如，可以根據症狀監測帕金森病患者的運動功能的改善，從而指示對治療的陽性反應。向具有發生神經元病症的風險的受試者施用採用本文所述方法的療法可以防止一種或多種症狀的發生或減緩其進展。

【0311】 在某些實施方案中，本文提供的包含轉基因和PV選擇性微小RNA結合位點以及可選的PV選擇性調節元件的表達盒、表達載體或病毒顆粒可用來治療已被診斷為神經元疾病的受試者，該疾病例如是與SCN1A單倍性不足相關的癲癇，例如Dravet綜合症。在各個實施方案中，本文公開的任何神經元疾病或病症是由已知的遺傳事件(例如，本領域中已知的任何*SCN1A*突變)引起的或具有未知的原因。

【0312】 在某些實施方案中，本文提供的包含轉基因和PV選擇性微小RNA結合位點以及可選的PV選擇性調節元件的表達盒、表達載體或病毒顆粒可用來治療具有發生疾病或病症的風險的受試者。在一些實施方案

中，可能已知受試者易患疾病，例如，神經元疾病(例如，與SCN1A單倍性不足相關的癲癇，例如Dravet綜合症)。在一些實施方案中，由於遺傳事件，或由於已知的危險因素，受試者可能易患疾病。例如，受試者可以在*SCN1A*中攜帶與癲癇或Dravet綜合症相關的突變，例如，插入、缺失、倒位、易位或置換(例如，點突變，包括無義突變和/或錯義突變)。

【0313】 在某些實施方案中，本文提供的包含轉基因和PV選擇性微小RNA結合位點以及可選的PV選擇性調節元件的表達盒、表達載體或病毒顆粒可用來減輕與疾病或病症相關的一種或多種症狀。例如，與Dravet綜合症相關的症狀包括癲癇發作、記憶缺陷、發育遲緩、肌張力差和/或認知問題。用本文提供的包含轉基因和PV選擇性微小RNA結合位點以及可選的PV選擇性調節元件的病毒載體、病毒顆粒或藥物組合物進行的治療可以導致一種或多種症狀得到改善，例如癲癇發作的次數、持續時間和/或強度減小。

【0314】 在某些實施方案中，本文所述的方法用於增加轉基因在PV神經元中的表達、基因療法或治療靈長類動物的疾病或病症。在某些實施方案中，該靈長類動物是人。在某些實施方案中，該靈長類動物為非人靈長類動物。在某些實施方案中，該非人靈長類動物是舊大陸猴、猩猩、大猩猩、黑猩猩、食蟹猴、恒河猴或豬尾猴。

【0315】 術語“受試者”和“個體”在本文中可互換使用，是指脊椎動物，優選哺乳動物，更優選人。本文所述的方法可用於人類治療、獸醫應用和/或在疾病或病況的動物模型中的臨床前研究。在各個實施方案中，可根據本文所述的方法治療的受試者是哺乳動物，例如小鼠、大鼠、倉鼠、豚鼠、沙鼠、牛、綿羊、豬、山羊、驢、馬、狗、貓、美洲駝、猴

(例如，舊大陸猴、絨猴或獼猴，如恒河猴、豬尾猴或食蟹猴(即食蟹獼猴))、猿(例如，猩猩、大猩猩或黑猩猩)或人。在示例性實施方案中，受試者是人。

【0316】 下表提供了本文公開的序列。

表1. 本文公開的示例性工程化轉錄因子。調節元件(RE)的序列在以下表2和表8中公開。對於RE，當示出m1時，意指在編碼區與聚A尾之間包括m1微小RNA結合位點(SEQ ID NO: 7，表8)。DNA結合域(DBD)的序列在以下表3中公開。對於DBD，工程化鋅指(eZF)表示該構建體具有SEQ ID NO: 147(表10)所示的通式；EGR1表示DBD衍生自野生型人EGR1(SEQ ID NO: 176；表12)；而EGR3表示DBD衍生自野生型人EGR3(SEQ ID NO: 175，表12)。靶位點的序列(例如，被DBD結合的序列)在以下表4中提供。轉錄啟動域(TAD)的序列在以下表5中公開。對於TAD，(c)表示TAD位於DBD的c-末端，(n)表示TAD位於DBD的n-末端，(n/c)表示存在位於DBD的n-末端和c-末端的TAD，2x CITED4(n)表示有2個拷貝的CITED4 TAD位於DBD的n-末端。全長工程化轉錄因子(DBD + TAD)的序列在以下表6中提供。

構建體	RE	DBD	靶位點	TAD (位置)	SEQ ID NO (DBD)	SEQ ID NO (DBD + TAD)
1	RE 1	eZF	Z13	VPR (c)	89	99
2	RE 1	eZF	Z1	VPR (c)	77	100
3	RE 1	eZF	Z13	VP64 (c)	89	101
4	RE 1	eZF	Z1	VP64 (c)	77	102
5	RE 1	EGR1	Z13	CITED4 (n/c)	93	103
6	RE 1	EGR1	Z13	CITED4 (n)	93	104
7	RE 1	EGR1	Z1	CITED4 (n/c)	92	105
8	RE 1	EGR1	Z1	CITED4 (n)	92	106
9	RE 1	EGR1	Z1	CITED4 (c)	92	107

構建體	RE	DBD	靶位點	TAD (位置)	SEQ ID NO (DBD)	SEQ ID NO (DBD + TAD)
10	RE 1	EGR1	Z1	CITED2 (c)	92	108
11	RE 1	EGR1	Z1	CITED2 (n)	92	109
12	RE 1	EGR3	Z1	CITED4 (n/c)	96	110
13	RE 1	EGR3	Z1	CITED4 (c)	96	111
14	RE 1	EGR3	Z1	CITED2 (c)	96	112
15	RE 1	EGR3	Z1	CITED2 (n)	96	113
16	CBA	EGR3	Z15	N/A	98	114
17	RE 1	EGR1	Z13	N/A	93	115
18	RE 1	EGR1	Z15	N/A	94	116
19	RE 1	EGR1	Z13	N/A	93	117
20	RE 1	EGR1	Z13	N/A	93	115
21	RE 1	EGR1	Z1	N/A	92	118
22	CBA	EGR3	Z13	N/A	97	119
23	RE 1	EGR1	Z17	N/A	95	120
24	RE 1	EGR1	Z13	N/A	93	121
25	CBA	EGR3	Z1	N/A	96	122
26	RE 1	EGR1	Z1	N/A	92	123
27	RE 1	EGR1	Z1	N/A	92	124
28	RE 1	eZF	Z8	VP64 (c)	84	125
29	RE 1	eZF	Z14	VP64 (c)	90	126
30	CBA	eZF	Z13	VPR (c)	89	99
31	RE 2 (m1)	eZF	Z1	VP64 (c)	77	102
32	RE 2	eZF	Z1	VP64 (c)	77	102
33	RE 2	eZF	Z1	VPR (c)	77	100
34	RE 2	eZF	Z1	VP64 (c)	77	127
35	RE 2 (m1)	eZF	Z1	VP64 (c)	77	127
36	RE 2	EGR1	Z1	CITED4 (n)	92	128
37	RE 2 (m1)	EGR1	Z1	CITED4 (n)	92	106
38	RE 2 (m1)	EGR1	Z1	CITED4 (n)	92	129
39	RE 2 (m1)	EGR1	Z1	CITED4 (n/c)	92	105
40	RE 2 (m1)	EGR1	Z1	2x CITED4 (n)	92	130
41	RE 2 (m1)	EGR1	Z1	2x CITED4 (n)	92	131

構建體	RE	DBD	靶位點	TAD (位置)	SEQ ID NO (DBD)	SEQ ID NO (DBD + TAD)
42	RE 2 (m2)	eZF	Z1	VP64 (c)	77	102
43	RE 2 (m3)	eZF	Z1	VP64 (c)	77	102
44	RE 2 (m2)	EGR1	Z1	CITED4 (n)	92	106
45	RE 2 (m3)	EGR1	Z1	CITED4 (n)	92	106
46	RE 1	EGR1	Z1	CITED4 (n)	92	205
47	RE 1	EGR1	Z1	2x CITED4 (n)	92	207
48	RE 1	EGR1	Z1	2x CITED4 (n)	92	209
49	CBA	eZF	Z1	CREB3 (n)	77	213
50	CBA	EGR1	Z1	CREB3 (n)	92 (無C-末端Lys)	217
51	CBA	EGR1	Z13	CREB3 (n)	93 (無C-末端Lys)	219
52	CBA	eZF	Z1	CREB3 (n)；在(c)處無TM結構域	77	221
53	CBA	EGR1	Z1	CREB3 (n)；在(c)處無TM結構域	92 (無C-末端Lys)	223

表2. 本文公開的各種調節元件(RE)的核酸序列。

RE	序列	SEQ ID NO
RE 1	GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTGGTACCGTGTGTATGCTCAGGGGCTGGGAAAGGAGGGGAGGGAGCTCCGGCTCAG	1
RE 2	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaacatattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaagggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtgtggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctgattggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatgaaagtcatttaagcacaactaactttttttttttttttttttttgagacagagtcttgctctgttgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcctaggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtggcccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcaccaggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagtgctgggattataggcgtcagccactatgcccaacccgaccaacctttttaaaataaatatttaaaaaattgg	2

RE	序列	SEQ ID NO
	tatttcacatatatactagtatttacatttatccacacaaaacggacgggcctccgctgaaccagtg aggccccagacgtgcgcataaataacccctgcgtgctgcaccacctggggagaggggaggg accacggtaaatggagcgagcgcatagcaaaagggacgcggggtccttttctctgccggtgg cactgggtagctgtggccaggtgtggtactttgatggggcccagggctggagctcaaggaag cgtcgcagggtcacagatctgggggaaccccggggaaaagcactgaggcaaaaccgccgc tcgtctcctacaatatatgggaggggggaggttgagtacgttctggattactcataagacctttttttt ttccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagaggcggtc aggcacctgcagaggagccccgccgctccgccgactagctgcccccgcgagcaacggcct cgtgatttccccgccgatccggtccccgcctccccactctgcccccgcctaccccggagccgt gcagccgcctctccgaatctctctcttctcctggcgctcgcgtgcgagagggaactagcgaga acgaggaagcagctggaggtgacgccgggcagattacgcctgtcagggccgagccgagcg gatcgctgggcgctgtgcagaggaaaggcgggagtgcccggctcgctgtcgcagagccga ggtgggtaagctagcgaccacctggacttcccagcgcccaaccgtggcttttcagccaggtcc tctcctcccgcggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatga gtgcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttccaggggc cgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggctagagtgcaaggtg actgtggttcttctctggccaagtccgagggagaacgtaaagatatgggccttttttcccctctca ccttgtctcaccaaagtccctagtccccggagcagttagcctctttctttccagggaattagccag acacaacaacgggaaccagacaccgaaccagacatgcccgccccgtgcgccctccccgctc gctgcctttcctccctcttgtctctccagagccggatcttcaaggggagcctccgtgccccccggc tgctcagtccctccggtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgc agcctgtttctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccga gctg	
RE 3	TCAACAGGGGGACACTTGGGAAAGAAGGATGGGGACAG AGCCGAGAGGACTGTTACACATTAGAGAAACATCAGTGA CTGTGCCAGCTTTGGGGTAGACTGCACAAAAGCCCTGAG GCAGCACAGGCAGGATCCAGTCTGCTGGTCCCAGGAAGC TAACCGTCTCAGACAGAGCACAAAGCACCGAGACATGTG CCACAAGGCTTGTGTAGAGAGGTCAGAGGACAGCGTACA GGTCCCAGAGATCAAACTCAACCTCACCAGGCTTGGCAG CAAGCCTTTACCAACCCACCCCCACCCCACCCACCCTGCA CGCGCCCCTCTCCCCTCCCCATGGTCTCCCATGGCTATCT CACTTGGCCCTAAAATGTTTAAGGATGACACTGGCTGCTG AGTGGAAATGAGACAGCAGAAGTCAACAGTAGATTTTAG GAAAGCCAGAGAAAAAGGCTTGTGCTGTTTTTAGAAAGC CAAGGGACAAGCTAAGATAGGGCCCAAGTAATGCTAGTA TTTACATTTATCCACACAAAACGGACGGGCCTCCGCTGAA CCAGTGAGGCCCCAGACGTGCGCATAAATAACCCCTGCG TGCTGCACCACCTGGGGAGAGGGGGAGGACCACGGTAAA TGGAGCGAGCGCATAGCAAAAGGGACGCGGGGTCCTTTT CTCTGCCGGTGGCACTGGGTAGCTGTGGCCAGGTGTGGT ACTTTGATGGGGCCCAGGGCTGGAGCTCAAGGAAGCGTC GCAGGGTCACAGATCTGGGGGAACCCCGGGGAAAAGCA CTGAGGCAAAACCGCCGCTCGTCTCCTACAATATATGGG AGGGGGAGGTTGAGTACGTTCTGGATTACTCATAAGACC TTTTTTTTTTCCTTCCGGGCGCAAAACCGTGAGCTGGATTT ATAATCGCCCTATAAAGCTCCAGAGGCGGTCAGGCACCT GCAGAGGAGCCCCGCCGCTCCGCCGACTAGCTGCCCCCG	3

RE	序列	SEQ ID NO
	CGAGCAACGGCCTCGTGATTTCCCCGCCGATCCGGTCCCC GCCTCCCCACTCTGCCCCCGCCTACCCCGGAGCCGTGCAG CCGCCTCTCCGAATCTCTCTCTTCTCCTGGCGCTCGCGTGC GAGAGGGAACTAGCGAGAACGAGGAAGCAGCTGGAGGT GACGCCGGGCAGATTACGCCTGTCAGGGCCGAGCCGAGC GGATCGCTGGGCGCTGTGCAGAGGAAAGGCGGGAGTGCC CGGCTCGCTGTCGCAGAGCCGAGGTGGGTAAGCTAGCGA CCACCTGGACTTCCCAGCGCCCAACCGTGGCTTTTCAGCC AGGTCCTCTCCTCCCGCGGCTTCTCAACCAACCCCATCCC AGCGCCGGCCACCCAACCTCCCGAAATGAGTGCTTCCTG CCCCAGCAGCCGAAGGCGCTACTAGGAACGGTAACCTGT TACTTTTCCAGGGGCCGTAGTCGACCCGCTGCCCGAGTTG CTGTGCGACTGCGCGCGCGGGGCTAGAGTGCAAGGTGAC TGTGGTTCTTCTCTGGCCAAGTCCGAGGGAGAACGTAAA GATATGGGCCTTTTTCCCCCTCTCACCTTGTCTCACCAAA GTCCCTAGTCCCCGGAGCAGTTAGCCTCTTTCTTTCCAGG GAATTAGCCAGACACAACAACGGGAACCAGACACCGAA CCAGACATGCCCGCCCCGTGCGCCCTCCCCGCTCGCTGCC TTTCCTCCCTCTTGTCTCTCCAGAGCCGGATCTTCAAGGG GAGCCTCCGTGCCCCCGGCTGCTCAGTCCCTCCGGTGTGC AGGACCCCGGAAGTCCTCCCCGCACAGCTCTCGCTTCTCT TTGCAGCCTGTTTCTGCGCCGGACCAGTCGAGGACTCTGG ACAGTAGAGGCCCCGGGACGACCGAGCTG	
RE 4	GCCCTCTAGGCCACCTGACCAGGTCCCCTCAGTCCCCCCC TTCCCACACTCCCACACTCAGCCCCCCTCCCCCCCCCCCG ACCCCTGCAGGATTATCCTGTCTGTGTTCCTGACTCAGCC TGGGAGCCACCTGGGCAGCAGGGGCCAAGGGTGTCCTAG AAGGGACCTGGAGTCCACGCTGGGCCAAGCCTGCCCTTT CTCCCTCTGTCTTCCGTCCCTGCTTGCGGTTCTGCTGAATG TGGTTATTTCTCTGGCTCCTTTTACAGAGAATGCTGCTGCT AATTTTATGTGGAGCTCTGAGGCAGTGTAATTGGAAGCC AGACACCCTGTCAGCAGTGGGCTCCCGTCCTGAGCTGCC ATGCTTCCTGCTCTCCTCCCGTCCCGGCTCCTCATTTCATG CAGCCACCTGTCCCAGGGAGAGAGGAGTCACCCAGGCCC CTCAGTCCGCCCCTTAAATAAGAAAGCCTCCGTTGCTCGG CACACATACCAAGCAGCCGCTGGTGCAATCT	4
CBA (CMV增強子+雞beta肌動蛋白啟動子)	CGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCG CCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTC AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAG TACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGA CGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA GTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATC TACGTATTAGTCATCGCTATTACCATGggtcgaggtgagccccacgtt ctgcttcactctccccatctccccccccctccccaccccccaattttgtatttatttatttttttaattattttg tgcagcgatgggggcgggggggggggggggcgcgcgccaggcggggcggggcgggggcg aggggcggggcggggcgaggcggagaggtgcggcggcagccaatcagagcggcgcgct ccgaaagtttccttttatggcgaggcggcggcggcggccctataaaaagcgaagcgcgc	5

RE	序列	SEQ ID NO
	ggcgggcgggagtcgctgcgttgccttcgccccgtgccccgctccgcgccgcctcgcgccg cccgccccggctctgactgaccgcgttactcccacaggtgagcgggcgggacggccccttctc ctccgggctgtaattagcgcttggtttaatgacggctcgtttcttttctgtggctgcgtgaaagcctt aaagggctccgggagggccctttgtgcgggggggagcggctcgggggtgcgtgcgtgtgt gtgtgcgtggggagcgccgcgtgcggcccgcgctgcccggcggctgtgagcgctgcgggc gcggcgcggggctttgtgcgctccgcgtgtgcgcgaggggagcgcggccgggggcggtgc cccgcggtgcgggggggctgcgagggggaacaaaggctgcgtgcggggtgtgtgcgtgggg gggtgagcaggggggtgtgggcgcggcggtcgggctgtaacccccccccctgcacccccctccc cgagttgctgagcacggcccggcttcgggtgcggggctccgtgcggggcgtggcgcgggg ctcgccgtgccgggcgggggggtggcggcaggtgggggtgccgggcgggggcgggggccgc ctcgggccggggagggctcgggggaggggcgcggcggcccccggagcgccggcggctgt cgaggcgcggcgagccgcagccattgccttttatggtaatcgtgcgagagggcgcagggact tcctttgtcccaaatctggcggagccgaaatctgggaggcgccgccgcacccccctctagcggg cgcgggcgaagcggtgcggcgccggcaggaaggaaatgggcggggagggccttcgtgcg tcgccgcgccgccgtcccccttctccatctccagcctcggggctgccgcaggggacggctgc cttcgggggggacggggcagggcggggttcggcttctggcgtgtgaccggcggctctagag cctctgctaaccatgttcatgccttcttcttttttcctacagctcctgggcaacgtgctggttgttgtgc tgtctcatcattttggcaaagaatt	
EF1alpha	GAGTAATTCATACAAAAGGACTCGCCCCTGCCTTGGGGA ATCCCAGGGACCGTCGTTAAACTCCCACTAACGTAGAAC CCAGAGATCGCTGCGTTCCCGCCCCCTCACCCGCCCGCTC TCGTCATCACTGAGGTGGAGAAGAGCATGCGTGAGGCTC CGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGT CCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGG TGCCTAGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGA TGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGGTGGGGGA GAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTT TTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCG TGTGTGGTTCCCGCGGGCCTGGCCTCTTTACGGGTTATGG CCCTTGCGTGCCTTGAATTACTTCCACGCCCCTGGCTGCA GTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGG TGGGAGAGTTCGAGGCCTTGCGCTTAAGGAGCCCCTTCG CCTCGTGCTTGAGTTGAGGCCTGGCTTGGGCGCTGGGGCC GCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGC TGCTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGA CCTGCTGCGACGCTTTTTTTCTGGCAAGATAGTCTTGTAA ATGCGGGCCAAGATCTGCACACTGGTATTTCGGTTTTTGG GGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAGCGCAC ATGTTCGGCGAGGCGGGGCCTGCGAGCGCGGCCACCGAG AATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCT GGTGCCTGGCCTCGCGCCGCCGTGTATCGCCCCGCCCTGG GCGGCAAGGCTGGCCCGGTCGGCACCAGTTGCGTGAGCG GAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCA AAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTGA GTCACCCACACAAAGGAAAAGGGCCTTTCCGTCCTCAGC CGTCGCTTCATGTGACTCCACGGAGTACCGGGCGCCGTCC AGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGT CTTTAGGTTGGGGGGAGGGGTTTTATGCGATGGAGTTTCC	6

RE	序列	SEQ ID NO
	CCACACTGAGTGGGTGGAGACTGAAGTTAGGCCAGCTTG GCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGAGT TTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCA AAGTTTTTTTCTTCCATTTCAGGTGTCGTGA	

表3. 本文提供的示例性DNA結合域(DBD)的氨基酸序列。對於DBD，工程化鋅指(eZF)表示該構建體具有SEQ ID NO: 147(表10)所示的通式；EGR1表示DBD衍生自野生型人EGR1(SEQ ID NO: 176；表12)；而EGR3表示DBD衍生自野生型人EGR3(SEQ ID NO: 175，表12)。靶位點是被DBD結合的序列，並在以下表4中提供。

DBD/靶位點	序列	SEQ ID NO
eZF/z1	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPEC GKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDELVRH QRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKC PECGKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNSTL TEHQRTHTGKKTS	77
eZF/z2	LEPGEKPYKCPECGKSFSTKNSLTEHQRTHTGEKPYKCPEC GKSFSRADNLTEHQRTHTGEKPYKCPECGKSFSQLAHLRA HQRTHTGEKPYKCPECGKSFSTKNSLTEHQRTHTGEKPYK CPECGKSFSQAGHLASHQRTHTGEKPYKCPECGKSFSTHLD LIRHQRTHTGKKTS	78
eZF/z3	LEPGEKPYKCPECGKSFSQAGHLASHQRTHTGEKPYKCPEC GKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSTSGNLTEH QRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCP ECGKSFSQKSSLIAHQRTHTGEKPYKCPECGKSFSQAGHLA SHQRTHTGKKTS	79
eZF/z4	LEPGEKPYKCPECGKSFSTTGNLTVHQRTHTGEKPYKCPEC GKSFSTSGELVRHQRTHTGEKPYKCPECGKSFSREDNLHTH QRTHTGEKPYKCPECGKSFSTSGNLTEHQRTHTGEKPYKCP ECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSQRANLR AHQRTHTGKKTS	80
eZF/z5	LEPGEKPYKCPECGKSFSSRRTCRAHQRTHTGEKPYKCPEC GKSFSTTGALTEHQRTHTGEKPYKCPECGKSFSRSDELVRH QRTHTGEKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKC PECGKSFSQSGDLRRHQRTHTGEKPYKCPECGKSFSTSHSL TEHQRTHTGKKTS	81
eZF/z6	LEPGEKPYKCPECGKSFSRKDNLKNHQRTHTGEKPYKCPE CGKSFSDPGALVRHQRTHTGEKPYKCPECGKSFSREDNLH THQRTHTGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPY	82

DBD/靶位點	序列	SEQ ID NO
	KCPECGKSFSTSGELVRHQRTHTGEKPYKCPECGKSFSRKDNLKNHQRTHTGKKTS	
eZF/z7	LEPGEKPYKCPECGKSFSSKKALTEHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTSGHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTS	83
eZF/z8	LEPGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTSGHLVRHQRTHTGKKTS	84
eZF/z9	LEPGEKPYKCPECGKSFSSKKALTEHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSGNLTEHQRTHTGKKTS	85
eZF/z10	LEPGEKPYKCPECGKSFSDCRDLARHQRTHTGEKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSDPGNLVRHQRTHTGEKPYKCPECGKSFSQRAHLERHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRTHTGKKTS	86
eZF/z11	LEPGEKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPECGKSFSSPADLTRHQRTHTGEKPYKCPECGKSFSDPGNLVRHQRTHTGEKPYKCPECGKSFSQRAHLERHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRTHTGKKTS	87
eZF/z12	LEPGEKPYKCPECGKSFSRNDALTEHQRTHTGEKPYKCPECGKSFSDPGHLVRHQRTHTGEKPYKCPECGKSFSTSGELVRHQRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCPECGKSFSSKKALTEHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEKPYKCPECGKSFSRSDHLTNHQRTHTGKKTS	88
eZF/z13	LEPGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRTHTGKKTS	89
eZF/z14	LEPGEKPYKCPECGKSFSDPGALVRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKCPECGKSFSQSGDLRRHQRTHTGEKPYKCPECGKSFSTHLDLIRHQRTHTGEKPYKCPECGKSFSTSGNLVRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGKKTS	90

DBD/靶位點	序列	SEQ ID NO
eZF/z15	LEPGEKPYKCPECGKSFSRRDELNVHQRTHTGEKPYKCPEC GKSFSRSDHLTNHQRTHTGEKPYKCPECGKSFSRSDDLVRH QRTHTGEKPYKCPECGKSFSRSDNLVRHQRTHTGEKPYKC PECGKSFSHRTTLTNHQRTHTGEKPYKCPECGKSFSREDNL HTHQRTHTGEKPYKCPECGKSFSTSHSLTEHQRTHTGEKPY KCPECGKSFSQSSSLVRHQRTHTGEKPYKCPECGKSFSRED NLHTHQRTHTGKKTS	91
EGR1/z1	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNF SREDNLHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHL RQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRIC MRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEH TKIHLRQKDK	92
EGR1/z13	RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNF SHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHL RQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQCRIC MRNFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNLHTH TKIHLRQKDK	93
EGR1/z15	RPYACPVESCDRRFSRRDELNVHIRIHTGQKPFQCRICMRN FSRSDHLTNHIRTHTGEKPFACDICGRKFARSDDLVRHTKIH LRQKDRPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRI CMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFAREDNLHT HTKIHLRQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKP FQCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFARED NLHTHTKIHLRQKD	94
EGr1/z17	RPYACPVESCDRRFSDPGALVRHIRIHTGQKPFQCRICMRN FSRSDNLVRHIRTHTGEKPFACDICGRKFAQSGDLRRHTKI HLRQKDRPYACPVESCDRRFSTHLDLIRHIRIHTGQKPFQCR ICMRNFSTSGNLVRHIRTHTGEKPFACDICGRKFARSDNLV RHTKIHLRQKDRPYACPVESCDRRFSQSGHLTEHIRIHTGQ KPFQCRICMRNFSERSHLREHIRTHTGEKPFACDICGRKFAQ AGHLASHTKIHLRQKD	95
EGR3/z1	RPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRS FSREDNLHTHIRTHTGEKPFACEFCGRKFARSDELVRHAKI HLKQKEHACPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRI CMRSFSTSGHLVRHIRTHTGEKPFACEFCGRKFAQNSTLTE HAKIHLKQKEK	96
EGR3/z13	RPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRICMRS FSHRTTLTNHIRTHTGEKPFACEFCGRKFAREDNLHTHAKI HLKQKEHACPAEGCDRRFSTSHSLTEHLRIHTGHKPFQCRI CMRSFSQSSSLVRHIRTHTGEKPFACEFCGRKFAREDNLHT HAKIHLKQKEK	97
EGR3/z15	RPHACPAEGCDRRFSRRDELNVHLRIHTGHKPFQCRICMRS FSRSDHLTNHIRTHTGEKPFACEFCGRKFARSDDLVRHAKI HLKQKEHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCRI CMRSFSHRTTLTNHIRTHTGEKPFACEFCGRKFAREDNLHT	98

DBD/靶位點	序列	SEQ ID NO
	HAKIHLKQKEHACPAEGCDRRFSTSHSLTEHLRIHTGHKPF QCRICMRSFSQSSSLVRHIRTHTGEKPFACEFCGRKFAREDN LHTHAKIHLKQKEK	

表4. 被本文公開的DNA結合域結合的示例性靶位點的靶位點序列和染色體位置。

染色體2起始位置	靶位點序列	靶位點	靶位點序列的SEQ ID NO
166149168	CTAGGTCAAGTGTAGGAG	z1	18
166149158	ACTTGACCTAGACAGCCT	z2	19
166073978	TGAATAACTCATTAGTGA	z3	20
166073933	AAAGTACATTAGGCTAAT	z4	21
166149199	CCAGCACTGGTGCTTCGT	z5	22
166149176	AAGGCTGTCTAGGTCAAG	z6	23
166149168	CTAGGTCAAGTGTAGGAGACACAC	z7	24
166149165	GGTCAAGTGTAGGAGACA	z8	25
166149162	CAAGTGTAGGAGACACAC	z9	26
166149160	AGTGTAGGAGACACACTGCTGGCC	z10	27
166149160	AGTGTAGGAGACACACTG	z11	28
166149155	AGGAGACACACTGCTGGCCTG	z12	29
166128025	TAGGTACCATAGAGTGAG	z13	30
166127991	GAGGATACTGCAGAGGTC	z14	31
166127999	TAGGTACCATAGAGTGAGGCGAGGATG	z15	32
166127991	ATAGAGTGAGGCGAGGATGAAGCCGAG	z16	33
166127974	TGAAGCCGAGAGGATACTGCAGAGGTC	z17	34

表5. 本文公開的示例性轉錄啟動域(TAD)的氨基酸序列。

TAD	序列	SEQ ID NO
VPR	DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGS DALDDFDLDMLINSRSSGSPKKKRKVGSQYLPDTDDRHRIE EKRKRTYETFKSIMKKSPFSGPTDPRPPPRRIAVPSRSSASVP KPAPQPYPFTSSLSTINYDEFPTMVFPSGQISQASALAPAPPQ VLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPPQAVAPPAP KPTQAGEGTLSEALLQLQFDDEDLGALLGNSTDPAVFTDLA SVDNSEFQQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQ	132

TAD	序列	SEQ ID NO
	RPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLGSGS GSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPG SPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAV TPEASHLLEDPDEETSQAVKALREMADTVIPQKEEAAICGQ MDLSHPPPRGHLDELTTTLESMTEDLNLDSPLTPELNEILDTF LNDECLLHAMHISTGLSIFDTSLF	
VP64	DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLGS DALDDFDLDML	133
CITED2	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMGMGQ FPSPHHHQQQQPQHAFNALMGEHIHYGAGNMNATSGIRHA MGPGTVNGGHPPSALAPAARFNNSQFMGPPVASQGGSLPA SMQLQKLNNQYFNHHPYPHNHYMPDLHPAAGHQMNGTNQ HFRDCNPKHSGGSSTPGGSGGSSTPGGSGSSSGGGAGSSNSG GGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLVIE MGLDRIKELPELWLGQNEFDFMTDFVCKQQPSRVSC	134
CITED4	ADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLD SGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPA AGIAHLQPVATPYPGRAAAPPNAPGGPPGPQPAPSAAAPPPP AHALGGMDAELIDEEALTSLELELGLHRVRELPELFLGQSEF DCFSDLGSAPPAGSVSC	135
CREB3	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWALP LSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCLVHHDHTYS LPRETVSMDLESESCRKEGTQMTPQHMEELAEQEIARLVLT DEEKSLLEKEGLILPETLPLTKTEEQILKRVR	224

表6. 本文公開的示例性工程化轉錄因子(DBD + TAD)的氨基酸序列。

構建體	序列	SEQ ID NO
1	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRT HTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKC PECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFSQS SSLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE EASGSGRADALDDFDLDMLGSDALDDFDLDMLGSDA LDDFDLDMLGSDALDDFDLDMLINSRSSGSPKKKRKV GSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPT DPRPPPRRIAVPSRSSASVPKPAPQPYPFTSSLSTINYDE FPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPAMV SALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLS EALLQLQFDDEDLGALLGNSTDPAVFTDLASVDNSEF QQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPP DPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLGSG SGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRP	99

構建體	序列	SEQ ID NO
	FHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVP QPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADT VIPQKEEAAICGQMDLSHPPPRGHLDELTTTLESMTED LNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDT SLF	
2	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE EASGSGRADALDDFDLDMLGSDALDDFDLDMLGSDA LDDFDLDMLGSDALDDFDLDMLINSRSSGSPKKKRKV GSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPT DPRPPPRRIAVPSRSSASVPKPAPQPYPFTSSLSTINYDE FPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPAMV SALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLS EALLQLQFDDEDLGALLGNSTDPAVFTDLASVDNSEF QQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPP DPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLGSG SGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRP FHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVP QPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADT VIPQKEEAAICGQMDLSHPPPRGHLDELTTTLESMTED LNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDT SLF	100
3	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRT HTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKC PECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFSQS SSLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	101
4	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	102
5	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS	103

構建體	序列	SEQ ID NO
	GGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSHRTT LTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHLR QKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQC RICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFARE DNLHTHTKIHLRQKDKLEMADHLMLAEGYRLVQRPP SAAAAHGPHALRTLPPYAGPGLDSGLRPRGAPLGPPP PRQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVA TPYPGRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALG GMDAELIDEEALTSLELELGLHRVRELPELFLGQSEFD CFSDLGSAPPAGSVSC	
6	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSHRTT LTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIHLR QKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPFQC RICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFARE DNLHTHTKIHLRQKDK	104
7	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDN LHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQC RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQN STLTEHTKIHLRQKDKLEMADHLMLAEGYRLVQRPPS AAAAHGPHALRTLPPYAGPGLDSGLRPRGAPLGPPPP RQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVAT PYPGRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGG MDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSC	105
8	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDN LHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQC	106

構建體	序列	SEQ ID NO
	RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQN STLTEHTKIHLRQKDK	
9	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDR RFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDNLHT HIRTHTGEKPFACDICGRKFARSDELVRHTKIHLRQKD RPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRIC MRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTL TEHTKIHLRQKDKLEMADHLMLAEGYRLVQRPPSAA AAHGPHALRTLPPYAGPGLDSGLRPRGAPLGPPPPRQP GALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVATPYP GRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGGMDA ELIDEEALTSLELELGLHRVRELPELFLGQSEFDCFSDL GSAPPAGSVSC	107
10	MQSQLIKPSRMRKYPNRPSKTPPHERPYACPVESCDR RFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDNLHT HIRTHTGEKPFACDICGRKFARSDELVRHTKIHLRQKD RPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRIC MRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTL TEHTKIHLRQKDKLEMSGLEMADHMMAMNHGRFPD GTNGLHHHPAHRMGMGQFPSPHHHQQQQPQHAFNA LMGEHIHYGAGNMNATSGIRHAMGPGTVNGGHPPSA LAPAARFNNSQFMGPPVASQGGSLPASMQLQKLNNQ YFNHHPYPHNHYMPDLHPAAGHQMNGTNQHFRDCN PKHSGGSSTPGGSGGSSTPGGSGSSSGGGAGSSNSGGG SGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMSLV IEMGLDRIKELPELWLGQNEFDFMTDFVCKQQPSRVS C	108
11	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMG MGQFPSPHHHQQQQPQHAFNALMGEHIHYGAGNMN ATSGVRHAMGPGTVNGGHPPSALAPAARFNNSQFMG PPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMP DLHPAAGHQMNGTNQHFRDCNPKHSGGSSTPGGSGG SSTPGGSGSSSGGGAGSSNSGGGSGSGNMPASVAHVP AAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIKELPELW LGQNEFDFMTDFVCKQQPSRVSCQSQLIKPSRMRKYP NRPSKTPPHERPYACPVESCDRRFSRSDNLVRHIRIHT GQKPFQCRICMRNFSREDNLHTHIRTHTGEKPFACDIC GRKFARSDELVRHTKIHLRQKDRPYACPVESCDRRFS QSGNLTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHIR THTGEKPFACDICGRKFAQNSTLTEHTKIHLRQKDK	109
12	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKP FQCRICMRSFSREDNLHTHIRTHTGEKPFACEFCGRKF	110

構建體	序列	SEQ ID NO
	ARSDELVRHAKIHLKQKEHACPAEGCDRRFSQSGNLT EHLRIHTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEK PFACEFCGRKFAQNSTLTEHAKIHLKQKEKLEMADHL MLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLD SGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAV PPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQPAP SAAAPPPPAHALGGMDAELIDEEALTSLELELGLHRV RELPELFLGQSEFDCFSDLGSAPPAGSVSC	
13	MRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCR ICMRSFSREDNLHTHIRTHTGEKPFACEFCGRKFARSD ELVRHAKIHLKQKEHACPAEGCDRRFSQSGNLTEHLR IHTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEKPFAC EFCGRKFAQNSTLTEHAKIHLKQKEKLEMADHLMLA EGYRLVQRPPSAAAAHGPHALRTLPPYAGPGLDSGLR PRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPAVPPPA AGIAHLQPVATPYPGRAAAPPNAPGGPPGPQPAPSAA APPPPAHALGGMDAELIDEEALTSLELELGLHRVRELP ELFLGQSEFDCFSDLGSAPPAGSVSC	111
14	MRPHACPAEGCDRRFSRSDNLVRHLRIHTGHKPFQCR ICMRSFSREDNLHTHIRTHTGEKPFACEFCGRKFARSD ELVRHAKIHLKQKEHACPAEGCDRRFSQSGNLTEHLR IHTGHKPFQCRICMRSFSTSGHLVRHIRTHTGEKPFAC EFCGRKFAQNSTLTEHAKIHLKQKEKKAEKGGAPSAS SAPPVSLAPVVTTCALEMSGLEMADHMMAMNHGRF PDGTNGLHHHPAHRMGMGQFPSPHHHQQQQPQHAF NALMGEHIHYGAGNMNATSGIRHAMGPGTVNGGHPP SALAPAARFNNSQFMGPPVASQGGSLPASMQLQKLN NQYFNHHPYPHNHYMPDLHPAAGHQMNGTNQHFRD CNPKHSGGSSTPGGSGGSSTPGGSGSSSGGGAGSSNSG GGSGSGNMPASVAHVPAAMLPPNVIDTDFIDEEVLMS LVIEMGLDRIKELPELWLGQNEFDFMTDFVCKQQPSR VSC	112
15	MSGLEMADHMMAMNHGRFPDGTNGLHHHPAHRMG MGQFPSPHHHQQQQPQHAFNALMGEHIHYGAGNMN ATSGVRHAMGPGTVNGGHPPSALAPAARFNNSQFMG PPVASQGGSLPASMQLQKLNNQYFNHHPYPHNHYMP DLHPAAGHQMNGTNQHFRDCNPKHSGGSSTPGGSGG SSTPGGSGSSSGGGAGSSNSGGGSGSGNMPASVAHVP AAMLPPNVIDTDFIDEEVLMSLVIEMGLDRIKELPELW LGQNEFDFMTDFVCKQQPSRVSCRPHACPAEGCDRRF SRSDNLVRHLRIHTGHKPFQCRICMRSFSREDNLHTHI RTHTGEKPFACEFCGRKFARSDELVRHAKIHLKQKEH ACPAEGCDRRFSQSGNLTEHLRIHTGHKPFQCRICMRS FSTSGHLVRHIRTHTGEKPFACEFCGRKFAQNSTLTEH AKIHLKQKEKKAEKGGAPSASSAPPVSLAPVVTTCA	113
16	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLFSG SSDSVVHYNQMATENVMDIGLTNEKPNPELSYSGSFQ	114

構建體	序列	SEQ ID NO
	PAPGNKTVTYLGKFAFDSPSNWCQDNIISLMSAGILGV PPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNC GDLYSEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNL FPMIPDYNLYHHPNDMGSIPEHKPFQGMDPIRVNPPPI TPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRKYPNR PSKTPLHERPHACPAEGCDRRFSRRDELNVHLRIHTGH KPFQCRICMRSFSRSDHLTNHIRTHTGEKPFACEFCGR KFARSDDLVRHAKIHLKQKEHACPAEGCDRRFSRSDN LVRHLRIHTGHKPFQCRICMRSFSHRTTLTNHIRTHTG EKPFACEFCGRKFAREDNLHTHAKIHLKQKEHACPAE GCDRRFSTSHSLTEHLRIHTGHKPFQCRICMRSFSQSSS LVRHIRTHTGEKPFACEFCGRKFAREDNLHTHAKIHL KQKEKKAEKGGAPSASSAPPVSLAPVVTTCA	
17	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSHR TTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIH LRQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPF QCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFA REDNLHTHTKIHLRQKDKKADKSVVASSATSSLSSYP SPVATSYPSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYP SPVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSF SASTGLSDMTATFSPRTIEIC	115
18	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSRRDELNVHIRIHTGQKPFQCRICMRNFSRS DHLTNHIRTHTGEKPFACDICGRKFARSDDLVRHTKIH LRQKDRPYACPVESCDRRFSRSDNLVRHIRIHTGQKPF QCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKFA REDNLHTHTKIHLRQKDRPYACPVESCDRRFSTSHSLT EHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKP FACDICGRKFAREDNLHTHTKIHLRQKDKKADKSVV ASSATSSLSSYPSPVATSYPSPVTTSYPSPATTSYPSPVP TSFSSPGSSTYPSPVHSGFPSPSVATTYSSVPPAFPAQV	116

構建體	序列	SEQ ID NO
	SSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	
19	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSHR TTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHIRTH TGEKPFACDICGRKFSTSHSLTEHIRIHTGQKPFQCRIC MRNFSQSSSLVRHIRTHTGEKPFACDICGRKFAREDNL HTHTKIHLRQKDKKADKSVVASSATSSLSSYPSPVATS YPSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYPSPVHSG FPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGL SDMTATFSPRTIEIC	117
20	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSHR TTLTNHIRTHTGEKPFACDICGRKFAREDNLHTHTKIH LRQKDRPYACPVESCDRRFSTSHSLTEHIRIHTGQKPF QCRICMRNFSQSSSLVRHIRTHTGEKPFACDICGRKFA REDNLHTHTKIHLRQKDKKADKSVVASSATSSLSSYP SPVATSYPSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYP SPVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSF SASTGLSDMTATFSPRTIEIC	115
21	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSRE DNLHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIH LRQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPF	118

構建體	序列	SEQ ID NO
	QCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFA QNSTLTEHTKIHLRQKDKKADKSVVASSATSSLSSYPS PVATSYPSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYPS PVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFS ASTGLSDMTATFSPRTIEIC	
22	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLFSG SSDSVVHYNQMATENVMDIGLTNEKPNPELSYSGSFQ PAPGNKTVTYLGKFAFDSPSNWCQDNIISLMSAGILGV PPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNC GDLYSEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNL FPMIPDYNLYHHPNDMGSIPEHKPFQGMDPIRVNPPPI TPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRKYPNR PSKTPLHERPHACPAEGCDRRFSRSDNLVRHLRIHTGH KPFQCRICMRSFSHRTTLTNHIRTHTGEKPFACEFCGR KFAREDNLHTHAKIHLKQKEHACPAEGCDRRFSTSHS LTEHLRIHTGHKPFQCRICMRSFSQSSSLVRHIRTHTGE KPFACEFCGRKFAREDNLHTHAKIHLKQKEKKAEKG GAPSASSAPPVSLAPVVTTCA	119
23	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSDPGALVRHIRIHTGQKPFQCRICMRNFSRS DNLVRHIRTHTGEKPFACDICGRKFAQSGDLRRHTKIH LRQKDRPYACPVESCDRRFSTHLDLIRHIRIHTGQKPF QCRICMRNFSTSGNLVRHIRTHTGEKPFACDICGRKFA RSDNLVRHTKIHLRQKDRPYACPVESCDRRFSQSGHL TEHIRIHTGQKPFQCRICMRNFSERSHLREHIRTHTGEK PFACDICGRKFAQAGHLASHTKIHLRQKDKKADKSVV ASSATSSLSSYPSPVATSYPSPVTTSYPSPATTSYPSPVP TSFSSPGSSTYPSPVHSGFPSPSVATTYSSVPPAFPAQV SSFPSSAVTNSFSASTGLSDMTATFSPRTIEIC	120
24	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSRSDNLTRHIRIHTGQKPFQCRICMRNFSHST	121

構建體	序列	SEQ ID NO
	TLTNHIRTHTGEKPFACDICGRKFARSDNRKTHIRTHT GEKPFACDICGRKFSTSHSLTEHIRIHTGQKPFQCRICM RNFSQSSSLTRHIRTHTGEKPFACDICGRKFARSDNRK THTKIHLRQKDKKADKSVVASSATSSLSSYPSPVATSY PSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYPSPVHSGF PSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGLS DMTATFSPRTIEIC	
25	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLFSG SSDSVVHYNQMATENVMDIGLTNEKPNPELSYSGSFQ PAPGNKTVTYLGKFAFDSPSNWCQDNIISLMSAGILGV PPASGALSTQTSTASMVQPPQGDVEAMYPALPPYSNC GDLYSEPVSFHDPQGNPGLAYSPQDYQSAKPALDSNL FPMIPDYNLYHHPNDMGSIPEHKPFQGMDPIRVNPPPI TPLETIKAFKDKQIHPGFGSLPQPPLTLKPIRPRKYPNR PSKTPLHERPHACPAEGCDRRFSRSDNLVRHLRIHTGH KPFQCRICMRSFSREDNLHTHIRTHTGEKPFACEFCGR KFARSDELVRHAKIHLKQKEHACPAEGCDRRFSQSGN LTEHLRIHTGHKPFQCRICMRSFSTSGHLVRHIRTHTG EKPFACEFCGRKFAQNSTLTEHAKIHLKQKEKKAEKG GAPSASSAPPVSLAPVVTTCA	122
26	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSRSDNLTRHIRIHTGQKPFQCRICMRNFSRS DNLTTHIRTHTGEKPFACDICGRKFARSDERKRHIRTH TGEKPFACDICGRKFSQSGNLTEHIRIHTGQKPFQCRIC MRNFSTSGHLTRHIRTHTGEKPFACDICGRKFAQSSTR KEHTKIHLRQKDKKADKSVVASSATSSLSSYPSPVATS YPSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYPSPVHSG FPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGL SDMTATFSPRTIEIC	123
27	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEE MMLLSNGAPQFLGAAGAPEGSGSNSSSSSGGGGGG GGGSNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNN EKVLVETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLW PEPLFSLVSGLVSMTNPPASSSSAPSPAASSASASQSPP LSCAVPSNDSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSA GTALQYPPPAYPAAKGGFQVPMIPDYLFPQQQGDLGL GTPDQKPFQGLESRTQQPSLTPLSTIKAFATQSGSQDL KALNTSYQSQLIKPSRMRKYPNRPSKTPPHERPYACPV ESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSRE	124

構建體	序列	SEQ ID NO
	DNLHTHIRTHTGEKPFACDICGRKFARSDELVRHIRTH TGEKPFACDICGRKFSQSGNLTEHIRIHTGQKPFQCRIC MRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQNSTL TEHTKIHLRQKDKKADKSVVASSATSSLSSYPSPVATS YPSPVTTSYPSPATTSYPSPVPTSFSSPGSSTYPSPVHSG FPSPSVATTYSSVPPAFPAQVSSFPSSAVTNSFSASTGL SDMTATFSPRTIEIC	
28	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSSP ADLTRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRT HTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKC PECGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQS GNLTEHQRTHTGEKPYKCPECGKSFSTSGHLVRHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	125
29	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSDP GALVRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRT HTGEKPYKCPECGKSFSQSGDLRRHQRTHTGEKPYKC PECGKSFSTHLDLIRHQRTHTGEKPYKCPECGKSFSTS GNLVRHQRTHTGEKPYKCPECGKSFSRSDNLVRHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	126
30	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSHRTTLTNHQRT HTGEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKC PECGKSFSTSHSLTEHQRTHTGEKPYKCPECGKSFSQS SSLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE EASGSGRADALDDFDLDMLGSDALDDFDLDMLGSDA LDDFDLDMLGSDALDDFDLDMLINSRSSGSPKKKRKV GSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPT DPRPPPRRIAVPSRSSASVPKPAPQPYPFTSSLSTINYDE FPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPAMV SALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLS EALLQLQFDDEDLGALLGNSTDPAVFTDLASVDNSEF QQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPP DPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLGSG SGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRP FHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVP QPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADT VIPQKEEAAICGQMDLSHPPPRGHLDELTTTLESMTED LNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDT SLF	99
31	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC	102

構建體	序列	SEQ ID NO
	PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	
32	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	102
33	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE EASGSGRADALDDFDLDMLGSDALDDFDLDMLGSDA LDDFDLDMLGSDALDDFDLDMLINSRSSGSPKKKRKV GSQYLPDTDDRHRIEEKRKRTYETFKSIMKKSPFSGPT DPRPPPRRIAVPSRSSASVPKPAPQPYPFTSSLSTINYDE FPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPAMV SALAQAPAPVPVLAPGPPQAVAPPAPKPTQAGEGTLS EALLQLQFDDEDLGALLGNSTDPAVFTDLASVDNSEF QQLLNQGIPVAPHTTEPMLMEYPEAITRLVTGAQRPP DPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLGSG SGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRP FHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVP QPLDPAPAVTPEASHLLEDPDEETSQAVKALREMADT VIPQKEEAAICGQMDLSHPPPRGHLDELTTTLESMTED LNLDSPLTPELNEILDTFLNDECLLHAMHISTGLSIFDT SLF	100
34	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT HTGKKTSKRPAATKKAGQAKKKKGSDALDDFDLDM LGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFD LDML	127
35	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT	127

構建體	序列	SEQ ID NO
	HTGKKTSKRPAATKKAGQAKKKKGSDALDDFDLDM LGSDALDDFDLDMLGSDALDDFDLDMLGSDALDDFD LDML	
36	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCQSQ LIKPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSRS DNLVRHIRIHTGQKPFQCRICMRNFSREDNLHTHIRTH TGEKPFACDICGRKFARSDELVRHTKIHLRQKDRPYA CPVESCDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNF STSGHLVRHIRTHTGEKPFACDICGRKFAQNSTLTEHT KIHLRQKDK	128
37	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDN LHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQC RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQN STLTEHTKIHLRQKDK	106
38	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRIC MRNFSREDNLHTHIRTHTGEKPFACDICGRKFARSDEL VRHTKIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRI HTGQKPFQCRICMRNFSTSGHLVRHIRTHTGEKPFAC DICGRKFAQNSTLTEHTKIHLRQKDK	129
39	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDN LHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQC RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQN STLTEHTKIHLRQKDKLEMADHLMLAEGYRLVQRPPS	105

構建體	序列	SEQ ID NO
	AAAAHGPHALRTLPPYAGPGLDSGLRPRGAPLGPPPP RQPGALAYGAFGPPSSFQPFPAVPPPAAGIAHLQPVAT PYPGRAAAPPNAPGGPPGPQPAPSAAAPPPPAHALGG MDAELIDEEALTSLELELGLHRVRELPELFLGQSEFDC FSDLGSAPPAGSVSC	
40	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCADH LMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGL DSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPA VPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQP APSAAAPPPPAHALGGMDAELIDEEALTSLELELGLHR VRELPELFLGQSEFDCFSDLGSAPPAGSVSCQSQLIKPS RMRKYPNRPSKTPPHERPYACPVESCDRRFSRSDNLV RHIRIHTGQKPFQCRICMRNFSREDNLHTHIRTHTGEK PFACDICGRKFARSDELVRHTKIHLRQKDRPYACPVES CDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHLR QKDK	130
41	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGADHLMLAEGYRLVQRPPSAAAAHGPHALRTL PPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPG GPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLE LELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC GGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYAC PVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFS REDNLHTHIRTHTGEKPFACDICGRKFARSDELVRHT KIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQ KPFQCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGR KFAQNSTLTEHTKIHLRQKDK	131
42	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	102
43	MAPKKKRKVGIHGVPAALEPGEKPYKCPECGKSFSRS DNLVRHQRTHTGEKPYKCPECGKSFSREDNLHTHQRT	102

構建體	序列	SEQ ID NO
	HTGEKPYKCPECGKSFSRSDELVRHQRTHTGEKPYKC PECGKSFSQSGNLTEHQRTHTGEKPYKCPECGKSFSTS GHLVRHQRTHTGEKPYKCPECGKSFSQNSTLTEHQRT HTGKKTSKRPAATKKAGQAKKKKGSYPYDVPDYALE DALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLD MLGSDALDDFDLDML	
44	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDN LHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQC RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQN STLTEHTKIHLRQKDK	106
45	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYACPVES CDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFSREDN LHTHIRTHTGEKPFACDICGRKFARSDELVRHTKIHLR QKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQKPFQC RICMRNFSTSGHLVRHIRTHTGEKPFACDICGRKFAQN STLTEHTKIHLRQKDK	106
46	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGGGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHE RPYACPVESCDRRFSRSDNLVRHIRIHTGQKPFQCRIC MRNFSREDNLHTHIRTHTGEKPFACDICGRKFARSDEL VRHTKIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRI HTGQKPFQCRICMRNFSTSGHLVRHIRTHTGEKPFAC DICGRKFAQNSTLTEHTKIHLRQKDK	205
47	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCADH LMLAEGYRLVQRPPSAAAAHGPHALRTLPPYAGPGL DSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSFQPFPA VPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPPGPQP	207

構建體	序列	SEQ ID NO
	APSAAAPPPPAHALGGMDAELIDEEALTSLELELGLHR VRELPELFLGQSEFDCFSDLGSAPPAGSVSCQSQLIKPS RMRKYPNRPSKTPPHERPYACPVESCDRRFSRSDNLV RHIRIHTGQKPFQCRICMRNFSREDNLHTHIRTHTGEK PFACDICGRKFARSDELVRHTKIHLRQKDRPYACPVES CDRRFSQSGNLTEHIRIHTGQKPFQCRICMRNFSTSGH LVRHIRTHTGEKPFACDICGRKFAQNSTLTEHTKIHLR QKDK	
48	MAADHLMLAEGYRLVQRPPSAAAAHGPHALRTLPPY AGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPPSSF QPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPGGPP GPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLELEL GLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSCGGS GGGSGADHLMLAEGYRLVQRPPSAAAAHGPHALRTL PPYAGPGLDSGLRPRGAPLGPPPPRQPGALAYGAFGPP SSFQPFPAVPPPAAGIAHLQPVATPYPGRAAAPPNAPG GPPGPQPAPSAAAPPPPAHALGGMDAELIDEEALTSLE LELGLHRVRELPELFLGQSEFDCFSDLGSAPPAGSVSC GGSGGGSGQSQLIKPSRMRKYPNRPSKTPPHERPYAC PVESCDRRFSRSDNLVRHIRIHTGQKPFQCRICMRNFS REDNLHTHIRTHTGEKPFACDICGRKFARSDELVRHT KIHLRQKDRPYACPVESCDRRFSQSGNLTEHIRIHTGQ KPFQCRICMRNFSTSGHLVRHIRTHTGEKPFACDICGR KFAQNSTLTEHTKIHLRQKDK	209
49	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLD WALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCL VHHDHTYSLPRETVSMDLESESCRKEGTQMTPQHME ELAEQEIARLVLTDEEKSLLEKEGLILPETLPLTKTEEQ ILKRVRLEPGEKPYKCPECGKSFSRSDNLVRHQRTHT GEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPE CGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSG NLTEHQRTHTGEKPYKCPECGKSFSTSGHLVRHQRTH TGEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTSVYV GGLESRVLKYTAQNMELQNKVQLLEEQNLSLLDQLR KLQAMVIEISNKTSSSSTCILVLLVSFCLLLVPAMYSSD TRGSLPAEHGVLSRQLRALPSEDPYQLELPALQSEVPK DSTHQWLDGSDCVLQAPGNTSCLLHYMPQAPSAEPP LEWPFPDLFSEPLCRGPILPLQANLTRKGGWLPTGSPS VILQDRYSG	213
50	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLD WALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCL VHHDHTYSLPRETVSMDLESESCRKEGTQMTPQHME ELAEQEIARLVLTDEEKSLLEKEGLILPETLPLTKTEEQ ILKRVRRPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSREDNLHTHIRTHTGEKPFACDICGRKF ARSDELVRHTKIHLRQKDRPYACPVESCDRRFSQSGN LTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE	217

構建體	序列	SEQ ID NO
	KPFACDICGRKFAQNSTLTEHTKIHLRQKDVYVGGLE SRVLKYTAQNMELQNKVQLLEEQNLSLLDQLRKLQA MVIEISNKTSSSSTCILVLLVSFCLLLVPAMYSSDTRGS LPAEHGVLSRQLRALPSEDPYQLELPALQSEVPKDSTH QWLDGSDCVLQAPGNTSCLLHYMPQAPSAEPPLEWP FPDLFSEPLCRGPILPLQANLTRKGGWLPTGSPSVILQD RYSG	
51	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLD WALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCL VHHDHTYSLPRETVSMDLESESCRKEGTQMTPQHME ELAEQEIARLVLTDEEKSLLEKEGLILPETLPLTKTEEQ ILKRVRRPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSHRTTLTNHIRTHTGEKPFACDICGRKF ARSDNLHTHTKIHLRQKDRPYACPVESCDRRFSTSHS LTEHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGE KPFACDICGRKFAREDNLHTHTKIHLRQKDVYVGGLE SRVLKYTAQNMELQNKVQLLEEQNLSLLDQLRKLQA MVIEISNKTSSSSTCILVLLVSFCLLLVPAMYSSDTRGS LPAEHGVLSRQLRALPSEDPYQLELPALQSEVPKDSTH QWLDGSDCVLQAPGNTSCLLHYMPQAPSAEPPLEWP FPDLFSEPLCRGPILPLQANLTRKGGWLPTGSPSVILQD RYSG	219
52	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLD WALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCL VHHDHTYSLPRETVSMDLESESCRKEGTQMTPQHME ELAEQEIARLVLTDEEKSLLEKEGLILPETLPLTKTEEQ ILKRVRLEPGEKPYKCPECGKSFSRSDNLVRHQRTHT GEKPYKCPECGKSFSREDNLHTHQRTHTGEKPYKCPE CGKSFSRSDELVRHQRTHTGEKPYKCPECGKSFSQSG NLTEHQRTHTGEKPYKCPECGKSFSTSGHLVRHQRTH TGEKPYKCPECGKSFSQNSTLTEHQRTHTGKKTSVYV GGLESRVLKYTAQNMELQNKVQLLEEQNLSLLDQLR KLQAMVIEIS	221
53	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLD WALPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCL VHHDHTYSLPRETVSMDLESESCRKEGTQMTPQHME ELAEQEIARLVLTDEEKSLLEKEGLILPETLPLTKTEEQ ILKRVRRPYACPVESCDRRFSRSDNLVRHIRIHTGQKP FQCRICMRNFSREDNLHTHIRTHTGEKPFACDICGRKF ARSDELVRHTKIHLRQKDRPYACPVESCDRRFSQSGN LTEHIRIHTGQKPFQCRICMRNFSTSGHLVRHIRTHTGE KPFACDICGRKFAQNSTLTEHTKIHLRQKDVYVGGLE SRVLKYTAQNMELQNKVQLLEEQNLSLLDQLRKLQA MVIEIS	223

表7. 編碼本文公開的示例性工程化轉錄因子的核酸序列。

構建體	序列	SEQ ID NO
31 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttttaaaataaa tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttttccccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtgcgccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgccccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag ctgGAATTCGCCACCATGGCCCCAAAGAAGAAGCGG AAGGTCGGTATCCACGGAGTCCCAGCAGCCCTCGA ACCAGGTGAAAAACCTTACAAATGTCCTGAATGTG GGAAATCATTCAGTCGCAGCGACAACCTGGTGAGA CATCAACGCACCCATACAGGAGAAAAACCTTATAA ATGTCCAGAATGTGGAAAGTCCTTCTCACGAGAGG ATAACTTGCACACTCATCAACGAACACATACTGGT GAAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCAA CGAACCCATACAGGCGAGAAGCCTTACAAATGTCC CGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATC TGACTGAGCATCAACGAACTCATACCGGGGAAAA ACCTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTT	67

構建體	序列	SEQ ID NO
	CCACAAGTGGACATCTGGTACGCCACCAGAGGACA CATACAGGGGAGAAGCCCTACAAATGCCCCGAAT GCGGTAAAAGTTTCTCTCAGAATAGTACCCTGACC GAACACCAGCGAACACACACTGGGAAAAAAACGA GTAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCA GGCAAAAAAGAAAAAGGGATCCTACCCATACGAC GTACCAGATTACGCTCTCGAGGACGCGCTGGACGA TTTCGATCTCGACATGCTGGGTTCTGATGCCCTCGA TGACTTTGACCTGGATATGTTGGGAAGCGACGCAT TGGATGACTTTGATCTGGACATGCTCGGCTCCGAT GCTCTGGACGATTTCGATCTCGATATGTTATAAACT AGTaaagagaccggttcactgtgacagtaaaagagaccggttcactgtgagaat gaaagagaccggttcactgtgatcggaaaagagaccggttcactgtgagcggcctt gaaacccagcagacaatgtagctcagtagaaacccagcagacaatgtagctgaat ggaaacccagcagacaatgtagcttcggagaaacccagcagacaatgtagctAA GCTTGGGTGGCATCCCTGTGACCCCTCCCCAGTGC CTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCC ACCAGCCTTGTCCTAATAAAATTAAGTTGCATCATT TTGTCTGACTAGGTGTCCTTCTATAATATTATGGGG TGGAGGGGGGTGGTATGGAGCAAGGGGCAAGTTG GGAAGACAACCTGTAGGGCCTGCGGGGTCTATTGG GAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCT CACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCT CCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGG CATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTT GGTAGAGACGGGGTTTCACCATATTGGCCAGGCTG GTCTCCAACTCCTAATCTCAGGTGATCTACCCACCT TGGCCTCCCAAATTGCTGGGATTACAGGCGTGAAC CACTGCTCCCTTCCCTGTCCTT	
32 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttttaaaataaa tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg	68

構建體	序列	SEQ ID NO
	cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttttcccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtgcgccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgccccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag ctgGAATTCGCCACCATGGCCCCAAAGAAGAAGCGG AAGGTCGGTATCCACGGAGTCCCAGCAGCCCTCGA ACCAGGTGAAAAACCTTACAAATGTCCTGAATGTG GGAAATCATTCAGTCGCAGCGACAACCTGGTGAGA CATCAACGCACCCATACAGGAGAAAAACCTTATAA ATGTCCAGAATGTGGAAAGTCCTTCTCACGAGAGG ATAACTTGCACACTCATCAACGAACACATACTGGT GAAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCAA CGAACCCATACAGGCGAGAAGCCTTACAAATGTCC CGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATC TGACTGAGCATCAACGAACTCATACCGGGGAAAA ACCTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTT CCACAAGTGGACATCTGGTACGCCACCAGAGGACA CATACAGGGGAGAAGCCCTACAAATGCCCCGAAT GCGGTAAAAGTTTCTCTCAGAATAGTACCCTGACC GAACACCAGCGAACACACACTGGGAAAAAAACGA GTAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCA GGCAAAAAAGAAAAAGGGATCCTACCCATACGAC GTACCAGATTACGCTCTCGAGGACGCGCTGGACGA TTTCGATCTCGACATGCTGGGTTCTGATGCCCTCGA TGACTTTGACCTGGATATGTTGGGAAGCGACGCAT TGGATGACTTTGATCTGGACATGCTCGGCTCCGAT GCTCTGGACGATTTCGATCTCGATATGTTATAAACT AGTAAGCTTGGGTGGCATCCCTGTGACCCCTCCCC AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAG TGCCCACCAGCCTTGTCCTAATAAAATTAAGTTGC ATCATTTTGTCTGACTAGGTGTCCTTCTATAATATT ATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGGC AAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTC TATTGGGAACCAAGCTGGAGTGCAGTGGCACAATC TTGGCTCACTGCAATCTCCGCCTCCTGGGTTCAAGC	

構建體	序列	SEQ ID NO
	GATTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATT CCAGGCATGCATGACCAGGCTCAGCTAATTTTTGT TTTTTTGGTAGAGACGGGGTTTCACCATATTGGCCA GGCTGGTCTCCAACTCCTAATCTCAGGTGATCTACC CACCTTGGCCTCCCAAATTGCTGGGATTACAGGCG TGAACCACTGCTCCCTTCCCTGTCCTT	
33 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttaaaataaa tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttcccccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtgcgccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgccccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag ctgGAATTCGCCACCATGGCCCCAAAGAAGAAGCGG AAGGTCGGTATCCACGGAGTCCCAGCAGCCCTCGA ACCAGGTGAAAAACCTTACAAATGTCCTGAATGTG GGAAATCATTCAGTCGCAGCGACAACCTGGTGAGA CATCAACGCACCCATACAGGAGAAAAACCTTATAA ATGTCCAGAATGTGGAAAGTCCTTCTCACGAGAGG	69

構建體	序列	SEQ ID NO
	ATAACTTGCACACTCATCAACGAACACATACTGGT GAAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCAA CGAACCCATACAGGCGAGAAGCCTTACAAATGTCC CGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATC TGACTGAGCATCAACGAACTCATACCGGGGAAAA ACCTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTT CCACAAGTGGACATCTGGTACGCCACCAGAGGACA CATACAGGGGAGAAGCCCTACAAATGCCCCGAAT GCGGTAAAAGTTTCTCTCAGAATAGTACCCTGACC GAACACCAGCGAACACACACTGGGAAAAAAACGA GTAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCA GGCAAAAAAGAAAAAGGGATCCTACCCATACGAC GTACCAGATTACGCTCTCGAGGAGGCCAGCGGTTC CGGACGGGCTGACGCATTGGACGATTTTGATCTGG ATATGCTGGGAAGTGACGCCCTCGATGATTTTGAC CTTGACATGCTTGGTTCGGATGCCCTTGATGACTTT GACCTCGACATGCTCGGCAGTGACGCCCTTGATGA TTTCGACCTGGACATGCTGATTAACTCTAGAAGTTC CGGATCTCCGAAAAAGAAACGCAAAGTTGGTAGC CAGTACCTGCCCGACACCGACGACCGGCACCGGAT CGAGGAAAAGCGGAAGCGGACCTACGAGACATTC AAGAGCATCATGAAGAAGTCCCCCTTCAGCGGCCC CACCGACCCTAGACCTCCACCTAGAAGAATCGCCG TGCCCAGCAGATCCAGCGCCAGCGTGCCAAAACCT GCCCCCCAGCCTTACCCCTTCACCAGCAGCCTGAG CACCATCAACTACGACGAGTTCCCTACCATGGTGT TCCCCAGCGGCCAGATCTCTCAGGCCTCTGCTCTG GCTCCAGCCCCTCCTCAGGTGCTGCCTCAGGCTCCT GCTCCTGCACCAGCTCCAGCCATGGTGTCTGCACT GGCTCAGGCACCAGCACCCGTGCCTGTGCTGGCTC CTGGACCTCCACAGGCTGTGGCTCCACCAGCCCCT AAACCTACACAGGCCGGCGAGGGCACACTGTCTGA AGCTCTGCTGCAGCTGCAGTTCGACGACGAGGATC TGGGAGCCCTGCTGGGAAACAGCACCGATCCTGCC GTGTTCACCGACCTGGCCAGCGTGGACAACAGCGA GTTCCAGCAGCTGCTGAACCAGGGCATCCCTGTGG CCCCTCACACCACCGAGCCCATGCTGATGGAATAC CCCGAGGCCATCACCCGGCTCGTGACAGGCGCTCA GAGGCCTCCTGATCCAGCTCCTGCCCCTCTGGGAG CACCAGGCCTGCCTAATGGACTGCTGTCTGGCGAC GAGGACTTCAGCTCTATCGCCGATATGGATTTCTC AGCCTTGCTGGGCTCTGGCAGCGGCAGCCGGGATT CCAGGGAAGGGATGTTTTTGCCGAAGCCTGAGGCC GGCTCCGCTATTAGTGACGTGTTTGAGGGCCGCGA GGTGTGCCAGCCAAAACGAATCCGGCCATTTCATC CTCCAGGAAGTCCATGGGCCAACCGCCCACTCCCC GCCAGCCTCGCACCAACACCAACCGGTCCAGTACA TGAGCCAGTCGGGTCACTGACCCCGGCACCAGTCC	

構建體	序列	SEQ ID NO
	CTCAGCCACTGGATCCAGCGCCCGCAGTGACTCCC GAGGCCAGTCACCTGTTGGAGGATCCCGATGAAGA GACGAGCCAGGCTGTCAAAGCCCTTCGGGAGATGG CCGATACTGTGATTCCCCAGAAGGAAGAGGCTGCA ATCTGTGGCCAAATGGACCTTTCCCATCCGCCCCC AAGGGGCCATCTGGATGAGCTGACAACCACACTTG AGTCCATGACCGAGGATCTGAACCTGGACTCACCC CTGACCCCGGAATTGAACGAGATTCTGGATACCTT CCTGAACGACGAGTGCCTCTTGCATGCCATGCATA TCAGCACAGGACTGTCCATCTTCGACACATCTCTGT TTTAAACTAGTaataaaagatctttattttcattagatctgtgtgttggtttttgt gtg	
34 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttaaaataaa tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttcccccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtgcgccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgcccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag	70

構建體	序列	SEQ ID NO
	ctgGAATTCGCCACCATGGCCCCAAAGAAGAAGCGG AAGGTCGGTATCCACGGAGTCCCAGCAGCCCTCGA ACCAGGTGAAAAACCTTACAAATGTCCTGAATGTG GGAAATCATTCAGTCGCAGCGACAACCTGGTGAGA CATCAACGCACCCATACAGGAGAAAAACCTTATAA ATGTCCAGAATGTGGAAAGTCCTTCTCACGAGAGG ATAACTTGCACACTCATCAACGAACACATACTGGT GAAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCAA CGAACCCATACAGGCGAGAAGCCTTACAAATGTCC CGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATC TGACTGAGCATCAACGAACTCATACCGGGGAAAA ACCTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTT CCACAAGTGGACATCTGGTACGCCACCAGAGGACA CATACAGGGGAGAAGCCCTACAAATGCCCCGAAT GCGGTAAAAGTTTCTCTCAGAATAGTACCCTGACC GAACACCAGCGAACACACACTGGGAAAAAAACGA GTAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCA GGCAAAAAAGAAAAAGGGATCCGACGCGCTGGAC GATTTCGATCTCGACATGCTGGGTTCTGATGCCCTC GATGACTTTGACCTGGATATGTTGGGAAGCGACGC ATTGGATGACTTTGATCTGGACATGCTCGGCTCCG ATGCTCTGGACGATTTCGATCTCGATATGTTATAAA AGCTTGGGTGGCATCCCTGTGACCCCTCCCCAGTG CCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCC ACCAGCCTTGTCCTAATAAAATTAAGTTGCATCATT TTGTCTGACTAGGTGTCCTTCTATAATATTATGGGG TGGAGGGGGGTGGTATGGAGCAAGGGGCAAGTTG GGAAGACAACCTGTAGGGCCTGCGGGGTCTATTGG GAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCT CACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCT CCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGG CATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTT GGTAGAGACGGGGTTTCACCATATTGGCCAGGCTG GTCTCCAACTCCTAATCTCAGGTGATCTACCCACCT TGGCCTCCCAAATTGCTGGGATTACAGGCGTGAAC CACTGCTCCCTTCCCTGTCCTT	
26 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttaaaataaa	71

構建體	序列	SEQ ID NO
	tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttttccccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtgcgccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgcccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag ctgGAATTCGCCACCATGGCCCCAAAGAAGAAGCGG AAGGTCGGTATCCACGGAGTCCCAGCAGCCCTCGA ACCAGGTGAAAAACCTTACAAATGTCCTGAATGTG GGAAATCATTCAGTCGCAGCGACAACCTGGTGAGA CATCAACGCACCCATACAGGAGAAAAACCTTATAA ATGTCCAGAATGTGGAAAGTCCTTCTCACGAGAGG ATAACTTGCACACTCATCAACGAACACATACTGGT GAAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCAA CGAACCCATACAGGCGAGAAGCCTTACAAATGTCC CGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATC TGACTGAGCATCAACGAACTCATACCGGGGAAAA ACCTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTT CCACAAGTGGACATCTGGTACGCCACCAGAGGACA CATACAGGGGAGAAGCCCTACAAATGCCCCGAAT GCGGTAAAAGTTTCTCTCAGAATAGTACCCTGACC GAACACCAGCGAACACACACTGGGAAAAAAACGA GTAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCA GGCAAAAAAGAAAAAGGGATCCGACGCGCTGGAC GATTTCGATCTCGACATGCTGGGTTCTGATGCCCTC GATGACTTTGACCTGGATATGTTGGGAAGCGACGC ATTGGATGACTTTGATCTGGACATGCTCGGCTCCG ATGCTCTGGACGATTTCGATCTCGATATGTTATAAa aagagaccggttcactgtgacagtaaaagagaccggttcactgtgagaatgaaaga	

構建體	序列	SEQ ID NO
	gaccggttcactgtgatcggaaaagagaccggttcactgtgagcggccttgaaacc cagcagacaatgtagctcagtagaaacccagcagacaatgtagctgaatggaaac ccagcagacaatgtagcttcggagaaacccagcagacaatgtagctAAGCTT GGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTC CTGGCCCTGGAAGTTGCCACTCCAGTGCCCACCAG CCTTGTCCTAATAAAATTAAGTTGCATCATTTTGTC TGACTAGGTGTCCTTCTATAATATTATGGGGTGGA GGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGAA GACAACCTGTAGGGCCTGCGGGGTCTATTGGGAAC CAAGCTGGAGTGCAGTGGCACAATCTTGGCTCACT GCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTG CCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCATG CATGACCAGGCTCAGCTAATTTTTGTTTTTTTGGTA GAGACGGGGTTTCACCATATTGGCCAGGCTGGTCT CCAACTCCTAATCTCAGGTGATCTACCCACCTTGGC CTCCCAAATTGCTGGGATTACAGGCGTGAACCACT GCTCCCTTCCCTGTCCTT	
40 (編碼)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCTA CAGACTGGTGCAGCGGCCTCCATCTGCCGCTGCCG CCCACGGCCCCCACGCCCTGAGAACACTGCCCCCC TACGCCGGCCCTGGTCTTGATAGCGGACTCAGACC TAGAGGCGCCCCTCTGGGCCCTCCACCTCCAAGAC AGCCTGGAGCCCTGGCCTACGGCGCCTTCGGCCCT CCTTCTAGCTTCCAGCCCTTCCCCGCCGTGCCTCCT CCAGCcGCTGGCATCGCCCACCTGCAGCCTGTGGC CACCCCTTACCCCGGAAGAGCCGCCGCCCCTCCAA ACGCCCCTGGCGGACCTCCTGGCCCCCAGCCTGCT CCAAGCGCCGCTGCCCCTCCACCTCCTGCTCATGC CCTGGGCGGCATGGACGCCGAGCTGATCGACGAG GAAGCCCTGACCAGCCTGGAACTGGAACTGGGCCT GCACAGAGTGCGGGAACTGCCTGAGCTGTTCCTGG GACAGAGCGAGTTCGACTGCTTCAGCGACCTGGGC AGCGCCCCTCCTGCCGGCTCTGTGTCCTGCgccgaccac ctgatgctcgccgagggctaccgcctggtgcagaggccgccgtccgccgccgcc gcccatggccctcatgcgctccggactctgccgccgtacgcgggcccgggcctgg acagtgggctgaggccgcgggggctccgctggggccgccgccgccccgccaa cccggggccctggcgtacggggccttcgggccgccgtcctccttccagccctttcc ggccgtgcctccgccggccgcgggcatcgcgcacctgcagcctgtggcgacgcc gtaccccggccgcgcCgccgcgcccccccaacgctccgggaggcccccccgggc ccgcagccggccccaagcgccgcagccccgccgccgcccgcgcacgccctgg gcggcatggacgccgaactcatcgacgaggaggcgctgacgtcgctggagctgg agctggggctgcaccgcgtgcgcgagctgcccgagctgttcctgggccagagcg agttcgactgcttctcggacttggggtccgcgccgcccgccggctccgtgagctgc cagtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccggcccagca agacgccccccccacgaacgcccttacgcttgcccagtggagtcctgtgatcgccgc ttctccCGCAGCGACAACCTGGTGAGAcacatccgcatccacac aggccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCGAGA GGATAACTTGCACACTcacatccgcacccacacaggcgaaaagcc	72

構建體	序列	SEQ ID NO
	cttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGATGAACT TGTCCGAcataccaagatccacttgcggcagaaggaccgcccttacgcttgc ccagtggagtcctgtgatcgccgcttctccCAATCAGGGAATCTGAC TGAGcacatccgcatccacacaggccagaagcccttccagtgccgcatctgcat gAGAaacttcagcACAAGTGGACATCTGGTACGCcacatcc gcacccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgcc CAGAATAGTACCCTGACCGAAcataccaagatccacttgcggc agaaggacaag	
41 (編碼)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCTA CAGACTGGTGCAGCGGCCTCCATCTGCCGCTGCCG CCCACGGCCCCCACGCCCTGAGAACACTGCCCCCC TACGCCGGCCCTGGTCTTGATAGCGGACTCAGACC TAGAGGCGCCCCTCTGGGCCCTCCACCTCCAAGAC AGCCTGGAGCCCTGGCCTACGGCGCCTTCGGCCCT CCTTCTAGCTTCCAGCCCTTCCCCGCCGTGCCTCCT CCAGCTGCTGGCATCGCCCACCTGCAGCCTGTGGC CACCCCTTACCCCGGAAGAGCCGCCGCCCCTCCAA ACGCCCCTGGCGGACCTCCTGGCCCCCAGCCTGCT CCAAGCGCCGCTGCCCCTCCACCTCCTGCTCATGC CCTGGGCGGCATGGACGCCGAGCTGATCGACGAG GAAGCCCTGACCAGCCTGGAACTGGAACTGGGCCT GCACAGAGTGCGGGAACTGCCTGAGCTGTTCCTGG GACAGAGCGAGTTCGACTGCTTCAGCGACCTGGGC AGCGCCCCTCCTGCCGGCTCTGTGTCCTGCGGCGG CAGCGGCGGCGGAAGCGGCgccgaccacctgatgctcgccgag ggctaccgcctggtgcagaggccgccgtccgccgccgccgcccatggccctcat gcgctccggactctgccgccgtacgcgggcccgggcctggacagtgggctgagg ccgcgggggctccgctggggccgccgccccgccaacccggggccctgg cgtacgggcctcgggccgccgtcctccttccagccctttccggccgtgcctccgc cggccgcgggcatcgcgcacctgcagcctgtggcgacgccgtaccccggccgc gcggccgcgccccccaacgctccgggaggcccccgggcccgcagccggccc caagcgccgcagccccgccgccgcccgcgcacgccctgggcggcatggacgcc gaactcatcgacgaggaggcgctgacgtcgctggagctggagctggggctgcac cgcgtgcgcgagctgcccgagctgttcctgggccagagcgagttcgactgcttctc ggacttggggtccgcgccgcccgccggctccgtgagctgcggtggttctggtggt ggttctggtcagtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccg gcccagcaagacgccccccccacgaacgcccttacgcttgcccagtggagtcctgt gatcgccgcttctccCGCAGCGACAACCTGGTGAGAcacatccg catccacacaggccagaagcccttccagtgccgcatctgcatgAGAaacttcag cCGAGAGGATAACTTGCACACTcacatccgcacccacacaggc gaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGAT GAACTTGTCCGAcataccaagatccacttgcggcagaaggaccgcccctt acgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAGGGAAT CTGACTGAGcacatccgcatccacacaggccagaagcccttccagtgccg catctgcatgAGAaacttcagcACAAGTGGACATCTGGTACG Ccacatccgcacccacacaggcgaaaagcccttcgcctgcgacatctgtggaaga aagtttgccCAGAATAGTACCCTGACCGAAcataccaagatcca cttgcggcagaaggacaag	73

構建體	序列	SEQ ID NO
42 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttttaaaataaa tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttttccccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtcgcccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgcccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag ctgGAATTCGCCACCATGGCCCCAAAGAAGAAGCGG AAGGTCGGTATCCACGGAGTCCCAGCAGCCCTCGA ACCAGGTGAAAAACCTTACAAATGTCCTGAATGTG GGAAATCATTCAGTCGCAGCGACAACCTGGTGAGA CATCAACGCACCCATACAGGAGAAAAACCTTATAA ATGTCCAGAATGTGGAAAGTCCTTCTCACGAGAGG ATAACTTGCACACTCATCAACGAACACATACTGGT GAAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCAA CGAACCCATACAGGCGAGAAGCCTTACAAATGTCC CGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATC TGACTGAGCATCAACGAACTCATACCGGGGAAAA ACCTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTT	74

構建體	序列	SEQ ID NO
	CCACAAGTGGACATCTGGTACGCCACCAGAGGACA CATACAGGGGAGAAGCCCTACAAATGCCCCGAAT GCGGTAAAAGTTTCTCTCAGAATAGTACCCTGACC GAACACCAGCGAACACACACTGGGAAAAAAACGA GTAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCA GGCAAAAAAGAAAAAGGGATCCTACCCATACGAC GTACCAGATTACGCTCTCGAGGACGCGCTGGACGA TTTCGATCTCGACATGCTGGGTTCTGATGCCCTCGA TGACTTTGACCTGGATATGTTGGGAAGCGACGCAT TGGATGACTTTGATCTGGACATGCTCGGCTCCGAT GCTCTGGACGATTTCGATCTCGATATGTTATAAACT AGTGAAACCCAGCAGACAATGTAGCTAGACCCAGT AGCCAGATGTAGCTAAAGAGACCGGTTCACTGTGA AAGCTTGGGTGGCATCCCTGTGACCCCTCCCCAGT GCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGC CCACCAGCCTTGTCCTAATAAAATTAAGTTGCATC ATTTTGTCTGACTAGGTGTCCTTCTATAATATTATG GGGTGGAGGGGGGTGGTATGGAGCAAGGGGCAAG TTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTAT TGGGAACCAAGCTGGAGTGCAGTGGCACAATCTTG GCTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGA TTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATTCC AGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTT TTTGGTAGAGACGGGGTTTCACCATATTGGCCAGG CTGGTCTCCAACTCCTAATCTCAGGTGATCTACCCA CCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTG AACCACTGCTCCCTTCCCTGTCCTT	
43 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttttaaaataaa tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc	75

構建體	序列	SEQ ID NO
	cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttttcccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtgcgccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgcccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag ctgGAATTCGCCACCATGGCCCCAAAGAAGAAGCGG AAGGTCGGTATCCACGGAGTCCCAGCAGCCCTCGA ACCAGGTGAAAAACCTTACAAATGTCCTGAATGTG GGAAATCATTCAGTCGCAGCGACAACCTGGTGAGA CATCAACGCACCCATACAGGAGAAAAACCTTATAA ATGTCCAGAATGTGGAAAGTCCTTCTCACGAGAGG ATAACTTGCACACTCATCAACGAACACATACTGGT GAAAAACCATACAAGTGTCCCGAATGTGGTAAAA GTTTTAGCCGGAGCGATGAACTTGTCCGACACCAA CGAACCCATACAGGCGAGAAGCCTTACAAATGTCC CGAGTGTGGCAAGAGCTTCTCACAATCAGGGAATC TGACTGAGCATCAACGAACTCATACCGGGGAAAA ACCTTACAAGTGTCCAGAGTGTGGGAAGAGCTTTT CCACAAGTGGACATCTGGTACGCCACCAGAGGACA CATACAGGGGAGAAGCCCTACAAATGCCCCGAAT GCGGTAAAAGTTTCTCTCAGAATAGTACCCTGACC GAACACCAGCGAACACACACTGGGAAAAAAACGA GTAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCA GGCAAAAAAGAAAAAGGGATCCTACCCATACGAC GTACCAGATTACGCTCTCGAGGACGCGCTGGACGA TTTCGATCTCGACATGCTGGGTTCTGATGCCCTCGA TGACTTTGACCTGGATATGTTGGGAAGCGACGCAT TGGATGACTTTGATCTGGACATGCTCGGCTCCGAT GCTCTGGACGATTTCGATCTCGATATGTTATAAACT AGTGAAACCCAGCAGACAATGTAGCTAGACCCAGT AGCCAGATGTAGCTAAAGAGACCGGTTCACTGTGA GAAACCCAGCAGACAATGTAGCTAGACCCAGTAG CCAGATGTAGCTAAAGAGACCGGTTCACTGTGAAA GCTTGGGTGGCATCCCTGTGACCCCTCCCCAGTGC CTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCC ACCAGCCTTGTCCTAATAAAATTAAGTTGCATCATT TTGTCTGACTAGGTGTCCTTCTATAATATTATGGGG TGGAGGGGGGTGGTATGGAGCAAGGGGCAAGTTG GGAAGACAACCTGTAGGGCCTGCGGGGTCTATTGG	

構建體	序列	SEQ ID NO
	GAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCT CACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCT CCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGG CATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTT GGTAGAGACGGGGTTTCACCATATTGGCCAGGCTG GTCTCCAACTCCTAATCTCAGGTGATCTACCCACCT TGGCCTCCCAAATTGCTGGGATTACAGGCGTGAAC CACTGCTCCCTTCCCTGTCCTT	
44 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacgggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttttaaaataaa tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttttcccccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtgcgccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgcccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag ctgGAATTCGCCACCATGGCCgccgaccacctgatgctcgccgag ggctaccgcctggtgcagaggccgccgtccgccgccgccgcccatggccctcat gcgctccggactctgccgccgtacgcgggcccgggcctggacagtgggctgagg ccgcgggggctccgctggggccgccgccgccccgccaacccggggccctgg	76

構建體	序列	SEQ ID NO
	cgtacggggccttcgggccgccgtcctccttccagccctttccggccgtgcctccgc cggccgcgggcatcgcgcacctgcagcctgtggcgacgccgtaccccggccgc gcggccgcgccccccaacgctccgggaggcccccccgggcccgcagccggccc caagcgccgcagccccgccgccgcccgcgcacgccctgggcggcatggacgcc gaactcatcgacgaggaggcgctgacgtcgctggagctggagctggggctgcac cgcgtgcgcgagctgcccgagctgttcctgggccagagcgagttcgactgcttctc ggacttggggtccgcgccgcccgccggctccgtgagctgcggtggttctggtggt ggttctggtcagtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccg gcccagcaagacgcccccccacgaacgcccttacgcttgcccagtggagtcctgt gatcgccgcttctccCGCAGCGACAACCTGGTGAGAcacatccg catccacacaggccagaagcccttccagtgccgcatctgcatgAGAaacttcag cCGAGAGGATAACTTGCACACTcacatccgcacccacacaggc gaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGAT GAACTTGTCCGAcataccaagatccacttgcggcagaaggaccgcccctt acgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAGGGAAT CTGACTGAGcacatccgcatccacacaggccagaagcccttccagtgccg catctgcatgAGAaacttcagcACAAGTGGACATCTGGTACG Ccacatccgcacccacacaggcgaaaagcccttcgcctgcgacatctgtggaaga aagtttgccCAGAATAGTACCCTGACCGAAcataccaagatcca cttgcggcagaaggacaagtaaCTCGAGGAAACCCAGCAGAC AATGTAGCTAGACCCAGTAGCCAGATGTAGCTAAA GAGACCGGTTCACTGTGAAAGCTTGGGTGGCATCC CTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTGGA AGTTGCCACTCCAGTGCCCACCAGCCTTGTCCTAAT AAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTC CTTCTATAATATTATGGGGTGGAGGGGGGTGGTAT GGAGCAAGGGGCAAGTTGGGAAGACAACCTGTAG GGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTG CAGTGGCACAATCTTGGCTCACTGCAATCTCCGCC TCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC GAGTTGTTGGGATTCCAGGCATGCATGACCAGGCT CAGCTAATTTTTGTTTTTTTGGTAGAGACGGGGTTT CACCATATTGGCCAGGCTGGTCTCCAACTCCTAAT CTCAGGTGATCTACCCACCTTGGCCTCCCAAATTGC TGGGATTACAGGCGTGAACCACTGCTCCCTTCCCT GTCCTT	
45 (RE + 編碼)	ggaggaagccatcaactaaactacaatgactgtaagatacaaaattgggaatggtaa catattttgaagttctgttgacataaagaatcatgatattaatgcccatggaaatgaaag ggcgatcaacactatggtttgaaaaggggaaattgtagagcacagatgtgttcgtg tggcagtgtgctgtctctagcaatactcagagaagagagagaacaatgaaattctga ttggccccagtgtgagcccagatgaggttcagctgccaactttctctttcacatcttatg aaagtcatttaagcacaactaacttttttttttttttttttttttgagacagagtcttgctctgt tgcccaggacagagtgcagtagtgactcaatctcggctcactgcagcctccacctcc taggctcaaacggtcctcctgcatcagcctcccaagtagctggaattacaggagtgg cccaccatgcccagctaatttttgtatttttaatagatacggggtttcaccatatcacc caggctggtctcgaactcctggcctcaagtgatccacctgcctcggcctcccaaagt gctgggattataggcgtcagccactatgcccaacccgaccaacctttttttaaaataaa tatttaaaaaattggtatttcacatatatactagtatttacatttatccacacaaaacggac	184

構建體	序列	SEQ ID NO
	gggcctccgctgaaccagtgaggccccagacgtgcgcataaataacccctgcgtg ctgcaccacctggggagaggggggaggaccacggtaaatggagcgagcgcatag caaaagggacgcggggtccttttctctgccggtggcactgggtagctgtggccagg tgtggtactttgatggggcccagggctggagctcaaggaagcgtcgcagggtcac agatctgggggaaccccggggaaaagcactgaggcaaaaccgccgctcgtctcct acaatatatgggaggggggaggttgagtacgttctggattactcataagacctttttttttt ccttccgggcgcaaaaccgtgagctggatttataatcgccctataaagctccagagg cggtcaggcacctgcagaggagccccgccgctccgccgactagctgcccccgcg agcaacggcctcgtgatttccccgccgatccggtccccgcctccccactctgcccc cgcctaccccggagccgtgcagccgcctctccgaatctctctcttctcctggcgctc gcgtgcgagagggaactagcgagaacgaggaagcagctggaggtgacgccgg gcagattacgcctgtcagggccgagccgagcggatcgctgggcgctgtgcagag gaaaggcgggagtgcccggctcgctgtcgcagagccgaggtgggtaagctagcg accacctggacttcccagcgcccaaccgtggcttttcagccaggtcctctcctcccg cggcttctcaaccaaccccatcccagcgccggccacccaacctcccgaaatgagt gcttcctgccccagcagccgaaggcgctactaggaacggtaacctgttacttttcca ggggccgtagtcgacccgctgcccgagttgctgtgcgactgcgcgcgcggggcta gagtgcaaggtgactgtggttcttctctggccaagtccgagggagaacgtaaagata tgggccttttttccccctctcaccttgtctcaccaaagtccctagtccccggagcagtta gcctctttctttccagggaattagccagacacaacaacgggaaccagacaccgaac cagacatgcccgccccgtgcgccctccccgctcgctgcctttcctccctcttgtctct ccagagccggatcttcaaggggagcctccgtgccccccggctgctcagtccctccg gtgtgcaggaccccggaagtcctccccgcacagctctcgcttctctttgcagcctgtt tctgcgccggaccagtcgaggactctggacagtagaggccccgggacgaccgag ctgGAATTCGCCACCATGGCCgccgaccacctgatgctcgccgag ggctaccgcctggtgcagaggccgccgtccgccgccgccgcccatggccctcat gcgctccggactctgccgccgtacgcgggcccgggcctggacagtgggctgagg ccgcgggggctccgctggggccgccgccgccccgccaacccggggccctgg cgtacggggccttcgggccgccgtcctccttccagccctttccggccgtgcctccgc cggccgcgggcatcgcgcacctgcagcctgtggcgacgccgtaccccggccgc gcggccgcgccccccaacgctccgggaggcccccgggcccgcagccggccc caagcgccgcagccccgccgccgcccgcgcacgccctgggcggcatggacgcc gaactcatcgacgaggaggcgctgacgtcgctggagctggagctggggctgcac cgcgtgcgcgagctgcccgagctgttcctgggccagagcgagttcgactgcttctc ggacttggggtccgcgccgcccgccggctccgtgagctgcggtggttctggtggt ggttctggtcagtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccg gcccagcaagacgccccccccacgaacgcccttacgcttgcccagtggagtcctgt gatcgccgcttctccCGCAGCGACAACCTGGTGAGAcacatccg catccacacaggccagaagcccttccagtgccgcatctgcatgAGAaacttcag cCGAGAGGATAACTTGCACACTcacatccgcacccacacaggc gaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGAT GAACTTGTCCGAcataccaagatccacttgcggcagaaggaccgcccctt acgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAGGGAAT CTGACTGAGcacatccgcatccacacaggccagaagcccttccagtgccg catctgcatgAGAaacttcagcACAAGTGGACATCTGGTACG Ccacatccgcacccacacaggcgaaaagcccttcgcctgcgacatctgtggaaga aagtttgccCAGAATAGTACCCTGACCGAAcataccaagatcca cttgcggcagaaggacaagtaaCTCGAGGAAACCCAGCAGAC AATGTAGCTAGACCCAGTAGCCAGATGTAGCTAAA	

構建體	序列	SEQ ID NO
	GAGACCGGTTCACTGTGAGAAACCCAGCAGACAAT GTAGCTAGACCCAGTAGCCAGATGTAGCTAAAGAG ACCGGTTCACTGTGAAAGCTTGGGTGGCATCCCTG TGACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGT TGCCACTCCAGTGCCCACCAGCCTTGTCCTAATAA AATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCT TCTATAATATTATGGGGTGGAGGGGGGTGGTATGG AGCAAGGGGCAAGTTGGGAAGACAACCTGTAGGG CCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCA GTGGCACAATCTTGGCTCACTGCAATCTCCGCCTCC TGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGA GTTGTTGGGATTCCAGGCATGCATGACCAGGCTCA GCTAATTTTTGTTTTTTTGGTAGAGACGGGGTTTCA CCATATTGGCCAGGCTGGTCTCCAACTCCTAATCTC AGGTGATCTACCCACCTTGGCCTCCCAAATTGCTG GGATTACAGGCGTGAACCACTGCTCCCTTCCCTGT CCTT	
8 (編碼)	ATGGCCgccgaccacctgatgctcgccgagggctaccgcctggtgcagagg ccgccgtccgccgccgccgcccatggccctcatgcgctccggactctgccgccgt acgcgggcccgggcctggacagtgggctgaggccgcgggggctccgctggg gccgccgccgccccgccaacccggggccctggcgtacggggccttcgggccgc cgtcctccttccagccctttccggccgtgcctccgccggccgcgggcatcgcgcac ctgcagcctgtggcgacgccgtaccccggccgcgcggccgcgccccccaacgct ccgggaggcccccccgggcccgcagccggccccaagcgccgcagccccgccgc cgcccgcgcacgccctgggcggcatggacgccgaactcatcgacgaggaggcg ctgacgtcgctggagctggagctggggctgcaccgcgtgcgcgagctgcccgag ctgttcctgggccagagcgagttcgactgcttctcggacttggggtccgcgccgccc gccggctccgtgagctgcggtggttctggtggtggttctggtcagtcccagctcatc aaacccagccgcatgcgcaagtaccccaaccggcccagcaagacgcccccccac gaacgcccttacgcttgcccagtggagtcctgtgatcgccgcttctccCGCAGC GACAACCTGGTGAGAcacatccgcatccacacaggccagaagcccct tccagtgccgcatctgcatgAGAaacttcagcCGAGAGGATAACTT GCACACTcacatccgcacccacacaggcgaaaagcccttcgcctgcgacat ctgtggaagaaagtttgccCGGAGCGATGAACTTGTCCGAcata ccaagatccacttgcggcagaaggaccgcccttacgcttgcccagtggagtcctgt gatcgccgcttctccCAATCAGGGAATCTGACTGAGcacatccg catccacacaggccagaagcccttccagtgccgcatctgcatgAGAaacttcag cACAAGTGGACATCTGGTACGCcacatccgcacccacacaggc gaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCAGAATAGT ACCCTGACCGAAcataccaagatccacttgcggcagaaggacaag	203
46 (編碼)	ATGGCCgccgaccacctgatgctcgccgagggctaccgcctggtgcagagg ccgccgtccgccgccgccgcccatggccctcatgcgctccggactctgccgccgt acgcgggcccgggcctggacagtgggctgaggccgcgggggctccgctggg gccgccgccgccccgccaacccggggccctggcgtacggggccttcgggccgc cgtcctccttccagccctttccggccgtgcctccgccggccgcgggcatcgcgcac ctgcagcctgtggcgacgccgtaccccggccgcgcggccgcgccccccaacgct ccgggaggcccccccgggcccgcagccggccccaagcgccgcagccccgccgc cgcccgcgcacgccctgggcggcatggacgccgaactcatcgacgaggaggcg	204

構建體	序列	SEQ ID NO
	ctgacgtcgctggagctggagctggggctgcaccgcgtgcgcgagctgcccgag ctgttcctgggccagagcgagttcgactgcttctcggacttggggtccgcgccgccc gccggctccgtgagctgcggtggttctggtggtggttctggtGGTGGCAGC GGGGGAGGTTCTGGTcagtcccagctcatcaaacccagccgcatgcg caagtaccccaaccggcccagcaagacgcccccccacgaacgcccttacgcttgc ccagtggagtcctgtgatcgccgcttctccCGCAGCGACAACCTGGT GAGAcacatccgcatccacacaggccagaagcccttccagtgccgcatctgca tgAGAaacttcagcCGAGAGGATAACTTGCACACTcacatcc gcacccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgcc CGGAGCGATGAACTTGTCCGAcataccaagatccacttgcggca gaaggaccgcccttacgcttgcccagtggagtcctgtgatcgccgcttctccCAA TCAGGGAATCTGACTGAGcacatccgcatccacacaggccagaa gcccttccagtgccgcatctgcatgAGAaacttcagcACAAGTGGACA TCTGGTACGCcacatccgcacccacacaggcgaaaagcccttcgcctgc gacatctgtggaagaaagtttgccCAGAATAGTACCCTGACCGA Acataccaagatccacttgcggcagaaggacaag	
47 (編碼)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCTA CAGACTGGTGCAGCGGCCTCCATCTGCCGCTGCCG CCCACGGCCCCCACGCCCTGAGAACACTGCCCCCC TACGCCGGCCCTGGTCTTGATAGCGGACTCAGACC TAGAGGCGCCCCTCTGGGCCCTCCACCTCCAAGAC AGCCTGGAGCCCTGGCCTACGGCGCCTTCGGCCCT CCTTCTAGCTTCCAGCCCTTCCCCGCCGTGCCTCCT CCAGCcGCTGGCATCGCCCACCTGCAGCCTGTGGC CACCCCTTACCCCGGAAGAGCCGCCGCCCCTCCAA ACGCCCCTGGCGGACCTCCTGGCCCCCAGCCTGCT CCAAGCGCCGCTGCCCCTCCACCTCCTGCTCATGC CCTGGGCGGCATGGACGCCGAGCTGATCGACGAG GAAGCCCTGACCAGCCTGGAACTGGAACTGGGCCT GCACAGAGTGCGGGAACTGCCTGAGCTGTTCCTGG GACAGAGCGAGTTCGACTGCTTCAGCGACCTGGGC AGCGCCCCTCCTGCCGGCTCTGTGTCCTGCgccgaccac ctgatgctcgccgagggctaccgcctggtgcagaggccgccgtccgccgccgcc gcccatggccctcatgcgctccggactctgccgccgtacgcgggcccgggcctgg acagtgggctgaggccgcgggggctccgctggggccgccgccgccccgccaa cccggggccctggcgtacggggccttcgggccgccgtcctccttccagccctttcc ggccgtgcctccgccggccgcgggcatcgcgcacctgcagcctgtggcgacgcc gtaccccggccgcgcCgccgcgcccccccaacgctccgggaggcccccccgggc ccgcagccggcccccaagcgccgcagccccgccgccgcccgcgcacgccctgg gcggcatggacgccgaactcatcgacgaggaggcgctgacgtcgctggagctgg agctggggctgcaccgcgtgcgcgagctgcccgagctgttcctgggccagagcg agttcgactgcttctcggacttggggtccgcgccgcccgccggctccgtgagctgc cagtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccggcccagca agacgcccccccacgaacgcccttacgcttgcccagtggagtcctgtgatcgccgc ttctccCGCAGCGACAACCTGGTGAGAcacatccgcatccacac aggccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCGAGA GGATAACTTGCACACTcacatccgcacccacacaggcgaaaagcc cttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGATGAACT	206

構建體	序列	SEQ ID NO
	TGTCCGAcataccaagatccacttgcggcagaaggaccgcccttacgcttgc ccagtggagtcctgtgatcgccgcttctccCAATCAGGGAATCTGAC TGAGcacatccgcatccacacaggccagaagcccttccagtgccgcatctgcat gAGAaacttcagcACAAGTGGACATCTGGTACGCcacatcc gcacccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgcc CAGAATAGTACCCTGACCGAAcataccaagatccacttgcggc agaaggacaag	
48 (編碼)	ATGGCCGCAGATCACCTGATGCTGGCTGAAGGCTA CAGACTGGTGCAGCGGCCTCCATCTGCCGCTGCCG CCCACGGCCCCCACGCCCTGAGAACACTGCCCCCC TACGCCGGCCCTGGTCTTGATAGCGGACTCAGACC TAGAGGCGCCCCTCTGGGCCCTCCACCTCCAAGAC AGCCTGGAGCCCTGGCCTACGGCGCCTTCGGCCCT CCTTCTAGCTTCCAGCCCTTCCCCGCCGTGCCTCCT CCAGCTGCTGGCATCGCCCACCTGCAGCCTGTGGC CACCCCTTACCCCGGAAGAGCCGCCGCCCCTCCAA ACGCCCCTGGCGGACCTCCTGGCCCCCAGCCTGCT CCAAGCGCCGCTGCCCCTCCACCTCCTGCTCATGC CCTGGGCGGCATGGACGCCGAGCTGATCGACGAG GAAGCCCTGACCAGCCTGGAACTGGAACTGGGCCT GCACAGAGTGCGGGAACTGCCTGAGCTGTTCCTGG GACAGAGCGAGTTCGACTGCTTCAGCGACCTGGGC AGCGCCCCTCCTGCCGGCTCTGTGTCCTGCGGCGG CAGCGGCGGCGGAAGCGGCgccgaccacctgatgctcgccgag ggctaccgcctggtgcagaggccgccgtccgccgccgccgcccatggccctcat gcgctccggactctgccgccgtacgcgggcccgggcctggacagtgggctgagg ccgcgggggctccgctggggccgccgccgccccgccaacccggggccctgg cgtacggggccttcgggccgccgtcctccttccagccctttccggccgtgcctccgc cggccgcgggcatcgcgcacctgcagcctgtggcgacgccgtaccccggccgc gcggccgcgccccccaacgctccgggaggcccccgggcccgcagccggccc caagcgccgcagccccgccgccgcccgcgcacgccctgggcggcatggacgcc gaactcatcgacgaggaggcgctgacgtcgctggagctggagctggggctgcac cgcgtgcgcgagctgcccgagctgttcctgggccagagcgagttcgactgcttctc ggacttggggtccgcgccgcccgccggctccgtgagctgcggtggttctggtggt ggttctggtcagtcccagctcatcaaacccagccgcatgcgcaagtaccccaaccg gcccagcaagacgccccccccacgaacgcccttacgcttgcccagtggagtcctgt gatcgccgcttctccCGCAGCGACAACCTGGTGAGAcacatccg catccacacaggccagaagcccttccagtgccgcatctgcatgAGAaacttcag cCGAGAGGATAACTTGCACACTcacatccgcacccacacaggc gaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCGGAGCGAT GAACTTGTCCGAcataccaagatccacttgcggcagaaggaccgcccctt acgcttgcccagtggagtcctgtgatcgccgcttctccCAATCAGGGAAT CTGACTGAGcacatccgcatccacacaggccagaagcccttccagtgccg catctgcatgAGAaacttcagcACAAGTGGACATCTGGTACG Ccacatccgcacccacacaggcgaaaagcccttcgcctgcgacatctgtggaaga aagtttgccCAGAATAGTACCCTGACCGAAcataccaagatcca cttgcggcagaaggacaag	208
49 (編碼)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaa	212

構建體	序列	SEQ ID NO
	agtggagatttggggacggcacccgatgaggccgtgagggccccactggactgg gcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctcc ctgctgagtcccccagcgtcgttgaacattctcagctcctccaaccccctgccttgtcc accatgaccacacctactccctcccacgggaaactgtctctatggatctagagagtg agagctgtagaaaagaggggacccagatgactccacagcatatggaggagctgg cagagcaggagattgctaggctagtactgacagatgaggagaagagtctattggag aaggaggggcttattctgcctgagacacttcctctcactaagacagaggaacaaatt ctgaaacgtgtgcggCTCGAACCAGGTGAAAAACCTTACA AATGTCCTGAATGTGGGAAATCATTCAGTCGCAGC GACAACCTGGTGAGACATCAACGCACCCATACAGG AGAAAAACCTTATAAATGTCCAGAATGTGGAAAGT CCTTCTCACGAGAGGATAACTTGCACACTCATCAA CGAACACATACTGGTGAAAAACCATACAAGTGTCC CGAATGTGGTAAAAGTTTTAGCCGGAGCGATGAAC TTGTCCGACACCAACGAACCCATACAGGCGAGAAG CCTTACAAATGTCCCGAGTGTGGCAAGAGCTTCTC ACAATCAGGGAATCTGACTGAGCATCAACGAACTC ATACCGGGGAAAAACCTTACAAGTGTCCAGAGTGT GGGAAGAGCTTTTCCACAAGTGGACATCTGGTACG CCACCAGAGGACACATACAGGGGAGAAGCCCTAC AAATGCCCCGAATGCGGTAAAAGTTTCTCTCAGAA TAGTACCCTGACCGAACACCAGCGAACACACACTG GGAAAAAAACGAGTgtgtaCgttgggggtttagagagcCgggtcttg aaatacacagcccagaatatggagcttcagaacaaagtacagcttctggaggaaca gaatttgtcccttctagatcaactgaggaaactccaggccatggtgattgagatCtca aacaaaaccagcagcagcagcacctgcatcttggtcctGctagtctccttctgcctc ctccttgtacctgctatgtactcctctgacacaaggggagcctgccagctgagcat ggagtgttgtcccgccagcttcgtgccctccccagtgaggacccttaccagctgga gctgcctgccctgcagtcagaagtgccgaaagacagcacacaccagtggttggac ggctcagactgtgtactccaggcccctggcaacacttcctgcctgctgcattacatgc ctcaggctcccagtgcagagcctcccctggagtggccCttccctgacctcttctcag agcctctctgccgaggtcccatcctcccccctgcaggcaaatctcacaaggaagggga ggatggcttcctactggtagcccctctgtcattttgcaggacagatactcaggc	
50 (編碼)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaa agtggagatttggggacggcacccgatgaggccgtgagggccccactggactgg gcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctcc ctgctgagtcccccagcgtcgttgaacattctcagctcctccaaccccctgccttgtcc accatgaccacacctactccctcccacgggaaactgtctctatggatctagagagtg agagctgtagaaaagaggggacccagatgactccacagcatatggaggagctgg cagagcaggagattgctaggctagtactgacagatgaggagaagagtctattggag aaggaggggcttattctgcctgagacacttcctctcactaagacagaggaacaaatt ctgaaacgtgtgcggcgcccttacgcttgcccagtggagtcctgtgatcgccgcttct ccCGCAGCGACAACCTGGTGAGAcacatccgcatccacacagg ccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCGAGAGG ATAACTTGCACACTcacatccgcacccacacaggcgaaaagcccttc gcctgcgacatctgtggaagaaagtttgccCGGAGCGATGAACTTG TCCGAcataccaagatccacttgcggcagaaggaccgcccttacgcttgccca gtggagtcctgtgatcgccgcttctccCAATCAGGGAATCTGACTG	216

構建體	序列	SEQ ID NO
	AGcacatccgcatccacacaggccagaagcccttccagtgccgcatctgcatgA GAaacttcagcACAAGTGGACATCTGGTACGCcacatccgca cccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCA GAATAGTACCCTGACCGAAcataccaagatccacttgcggcaga aggacgtgtaCgttgggggtttagagagcCgggtcttgaaatacacagcccagaa tatggagcttcagaacaaagtacagcttctggaggaacagaatttgtcccttctagat caactgaggaaactccaggccatggtgattgagatCtcaaacaaaaccagcagca gcagcacctgcatcttggtcctGctagtctccttctgcctcctccttgtacctgctatgt actcctctgacacaaggggagcctgccagctgagcatggagtgttgtcccgccag cttcgtgccctccccagtgaggacccttaccagctggagctgcctgccctgcagtca gaagtgccgaaagacagcacacaccagtggttggacggctcagactgtgtactcc aggcccctggcaacacttcctgcctgctgcattacatgcctcaggctcccagtgcag agcctccccctggagtggccCttccctgacctcttctcagagcctctctgccgaggtc ccatcctcccccctgcaggcaaatctcacaaggaagggaggatggcttcctactggt agcccctctgtcattttgcaggacagatactcaggc	
51 (編碼)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaa agtggagatttggggacggcacccgatgaggccgtgagggccccactggactgg gcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctcc ctgctgagtcccccagcgtcgttgaacattctcagctcctccaacccctgccttgtcc accatgaccacacctactccctcccacgggaaactgtctctatggatctagagagtg agagctgtagaaaagaggggacccagatgactccacagcatatggaggagctgg cagagcaggagattgctaggctagtactgacagatgaggagaagagtctattggag aaggaggggcttattctgcctgagacacttcctctcactaagacagaggaacaaatt ctgaaacgtgtgcggcgcccttacgcttgcccagtggagtcctgtgatcgccgcttct ccCGCTCAGACAACCTCGTTCGAcacatccgcatccacacagg ccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCACCGGA CTACACTCACGAACcacatccgcacccacacaggcgaaaagcccttc gcctgcgacatctgtggaagaaagtttgccAGAGAAGACAATCTCC ATACTcataccaagatccacttgcggcagaaggaccgcccttacgcttgccca gtggagtcctgtgatcgccgcttctccACCAGCCATTCTCTCACTG AAcacatccgcatccacacaggccagaagcccttccagtgccgcatctgcatgA GAaacttcagcCAGTCTAGCTCACTGGTGAGGcacatccgca cccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgccAG GGAGGATAACCTGCATACGcataccaagatccacttgcggcaga aggacgtgtaCgttgggggtttagagagcCgggtcttgaaatacacagcccagaa tatggagcttcagaacaaagtacagcttctggaggaacagaatttgtcccttctagat caactgaggaaactccaggccatggtgattgagatCtcaaacaaaaccagcagca gcagcacctgcatcttggtcctGctagtctccttctgcctcctccttgtacctgctatgt actcctctgacacaaggggagcctgccagctgagcatggagtgttgtcccgccag cttcgtgccctccccagtgaggacccttaccagctggagctgcctgccctgcagtca gaagtgccgaaagacagcacacaccagtggttggacggctcagactgtgtactcc aggcccctggcaacacttcctgcctgctgcattacatgcctcaggctcccagtgcag agcctccccctggagtggccCttccctgacctcttctcagagcctctctgccgaggtc ccatcctcccccctgcaggcaaatctcacaaggaagggaggatggcttcctactggt agcccctctgtcattttgcaggacagatactcaggc	218
52 (編碼)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaa agtggagatttggggacggcacccgatgaggccgtgagggccccactggactgg gcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctcc	220

構建體	序列	SEQ ID NO
	ctgctgagtcccccagcgtcgttgaacattctcagctcctccaacccctgccttgtcc accatgaccacacctactccctcccacgggaaactgtctctatggatctagagagtg agagctgtagaaaagaggggacccagatgactccacagcatatggaggagctgg cagagcaggagattgctaggctagtactgacagatgaggagaagagtctattggag aaggaggggcttattctgcctgagacacttcctctcactaagacagaggaacaaatt ctgaaacgtgtgcggCTCGAACCAGGTGAAAAACCTTACA AATGTCCTGAATGTGGGAAATCATTCAGTCGCAGC GACAACCTGGTGAGACATCAACGCACCCATACAGG AGAAAAACCTTATAAATGTCCAGAATGTGGAAAGT CCTTCTCACGAGAGGATAACTTGCACACTCATCAA CGAACACATACTGGTGAAAAACCATACAAGTGTCC CGAATGTGGTAAAAGTTTTAGCCGGAGCGATGAAC TTGTCCGACACCAACGAACCCATACAGGCGAGAAG CCTTACAAATGTCCCGAGTGTGGCAAGAGCTTCTC ACAATCAGGGAATCTGACTGAGCATCAACGAACTC ATACCGGGGAAAAACCTTACAAGTGTCCAGAGTGT GGGAAGAGCTTTTCCACAAGTGGACATCTGGTACG CCACCAGAGGACACATACAGGGGAGAAGCCCTAC AAATGCCCCGAATGCGGTAAAAGTTTCTCTCAGAA TAGTACCCTGACCGAACACCAGCGAACACACACTG GGAAAAAAACGAGTgtgtaCgttgggggtttagagagcCgggtcttg aaatacacagcccagaatatggagcttcagaacaaagtacagcttctggaggaaca gaatttgtcccttctagatcaactgaggaaactccaggccatggtgattgagatatca	
53 (編碼)	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaa agtggagatttggggacggcacccgatgaggccgtgagggccccactggactgg gcgctgccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctcc ctgctgagtcccccagcgtcgttgaacattctcagctcctccaacccctgccttgtcc accatgaccacacctactccctcccacgggaaactgtctctatggatctagagagtg agagctgtagaaaagaggggacccagatgactccacagcatatggaggagctgg cagagcaggagattgctaggctagtactgacagatgaggagaagagtctattggag aaggaggggcttattctgcctgagacacttcctctcactaagacagaggaacaaatt ctgaaacgtgtgcggcgcccttacgcttgcccagtggagtcctgtgatcgccgcttct ccCGCAGCGACAACCTGGTGAGAcacatccgcatccacacagg ccagaagcccttccagtgccgcatctgcatgAGAaacttcagcCGAGAGG ATAACTTGCACACTcacatccgcacccacacaggcgaaaagcccttc gcctgcgacatctgtggaagaaagtttgccCGGAGCGATGAACTTG TCCGAcataccaagatccacttgcggcagaaggaccgcccttacgcttgccca gtggagtcctgtgatcgccgcttctccCAATCAGGGAATCTGACTG AGcacatccgcatccacacaggccagaagcccttccagtgccgcatctgcatgA GAaacttcagcACAAGTGGACATCTGGTACGCcacatccgca cccacacaggcgaaaagcccttcgcctgcgacatctgtggaagaaagtttgccCA GAATAGTACCCTGACCGAAcataccaagatccacttgcggcaga aggacgtgtaCgttgggggtttagagagcCgggtcttgaaatacacagcccagaa tatggagcttcagaacaaagtacagcttctggaggaacagaatttgtcccttctagat caactgaggaaactccaggccatggtgattgagatCtca	222

表8. 示例性微小RNA和微小RNA結合位點的核酸序列。

描述	序列	SEQ ID NO
M1結合位點	aaagagaccggttcactgtgacagtaaaagagaccggttcactgtgagaatgaaagag accggttcactgtgatcggaaaagagaccggttcactgtgagcggccttgaaacccagc agacaatgtagctcagtagaaacccagcagacaatgtagctgaatggaaacccagcag acaatgtagcttcggagaaacccagcagacaatgtagct	7
miR128序列	UCACAGUGAACCGGUCUCUUU	8
miR128結合位點	AAAGAGACCGGTTCACTGTGA	9
miR221序列	AGCUACAUUGUCUGCUGGGUUUC	10
miR221結合位點	GAAACCCAGCAGACAATGTAGCT	11
miR222序列	AGCUACAUCUGGCUACUGGGUCU	12
miR222結合位點	AGACCCAGTAGCCAGATGTAGCT	13
M2結合位點	GAAACCCAGCAGACAATGTAGCTAGACCCAGTAGCC AGATGTAGCTAAAGAGACCGGTTCACTGTGA	14
M3結合位點	GAAACCCAGCAGACAATGTAGCTAGACCCAGTAGCC AGATGTAGCTAAAGAGACCGGTTCACTGTGAGAAAC CCAGCAGACAATGTAGCTAGACCCAGTAGCCAGATG TAGCTAAAGAGACCGGTTCACTGTGA	15

表9. 鋅指結構的不同類型和用於生成eTF的示例性鋅指蛋白。

ZF類型名稱	**SEQ ID NO**	**ZF結構 (其中每個x可以獨立地為任何殘基)**	**可以充當eTF的蛋白質平臺或本文公開的eTF的DNA結合域的示例性蛋白質**
鋅指C2H2-型 (ZNF)	136	C-x-C-x-H-x-H	KLF4, KLF5, EGR3, ZFP637, SLUG, ZNF750, ZNF281, ZBP89, GLIS1, GLIS3
環指蛋白(RNF)	137	C-x-C-x-C-x-H-xxx-C-x-C-x-C-x-C	MDM2, BRCA1, ZNF179
PHD指蛋白 (PHF)	138	C-x-C-x-C-x-C-xxx-H-x-C-x-C-x-C	KDM2A, PHF1, ING1
含LIM結構域	139	C-x-C-x-H-x-C-x-C-x-C-x-C-x-(C,H,D)	ZNF185, LIMK1, PXN
核激素受體 (NR)	140	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	VDR, ESR1, NR4A1
鋅指CCCH-型 (ZC3H)	141	C-x-C-x-C-x-H	RC3H1, HELZ, MBNL1, ZFP36, ZFP36L1
鋅指FYVE-型	140	C-x-C-x-C-x-C-xxx-C-x-	EEA1, HGS, PIKFYVE

ZF類型名稱	SEQ ID NO	ZF結構 (其中每個x可以獨立地為任何殘基)	可以充當eTF的蛋白質平臺或本文公開的eTF的DNA結合域的示例性蛋白質
(ZFYVE)		C-x-C-x-C	
鋅指CCHC-型(ZCCHC)	142	C-x-C-x-H-x-C	CNBP, SF1, LIN28A
鋅指DHHC-型(ZDHHC)	143	C-x-C-x-H-x-C-xxx-C-x-C-x-H-x-C	ZDHHC2, ZDHHC8, ZDHHC9
鋅指MYND-型(ZMYND)	144	C-x-C-x-C-x-C-xxx-C-x-C-x-H-x-C	PDCD2, RUNX1T1, SMYD2, SMYD1
鋅指RANBP2-型(ZRANB)	145	C-x-C-x-C-x-C	YAF2, SHARPIN, EWSR1
鋅指ZZ-型(ZZZ)	145	C-x-C-x-C-x-C	HERC2, NBR1, CREBBP
鋅指C2HC-型(ZC2HC)	142	C-x-C-x-H-x-C	IKBKG, L3MBTL1, ZNF746
含GATA鋅指結構域(GATAD)	145	C-x-C-x-C-x-C	GATA4, GATA6, MTA1
ZF類同源框和假基因	136	C-x-C-x-H-x-H	ADNP, ZEB1, ZHX1
含THAP結構域(THAP)	141	C-x-C-x-C-x-H	THAP1, THAP4, THAP11
鋅指CXXC-型(CXXC)	140	C-x-C-x-C-x-C-xxx-C-x-C-x-C-x-C	CXXC1, CXXC5, MBD1, DNMT1
鋅指SWIM-型(ZSWIM)	141	C-x-C-x-C-x-H	MAP3K1, ZSWIM5, ZSWIM6
鋅指AN1-型(ZFAND)	146	C-x-C-x-C-x-C-xxx-C-x-H-x-H-x-C	ZFAND3, ZFAND6, IGHMBP2
鋅指3CxxC-型(Z3CXXC)	142	C-x-C-x-H-x-C	ZAR1, RTP1, RTP4
鋅指CW-型(ZCW)	145	C-x-C-x-C-x-C	MORC1, ZCWPW1, KDM1B
鋅指GRF-型(ZGRF)	145	C-x-C-x-C-x-C	TTF2, NEIL3, TOP3A
鋅指MIZ-型(ZMIZ)	142	C-x-C-x-H-x-C	PIAS1, PIAS3, PIAS4
鋅指BED-型(ZBED)	136	C-x-C-x-H-x-H	ZBED1, ZBED4, ZBED6

ZF類型名稱	SEQ ID NO	ZF結構 (其中每個x可以獨立地為任何殘基)	可以充當eTF的蛋白質平臺或本文公開的eTF的DNA結合域的示例性蛋白質
鋅指HIT-型(ZNHIT)	144	C-x-C-x-C-x-C-xxx-C-x-C-x-H-x-C	ZNHIT3, DDX59, INO80B
鋅指MYM-型(ZMYM)	145	C-x-C-x-C-x-C	ZMYM2, ZMYM3, ZMYM4
鋅指基質蛋白型(ZMAT)	136	C-x-C-x-H-x-H	ZNF638, ZMAT1, ZMAT3, ZMAT5
鋅指C2H2C-型	136	C-x-C-x-H-x-H	MYT1, MYT1L, ST18
鋅指DBF-型(ZDBF)	136	C-x-C-x-H-x-H	DBF4, DBF4B, ZDBF2
鋅指PARP-型	142	C-x-C-x-H-x-C	LIG3, PARP1

表10. 示例性鋅指DNA結合域的氨基酸序列。

DBD/靶位點	序列	SEQ ID NO
eZF	LEPGEKP – [YKCPECGKSFS X HQRTH TGEKP]n - YKCPECGKSFS X HQRTH – TGKKTS，其中n為1-15的整數，並且每個X是能夠與靶序列的3 bp結合的識別序列	147
Z1靶位點	RSDNLVR x REDNLHT x RSDELVR x QSGNLTE x TSGHLVR x QNSTLTE，其中每個x是包含1-50個氨基酸殘基的連接體	148
Z13靶位點	RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT，其中每個x是包含1-50個氨基酸殘基的連接體	149
Z14靶位點	DPGALVR x RSDNLVR x QSGDLRR x THLDLIR x TSGNLVR x RSDNLVR，其中每個x是包含1-50個氨基酸殘基的連接體	150
Z15靶位點	RRDELNV x RSDHLTN x RSDDLVR x RSDNLVR x HRTTLTN x REDNLHT x TSHSLTE x QSSSLVR x REDNLHT，其中每個x是包含1-50個氨基酸殘基的連接體	151

表11. 本文公開的示例性鋅指識別序列的氨基酸序列。

序列	SEQ ID NO
RSDNLVR	152
REDNLHT	153
RSDELVR	154
QSGNLTE	155

序列	SEQ ID NO
TSGHLVR	156
QNSTLTE	157
DPGALVR	158
HRTTLTN	159
QSGDLRR	160
TSHSLTE	161
THLDLIR	162
QSSSLVR	163
TSGNLVR	164
RRDELNV	165
RSDDLVR	166
RSDHLTN	167

表12. 本文公開的其他核苷酸和氨基酸序列。

描述	序列	SEQ ID NO
EGR1 NLS結構域	LIKPSRMRKYPNRPSK	168
SV40 NLS	PKKKRKV	169
核質蛋白NLS	KRPAATKKAGQAKKKK	170
HA標籤	YPYDVPDYA	171
spA(合成的聚A)	AATAAAAGATCTTTATTTTCATTAGATCTGTGTGTTGGTTTTTTGTGTGCGGACCGCACGTG	16
hGH(人生長激素聚A)	GGGTGGCATCCCTGTGACCCCTCCCCAGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCCACCAGCCTTGTCCTAATAAAATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTTCTATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGGCAAGTTGGGAAGACAACCTGTAGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCACAATCTTGGCTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTTGGTAGAGACGGGGTTTCACCATATTGGCCAGGCTGGTCTCCAACTCCTAATCTCAGGTGATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCACTGCTCCCTTCCCTGTCCTT	17
SCN1A蛋白	MEQTVLVPPGPDSFNFFTRESLAAIERRIAEEKAKNPKPDKKDDDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLEDLDPYYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKI	172

描述	序列	SEQ ID NO
	AIKILVHSLFSMLIMCTILTNCVFMTMSNPPDWTKNVEY TFTGIYTFESLIKIIARGFCLEDFTFLRDPWNWLDFTVITF AYVTEFVDLGNVSALRTFRVLRALKTISVIPGLKTIVGAL IQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCIQ WPPTNASLEEHSIEKNITVNYNGTLINETVFEFDWKSYIQ DSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGRN PNYGYTSFDTFSWAFLSLFRLMTQDFWENLYQLTLRAA GKTYMIFFVLVIFLGSFYLINLILAVVAMAYEEQNQATLE EAEQKEAEFQQMIEQLKKQQEAAQQAATATASEHSREP SAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGG EEKDEDEFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSSP HQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDVGSENDFAD DEHSTFEDNESRRDSLFVPRRHGERRNSNLSQTSRSSRM LAVFPANGKMHSTVDCNGVVSLVGGPSVPTSPVGQLLP EVIIDKPATDDNGTTTETEMRKRRSSSFHVSMDFLEDPSQ RQRAMSIASILTNTVEELEESRQKCPPCWYKFSNIFLIWD CSPYWLKVKHVVNLVVMDPFVDLAITICIVLNTLFMAM EHYPMTDHFNNVLTVGNLVFTGIFTAEMFLKIIAMDPYY YFQEGWNIFDGFIVTLSLVELGLANVEGLSVLRSFRLLRV FKLAKSWPTLNMLIKIIGNSVGALGNLTLVLAIIVFIFAVV GMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVF RVLCGEWIETMWDCMEVAGQAMCLTVFMMVMVIGNL VVLNLFLALLLSSFSADNLAATDDDNEMNNLQIAVDRM HKGVAYVKRKIYEFIQQSFIRKQKILDEIKPLDDLNNKKD SCMSNHTAEIGKDLDYLKDVNGTTSGIGTGSSVEKYIIDE SDYMSFINNPSLTVTVPIAVGESDFENLNTEDFSSESDLEE SKEKLNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACF TEGCVQRFKCCQINVEEGRGKQWWNLRRTCFRIVEHNW FETFIVFMILLSSGALAFEDIYIDQRKTIKTMLEYADKVFT YIFILEMLLKWVAYGYQTYFTNAWCWLDFLIVDVSLVS LTANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVV NALLGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFYHCI NTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNF DNVGFGYLSLLQVATFKGWMDIMYAAVDSRNVELQPK YEESLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKFG GQDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFQG MVFDFVTRQVFDISIMILICLNMVTMMVETDDQSEYVTT ILSRINLVFIVLFTGECVLKLISLRHYYFTIGWNIFDFVVVI LSIVGMFLAELIEKYFVSPTLFRVIRLARIGRILRLIKGAK GIRTLLFALMMSLPALFNIGLLLFLVMFIYAIFGMSNFAY VKREVGIDDMFNFETFGNSMICLFQITTSAGWDGLLAPIL NSKPPDCDPNKVNPGSSVKGDCGNPSVGIFFFVSYIIISFL VVVNMYIAVILENFSVATEESAEPLSEDDFEMFYEVWEK FDPDATQFMEFEKLSQFAAALEPPLNLPQPNKLQLIAMD LPMVSGDRIHCLDILFAFTKRVLGESGEMDALRIQMEER FMASNPSKVSYQPITTTLKRKQEEVSAVIIQRAYRRHLLK RTVKQASFTYNKNKIKGGANLLIKEDMIIDRINENSITEK TDLTMSTAACPPSYDRVTKPIVEKHEQEGKDEKAKGK	

描述	序列	SEQ ID NO
dCAS蛋白	KRNYILGLAIGITSVGYGIIDYETRDVIDAGVRLFKEANV ENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTD HSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGV HNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLER LKKDGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQ SFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLM GHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDEN EKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKG YRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIA KILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTH NLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQ KEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIE LAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKEN AKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEV DHIIPRSVSFDNSFNNKVLVKQEEASKKGNRTPFQYLSSS DSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRF SVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKV KSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANAD FIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEY KEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYS TRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLL MYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYL TKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNK VVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYE VNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRV IGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQIIKKG	173
dCAS9-VP64構建體	MAPKKKRKVGIHGVPAAKRNYILGLAIGITSVGYGIIDYE TRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRH RIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSE EEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRN SKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEA KQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGS PFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLY NALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPT LKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDI TARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQ EEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIF NRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIK VINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNR QTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLE AIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEE ASKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRIS KTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMN LLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNK GYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFE EKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHR VDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKD	174

描述	序列	SEQ ID NO
	NDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDE KNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLN AHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFV TVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIAS FYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYL ENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKK HPQIIKKGKRPAATKKAGQAKKKKGSYPYDVPDYALED ALDDFDLDMLGSDALDDFDLDMLGSDALDDFDLDMLG SDALDDFDLDML	
WT EGR3蛋白(人)	MTGKLAEKLPVTMSSLLNQLPDNLYPEEIPSALNLFSGSS DSVVHYNQMATENVMDIGLTNEKPNPELSYSGSFQPAP GNKTVTYLGKFAFDSPSNWCQDNIISLMSAGILGVPPAS GALSTQTSTASMVQPPQGDVEAMYPALPPYSNCGDLYS EPVSFHDPQGNPGLAYSPQDYQSAKPALDSNLFPMIPDY NLYHHPNDMGSIPEHKPFQGMDPIRVNPPPITPLETIKAF KDKQIHPGFGSLPQPPLTLKPIRPRKYPNRPSKTPLHERPH ACPAEGCDRRFSRSDELTRHLRIHTGHKPFQCRICMRSFS RSDHLTTHIRTHTGEKPFACEFCGRKFARSDERKRHAKI HLKQKEKKAEKGGAPSASSAPPVSLAPVVTTCA	175
WT EGR1蛋白(人)	MAAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEM MLLSNGAPQFLGAAGAPEGSGSNSSSSSSGGGGGGGGG SNSSSSSSTFNPQADTGEQPYEHLTAESFPDISLNNEKVL VETSYPSQTTRLPPITYTGRFSLEPAPNSGNTLWPEPLFSL VSGLVSMTNPPASSSSAPSPAASSASASQSPPLSCAVPSN DSSPIYSAAPTFPTPNTDIFPEPQSQAFPGSAGTALQYPPP AYPAAKGGFQVPMIPDYLFPQQQGDLGLGTPDQKPFQG LESRTQQPSLTPLSTIKAFATQSGSQDLKALNTSYQSQLI KPSRMRKYPNRPSKTPPHERPYACPVESCDRRFSRSDEL TRHIRIHTGQKPFQCRICMRNFSRSDHLTTHIRTHTGEKP FACDICGRKFARSDERKRHTKIHLRQKDKKADKSVVASS ATSSLSSYPSPVATSYPSPVTTSYPSPATTSYPSPVPTSFSS PGSSTYPSPVHSGFPSPSVATTYSSVPPAFPAQVSSFPSSA VTNSFSASTGLSDMTATFSPRTIEIC	176
連接體	GGSGGGSG	177
連接體	GGSGGGSGGGSGGGSG	178
連接體	[GGGS]n	179
連接體	[GGGGS]n	180
連接體	[GGSG]n	181
針對WT EGR1或EGR3的識別位點	GCG(T/G)GGGCG	182
sgRNA支架	GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAG GCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAG	183

描述	序列	SEQ ID NO
	A	
人CREB3編碼序列	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaaagt ggagatttggggacggcacccgatgaggccgtgagggccccactggactgggcgctg ccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctccctgctgagtc ccccagcgtcgttgaacattctcagctcctccaacccctgccttgtccaccatgaccacac ctactccctcccacgggaaactgtctctatggatctagagagtgagagctgtagaaaaga ggggacccagatgactccacagcatatggaggagctggcagagcaggagattgctagg ctagtactgacagatgaggagaagagtctattggagaaggaggggcttattctgcctgag acacttcctctcactaagacagaggaacaaattctgaaacgtgtgcggaggaagattcga aataaaagatctgctcaagagagccgcaggaaaaagaaggtgtaCgttgggggtttaga gagcCgggtcttgaaatacacagcccagaatatggagcttcagaacaaagtacagcttct ggaggaacagaatttgtcccttctagatcaactgaggaaactccaggccatggtgattgag atCtcaaacaaaaccagcagcagcagcacctgcatcttggtcctGctagtctccttctgc ctcctccttgtacctgctatgtactcctctgacacaaggggagcctgccagctgagcatg gagtgttgtcccgccagcttcgtgccctccccagtgaggacccttaccagctggagctgc ctgccctgcagtcagaagtgccgaaagacagcacacaccagtggttggacggctcaga ctgtgtactccaggcccctggcaacacttcctgcctgctgcattacatgcctcaggctccca gtgcagagcctcccctggagtggccCttccctgacctcttctcagagcctctctgccgag gtcccatcctcccccctgcaggcaaatctcacaaggaagggaggatggcttcctactggta gcccctctgtcattttgcaggacagatactcaggc	210
人CREB3氨基酸序列	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWA LPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCLVHHD HTYSLPRETVSMDLESESCRKEGTQMTPQHMEELAEQEI ARLVLTDEEKSLLEKEGLILPETLPLTKTEEQILKRVRRKI RNKRSAQESRRKKKVYVGGLESRVLKYTAQNMELQNK VQLLEEQNLSLLDQLRKLQAMVIEISNKTSSSSTCILVLL VSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALPSEDP YQLELPALQSEVPKDSTHQWLDGSDCVLQAPGNTSCLL HYMPQAPSAEPPLEWPFPDLFSEPLCRGPILPLQANLTRK GGWLPTGSPSVILQDRYSG	211
具有跨膜區的人CREB3 C-末端結構域	VYVGGLESRVLKYTAQNMELQNKVQLLEEQNLSLLDQL RKLQAMVIEISNKTSSSSTCILVLLVSFCLLLVPAMYSSD TRGSLPAEHGVLSRQLRALPSEDPYQLELPALQSEVPKD STHQWLDGSDCVLQAPGNTSCLLHYMPQAPSAEPPLEW PFPDLFSEPLCRGPILPLQANLTRKGGWLPTGSPSVILQD RYSG	225
沒有跨膜區的人CREB3 C-末端結構域	VYVGGLESRVLKYTAQNMELQNKVQLLEEQNLSLLDQL RKLQAMVIEIS	226
CREB3-TRE編碼序列	atggagctggaattggatgctggtgaccaagacctgctggccttcctgctagaggaaagt ggagatttggggacggcacccgatgaggccgtgagggccccactggactgggcgctg ccgctttctgaggtGccgagcgactgggaagtagatgatttgctgtgctccctgctgagtc ccccagcgtcgttgaacattctcagctcctccaacccctgccttgtccaccatgaccacac ctactccctcccacgggaaactgtctctatggatctagagagtgagagctgtagaaaaga ggggacccagatgactccacagcatatggaggagctggcagagcaggagattgctagg ctagtactgacagatgaggagaagagtctattggagaaggaggggcttattctgcctgag	214

描述	序列	SEQ ID NO
	acacttcctctcactaagacagaggaacaaattctgaaacgtgtgcggCTTGAGCC CGGAGAGAAGCCGTACAAGTGCCCTGAGTGCGGCAA GTCTTTTAGCAGAAGAGACGAACTTAATGTCCACCAG CGAACGCATACTGGTGAAAAGCCCTATAAATGTCCTG AATGTGGGAAATCATTCTCCAGCCGCAGAACCTGTAG GGCTCACCAGCGAACACACACCGGCGAAAAACCATA CAAATGTCCAGAATGCGGGAAATCCTTTTCTCAGTCAT CCAACTTGGTGAGACATCAACGCACGCACACTGGAGA AAAGCCTTACAAATGCCCGGAATGTGGAAAGTCTTTT TCCCAATTGGCCCATTTGCGAGCCCATCAGAGGACTC ACACGGGCGAGAAACCTTACAAATGCCCGGAATGCGG GAAATCTTTTTCAACGAGTGGCAACCTCGTAAGACAC CAAAGAACGCATACAGGCGAAAAGCCATATAAGTGTC CTGAGTGTGGTAAATCATTCTCACACAGGACCACCCT GACAAATCACCAGCGCACGCACACCGGCAAGAAGAC AAGCgtgtaCgttgggggtttagagagcCgggtcttgaaatacacagcccagaatat ggagcttcagaacaaagtacagcttctggaggaacagaatttgtcccttctagatcaactg aggaaactccaggccatggtgattgagatCtcaaacaaaaccagcagcagcagcacct gcatcttggtcctGctagtctccttctgcctcctccttgtacctgctatgtactcctctgacac aaggggagcctgccagctgagcatggagtgttgtcccgccagcttcgtgccctcccca gtgaggacccttaccagctggagctgcctgccctgcagtcagaagtgccgaaagacagc acacaccagtggttggacggctcagactgtgtactccaggcccctggcaacacttcctgc ctgctgcattacatgcctcaggctcccagtgcagagcctcccctggagtggccCttccct gacctcttctcagagcctctctgccgaggtcccatcctccccctgcaggcaaatctcacaa ggaagggaggatggcttcctactggtagcccctctgtcattttgcaggacagatactcagg c	
CREB3-TRE	MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWA LPLSEVPSDWEVDDLLCSLLSPPASLNILSSSNPCLVHHD HTYSLPRETVSMDLESESCRKEGTQMTPQHMEELAEQEI ARLVLTDEEKSLLEKEGLILPETLPLTKTEEQILKRVRLEP GEKPYKCPECGKSFSRRDELNVHQRTHTGEKPYKCPEC GKSFSSRRTCRAHQRTHTGEKPYKCPECGKSFSQSSNLV RHQRTHTGEKPYKCPECGKSFSQLAHLRAHQRTHTGEK PYKCPECGKSFSTSGNLVRHQRTHTGEKPYKCPECGKSF SHRTTLTNHQRTHTGKKTSVYVGGLESRVLKYTAQNME LQNKVQLLEEQNLSLLDQLRKLQAMVIEISNKTSSSSTCI LVLLVSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALP SEDPYQLELPALQSEVPKDSTHQWLDGSDCVLQAPGNTS CLLHYMPQAPSAEPPLEWPFPDLFSEPLCRGPILPLQANL TRKGGWLPTGSPSVILQDRYSG	215

實施例

【0317】包括以下實施例以進一步描述本公開的一些方面，並且不應用來限制本發明的範圍。

實施例1

使用SCN1A特異性轉錄啟動物，能夠上調SCN1A的靶區域的鑒定

【0318】 為了鑒定能夠上調內源SCN1A表達的基因組區域，設計了各種工程化轉錄因子(鋅指核酸酶或gRNA/daCas9構建體)，它們靶向如以上表4和表13所示的基因組的各個區域。對於gRNA/daCas9構建體，由於將gRNA設計為靶向互補基因組鏈，因此該gRNA具有與靶區域相同的序列。該dCas9蛋白是dCAS9-VP64構建體(SEQ ID NO: 174)。

【0319】 按照標準方法培養HEK293細胞，並在6孔板的每個孔中用3ug攜帶工程化轉錄因子或對照構建體的質粒進行轉染(FugeneHD，Promega)。用表達以下表13中所示的構建體的質粒轉染細胞。轉染後48h，收集細胞並分離RNA(Qiagen RNeasy Mini試劑盒)，並進行DNA酶處理。使用OligoDT引物(Superscript IV，Invitrogen)對RNA(3ug)進行逆轉錄。使用Phusion聚合酶(New England Biolabs)和SYBR Green I，通過qPCR分析cDNA樣品：(98°C下30s，40x[98°C下10 sec，66°C下15 sec，72°C下15 sec])。使用針對SCN1A的引物(5'-TGTCTCGGCATTGAGAACATTC-3'(SEQ ID NO: 185)；5'-ATTGGTGGGAGGCCATTGTAT-3'(SEQ ID NO: 186))來量化內源SCN1A轉錄物的水準，並以GAPDH為參考基因通過delta-delta Ct方法(5'-ACCACAGTCCATGCCATCAC'-3'(SEQ ID NO: 187)；5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO: 188))確定SCN1A表達的相對水準。資料以相對於對照條件的變化倍數呈現。

【0320】 結果以SCN1A轉錄相對於對照條件(例如，EGFP-KASH報導構建體)的變化倍數示於圖1和以下的表13中。表13報告了導致轉錄相對

於對照條件增加至少1.5倍的構建體的值。

表13. 不同基因組靶位點和相應eTF對轉錄的影響。CON表示在實驗中使用的鋅指構建體(參見表1)。對於gRNA構建體，靶位點和gRNA序列是相同的，因為gRNA被設計為靶向互補DNA鏈。

CON	染色體2起始位置	靶序列	來自圖1的靶位點	靶位點的SEQ ID NO	平均值	T檢驗
4	166149168	ctaggtcaagtgtaggag	z1	18	5.12090848	0.0096628
28	166149165	GGTCAAGTGTAGGAGACA	z8	25	1.52068773	0.62403349
3	166128025	taggtaccatagagtgag	z13	30	25.4730028	0.14942042
29	166127991	gaggatactgcagaggtc	z14	31	8.6766618	0.16432794
	166149176	aaggctgtctaggtcaagtgt	g9	35	1.36378425	0.18753821
	166149118	tgttcctccagattaacactt	g10	36	1.63040825	0.46710683
	166128002	gatgaagccgagaggatactg	g11	37	18.7579211	0.13148732
	166128037	gctgatttgtattaggtacca	g12	38	22.4892633	0.09291316
	166177299	AGAAAGCTGATACAGATACAA	g15	39	1.7542842	0.34519408
	166178880	ggtacgggcaaagatttcttg	g17	40	1.36801947	0.48762102
	166177299	AGAAAGCTGATACAGATACAA	g19	41	1.45636874	0.44464045
	166177369	ACACAATGAGCCACCTACAAG	g20	42	1.31187425	0.42605224
	166177362	GTGGCTCATTGTGTGTGTGCC	g22	43	1.25217773	0.26572657
	166155264	CATATCCCTGCAGGTTCAGAA	g24	44	1.75688991	0.28984533
	166155099	agagagagagagagagagaga	g25	45	2.05701745	0.42409102
	166155393	TTCTCAGTTTTGAAATTAAAA	g26	46	1.64498972	0.21705582
	166155255	TGGATTCTCTTCTGAACCTGC	g27	47	2.27026665	0.43195546
	166148361	TGCTGAGGCAGGACACAGTGT	g29	48	2.4290169	0.30364553
	166148843	ATCATCTGTAACCATCAAGGA	g30	49	2.58328006	0.0748197
	166148565	TCCTGCCTACTTAGTTTCAAG	g31	50	1.97097781	0.25980859
	166148953	ATTACAGTTCTGTCAGCATGC	g32	51	1.34500323	0.32186367
	166149373	TGGTCTCATTCTTTTTGTGGG	g33	52	1.71471378	0.32104302
	166142239	CGATATTTTCATGGATTCCTT	g34	53	1.7735976	0.21954265
	166142391	CTGACACTTACTTTGTCTAAA	g35	54	1.95513108	0.02069095
	166142219	AAAACTGGAACCGCATTCCCA	g36	55	2.08698135	0.0454403
	166142396	ACAAAGTAAGTGTCAGTGTGG	g37	56	1.30739959	0.72725347
	166142344	ATAATAGTTGTGTCTTTATAA	g38	57	1.55783618	0.29846459

CON	染色體2起始位置	靶序列	來自圖1的靶位點	靶位點的SEQ ID NO	平均值	T檢驗
	166141162	TGTACAAGCAGGGCTGCAAAG	g39	58	1.4663605	0.02946062
	166140590	GTTAACAAATACACTAAACAC	g40	59	1.37399196	0.33638238
	166140928	ttcaacaagctcccaagaagt	g42	60	1.46929899	0.24465271
	166141090	ATGTTCAAGGTGCAGAAGGAA	g43	61	2.04547409	0.09880194
	165990246	TGTTTGCTCAAACGTGCACCA	g44	62	2.13402102	0.25583999
	165989684	AAATAAGACATGAAAACAAGA	g45	63	1.27016182	0.32368695
	165990193	AAATATGTACCACAAGAAATG	g46	64	2.29522738	0.41829497
	165990076	TATCTGGTTTCTCTCACTGCT	g47	65	1.44542116	0.0947106
	165989571	ATTGCAAAGCATAATTTGGAT	g48	66	1.42246971	0.18117243

實施例2

使用SCN1A特異性轉錄因子，HEK293細胞中內源SCN1A的上調

【0321】 按照標準方法培養HEK293細胞，並將其接種到6孔板中。用3ug攜帶單個工程化轉錄因子構建體、WT人CREB3(SEQ ID NO: 211)或EGFP對照構建體的質粒轉染(FugeneHD，Promega)每個孔中的細胞。測試的工程化轉錄因子構建體包括：構建體1-27和46-53(表1)和表達CREB3-TRE(SEQ ID NO: 215；bZIP DNA結合域被TET啟動子靶向的合成ZF結構域代替的CREB3)的質粒(各自單獨測試)。轉染後48h，收集細胞並分離RNA(Qiagen RNeasy Mini試劑盒)，並進行DNA酶處理。使用OligoDT引物(Superscript IV，Invitrogen)對RNA(3ug)進行逆轉錄。使用Phusion聚合酶(New England Biolabs)和SYBR Green I，通過qPCR分析cDNA樣品：(98°C下30s，40x[98°C下10 sec，66°C下15 sec，72°C下15 sec])。使用針對SCN1A的引物(5'-TGTCTCGGCATTGAGAACATTC-3'(SEQ ID NO: 185)；5'-ATTGGTGGGAGGCCATTGTAT-3'(SEQ ID NO: 186))來量化內源SCN1A轉錄物的水準，並以GAPDH為參考基因通

過delta-delta Ct方法(5’-ACCACAGTCCATGCCATCAC’3’(SEQ ID NO: 187)；5’-TCCACCACCCTGTTGCTGTA-3’(SEQ ID NO: 188))確定SCN1A表達的相對水準。資料以相對於對照條件的變化倍數呈現(參見圖2A、圖2B和圖2C)。對照構建體由在RE 1(SEQ ID NO: 1)控制下表達的EGFP組成。相對於EGFP條件，工程化轉錄因子的遞送誘導內源SCN1A轉錄物的不同程度的上調。

實施例3

使用SCN1A特異性轉錄因子，GABA神經元中內源SCN1A的上調

【0322】將iCell GABA神經元(Cellular Dynamics)接種於6孔板中(約1E6個細胞/孔)，並且按照製造商推薦的方案進行維持。接種後72 h，以大約2E11個基因組拷貝/孔，將在普遍存在的啟動子(CBA啟動子)控制下表達EGFP或啟動物(圖3A中的構建體30或圖3B中的構建體25或構建體16)的重組AAV(血清型AAV-DJ)添加到培養基中。感染後一周(圖3A)或兩周(圖3B)，從培養的細胞中分離RNA(Qiagen RNeasy Mini試劑盒)，並進行DNA酶處理。使用OligoDT引物(Superscript IV，Invitrogen)對回收的RNA進行逆轉錄。使用Phusion聚合酶(New England Biolabs)和SYBR Green I，通過qPCR分析cDNA樣品：(98°C下30s，40x[98°C下10 sec，66°C下15 sec，72°C下15 sec])。使用針對SCN1A的引物(5’-TGTCTCGGCATTGAGAACATTC-3’(SEQ ID NO: 185)；5’-ATTGGTGGGAGGCCATTGTAT-3’(SEQ ID NO: 186))來量化內源SCN1A轉錄物的水準，並以GAPDH為參考基因通過delta-delta Ct方法(5’-ACCACAGTCCATGCCATCAC’3’(SEQ ID NO: 187)；5’-TCCACCACCCTGTTGCTGTA-3’(SEQ ID NO: 188))確定SCN1A表達的

相對水準。資料以相對於對照條件的變化倍數呈現(參見圖3A和圖3B)。AAV驅動的工程化轉錄因子的表達在培養的iPS衍生的GABA神經元中產生內源SCN1A轉錄物的顯著上調。

實施例4

使用SCN1A特異性轉錄因子，GABA神經元中內源SCN1A的特異性上調

【0323】 將iCell GABA神經元(Cellular Dynamics)接種於6孔板中(約1E6個細胞/孔)，並且按照製造商推薦的方案進行維持。接種後72h，以大約2E11個基因組拷貝/孔，將在CBA啟動子控制下表達EGFP或啟動物(構建體30，其包含與被CBA啟動子驅動的VPR TAD融合的鋅指DBD)的重組AAV(血清型AAV-DJ)添加到培養基中。

【0324】 感染後一周，從培養的細胞中分離RNA(Qiagen RNeasy Mini試劑盒)，並進行DNA酶處理。使用TruSeq Stranded mRNA文庫試劑盒(Illumina)從回收的RNA製備RNAseq文庫，並在Illumina NextSeq上進行測序(2 x 75個迴圈配對末端測序)。將測序讀取結果與人類基因組(RNASTAR)進行比對，並使用DESeq2進行差異表達分析。資料以相對於對照(AAVDJ-CBA-EGFP)樣品的變化倍數來呈現(參見圖4)。結果示於表14中，圖4示出了內源SCN1A和40個最近鄰基因轉錄物的相對表達，其以相對於對照條件的變化倍數呈現。與檢查的其他基因相比，如表1所述的構建體30能夠特異性增加SCN1A基因或Nav1.1蛋白的表達。這表明，被構建體30的轉錄啟動物識別的靶位點對於SCN1A基因是特異性的，因此導致GABA神經元中SCN1A基因表達的增加。

表14. 用SCN1A特異性轉錄因子(構建體30)處理的GABA神經元中對內源SCN1A和40個最近鄰基因的轉錄的影響。

基因名稱	染色體2起點	染色體2終點	染色體鏈	相對於對照的變化倍數
PLA2R1	160788518	160919121	-	0.16367458
ITGB6	160956176	161128399	-	0.20679884
RBMS1	161128661	161350305	-	1.63514667
TANK	161993418	162092732	+	0.90946407
PSMD14	162164548	162268228	+	0.92699237
TBR1	162272604	162282381	+	0.53199642
SLC4A10	162280842	162841792	+	1.89407328
DPP4	162848750	162931052	-	2.82345284
FAP	163027193	163101661	-	2.26977379
IFIH1	163123588	163175213	-	1.46146481
GCA	163175349	163228105	+	2.58702426
FIGN	164449905	164592522	-	0.46785861
GRB14	165349321	165478358	-	0.5631965
COBLL1	165510133	165700189	-	0.43199257
SLC38A11	165752695	165812035	-	4.06730119
SCN3A	165944031	166060577	-	1.0807866
SCN2A	166095911	166248818	+	1.24475196
CSRNP3	166326156	166545917	+	0.82971233
GALNT3	166604100	166651192	-	0.33804418
TTC21B	166713984	166810353	-	1.58661143
SCN1A	166845669	166984523	-	62.9552975
SCN9A	167051694	167232503	-	1.71659087
SCN7A	167260082	167350757	-	0.29331967
B3GALT1	168675181	168730551	+	0.64436013
STK39	168810529	169104651	-	1.19821739
CERS6	169312371	169631644	+	0.86828378
NOSTRIN	169643048	169722024	+	1.82142718
SPC25	169690641	169769881	-	0.86880697
ABCB11	169779447	169887832	-	3.1441368
DHRS9	169921298	169952677	+	1.10381777
BBS5	170335687	170382432	+	0.65476347
KLHL41	170366211	170382772	+	0.87373377
FASTKD1	170386258	170430385	-	1.02786927

基因名稱	染色體2起點	染色體2終點	染色體鏈	相對於對照的變化倍數
PPIG	170440849	170497916	+	1.09866236
CCDC173	170501934	170550943	-	0.67290779
PHOSPHO2	170550974	170558218	+	0.91339152
KLHL23	170550997	170633499	+	0.73926347
SSB	170648442	170668574	+	1.00631994
METTL5	170666590	170681441	-	1.21271497
UBR3	170683967	170940641	+	1.21350908
MYO3B	171034654	171511681	+	0.52839217

實施例5

SCN1A從表達盒的體內表達

【0325】為了測試SCN1A的轉錄啟動物的體內表達，由Vector Biolabs(Malvern，PA)生成了重組AAV9載體。向雄性C57Bl/6小鼠(每組N=5，7-8周齡)的背側海馬體(AP -2.0 mm，側向±1.5， DV -1.4 mm，距硬腦膜)和腹側海馬體(AP -3.1 mm，側向±2.8， DV -3.8 mm，距硬腦膜)中雙側輸注1.5ul純化的AAV載體，總共4個注射部位。以0.3 ul/分鐘的速率遞送AAV，每次注射後有4m的休息期。治療後四周，對小鼠實施安樂死並解剖海馬組織。對於每組，除了一隻動物僅收集並均質化左側海馬體外，大多數動物(N=4)均收集來自左側和右側海馬組織的組織，併合並以進行均質化。從勻漿中分離RNA(Qiagen RNeasy Mini試劑盒)，並進行DNA酶處理。使用OligoDT引物(Superscript IV，Invitrogen)對RNA(3µg)進行逆轉錄。使用Phusion聚合酶(New England Biolabs)和SYBR Green I，通過qPCR分析cDNA樣品中的小鼠SCN1A表達：98°C下30s，40x [98°C下10 sec，64°C下15 sec，72°C下15 sec]。使用針對小鼠SCN1A的引物(5'-CAAAAAAGCCACAAAAGCCT-3'(SEQ ID NO: 189)；5'-TTAGCTCCGCAAGAAACATC-3'(SEQ ID NO: 190))來量化內

源SCN1A轉錄物的水準，並以GAPDH為參考基因，通過delta-delta Ct法(5'-ACCACAGTCCATGCCATCAC'3'(SEQ ID NO: 187)；5'-TCCACCACCCTGTTGCTGTA-3'(SEQ ID NO: 188))確定體內SCN1A表達的相對水準。

【0326】圖5A和圖5B示出了五隻動物的平均結果，每只動物注射AAV9構建體。eGFP對照構建體包含eGFP報導轉基因。構建體4(參見表1)包含轉錄啟動物，該轉錄啟動物識別包含SEQ ID NO: 18的靶序列，如以上表1中所述。圖5A示出了SCN1A在體內的相對表達。圖5B以平均eGFP表達百分比示出了體內SCN1A表達的變化。這些結果表明，表達盒A的SCN1A轉錄啟動物導致體內SCN1A表達上調約20%-30%。

【0327】此類表達盒可適合用於人類，以治療Dravet綜合症、癲癇、癲癇發作、阿爾茨海默病、帕金森病和/或與SCN1A缺乏和/或活性受損相關的其他任何疾病或病況。

實施例6

在Dravet綜合症小鼠模型中的高熱癲癇發作(HTS)試驗

【0328】**A. 雜合Scn1a敲除小鼠模型**

【0329】在Scn1a^{tm1Kea}小鼠品系中測試了使用表達盒對Dravet綜合症和/或其症狀的治療。該小鼠品系是已確立的Dravet綜合症的小鼠模型。Scn1a^{tm1Kea}小鼠品系不需要CRE重組酶。Scn1a^{tm1Kea}小鼠(可從Jackson Laboratory獲得；在Hawkins等人, Scientific Reports, vol. 7: 15327 (2017)中描述)包含SCN1A的第一個編碼外顯子的缺失。對於SCN1A敲除等位基因而言純合的小鼠的特徵為震顫、共濟失調、癲癇發作，並在出生後第16天死亡。C57BL/6背景的雜合小鼠發展出自發性癲癇

發作，並且大多數在數周內死亡。這樣的小鼠品種可用來研究癲癇和Dravet綜合症治療的安全性和功效。其他資訊參見Miller等人, Genes Brain Behav. 2014年2月;13(2):163-72。

【0330】 為了測試轉錄啟動物在Scn1a^{tm1Kea}小鼠品系中的功效，在P1經雙側ICV，向雄性Scn1a +/-與雌性C57Bl/6J雜交繁殖產生的幼崽給予AAV載體。向小鼠給予構建體31-34(表1)。使母鼠不受干擾，直到在P18斷奶，然後再次使小鼠不受干擾，直到P26-P28開始高熱癲癇發作(HTS)試驗。給藥的P1小鼠的單獨幼崽在P18斷奶，並每天觀察死亡率。在混合129Stac X C57BL/6背景中的P26-P28雜合(HET)和WT Scn1a小鼠中，進行高熱癲癇發作誘發。在試驗之前，向小鼠插入潤滑的直腸溫度探頭(Ret-4)，其連接至與加熱燈(HL-1)串聯連接的溫度控制模組(TCAT 2DF，Physitemp)。然後將小鼠放入一個大玻璃燒杯中，使其短暫適應環境。此後，每2分鐘將體溫升高約0.5°C，直到出現第一次伴有姿勢喪失的強直-陣攣性癲癇發作或達到43°C為止。如果小鼠發生癲癇發作伴姿勢喪失，則結束實驗並記錄小鼠的內部體溫。如果在整個實驗過程中沒有檢測到癲癇發作伴姿勢喪失，則認為該小鼠沒有癲癇發作，並結束試驗。在P1從小鼠獲得組織樣品，並且在實驗過程中使用即時PCR對小鼠進行基因型分型。完成試驗後，對基因型分型進行揭盲，並將小鼠的狀態HET或WT與獲得的資料進行關聯。將資料繪製成Kaplan-Meier存活曲線，並通過Mantel-Cox檢驗確定顯著性。結果示於表15和表16以及圖6A-E中。

表15. 實施例6中使用的條件的總結

構建體	**劑量(gc/小鼠)**
構建體31 (圖6A和6E)	6.0E+10
構建體32 (圖6B、6D、6E)	3.1E+11

構建體	劑量(gc/小鼠)
構建體33 (圖6C)	5.8E+10
構建體34 (圖6D)	4.9E+13

表16. 高熱癲癇發作試驗的結果的總結。

eTF構建體(圖)	對照動物數(PBS處理的)	治療的動物數	在42.6 °C時無癲癇發作的%	P值
EGFP報導分子	16	N/A	44%	
構建體31 (圖6A和6E)	16	18	95	P<0.0001
構建體32 (圖6B、6D、6E)	16	21	76	P<0.05
構建體33 (圖6C)	16	14	93	P<0.001
構建體34 (圖6D)	16	12	83	P<0.05

【0331】 如上所述，在Scn1a^{tm1Kea}小鼠中進行了另外的實驗，以在HTS試驗中測試構建體42和43對癲癇發作的影響。在這些實驗中，構建體42在P1通過雙側ICV以9x10^{10} gc/小鼠給藥，而構建體43在P1或P5通過雙側ICV以6x10^{10} gc/小鼠給藥。結果示於圖6F(構建體42)和圖6G(構建體43)中。與EGFP對照相比，這兩種構建體均顯示癲癇發作顯著減少(對於構建體42和43，均為P < 0.0001)。

【0332】 **B. 雜合Scn1a^{RX}突變小鼠模型**

【0333】 在Scn1a^{RX}小鼠品系中測試了使用本公開的表達盒對Dravet綜合症和/或其症狀的治療。該小鼠品系是已確立的Dravet綜合症的小鼠模型。Scn1a^{RX}小鼠品系不需要CRE重組酶。Scn1a^{RX}小鼠(可從Jackson Laboratory獲得；在Ogiwara等人, J. Neuroscience, vol. 27: 5903-5914 (2007)中描述)包含SCN1A基因外顯子21的功能喪失單堿基無義突變(CgG突變為TgA；R1407X)。C57BL/6背景的雜合小鼠發展出自發性癲癇發作，並且大多數在數周內死亡。

【0334】 為了測試轉錄啟動物在Scn1a^{RX}小鼠品系中的功效，在P1

通過雙側ICV，向通過雄性Scn1a$^{RX/+}$精子與雌性C57Bl/6J卵細胞的IVF雜交並將胚胎植入CD-1母鼠中而獲得的幼崽以5.1x10^{10}個基因組拷貝(gc)/小鼠給予AAV載體(構建體31)或給予PBS對照。使母鼠不受干擾，直到在P18斷奶，然後再次使小鼠不受干擾，直到P26-P28開始HTS試驗。給藥的P1小鼠的單獨幼崽在P18斷奶，並每天觀察死亡率。在C57BL/6背景中的P26-P28雜合(HET)和WT Scn1a小鼠中，進行高熱癲癇發作誘發。在試驗之前，向小鼠插入潤滑的直腸溫度探頭(Ret-4)，其連接至與加熱燈(HL-1)串聯連接的溫度控制模組(TCAT 2DF，Physitemp)。然後將小鼠放入一個大玻璃燒杯中，使其短暫適應環境。此後，每2分鐘將體溫升高約0.5°C，直到出現第一次伴有姿勢喪失的強直-陣攣性癲癇發作或達到43°C為止。如果小鼠發生癲癇發作伴姿勢喪失，則結束實驗並記錄小鼠的內部體溫。如果在整個實驗過程中沒有檢測到癲癇發作伴姿勢喪失，則認為該小鼠沒有癲癇發作，並結束試驗。在P1從小鼠獲得組織樣品，並且在實驗過程中使用即時PCR對小鼠進行基因型分型。完成試驗後，對基因型分型進行揭盲，並將小鼠的狀態HET或WT與獲得的資料進行關聯。測試的WT Scna1小鼠均未出現癲癇發作。將關於構建體31處理的(n=13)和PBS對照處理的(n=14)HET小鼠的資料繪製成Kaplan-Meier存活曲線，並通過Mantel-Cox檢驗確定顯著性。如圖6H所示，與PBS對照處理的HET小鼠相比，構建體31處理的HET小鼠顯示出高熱癲癇發作誘發顯著減少($P < 0.01$)。

實施例7

在Dravet綜合症小鼠模型中的存活試驗

【0335】 **A. 雜合Scn1a敲除小鼠模型**

【0336】 為了測試轉錄啟動物在Scn1a^{tm1Kea}小鼠品系中的功效，在P1經雙側ICV，向雄性Scn1a +/-與雌性C57Bl/6J雜交繁殖產生的幼崽給予AAV載體。使母鼠不受干擾，直到斷奶。在P18斷奶後，每天對Scn1a +/-小鼠的健康狀態進行觀察。對於發現因任何原因死於飼養籠中的小鼠，記錄日期。將資料繪製成Kaplan-Meier存活曲線，並通過Mantel-Cox檢驗確定顯著性。

【0337】 結果示於表17和圖7A-D中。

表17. 存活試驗的條件和結果的總結。

SEQ ID	**劑量(gc/小鼠)**	**對照動物數(PBS處理的)**	**治療的動物數**	**P100時的%存活(*P83時)**	**P值**
PBS	N/A	53	N/A	49%	
構建體33 (圖7C和7D)	5.8E+10	53	29	76%	P<0.05
構建體31 (圖7B和7D)	6.0E+10	53	34	97%	P<0.0001

【0338】 **B. 雜合Scn1a^{RX}突變小鼠模型**

【0339】 為了測試轉錄啟動物在Scn1a^{RX}小鼠品系中的功效，在P1通過雙側ICV，向通過雄性Scn1a$^{RX/+}$精子與雌性C57Bl/6J卵細胞的IVF雜交並將胚胎植入CD-1母鼠中而獲得的幼崽以5.1x10^{10}個基因組拷貝(gc)/小鼠給予AAV載體(構建體31)或給予PBS對照。使母鼠不受干擾，直到斷奶。在P18斷奶後，每天對Scn1a$^{RX/+}$小鼠的健康狀態進行觀察。對於發現因任何原因死於飼養籠中的小鼠，記錄日期。將關於構建體31處理的(n=27)和對照處理的(n=18)小鼠的資料繪製成Kaplan-Meier存活曲線，並通過Mantel-Cox檢驗確定顯著性。

【0340】 如圖7E所示，與PBS對照處理的Scn1a$^{RX/+}$小鼠相比，構建體31處理的Scn1a $^{RX/+}$小鼠具有增加的存活率($P < 0.0001$)。

實施例8

用編碼SCN1A特異性轉錄因子的AAV處理後，非人靈長類動物中的SCN1A轉錄水準

【0341】 該研究使用了2至3歲的雄性食蟹猴(macaca fascicularis)。在加入研究之前，通過基於細胞的中和抗體試驗，針對AAV9的交叉反應性抗體對動物進行預篩選。將表達SCN1A特異性轉錄因子(構建體33)的AAV9或對照在PBS中稀釋，並以1.2E12 gc/動物進行實質內注射。在每個半球中確定了三個不同的立體定位座標以供注射，每只動物六個注射部位。每個部位注射10 ul體積。右半球的注射與左半球的注射對稱。兩隻未經治療的動物用作對照。

【0342】 為了評估Scn1A mRNA表達，進行了逆轉錄，隨後進行qPCR方法。給藥後28天屍檢時，從對照和經治療動物的大腦各個區域(額葉皮質、頂葉皮質、顳葉皮質、枕葉皮質、海馬體、髓質、小腦；各200 mg)採集組織切片，置於RNAlater中，然後冷凍。簡言之，切下30 mg組織，提取RNA(使用Qiagen Rneasy Lipid tissue mini試劑盒，目錄號1023539)，通過逆轉錄(使用Applied Biosystems高容量cDNA逆轉錄試劑盒，目錄號4368814)轉化為cDNA，並使用針對Scn1A和管家基因GAPDH的引物/探針組(Applied Biosystems，目錄號Rh02621745-gI FAM)進行qPCR。

【0343】 針對SCN1A的引物/探針組在下面給出。

表18. 在實施例8中使用的引物序列。

基因	SEQ ID NO	序列(5'-3')	注意
Scn1A	191	CCATGGAACTGGCTCGATTTCAC	正向引物
	192	ATTGGTGGGAGGCCACTGTAT	反向引物

	193	AGGCCTGAAAACCATTGTGGGAGCCCT	探針(FAM)

【0344】 使用比較Ct(ΔCt)法，通過相對定量(RQ)確定每個測試樣品中Scn1A的基因表達。該方法測量靶基因與管家基因之間的Ct差異(ΔCt)，然後將治療樣品的ΔCt值與對照樣品進行比較。

【0345】 ΔCt = 靶基因的Ct平均值 – 管家基因的Ct平均值

【0346】 ΔΔCt = 治療樣品的ΔCt – 對照樣品的ΔCt

【0347】 相對表達(治療樣品)= $2^{-\Delta\Delta Ct}$

【0348】 資料被報告為目標mRNA在來自腦的不同組織切片中的歸一化表達(參見圖8)。如圖8所示，在實質內注射部位近端的大腦部位顯示出最高水準的SCN1A轉錄物表達。

實施例9

用具有PV選擇性啟動子和微小RNA結合位點的AAV處理後，在非人靈長類動物PV神經元中的選擇性轉基因表達

【0349】 該研究使用了六隻絨猴(普通狨(Callithrix jacchus))，它們在加入該研究之前，針對AAV9的交叉反應性抗體進行預篩選。兩隻猴子用含有在EF1alpha啟動子控制下的EGFP轉基因的AAV9處理，兩隻猴子用含有在RE 2(SEQ ID NO: 2)控制下的EGFP轉基因的AAV9處理，而兩隻猴子用含有在RE 2(SEQ ID NO: 2)控制下的EGFP轉基因並且還含有位於EGFP編碼區與聚A尾之間的微小RNA結合位點(SEQ ID NO: 7)的AAV9處理。將AAV9載體在PBS中稀釋，然後通過向每個腦半球的海馬體/內嗅皮質進行三次腦內注射(各2 uL)——每只動物總共6個注射部位——來處理動物。用含有EF1alpha-EGFP的AAV9載體處理的兩隻動物各自接受5.8E+11 gc/動物的總劑量，用含有RE 2-EGFP的AAV9載體處理的兩隻動物各自接受3.0E+11 gc/動物的總劑量，而用含有RE 2 + m1-EGFP的

AAV9載體處理的兩隻動物各自接受2.3E+11 gc/動物的總劑量。

【0350】 使用免疫組織化學來評估小白蛋白(PV)選擇性表達。在給藥後28天屍檢時，收集來自大腦各個區域的組織切片。將漂浮的絨猴腦切片(35um)在4%低聚甲醛中固定，用緩衝液(PBS，3% BSA，3%驢血清，0.3% Triton-X 100，0.2%吐溫-20)封閉，然後用抗GFP(Abcam ab290)、隨後用與Alexa-488(Thermo A21206)綴合的第二抗體染色。隨後是抗PV抗體(Swant)以及與Alexa-647(Thermo A31571)和4',6-二脒基-2-苯基吲哚(DAPI)綴合的第二抗體。封固切片，並使用PerkinElmer Vectra3進行成像。

【0351】 結果示於圖9A-F和圖10A-L中。

實施例10

eTF^{SCN1A}生物分佈

【0352】 該研究的目的是比較當通過單側腦室內(ICV)注射以4.8E+13劑量施用時，eTF^{SCN1A}在青年食蟹猴的中樞神經系統(CNS)中的生物分佈。向每只動物注射含有在GABA選擇性調節元件控制下編碼eTF^{SCN1A}(RE^{GABA}-eTF^{SCN1A})的表達盒的AAV9。將AAV9顆粒在PBS + 0.001%普朗尼克酸(pluronic acid)中配製，並以4.8E+13或8E+13 vg/動物的劑量施用。向每只動物施用2 ml體積的經配製的病毒顆粒。研究設計示於表19中。

【0353】 如表19所示，將二十四月齡的食蟹猴分組。在開始研究之前，使用上述NAb滴度測定來檢測來自動物的血液樣品中針對AAV9的中和抗體滴度水準。選擇抗體結果低或陰性的動物用於研究。使用標準外科手術程式通過ICV注射來施用樣品。解凍後的給藥材料短暫地儲存在濕冰

上，並在臨給藥前使其升溫至室溫。將動物麻醉，進行手術準備，並固定在MRI相容的立體定位架(Kopf)中。進行基線MRI以建立目標座標。製作切口，並在目標位置上方穿過顱骨鑽出一個洞。用樣品填充附接至36”微孔擴展套件上的3 mL BD注射器，並將其置於輸注泵中。準備好延長線。打開硬腦膜，並將給藥針推進至距軟腦膜13.0至18.1 mm的深度。採用造影劑注射和螢光檢查來確認脊椎穿刺針向右側腦室中的放置。在附接準備好的延長線和注射器之前，向3.0” 22g Quinke BD脊椎休伯尖型針填充造影劑，以確定放置。泵設置為0.1 mL/分鐘，持續19至20分鐘。給藥後用手推動緩衝液以清除延長線。輸注完成後，將針頭留在原處1至2分鐘，然後將針頭取出。在第1天施用媒介物和測試物一次，並使受試者維持27或29天的恢復期。

表19. 生物分佈研究設計

組	性別	ID	劑量(VG/動物)
組1 (緩衝液對照)	M	21001	
	F	11501	
組2 (RE^{GABA}-eTF^{SCN1A})	M	2001	4.8E+13
	F	2501	4.8E+13
	M	3001	4.8E+13
	M	3002	8E+13

【0354】 給藥後，在整個研究持續期間常規監測動物，並定期抽取血液樣品。eTF^{SCN1A}施用與任何意外死亡、臨床發現或宏觀觀察結果均無關。經AAV9-RE^{GABA}-eTF^{SCN1A}處理的動物直到第28 ± 2天計畫屍檢時仍存活。在每日或每週體格檢查中未觀察到臨床或行為體徵、體溫升高或體重減輕。在AAV9-RE^{GABA}-eTF^{SCN1A}處理的動物中觀察到肝臟轉氨酶(ALT和AST)暫態升高，但未經免疫調節到研究結束時完全恢復，並且血

清膽紅素或鹼性磷酸酶未見伴隨升高。測量的其他臨床化學終點均不顯著。在肝臟組織病理學研究中未報告顯微鏡觀察結果。相對於治療前的值，CSF白細胞在終末收集物中升高，但在對照與AAV9-RE^{GABA}-eTF^{SCN1A}處理的動物之間相當。沒有觀察到與AAV9-RE^{GABA}-eTF^{SCN1A}相關的腦脊液細胞增多。在所有動物中進行的宏觀觀察和非神經元組織的詳細顯微組織病理學檢查均不顯著。組織包括主要的周圍器官(即心臟、肺、脾臟、肝臟和性腺)。宏觀觀察和神經元組織的詳細顯微組織病理學未顯示任何顯著發現。組織包括腦、脊髓和相關的背根神經節(起自頸區、胸區和腰區)。研究由三名獨立的病理學家進行，其中包括一名在專門神經病理學場所的病理學家。

【0355】 由於在給藥後四周觀察到抗AAV9衣殼中和抗體，因此AAV9的ICV施用不能防止血清中的給藥後免疫應答。然而，CSF中的抗AAV9中和抗體水準保持不變，並且與給藥前水準相當(表20)。

表20. AAV9血清NAb滴度

受試者編號	AAV9血清NAb滴度			AAV9 CSF NAb滴度	
	預篩選	注射時	注射後4周	注射時	注射後4周
21001	1:5	<1:5	<1:5	<1:5	<1:5
11501	<1:5	<1:5	<1:5	<1:5	<1:5
2001	<1:5	<1:5	1:405	<1:5	1:5
2501	<1:5	<1:5	1:135	<1:5	1:5
3001	<1:5	<1:5	1:1215	<1:5	<1:5
3002	<1:5	<1:5	1:135	<1:5	<1:5

【0356】 給藥後27-29天，在計畫的屍檢期間，從主要器官(心室、肝葉、肺心葉、腎臟、脾臟、胰腺和頸淋巴結)收集樣品。通過八毫米的沖頭收集鑽取物，並如下所述進一步處理。

實施例11

eTF^{SCN1A}在腦中的生物分佈

【0357】使用ddPCR來測量eTF^{SCN1A}在腦中的生物分佈。測量了來自食蟹猴腦組織各個區域(FC：額葉皮質；PC：頂葉皮質；TC：顳葉皮質；Hip：海馬體；Med：髓質；OC：枕葉皮質)的樣品的載體拷貝數，以評估當通過單側ICV在AAV9中施用時，在GABA選擇性調節元件控制下的eTF^{SCN1A}(RE^{GABA}-eTF^{SCN1A})的生物分佈。用DNeasy Blood & Tissues試劑盒(Qiagen)分離組織DNA。測定DNA量並使用紫外分光光度計進行歸一化。將20納克組織DNA與ddPCR Super Mix for Probes(無dUTP)(Bio-Rad)以及針對eTF^{SCN1A}序列區域的TaqMan引物和探針一起添加到20微升反應中。生成小液滴，並使用自動化小液滴發生器和熱迴圈儀(Bio-Rad)擴增範本。PCR步驟後，將板載入並通過QX2000 Droplet Reader讀取，以確定組織中的載體拷貝數。猴白蛋白(MfAlb)基因充當用於基因組DNA含量歸一化的內部對照，並在相同反應中擴增。表21中列出了針對eTF^{SCN1A}和MfAlb的引物和探針。

表21. 針對eTF^{SCN1A}和MfAlb的引物和探針

引物/探針名稱		**序列(5'-3')**
eTF^{SCN1A}	eTF^{SCN1A}正向引物	GAATGTGGGAAATCATTCAGTCGC (SEQ ID NO: 194)
	eTF^{SCN1A}反向引物	GCAAGTTATCCTCTCGTGAGAAGG (SEQ ID NO: 195)
	eTF^{SCN1A}探針	GCGACAACCTGGTGAGACATCAACGCACC (SEQ ID NO: 196)
MfAlb	MfAlb正向引物	GCTGTTATCTCTTGTGGGCTGT (SEQ ID NO: 197)
	MfAlb反向引物	AAACTCATGGGAGCTGCCGGTT (SEQ ID NO: 198)
	MfAlb探針	CCACACAAATCTCTCCCTGGCATTG (SEQ ID NO: 199)

【0358】當以1.3-3.5 VG/二倍體基因組的平均值，以每只動物4.8E+13個病毒基因組給藥時，eTF^{SCN1A}廣泛分佈於整個腦中(圖11)。另外，當比較以每只動物4.8E+13個病毒基因組給藥的RE^{GABA}-eTF^{SCN1A}在整個腦中的基因轉移與以不同劑量通過ICV給藥的eGFP在整個腦中的基因轉移時，觀察到VG/二倍體基因組隨劑量增加而增加。這表明，當通過ICV在AAV9中施用時，腦中的基因轉移以劑量依賴性方式發生。

實施例12

腦中的eTF^{SCN1A}轉錄

【0359】通過使用基於ddPCR的基因表達測定測量eTF^{SCN1A} mRNA，來評估在GABA選擇性調節元件RE^{GABA}控制下的eTF^{SCN1A}(RE^{GABA}-eTF^{SCN1A})的轉錄。用針對腦組織的RNeasy Plus Mini試劑盒(Qiagen)或RNeasy Lipid Tissue Mini試劑盒(Qiagen)分離組織RNA。確定RNA量，並使用紫外分光光度計進行歸一化，並使用Bioanalyzer RNA Chip檢查RNA品質(RIN)。採用SuperScript VILO cDNA合成試劑盒和ezDNase™ Enzyme試劑盒(Thermo Fisher)，用一微克組織RNA進行DNA酶處理和cDNA合成。將50微克RNA轉化為cDNA。將cDNA與ddPCR Super Mix for Probes(無dUTP)(Bio-Rad)以及針對eTF^{SCN1A}序列區域的TaqMan引物和探針(表22)一起添加到20微升反應中。生成小液滴，並使用自動化小液滴發生器和熱迴圈儀(Bio-Rad)擴增範本。PCR擴增後，將板載入並通過QX2000 Droplet Reader讀取，以提供組織中的基因表達水準。猴基因*ARFGAP2*(Mf*ARFGAP2*)(Thermo Fisher Scientific)充當用於基因表達水準歸一化的內源對照，並在相同的反應中擴增。*ARFGAP2*的平均轉錄物為1.85E+6/ug RNA(圖12，上邊

界)。下邊界表示檢測極限。

【0360】 在所有動物的整個腦中均觀察到eTF^{SCN1A} mRNA，表明對於所有經AAV9-RE^{GABA}-eTF^{SCN1A}處理的獼猴，GABA選擇性啟動子RE^{GABA}在腦組織中具有轉錄活性(圖12)。FC：額葉皮質；PC：頂葉皮質；TC：顳葉皮質；Hip：海馬體；Med：髓質；OC：枕葉皮質。

表22. 針對eTF^{SCN1A}序列區域的TaqMan引物和探針

引物/探針名稱	**描述**	**序列(5'-3')**
eTF^{SCN1A}	eTF^{SCN1A}正向引物	GAATGTGGGAAATCATTCAGTCGC (SEQ ID NO: 200)
	eTF^{SCN1A}反向引物	GCAAGTTATCCTCTCGTGAGAAGG (SEQ ID NO: 201)
	eTF^{SCN1A}探針	GCGACAACCTGGTGAGACATCAACGCACC (SEQ ID NO: 202)
Mf*ARFGAP2*	正向引物，反向引物，探針	Thermo Fisher (目錄號：4448491)

實施例13

eTF^{SCN1A}在周圍組織中的生物分佈和轉錄

【0361】 在各個器官中進一步測量載體拷貝數，以評估當通過單側ICV在AAV9中施用時，RE^{GABA}-eTF^{SCN1A}在全身組織中的轉導。還通過ddPCR測量了eTF^{SCN1A}的轉錄水準，以評估當通過單側ICV在AAV9中施用時，在GABA選擇性調節元件RE^{GABA}控制下的eTF^{SCN1A}在全身組織中的轉錄活性。這兩種方法均如以上所概述的那樣進行。在脊髓(SC)和背根神經節(DRG)中的RE^{GABA}-eTF^{SCN1A}轉導和eTF^{SCN1A}轉錄與在腦中觀察到的水準相當。除肝臟外，RE^{GABA}-eTF^{SCN1A}轉導在腦外的周圍組織中較低(圖13)。肝臟中的RE^{GABA}-eTF^{SCN1A}轉導高於腦中。在包括心臟、肺和性腺在內的周圍組織中未檢測到eTF^{SCN1A}的轉錄。然而，肝臟中的eTF^{SCN1A}

轉錄物水準與腦中測得的eTF^{SCN1A}水準相當。此外，當相對於存在的載體拷貝數進行歸一化時，肝臟中的eTF^{SCN1A}轉錄極低(與腦中eTF^{SCN1A}的轉錄相比，大約為1/1000)。總之，這證明在GABA選擇性調節元件RE^{GABA}控制下的eTF^{SCN1A}的轉錄局限於CNS。

【序列表】

<110> 美商編碼製藥公司 (ENCODED THERAPEUTICS, INC.)

<120> 用於選擇性基因調節的組合物和方法

<140> TW109118224
<141> 2020-05-29

<150> US63/008,569
<151> 2020-04-10

<150> US62/857,727
<151> 2019-06-05

<150> US62/854,238
<151> 2019-05-29

<160> 226

<170> PatentIn version 3.5

<210> 1
<211> 259
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 1
gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt 60

ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac 120

tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg 180

tgggaggtct atataagcag agctggtacc gtgtgtatgc tcaggggctg ggaaaggagg 240

ggagggagct ccggctcag 259

<210> 2
<211> 2051
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 2

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttte cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct g 2051

<210> 3
<211> 1878
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 3
tcaacagggg gacacttggg aaagaaggat ggggacagag ccgagaggac tgttacacat 60

tagagaaaca tcagtgactg tgccagcttt ggggtagact gcacaaaagc cctgaggcag 120

cacaggcagg atccagtctg ctggtcccag gaagctaacc gtctcagaca gagcacaaag 180

caccgagaca tgtgccacaa ggcttgtgta gagaggtcag aggacagcgt acaggtccca 240

gagatcaaac tcaacctcac caggcttggc agcaagcctt taccaaccca cccccacccc 300

acccaccctg cacgcgcccc tctcccctcc ccatggtctc ccatggctat ctcacttggc 360

cctaaaatgt ttaaggatga cactggctgc tgagtggaaa tgagacagca gaagtcaaca 420

gtagatttta ggaaagccag agaaaaaggc ttgtgctgtt tttagaaagc caagggacaa 480

gctaagatag ggcccaagta atgctagtat ttacatttat ccacacaaaa cggacgggcc 540

tccgctgaac cagtgaggcc ccagacgtgc gcataaataa cccctgcgtg ctgcaccacc 600

tggggagagg gggaggacca cggtaaatgg agcgagcgca tagcaaaagg gacgcggggt 660

ccttttctct gccggtggca ctgggtagct gtggccaggt gtggtacttt gatggggccc 720

agggctggag ctcaaggaag cgtcgcaggg tcacagatct gggggaaccc cggggaaaag 780

cactgaggca aaaccgccgc tcgtctccta caatatatgg gaggggggagg ttgagtacgt 840

tctggattac tcataagacc ttttttttttt ccttccgggc gcaaaaccgt gagctggatt 900

tataatcgcc ctataaagct ccagaggcgg tcaggcacct gcagaggagc cccgccgctc 960

cgccgactag ctgcccccgc gagcaacggc ctcgtgattt ccccgccgat ccggtccccg 1020

cctccccact ctgcccccgc ctaccccgga gccgtgcagc cgcctctccg aatctctctc 1080

ttctcctggc gctcgcgtgc gagagggaac tagcgagaac gaggaagcag ctggaggtga 1140

cgccgggcag attacgcctg tcagggccga gccgagcgga tcgctgggcg ctgtgcagag 1200

gaaaggcggg agtgcccggc tcgctgtcgc agagccgagg tgggtaagct agcgaccacc 1260

tggacttccc agcgcccaac cgtggctttt cagccaggtc ctctcctccc gcggcttctc 1320

aaccaacccc atcccagcgc cggccaccca acctcccgaa atgagtgctt cctgccccag 1380

cagccgaagg cgctactagg aacggtaacc tgttacttttt ccaggggccg tagtcgaccc 1440

gctgcccgag ttgctgtgcg actgcgcgcg cggggctaga gtgcaaggtg actgtggttc 1500

ttctctggcc aagtccgagg gagaacgtaa agatatgggc ctttttcccc ctctcacctt 1560

gtctcaccaa agtccctagt ccccggagca gttagcctct ttctttccag ggaattagcc 1620

agacacaaca acgggaacca gacaccgaac cagacatgcc cgccccgtgc gccctccccg 1680

ctcgctgcct ttcctccctc ttgtctctcc agagccggat cttcaagggg agcctccgtg 1740

cccccggctg ctcagtccct ccggtgtgca ggaccccgga agtcctcccc gcacagctct 1800

cgcttctctt tgcagcctgt ttctgcgccg gaccagtcga ggactctgga cagtagaggc 1860

cccgggacga ccgagctg 1878

<210> 4
<211> 509
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 4
gccctctagg ccacctgacc aggtcccctc agtccccccc ttcccacact cccacactca 60

gccccccctcc ccccccccccg acccctgcag gattatcctg tctgtgttcc tgactcagcc 120

tgggagccac ctgggcagca ggggccaagg gtgtcctaga agggacctgg agtccacgct 180

gggccaagcc tgccctttct ccctctgtct tccgtccctg cttgcggttc tgctgaatgt 240

ggttatttct ctggctcctt ttacagagaa tgctgctgct aattttatgt ggagctctga 300

ggcagtgtaa ttggaagcca gacaccctgt cagcagtggg ctcccgtcct gagctgccat 360

gcttcctgct ctcctcccgt cccggctcct catttcatgc agccacctgt cccagggaga 420

gaggagtcac ccaggccccct cagtccgccc cttaaataag aaagcctccg ttgctcggca 480

cacataccaa gcagccgctg gtgcaatct 509

<210> 5
<211> 1644
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 5
cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt 60

gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca 120

atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc 180

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta 240

catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac 300

catgggtcga ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac 360

ccccaatttt gtatttattt atttttaat tattttgtgc agcgatgggg gcgggggggg 420

gggggcgcgg cgccaggcgg ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga 480

ggtgcggcgg cagccaatca gagcggcgcg ctccgaaagt ttccttttat ggcgaggcgg 540

cggcggcggc ggccctataa aaagcgaagc gcgcggcggg cgggagtcgc tgcgttgcct 600

tcgccccgtg ccccgctccg cgccgcctcg cgccgcccgc cccggctctg actgaccgcg 660

ttactcccac aggtgagcgg gcgggacggc ccttctcctc cgggctgtaa ttagcgcttg 720

gtttaatgac ggctcgtttc ttttctgtgg ctgcgtgaaa gccttaaagg gctccgggag 780

ggccctttgt gcgggggggga gcggctcggg gggtgcgtgc gtgtgtgtgt gcgtggggag 840

cgccgcgtgc ggcccgcgct gcccggcggc tgtgagcgct gcgggcgcgg cgcggggctt 900

tgtgcgctcc gcgtgtgcgc gaggggagcg cggccggggg cggtgccccg cggtgcgggg 960

gggctgcgag gggaacaaag gctgcgtgcg gggtgtgtgc gtgggggggt gagcaggggg 1020

tgtgggcgcg gcggtcgggc tgtaaccccc ccctgcaccc ccctccccga gttgctgagc 1080

acggcccggc ttcgggtgcg gggctccgtg cggggcgtgg cgcggggctc gccgtgccgg 1140

gcgggggggtg gcggcaggtg ggggtgccgg gcgggggcggg gccgcctcgg gccgggggagg 1200

gctcgggggga ggggcgcggc ggccccggag cgccggcggc tgtcgaggcg cggcgagccg 1260

cagccattgc cttttatggt aatcgtgcga gagggcgcag ggacttcctt tgtcccaaat 1320

ctggcggagc cgaaatctgg gaggcgccgc cgcacccccct ctagcgggcg cgggcgaagc 1380

ggtgcggcgc cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc 1440

gtcccccttct ccatctccag cctcggggct gccgcagggg gacggctgcc ttcgggggggg 1500

acggggcagg gcggggttcg gcttctggcg tgtgaccggc ggctctagag cctctgctaa 1560

ccatgttcat gccttcttct ttttcctaca gctcctgggc aacgtgctgg ttgttgtgct 1620

gtctcatcat tttggcaaag aatt 1644

<210> 6
<211> 1335
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 6

gagtaattca tacaaaagga ctcgcccctg ccttggggaa tcccagggac cgtcgttaaa 60

ctcccactaa cgtagaaccc agagatcgct gcgttcccgc cccctcaccc gcccgctctc 120

gtcatcactg aggtggagaa gagcatgcgt gaggctccgg tgcccgtcag tgggcagagc 180

gcacatcgcc cacagtcccc gagaagttgg ggggaggggt cggcaattga accggtgcct 240

agagaaggtg gcgcggggta aactgggaaa gtgatgtcgt gtactggctc cgccttttc 300

ccgagggtgg gggagaaccg tatataagtg cagtagtcgc cgtgaacgtt ctttttcgca 360

acgggtttgc cgccagaaca caggtaagtg ccgtgtgtgg ttcccgcggg cctggcctct 420

ttacgggtta tggcccttgc gtgccttgaa ttacttccac gcccctggct gcagtacgtg 480

attcttgatc ccgagcttcg ggttggaagt gggtgggaga gttcgaggcc ttgcgcttaa 540

ggagcccctt cgcctcgtgc ttgagttgag gcctggcttg ggcgctgggg ccgccgcgtg 600

cgaatctggt ggcaccttcg cgcctgtctc gctgctttcg ataagtctct agccatttaa 660

aatttttgat gacctgctgc gacgcttttt ttctggcaag atagtcttgt aaatgcgggc 720

caagatctgc acactggtat ttcggttttt ggggccgcgg gcggcgacgg ggcccgtgcg 780

tcccagcgca catgttcggc gaggcggggc ctgcgagcgc ggccaccgag aatcggacgg 840

gggtagtctc aagctggccg gcctgctctg gtgcctggcc tcgcgccgcc gtgtatcgcc 900

ccgccctggg cggcaaggct ggcccggtcg gcaccagttg cgtgagcgga aagatggccg 960

cttcccggcc ctgctgcagg gagctcaaaa tggaggacgc ggcgctcggg agagcgggcg 1020

ggtgagtcac ccacacaaag gaaaagggcc tttccgtcct cagccgtcgc ttcatgtgac 1080

tccacggagt accgggcgcc gtccaggcac ctcgattagt tctcgagctt ttggagtacg 1140

tcgtctttag gttgggggga ggggttttat gcgatggagt ttccccacac tgagtgggtg 1200

gagactgaag ttaggccagc ttggcacttg atgtaattct ccttggaatt tgccctttt 1260

gagtttggat cttggttcat tctcaagcct cagacagtgg ttcaaagttt ttttcttcca 1320

tttcaggtgt cgtga 1335

<210> 7
<211> 214
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 7
aaagagaccg gttcactgtg acagtaaaag agaccggttc actgtgagaa tgaaagagac 60

cggttcactg tgatcggaaa agagaccggt tcactgtgag cggccttgaa acccagcaga 120

caatgtagct cagtagaaac ccagcagaca atgtagctga atggaaaccc agcagacaat 180

gtagcttcgg agaaacccag cagacaatgt agct 214

<210> 8
<211> 21
<212> RNA
<213> 智人（Homo sapiens）

<400> 8
ucacagugaa ccggucucuu u 21

<210> 9
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成寡核苷酸

<400> 9
aaagagaccg gttcactgtg a 21

<210> 10
<211> 23
<212> RNA
<213> 智人（Homo sapiens）

<400> 10
agcuacauug ucugcugggu uuc 23

<210> 11
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成寡核苷酸

<400> 11
gaaacccagc agacaatgta gct 23

<210> 12
<211> 23
<212> RNA
<213> 智人（Homo sapiens）

<400> 12
agcuacaucu ggcuacuggg ucu 23

<210> 13
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成寡核苷酸

<400> 13
agacccagta gccagatgta gct 23

<210> 14
<211> 67
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成寡核苷酸

<400> 14
gaaacccagc agacaatgta gctagaccca gtagccagat gtagctaaag agaccggttc 60

actgtga 67

<210> 15
<211> 134
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 15
gaaacccagc agacaatgta gctagaccca gtagccagat gtagctaaag agaccggttc 60

actgtgagaa acccagcaga caatgtagct agacccagta gccagatgta gctaaagaga 120

ccggttcact gtga 134

<210> 16
<211> 62
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成寡核苷酸

<400> 16
aataaaagat ctttattttc attagatctg tgtgttggtt ttttgtgtgc ggaccgcacg 60

tg 62

<210> 17
<211> 477
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 17
gggtggcatc cctgtgaccc ctccccagtg cctctcctgg ccctggaagt tgccactcca 60

gtgcccacca gccttgtcct aataaaatta agttgcatca ttttgtctga ctaggtgtcc 120

ttctataata ttatggggtg gagggggggtg gtatggagca aggggcaagt tgggaagaca 180

acctgtaggg cctgcggggt ctattgggaa ccaagctgga gtgcagtggc acaatcttgg 240

ctcactgcaa tctccgcctc ctgggttcaa gcgattctcc tgcctcagcc tcccgagttg 300

ttgggattcc aggcatgcat gaccaggctc agctaatttt tgtttttttg gtagagacgg 360

ggtttcacca tattggccag gctggtctcc aactcctaat ctcaggtgat ctacccacct 420

tggcctccca aattgctggg attacaggcg tgaaccactg ctcccttccc tgtcctt 477

<210> 18
<211> 18

<212> DNA
<213> 智人（Homo sapiens）

<400> 18
ctaggtcaag tgtaggag 18

<210> 19
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 19
acttgaccta gacagcct 18

<210> 20
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 20
tgaataactc attagtga 18

<210> 21
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 21
aaagtacatt aggctaat 18

<210> 22
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 22
ccagcactgg tgcttcgt 18

<210> 23
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 23
aaggctgtct aggtcaag 18

<210> 24
<211> 24
<212> DNA
<213> 智人（Homo sapiens）

<400> 24
ctaggtcaag tgtaggagac acac 24

<210> 25
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 25
ggtcaagtgt aggagaca 18

<210> 26
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 26
caagtgtagg agacacac 18

<210> 27
<211> 24
<212> DNA
<213> 智人（Homo sapiens）

<400> 27
agtgtaggag acacactgct ggcc 24

<210> 28
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 28
agtgtaggag acacactg 18

<210> 29
<211> 21
<212> DNA

<213> 智人（Homo sapiens）

<400> 29
aggagacaca ctgctggcct g 21

<210> 30
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 30
taggtaccat agagtgag 18

<210> 31
<211> 18
<212> DNA
<213> 智人（Homo sapiens）

<400> 31
gaggatactg cagaggtc 18

<210> 32
<211> 27
<212> DNA
<213> 智人（Homo sapiens）

<400> 32
taggtaccat agagtgaggc gaggatg 27

<210> 33
<211> 27
<212> DNA
<213> 智人（Homo sapiens）

<400> 33
atagagtgag gcgaggatga agccgag 27

<210> 34
<211> 27
<212> DNA
<213> 智人（Homo sapiens）

<400> 34
tgaagccgag aggatactgc agaggtc 27

<210> 35
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 35
aaggctgtct aggtcaagtg t 21

<210> 36
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 36
tgttcctcca gattaacact t 21

<210> 37
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 37
gatgaagccg agaggatact g 21

<210> 38
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 38
gctgatttgt attaggtacc a 21

<210> 39
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 39
agaaagctga tacagataca a 21

<210> 40
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 40
ggtacgggca aagatttctt g 21

<210> 41
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 41
agaaagctga tacagataca a 21

<210> 42
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 42
acacaatgag ccacctacaa g 21

<210> 43
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 43
gtggctcatt gtgtgtgtgc c 21

<210> 44
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 44
catatccctg caggttcaga a 21

<210> 45
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 45
agagagagag agagagagag a 21

<210> 46
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 46
ttctcagttt tgaaattaaa a 21

<210> 47
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 47
tggattctct tctgaacctg c 21

<210> 48
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 48
tgctgaggca ggacacagtg t 21

<210> 49
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 49
atcatctgta accatcaagg a 21

<210> 50
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 50
tcctgcctac ttagtttcaa g 21

<210> 51
<211> 21
<212> DNA
<213> 智人 (Homo sapiens)

<400> 51
attacagttc tgtcagcatg c 21

<210> 52
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 52
tggtctcatt ctttttgtgg g 21

<210> 53
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 53
cgatattttc atggattcct t 21

<210> 54
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 54
ctgacactta ctttgtctaa a 21

<210> 55
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 55
aaaactggaa ccgcattccc a 21

<210> 56
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 56
acaaagtaag tgtcagtgtg g 21

<210> 57

<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 57
ataatagttg tgtctttata a 21

<210> 58
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 58
tgtacaagca gggctgcaaa g 21

<210> 59
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 59
gttaacaaat acactaaaca c 21

<210> 60
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 60
ttcaacaagc tcccaagaag t 21

<210> 61
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 61
atgttcaagg tgcagaagga a 21

<210> 62
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 62

tgtttgctca aacgtgcacc a 21

<210> 63
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 63
aaataagaca tgaaaacaag a 21

<210> 64
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 64
aaatatgtac cacaagaaat g 21

<210> 65
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 65
tatctggttt ctctcactgc t 21

<210> 66
<211> 21
<212> DNA
<213> 智人（Homo sapiens）

<400> 66
attgcaaagc ataatttgga t 21

<210> 67
<211> 3585
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 67
ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttttt ttttttttttt ttttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttte ccccteteac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760

tgggttctga tgccctcgat gactttgacc tggatatgtt gggaagcgac gcattggatg 2820

actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880

aaactagtaa agagaccggt tcactgtgac agtaaaagag accggttcac tgtgagaatg 2940

aaagagaccg gttcactgtg atcggaaaag agaccggttc actgtgagcg gccttgaaac 3000

ccagcagaca atgtagctca gtagaaaccc agcagacaat gtagctgaat ggaaacccag 3060

cagacaatgt agcttcggag aaacccagca gacaatgtag ctaagcttgg gtggcatccc 3120

tgtgacccct ccccagtgcc tctcctggcc ctggaagttg ccactccagt gcccaccagc 3180

cttgtcctaa taaaattaag ttgcatcatt ttgtctgact aggtgtcctt ctataatatt 3240

atggggtgga ggggggtggt atggagcaag gggcaagttg ggaagacaac ctgtagggcc 3300

tgcggggtct attgggaacc aagctggagt gcagtggcac aatcttggct cactgcaatc 3360

tccgcctcct gggttcaagc gattctcctg cctcagcctc ccgagttgtt gggattccag 3420

gcatgcatga ccaggctcag ctaatttttg tttttttggt agagacgggg tttcaccata 3480

ttggccaggc tggtctccaa ctcctaatct caggtgatct acccaccttg gcctcccaaa 3540

ttgctgggat tacaggcgtg aaccactgct cccttccctg tcctt 3585

<210> 68
<211> 3371
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 68
ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttte cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760

tgggttctga tgccctcgat gactttgacc tggatatgtt gggaagcgac gcattggatg 2820

actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880

aaactagtaa gcttgggtgg catccctgtg acccctcccc agtgcctctc ctggccctgg 2940

aagttgccac tccagtgccc accagccttg tcctaataaa attaagttgc atcattttgt 3000

ctgactaggt gtccttctat aatattatgg ggtggagggg ggtggtatgg agcaaggggc 3060

aagttgggaa gacaacctgt agggcctgcg gggtctattg ggaaccaagc tggagtgcag 3120

tggcacaatc ttggctcact gcaatctccg cctcctgggt tcaagcgatt ctcctgcctc 3180

agcctcccga gttgttggga ttccaggcat gcatgaccag gctcagctaa tttttgtttt 3240

tttggtagag acggggtttc accatattgg ccaggctggt ctccaactcc taatctcagg 3300

tgatctaccc accttggcct cccaaattgc tgggattaca ggcgtgaacc actgctccct 3360

tccctgtcct t 3371

<210> 69
<211> 4380
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 69

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttte cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg aggccagcgg ttccggacgg gctgacgcat 2760

tggacgattt tgatctggat atgctgggaa gtgacgccct cgatgatttt gaccttgaca 2820

tgcttggttc ggatgcccct gatgactttg acctcgacat gctcggcagt gacgcccttg 2880

atgatttcga cctggacatg ctgattaact ctagaagttc cggatctccg aaaaagaaac 2940

gcaaagttgg tagccagtac ctgcccgaca ccgacgaccg gcaccggatc gaggaaaagc 3000

ggaagcggac ctacgagaca ttcaagagca tcatgaagaa gtccccccttc agcggcccca 3060

ccgaccctag acctccacct agaagaatcg ccgtgcccag cagatccagc gccagcgtgc 3120

caaaacctgc cccccagcct taccccttca ccagcagcct gagcaccatc aactacgacg 3180

agttccctac catggtgttc cccagcggcc agatctctca ggcctctgct ctggctccag 3240

cccctcctca ggtgctgcct caggctcctg ctcctgcacc agctccagcc atggtgtctg 3300

cactggctca ggcaccagca cccgtgcctg tgctggctcc tggacctcca caggctgtgg 3360

ctccaccagc ccctaaacct acacaggccg gcgagggcac actgtctgaa gctctgctgc 3420

agctgcagtt cgacgacgag gatctgggag ccctgctggg aaacagcacc gatcctgccg 3480

tgttcaccga cctggccagc gtggacaaca gcgagttcca gcagctgctg aaccagggca 3540

tccctgtggc ccctcacacc accgagccca tgctgatgga ataccccgag gccatcaccc 3600

ggctcgtgac aggcgctcag aggcctcctg atccagctcc tgcccctctg ggagcaccag 3660

gcctgcctaa tggactgctg tctggcgacg aggacttcag ctctatcgcc gatatggatt 3720

tctcagcctt gctgggctct ggcagcggca gccgggattc cagggaaggg atgtttttgc 3780

cgaagcctga ggccggctcc gctattagtg acgtgtttga gggccgcgag gtgtgccagc 3840

caaaacgaat ccggccattt catcctccag gaagtccatg ggccaaccgc ccactccccg 3900

ccagcctcgc accaacacca accggtccag tacatgagcc agtcgggtca ctgaccccgg 3960

caccagtccc tcagccactg gatccagcgc ccgcagtgac tcccgaggcc agtcacctgt 4020

tggaggatcc cgatgaagag acgagccagg ctgtcaaagc ccttcgggag atggccgata 4080

ctgtgattcc ccagaaggaa gaggctgcaa tctgtggcca aatggacctt tcccatccgc 4140

ccccaagggg ccatctggat gagctgacaa ccacacttga gtccatgacc gaggatctga 4200

acctggactc acccctgacc ccggaattga acgagattct ggataccttc ctgaacgacg 4260

agtgcctctt gcatgccatg catatcagca caggactgtc catcttcgac acatctctgt 4320

tttaaactag taataaaaga tctttatttt cattagatct gtgtgttggt tttttgtgtg 4380

<210> 70

<211> 3332
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 70
ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttte ccccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatccgacg 2700

cgctggacga tttcgatctc gacatgctgg gttctgatgc cctcgatgac tttgacctgg 2760

atatgttggg aagcgacgca ttggatgact ttgatctgga catgctcggc tccgatgctc 2820

tggacgattt cgatctcgat atgttataaa agcttgggtg gcatccctgt gacccctccc 2880

cagtgcctct cctggccctg gaagttgcca ctccagtgcc caccagcctt gtcctaataa 2940

aattaagttg catcattttg tctgactagg tgtccttcta taatattatg gggtggaggg 3000

gggtggtatg gagcaagggg caagttggga agacaacctg tagggcctgc ggggtctatt 3060

gggaaccaag ctggagtgca gtggcacaat cttggctcac tgcaatctcc gcctcctggg 3120

ttcaagcgat tctcctgcct cagcctcccg agttgttggg attccaggca tgcatgacca 3180

ggctcagcta atttttgttt ttttggtaga gacggggttt caccatattg gccaggctgg 3240

tctccaactc ctaatctcag gtgatctacc caccttggcc tcccaaattg ctgggattac 3300

aggcgtgaac cactgctccc ttccctgtcc tt 3332

<210> 71
<211> 3546
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 71
ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgccccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatccgacg 2700

cgctggacga tttcgatctc gacatgctgg gttctgatgc cctcgatgac tttgacctgg 2760

atatgttggg aagcgacgca ttggatgact ttgatctgga catgctcggc tccgatgctc 2820

tggacgattt cgatctcgat atgttataaa aagagaccgg ttcactgtga cagtaaaaga 2880

gaccggttca ctgtgagaat gaaagagacc ggttcactgt gatcggaaaa gagaccggtt 2940

cactgtgagc ggccttgaaa cccagcagac aatgtagctc agtagaaacc cagcagacaa 3000

tgtagctgaa tggaaaccca gcagacaatg tagcttcgga gaaacccagc agacaatgta 3060

gctaagcttg ggtggcatcc ctgtgacccc tccccagtgc ctctcctggc cctggaagtt 3120

gccactccag tgcccaccag ccttgtccta ataaaattaa gttgcatcat tttgtctgac 3180

taggtgtcct tctataatat tatggggtgg aggggggtgg tatggagcaa ggggcaagtt 3240

gggaagacaa cctgtagggc ctgcggggtc tattgggaac caagctggag tgcagtggca 3300

caatcttggc tcactgcaat ctccgcctcc tgggttcaag cgattctcct gcctcagcct 3360

cccgagttgt tgggattcca ggcatgcatg accaggctca gctaattttt gtttttttgg 3420

tagagacggg gtttcaccat attggccagg ctggtctcca actcctaatc tcaggtgatc 3480

tacccacctt ggcctcccaa attgctggga ttacaggcgt gaaccactgc tcccttccct 3540

gtcctt 3546

<210> 72
<211> 1707
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 72
atggccgcag atcacctgat gctggctgaa ggctacagac tggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgccctg agaacactgc ccccctacgc cggccctggt 120

cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180

ggagccctgg cctacggcgc cttcggccct ccttctagct tccagccctt ccccgccgtg 240

cctcctccag ccgctggcat cgcccacctg cagcctgtgg ccacccctta ccccggaaga 300

gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360

gctgcccctc cacctcctgc tcatgccctg ggcggcatgg acgccgagct gatcgacgag 420

gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480

ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540

ggctctgtgt cctgcgccga ccacctgatg ctcgccgagg gctaccgcct ggtgcagagg 600

ccgccgtccg ccgccgccgc ccatggccct catgcgctcc ggactctgcc gccgtacgcg 660

ggcccgggcc tggacagtgg gctgaggccg cgggggggctc cgctggggcc gccgccgccc 720

cgccaacccg gggccctggc gtacggggcc ttcgggccgc cgtcctcctt ccagcccttt 780

ccggccgtgc ctccgccggc cgcgggcatc gcgcacctgc agcctgtggc gacgccgtac 840

cccggccgcg ccgccgcgcc ccccaacgct ccgggaggcc ccccgggccc gcagccggcc 900

ccaagcgccg cagccccgcc gccgcccgcg cacgccctgg gcggcatgga cgccgaactc 960

atcgacgagg aggcgctgac gtcgctggag ctggagctgg ggctgcaccg cgtgcgcgag 1020

ctgcccgagc tgttcctggg ccagagcgag ttcgactgct tctcggactt ggggtccgcg 1080

ccgcccgccg gctccgtgag ctgccagtcc cagctcatca aacccagccg catgcgcaag 1140

taccccaacc ggcccagcaa gacgcccccc cacgaacgcc cttacgcttg cccagtggag 1200

tcctgtgatc gccgcttctc ccgcagcgac aacctggtga gacacatccg catccacaca 1260

ggccagaagc ccttccagtg ccgcatctgc atgagaaact tcagccgaga ggataacttg 1320

cacactcaca tccgcaccca cacaggcgaa aagcccttcg cctgcgacat ctgtggaaga 1380

aagtttgccc ggagcgatga acttgtccga cataccaaga tccacttgcg gcagaaggac 1440

cgcccttacg cttgcccagt ggagtcctgt gatcgccgct tctcccaatc agggaatctg 1500

actgagcaca tccgcatcca cacaggccag aagcccttcc agtgccgcat ctgcatgaga 1560

aacttcagca caagtggaca tctggtacgc cacatccgca cccacacagg cgaaaagccc 1620

ttcgcctgcg acatctgtgg aagaaagttt gcccagaata gtaccctgac cgaacatacc 1680

aagatccact tgcggcagaa ggacaag 1707

<210> 73
<211> 1755
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 73
atggccgcag atcacctgat gctggctgaa ggctacagac tggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgccctg agaacactgc ccccctacgc cggccctggt 120

cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180

ggagccctgg cctacggcgc cttcggccct ccttctagct tccagccctt ccccgccgtg 240

cctcctccag ctgctggcat cgcccacctg cagcctgtgg ccacccctta ccccggaaga 300

gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360

gctgcccctc cacctcctgc tcatgccctg ggcggcatgg acgccgagct gatcgacgag 420

gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480

ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540

ggctctgtgt cctgcggcgg cagcggcggc ggaagcggcg ccgaccacct gatgctcgcc 600

gagggctacc gcctggtgca gaggccgccg tccgccgccg ccgcccatgg ccctcatgcg 660

ctccggactc tgccgccgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg 720

gctccgctgg ggccgccgcc gccccgccaa cccggggccc tggcgtacgg ggccttcggg 780

ccgccgtcct ccttccagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac 840

ctgcagcctg tggcgacgcc gtaccccggc cgcgcggccg cgccccccaa cgctccggga 900

ggcccccccgg gcccgcagcc ggccccaagc gccgcagccc cgccgccgcc cgcgcacgcc 960

ctgggcggca tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag 1020

ctggggctgc accgcgtgcg cgagctgccc gagctgttcc tgggccagag cgagttcgac 1080

tgcttctcgg acttggggtc cgcgccgccc gccggctccg tgagctgcgg tggttctggt 1140

ggtggttctg gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg 1200

cccagcaaga cgccccccca cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc 1260

cgcttctccc gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc 1320

ttccagtgcc gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc 1380

cgcacccaca caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg 1440

agcgatgaac ttgtccgaca taccaagatc cacttgcggc agaaggaccg cccttacgct 1500

tgcccagtgg agtcctgtga tcgccgcttc tcccaatcag ggaatctgac tgagcacatc 1560

cgcatccaca caggccagaa gcccttccag tgccgcatct gcatgagaaa cttcagcaca 1620

agtggacatc tggtacgcca catccgcacc cacacaggcg aaaagcccct cgcctgcgac 1680

atctgtggaa gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg 1740

cggcagaagg acaag 1755

<210> 74
<211> 3438
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 74
ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttttt ttttttttttt ttttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttc ccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760

tgggttctga tgccctcgat gactttgacc tggatatgtt gggaagcgac gcattggatg 2820

actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880

aaactagtga aacccagcag acaatgtagc tagacccagt agccagatgt agctaaagag 2940

accggttcac tgtgaaagct tgggtggcat ccctgtgacc cctccccagt gcctctcctg 3000

gccctggaag ttgccactcc agtgcccacc agccttgtcc taataaaatt aagttgcatc 3060

attttgtctg actaggtgtc cttctataat attatggggt ggaggggggt ggtatggagc 3120

aaggggcaag ttgggaagac aacctgtagg gcctgcgggg tctattggga accaagctgg 3180

agtgcagtgg cacaatcttg gctcactgca atctccgcct cctgggttca agcgattctc 3240

ctgcctcagc ctcccgagtt gttgggattc caggcatgca tgaccaggct cagctaattt 3300

ttgtttttt ggtagagacg gggtttcacc atattggcca ggctggtctc caactcctaa 3360

tctcaggtga tctacccacc ttggcctccc aaattgctgg gattacaggc gtgaaccact 3420

gctcccttcc ctgtcctt 3438

<210> 75
<211> 3505
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 75
ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

aggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgcccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccc caaagaagaa gcggaaggtc ggtatccacg 2100

gagtcccagc agccctcgaa ccaggtgaaa aaccttacaa atgtcctgaa tgtgggaaat 2160

cattcagtcg cagcgacaac ctggtgagac atcaacgcac ccatacagga gaaaaacctt 2220

ataaatgtcc agaatgtgga aagtccttct cacgagagga taacttgcac actcatcaac 2280

gaacacatac tggtgaaaaa ccatacaagt gtcccgaatg tggtaaaagt tttagccgga 2340

gcgatgaact tgtccgacac caacgaaccc atacaggcga gaagccttac aaatgtcccg 2400

agtgtggcaa gagcttctca caatcaggga atctgactga gcatcaacga actcataccg 2460

gggaaaaacc ttacaagtgt ccagagtgtg ggaagagctt ttccacaagt ggacatctgg 2520

tacgccacca gaggacacat acaggggaga agccctacaa atgccccgaa tgcggtaaaa 2580

gtttctctca gaatagtacc ctgaccgaac accagcgaac acacactggg aaaaaaacga 2640

gtaaaaggcc ggcggccacg aaaaaggccg gccaggcaaa aaagaaaaag ggatcctacc 2700

catacgacgt accagattac gctctcgagg acgcgctgga cgatttcgat ctcgacatgc 2760

tgggttctga tgccctcgat gactttgacc tggatatgtt gggaagcgac gcattggatg 2820

actttgatct ggacatgctc ggctccgatg ctctggacga tttcgatctc gatatgttat 2880

aaactagtga aacccagcag acaatgtagc tagacccagt agccagatgt agctaaagag 2940

accggttcac tgtgagaaac ccagcagaca atgtagctag acccagtagc cagatgtagc 3000

taaagagacc ggttcactgt gaaagcttgg gtggcatccc tgtgaccccc ccccagtgcc 3060

tctcctggcc ctggaagttg ccactccagt gcccaccagc cttgtcctaa taaaattaag 3120

ttgcatcatt ttgtctgact aggtgtcctt ctataatatt atggggtgga gggggggtggt 3180

atggagcaag gggcaagttg ggaagacaac ctgtagggcc tgcggggtct attgggaacc 3240

aagctggagt gcagtggcac aatcttggct cactgcaatc tccgcctcct gggttcaagc 3300

gattctcctg cctcagcctc ccgagttgtt gggattccag gcatgcatga ccaggctcag 3360

ctaatttttg ttttttggt agagacgggg tttcaccata ttggccaggc tggtctccaa 3420

ctcctaatct caggtgatct acccaccttg gcctcccaaa ttgctgggat tacaggcgtg 3480

aaccactgct cccttccctg tcctt 3505

<210> 76
<211> 3804
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 76

ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttt ttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttte cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgccccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccg ccgaccacct gatgctcgcc gaggggctacc 2100

gcctggtgca gaggccgccg tccgccgccg ccgcccatgg ccctcatgcg ctccggactc 2160

tgccgccgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg gctccgctgg 2220

ggccgccgcc gccccgccaa cccggggccc tggcgtacgg ggccttcggg ccgccgtcct 2280

ccttccagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac ctgcagcctg 2340

tggcgacgcc gtaccccggc cgcgcggccg cgccccccaa cgctccggga ggcccccggg 2400

gcccgcagcc ggccccaagc gccgcagccc cgccgccgcc cgcgcacgcc ctgggcggca 2460

tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag ctggggctgc 2520

accgcgtgcg cgagctgccc gagctgttcc tgggccagag cgagttcgac tgcttctcgg 2580

acttggggtc cgcgccgccc gccggctccg tgagctgcgg tggttctggt ggtggttctg 2640

gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg cccagcaaga 2700

cgccccccca cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc cgcttctccc 2760

gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc ttccagtgcc 2820

gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc cgcacccaca 2880

caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg agcgatgaac 2940

ttgtccgaca taccaagatc cacttgcggc agaaggaccg cccttacgct tgcccagtgg 3000

agtcctgtga tcgccgcttc tcccaatcag ggaatctgac tgagcacatc cgcatccaca 3060

caggccagaa gcccttccag tgccgcatct gcatgagaaa cttcagcaca agtggacatc 3120

tggtacgcca catccgcacc cacacaggcg aaaagcccct cgcctgcgac atctgtggaa 3180

gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg cggcagaagg 3240

acaagtaact cgaggaaacc cagcagacaa tgtagctaga cccagtagcc agatgtagct 3300

aaagagaccg gttcactgtg aaagcttggg tggcatccct gtgacccctc cccagtgcct 3360

ctcctggccc tggaagttgc cactccagtg cccaccagcc ttgtcctaat aaaattaagt 3420

tgcatcattt tgtctgacta ggtgtccttc tataatatta tggggtggag gggggtggta 3480

tggagcaagg ggcaagttgg gaagacaacc tgtagggcct gcggggtcta ttgggaacca 3540

agctggagtg cagtggcaca atcttggctc actgcaatct ccgcctcctg ggttcaagcg 3600

attctcctgc ctcagcctcc cgagttgttg ggattccagg catgcatgac caggctcagc 3660

taatttttgt ttttttggta gagacggggt ttcaccatat tggccaggct ggtctccaac 3720

tcctaatctc aggtgatcta cccaccttgg cctcccaaat tgctgggatt acaggcgtga 3780

accactgctc ccttccctgt cctt 3804

<210> 77
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 77
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
35 40 45

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn
145 150 155 160

Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 78
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 78
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Thr Lys Asn Ser Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ala
35 40 45

Asp Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Leu Ala His Leu Arg
65 70 75 80

Ala His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr Lys Asn Ser Leu Thr Glu His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ala Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr His
145 150 155 160

Leu Asp Leu Ile Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 79
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 79
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Gln Ala Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
35 40 45

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly Asn Leu Thr
65 70 75 80

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Lys Ser Ser Leu Ile Ala His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ala
145 150 155 160

Gly His Leu Ala Ser His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 80
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 80
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser

1 5 10 15

Phe Ser Thr Thr Gly Asn Leu Thr Val His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser
35 40 45

Gly Glu Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr Ser Gly Asn Leu Thr Glu His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg
145 150 155 160

Ala Asn Leu Arg Ala His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 81
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 81

```
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1               5                  10                  15

Phe Ser Ser Arg Arg Thr Cys Arg Ala His Gln Arg Thr His Thr Gly
            20                  25                  30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Thr
        35                  40                  45

Gly Ala Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
    50                  55                  60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val
65                  70                  75                  80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
                85                  90                  95

Cys Gly Lys Ser Phe Ser Arg Asn Asp Ala Leu Thr Glu His Gln Arg
            100                 105                 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
        115                 120                 125

Phe Ser Gln Ser Gly Asp Leu Arg Arg His Gln Arg Thr His Thr Gly
    130                 135                 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser
145                 150                 155                 160

His Ser Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
                165                 170                 175
```

<210> 82
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 82

```
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1               5                  10                  15

Phe Ser Arg Lys Asp Asn Leu Lys Asn His Gln Arg Thr His Thr Gly
            20                  25                  30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro
        35                  40                  45

Gly Ala Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
    50                  55                  60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
65                  70                  75                  80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
                85                  90                  95

Cys Gly Lys Ser Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg
            100                 105                 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
        115                 120                 125

Phe Ser Thr Ser Gly Glu Leu Val Arg His Gln Arg Thr His Thr Gly
    130                 135                 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Lys
145                 150                 155                 160

Asp Asn Leu Lys Asn His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
                165                 170                 175
```

<210> 83

<211> 232
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 83

```
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1               5                  10                  15

Phe Ser Ser Lys Lys Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
            20                  25                  30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Pro
        35                  40                  45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
    50                  55                  60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Asn Leu Val
65                  70                  75                  80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
                85                  90                  95

Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg
            100                 105                 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
        115                 120                 125

Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr His Thr Gly
    130                 135                 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser
145                 150                 155                 160

Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
                165                 170                 175
```

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly His Leu Val
180 185 190

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
195 200 205

Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu His Gln Arg
210 215 220

Thr His Thr Gly Lys Lys Thr Ser
225 230

<210> 84
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 84
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
35 40 45

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser
145 150 155 160

Gly His Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 85
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 85
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Ser Lys Lys Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Pro
35 40 45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Asn Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser
145 150 155 160

Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 86
<211> 232
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 86
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Asp Cys Arg Asp Leu Ala Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Asn
35 40 45

Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Asn Asp Ala Leu Thr

65 70 75 80

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Asp Pro Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Arg
145 150 155 160

Ala His Leu Glu Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Ser Ser Leu Val
180 185 190

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
195 200 205

Cys Gly Lys Ser Phe Ser His Arg Thr Thr Leu Thr Asn His Gln Arg
210 215 220

Thr His Thr Gly Lys Lys Thr Ser
225 230

<210> 87
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 87

Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Asn Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Pro
35 40 45

Ala Asp Leu Thr Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro Gly Asn Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Gln Arg Ala His Leu Glu Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg
145 150 155 160

Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 88
<211> 204
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 88
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Asn Asp Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Asp Pro
35 40 45

Gly His Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly Glu Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Ser Lys Lys Ala Leu Thr Glu His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Leu
145 150 155 160

Ala His Leu Arg Ala His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp His Leu Thr

180 185 190

Asn His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
195 200

<210> 89
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 89
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg
35 40 45

Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His
65 70 75 80

Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
145 150 155 160

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 90
<211> 176
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 90
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
35 40 45

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asp Leu Arg
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
145 150 155 160

Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr Ser
165 170 175

<210> 91
<211> 260
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 91
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser Arg Arg Asp Glu Leu Asn Val His Gln Arg Thr His Thr Gly
20 25 30

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser
35 40 45

Asp His Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
50 55 60

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Asp Leu Val
65 70 75 80

Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg
100 105 110

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
115 120 125

Phe Ser His Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly
130 135 140

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu
145 150 155 160

Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr
180 185 190

Glu His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
195 200 205

Cys Gly Lys Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg
210 215 220

Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
225 230 235 240

Phe Ser Arg Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly
245 250 255

Lys Lys Thr Ser
260

<210> 92
<211> 177
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 92
Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg

1 5 10 15

Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro
20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu
35 40 45

His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
50 55 60

Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr
65 70 75 80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu
85 90 95

Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile
100 105 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
115 120 125

Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr
130 135 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln
145 150 155 160

Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys Asp
165 170 175

Lys

<210> 93
<211> 177
<212> PRT

<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 93

```
Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg
1               5                  10                  15

Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro
            20                  25                  30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu
        35                  40                  45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
    50                  55                  60

Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr
65                  70                  75                  80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu
                85                  90                  95

Ser Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Ile
            100                 105                 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
        115                 120                 125

Asn Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr
    130                 135                 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
145                 150                 155                 160

Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp
                165                 170                 175
```

Lys

<210> 94
<211> 264
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 94

```
Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg
1               5                  10                  15

Arg Asp Glu Leu Asn Val His Ile Arg Ile His Thr Gly Gln Lys Pro
            20                  25                  30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu
        35                  40                  45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
    50                  55                  60

Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Thr
65                  70                  75                  80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu
                85                  90                  95

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
            100                 105                 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
        115                 120                 125

Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr
    130                 135                 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
```

145 150 155 160

Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp
165 170 175

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Thr
180 185 190

Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
195 200 205

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser Leu
210 215 220

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
225 230 235 240

Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr
245 250 255

Lys Ile His Leu Arg Gln Lys Asp
260

<210> 95
<211> 264
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 95
Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Asp
1 5 10 15

Pro Gly Ala Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro
20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn Leu
35 40 45

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
50 55 60

Ile Cys Gly Arg Lys Phe Ala Gln Ser Gly Asp Leu Arg Arg His Thr
65 70 75 80

Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu
85 90 95

Ser Cys Asp Arg Arg Phe Ser Thr His Leu Asp Leu Ile Arg His Ile
100 105 110

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
115 120 125

Asn Phe Ser Thr Ser Gly Asn Leu Val Arg His Ile Arg Thr His Thr
130 135 140

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
145 150 155 160

Ser Asp Asn Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
165 170 175

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
180 185 190

Ser Gly His Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
195 200 205

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Glu Arg Ser His Leu
210 215 220

Arg Glu His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
225 230 235 240

Ile Cys Gly Arg Lys Phe Ala Gln Ala Gly His Leu Ala Ser His Thr
245 250 255

Lys Ile His Leu Arg Gln Lys Asp
260

<210> 96
<211> 175
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 96
Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
1 5 10 15

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro
20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn Leu
35 40 45

His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
50 55 60

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Ala
65 70 75 80

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
85 90 95

Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg Ile
100 105 110

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
115 120 125

Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu

130 135 140

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn Ser
145 150 155 160

Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
165 170 175

<210> 97
<211> 175
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 97
Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
1 5 10 15

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro
20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser His Arg Thr Thr Leu
35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
50 55 60

Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ala
65 70 75 80

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
85 90 95

Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Leu Arg Ile
100 105 110

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
115 120 125

Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu
130 135 140

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
145 150 155 160

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
165 170 175

<210> 98
<211> 261
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 98
Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
1 5 10 15

Arg Asp Glu Leu Asn Val His Leu Arg Ile His Thr Gly His Lys Pro
20 25 30

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Ser Asp His Leu
35 40 45

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
50 55 60

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Ala
65 70 75 80

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
85 90 95

Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile
100 105 110

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
115 120 125

Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly Glu
130 135 140

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
145 150 155 160

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu His Ala
165 170 175

Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu
180 185 190

Thr Glu His Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg
195 200 205

Ile Cys Met Arg Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile
210 215 220

Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg
225 230 235 240

Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ala Lys Ile His Leu
245 250 255

Lys Gln Lys Glu Lys
260

<210> 99
<211> 753
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 99

```
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1               5                  10                  15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
            20                  25                  30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
        35                  40                  45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
    50                  55                  60

Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro
65                  70                  75                  80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu
                85                  90                  95

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
            100                 105                 110

Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln
        115                 120                 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
    130                 135                 140

Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr
145                 150                 155                 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
                165                 170                 175

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Lys Lys Thr
            180                 185                 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
        195                 200                 205
```

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Glu Ala
210 215 220

Ser Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met
225 230 235 240

Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
245 250 255

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
260 265 270

Asp Asp Phe Asp Leu Asp Met Leu Ile Asn Ser Arg Ser Ser Gly Ser
275 280 285

Pro Lys Lys Lys Arg Lys Val Gly Ser Gln Tyr Leu Pro Asp Thr Asp
290 295 300

Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe
305 310 315 320

Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg
325 330 335

Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg Ser Ser Ala Ser Val
340 345 350

Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr
355 360 365

Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile
370 375 380

Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln
385 390 395 400

Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln
405 410 415

Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val
420 425 430

Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser
435 440 445

Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu
450 455 460

Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr Asp Leu Ala Ser Val
465 470 475 480

Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala
485 490 495

Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr
500 505 510

Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro
515 520 525

Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp
530 535 540

Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Gly Ser Gly
545 550 555 560

Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro Glu
565 570 575

Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys Gln
580 585 590

Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala Asn
595 600 605

Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val His
610 615 620

Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro Leu Asp
625 630 635 640

Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro
645 650 655

Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp
660 665 670

Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met Asp
675 680 685

Leu Ser His Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr Thr Thr
690 695 700

Leu Glu Ser Met Thr Glu Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro
705 710 715 720

Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu
725 730 735

His Ala Met His Ile Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu
740 745 750

Phe

<210> 100
<211> 753
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 100

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu
85 90 95

Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110

Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln
115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140

Ser Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln
165 170 175

Asn Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys

195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Glu Ala
210 215 220

Ser Gly Ser Gly Arg Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met
225 230 235 240

Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
245 250 255

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
260 265 270

Asp Asp Phe Asp Leu Asp Met Leu Ile Asn Ser Arg Ser Ser Gly Ser
275 280 285

Pro Lys Lys Lys Arg Lys Val Gly Ser Gln Tyr Leu Pro Asp Thr Asp
290 295 300

Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe
305 310 315 320

Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg
325 330 335

Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg Ser Ser Ala Ser Val
340 345 350

Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr
355 360 365

Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile
370 375 380

Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln
385 390 395 400

Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln
405 410 415

Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val
420 425 430

Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser
435 440 445

Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu
450 455 460

Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr Asp Leu Ala Ser Val
465 470 475 480

Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala
485 490 495

Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr
500 505 510

Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro
515 520 525

Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp
530 535 540

Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu Leu Gly Ser Gly
545 550 555 560

Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro Glu
565 570 575

Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys Gln
580 585 590

Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala Asn

595 600 605

Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val His
610 615 620

Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro Leu Asp
625 630 635 640

Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp Pro
645 650 655

Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala Asp
660 665 670

Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met Asp
675 680 685

Leu Ser His Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr Thr Thr
690 695 700

Leu Glu Ser Met Thr Glu Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro
705 710 715 720

Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu
725 730 735

His Ala Met His Ile Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu
740 745 750

Phe

<210> 101
<211> 272
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 101

Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
50 55 60

Arg Thr Thr Leu Thr Asn His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu
85 90 95

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110

Glu Cys Gly Lys Ser Phe Ser Thr Ser His Ser Leu Thr Glu His Gln
115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140

Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
165 170 175

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
260 265 270

<210> 102
<211> 272
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 102
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu

85 90 95

Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110

Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln
115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140

Ser Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln
165 170 175

Asn Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
260 265 270

<210> 103
<211> 580
<212> PRT

<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 103

```
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1               5                  10                  15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
            20                  25                  30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
        35                  40                  45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
    50                  55                  60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65                  70                  75                  80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
                85                  90                  95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
            100                 105                 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
        115                 120                 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
    130                 135                 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145                 150                 155                 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
                165                 170                 175
```

```
Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
            180                 185                 190

Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
        195                 200                 205

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
    210                 215                 220

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
225                 230                 235                 240

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
                245                 250                 255

Arg Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His
            260                 265                 270

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
        275                 280                 285

Arg Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys
    290                 295                 300

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
305                 310                 315                 320

Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
                325                 330                 335

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser
            340                 345                 350

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
        355                 360                 365

Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His
    370                 375                 380
```

Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ala Asp His
385 390 395 400

Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala
405 410 415

Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala
420 425 430

Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly
435 440 445

Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly
450 455 460

Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala
465 470 475 480

Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala
485 490 495

Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala
500 505 510

Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met
515 520 525

Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu
530 535 540

Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln
545 550 555 560

Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly
565 570 575

Ser Val Ser Cys
580

<210> 104
<211> 394
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 104
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu

145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
195 200 205

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
210 215 220

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
225 230 235 240

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
245 250 255

Arg Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His
260 265 270

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
275 280 285

Arg Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys
290 295 300

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
305 310 315 320

Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
325 330 335

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser
340 345 350

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
355 360 365

Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His
370 375 380

Thr Lys Ile His Leu Arg Gln Lys Asp Lys
385 390

<210> 105
<211> 580
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 105
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
195 200 205

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
210 215 220

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
225 230 235 240

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
245 250 255

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His
260 265 270

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
275 280 285

Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys
290 295 300

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
305 310 315 320

Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
325 330 335

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
340 345 350

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
355 360 365

Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
370 375 380

Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ala Asp His
385 390 395 400

Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala
405 410 415

Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala
420 425 430

Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly
435 440 445

Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly
450 455 460

Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala
465 470 475 480

Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala
485 490 495

Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala
500 505 510

Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met
515 520 525

Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu
530 535 540

Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln
545 550 555 560

Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly
565 570 575

Ser Val Ser Cys
580

<210> 106
<211> 394
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 106
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro

85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
195 200 205

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
210 215 220

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
225 230 235 240

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
245 250 255

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His
260 265 270

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
275 280 285

Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys
290 295 300

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
305 310 315 320

Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
325 330 335

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
340 345 350

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
355 360 365

Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
370 375 380

Thr Lys Ile His Leu Arg Gln Lys Asp Lys
385 390

<210> 107
<211> 388
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 107
Met Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
1 5 10 15

Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
20 25 30

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
35 40 45

```
Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
    50                  55                  60

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His
65                  70                  75                  80

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
                85                  90                  95

Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys
            100                 105                 110

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
        115                 120                 125

Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
    130                 135                 140

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
145                 150                 155                 160

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
                165                 170                 175

Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
            180                 185                 190

Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ala Asp His
        195                 200                 205

Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala
    210                 215                 220

Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala
225                 230                 235                 240

Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly
                245                 250                 255
```

Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly
260 265 270

Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala
275 280 285

Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala
290 295 300

Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala
305 310 315 320

Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met
325 330 335

Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu
340 345 350

Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln
355 360 365

Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly
370 375 380

Ser Val Ser Cys
385

<210> 108
<211> 479
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 108
Met Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn
1 5 10 15

```
Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val
            20                  25                  30

Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His
        35                  40                  45

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
    50                  55                  60

Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His
65                  70                  75                  80

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
                85                  90                  95

Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys
            100                 105                 110

Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
        115                 120                 125

Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys
    130                 135                 140

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His
145                 150                 155                 160

Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
                165                 170                 175

Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His
            180                 185                 190

Thr Lys Ile His Leu Arg Gln Lys Asp Lys Leu Glu Met Ser Gly Leu
        195                 200                 205

Glu Met Ala Asp His Met Met Ala Met Asn His Gly Arg Phe Pro Asp
```

210 215 220

Gly Thr Asn Gly Leu His His His Pro Ala His Arg Met Gly Met Gly
225 230 235 240

Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln Pro Gln His Ala
245 250 255

Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly Ala Gly Asn Met
260 265 270

Asn Ala Thr Ser Gly Ile Arg His Ala Met Gly Pro Gly Thr Val Asn
275 280 285

Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala Arg Phe Asn Asn
290 295 300

Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly Gly Ser Leu Pro
305 310 315 320

Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr Phe Asn His His
325 330 335

Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His Pro Ala Ala Gly
340 345 350

His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp Cys Asn Pro Lys
355 360 365

His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly Gly Ser Ser Thr
370 375 380

Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala Gly Ser Ser Asn
385 390 395 400

Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala Ser Val Ala His
405 410 415

Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp Thr Asp Phe Ile
420 425 430

Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met Gly Leu Asp Arg
435 440 445

Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn Glu Phe Asp Phe
450 455 460

Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg Val Ser Cys
465 470 475

<210> 109
<211> 476
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 109
Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly
1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg
20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln
35 40 45

Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly
50 55 60

Ala Gly Asn Met Asn Ala Thr Ser Gly Val Arg His Ala Met Gly Pro
65 70 75 80

Gly Thr Val Asn Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala
85 90 95

```
Arg Phe Asn Asn Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly
            100                 105                 110

Gly Ser Leu Pro Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr
        115                 120                 125

Phe Asn His His Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His
    130                 135                 140

Pro Ala Ala Gly His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp
145                 150                 155                 160

Cys Asn Pro Lys His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly
                165                 170                 175

Gly Ser Ser Thr Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala
            180                 185                 190

Gly Ser Ser Asn Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala
        195                 200                 205

Ser Val Ala His Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp
    210                 215                 220

Thr Asp Phe Ile Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met
225                 230                 235                 240

Gly Leu Asp Arg Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn
                245                 250                 255

Glu Phe Asp Phe Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg
            260                 265                 270

Val Ser Cys Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr
        275                 280                 285

Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys
    290                 295                 300
```

Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val
305 310 315 320

Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile
325 330 335

Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg
340 345 350

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
355 360 365

Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg
370 375 380

Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg
385 390 395 400

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
405 410 415

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
420 425 430

Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
435 440 445

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr
450 455 460

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys
465 470 475

<210> 110
<211> 554
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 110

```
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1               5                  10                  15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
            20                  25                  30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
        35                  40                  45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
    50                  55                  60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65                  70                  75                  80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
                85                  90                  95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
            100                 105                 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
        115                 120                 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
    130                 135                 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145                 150                 155                 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
                165                 170                 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
```

180 185 190

Gly Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser
195 200 205

Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys
210 215 220

Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn
225 230 235 240

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
245 250 255

Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
260 265 270

Ala Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly
275 280 285

Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg
290 295 300

Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser
305 310 315 320

Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly
325 330 335

Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn
340 345 350

Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
355 360 365

Leu Glu Met Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val
370 375 380

Gln Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg
385 390 395 400

Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro
405 410 415

Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu
420 425 430

Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala
435 440 445

Val Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr
450 455 460

Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro
465 470 475 480

Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala
485 490 495

His Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu
500 505 510

Thr Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro
515 520 525

Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly
530 535 540

Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
545 550

<210> 111
<211> 362
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 111

```
Met Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser
1               5                   10                  15

Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys
            20                  25                  30

Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn
        35                  40                  45

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
    50                  55                  60

Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
65                  70                  75                  80

Ala Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly
                85                  90                  95

Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg
            100                 105                 110

Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser
        115                 120                 125

Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly
    130                 135                 140

Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn
145                 150                 155                 160

Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
                165                 170                 175

Leu Glu Met Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val
            180                 185                 190
```

Gln Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg
195 200 205

Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro
210 215 220

Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu
225 230 235 240

Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala
245 250 255

Val Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr
260 265 270

Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro
275 280 285

Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala
290 295 300

His Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu
305 310 315 320

Thr Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro
325 330 335

Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly
340 345 350

Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
355 360

<210> 112
<211> 479
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 112

```
Met Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser
1               5                  10                  15

Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys
            20                  25                  30

Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn
        35                  40                  45

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
    50                  55                  60

Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
65                  70                  75                  80

Ala Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly
                85                  90                  95

Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg
            100                 105                 110

Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser
        115                 120                 125

Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly
    130                 135                 140

Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn
145                 150                 155                 160

Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys
                165                 170                 175

Lys Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro Pro Val
```

180 185 190

Ser Leu Ala Pro Val Val Thr Thr Cys Ala Leu Glu Met Ser Gly Leu
195 200 205

Glu Met Ala Asp His Met Met Ala Met Asn His Gly Arg Phe Pro Asp
210 215 220

Gly Thr Asn Gly Leu His His His Pro Ala His Arg Met Gly Met Gly
225 230 235 240

Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln Pro Gln His Ala
245 250 255

Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly Ala Gly Asn Met
260 265 270

Asn Ala Thr Ser Gly Ile Arg His Ala Met Gly Pro Gly Thr Val Asn
275 280 285

Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala Arg Phe Asn Asn
290 295 300

Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly Gly Ser Leu Pro
305 310 315 320

Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr Phe Asn His His
325 330 335

Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His Pro Ala Ala Gly
340 345 350

His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp Cys Asn Pro Lys
355 360 365

His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly Gly Ser Ser Thr
370 375 380

Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala Gly Ser Ser Asn
385 390 395 400

Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala Ser Val Ala His
405 410 415

Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp Thr Asp Phe Ile
420 425 430

Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met Gly Leu Asp Arg
435 440 445

Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn Glu Phe Asp Phe
450 455 460

Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg Val Ser Cys
465 470 475

<210> 113
<211> 476
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 113
Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly
1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg
20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln
35 40 45

Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly
50 55 60

Ala Gly Asn Met Asn Ala Thr Ser Gly Val Arg His Ala Met Gly Pro
65 70 75 80

Gly Thr Val Asn Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala
85 90 95

Arg Phe Asn Asn Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly
100 105 110

Gly Ser Leu Pro Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr
115 120 125

Phe Asn His His Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His
130 135 140

Pro Ala Ala Gly His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp
145 150 155 160

Cys Asn Pro Lys His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly
165 170 175

Gly Ser Ser Thr Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala
180 185 190

Gly Ser Ser Asn Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala
195 200 205

Ser Val Ala His Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp
210 215 220

Thr Asp Phe Ile Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met
225 230 235 240

Gly Leu Asp Arg Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn
245 250 255

Glu Phe Asp Phe Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg
260 265 270

Val Ser Cys Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg
275 280 285

Phe Ser Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly
290 295 300

His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu
305 310 315 320

Asp Asn Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
325 330 335

Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val
340 345 350

Arg His Ala Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala
355 360 365

Glu Gly Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
370 375 380

Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met
385 390 395 400

Arg Ser Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
405 410 415

Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala
420 425 430

Gln Asn Ser Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys
435 440 445

Glu Lys Lys Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro
450 455 460

Pro Val Ser Leu Ala Pro Val Val Thr Thr Cys Ala
465 470 475

<210> 114
<211> 559
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 114
Met Thr Gly Lys Leu Ala Glu Lys Leu Pro Val Thr Met Ser Ser Leu
1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala
20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln
35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro
50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn
65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
85 90 95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
100 105 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
115 120 125

Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu
130 135 140

Pro Pro Tyr Ser Asn Cys Gly Asp Leu Tyr Ser Glu Pro Val Ser Phe

145 150 155 160

His Asp Pro Gln Gly Asn Pro Gly Leu Ala Tyr Ser Pro Gln Asp Tyr
165 170 175

Gln Ser Ala Lys Pro Ala Leu Asp Ser Asn Leu Phe Pro Met Ile Pro
180 185 190

Asp Tyr Asn Leu Tyr His His Pro Asn Asp Met Gly Ser Ile Pro Glu
195 200 205

His Lys Pro Phe Gln Gly Met Asp Pro Ile Arg Val Asn Pro Pro Pro
210 215 220

Ile Thr Pro Leu Glu Thr Ile Lys Ala Phe Lys Asp Lys Gln Ile His
225 230 235 240

Pro Gly Phe Gly Ser Leu Pro Gln Pro Pro Leu Thr Leu Lys Pro Ile
245 250 255

Arg Pro Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Leu His Glu
260 265 270

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
275 280 285

Arg Asp Glu Leu Asn Val His Leu Arg Ile His Thr Gly His Lys Pro
290 295 300

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Ser Asp His Leu
305 310 315 320

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
325 330 335

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Ala
340 345 350

```
Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
        355                 360                 365

Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Leu Arg Ile
    370                 375                 380

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
385                 390                 395                 400

Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly Glu
                405                 410                 415

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
            420                 425                 430

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu His Ala
        435                 440                 445

Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu
    450                 455                 460

Thr Glu His Leu Arg Ile His Thr Gly His Lys Pro Phe Gln Cys Arg
465                 470                 475                 480

Ile Cys Met Arg Ser Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile
                485                 490                 495

Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg
            500                 505                 510

Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ala Lys Ile His Leu
        515                 520                 525

Lys Gln Lys Glu Lys Lys Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser
    530                 535                 540

Ser Ala Pro Pro Val Ser Leu Ala Pro Val Val Thr Thr Cys Ala
```

545 550 555

<210> 115
<211> 631
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 115
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
340 345 350

```
Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
        355                 360                 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu Thr
    370                 375                 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385                 390                 395                 400

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr Lys
                405                 410                 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser
            420                 425                 430

Cys Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg
        435                 440                 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn
    450                 455                 460

Phe Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly
465                 470                 475                 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu
                485                 490                 495

Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys
            500                 505                 510

Lys Ala Asp Lys Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser
        515                 520                 525

Ser Tyr Pro Ser Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr
    530                 535                 540

Ser Tyr Pro Ser Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr
545                 550                 555                 560
```

Ser Phe Ser Ser Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser
565 570 575

Gly Phe Pro Ser Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro
580 585 590

Ala Phe Pro Ala Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn
595 600 605

Ser Phe Ser Ala Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser
610 615 620

Pro Arg Thr Ile Glu Ile Cys
625 630

<210> 116
<211> 719
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 116
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro

275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Arg
340 345 350

Asp Glu Leu Asn Val His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu Thr
370 375 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Asp Leu Val Arg His Thr Lys
405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser
420 425 430

Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile Arg
435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn
450 455 460

Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg Thr His Thr Gly
465 470 475 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu
485 490 495

Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Arg
500 505 510

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Thr Ser
515 520 525

His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
530 535 540

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser Ser Ser Leu Val
545 550 555 560

Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
565 570 575

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Thr Lys
580 585 590

Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala
595 600 605

Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr
610 615 620

Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr
625 630 635 640

Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser
645 650 655

Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala
660 665 670

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser

675 680 685

Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu
690 695 700

Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys
705 710 715

<210> 117
<211> 627
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 117
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

```
Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
                325                 330                 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
            340                 345                 350

Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
        355                 360                 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu Thr
    370                 375                 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385                 390                 395                 400

Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ile Arg
                405                 410                 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
            420                 425                 430

Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly
        435                 440                 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser
    450                 455                 460

Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
465                 470                 475                 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His
                485                 490                 495

Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
            500                 505                 510

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
        515                 520                 525
```

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
530 535 540

Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
595 600 605

Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620

Glu Ile Cys
625

<210> 118
<211> 631
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 118
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu

245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
340 345 350

Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His
370 375 380

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys
405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser
420 425 430

Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg
435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn
450 455 460

Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly
465 470 475 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn
485 490 495

Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys
500 505 510

Lys Ala Asp Lys Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser
515 520 525

Ser Tyr Pro Ser Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr
530 535 540

Ser Tyr Pro Ser Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr
545 550 555 560

Ser Phe Ser Ser Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser
565 570 575

Gly Phe Pro Ser Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro
580 585 590

Ala Phe Pro Ala Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn
595 600 605

Ser Phe Ser Ala Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser
610 615 620

Pro Arg Thr Ile Glu Ile Cys
625 630

<210> 119

<211> 473
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 119

```
Met Thr Gly Lys Leu Ala Glu Lys Leu Pro Val Thr Met Ser Ser Leu
1               5                  10                  15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala
            20                  25                  30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln
        35                  40                  45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro
    50                  55                  60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn
65                  70                  75                  80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
                85                  90                  95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
            100                 105                 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
        115                 120                 125

Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu
    130                 135                 140

Pro Pro Tyr Ser Asn Cys Gly Asp Leu Tyr Ser Glu Pro Val Ser Phe
145                 150                 155                 160

His Asp Pro Gln Gly Asn Pro Gly Leu Ala Tyr Ser Pro Gln Asp Tyr
                165                 170                 175
```

```
Gln Ser Ala Lys Pro Ala Leu Asp Ser Asn Leu Phe Pro Met Ile Pro
            180                 185                 190

Asp Tyr Asn Leu Tyr His His Pro Asn Asp Met Gly Ser Ile Pro Glu
        195                 200                 205

His Lys Pro Phe Gln Gly Met Asp Pro Ile Arg Val Asn Pro Pro Pro
    210                 215                 220

Ile Thr Pro Leu Glu Thr Ile Lys Ala Phe Lys Asp Lys Gln Ile His
225                 230                 235                 240

Pro Gly Phe Gly Ser Leu Pro Gln Pro Pro Leu Thr Leu Lys Pro Ile
                245                 250                 255

Arg Pro Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Leu His Glu
            260                 265                 270

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
        275                 280                 285

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro
    290                 295                 300

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser His Arg Thr Thr Leu
305                 310                 315                 320

Thr Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
                325                 330                 335

Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His Thr His Ala
            340                 345                 350

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
        355                 360                 365
```

Asp Arg Arg Phe Ser Thr Ser His Ser Leu Thr Glu His Leu Arg Ile
370 375 380

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
385 390 395 400

Ser Gln Ser Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu
405 410 415

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Arg Glu Asp
420 425 430

Asn Leu His Thr His Ala Lys Ile His Leu Lys Gln Lys Glu Lys Lys
435 440 445

Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro Pro Val Ser
450 455 460

Leu Ala Pro Val Val Thr Thr Cys Ala
465 470

<210> 120
<211> 719
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 120
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser

50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Asp Pro
340 345 350

Gly Ala Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn Leu Val
370 375 380

Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Gln Ser Gly Asp Leu Arg Arg His Thr Lys
405 410 415

Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser
420 425 430

Cys Asp Arg Arg Phe Ser Thr His Leu Asp Leu Ile Arg His Ile Arg
435 440 445

Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn

450 455 460

Phe Ser Thr Ser Gly Asn Leu Val Arg His Ile Arg Thr His Thr Gly
465 470 475 480

Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser
485 490 495

Asp Asn Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp Arg
500 505 510

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser
515 520 525

Gly His Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
530 535 540

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Glu Arg Ser His Leu Arg
545 550 555 560

Glu His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
565 570 575

Cys Gly Arg Lys Phe Ala Gln Ala Gly His Leu Ala Ser His Thr Lys
580 585 590

Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala
595 600 605

Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr
610 615 620

Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr
625 630 635 640

Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser
645 650 655

Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala
660 665 670

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser
675 680 685

Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu
690 695 700

Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys
705 710 715

<210> 121
<211> 627
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 121
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
340 345 350

Asp Asn Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser His Ser Thr Thr Leu Thr
370 375 380

Asn His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Asn Arg Lys Thr His Ile Arg
405 410 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
420 425 430

Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly
435 440 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser
450 455 460

Ser Ser Leu Thr Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
465 470 475 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Asn Arg Lys
485 490 495

Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
500 505 510

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
515 520 525

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
530 535 540

Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
595 600 605

Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620

Glu Ile Cys
625

<210> 122
<211> 473
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 122
Met Thr Gly Lys Leu Ala Glu Lys Leu Pro Val Thr Met Ser Ser Leu
1 5 10 15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala

20 25 30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln
35 40 45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro
50 55 60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn
65 70 75 80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
85 90 95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
100 105 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
115 120 125

Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu
130 135 140

Pro Pro Tyr Ser Asn Cys Gly Asp Leu Tyr Ser Glu Pro Val Ser Phe
145 150 155 160

His Asp Pro Gln Gly Asn Pro Gly Leu Ala Tyr Ser Pro Gln Asp Tyr
165 170 175

Gln Ser Ala Lys Pro Ala Leu Asp Ser Asn Leu Phe Pro Met Ile Pro
180 185 190

Asp Tyr Asn Leu Tyr His His Pro Asn Asp Met Gly Ser Ile Pro Glu
195 200 205

His Lys Pro Phe Gln Gly Met Asp Pro Ile Arg Val Asn Pro Pro Pro
210 215 220

Ile Thr Pro Leu Glu Thr Ile Lys Ala Phe Lys Asp Lys Gln Ile His
225 230 235 240

Pro Gly Phe Gly Ser Leu Pro Gln Pro Pro Leu Thr Leu Lys Pro Ile
245 250 255

Arg Pro Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Leu His Glu
260 265 270

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
275 280 285

Ser Asp Asn Leu Val Arg His Leu Arg Ile His Thr Gly His Lys Pro
290 295 300

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Glu Asp Asn Leu
305 310 315 320

His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
325 330 335

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Ala
340 345 350

Lys Ile His Leu Lys Gln Lys Glu His Ala Cys Pro Ala Glu Gly Cys
355 360 365

Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His Leu Arg Ile
370 375 380

His Thr Gly His Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Ser Phe
385 390 395 400

Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu
405 410 415

Lys Pro Phe Ala Cys Glu Phe Cys Gly Arg Lys Phe Ala Gln Asn Ser

420 425 430

Thr Leu Thr Glu His Ala Lys Ile His Leu Lys Gln Lys Glu Lys Lys
435 440 445

Ala Glu Lys Gly Gly Ala Pro Ser Ala Ser Ser Ala Pro Pro Val Ser
450 455 460

Leu Ala Pro Val Val Thr Thr Cys Ala
465 470

<210> 123
<211> 627
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 123
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

```
Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305                 310                 315                 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
                325                 330                 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
            340                 345                 350

Asp Asn Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
        355                 360                 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp Asn Leu Thr
    370                 375                 380

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385                 390                 395                 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Ile Arg
                405                 410                 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
            420                 425                 430

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
        435                 440                 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
    450                 455                 460

Gly His Leu Thr Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
465                 470                 475                 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Ser Ser Thr Arg Lys
                485                 490                 495

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
            500                 505                 510
```

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
515 520 525

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
530 535 540

Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
595 600 605

Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620

Glu Ile Cys
625

<210> 124
<211> 627
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 124
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

```
Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
        35                  40                  45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
    50                  55                  60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65                  70                  75                  80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
                85                  90                  95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
            100                 105                 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
        115                 120                 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
    130                 135                 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145                 150                 155                 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
                165                 170                 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
            180                 185                 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
        195                 200                 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
    210                 215                 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
```

225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
340 345 350

Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His
370 375 380

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His Ile Arg
405 410 415

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
420 425 430

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
435 440 445

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
450 455 460

Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
465 470 475 480

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr
485 490 495

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys
500 505 510

Ser Val Val Ala Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser
515 520 525

Pro Val Ala Thr Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser
530 535 540

Pro Ala Thr Thr Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser
545 550 555 560

Pro Gly Ser Ser Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser
565 570 575

Pro Ser Val Ala Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala
580 585 590

Gln Val Ser Ser Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala
595 600 605

Ser Thr Gly Leu Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile
610 615 620

Glu Ile Cys

625

<210> 125
<211> 272
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 125

```
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1               5                  10                  15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
            20                  25                  30

Ser Phe Ser Ser Pro Ala Asp Leu Thr Arg His Gln Arg Thr His Thr
        35                  40                  45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
    50                  55                  60

Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro
65                  70                  75                  80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu
                85                  90                  95

His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
            100                 105                 110

Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln
        115                 120                 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
    130                 135                 140

Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr
145                 150                 155                 160
```

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr
165 170 175

Ser Gly His Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
260 265 270

<210> 126
<211> 272
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 126
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Asp Pro Gly Ala Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60

Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asp Leu
85 90 95

Arg Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110

Glu Cys Gly Lys Ser Phe Ser Thr His Leu Asp Leu Ile Arg His Gln
115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140

Ser Phe Ser Thr Ser Gly Asn Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
165 170 175

Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205

Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu Asp Ala
210 215 220

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
225 230 235 240

Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu

245 250 255

Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu
260 265 270

<210> 127
<211> 261
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 127
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1 5 10 15

Ala Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
20 25 30

Ser Phe Ser Arg Ser Asp Asn Leu Val Arg His Gln Arg Thr His Thr
35 40 45

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg
50 55 60

Glu Asp Asn Leu His Thr His Gln Arg Thr His Thr Gly Glu Lys Pro
65 70 75 80

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu
85 90 95

Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
100 105 110

Glu Cys Gly Lys Ser Phe Ser Gln Ser Gly Asn Leu Thr Glu His Gln
115 120 125

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
130 135 140

Ser Phe Ser Thr Ser Gly His Leu Val Arg His Gln Arg Thr His Thr
145 150 155 160

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln
165 170 175

Asn Ser Thr Leu Thr Glu His Gln Arg Thr His Thr Gly Lys Lys Thr
180 185 190

Ser Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys
195 200 205

Lys Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
210 215 220

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
225 230 235 240

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
245 250 255

Asp Leu Asp Met Leu
260

<210> 128
<211> 386
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 128
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gln Ser Gln Leu Ile Lys Pro
180 185 190

Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His
195 200 205

Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
210 215 220

Arg Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys

225 230 235 240

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn
245 250 255

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
260 265 270

Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
275 280 285

Thr Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val
290 295 300

Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
305 310 315 320

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
325 330 335

Arg Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
340 345 350

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
355 360 365

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys
370 375 380

Asp Lys
385

<210> 129
<211> 402
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 129

```
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1               5                  10                  15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
            20                  25                  30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
        35                  40                  45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
    50                  55                  60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65                  70                  75                  80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
                85                  90                  95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
            100                 105                 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
        115                 120                 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
    130                 135                 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145                 150                 155                 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
                165                 170                 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
            180                 185                 190
```

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gln Ser Gln Leu Ile Lys Pro
195 200 205

Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His
210 215 220

Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
225 230 235 240

Arg Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys
245 250 255

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn
260 265 270

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
275 280 285

Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
290 295 300

Thr Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val
305 310 315 320

Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
325 330 335

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
340 345 350

Arg Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
355 360 365

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
370 375 380

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys
385 390 395 400

Asp Lys

<210> 130
<211> 569
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 130
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Ala Asp His Leu Met Leu Ala
180 185 190

Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala Ala Ala Ala His
195 200 205

Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu
210 215 220

Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro
225 230 235 240

Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser
245 250 255

Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala Gly Ile Ala His
260 265 270

Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro
275 280 285

Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala
290 295 300

Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met Asp Ala Glu Leu
305 310 315 320

Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu Leu Gly Leu His
325 330 335

Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp

340 345 350

Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
355 360 365

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg
370 375 380

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
385 390 395 400

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
405 410 415

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
420 425 430

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr
435 440 445

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
450 455 460

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
465 470 475 480

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
485 490 495

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
500 505 510

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu
515 520 525

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
530 535 540

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
545 550 555 560

Lys Ile His Leu Arg Gln Lys Asp Lys
565

<210> 131
<211> 585
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 131
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1 5 10 15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
20 25 30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg
195 200 205

Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu
210 215 220

Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly
225 230 235 240

Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr
245 250 255

Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro
260 265 270

Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr
275 280 285

Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly
290 295 300

Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala
305 310 315 320

Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser
325 330 335

Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu
340 345 350

Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala
355 360 365

Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser Gly
370 375 380

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg
385 390 395 400

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
405 410 415

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
420 425 430

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
435 440 445

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr
450 455 460

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
465 470 475 480

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
485 490 495

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
500 505 510

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
515 520 525

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu
530 535 540

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
545 550 555 560

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
565 570 575

Lys Ile His Leu Arg Gln Lys Asp Lys
580 585

<210> 132
<211> 523
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 132
Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
20 25 30

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp
35 40 45

Met Leu Ile Asn Ser Arg Ser Ser Gly Ser Pro Lys Lys Lys Arg Lys
50 55 60

Val Gly Ser Gln Tyr Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu
65 70 75 80

Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys
85 90 95

Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile

100 105 110

Ala Val Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln
115 120 125

Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe
130 135 140

Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu
145 150 155 160

Ala Pro Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala Pro Ala Pro
165 170 175

Ala Pro Ala Met Val Ser Ala Leu Ala Gln Ala Pro Ala Pro Val Pro
180 185 190

Val Leu Ala Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala Pro Lys
195 200 205

Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu
210 215 220

Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp
225 230 235 240

Pro Ala Val Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu Phe Gln
245 250 255

Gln Leu Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr Glu Pro
260 265 270

Met Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val Thr Gly Ala
275 280 285

Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu
290 295 300

Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp
305 310 315 320

Met Asp Phe Ser Ala Leu Leu Gly Ser Gly Ser Gly Ser Arg Asp Ser
325 330 335

Arg Glu Gly Met Phe Leu Pro Lys Pro Glu Ala Gly Ser Ala Ile Ser
340 345 350

Asp Val Phe Glu Gly Arg Glu Val Cys Gln Pro Lys Arg Ile Arg Pro
355 360 365

Phe His Pro Pro Gly Ser Pro Trp Ala Asn Arg Pro Leu Pro Ala Ser
370 375 380

Leu Ala Pro Thr Pro Thr Gly Pro Val His Glu Pro Val Gly Ser Leu
385 390 395 400

Thr Pro Ala Pro Val Pro Gln Pro Leu Asp Pro Ala Pro Ala Val Thr
405 410 415

Pro Glu Ala Ser His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln
420 425 430

Ala Val Lys Ala Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys
435 440 445

Glu Glu Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro
450 455 460

Arg Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu
465 470 475 480

Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile Leu
485 490 495

Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser

500 505 510

Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe
515 520

<210> 133
<211> 50
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 133
Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
20 25 30

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp
35 40 45

Met Leu
50

<210> 134
<211> 275
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 134
Met Ser Gly Leu Glu Met Ala Asp His Met Met Ala Met Asn His Gly
1 5 10 15

Arg Phe Pro Asp Gly Thr Asn Gly Leu His His His Pro Ala His Arg
20 25 30

Met Gly Met Gly Gln Phe Pro Ser Pro His His His Gln Gln Gln Gln

35 40 45

Pro Gln His Ala Phe Asn Ala Leu Met Gly Glu His Ile His Tyr Gly
50 55 60

Ala Gly Asn Met Asn Ala Thr Ser Gly Ile Arg His Ala Met Gly Pro
65 70 75 80

Gly Thr Val Asn Gly Gly His Pro Pro Ser Ala Leu Ala Pro Ala Ala
85 90 95

Arg Phe Asn Asn Ser Gln Phe Met Gly Pro Pro Val Ala Ser Gln Gly
100 105 110

Gly Ser Leu Pro Ala Ser Met Gln Leu Gln Lys Leu Asn Asn Gln Tyr
115 120 125

Phe Asn His His Pro Tyr Pro His Asn His Tyr Met Pro Asp Leu His
130 135 140

Pro Ala Ala Gly His Gln Met Asn Gly Thr Asn Gln His Phe Arg Asp
145 150 155 160

Cys Asn Pro Lys His Ser Gly Gly Ser Ser Thr Pro Gly Gly Ser Gly
165 170 175

Gly Ser Ser Thr Pro Gly Gly Ser Gly Ser Ser Ser Gly Gly Gly Ala
180 185 190

Gly Ser Ser Asn Ser Gly Gly Gly Ser Gly Ser Gly Asn Met Pro Ala
195 200 205

Ser Val Ala His Val Pro Ala Ala Met Leu Pro Pro Asn Val Ile Asp
210 215 220

Thr Asp Phe Ile Asp Glu Glu Val Leu Met Ser Leu Val Ile Glu Met
225 230 235 240

Gly Leu Asp Arg Ile Lys Glu Leu Pro Glu Leu Trp Leu Gly Gln Asn
245 250 255

Glu Phe Asp Phe Met Thr Asp Phe Val Cys Lys Gln Gln Pro Ser Arg
260 265 270

Val Ser Cys
275

<210> 135
<211> 183
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 135
Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg Pro
1 5 10 15

Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu Pro
20 25 30

Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly Ala
35 40 45

Pro Leu Gly Pro Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr Gly
50 55 60

Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro Pro
65 70 75 80

Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr Pro
85 90 95

Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly Pro
100 105 110

Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala Leu
115 120 125

Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser Leu
130 135 140

Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu Phe
145 150 155 160

Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala Pro
165 170 175

Pro Ala Gly Ser Val Ser Cys
180

<210> 136
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<400> 136
Cys Xaa Cys Xaa His Xaa His
1 5

<210> 137

<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (8)..(10)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (12)..(12)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (14)..(14)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (16)..(16)
<223> 任意氨基酸

<400> 137
Cys Xaa Cys Xaa Cys Xaa His Xaa Xaa Xaa Cys Xaa Cys Xaa Cys Xaa
1 5 10 15

Cys

<210> 138
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (8)..(10)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (12)..(12)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (14)..(14)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (16)..(16)
<223> 任意氨基酸

<400> 138
Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa Xaa His Xaa Cys Xaa Cys Xaa
1 5 10 15

Cys

<210> 139
<211> 15
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (8)..(8)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (10)..(10)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (12)..(12)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (14)..(14)
<223> 任意氨基酸

<220>

<221> MOD_RES
<222> (15)..(15)
<223> Cys、His或Asp

<400> 139
Cys Xaa Cys Xaa His Xaa Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa
1 5 10 15

<210> 140
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (8)..(10)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (12)..(12)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (14)..(14)
<223> 任意氨基酸

<220>
<221> MOD_RES

<222> (16)..(16)
<223> 任意氨基酸

<400> 140
Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa Xaa Cys Xaa Cys Xaa Cys Xaa
1 5 10 15

Cys

<210> 141
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<400> 141
Cys Xaa Cys Xaa Cys Xaa His
1 5

<210> 142
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<400> 142
Cys Xaa Cys Xaa His Xaa Cys
1   5

<210> 143
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (8)..(10)
<223> 任意氨基酸

<220>

<221> MOD_RES
<222> (12)..(12)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (14)..(14)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (16)..(16)
<223> 任意氨基酸

<400> 143
Cys Xaa Cys Xaa His Xaa Cys Xaa Xaa Xaa Cys Xaa Cys Xaa His Xaa
1 5 10 15

Cys

<210> 144
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (8)..(10)

<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (12)..(12)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (14)..(14)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (16)..(16)
<223> 任意氨基酸

<400> 144
Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa Xaa Cys Xaa Cys Xaa His Xaa
1 5 10 15

Cys

<210> 145
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<400> 145
Cys Xaa Cys Xaa Cys Xaa Cys
1 5

<210> 146
<211> 17
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<221> MOD_RES
<222> (2)..(2)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (4)..(4)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (6)..(6)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (8)..(10)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (12)..(12)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (14)..(14)
<223> 任意氨基酸

<220>
<221> MOD_RES
<222> (16)..(16)
<223> 任意氨基酸

<400> 146

Cys Xaa Cys Xaa Cys Xaa Cys Xaa Xaa Xaa Cys Xaa His Xaa His Xaa
1 5 10 15

Cys

<210> 147
<211> 344
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<220>
<221> MISC_FEATURE
<222> (8)..(322)
<223> 該區域可包含1-15個"Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro"重複單元

<220>
<221> NON_CONS
<222> (18)..(19)

<220>
<221> NON_CONS
<222> (39)..(40)

<220>
<221> NON_CONS
<222> (60)..(61)

<220>
<221> NON_CONS
<222> (81)..(82)

<220>
<221> NON_CONS
<222> (102)..(103)

<220>
<221> NON_CONS
<222> (123)..(124)

<220>
<221> NON_CONS
<222> (144)..(145)

<220>
<221> NON_CONS
<222> (165)..(166)

<220>
<221> NON_CONS
<222> (186)..(187)

<220>
<221> NON_CONS
<222> (207)..(208)

<220>
<221> NON_CONS
<222> (228)..(229)

<220>
<221> NON_CONS
<222> (249)..(250)

<220>
<221> NON_CONS
<222> (270)..(271)

<220>
<221> NON_CONS
<222> (291)..(292)

<220>
<221> NON_CONS
<222> (312)..(313)

<220>
<221> NON_CONS
<222> (333)..(334)

<220>
<223> 關於置換和優選實施方案的詳細描述，參見提交的說明書

<400> 147
Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser
1 5 10 15

Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro
20 25 30

Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys

35 40 45

Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr
50 55 60

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
65 70 75 80

Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu
85 90 95

Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro
100 105 110

Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His
115 120 125

Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser
130 135 140

His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
145 150 155 160

Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr
165 170 175

Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr
180 185 190

Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His
195 200 205

Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly
210 215 220

Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
225 230 235 240

Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly
245 250 255

Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln
260 265 270

Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
275 280 285

Ser Phe Ser His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys
290 295 300

Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg Thr His Thr Gly Glu
305 310 315 320

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser His Gln Arg
325 330 335

Thr His Thr Gly Lys Lys Thr Ser
340

<210> 148
<211> 292
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<220>
<221> MOD_RES
<222> (8)..(57)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (8)..(57)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES

<222> (65)..(114)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (65)..(114)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (122)..(171)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (122)..(171)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (179)..(228)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (179)..(228)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (236)..(285)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (236)..(285)
<223> 該區域可包含1-50個殘基

<400> 148
Arg Ser Asp Asn Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp Asn Leu His Thr
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Arg Ser Asp Glu Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln Ser Gly Asn Leu
165 170 175

Thr Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
180 185 190

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
195 200 205

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
210 215 220

Xaa Xaa Xaa Xaa Thr Ser Gly His Leu Val Arg Xaa Xaa Xaa Xaa Xaa
225 230 235 240

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

245 250 255

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
260 265 270

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln Asn Ser
275 280 285

Thr Leu Thr Glu
290

<210> 149
<211> 292
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<220>
<221> MOD_RES
<222> (8)..(57)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (8)..(57)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (65)..(114)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (65)..(114)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (122)..(171)
<223> 任意氨基酸

<220>

<221> MISC_FEATURE
<222> (122)..(171)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (179)..(228)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (179)..(228)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (236)..(285)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (236)..(285)
<223> 該區域可包含1-50個殘基

<400> 149
Arg Ser Asp Asn Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa His Arg Thr Thr Leu Thr Asn
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

100 105 110

Xaa Xaa Arg Glu Asp Asn Leu His Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser His Ser Leu
165 170 175

Thr Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
180 185 190

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
195 200 205

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
210 215 220

Xaa Xaa Xaa Xaa Gln Ser Ser Ser Leu Val Arg Xaa Xaa Xaa Xaa Xaa
225 230 235 240

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
245 250 255

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
260 265 270

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp
275 280 285

Asn Leu His Thr
290

<210> 150
<211> 292
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<220>
<221> MOD_RES
<222> (8)..(57)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (8)..(57)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (65)..(114)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (65)..(114)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (122)..(171)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (122)..(171)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (179)..(228)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (179)..(228)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (236)..(285)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (236)..(285)
<223> 該區域可包含1-50個殘基

<400> 150
Asp Pro Gly Ala Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp Asn Leu Val Arg
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Gln Ser Gly Asp Leu Arg Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr His Leu Asp Leu
165 170 175

Ile Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
180 185 190

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
195 200 205

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
210 215 220

Xaa Xaa Xaa Xaa Thr Ser Gly Asn Leu Val Arg Xaa Xaa Xaa Xaa Xaa
225 230 235 240

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
245 250 255

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
260 265 270

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp
275 280 285

Asn Leu Val Arg
290

<210> 151
<211> 463
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<220>
<221> MOD_RES
<222> (8)..(57)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (8)..(57)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (65)..(114)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (65)..(114)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (122)..(171)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (122)..(171)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (179)..(228)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (179)..(228)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (236)..(285)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (236)..(285)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (293)..(342)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (293)..(342)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (350)..(399)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (350)..(399)
<223> 該區域可包含1-50個殘基

<220>
<221> MOD_RES
<222> (407)..(456)
<223> 任意氨基酸

<220>
<221> MISC_FEATURE
<222> (407)..(456)
<223> 該區域可包含1-50個殘基

<400> 151
Arg Arg Asp Glu Leu Asn Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp His Leu Thr Asn
50 55 60

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
65 70 75 80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
85 90 95

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Xaa Arg Ser Asp Asp Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa
115 120 125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
145 150 155 160

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Ser Asp Asn Leu
165 170 175

Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
180 185 190

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
195 200 205

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
210 215 220

Xaa Xaa Xaa Xaa His Arg Thr Thr Leu Thr Asn Xaa Xaa Xaa Xaa Xaa
225 230 235 240

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
245 250 255

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
260 265 270

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp
275 280 285

Asn Leu His Thr Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

290 295 300

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
305 310 315 320

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
325 330 335

Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser His Ser Leu Thr Glu Xaa Xaa Xaa
340 345 350

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
355 360 365

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
370 375 380

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gln
385 390 395 400

Ser Ser Ser Leu Val Arg Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
405 410 415

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
420 425 430

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
435 440 445

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Arg Glu Asp Asn Leu His Thr
450 455 460

<210> 152
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 152
Arg Ser Asp Asn Leu Val Arg
1 5

<210> 153
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 153
Arg Glu Asp Asn Leu His Thr
1 5

<210> 154
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 154
Arg Ser Asp Glu Leu Val Arg
1 5

<210> 155
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 155
Gln Ser Gly Asn Leu Thr Glu
1 5

<210> 156
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 156
Thr Ser Gly His Leu Val Arg
1 5

<210> 157
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 157
Gln Asn Ser Thr Leu Thr Glu
1 5

<210> 158
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 158
Asp Pro Gly Ala Leu Val Arg
1 5

<210> 159
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 159
His Arg Thr Thr Leu Thr Asn
1 5

<210> 160
<211> 7
<212> PRT

<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 160
Gln Ser Gly Asp Leu Arg Arg
1 5

<210> 161
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 161
Thr Ser His Ser Leu Thr Glu
1 5

<210> 162
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 162
Thr His Leu Asp Leu Ile Arg
1 5

<210> 163
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 163
Gln Ser Ser Ser Leu Val Arg
1 5

<210> 164

<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 164
Thr Ser Gly Asn Leu Val Arg
1 5

<210> 165
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 165
Arg Arg Asp Glu Leu Asn Val
1 5

<210> 166
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 166
Arg Ser Asp Asp Leu Val Arg
1 5

<210> 167
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 167
Arg Ser Asp His Leu Thr Asn
1 5

<210> 168
<211> 16
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 168
Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys
1 5 10 15

<210> 169
<211> 7
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 169
Pro Lys Lys Lys Arg Lys Val
1 5

<210> 170
<211> 16
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 170
Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 171
<211> 9
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 171
Tyr Pro Tyr Asp Val Pro Asp Tyr Ala

1 5

<210> 172
<211> 2009
<212> PRT
<213> 智人（Homo sapiens）

<400> 172
Met Glu Gln Thr Val Leu Val Pro Pro Gly Pro Asp Ser Phe Asn Phe
1 5 10 15

Phe Thr Arg Glu Ser Leu Ala Ala Ile Glu Arg Arg Ile Ala Glu Glu
20 25 30

Lys Ala Lys Asn Pro Lys Pro Asp Lys Lys Asp Asp Asp Glu Asn Gly
35 40 45

Pro Lys Pro Asn Ser Asp Leu Glu Ala Gly Lys Asn Leu Pro Phe Ile
50 55 60

Tyr Gly Asp Ile Pro Pro Glu Met Val Ser Glu Pro Leu Glu Asp Leu
65 70 75 80

Asp Pro Tyr Tyr Ile Asn Lys Lys Thr Phe Ile Val Leu Asn Lys Gly
85 90 95

Lys Ala Ile Phe Arg Phe Ser Ala Thr Ser Ala Leu Tyr Ile Leu Thr
100 105 110

Pro Phe Asn Pro Leu Arg Lys Ile Ala Ile Lys Ile Leu Val His Ser
115 120 125

Leu Phe Ser Met Leu Ile Met Cys Thr Ile Leu Thr Asn Cys Val Phe
130 135 140

Met Thr Met Ser Asn Pro Pro Asp Trp Thr Lys Asn Val Glu Tyr Thr
145 150 155 160

Phe Thr Gly Ile Tyr Thr Phe Glu Ser Leu Ile Lys Ile Ile Ala Arg

165 170 175

Gly Phe Cys Leu Glu Asp Phe Thr Phe Leu Arg Asp Pro Trp Asn Trp
180 185 190

Leu Asp Phe Thr Val Ile Thr Phe Ala Tyr Val Thr Glu Phe Val Asp
195 200 205

Leu Gly Asn Val Ser Ala Leu Arg Thr Phe Arg Val Leu Arg Ala Leu
210 215 220

Lys Thr Ile Ser Val Ile Pro Gly Leu Lys Thr Ile Val Gly Ala Leu
225 230 235 240

Ile Gln Ser Val Lys Lys Leu Ser Asp Val Met Ile Leu Thr Val Phe
245 250 255

Cys Leu Ser Val Phe Ala Leu Ile Gly Leu Gln Leu Phe Met Gly Asn
260 265 270

Leu Arg Asn Lys Cys Ile Gln Trp Pro Pro Thr Asn Ala Ser Leu Glu
275 280 285

Glu His Ser Ile Glu Lys Asn Ile Thr Val Asn Tyr Asn Gly Thr Leu
290 295 300

Ile Asn Glu Thr Val Phe Glu Phe Asp Trp Lys Ser Tyr Ile Gln Asp
305 310 315 320

Ser Arg Tyr His Tyr Phe Leu Glu Gly Phe Leu Asp Ala Leu Leu Cys
325 330 335

Gly Asn Ser Ser Asp Ala Gly Gln Cys Pro Glu Gly Tyr Met Cys Val
340 345 350

Lys Ala Gly Arg Asn Pro Asn Tyr Gly Tyr Thr Ser Phe Asp Thr Phe
355 360 365

Ser Trp Ala Phe Leu Ser Leu Phe Arg Leu Met Thr Gln Asp Phe Trp
370 375 380

Glu Asn Leu Tyr Gln Leu Thr Leu Arg Ala Ala Gly Lys Thr Tyr Met
385 390 395 400

Ile Phe Phe Val Leu Val Ile Phe Leu Gly Ser Phe Tyr Leu Ile Asn
405 410 415

Leu Ile Leu Ala Val Val Ala Met Ala Tyr Glu Glu Gln Asn Gln Ala
420 425 430

Thr Leu Glu Glu Ala Glu Gln Lys Glu Ala Glu Phe Gln Gln Met Ile
435 440 445

Glu Gln Leu Lys Lys Gln Gln Glu Ala Ala Gln Gln Ala Ala Thr Ala
450 455 460

Thr Ala Ser Glu His Ser Arg Glu Pro Ser Ala Ala Gly Arg Leu Ser
465 470 475 480

Asp Ser Ser Ser Glu Ala Ser Lys Leu Ser Ser Lys Ser Ala Lys Glu
485 490 495

Arg Arg Asn Arg Arg Lys Lys Arg Lys Gln Lys Glu Gln Ser Gly Gly
500 505 510

Glu Glu Lys Asp Glu Asp Glu Phe Gln Lys Ser Glu Ser Glu Asp Ser
515 520 525

Ile Arg Arg Lys Gly Phe Arg Phe Ser Ile Glu Gly Asn Arg Leu Thr
530 535 540

Tyr Glu Lys Arg Tyr Ser Ser Pro His Gln Ser Leu Leu Ser Ile Arg
545 550 555 560

Gly Ser Leu Phe Ser Pro Arg Arg Asn Ser Arg Thr Ser Leu Phe Ser

565 570 575

Phe Arg Gly Arg Ala Lys Asp Val Gly Ser Glu Asn Asp Phe Ala Asp
580 585 590

Asp Glu His Ser Thr Phe Glu Asp Asn Glu Ser Arg Arg Asp Ser Leu
595 600 605

Phe Val Pro Arg Arg His Gly Glu Arg Arg Asn Ser Asn Leu Ser Gln
610 615 620

Thr Ser Arg Ser Ser Arg Met Leu Ala Val Phe Pro Ala Asn Gly Lys
625 630 635 640

Met His Ser Thr Val Asp Cys Asn Gly Val Val Ser Leu Val Gly Gly
645 650 655

Pro Ser Val Pro Thr Ser Pro Val Gly Gln Leu Leu Pro Glu Val Ile
660 665 670

Ile Asp Lys Pro Ala Thr Asp Asp Asn Gly Thr Thr Thr Glu Thr Glu
675 680 685

Met Arg Lys Arg Arg Ser Ser Ser Phe His Val Ser Met Asp Phe Leu
690 695 700

Glu Asp Pro Ser Gln Arg Gln Arg Ala Met Ser Ile Ala Ser Ile Leu
705 710 715 720

Thr Asn Thr Val Glu Glu Leu Glu Glu Ser Arg Gln Lys Cys Pro Pro
725 730 735

Cys Trp Tyr Lys Phe Ser Asn Ile Phe Leu Ile Trp Asp Cys Ser Pro
740 745 750

Tyr Trp Leu Lys Val Lys His Val Val Asn Leu Val Val Met Asp Pro
755 760 765

```
Phe Val Asp Leu Ala Ile Thr Ile Cys Ile Val Leu Asn Thr Leu Phe
    770                 775                 780

Met Ala Met Glu His Tyr Pro Met Thr Asp His Phe Asn Asn Val Leu
785                 790                 795                 800

Thr Val Gly Asn Leu Val Phe Thr Gly Ile Phe Thr Ala Glu Met Phe
                805                 810                 815

Leu Lys Ile Ile Ala Met Asp Pro Tyr Tyr Tyr Phe Gln Glu Gly Trp
            820                 825                 830

Asn Ile Phe Asp Gly Phe Ile Val Thr Leu Ser Leu Val Glu Leu Gly
        835                 840                 845

Leu Ala Asn Val Glu Gly Leu Ser Val Leu Arg Ser Phe Arg Leu Leu
    850                 855                 860

Arg Val Phe Lys Leu Ala Lys Ser Trp Pro Thr Leu Asn Met Leu Ile
865                 870                 875                 880

Lys Ile Ile Gly Asn Ser Val Gly Ala Leu Gly Asn Leu Thr Leu Val
                885                 890                 895

Leu Ala Ile Ile Val Phe Ile Phe Ala Val Val Gly Met Gln Leu Phe
            900                 905                 910

Gly Lys Ser Tyr Lys Asp Cys Val Cys Lys Ile Ala Ser Asp Cys Gln
        915                 920                 925

Leu Pro Arg Trp His Met Asn Asp Phe Phe His Ser Phe Leu Ile Val
    930                 935                 940

Phe Arg Val Leu Cys Gly Glu Trp Ile Glu Thr Met Trp Asp Cys Met
945                 950                 955                 960

Glu Val Ala Gly Gln Ala Met Cys Leu Thr Val Phe Met Met Val Met
```

965 970 975

Val Ile Gly Asn Leu Val Val Leu Asn Leu Phe Leu Ala Leu Leu Leu
980 985 990

Ser Ser Phe Ser Ala Asp Asn Leu Ala Ala Thr Asp Asp Asp Asn Glu
995 1000 1005

Met Asn Asn Leu Gln Ile Ala Val Asp Arg Met His Lys Gly Val
1010 1015 1020

Ala Tyr Val Lys Arg Lys Ile Tyr Glu Phe Ile Gln Gln Ser Phe
1025 1030 1035

Ile Arg Lys Gln Lys Ile Leu Asp Glu Ile Lys Pro Leu Asp Asp
1040 1045 1050

Leu Asn Asn Lys Lys Asp Ser Cys Met Ser Asn His Thr Ala Glu
1055 1060 1065

Ile Gly Lys Asp Leu Asp Tyr Leu Lys Asp Val Asn Gly Thr Thr
1070 1075 1080

Ser Gly Ile Gly Thr Gly Ser Ser Val Glu Lys Tyr Ile Ile Asp
1085 1090 1095

Glu Ser Asp Tyr Met Ser Phe Ile Asn Asn Pro Ser Leu Thr Val
1100 1105 1110

Thr Val Pro Ile Ala Val Gly Glu Ser Asp Phe Glu Asn Leu Asn
1115 1120 1125

Thr Glu Asp Phe Ser Ser Glu Ser Asp Leu Glu Glu Ser Lys Glu
1130 1135 1140

Lys Leu Asn Glu Ser Ser Ser Ser Ser Glu Gly Ser Thr Val Asp
1145 1150 1155

Ile Gly Ala Pro Val Glu Glu Gln Pro Val Val Glu Pro Glu Glu
1160 1165 1170

Thr Leu Glu Pro Glu Ala Cys Phe Thr Glu Gly Cys Val Gln Arg
1175 1180 1185

Phe Lys Cys Cys Gln Ile Asn Val Glu Glu Gly Arg Gly Lys Gln
1190 1195 1200

Trp Trp Asn Leu Arg Arg Thr Cys Phe Arg Ile Val Glu His Asn
1205 1210 1215

Trp Phe Glu Thr Phe Ile Val Phe Met Ile Leu Leu Ser Ser Gly
1220 1225 1230

Ala Leu Ala Phe Glu Asp Ile Tyr Ile Asp Gln Arg Lys Thr Ile
1235 1240 1245

Lys Thr Met Leu Glu Tyr Ala Asp Lys Val Phe Thr Tyr Ile Phe
1250 1255 1260

Ile Leu Glu Met Leu Leu Lys Trp Val Ala Tyr Gly Tyr Gln Thr
1265 1270 1275

Tyr Phe Thr Asn Ala Trp Cys Trp Leu Asp Phe Leu Ile Val Asp
1280 1285 1290

Val Ser Leu Val Ser Leu Thr Ala Asn Ala Leu Gly Tyr Ser Glu
1295 1300 1305

Leu Gly Ala Ile Lys Ser Leu Arg Thr Leu Arg Ala Leu Arg Pro
1310 1315 1320

Leu Arg Ala Leu Ser Arg Phe Glu Gly Met Arg Val Val Val Asn
1325 1330 1335

Ala Leu Leu Gly Ala Ile Pro Ser Ile Met Asn Val Leu Leu Val

1340 1345 1350

Cys Leu Ile Phe Trp Leu Ile Phe Ser Ile Met Gly Val Asn Leu
1355 1360 1365

Phe Ala Gly Lys Phe Tyr His Cys Ile Asn Thr Thr Thr Gly Asp
1370 1375 1380

Arg Phe Asp Ile Glu Asp Val Asn Asn His Thr Asp Cys Leu Lys
1385 1390 1395

Leu Ile Glu Arg Asn Glu Thr Ala Arg Trp Lys Asn Val Lys Val
1400 1405 1410

Asn Phe Asp Asn Val Gly Phe Gly Tyr Leu Ser Leu Leu Gln Val
1415 1420 1425

Ala Thr Phe Lys Gly Trp Met Asp Ile Met Tyr Ala Ala Val Asp
1430 1435 1440

Ser Arg Asn Val Glu Leu Gln Pro Lys Tyr Glu Glu Ser Leu Tyr
1445 1450 1455

Met Tyr Leu Tyr Phe Val Ile Phe Ile Ile Phe Gly Ser Phe Phe
1460 1465 1470

Thr Leu Asn Leu Phe Ile Gly Val Ile Ile Asp Asn Phe Asn Gln
1475 1480 1485

Gln Lys Lys Lys Phe Gly Gly Gln Asp Ile Phe Met Thr Glu Glu
1490 1495 1500

Gln Lys Lys Tyr Tyr Asn Ala Met Lys Lys Leu Gly Ser Lys Lys
1505 1510 1515

Pro Gln Lys Pro Ile Pro Arg Pro Gly Asn Lys Phe Gln Gly Met
1520 1525 1530

```
Val Phe  Asp Phe Val Thr Arg  Gln Val Phe Asp Ile  Ser Ile Met
    1535                1540                1545

Ile Leu  Ile Cys Leu Asn Met  Val Thr Met Met Val  Glu Thr Asp
    1550                1555                1560

Asp Gln  Ser Glu Tyr Val Thr  Thr Ile Leu Ser Arg  Ile Asn Leu
    1565                1570                1575

Val Phe  Ile Val Leu Phe Thr  Gly Glu Cys Val Leu  Lys Leu Ile
    1580                1585                1590

Ser Leu  Arg His Tyr Tyr Phe  Thr Ile Gly Trp Asn  Ile Phe Asp
    1595                1600                1605

Phe Val  Val Val Ile Leu Ser  Ile Val Gly Met Phe  Leu Ala Glu
    1610                1615                1620

Leu Ile  Glu Lys Tyr Phe Val  Ser Pro Thr Leu Phe  Arg Val Ile
    1625                1630                1635

Arg Leu  Ala Arg Ile Gly Arg  Ile Leu Arg Leu Ile  Lys Gly Ala
    1640                1645                1650

Lys Gly  Ile Arg Thr Leu Leu  Phe Ala Leu Met Met  Ser Leu Pro
    1655                1660                1665

Ala Leu  Phe Asn Ile Gly Leu  Leu Leu Phe Leu Val  Met Phe Ile
    1670                1675                1680

Tyr Ala  Ile Phe Gly Met Ser  Asn Phe Ala Tyr Val  Lys Arg Glu
    1685                1690                1695

Val Gly  Ile Asp Asp Met Phe  Asn Phe Glu Thr Phe  Gly Asn Ser
    1700                1705                1710

Met Ile  Cys Leu Phe Gln Ile  Thr Thr Ser Ala Gly  Trp Asp Gly
```

1715 1720 1725

Leu Leu Ala Pro Ile Leu Asn Ser Lys Pro Pro Asp Cys Asp Pro
1730 1735 1740

Asn Lys Val Asn Pro Gly Ser Ser Val Lys Gly Asp Cys Gly Asn
1745 1750 1755

Pro Ser Val Gly Ile Phe Phe Phe Val Ser Tyr Ile Ile Ile Ser
1760 1765 1770

Phe Leu Val Val Val Asn Met Tyr Ile Ala Val Ile Leu Glu Asn
1775 1780 1785

Phe Ser Val Ala Thr Glu Glu Ser Ala Glu Pro Leu Ser Glu Asp
1790 1795 1800

Asp Phe Glu Met Phe Tyr Glu Val Trp Glu Lys Phe Asp Pro Asp
1805 1810 1815

Ala Thr Gln Phe Met Glu Phe Glu Lys Leu Ser Gln Phe Ala Ala
1820 1825 1830

Ala Leu Glu Pro Pro Leu Asn Leu Pro Gln Pro Asn Lys Leu Gln
1835 1840 1845

Leu Ile Ala Met Asp Leu Pro Met Val Ser Gly Asp Arg Ile His
1850 1855 1860

Cys Leu Asp Ile Leu Phe Ala Phe Thr Lys Arg Val Leu Gly Glu
1865 1870 1875

Ser Gly Glu Met Asp Ala Leu Arg Ile Gln Met Glu Glu Arg Phe
1880 1885 1890

Met Ala Ser Asn Pro Ser Lys Val Ser Tyr Gln Pro Ile Thr Thr
1895 1900 1905

Thr Leu Lys Arg Lys Gln Glu Glu Val Ser Ala Val Ile Ile Gln
1910 1915 1920

Arg Ala Tyr Arg Arg His Leu Leu Lys Arg Thr Val Lys Gln Ala
1925 1930 1935

Ser Phe Thr Tyr Asn Lys Asn Lys Ile Lys Gly Gly Ala Asn Leu
1940 1945 1950

Leu Ile Lys Glu Asp Met Ile Ile Asp Arg Ile Asn Glu Asn Ser
1955 1960 1965

Ile Thr Glu Lys Thr Asp Leu Thr Met Ser Thr Ala Ala Cys Pro
1970 1975 1980

Pro Ser Tyr Asp Arg Val Thr Lys Pro Ile Val Glu Lys His Glu
1985 1990 1995

Gln Glu Gly Lys Asp Glu Lys Ala Lys Gly Lys
2000 2005

<210> 173
<211> 1052
<212> PRT
<213> 智人（Homo sapiens）

<400> 173
Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val Gly
1 5 10 15

Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly Val
20 25 30

Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg Ser
35 40 45

Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg Arg His Arg Ile Gln
50 55 60

```
Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His Ser
65                  70                  75                  80

Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu Ser
                85                  90                  95

Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu Ala
            100                 105                 110

Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr Gly
        115                 120                 125

Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala Leu
    130                 135                 140

Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys Asp
145                 150                 155                 160

Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr Val
                165                 170                 175

Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln Leu
            180                 185                 190

Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg Arg
        195                 200                 205

Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys Asp
    210                 215                 220

Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe Pro
225                 230                 235                 240

Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr Asn
                245                 250                 255

Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn Glu
```

260 265 270

Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe Lys
275 280 285

Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu Val
290 295 300

Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys Pro
305 310 315 320

Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr Ala
325 330 335

Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala Lys
340 345 350

Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu Thr
355 360 365

Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser Asn
370 375 380

Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile Asn
385 390 395 400

Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala Ile
405 410 415

Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln Gln
420 425 430

Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro Val
435 440 445

Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile Ile
450 455 460

Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg Glu
465 470 475 480

Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys Arg
485 490 495

Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr Gly
500 505 510

Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp Met
515 520 525

Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu Asp
530 535 540

Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro Arg
545 550 555 560

Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys Gln
565 570 575

Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu Ser
580 585 590

Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile Leu
595 600 605

Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu Tyr
610 615 620

Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp Phe
625 630 635 640

Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu Met
645 650 655

Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys Val

660 665 670

Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp Lys
675 680 685

Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp Ala
690 695 700

Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys Leu
705 710 715 720

Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys Gln
725 730 735

Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu Ile
740 745 750

Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp Tyr
755 760 765

Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile Asn
770 775 780

Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu Ile
785 790 795 800

Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu Lys
805 810 815

Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His Asp
820 825 830

Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly Asp
835 840 845

Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr Leu
850 855 860

Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile Lys
865 870 875 880

Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp Tyr
885 890 895

Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr Arg
900 905 910

Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val Lys
915 920 925

Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser Lys
930 935 940

Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala Glu
945 950 955 960

Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly Glu
965 970 975

Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu
980 985 990

Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met Asn
995 1000 1005

Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys Thr
1010 1015 1020

Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu Tyr
1025 1030 1035

Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
1040 1045 1050

<210> 174

<211> 1148
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 174

```
Met Ala Pro Lys Lys Lys Arg Lys Val Gly Ile His Gly Val Pro Ala
1               5                  10                  15

Ala Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
            20                  25                  30

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
        35                  40                  45

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
    50                  55                  60

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
65                  70                  75                  80

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
                85                  90                  95

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
            100                 105                 110

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
        115                 120                 125

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
    130                 135                 140

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
145                 150                 155                 160

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys
                165                 170                 175
```

Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr
180 185 190

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln
195 200 205

Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg
210 215 220

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys
225 230 235 240

Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
245 250 255

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
260 265 270

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
275 280 285

Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
290 295 300

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
305 310 315 320

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
325 330 335

Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
340 345 350

Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
355 360 365

Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
370 375 380

Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
385 390 395 400

Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
405 410 415

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
420 425 430

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
435 440 445

Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
450 455 460

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
465 470 475 480

Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
485 490 495

Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
500 505 510

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
515 520 525

Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
530 535 540

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
545 550 555 560

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
565 570 575

Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
580 585 590

Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
595 600 605

Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
610 615 620

Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
625 630 635 640

Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
645 650 655

Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
660 665 670

Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
675 680 685

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
690 695 700

Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
705 710 715 720

Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
725 730 735

Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
740 745 750

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
755 760 765

```
Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
    770                 775                 780

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
785                 790                 795                 800

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
                805                 810                 815

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
            820                 825                 830

Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
        835                 840                 845

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
    850                 855                 860

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
865                 870                 875                 880

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
                885                 890                 895

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
            900                 905                 910

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
        915                 920                 925

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
    930                 935                 940

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
945                 950                 955                 960

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
                965                 970                 975
```

Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
980 985 990

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
995 1000 1005

Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn
1010 1015 1020

Met Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser
1025 1030 1035

Lys Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn
1040 1045 1050

Leu Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys
1055 1060 1065

Gly Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys
1070 1075 1080

Lys Lys Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Leu Glu
1085 1090 1095

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala
1100 1105 1110

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp
1115 1120 1125

Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
1130 1135 1140

Asp Leu Asp Met Leu
1145

<210> 175
<211> 387
<212> PRT
<213> 智人（Homo sapiens）

<400> 175

```
Met Thr Gly Lys Leu Ala Glu Lys Leu Pro Val Thr Met Ser Ser Leu
1               5                  10                  15

Leu Asn Gln Leu Pro Asp Asn Leu Tyr Pro Glu Glu Ile Pro Ser Ala
            20                  25                  30

Leu Asn Leu Phe Ser Gly Ser Ser Asp Ser Val Val His Tyr Asn Gln
        35                  40                  45

Met Ala Thr Glu Asn Val Met Asp Ile Gly Leu Thr Asn Glu Lys Pro
    50                  55                  60

Asn Pro Glu Leu Ser Tyr Ser Gly Ser Phe Gln Pro Ala Pro Gly Asn
65                  70                  75                  80

Lys Thr Val Thr Tyr Leu Gly Lys Phe Ala Phe Asp Ser Pro Ser Asn
                85                  90                  95

Trp Cys Gln Asp Asn Ile Ile Ser Leu Met Ser Ala Gly Ile Leu Gly
            100                 105                 110

Val Pro Pro Ala Ser Gly Ala Leu Ser Thr Gln Thr Ser Thr Ala Ser
        115                 120                 125

Met Val Gln Pro Pro Gln Gly Asp Val Glu Ala Met Tyr Pro Ala Leu
    130                 135                 140

Pro Pro Tyr Ser Asn Cys Gly Asp Leu Tyr Ser Glu Pro Val Ser Phe
145                 150                 155                 160

His Asp Pro Gln Gly Asn Pro Gly Leu Ala Tyr Ser Pro Gln Asp Tyr
                165                 170                 175
```

Gln Ser Ala Lys Pro Ala Leu Asp Ser Asn Leu Phe Pro Met Ile Pro
180 185 190

Asp Tyr Asn Leu Tyr His His Pro Asn Asp Met Gly Ser Ile Pro Glu
195 200 205

His Lys Pro Phe Gln Gly Met Asp Pro Ile Arg Val Asn Pro Pro Pro
210 215 220

Ile Thr Pro Leu Glu Thr Ile Lys Ala Phe Lys Asp Lys Gln Ile His
225 230 235 240

Pro Gly Phe Gly Ser Leu Pro Gln Pro Pro Leu Thr Leu Lys Pro Ile
245 250 255

Arg Pro Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Leu His Glu
260 265 270

Arg Pro His Ala Cys Pro Ala Glu Gly Cys Asp Arg Arg Phe Ser Arg
275 280 285

Ser Asp Glu Leu Thr Arg His Leu Arg Ile His Thr Gly His Lys Pro
290 295 300

Phe Gln Cys Arg Ile Cys Met Arg Ser Phe Ser Arg Ser Asp His Leu
305 310 315 320

Thr Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Glu
325 330 335

Phe Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Ala
340 345 350

Lys Ile His Leu Lys Gln Lys Glu Lys Lys Ala Glu Lys Gly Gly Ala
355 360 365

Pro Ser Ala Ser Ser Ala Pro Pro Val Ser Leu Ala Pro Val Val Thr
370 375 380

Thr Cys Ala
385

<210> 176
<211> 543
<212> PRT
<213> 智人（Homo sapiens）

<400> 176
Met Ala Ala Ala Lys Ala Glu Met Gln Leu Met Ser Pro Leu Gln Ile
1 5 10 15

Ser Asp Pro Phe Gly Ser Phe Pro His Ser Pro Thr Met Asp Asn Tyr
20 25 30

Pro Lys Leu Glu Glu Met Met Leu Leu Ser Asn Gly Ala Pro Gln Phe
35 40 45

Leu Gly Ala Ala Gly Ala Pro Glu Gly Ser Gly Ser Asn Ser Ser Ser
50 55 60

Ser Ser Ser Gly Gly Gly Gly Gly Gly Gly Gly Gly Ser Asn Ser Ser
65 70 75 80

Ser Ser Ser Ser Thr Phe Asn Pro Gln Ala Asp Thr Gly Glu Gln Pro
85 90 95

Tyr Glu His Leu Thr Ala Glu Ser Phe Pro Asp Ile Ser Leu Asn Asn
100 105 110

Glu Lys Val Leu Val Glu Thr Ser Tyr Pro Ser Gln Thr Thr Arg Leu
115 120 125

Pro Pro Ile Thr Tyr Thr Gly Arg Phe Ser Leu Glu Pro Ala Pro Asn
130 135 140

Ser Gly Asn Thr Leu Trp Pro Glu Pro Leu Phe Ser Leu Val Ser Gly
145 150 155 160

Leu Val Ser Met Thr Asn Pro Pro Ala Ser Ser Ser Ser Ala Pro Ser
165 170 175

Pro Ala Ala Ser Ser Ala Ser Ala Ser Gln Ser Pro Pro Leu Ser Cys
180 185 190

Ala Val Pro Ser Asn Asp Ser Ser Pro Ile Tyr Ser Ala Ala Pro Thr
195 200 205

Phe Pro Thr Pro Asn Thr Asp Ile Phe Pro Glu Pro Gln Ser Gln Ala
210 215 220

Phe Pro Gly Ser Ala Gly Thr Ala Leu Gln Tyr Pro Pro Pro Ala Tyr
225 230 235 240

Pro Ala Ala Lys Gly Gly Phe Gln Val Pro Met Ile Pro Asp Tyr Leu
245 250 255

Phe Pro Gln Gln Gln Gly Asp Leu Gly Leu Gly Thr Pro Asp Gln Lys
260 265 270

Pro Phe Gln Gly Leu Glu Ser Arg Thr Gln Gln Pro Ser Leu Thr Pro
275 280 285

Leu Ser Thr Ile Lys Ala Phe Ala Thr Gln Ser Gly Ser Gln Asp Leu
290 295 300

Lys Ala Leu Asn Thr Ser Tyr Gln Ser Gln Leu Ile Lys Pro Ser Arg
305 310 315 320

Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His Glu Arg
325 330 335

Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser
340 345 350

Asp Glu Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe
355 360 365

Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu Thr
370 375 380

Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile
385 390 395 400

Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Thr Lys
405 410 415

Ile His Leu Arg Gln Lys Asp Lys Lys Ala Asp Lys Ser Val Val Ala
420 425 430

Ser Ser Ala Thr Ser Ser Leu Ser Ser Tyr Pro Ser Pro Val Ala Thr
435 440 445

Ser Tyr Pro Ser Pro Val Thr Thr Ser Tyr Pro Ser Pro Ala Thr Thr
450 455 460

Ser Tyr Pro Ser Pro Val Pro Thr Ser Phe Ser Ser Pro Gly Ser Ser
465 470 475 480

Thr Tyr Pro Ser Pro Val His Ser Gly Phe Pro Ser Pro Ser Val Ala
485 490 495

Thr Thr Tyr Ser Ser Val Pro Pro Ala Phe Pro Ala Gln Val Ser Ser
500 505 510

Phe Pro Ser Ser Ala Val Thr Asn Ser Phe Ser Ala Ser Thr Gly Leu
515 520 525

Ser Asp Met Thr Ala Thr Phe Ser Pro Arg Thr Ile Glu Ile Cys
530 535 540

<210> 177
<211> 8

<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 177
Gly Gly Ser Gly Gly Gly Ser Gly
1 5

<210> 178
<211> 16
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<400> 178
Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly
1 5 10 15

<210> 179
<211> 4
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<223> 關於置換和優選實施方案的詳細描述，參見提交的說明書

<400> 179
Gly Gly Gly Ser
1

<210> 180
<211> 5
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<223> 關於置換和優選實施方案的詳細描述，參見提交的說明書

<400> 180
Gly Gly Gly Gly Ser
1 5

<210> 181
<211> 4
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成肽

<220>
<223> 關於置換和優選實施方案的詳細描述，參見提交的說明書

<400> 181
Gly Gly Ser Gly
1

<210> 182
<211> 9
<212> DNA
<213> 智人（Homo sapiens）

<400> 182
gcgkgggcg 9

<210> 183
<211> 76
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成寡核苷酸

<400> 183
gttttagtac tctggaaaca gaatctacta aaacaaggca aaatgccgtg tttatctcgt 60

caacttgttg gcgaga 76

<210> 184
<211> 3871
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 184
ggaggaagcc atcaactaaa ctacaatgac tgtaagatac aaaattggga atggtaacat 60

attttgaagt tctgttgaca taaagaatca tgatattaat gcccatggaa atgaaagggc 120

gatcaacact atggtttgaa aaggggggaaa ttgtagagca cagatgtgtt cgtgtggcag 180

tgtgctgtct ctagcaatac tcagagaaga gagagaacaa tgaaattctg attggcccca 240

gtgtgagccc agatgaggtt cagctgccaa ctttctcttt cacatcttat gaaagtcatt 300

taagcacaac taacttttttt ttttttttttt ttttttgag acagagtctt gctctgttgc 360

ccaggacaga gtgcagtagt gactcaatct cggctcactg cagcctccac ctcctaggct 420

caaacggtcc tcctgcatca gcctcccaag tagctggaat tacaggagtg gcccaccatg 480

cccagctaat ttttgtattt ttaatagata cgggggtttc accatatcac ccaggctggt 540

ctcgaactcc tggcctcaag tgatccacct gcctcggcct cccaaagtgc tgggattata 600

ggcgtcagcc actatgccca acccgaccaa cctttttttaa aataaatatt taaaaaattg 660

gtatttcaca tatatactag tatttacatt tatccacaca aaacggacgg gcctccgctg 720

aaccagtgag gccccagacg tgcgcataaa taacccctgc gtgctgcacc acctggggag 780

agggggagga ccacggtaaa tggagcgagc gcatagcaaa agggacgcgg ggtccttttc 840

tctgccggtg gcactgggta gctgtggcca ggtgtggtac tttgatgggg cccagggctg 900

gagctcaagg aagcgtcgca gggtcacaga tctgggggaa ccccggggaa aagcactgag 960

gcaaaaccgc cgctcgtctc ctacaatata tgggaggggg aggttgagta cgttctggat 1020

tactcataag acctttttttt tttccttccg ggcgcaaaac cgtgagctgg atttataatc 1080

gccctataaa gctccagagg cggtcaggca cctgcagagg agccccgccg ctccgccgac 1140

tagctgcccc cgcgagcaac ggcctcgtga tttccccgcc gatccggtcc ccgcctcccc 1200

actctgcccc cgcctacccc ggagccgtgc agccgcctct ccgaatctct ctcttctcct 1260

ggcgctcgcg tgcgagaggg aactagcgag aacgaggaag cagctggagg tgacgccggg 1320

cagattacgc ctgtcagggc cgagccgagc ggatcgctgg gcgctgtgca gaggaaaggc 1380

gggagtgccc ggctcgctgt cgcagagccg aggtgggtaa gctagcgacc acctggactt 1440

cccagcgccc aaccgtggct tttcagccag gtcctctcct cccgcggctt ctcaaccaac 1500

cccatcccag cgccggccac ccaacctccc gaaatgagtg cttcctgccc cagcagccga 1560

aggcgctact aggaacggta acctgttact tttccagggg ccgtagtcga cccgctgccc 1620

gagttgctgt gcgactgcgc gcgcggggct agagtgcaag gtgactgtgg ttcttctctg 1680

gccaagtccg agggagaacg taaagatatg ggccttttttc cccctctcac cttgtctcac 1740

caaagtccct agtccccgga gcagttagcc tctttctttc cagggaatta gccagacaca 1800

acaacgggaa ccagacaccg aaccagacat gcccgccccg tgcgccctcc ccgctcgctg 1860

cctttcctcc ctcttgtctc tccagagccg gatcttcaag gggagcctcc gtgccccccgg 1920

ctgctcagtc cctccggtgt gcaggacccc ggaagtcctc cccgcacagc tctcgcttct 1980

ctttgcagcc tgtttctgcg ccggaccagt cgaggactct ggacagtaga ggccccggga 2040

cgaccgagct ggaattcgcc accatggccg ccgaccacct gatgctcgcc gagggctacc 2100

gcctggtgca gaggccgccg tccgccgccg ccgcccatgg ccctcatgcg ctccggactc 2160

tgccgccgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg gctccgctgg 2220

ggccgccgcc gccccgccaa cccggggccc tggcgtacgg ggccttcggg ccgccgtcct 2280

ccttccagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac ctgcagcctg 2340

tggcgacgcc gtaccccggc cgcgcggccg cgcccccccaa cgctccggga ggcccccccgg 2400

gcccgcagcc ggccccaagc gccgcagccc cgccgccgcc cgcgcacgcc ctgggcggca 2460

tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag ctggggctgc 2520

accgcgtgcg cgagctgccc gagctgttcc tgggccagag cgagttcgac tgcttctcgg 2580

acttggggtc cgcgccgccc gccggctccg tgagctgcgg tggttctggt ggtggttctg 2640

gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg cccagcaaga 2700

cgccccccca cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc cgcttctccc 2760

gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc ttccagtgcc 2820

gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc cgcacccaca 2880

caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg agcgatgaac 2940

ttgtccgaca taccaagatc cacttgcggc agaaggaccg cccttacgct tgcccagtgg 3000

agtcctgtga tcgccgcttc tcccaatcag ggaatctgac tgagcacatc cgcatccaca 3060

caggccagaa gcccttccag tgccgcatct gcatgagaaa cttcagcaca agtggacatc 3120

tggtacgcca catccgcacc cacacaggcg aaaagcccct cgcctgcgac atctgtggaa 3180

gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg cggcagaagg 3240

acaagtaact cgaggaaacc cagcagacaa tgtagctaga cccagtagcc agatgtagct 3300

aaagagaccg gttcactgtg agaaacccag cagacaatgt agctagaccc agtagccaga 3360

tgtagctaaa gagaccggtt cactgtgaaa gcttgggtgg catccctgtg acccctcccc 3420

agtgcctctc ctggccctgg aagttgccac tccagtgccc accagccttg tcctaataaa 3480

attaagttgc atcattttgt ctgactaggt gtccttctat aatattatgg ggtggagggg 3540

ggtggtatgg agcaaggggc aagttgggaa gacaacctgt agggcctgcg gggtctattg 3600

ggaaccaagc tggagtgcag tggcacaatc ttggctcact gcaatctccg cctcctgggt 3660

tcaagcgatt ctcctgcctc agcctcccga gttgttggga ttccaggcat gcatgaccag 3720

gctcagctaa tttttgtttt tttggtagag acggggtttc accatattgg ccaggctggt 3780

ctccaactcc taatctcagg tgatctaccc accttggcct cccaaattgc tgggattaca 3840

ggcgtgaacc actgctccct tccctgtcct t 3871

<210> 185
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 185
tgtctcggca ttgagaacat tc 22

<210> 186
<211> 21
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 186
attggtggga ggccattgta t 21

<210> 187
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 187
accacagtcc atgccatcac 20

<210> 188
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 188
tccaccaccc tgttgctgta 20

<210> 189
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 189
caaaaaagcc acaaaagcct 20

<210> 190
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 190
ttagctccgc aagaaacatc 20

<210> 191
<211> 23
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 191
ccatggaact ggctcgattt cac 23

<210> 192
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 192
attggtggga ggccactgta t 21

<210> 193
<211> 27
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成探針

<400> 193
aggcctgaaa accattgtgg gagccct 27

<210> 194
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 194
gaatgtggga aatcattcag tcgc 24

<210> 195
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 195
gcaagttatc ctctcgtgag aagg 24

<210> 196
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成探針

<400> 196
gcgacaacct ggtgagacat caacgcacc 29

<210> 197
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 197
gctgttatct cttgtgggct gt 22

<210> 198
<211> 22
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 198

aaactcatgg gagctgccgg tt 22

<210> 199
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成探針

<400> 199
ccacacaaat ctctccctgg cattg 25

<210> 200
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 200
gaatgtggga aatcattcag tcgc 24

<210> 201
<211> 24
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成引物

<400> 201
gcaagttatc ctctcgtgag aagg 24

<210> 202
<211> 29
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成探針

<400> 202
gcgacaacct ggtgagacat caacgcacc 29

<210> 203
<211> 1182
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 203
atggccgccg accacctgat gctcgccgag ggctaccgcc tggtgcagag gccgccgtcc 60

gccgccgccg cccatggccc tcatgcgctc cggactctgc cgccgtacgc gggccccggc 120

ctggacagtg ggctgaggcc gcgggggggct ccgctggggc cgccgccgcc ccgccaaccc 180

ggggccctgg cgtacggggc cttcgggccg ccgtcctcct tccagccctt tccggccgtg 240

cctccgccgg ccgcgggcat cgcgcacctg cagcctgtgg cgacgccgta ccccggccgc 300

gcggccgcgc cccccaacgc tccgggaggc cccccgggcc cgcagccggc cccaagcgcc 360

gcagccccgc cgccgccccgc gcacgccctg ggcggcatgg acgccgaact catcgacgag 420

gaggcgctga cgtcgctgga gctggagctg gggctgcacc gcgtgcgcga gctgcccgag 480

ctgttcctgg gccagagcga gttcgactgc ttctcggact tggggtccgc gccgcccgcc 540

ggctccgtga gctgcggtgg ttctggtggt ggttctggtc agtcccagct catcaaaccc 600

agccgcatgc gcaagtaccc caaccggccc agcaagacgc ccccccacga acgcccttac 660

gcttgcccag tggagtcctg tgatcgccgc ttctcccgca gcgacaacct ggtgagacac 720

atccgcatcc acacaggcca gaagcccttc cagtgccgca tctgcatgag aaacttcagc 780

cgagaggata acttgcacac tcacatccgc acccacacag gcgaaaagcc cttcgcctgc 840

gacatctgtg gaagaaagtt tgcccggagc gatgaacttg tccgacatac caagatccac 900

ttgcggcaga aggaccgccc ttacgcttgc ccagtggagt cctgtgatcg ccgcttctcc 960

caatcaggga atctgactga gcacatccgc atccacacag gccagaagcc cttccagtgc 1020

cgcatctgca tgagaaactt cagcacaagt ggacatctgg tacgccacat ccgcacccac 1080

acaggcgaaa agcccttcgc ctgcgacatc tgtggaagaa agtttgccca gaatagtacc 1140

ctgaccgaac ataccaagat ccacttgcgg cagaaggaca ag 1182

<210> 204
<211> 1206
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 204
atggccgccg accacctgat gctcgccgag ggctaccgcc tggtgcagag gccgccgtcc 60

gccgccgccg cccatggccc tcatgcgctc cggactctgc cgccgtacgc gggccccggc 120

ctggacagtg ggctgaggcc gcgggggggct ccgctggggc cgccgccgcc ccgccaaccc 180

ggggccctgg cgtacggggc cttcgggccg ccgtcctcct tccagccctt tccggccgtg 240

cctccgccgg ccgcgggcat cgcgcacctg cagcctgtgg cgacgccgta ccccggccgc 300

gcggccgcgc cccccaacgc tccgggaggc cccccgggcc cgcagccggc cccaagcgcc 360

gcagccccgc cgccgccccgc gcacgccctg ggcggcatgg acgccgaact catcgacgag 420

gaggcgctga cgtcgctgga gctggagctg gggctgcacc gcgtgcgcga gctgcccgag 480

ctgttcctgg gccagagcga gttcgactgc ttctcggact tggggtccgc gccgcccgcc 540

ggctccgtga gctgcggtgg ttctggtggt ggttctggtg gtggcagcgg gggaggttct 600

ggtcagtccc agctcatcaa acccagccgc atgcgcaagt accccaaccg gcccagcaag 660

acgccccccc acgaacgccc ttacgcttgc ccagtggagt cctgtgatcg ccgcttctcc 720

cgcagcgaca acctggtgag acacatccgc atccacacag gccagaagcc cttccagtgc 780

cgcatctgca tgagaaactt cagccgagag gataacttgc acactcacat ccgcacccac 840

acaggcgaaa agcccttcgc ctgcgacatc tgtggaagaa agtttgcccg gagcgatgaa 900

cttgtccgac ataccaagat ccacttgcgg cagaaggacc gcccttacgc ttgcccagtg 960

gagtcctgtg atcgccgctt ctcccaatca gggaatctga ctgagcacat ccgcatccac 1020

acaggccaga agcccttcca gtgccgcatc tgcatgagaa acttcagcac aagtggacat 1080

ctggtacgcc acatccgcac ccacacaggc gaaaagccct tcgcctgcga catctgtgga 1140

agaaagtttg cccagaatag taccctgacc gaacatacca agatccactt gcggcagaag 1200

gacaag 1206

<210> 205
<211> 402
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 205

```
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1               5                  10                  15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
            20                  25                  30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
        35                  40                  45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
    50                  55                  60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65                  70                  75                  80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
                85                  90                  95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
            100                 105                 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
        115                 120                 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
    130                 135                 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145                 150                 155                 160
```

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Gly Gly Ser Gly Gly Gly Ser Gly Gln Ser Gln Leu Ile Lys Pro
195 200 205

Ser Arg Met Arg Lys Tyr Pro Asn Arg Pro Ser Lys Thr Pro Pro His
210 215 220

Glu Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser
225 230 235 240

Arg Ser Asp Asn Leu Val Arg His Ile Arg Ile His Thr Gly Gln Lys
245 250 255

Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Glu Asp Asn
260 265 270

Leu His Thr His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys
275 280 285

Asp Ile Cys Gly Arg Lys Phe Ala Arg Ser Asp Glu Leu Val Arg His
290 295 300

Thr Lys Ile His Leu Arg Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val
305 310 315 320

Glu Ser Cys Asp Arg Arg Phe Ser Gln Ser Gly Asn Leu Thr Glu His
325 330 335

Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met
340 345 350

Arg Asn Phe Ser Thr Ser Gly His Leu Val Arg His Ile Arg Thr His
355 360 365

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
370 375 380

Gln Asn Ser Thr Leu Thr Glu His Thr Lys Ile His Leu Arg Gln Lys
385 390 395 400

Asp Lys

<210> 206
<211> 1707
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 206
atggccgcag atcacctgat gctggctgaa ggctacagac tggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgccctg agaacactgc ccccctacgc cggccctggt 120

cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180

ggagccctgg cctacggcgc cttcggccct ccttctagct tccagccctt ccccgccgtg 240

cctcctccag ccgctggcat cgcccacctg cagcctgtgg ccacccctta ccccggaaga 300

gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360

gctgcccctc cacctcctgc tcatgccctg ggcggcatgg acgccgagct gatcgacgag 420

gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480

ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540

ggctctgtgt cctgcgccga ccacctgatg ctcgccgagg gctaccgcct ggtgcagagg 600

ccgccgtccg ccgccgccgc ccatggccct catgcgctcc ggactctgcc gccgtacgcg 660

ggcccgggcc tggacagtgg gctgaggccg cgggggggctc cgctggggcc gccgccgccc 720

cgccaacccg gggccctggc gtacggggcc ttcgggccgc cgtcctcctt ccagcccttt 780

ccggccgtgc ctccgccggc cgcgggcatc gcgcacctgc agcctgtggc gacgccgtac 840

cccggccgcg ccgccgcgcc ccccaacgct ccgggaggcc ccccgggccc gcagccggcc 900

ccaagcgccg cagccccgcc gccgcccgcg cacgccctgg gcggcatgga cgccgaactc 960

atcgacgagg aggcgctgac gtcgctggag ctggagctgg ggctgcaccg cgtgcgcgag 1020

ctgcccgagc tgttcctggg ccagagcgag ttcgactgct tctcggactt ggggtccgcg 1080

ccgcccgccg gctccgtgag ctgccagtcc cagctcatca aacccagccg catgcgcaag 1140

taccccaacc ggcccagcaa gacgcccccc cacgaacgcc cttacgcttg cccagtggag 1200

tcctgtgatc gccgcttctc ccgcagcgac aacctggtga gacacatccg catccacaca 1260

ggccagaagc ccttccagtg ccgcatctgc atgagaaact tcagccgaga ggataacttg 1320

cacactcaca tccgcaccca cacaggcgaa aagcccttcg cctgcgacat ctgtggaaga 1380

aagtttgccc ggagcgatga acttgtccga ataccaaaga tccacttgcg gcagaaggac 1440

cgcccttacg cttgcccagt ggagtcctgt gatcgccgct tctcccaatc agggaatctg 1500

actgagcaca tccgcatcca cacaggccag aagcccttcc agtgccgcat ctgcatgaga 1560

aacttcagca caagtggaca tctggtacgc cacatccgca cccacacagg cgaaaagccc 1620

ttcgcctgcg acatctgtgg aagaaagttt gcccagaata gtaccctgac cgaacatacc 1680

aagatccact tgcggcagaa ggacaag 1707

<210> 207
<211> 569
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 207

Met	Ala	Ala	Asp	His	Leu	Met	Leu	Ala	Glu	Gly	Tyr	Arg	Leu	Val	Gln
1				5				10				15			

Arg	Pro	Pro	Ser	Ala	Ala	Ala	Ala	His	Gly	Pro	His	Ala	Leu	Arg	Thr
			20				25				30				

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
35 40 45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
50 55 60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65 70 75 80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
85 90 95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
100 105 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
115 120 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
130 135 140

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Ala Asp His Leu Met Leu Ala
180 185 190

Glu Gly Tyr Arg Leu Val Gln Arg Pro Pro Ser Ala Ala Ala Ala His
195 200 205

Gly Pro His Ala Leu Arg Thr Leu Pro Pro Tyr Ala Gly Pro Gly Leu
210 215 220

Asp Ser Gly Leu Arg Pro Arg Gly Ala Pro Leu Gly Pro Pro Pro Pro

225 230 235 240

Arg Gln Pro Gly Ala Leu Ala Tyr Gly Ala Phe Gly Pro Pro Ser Ser
245 250 255

Phe Gln Pro Phe Pro Ala Val Pro Pro Pro Ala Ala Gly Ile Ala His
260 265 270

Leu Gln Pro Val Ala Thr Pro Tyr Pro Gly Arg Ala Ala Ala Pro Pro
275 280 285

Asn Ala Pro Gly Gly Pro Pro Gly Pro Gln Pro Ala Pro Ser Ala Ala
290 295 300

Ala Pro Pro Pro Pro Ala His Ala Leu Gly Gly Met Asp Ala Glu Leu
305 310 315 320

Ile Asp Glu Glu Ala Leu Thr Ser Leu Glu Leu Glu Leu Gly Leu His
325 330 335

Arg Val Arg Glu Leu Pro Glu Leu Phe Leu Gly Gln Ser Glu Phe Asp
340 345 350

Cys Phe Ser Asp Leu Gly Ser Ala Pro Pro Ala Gly Ser Val Ser Cys
355 360 365

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg
370 375 380

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
385 390 395 400

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
405 410 415

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
420 425 430

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr
435 440 445

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
450 455 460

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
465 470 475 480

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
485 490 495

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
500 505 510

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu
515 520 525

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
530 535 540

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
545 550 555 560

Lys Ile His Leu Arg Gln Lys Asp Lys
565

<210> 208
<211> 1755
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 208
atggccgcag atcacctgat gctggctgaa ggctacagac tggtgcagcg gcctccatct 60

gccgctgccg cccacggccc ccacgccctg agaacactgc ccccctacgc cggccctggt 120

cttgatagcg gactcagacc tagaggcgcc cctctgggcc ctccacctcc aagacagcct 180

ggagccctgg cctacggcgc cttcggccct ccttctagct tccagccctt ccccgccgtg 240

cctcctccag ctgctggcat cgcccacctg cagcctgtgg ccacccctta ccccggaaga 300

gccgccgccc ctccaaacgc ccctggcgga cctcctggcc cccagcctgc tccaagcgcc 360

gctgcccctc cacctcctgc tcatgccctg ggcggcatgg acgccgagct gatcgacgag 420

gaagccctga ccagcctgga actggaactg ggcctgcaca gagtgcggga actgcctgag 480

ctgttcctgg gacagagcga gttcgactgc ttcagcgacc tgggcagcgc ccctcctgcc 540

ggctctgtgt cctgcggcgg cagcggcggc ggaagcggcg ccgaccacct gatgctcgcc 600

gagggctacc gcctggtgca gaggccgccg tccgccgccg ccgcccatgg ccctcatgcg 660

ctccggactc tgccgccgta cgcgggcccg ggcctggaca gtgggctgag gccgcggggg 720

gctccgctgg ggccgccgcc gccccgccaa cccgggggccc tggcgtacgg ggccttcggg 780

ccgccgtcct ccttccagcc ctttccggcc gtgcctccgc cggccgcggg catcgcgcac 840

ctgcagcctg tggcgacgcc gtaccccggc cgcgcggccg cgccccccaa cgctccggga 900

ggcccccccgg gcccgcagcc ggccccaagc gccgcagccc cgccgccgcc cgcgcacgcc 960

ctgggcggca tggacgccga actcatcgac gaggaggcgc tgacgtcgct ggagctggag 1020

ctggggctgc accgcgtgcg cgagctgccc gagctgttcc tgggccagag cgagttcgac 1080

tgcttctcgg acttggggtc cgcgccgccc gccggctccg tgagctgcgg tggttctggt 1140

ggtggttctg gtcagtccca gctcatcaaa cccagccgca tgcgcaagta ccccaaccgg 1200

cccagcaaga cgccccccca cgaacgccct tacgcttgcc cagtggagtc ctgtgatcgc 1260

cgcttctccc gcagcgacaa cctggtgaga cacatccgca tccacacagg ccagaagccc 1320

ttccagtgcc gcatctgcat gagaaacttc agccgagagg ataacttgca cactcacatc 1380

cgcacccaca caggcgaaaa gcccttcgcc tgcgacatct gtggaagaaa gtttgcccgg 1440

agcgatgaac ttgtccgaca taccaagatc cacttgcggc agaaggaccg cccttacgct 1500

tgcccagtgg agtcctgtga tcgccgcttc tcccaatcag ggaatctgac tgagcacatc 1560

cgcatccaca caggccagaa gcccttccag tgccgcatct gcatgagaaa cttcagcaca 1620

agtggacatc tggtacgcca catccgcacc cacacaggcg aaaagccctt cgcctgcgac 1680

atctgtggaa gaaagtttgc ccagaatagt accctgaccg aacataccaa gatccacttg 1740

cggcagaagg acaag 1755

<210> 209
<211> 585
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 209

```
Met Ala Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln
1               5                  10                  15

Arg Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr
            20                  25                  30

Leu Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg
        35                  40                  45

Gly Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala
    50                  55                  60

Tyr Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val
65                  70                  75                  80

Pro Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro
                85                  90                  95

Tyr Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro
            100                 105                 110

Gly Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His
        115                 120                 125

Ala Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr
    130                 135                 140
```

Ser Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu
145 150 155 160

Leu Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser
165 170 175

Ala Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser
180 185 190

Gly Ala Asp His Leu Met Leu Ala Glu Gly Tyr Arg Leu Val Gln Arg
195 200 205

Pro Pro Ser Ala Ala Ala Ala His Gly Pro His Ala Leu Arg Thr Leu
210 215 220

Pro Pro Tyr Ala Gly Pro Gly Leu Asp Ser Gly Leu Arg Pro Arg Gly
225 230 235 240

Ala Pro Leu Gly Pro Pro Pro Pro Arg Gln Pro Gly Ala Leu Ala Tyr
245 250 255

Gly Ala Phe Gly Pro Pro Ser Ser Phe Gln Pro Phe Pro Ala Val Pro
260 265 270

Pro Pro Ala Ala Gly Ile Ala His Leu Gln Pro Val Ala Thr Pro Tyr
275 280 285

Pro Gly Arg Ala Ala Ala Pro Pro Asn Ala Pro Gly Gly Pro Pro Gly
290 295 300

Pro Gln Pro Ala Pro Ser Ala Ala Ala Pro Pro Pro Pro Ala His Ala
305 310 315 320

Leu Gly Gly Met Asp Ala Glu Leu Ile Asp Glu Glu Ala Leu Thr Ser
325 330 335

Leu Glu Leu Glu Leu Gly Leu His Arg Val Arg Glu Leu Pro Glu Leu
340 345 350

Phe Leu Gly Gln Ser Glu Phe Asp Cys Phe Ser Asp Leu Gly Ser Ala
355 360 365

Pro Pro Ala Gly Ser Val Ser Cys Gly Gly Ser Gly Gly Gly Ser Gly
370 375 380

Gln Ser Gln Leu Ile Lys Pro Ser Arg Met Arg Lys Tyr Pro Asn Arg
385 390 395 400

Pro Ser Lys Thr Pro Pro His Glu Arg Pro Tyr Ala Cys Pro Val Glu
405 410 415

Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val Arg His Ile
420 425 430

Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg
435 440 445

Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg Thr His Thr
450 455 460

Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg
465 470 475 480

Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg Gln Lys Asp
485 490 495

Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Gln
500 505 510

Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly Gln Lys Pro
515 520 525

Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser Gly His Leu
530 535 540

Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp
545 550 555 560

Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr Glu His Thr
565 570 575

Lys Ile His Leu Arg Gln Lys Asp Lys
580 585

<210> 210
<211> 1113
<212> DNA
<213> 智人（Homo sapiens）

<400> 210
atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggaggaa gattcgaaat 480

aaaagatctg ctcaagagag ccgcaggaaa aagaaggtgt acgttggggg tttagagagc 540

cgggtcttga aatacacagc ccagaatatg gagcttcaga acaaagtaca gcttctggag 600

gaacagaatt tgtcccttct agatcaactg aggaaactcc aggccatggt gattgagatc 660

tcaaacaaaa ccagcagcag cagcacctgc atcttggtcc tgctagtctc cttctgcctc 720

ctccttgtac ctgctatgta ctcctctgac acaaggggga gcctgccagc tgagcatgga 780

gtgttgtccc gccagcttcg tgccctcccc agtgaggacc cttaccagct ggagctgcct 840

gccctgcagt cagaagtgcc gaaagacagc acacaccagt ggttggacgg ctcagactgt 900

gtactccagg cccctggcaa cacttcctgc ctgctgcatt acatgcctca ggctcccagt 960

gcagagcctc ccctggagtg gcccttccct gacctcttct cagagcctct ctgccgaggt 1020

cccatcctcc ccctgcaggc aaatctcaca aggaagggag gatggcttcc tactggtagc 1080

ccctctgtca ttttgcagga cagatactca ggc 1113

<210> 211
<211> 371
<212> PRT
<213> 智人（Homo sapiens）

<400> 211

```
Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1               5                  10                  15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
            20                  25                  30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
        35                  40                  45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
    50                  55                  60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65                  70                  75                  80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
                85                  90                  95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
            100                 105                 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
        115                 120                 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
    130                 135                 140
```

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Lys Ile Arg Asn
145 150 155 160

Lys Arg Ser Ala Gln Glu Ser Arg Arg Lys Lys Lys Val Tyr Val Gly
165 170 175

Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu
180 185 190

Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp
195 200 205

Gln Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr
210 215 220

Ser Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu
225 230 235 240

Leu Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro
245 250 255

Ala Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu
260 265 270

Asp Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys
275 280 285

Asp Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala
290 295 300

Pro Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser
305 310 315 320

Ala Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro
325 330 335

Leu Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys

340 345 350

Gly Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg
355 360 365

Tyr Ser Gly
370

<210> 212
<211> 1590
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 212
atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggctcga accaggtgaa 480

aaaccttaca aatgtcctga atgtgggaaa tcattcagtc gcagcgacaa cctggtgaga 540

catcaacgca cccatacagg agaaaaacct tataaatgtc cagaatgtgg aaagtccttc 600

tcacgagagg ataacttgca cactcatcaa cgaacacata ctggtgaaaa accatacaag 660

tgtcccgaat gtggtaaaag ttttagccgg agcgatgaac ttgtccgaca ccaacgaacc 720

catacaggcg agaagcctta caaatgtccc gagtgtggca agagcttctc acaatcaggg 780

aatctgactg agcatcaacg aactcatacc ggggaaaaac cttacaagtg tccagagtgt 840

gggaagagct tttccacaag tggacatctg gtacgccacc agaggacaca tacaggggag 900

aagccctaca aatgccccga atgcggtaaa agtttctctc agaatagtac cctgaccgaa 960

caccagcgaa cacacactgg gaaaaaaacg agtgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

aacaaaacca gcagcagcag cacctgcatc ttggtcctgc tagtctcctt ctgcctcctc 1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260

ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320

ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380

ctccaggccc ctggcaacac ttcctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440

gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500

atcctccccc tgcaggcaaa tctcacaagg aagggaggat ggcttcctac tggtagcccc 1560

tctgtcattt tgcaggacag atactcaggc 1590

<210> 213
<211> 530
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 213
Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
50 55 60

```
Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65                  70                  75                  80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
                85                  90                  95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
            100                 105                 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
        115                 120                 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
    130                 135                 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Leu Glu Pro Gly Glu
145                 150                 155                 160

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp
                165                 170                 175

Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
            180                 185                 190

Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr
        195                 200                 205

His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
    210                 215                 220

Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr
225                 230                 235                 240

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
                245                 250                 255

Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu
            260                 265                 270
```

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly
275 280 285

His Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
290 295 300

Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu
305 310 315 320

His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser
370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu
385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala
405 410 415

Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp
420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp
435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro
450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala
465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu
485 490 495

Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly
500 505 510

Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr
515 520 525

Ser Gly
530

<210> 214
<211> 1590
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 214
atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcttga gcccggagag 480

aagccgtaca agtgccctga gtgcggcaag tcttttagca gaagagacga acttaatgtc 540

caccagcgaa cgcatactgg tgaaaagccc tataaatgtc ctgaatgtgg gaaatcattc 600

tccagccgca gaacctgtag ggctcaccag cgaacacaca ccggcgaaaa accatacaaa 660

tgtccagaat gcgggaaatc cttttctcag tcatccaact tggtgagaca tcaacgcacg 720

cacactggag aaaagcctta caaatgcccg gaatgtggaa agtctttttc ccaattggcc 780

catttgcgag cccatcagag gactcacacg ggcgagaaac cttacaaatg cccggaatgc 840

gggaaatctt tttcaacgag tggcaacctc gtaagacacc aaagaacgca tacaggcgaa 900

aagccatata agtgtcctga gtgtggtaaa tcattctcac acaggaccac cctgacaaat 960

caccagcgca cgcacaccgg caagaagaca agcgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

aacaaaacca gcagcagcag cacctgcatc ttggtcctgc tagtctcctt ctgcctcctc 1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260

ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320

ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380

ctccaggccc ctggcaacac ttcctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440

gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500

atcctccccc tgcaggcaaa tctcacaagg aagggaggat ggcttcctac tggtagcccc 1560

tctgtcattt tgcaggacag atactcaggc 1590

<210> 215
<211> 530
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 215
Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
100 105 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
115 120 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
130 135 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Leu Glu Pro Gly Glu
145 150 155 160

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Arg Asp
165 170 175

Glu Leu Asn Val His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
180 185 190

Cys Pro Glu Cys Gly Lys Ser Phe Ser Ser Arg Arg Thr Cys Arg Ala
195 200 205

His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
210 215 220

Gly Lys Ser Phe Ser Gln Ser Ser Asn Leu Val Arg His Gln Arg Thr

225 230 235 240

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
245 250 255

Ser Gln Leu Ala His Leu Arg Ala His Gln Arg Thr His Thr Gly Glu
260 265 270

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly
275 280 285

Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
290 295 300

Cys Pro Glu Cys Gly Lys Ser Phe Ser His Arg Thr Thr Leu Thr Asn
305 310 315 320

His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser
370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu
385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala
405 410 415

Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp
420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp
435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro
450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala
465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu
485 490 495

Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly
500 505 510

Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr
515 520 525

Ser Gly
530

<210> 216
<211> 1590
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 216
atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc 480

ccagtggagt cctgtgatcg ccgcttctcc cgcagcgaca acctggtgag acacatccgc 540

atccacacag gccagaagcc cttccagtgc cgcatctgca tgagaaactt cagccgagag 600

gataacttgc acactcacat ccgcacccac acaggcgaaa agcccttcgc ctgcgacatc 660

tgtggaagaa agtttgcccg gagcgatgaa cttgtccgac ataccaagat ccacttgcgg 720

cagaaggacc gcccttacgc ttgcccagtg gagtcctgtg atcgccgctt ctcccaatca 780

gggaatctga ctgagcacat ccgcatccac acaggccaga agcccttcca gtgccgcatc 840

tgcatgagaa acttcagcac aagtggacat ctggtacgcc acatccgcac ccacacaggc 900

gaaaagccct tcgcctgcga catctgtgga agaaagtttg cccagaatag taccctgacc 960

gaacatacca agatccactt gcggcagaag gacgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

aacaaaacca gcagcagcag cacctgcatc ttggtcctgc tagtctcctt ctgcctcctc 1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260

ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320

ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380

ctccaggccc ctggcaacac ttcctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440

gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500

atcctccccc tgcaggcaaa tctcacaagg aagggaggat ggcttcctac tggtagcccc 1560

tctgtcattt tgcaggacag atactcaggc 1590

<210> 217
<211> 530
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 217

```
Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1               5                  10                  15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
            20                  25                  30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
        35                  40                  45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
    50                  55                  60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65                  70                  75                  80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
                85                  90                  95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
            100                 105                 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
        115                 120                 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
    130                 135                 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Pro Tyr Ala Cys
145                 150                 155                 160

Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val
                165                 170                 175

Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile
            180                 185                 190
```

Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg
195 200 205

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
210 215 220

Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg
225 230 235 240

Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg
245 250 255

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
260 265 270

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
275 280 285

Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
290 295 300

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr
305 310 315 320

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser
370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu
385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala
405 410 415

Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp
420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp
435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro
450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala
465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu
485 490 495

Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly
500 505 510

Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr
515 520 525

Ser Gly
530

<210> 218
<211> 1590
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 218
atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc 480

ccagtggagt cctgtgatcg ccgcttctcc cgctcagaca acctcgttcg acacatccgc 540

atccacacag gccagaagcc cttccagtgc cgcatctgca tgagaaactt cagccaccgg 600

actacactca cgaaccacat ccgcacccac acaggcgaaa agcccttcgc ctgcgacatc 660

tgtggaagaa agtttgccag agaagacaat ctccatactc ataccaagat ccacttgcgg 720

cagaaggacc gcccttacgc ttgcccagtg gagtcctgtg atcgccgctt ctccaccagc 780

cattctctca ctgaacacat ccgcatccac acaggccaga agcccttcca gtgccgcatc 840

tgcatgagaa acttcagcca gtctagctca ctggtgaggc acatccgcac ccacacaggc 900

gaaaagccct tcgcctgcga catctgtgga agaaagtttg ccagggagga taacctgcat 960

acgcatacca agatccactt gcggcagaag gacgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

aacaaaacca gcagcagcag cacctgcatc ttggtcctgc tagtctcctt ctgcctcctc 1200

cttgtacctg ctatgtactc ctctgacaca agggggagcc tgccagctga gcatggagtg 1260

ttgtcccgcc agcttcgtgc cctccccagt gaggaccctt accagctgga gctgcctgcc 1320

ctgcagtcag aagtgccgaa agacagcaca caccagtggt tggacggctc agactgtgta 1380

ctccaggccc ctggcaacac ttcctgcctg ctgcattaca tgcctcaggc tcccagtgca 1440

gagcctcccc tggagtggcc cttccctgac ctcttctcag agcctctctg ccgaggtccc 1500

atcctccccc tgcaggcaaa tctcacaagg aagggaggat ggcttcctac tggtagcccc 1560

tctgtcattt tgcaggacag atactcaggc 1590

<210> 219
<211> 530
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 219

```
Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1               5                  10                  15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
            20                  25                  30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
        35                  40                  45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
    50                  55                  60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65                  70                  75                  80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
                85                  90                  95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
            100                 105                 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
        115                 120                 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
    130                 135                 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Pro Tyr Ala Cys
145                 150                 155                 160
```

Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val
165 170 175

Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile
180 185 190

Cys Met Arg Asn Phe Ser His Arg Thr Thr Leu Thr Asn His Ile Arg
195 200 205

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
210 215 220

Phe Ala Arg Glu Asp Asn Leu His Thr His Thr Lys Ile His Leu Arg
225 230 235 240

Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg
245 250 255

Phe Ser Thr Ser His Ser Leu Thr Glu His Ile Arg Ile His Thr Gly
260 265 270

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Gln Ser
275 280 285

Ser Ser Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
290 295 300

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Glu Asp Asn Leu His
305 310 315 320

Thr His Thr Lys Ile His Leu Arg Gln Lys Asp Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln

355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr Ser
370 375 380

Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu Leu
385 390 395 400

Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala
405 410 415

Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp
420 425 430

Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys Asp
435 440 445

Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala Pro
450 455 460

Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser Ala
465 470 475 480

Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu
485 490 495

Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly
500 505 510

Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr
515 520 525

Ser Gly
530

<210> 220
<211> 1140
<212> DNA

<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 220
atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggctcga accaggtgaa 480

aaaccttaca aatgtcctga atgtgggaaa tcattcagtc gcagcgacaa cctggtgaga 540

catcaacgca cccatacagg agaaaaacct tataaatgtc cagaatgtgg aaagtccttc 600

tcacgagagg ataacttgca cactcatcaa cgaacacata ctggtgaaaa accatacaag 660

tgtcccgaat gtggtaaaag ttttagccgg agcgatgaac ttgtccgaca ccaacgaacc 720

catacaggcg agaagcctta caaatgtccc gagtgtggca agagcttctc acaatcaggg 780

aatctgactg agcatcaacg aactcatacc ggggaaaaac cttacaagtg tccagagtgt 840

gggaagagct tttccacaag tggacatctg gtacgccacc agaggacaca tacaggggag 900

aagccctaca aatgccccga atgcggtaaa agtttctctc agaatagtac cctgaccgaa 960

caccagcgaa cacacactgg gaaaaaaacg agtgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatatca 1140

<210> 221
<211> 380
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 221

```
Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1               5                  10                  15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
            20                  25                  30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
        35                  40                  45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
    50                  55                  60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65                  70                  75                  80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
                85                  90                  95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
            100                 105                 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
        115                 120                 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
    130                 135                 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Leu Glu Pro Gly Glu
145                 150                 155                 160

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp
                165                 170                 175

Asn Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
            180                 185                 190
```

Cys Pro Glu Cys Gly Lys Ser Phe Ser Arg Glu Asp Asn Leu His Thr
195 200 205

His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys
210 215 220

Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Val Arg His Gln Arg Thr
225 230 235 240

His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe
245 250 255

Ser Gln Ser Gly Asn Leu Thr Glu His Gln Arg Thr His Thr Gly Glu
260 265 270

Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys Ser Phe Ser Thr Ser Gly
275 280 285

His Leu Val Arg His Gln Arg Thr His Thr Gly Glu Lys Pro Tyr Lys
290 295 300

Cys Pro Glu Cys Gly Lys Ser Phe Ser Gln Asn Ser Thr Leu Thr Glu
305 310 315 320

His Gln Arg Thr His Thr Gly Lys Lys Thr Ser Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser
370 375 380

<210> 222
<211> 1140
<212> DNA
<213> 人工序列

<220>
<223> 人工序列的描述：合成多核苷酸

<400> 222
atggagctgg aattggatgc tggtgaccaa gacctgctgg ccttcctgct agaggaaagt 60

ggagatttgg ggacggcacc cgatgaggcc gtgagggccc cactggactg ggcgctgccg 120

ctttctgagg tgccgagcga ctgggaagta gatgatttgc tgtgctccct gctgagtccc 180

ccagcgtcgt tgaacattct cagctcctcc aacccctgcc ttgtccacca tgaccacacc 240

tactccctcc cacgggaaac tgtctctatg gatctagaga gtgagagctg tagaaaagag 300

gggacccaga tgactccaca gcatatggag gagctggcag agcaggagat tgctaggcta 360

gtactgacag atgaggagaa gagtctattg gagaaggagg ggcttattct gcctgagaca 420

cttcctctca ctaagacaga ggaacaaatt ctgaaacgtg tgcggcgccc ttacgcttgc 480

ccagtggagt cctgtgatcg ccgcttctcc cgcagcgaca acctggtgag acacatccgc 540

atccacacag gccagaagcc cttccagtgc cgcatctgca tgagaaactt cagccgagag 600

gataacttgc acactcacat ccgcacccac acaggcgaaa agcccttcgc ctgcgacatc 660

tgtggaagaa agtttgcccg gagcgatgaa cttgtccgac ataccaagat ccacttgcgg 720

cagaaggacc gcccttacgc ttgcccagtg gagtcctgtg atcgccgctt ctcccaatca 780

gggaatctga ctgagcacat ccgcatccac acaggccaga agcccttcca gtgccgcatc 840

tgcatgagaa acttcagcac aagtggacat ctggtacgcc acatccgcac ccacacaggc 900

gaaaagccct tcgcctgcga catctgtgga agaaagtttg cccagaatag taccctgacc 960

gaacatacca agatccactt gcggcagaag gacgtgtacg ttgggggttt agagagccgg 1020

gtcttgaaat acacagccca gaatatggag cttcagaaca aagtacagct tctggaggaa 1080

cagaatttgt cccttctaga tcaactgagg aaactccagg ccatggtgat tgagatctca 1140

<210> 223
<211> 380

<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 223

```
Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1               5                  10                  15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
            20                  25                  30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
        35                  40                  45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
    50                  55                  60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65                  70                  75                  80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
                85                  90                  95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
            100                 105                 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
        115                 120                 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
    130                 135                 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Pro Tyr Ala Cys
145                 150                 155                 160

Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp Asn Leu Val
                165                 170                 175
```

Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln Cys Arg Ile
180 185 190

Cys Met Arg Asn Phe Ser Arg Glu Asp Asn Leu His Thr His Ile Arg
195 200 205

Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys
210 215 220

Phe Ala Arg Ser Asp Glu Leu Val Arg His Thr Lys Ile His Leu Arg
225 230 235 240

Gln Lys Asp Arg Pro Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg
245 250 255

Phe Ser Gln Ser Gly Asn Leu Thr Glu His Ile Arg Ile His Thr Gly
260 265 270

Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser Thr Ser
275 280 285

Gly His Leu Val Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe
290 295 300

Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Gln Asn Ser Thr Leu Thr
305 310 315 320

Glu His Thr Lys Ile His Leu Arg Gln Lys Asp Val Tyr Val Gly Gly
325 330 335

Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu Gln
340 345 350

Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp Gln
355 360 365

Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser

370 375 380

<210> 224
<211> 155
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 224
Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
1 5 10 15

Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
20 25 30

Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
35 40 45

Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
50 55 60

Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
65 70 75 80

Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
85 90 95

Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
100 105 110

Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
115 120 125

Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
130 135 140

Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg
145 150 155

<210> 225
<211> 199
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 225

```
Val Tyr Val Gly Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln
1               5                  10                  15

Asn Met Glu Leu Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu
            20                  25                  30

Ser Leu Leu Asp Gln Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile
        35                  40                  45

Ser Asn Lys Thr Ser Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val
    50                  55                  60

Ser Phe Cys Leu Leu Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg
65                  70                  75                  80

Gly Ser Leu Pro Ala Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala
                85                  90                  95

Leu Pro Ser Glu Asp Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser
            100                 105                 110

Glu Val Pro Lys Asp Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys
        115                 120                 125

Val Leu Gln Ala Pro Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro
    130                 135                 140

Gln Ala Pro Ser Ala Glu Pro Pro Leu Glu Trp Pro Phe Pro Asp Leu
145                 150                 155                 160
```

Phe Ser Glu Pro Leu Cys Arg Gly Pro Ile Leu Pro Leu Gln Ala Asn
165 170 175

Leu Thr Arg Lys Gly Gly Trp Leu Pro Thr Gly Ser Pro Ser Val Ile
180 185 190

Leu Gln Asp Arg Tyr Ser Gly
195

<210> 226
<211> 49
<212> PRT
<213> 人工序列

<220>
<223> 人工序列的描述：合成多肽

<400> 226
Val Tyr Val Gly Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln
1 5 10 15

Asn Met Glu Leu Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu
20 25 30

Ser Leu Leu Asp Gln Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile
35 40 45

Ser

【發明申請專利範圍】

【請求項1】

一種多核苷酸，其包含可操作地與編碼工程化的轉錄因子的核酸序列連接的調節元件，從而形成表達盒，其中該調節元件包含SEQ ID NO: 2且係在編碼該轉錄因子的核酸序列的上游，且其中該工程化的轉錄因子包含SEQ ID NO: 127之胺基酸序列。

【請求項2】

如請求項1之多核苷酸，其中該表達盒進一步包含抑制該工程化的轉錄因子在興奮性神經元中之表現的元件。

【請求項3】

如請求項2之多核苷酸，其中該抑制該工程化的轉錄因子在興奮性神經元中之表現的元件為促進mRNA降解之元件。

【請求項4】

如請求項3之多核苷酸，其中該抑制該工程化的轉錄因子在興奮性神經元中之表現的元件包含微小RNA結合位點。

【請求項5】

如請求項4之多核苷酸，其中該微小RNA結合位點包含SEQ ID NO: 9、11或13之序列。

【請求項6】

如請求項4之多核苷酸，其中該微小RNA結合位點包含SEQ ID NO: 9及SEQ ID NO: 11。

【請求項7】

如請求項4之多核苷酸，其中該抑制該工程化的轉錄因子之表現的元

件包含至少6個微小RNA結合位點。

【請求項8】

如請求項4之多核苷酸，其中該微小RNA結合位點包含SEQ ID NO: 7、14或15。

【請求項9】

如請求項4之多核苷酸，其中該微小RNA結合位點包含SEQ ID NO: 7。

【請求項10】

如請求項2之多核苷酸，其中該多核苷酸包含SEQ ID NO: 71之核酸序列。

【請求項11】

一種表達載體，其包含如請求項1至10中任一項之多核苷酸。

【請求項12】

如請求項11之表達載體，其中該表達載體為病毒載體。

【請求項13】

如請求項12之表達載體，其中該病毒載體為腺伴隨病毒(AAV)病毒。

【請求項14】

如請求項13之表達載體，其中該AAV具有選自以下組成之群之血清型：AAV1、AAV2、AAV3、AAV4、AAV5、AAV6、AAV7、AAV8、AAV9、AAV10、AAV11、AAV12、rh10、禽AAV、牛AAV、犬AAV、馬AAV、靈長類動物AAV、羊AAV、scAAV及其任一者之雜合。

【請求項15】

如請求項13之表達載體，其中該表達載體進一步包含5' AAV反向末端重複(ITR)序列及3' AAV ITR序列。

【請求項16】

如請求項15之表達載體，其中該5' AAV ITR序列及該3' AAV ITR序列各獨立地為衍生自AAV1、AAV2、AAV5、AAV8或AAV9之ITR序列。

【請求項17】

如請求項11至16中任一項之表達載體，其中該表達載體包含SEQ ID NO: 71。

【請求項18】

如請求項17之表達載體，其中該表達載體包含具有5' AAV反向末端重複(ITR)序列及3' AAV ITR序列之AAV，其中該AAV為AAV9，且其中該5' AAV ITR序列及該3' AAV ITR序列係衍生自AAV2。

【請求項19】

一種藥物組合物，其包含如請求項1至10中任一項之多核苷酸或如請求項11至18中任一項之表達載體，以及藥學上可接受的載體。

【請求項20】

一種細胞，其包含如請求項1至10中任一項之多核苷酸或如請求項11至18中任一項之表達載體。

【請求項21】

如請求項20之細胞，其中該細胞為293細胞、A549細胞或HeLa細胞。

【請求項22】

一種製備腺伴隨病毒(AAV)病毒體之方法，其包含：

a) 在產生重組AAV病毒體之條件下培養如請求項20或21之細胞；

b) 收穫細胞培養物；及

c) 純化該細胞所產生之重組AAV病毒體。

【請求項23】

一種增加細胞中之SCN1A表現量之活體外方法，其包含使該細胞與如請求項1至10中任一項之多核苷酸或如請求項11至18中任一項之表達載體接觸。

【請求項24】

一種如請求項1至10中任一項之多核苷酸或如請求項11至18中任一項之表達載體之用途，其係用於製備增加細胞中之SCN1A表現量之醫藥品。

【請求項25】

如請求項24之用途，其中該細胞係在人類個體中。

【請求項26】

如請求項24或25之用途，其中增加細胞中之SCN1A的表現治療與SCN1A相關之疾病或病症。

【請求項27】

如請求項26之用途，其中該疾病或病症是中樞神經系統病症。

【請求項28】

如請求項27之用途，其中該病症是與SCN1A單倍性不足相關之癲癇。

【請求項29】

一種如請求項1至10中任一項之多核苷酸或如請求項11至18中任一項之表達載體之用途，其係用於製備治療Dravet綜合症之醫藥品。

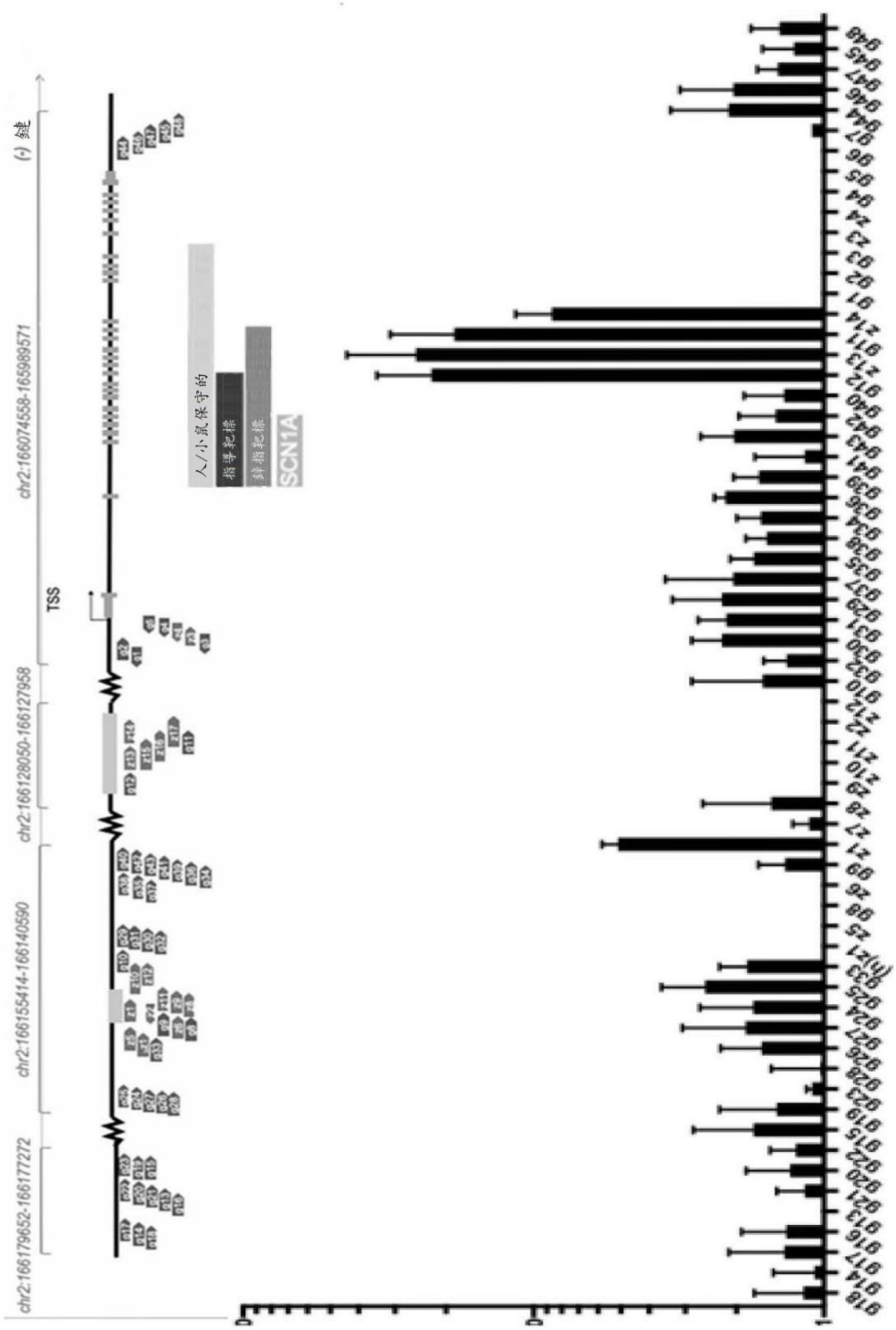
【請求項30】

一種如請求項1至10中任一項之多核苷酸或如請求項11至18中任一項之表達載體之用途，其係用於製備治療阿爾茨海默病(Alzheimer's disease)之醫藥品。

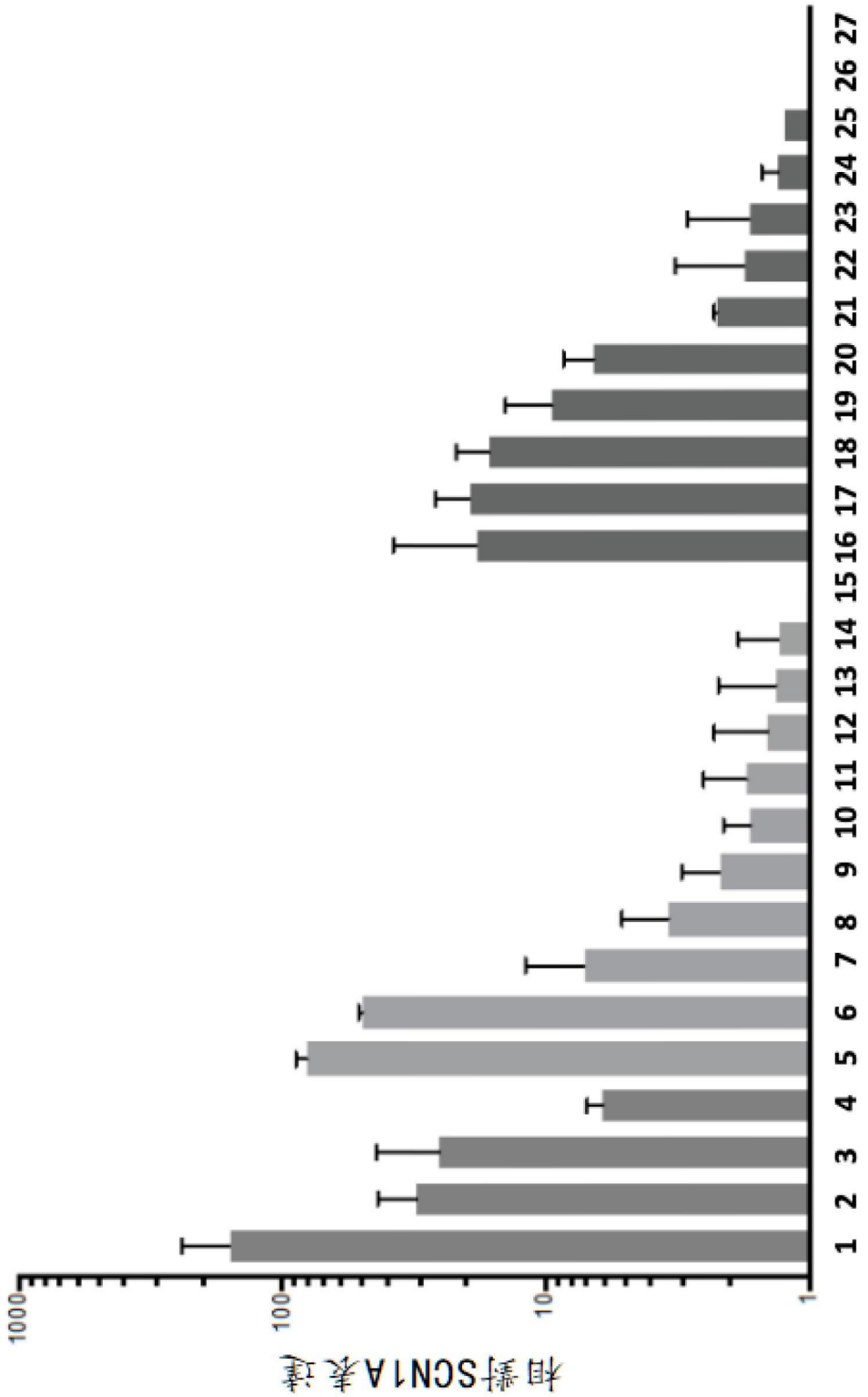
【請求項31】

一種如請求項1至10中任一項之多核苷酸或如請求項11至18中任一項之表達載體之用途，其係用於製備降低癲癇發作嚴重程度、癲癇發作頻率或癲癇發作嚴重程度及頻率之醫藥品。

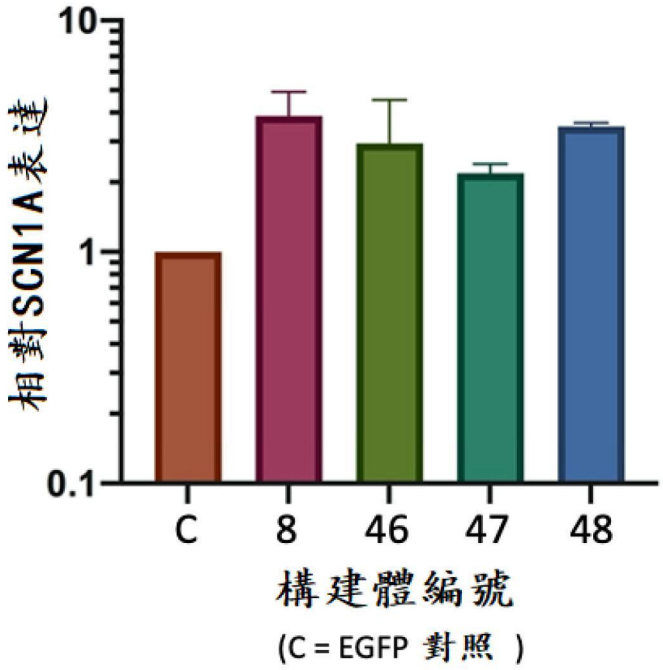
【發明圖式】


(-) 鏈
chr2:166074558-165989571
chr2:166128050-166127958
chr2:166155414-166140590
chr2:166179652-166177272
TSS
人/小鼠保守的
指導靶標
非指靶標
SCN1A
不同靶位點

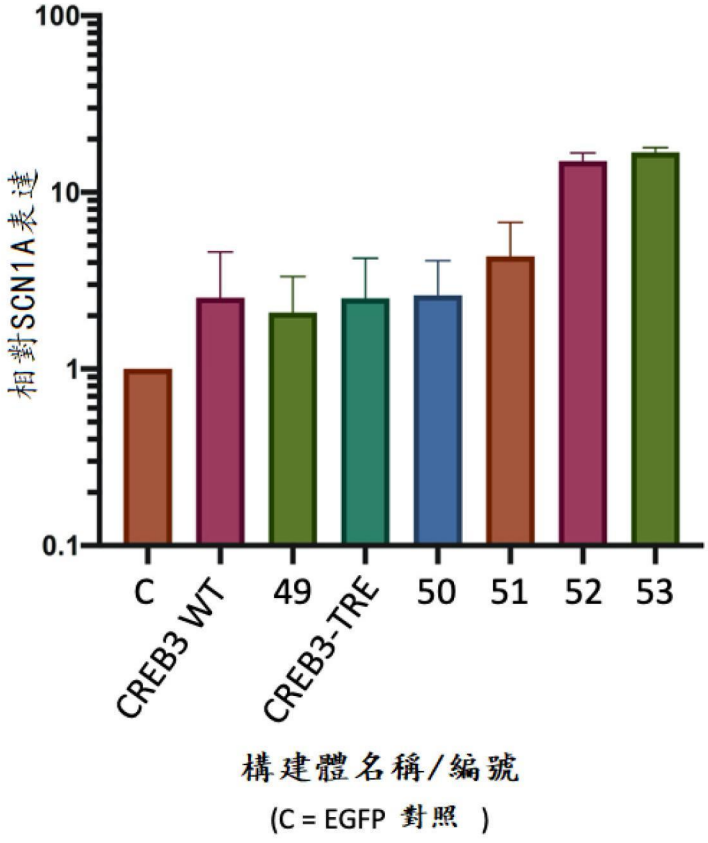
【圖1】


相對SCN1A表達
1000
100
10
1
構建體編號
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27

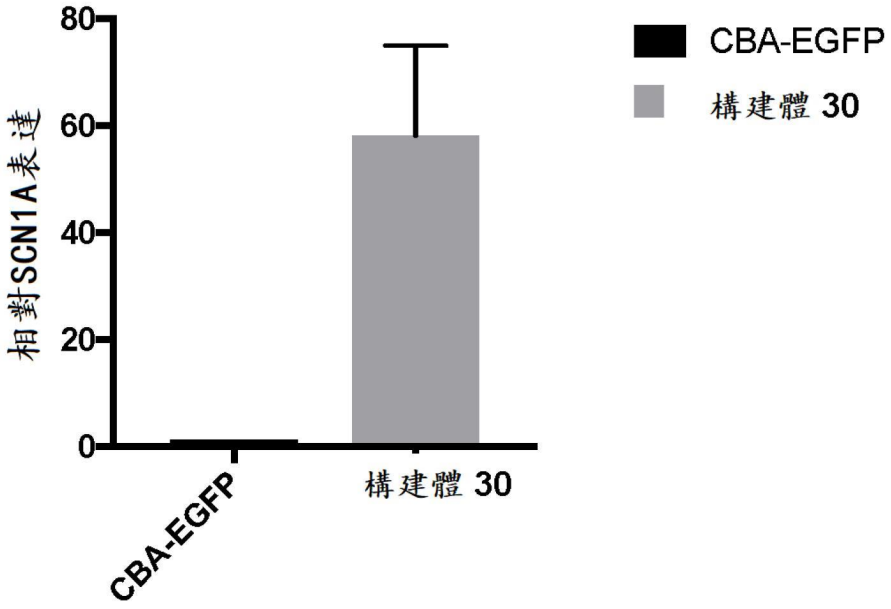
【圖2A】


10
1
0.1
相對SCN1A表達
C
8
46
47
48
構建體編號
(C = EGFP 對照)

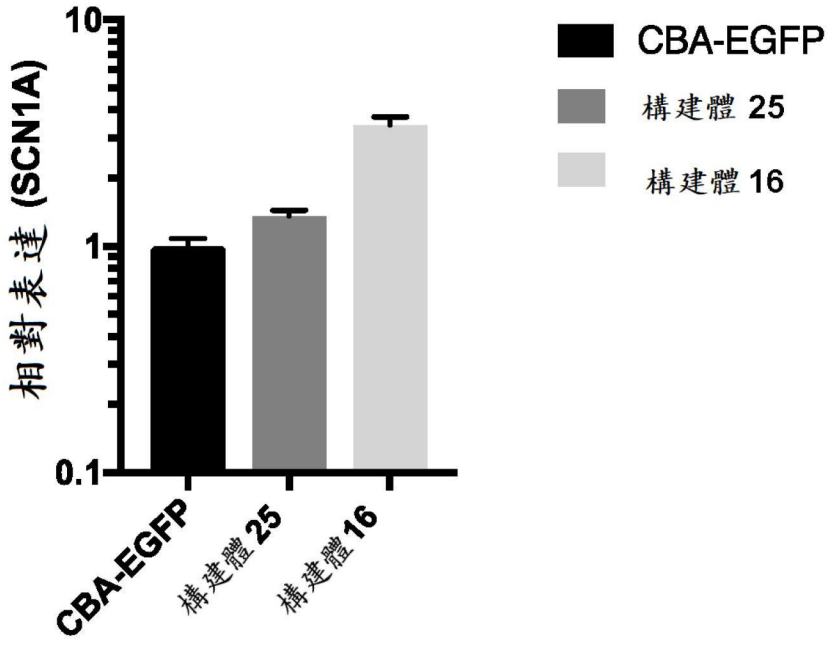
【圖2B】


100
10
1
0.1
相對SCN1A表達
C
CREB3 WT
49
CREB3-TRE
50
51
52
53
構建體名稱/編號
(C = EGFP 對照)

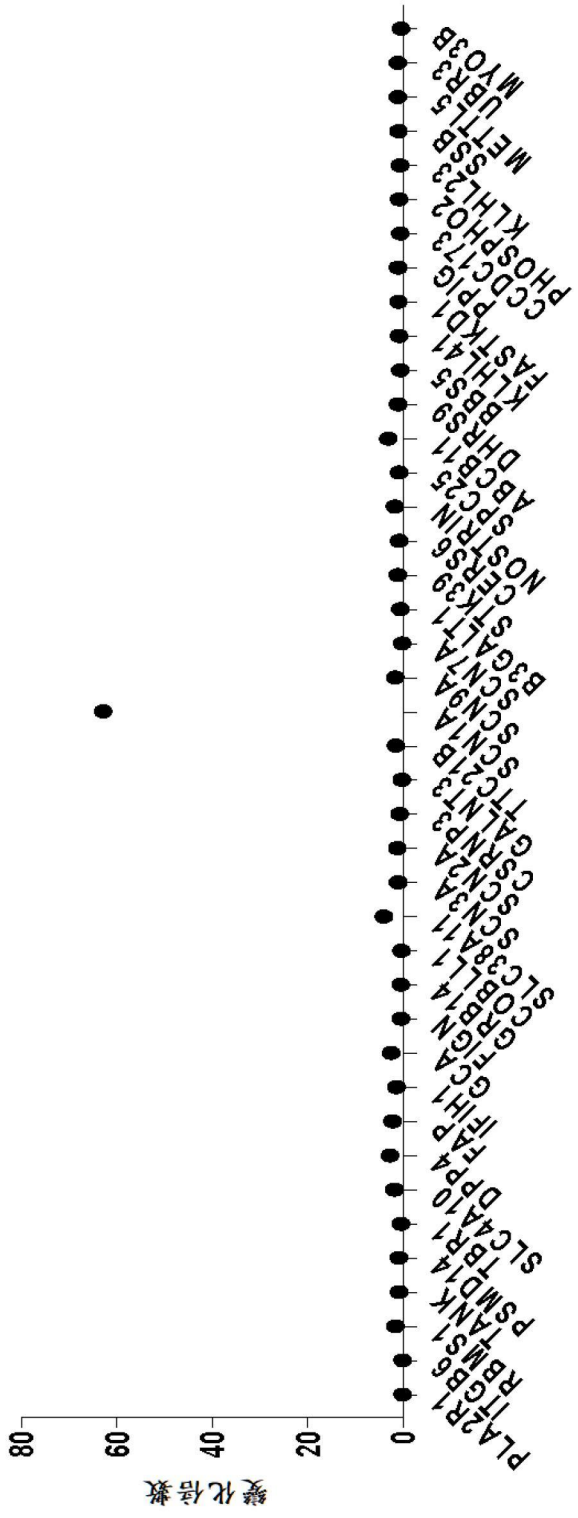
【圖2C】


80
60
40
20
0
相對SCN1A表達
CBA-EGFP
構建體 30
CBA-EGFP
構建體 30

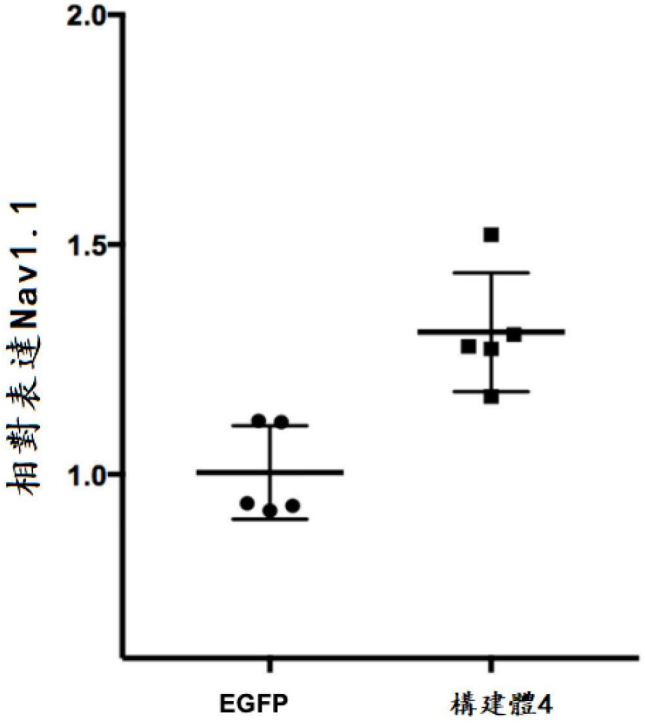
【圖3A】


10
1
0.1
相對表達 (SCN1A)
CBA-EGFP
構建體 25
構建體 16
CBA-EGFP
構建體25
構建體16

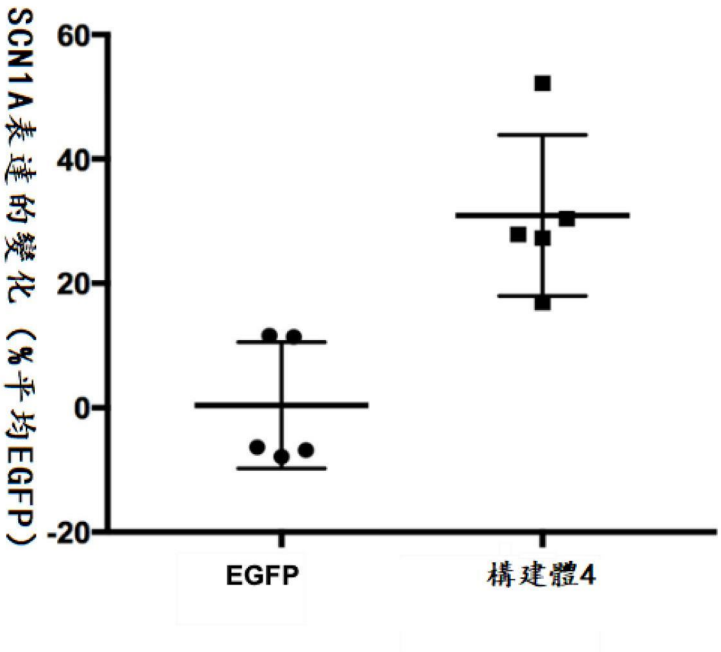
【圖3B】



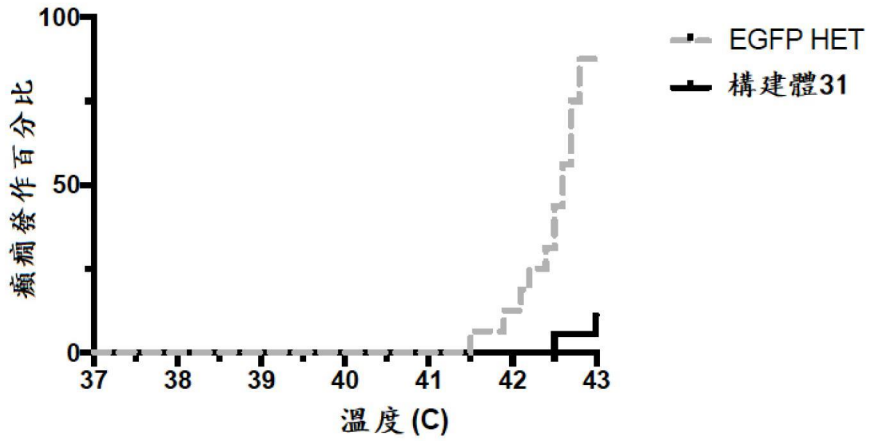
【圖4】


2.0
1.5
1.0
相對表達Nav1.1
EGFP
構建體4

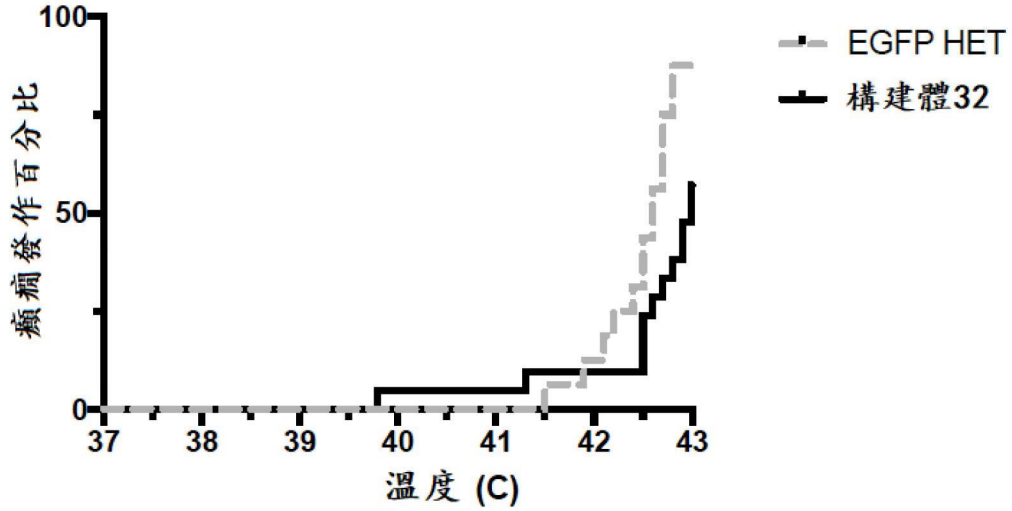
【圖5A】


60
40
20
0
-20
SCN1A表達的變化（%平均EGFP）
EGFP
構建體4

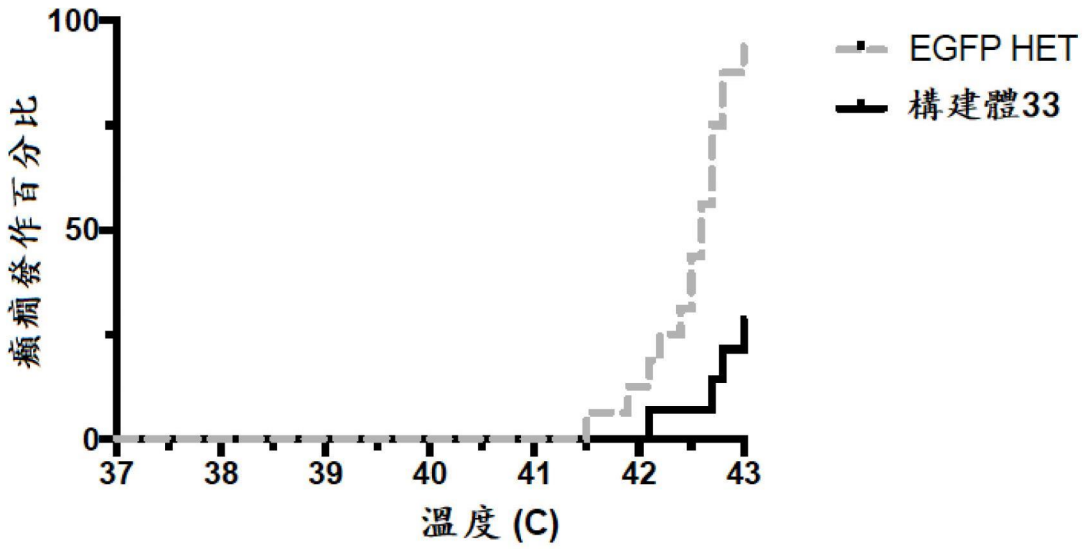
【圖5B】


100
50
0
癲癇發作百分比
37
38
39
40
41
42
43
溫度 (C)
EGFP HET
構建體31

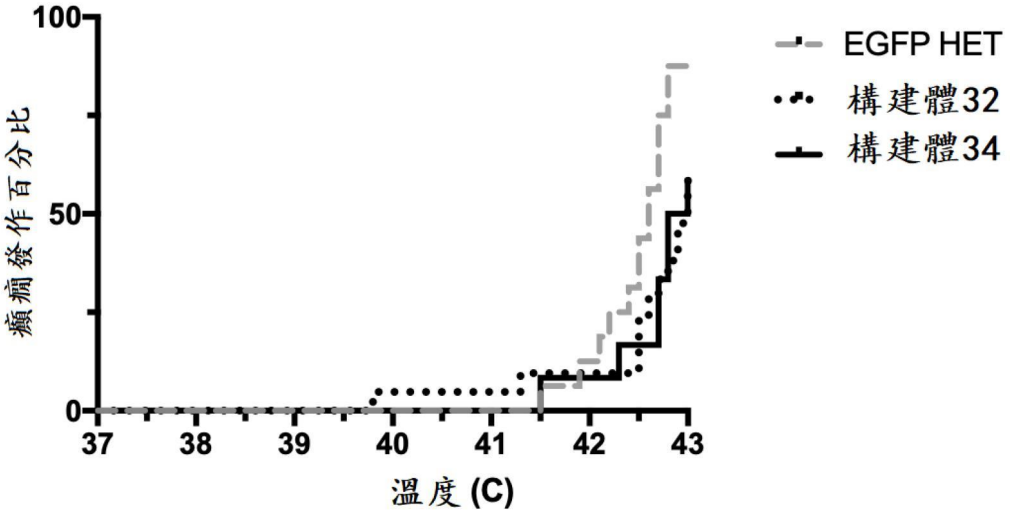
【圖6A】


100
50
0
癲癇發作百分比
37
38
39
40
41
42
43
溫度 (C)
EGFP HET
構建體32

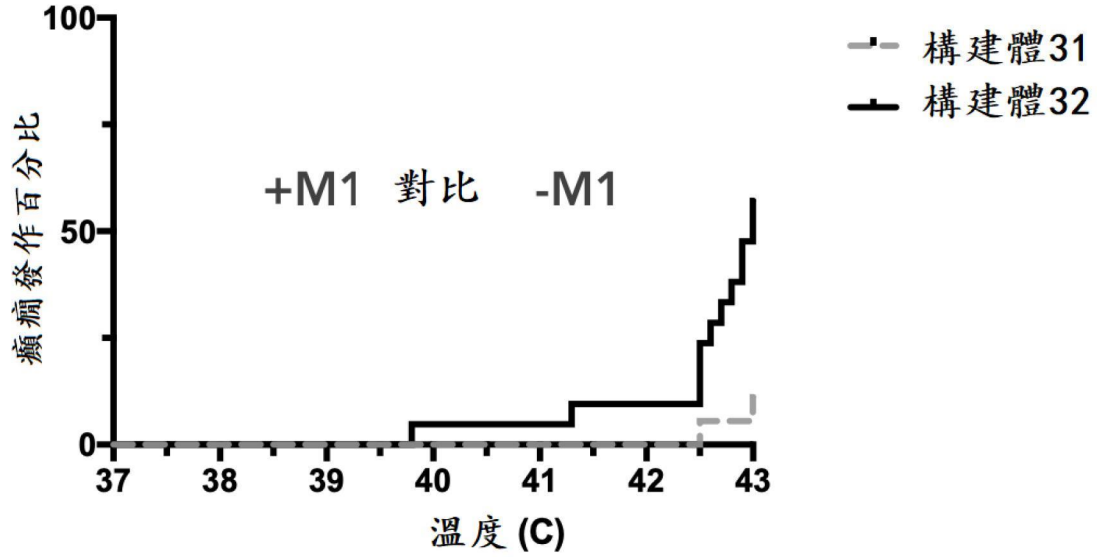
【圖6B】


100
50
0
癲癇發作百分比
37
38
39
40
41
42
43
溫度 (C)
EGFP HET
構建體33

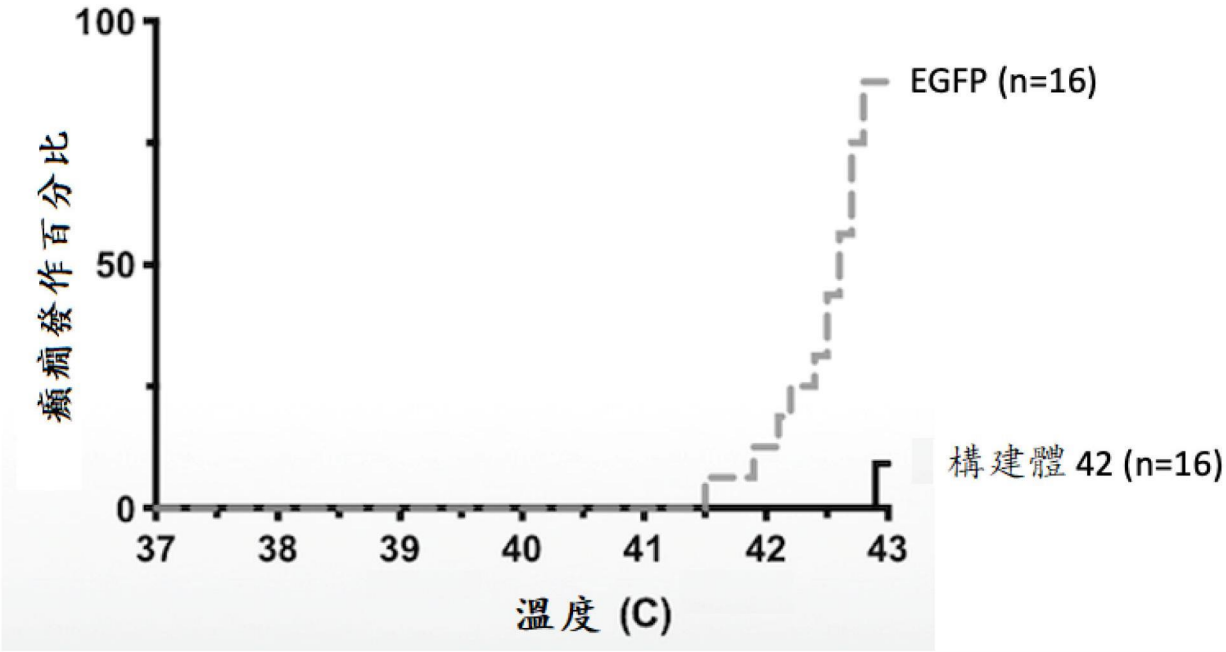
【圖6C】


100
50
0
癲癇發作百分比
37
38
39
40
41
42
43
溫度 (C)
EGFP HET
構建體32
構建體34

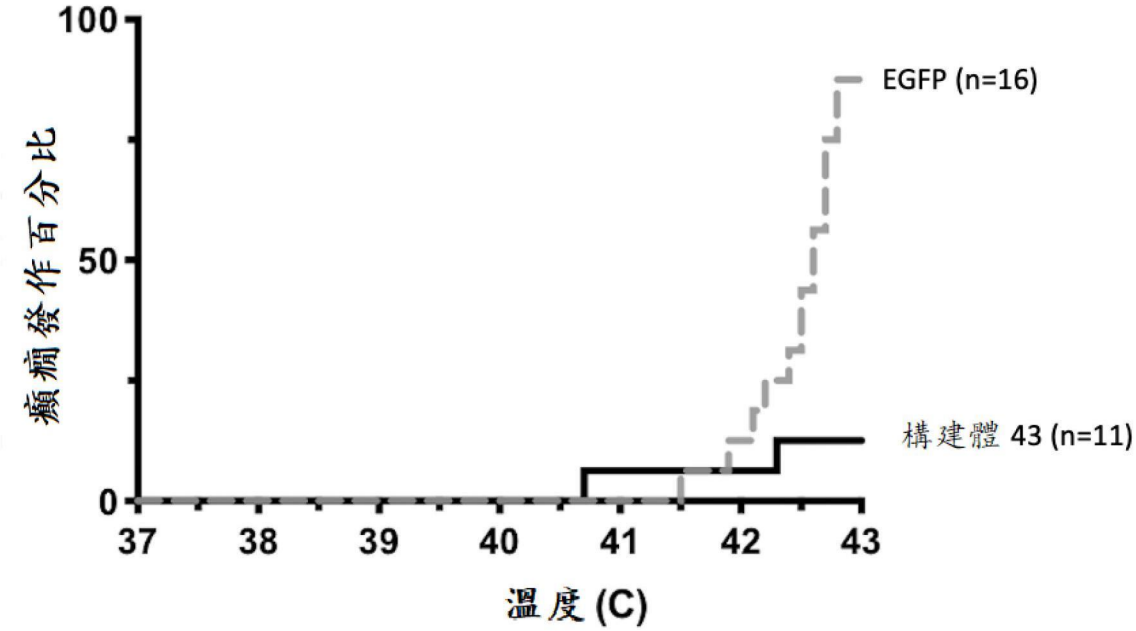
【圖6D】


100
50
0
癲癇發作百分比
+M1 對比 -M1
37
38
39
40
41
42
43
溫度 (C)
構建體31
構建體32

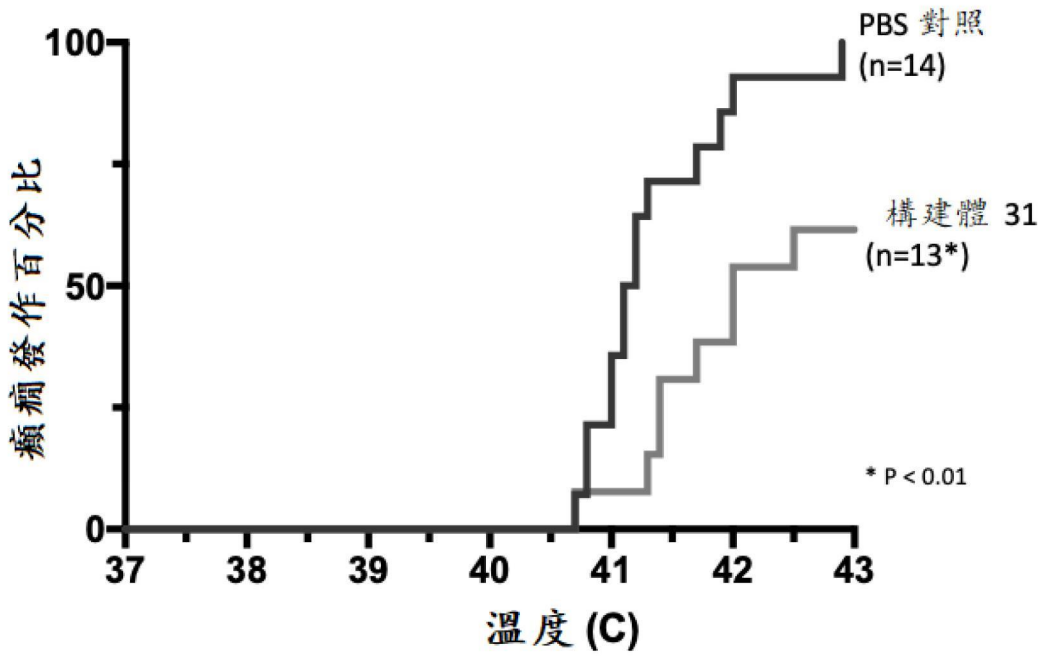
【圖6E】


100
50
0
癲癇發作百分比
37
38
39
40
41
42
43
溫度 (C)
EGFP (n=16)
構建體 42 (n=16)

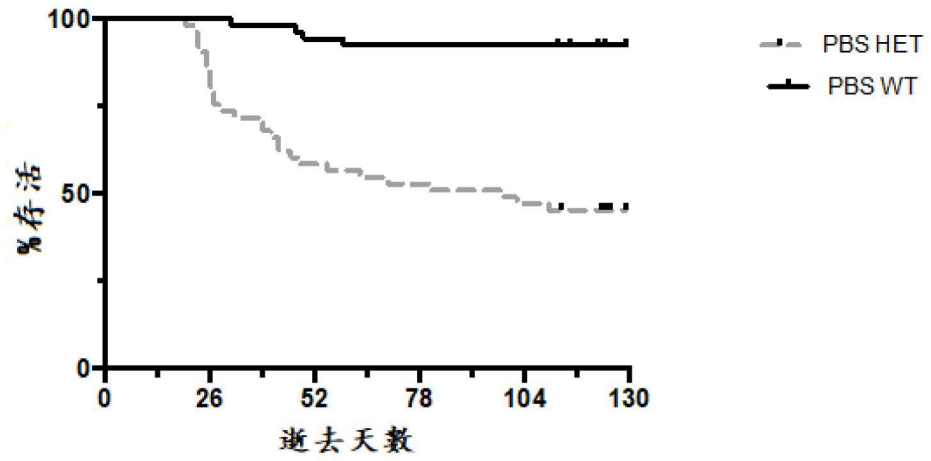
【圖6F】


100
50
0
癲癇發作百分比
37
38
39
40
41
42
43
溫度 (C)
EGFP (n=16)
構建體 43 (n=11)

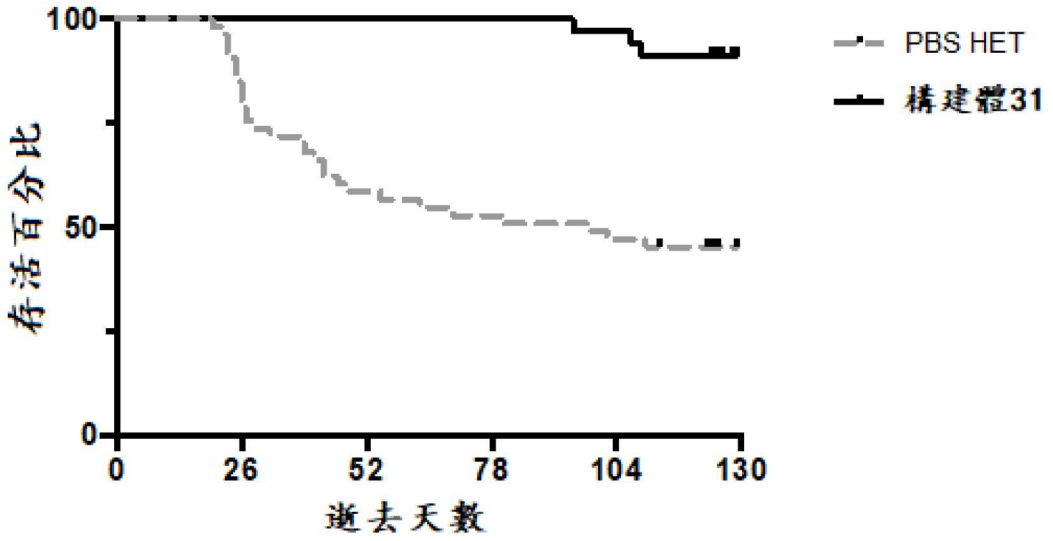
【圖6G】


100
50
0
癲癇發作百分比
37
38
39
40
41
42
43
溫度 (C)
PBS 對照
(n=14)
構建體 31
(n=13*)
* P < 0.01

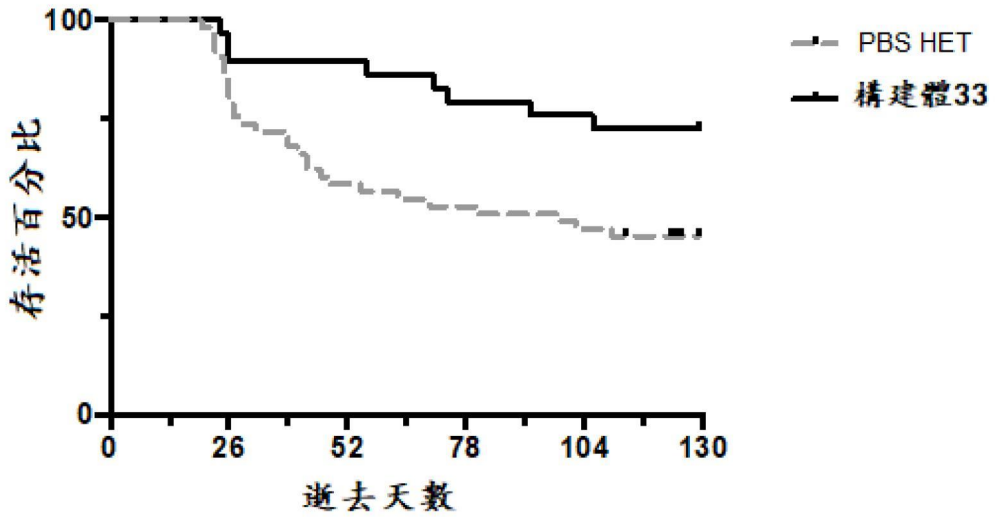
【圖6H】


100
50
0
%存活
0
26
52
78
104
130
逝去天數
PBS HET
PBS WT

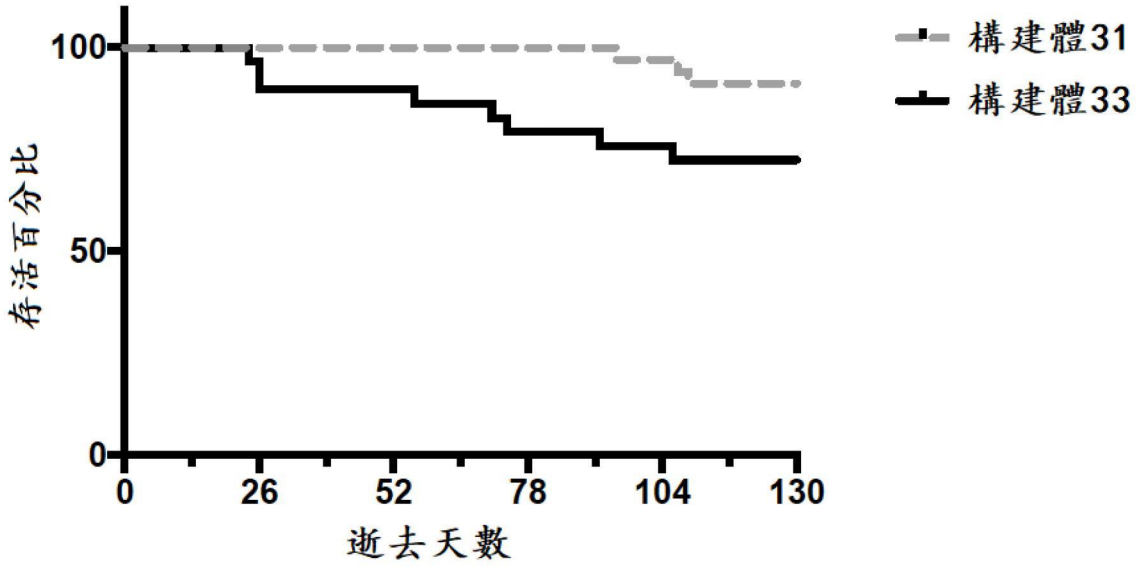
【圖7A】


100
50
0
存活百分比
0
26
52
78
104
130
逝去天數
PBS HET
構建體31

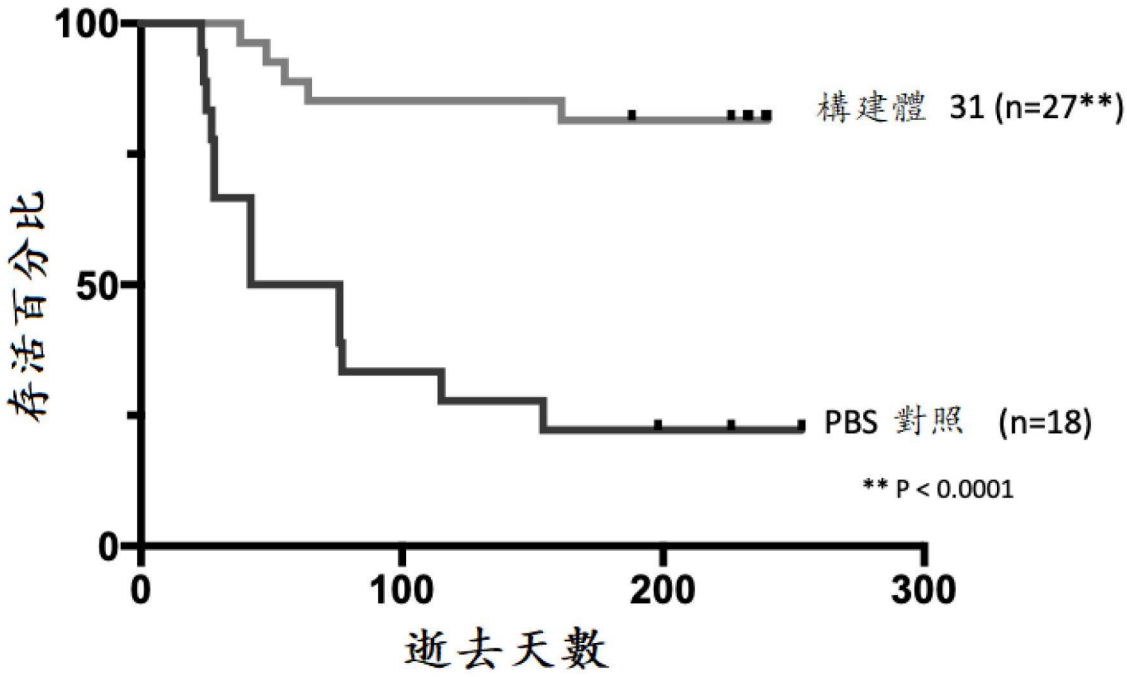
【圖7B】


100
50
0
存活百分比
0
26
52
78
104
130
逝去天數
PBS HET
構建體33

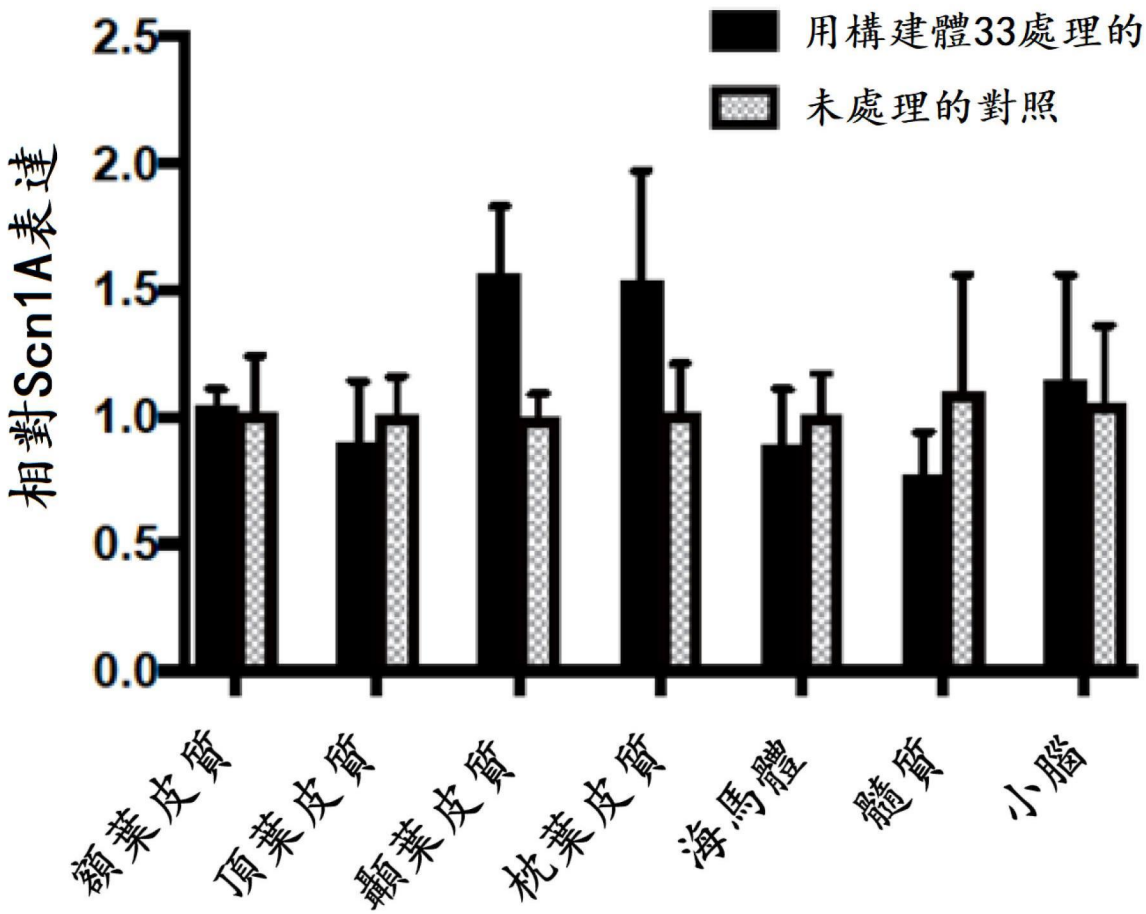
【圖7C】


100
50
0
存活百分比
0
26
52
78
104
130
逝去天數
構建體31
構建體33

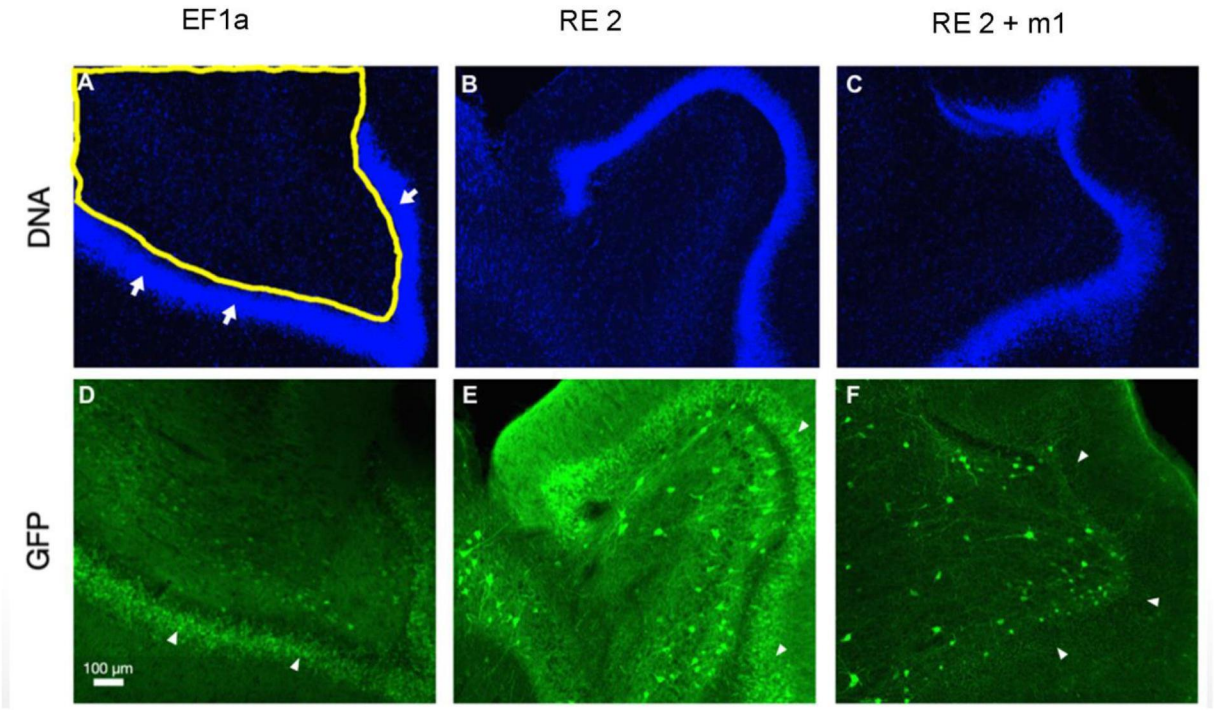
【圖7D】



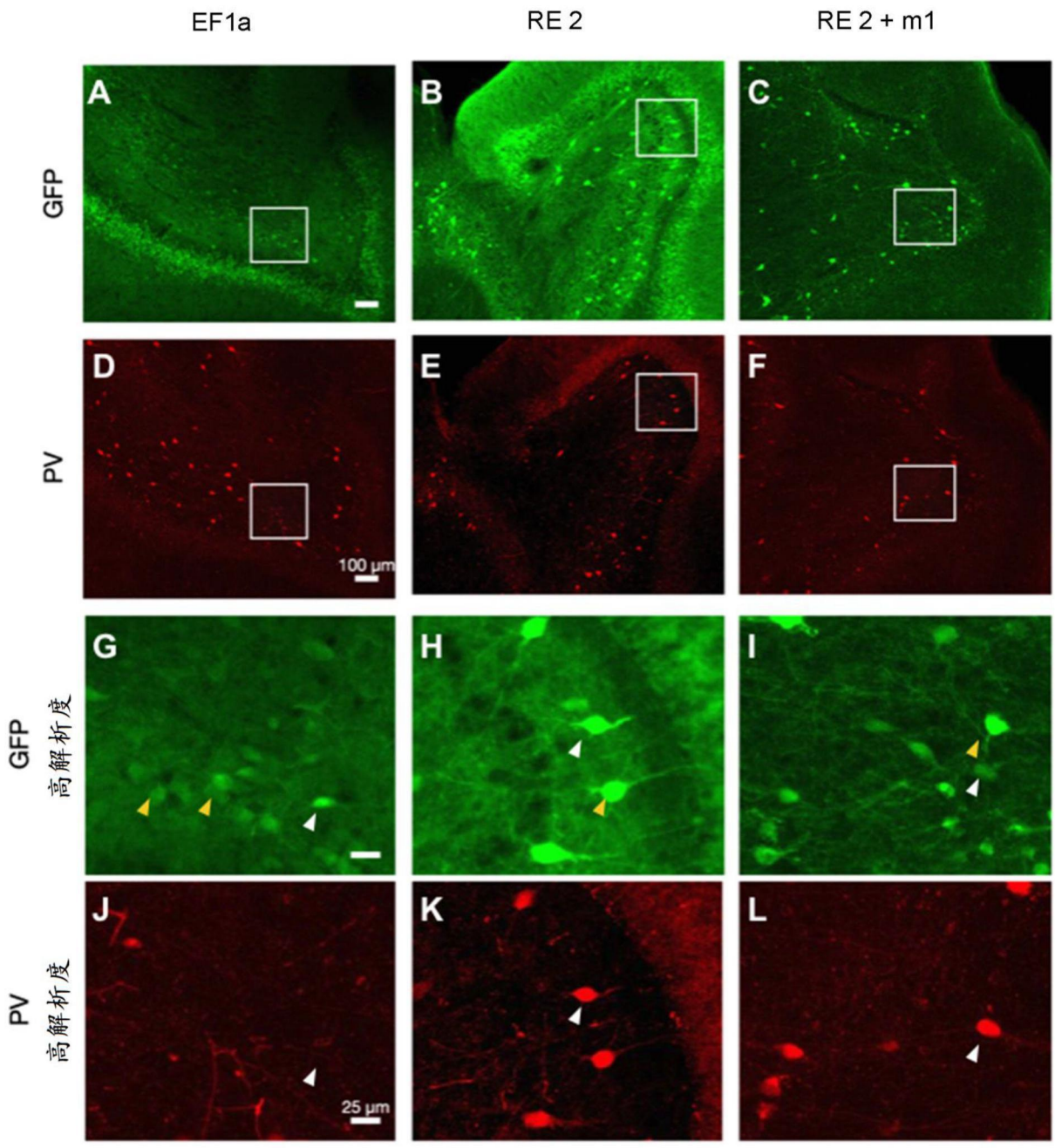
【圖7E】


用構建體33處理的
未處理的對照
相對Scn1A表達
2.5
2.0
1.5
1.0
0.5
0.0
額葉皮質
頂葉皮質
顳葉皮質
枕葉皮質
海馬體
髓質
小腦

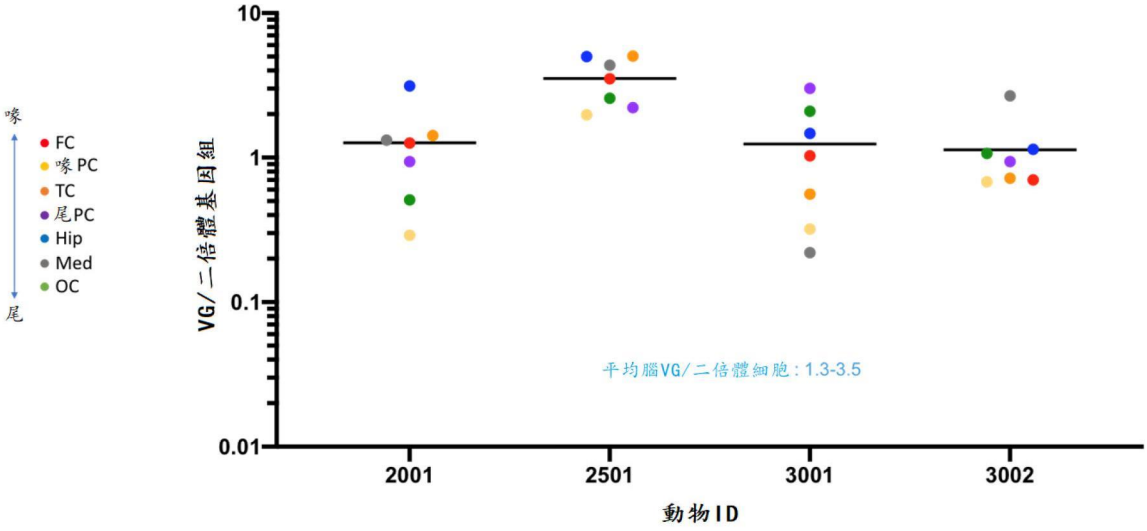
【圖8】



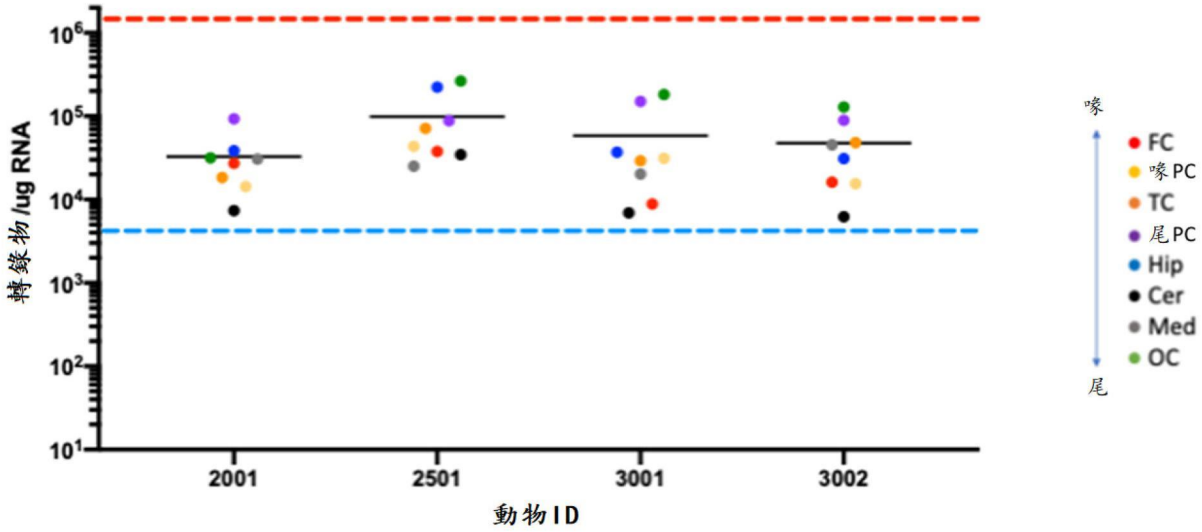
【圖9】


EF1a
RE 2
RE 2 + m1
GFP
PV
GFP 高解析度
PV 高解析度
A
B
C
D
E
F
G
H
I
J
K
L
100 μm
25 μm

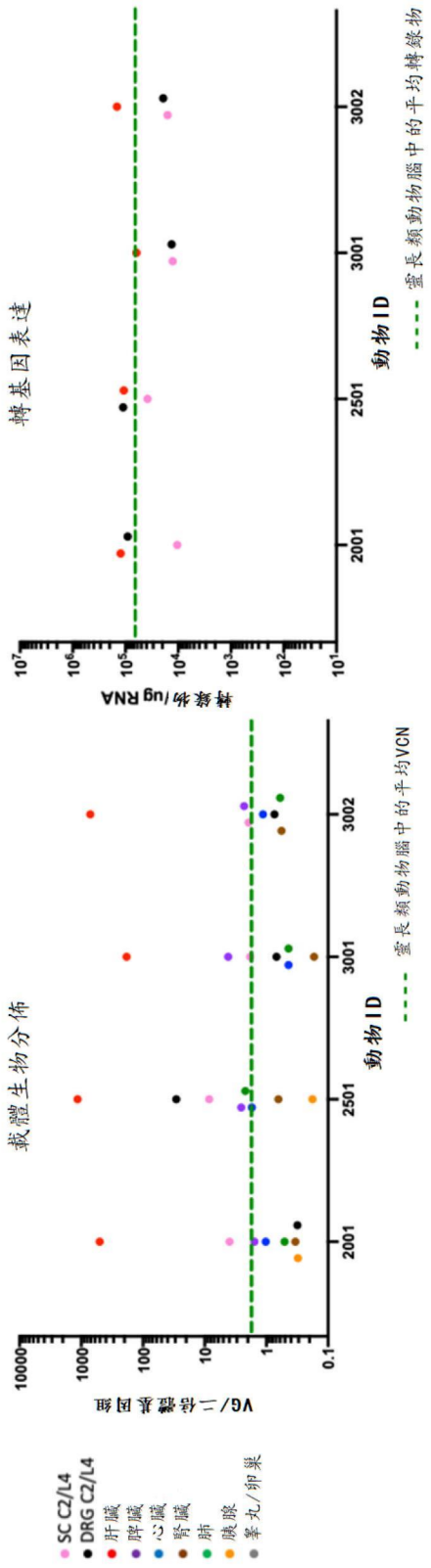
【圖10】


喙
FC
喙 PC
TC
尾PC
Hip
Med
OC
尾
VG/二倍體基因組
10
1
0.1
0.01
平均腦VG/二倍體細胞：1.3-3.5
2001
2501
3001
3002
動物ID

【圖11】


轉錄物/ug RNA
10^6
10^5
10^4
10^3
10^2
10^1
2001
2501
3001
3002
動物ID
喙
FC
喙PC
TC
尾PC
Hip
Cer
Med
OC
尾

【圖12】


載體生物分佈
VG/二倍體基因組
10000
1000
100
10
1
0.1
2001
2501
3001
3002
動物ID
--- 靈長類動物腦中的平均VCN
SC C2/L4
DRG C2/L4
肝臟
脾臟
心臟
腎臟
肺
胰腺
睾丸/卵巢
轉基因表達
轉錄物/ug RNA
10^7
10^6
10^5
10^4
10^3
10^2
10^1
2001
2501
3001
3002
動物ID
--- 靈長類動物腦中的平均轉錄物

【圖13】