(19) 中华人民共和国国家知识产权局



(12) 发明专利申请



(10) 申请公布号 CN 113122517 A (43) 申请公布日 2021.07.16

C12Q 1/6869 (2018.01) C12Q 1/6844 (2018.01) C12R 1/19 (2006.01)

(21)申请号 202110316008.9

(22) 申请日 2021.03.24

(71) 申请人 深圳清华大学研究院 地址 518000 广东省深圳市南山区科技园 高新南七道19号清华研究院 申请人 安序源生物科技(深圳)有限公司

(72) 发明人 高亚平 何筠 陈丽伊 田晖

(74) 专利代理机构 广州嘉权专利商标事务所有 限公司 44205

代理人 刘燚

(51) Int.CI.

C12N 9/12 (2006.01)

C12N 15/54 (2006.01)

C12N 15/70 (2006.01)

C12N 1/21 (2006.01)

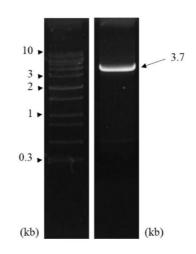
权利要求书1页 说明书10页 序列表16页 附图3页

(54) 发明名称

聚合酶突变体及其应用

(57) 摘要

本申请公开了聚合酶突变体及其应用。该聚合酶突变体包括如SEQ ID No.1所示的氨基酸序列经过以下突变后的氨基酸序列:第368位的苏氨酸突变为苯丙氨酸、第372位的苏氨酸突变为亮氨酸、第375位的谷氨酸突变为丝氨酸或赖氨酸、第478位的赖氨酸突变为酪氨酸、第484位的丙氨酸突变为谷氨酸、第512位的赖氨酸突变为酪氨酸。该聚合酶突变体至少具有如下有益效果:发明人意外地发现,具有如SEQ ID No.1所示氨基酸序列的Phi29聚合酶的上述位点发生突变后获得的突变体对于分子结构更复杂、分子量更大的核苷酸类似物底物具有更高的利用效率,可以有效利用核苷酸类似物进行模板DNA的复制和21延伸。



1.聚合酶突变体,其特征在于,包括如SEQ ID No.1所示的氨基酸序列经过以下突变后的氨基酸序列:第368位的苏氨酸突变为苯丙氨酸、第372位的苏氨酸突变为亮氨酸、第375位的谷氨酸突变为丝氨酸或赖氨酸、第478位的赖氨酸突变为酪氨酸、第484位的丙氨酸突变为谷氨酸、第512位的赖氨酸突变为酪氨酸:

优选的,所述聚合酶突变体的氨基酸序列如SEQ ID No.2~3中任一种所示。

- 2.分离的多核苷酸,其特征在于,包括:
- (a) 编码权利要求1所述的聚合酶突变体的核苷酸序列:或
- (b) 与(a) 所述的核苷酸序列互补的核苷酸序列:

优选的,所述多核苷酸的序列如SEQ ID No.5~6中任一种所示。

- 3. 重组载体, 其特征在于, 包括权利要求2所述的多核苷酸。
- 4.宿主细胞,其特征在于,包括权利要求3所述重组载体,或包括权利要求2所述的多核苷酸。
 - 5.权利要求1所述的聚合酶突变体在核酸的复制、扩增或测序中的应用。
 - 6.DNA模板的测序方法,其特征在于,包括以下步骤:
 - (1)提供DNA模板、引物、核苷酸底物及权利要求1所述的聚合酶突变体;
- (2) 在所述聚合酶突变体催化下,引导所述核苷酸底物掺入所述DNA模板,确定所述核苷酸底物的掺入顺序。
- 7.根据权利要求6所述的测序方法,其特征在于,所述核苷酸底物为核苷酸、核苷酸类似物中的至少一种。
- 8.根据权利要求7所述的测序方法,其特征在于,所述核苷酸类似物的分子量在1500~ 15000。
 - 9.组合物,其特征在于,包括权利要求1所述的聚合酶突变体。
 - 10.测序系统,其特征在于,所述测序系统包括权利要求1所述的聚合酶突变体。

聚合酶突变体及其应用

技术领域

[0001] 本申请涉及测序技术领域,尤其是涉及聚合酶突变体及其应用。

背景技术

[0002] 自1977年第一代测序技术的发明,测序技术已历经三代更迭发展。为适应广大的市场需求,纳米孔测序技术,凭借其小型化、自动化、低成本、高通量、高灵敏度等诸多优势,成为世界公认的第四代测序技术。与传统的测序技术不同,纳米孔测序可以在不对DNA进行生物或化学处理的前提下,直接采用物理办法读出DNA序列。其原理可以描述为:单个碱基通过纳米孔通道时,会引起通道电学参数的变化,对于A、C、T、G四种不同的碱基,其化学性质的差异所引起的电学参数的变化量也不同,对这些变化进行检测可以得到相应碱基的类型从而实现测序。

[0003] DNA链高速通过纳米孔的特性使得纳米孔测序成为可能,但其速度一般在微秒级别,过快的速度使得检测的信号质量较低甚至无法被检测到。相比较而言,DNA与DNA酶发生结合时,受酶催化反应的速度限制,其速度在毫秒级别左右,这提示可以通过DNA酶来控制DNA链在纳米孔中的通过速率,从而提高检测信号质量。因此,研发人员尝试将DNA外切酶与纳米孔复合,当DNA链靠近纳米孔时,DNA外切酶将核苷酸逐个剪下,剪下的核苷酸在电压驱动下通过纳米孔,相比于直接以DNA链通过的方式,速率放缓,对不同核苷酸的检测信号更容易区分。但利用外切酶测序过程中仍有许多需要改进的地方。因此,更多的研究针对DNA聚合酶控制DNA的通过速率。例如,有报道提出一种使用经过修饰的核苷酸类似物和纳米孔进行边合成边测序的方法,在该方法中,DNA聚合酶和纳米孔复合,当DNA模板链上具有部分杂交的引物并与纳米孔接触时,DNA聚合酶催化经过修饰的核苷酸类似物合成互补链,通过对该过程中引起的纳米孔电学参数的变化进行检测以完成测序。

[0004] Phi 29 DNA聚合酶即嗜热脂肪芽孢杆菌的噬菌体Phi 29中的p2蛋白。体外条件下,Phi 29能够催化寡核苷酸引物起始的延伸反应,而且特有的结构域组成和折叠使其具有高保真性和卓越的链置换能力,在众多的商业化酶中,持续合成能力最高(>70kb)、延伸速率最快(50-200bp/s)。然而,相比于一般的核苷酸,上述纳米孔测序方法中所采用的核苷酸类似物的分子结构与天然核苷酸底物明显不同,野生型Phi 29 DNA聚合酶对经修饰的核苷酸类似物底物的利用效率较低。

发明内容

[0005] 本申请旨在至少解决现有技术中存在的技术问题之一。为此,本申请提出一种能够对核苷酸类似物具有较高利用效率的聚合酶突变体及其应用。

[0006] 本申请的第一方面,提供聚合酶突变体,该聚合酶突变体包括如SEQ ID No.1所示的氨基酸序列经过以下突变后的氨基酸序列:第368位的苏氨酸突变为苯丙氨酸(T368F)、第372位的苏氨酸突变为亮氨酸(T372L)、第375位的谷氨酸突变为丝氨酸或赖氨酸(E375S或E375K)、第478位的赖氨酸突变为酪氨酸(K478Y)、第484位的丙氨酸突变为谷氨酸

(A484E)、第512位的赖氨酸突变为酪氨酸(K512Y)。

[0007] 其中,SEQ ID No.1所示的氨基酸序列如下:

[0008] MKHMPRKMYSCDFETTTKVEDCRVWAYGYMNIEDHSEYKIGNSLDEFMAWVLKVQADLYFHNLKFDGA
FIINWLERNGFKWSADGLPNTYNTIISRMGQWYMIDICLGYKGKRKIHTVIYDSLKKLPFPVKKIAKDFKLTVLKG
DIDYHKERPVGYKITPEEYAYIKNDIQIIAEALLIQFKQGLDRMTAGSDSLKGFKDIITTKKFKKVFPTLSLGLDK
EVRYAYRGGFTWLNDRFKEKEIGEGMVFDVNSLYPAQMYSRLLPYGEPIVFEGKYVWDEDYPLHIQHIRCEFELKE
GYIPTIQIKRSRFYKGNEYLKSSGGEIADLWLSNVDLELMKEHYDLYNVEYISGLKFKATTGLFKDFIDKWTYIKT
TSEGAIKQLAKLMLNSLYGKFASNPDVTGKVPYLKENGALGFRLGEEETKDPVYTPMGVFITAWARYTTITAAQAC
YDRIIYCDTDSIHLTGTEIPDVIKDIVDPKKLGYWAHESTFKRAKYLRQKTYIQDIYMKEVDGKLVEGSPDDYTDI
KFSVKCAGMTDKIKKEVTFENFKVGFSRKMKPKPVQVPGGVVLVDDTFTIKGTGSGALKTLESIVGDLEKADELKR
KYGSASAVRRLPVEELRELGFSDDEIAEIKGIPKKLREAFDLETAAELYERYGSLKEIGRRLSYDDLLELGATPKA
AAEIKGPEFKFLLNIEGVGPKLAERILEAVDYDLERLASLNPEELAEKVEGLGEELAERVVYAARERVESRRKSGR
QERSEEEWKEWLERKVGEGRARRLIEYFGSAGEVGKLVENAEVSKLLEVPGIGDEAVARLVPGYKTLRDAGLTPAE
AERVLKRYGSVSKVQEGATPDELRELGLGDAKIARILGLRSLVNKRLDVDTAYELKRRYGSVSAVRKAPVKELREL
GLSDRKIARIKGIPETMLQVRGMSVEKAERLLERFDTWTKVKEAPVSELVRVPGVGLSLVKEIKAQVDPAWKALLD
VKGVSPELADRLVEELGSPYRVLTAKKSDLMRVERVGPKLAERIRAAG。

[0009] 根据本申请实施例的聚合酶突变体,至少具有如下有益效果:

[0010] 发明人意外地发现,具有如SEQ ID No.1所示氨基酸序列的Phi29EL聚合酶的上述位点发生突变后获得的突变体对于分子结构更复杂、分子量更大的核苷酸类似物底物具有更高的利用效率,可以有效利用核苷酸类似物进行模板DNA的复制和延伸。

[0011] 根据本申请的一些实施方式,该聚合酶突变体的氨基酸序列如SEQ ID No.2~3中任一种所示。

[0012] 其中, SEQ ID No. 2所示的氨基酸序列如下:

[0013] MKHMPRKMYSCDFETTTKVEDCRVWAYGYMNIEDHSEYKIGNSLDEFMAWVLKVQADLYFHNLKFDGA
FIINWLERNGFKWSADGLPNTYNTIISRMGQWYMIDICLGYKGKRKIHTVIYDSLKKLPFPVKKIAKDFKLTVLKG
DIDYHKERPVGYKITPEEYAYIKNDIQIIAEALLIQFKQGLDRMTAGSDSLKGFKDIITTKKFKKVFPTLSLGLDK
EVRYAYRGGFTWLNDRFKEKEIGEGMVFDVNSLYPAQMYSRLLPYGEPIVFEGKYVWDEDYPLHIQHIRCEFELKE
GYIPTIQIKRSRFYKGNEYLKSSGGEIADLWLSNVDLELMKEHYDLYNVEYISGLKFKATTGLFKDFIDKWFYIKL
TSSGAIKQLAKLMLNSLYGKFASNPDVTGKVPYLKENGALGFRLGEEETKDPVYTPMGVFITAWARYTTITAAQAC
YDRIIYCDTDSIHLTGTEIPDVIKDIVDPYKLGYWEHESTFKRAKYLRQKTYIQDIYMKEVDGYLVEGSPDDYTDI
KFSVKCAGMTDKIKKEVTFENFKVGFSRKMKPKPVQVPGGVVLVDDTFTIKGTGSGALKTLESIVGDLEKADELKR
KYGSASAVRRLPVEELRELGFSDDEIAEIKGIPKKLREAFDLETAAELYERYGSLKEIGRRLSYDDLLELGATPKA
AAEIKGPEFKFLLNIEGVGPKLAERILEAVDYDLERLASLNPEELAEKVEGLGEELAERVVYAARERVESRRKSGR
QERSEEEWKEWLERKVGEGRARRLIEYFGSAGEVGKLVENAEVSKLLEVPGIGDEAVARLVPGYKTLRDAGLTPAE
AERVLKRYGSVSKVQEGATPDELRELGLGDAKIARILGLRSLVNKRLDVDTAYELKRRYGSVSAVRKAPVKELREL
GLSDRKIARIKGIPETMLQVRGMSVEKAERLLERFDTWTKVKEAPVSELVRVPGVGLSLVKEIKAQVDPAWKALLD
VKGVSPELADRLVEELGSPYRVLTAKKSDLMRVERVGPKLAERIRAAG。

[0014] SEQ ID No.3所示的氨基酸序列如下:

[0015] MKHMPRKMYSCDFETTTKVEDCRVWAYGYMNIEDHSEYKIGNSLDEFMAWVLKVQADLYFHNLKFDGA

FIINWLERNGFKWSADGLPNTYNTIISRMGQWYMIDICLGYKGKRKIHTVIYDSLKKLPFPVKKIAKDFKLTVLKG
DIDYHKERPVGYKITPEEYAYIKNDIQIIAEALLIQFKQGLDRMTAGSDSLKGFKDIITTKKFKKVFPTLSLGLDK
EVRYAYRGGFTWLNDRFKEKEIGEGMVFDVNSLYPAQMYSRLLPYGEPIVFEGKYVWDEDYPLHIQHIRCEFELKE
GYIPTIQIKRSRFYKGNEYLKSSGGEIADLWLSNVDLELMKEHYDLYNVEYISGLKFKATTGLFKDFIDKWFYIKL
TSKGAIKQLAKLMLNSLYGKFASNPDVTGKVPYLKENGALGFRLGEEETKDPVYTPMGVFITAWARYTTITAAQAC
YDRIIYCDTDSIHLTGTEIPDVIKDIVDPYKLGYWEHESTFKRAKYLRQKTYIQDIYMKEVDGYLVEGSPDDYTDI
KFSVKCAGMTDKIKKEVTFENFKVGFSRKMKPKPVQVPGGVVLVDDTFTIKGTGSGALKTLESIVGDLEKADELKR
KYGSASAVRRLPVEELRELGFSDDEIAEIKGIPKKLREAFDLETAAELYERYGSLKEIGRRLSYDDLLELGATPKA
AAEIKGPEFKFLLNIEGVGPKLAERILEAVDYDLERLASLNPEELAEKVEGLGEELAERVVYAARERVESRRKSGR
QERSEEEWKEWLERKVGEGRARRLIEYFGSAGEVGKLVENAEVSKLLEVPGIGDEAVARLVPGYKTLRDAGLTPAE
AERVLKRYGSVSKVQEGATPDELRELGLGDAKIARILGLRSLVNKRLDVDTAYELKRRYGSVSAVRKAPVKELREL
GLSDRKIARIKGIPETMLQVRGMSVEKAERLLERFDTWTKVKEAPVSELVRVPGVGLSLVKEIKAQVDPAWKALLD
VKGVSPELADRLVEELGSPYRVLTAKKSDLMRVERVGPKLAERIRAAG。

[0016] 其中,上述的聚合酶突变体中,突变前的Phi29EL聚合酶(SEQ ID No.1)中,在Phi29聚合酶的C端通过连接肽嵌合有坎德勒氏甲烷嗜热菌拓扑异构酶V的E~L的HhH基序。该拓扑异构酶V的E~L的HhH基序的氨基酸序列如SEQ ID No.4所示:

[0017] LKTLESIVGDLEKADELKRKYGSASAVRRLPVEELRELGFSDDEIAEIKGIPKKLREAFDLETAAELY ERYGSLKEIGRRLSYDDLLELGATPKAAAEIKGPEFKFLLNIEGVGPKLAERILEAVDYDLERLASLNPEELAEKV EGLGEELAERVVYAARERVESRRKSGRQERSEEEWKEWLERKVGEGRARRLIEYFGSAGEVGKLVENAEVSKLLEV PGIGDEAVARLVPGYKTLRDAGLTPAEAERVLKRYGSVSKVQEGATPDELRELGLGDAKIARILGLRSLVNKRLDV DTAYELKRRYGSVSAVRKAPVKELRELGLSDRKIARIKGIPETMLQVRGMSVEKAERLLERFDTWTKVKEAPVSEL VRVPGVGLSLVKEIKAQVDPAWKALLDVKGVSPELADRLVEELGSPYRVLTAKKSDLMRVERVGPKLAERIRAAG。

[0018] 在利用聚合酶进行纳米孔测序时,较高的盐浓度对于减缓过孔速率有重要作用,能够有效提高检测分辨率。但是,常规的聚合酶在高盐浓度条件下活性降低,因此,通过引入坎德勒氏甲烷嗜热菌拓扑异构酶V的HhH基序提高Phi29聚合酶对高盐溶液的耐受性,保证测序质量和速度。

[0019] 本申请的第二方面,提供分离的多核苷酸,该多核苷酸包括:

[0020] (a) 编码上述的聚合酶突变体的核苷酸序列;

[0021] (b) 与(a) 所述的核苷酸序列互补的核苷酸序列。

[0022] 根据本申请的一些实施方式,分离的多核苷酸的序列如SEQ ID No.5~6中任一种所示。

[0023] 本申请的第三方面,提供重组载体,该重组载体包括上述的多核苷酸。

[0024] 本申请的第四方面,提供宿主细胞,该宿主细胞包括上述的重组载体,或包括上述的多核苷酸。

[0025] 本申请的第五方面,提供上述的聚合酶突变体在核酸的复制、扩增或测序中的应用。

[0026] 本申请的第六方面,提供DNA模板的测序方法,该测序方法包括以下步骤:

[0027] (1)提供DNA模板、引物、核苷酸底物及上述的聚合酶突变体;

[0028] (2) 在聚合酶突变体催化下,引导核苷酸底物掺入DNA模板,确定核苷酸底物的掺

入顺序。

[0029] 其中,引物是指天然或人工合成的寡核苷酸,在适当条件下(例如在缓冲液中提供核苷酸底物、聚合酶等反应原料并提供特定的温度条件)能够作为核酸合成的起始点进行延伸,从而使核苷酸底物掺入DNA模板,合成DNA模板的互补链。核苷酸底物是指由戊糖、磷酸酯及含氮杂环碱基形成的核苷酸单元或这些核苷酸单元的修饰产物。修饰产物具体是指通过对核苷酸进行修饰标记,以使不同类别的核苷酸(包括脱氧腺苷、脱氧胸苷、脱氧胞苷、脱氧鸟苷、脱氧尿苷)的大小、质量、电荷量等产生明显区别,从而提高检测过程中不同类别核苷酸之间的区分度。一方面,核苷酸类似物可以是包含多个磷酸基团的核苷酸,具体可以是二磷酸核苷酸、三磷酸核苷酸、四磷酸核苷酸、五磷酸核苷酸、六磷酸核苷酸等多磷酸核苷酸。另一方面,为了增强对核苷酸类似物的鉴别,多磷酸核苷酸上还可以进一步针对不同碱基在磷酸基团上引入不同长度/质量的标签,具体的标签可以是任选的聚合物接头如聚乙二醇或其衍生物,以及染料分子如香豆素等。

[0030] 根据本申请的一些实施方式,在掺入过程中,聚合酶突变体复合在纳米孔上或与纳米孔处于相对接近的位置。其中,纳米孔 (nanopore) 是指具有纳米尺度孔隙的材料,具体可以包括生物纳米孔、固态纳米孔、以及生物纳米孔和固态纳米孔的复合纳米孔。生物纳米孔具体可以例举出包括但不限于金黄色葡萄球菌α-溶血素蛋白纳米孔(α-HL)、耻垢分枝杆菌孔道蛋白A (MspA)、噬菌体Phi 29连接器 (Phi 29 connector)。固态纳米孔具体可以例举出包括但不限于以氮化硅、二氧化硅、氧化铝、玻璃毛细管、二硫化钼、石墨烯等材料制备得到的纳米孔。

[0031] 根据本申请的一些实施方式,确定核苷酸底物的掺入顺序的具体方法包括但不限于根据不同的核苷酸底物的化学性质的差异,分析掺入过程中体系的电学信号和/或光学信号,从而得到不同的核苷酸底物的掺入顺序。其中,电学信号的非限制性实例包括电流信号,具体的检测方式例如检测离子电流、隧道电流等。光学信号的非限制性实例包括荧光信号,具体的检测方式例如通过荧光共振能量转移检测等。

[0032] 根据本申请的一些实施方式,核苷酸底物为核苷酸、核苷酸类似物中的至少一种。 [0033] 核苷酸类似物作为底物在聚合酶催化作用下合成与DNA模板链互补的产物链时, 修饰的标签通过聚合酶从核苷酸类似物中释放。释放出的标签通过跨孔电压等方式的介导,平移通过纳米孔,从而改变跨孔电流。通过检测其电流变化,可以推导出DNA链的碱基序列。

[0034] 根据本申请的一些实施方式,核苷酸类似物的分子量在1500~15000。当核苷酸类似物的分子量达到1500~15000时,标签的长度足够长,使过孔速率大大放缓,对纳米孔跨孔电流信号的检测分辨率明显提高。但同时,对于聚合酶对核苷酸类似物底物的利用效率也提出了更高的要求。另外,不同的核苷酸类似物携带有不同的标签,而不同标签的长度和结构的不同使其过孔时的阻滞电流和阻滞时间也不相同,从而能够区别不同的底物类型。优选的,核苷酸类似物的分子量为2000~15000、3000~15000、5000~15000。

[0035] 根据本申请的一些实施方式,核苷酸类似物为多磷酸核苷酸。

[0036] 根据本申请的一些实施方式,核苷酸类似物的磷酸基团在四个以上。

[0037] 根据本申请的一些实施方式,核苷酸类似物的磷酸基团在六个以上。磷酸基团越多,对于聚合酶的要求也越高,当核苷酸类似物的磷酸基团在六个以上时,本申请实施例中

所提供的聚合酶突变体仍然对其有较好的利用率。

[0038] 本申请的第七方面,提供组合物,该组合物包括上述的聚合酶突变体。该组合物具体可以是纳米孔和聚合酶突变体的组合物,纳米孔和聚合酶突变体通过共价连接或其它本领域熟知的方式结合到一起。在另外一些情况下,该组合物还可以是核苷酸底物和聚合酶突变体的组合物。在进行纳米孔测序或其它测序过程中,加入包含聚合酶突变体和核苷酸底物的混合液进行反应。

[0039] 本申请的第八方面,提供测序系统,该测序系统包括上述的聚合酶突变体。该测序系统能够以修饰后的核苷酸类似物作为底物参与待测序DNA模板的互补链的合成,有效进行延伸反应,从而完成测序。

[0040] 根据本申请的一些实施方式,该测序系统为纳米孔测序系统,该纳米孔测序系统 在纳米孔上复合有上述的聚合酶突变体。

[0041] 本申请的附加方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本申请的实践了解到。

附图说明

[0042] 图1是本申请的一个实施例的pGH-18B的PCR产物的电泳结果,左侧为marker,右侧为pGH-18B的条带。

[0043] 图2是本申请的一个实施例的pET28a-Phi29EL、pGH-14B和pGH-18B的质粒双酶切产物的电泳结果,(A) 为pET28a-Phi29EL,(B) 为pGH-14B,(C) 为pGH-18B; 左侧为marker,右侧为对应质粒的条带。

[0044] 图3是本申请的一个实施例的表达载体pET28a-Phi29EL-14B和pET28a-Phi29EL-18B的质粒双酶切产物的电泳结果,(A) 为pET28a-Phi29EL-14B,(B) 为pET28a-Phi29EL-18B; 左侧为marker,右侧为对应表达载体的条带。

[0045] 图4是本申请的一个实施例的聚合酶突变体的12% SDS-PAGE电泳结果,(A)为 Phi 29EL-14B聚合酶,(B)为Phi 29EL-18B聚合酶。

[0046] 图5是本申请的一个实施例中部分核苷酸类似物的结构式示意图。

[0047] 图6是本申请的一个实施例的发卡模板示意图。

[0048] 图7是本申请的一个实施例中发卡延伸反应实验结果的电泳图。

具体实施方式

[0049] 以下将结合实施例对本申请的构思及产生的技术效果进行清楚、完整地描述,以充分地理解本申请的目的、特征和效果。显然,所描述的实施例只是本申请的一部分实施例,而不是全部实施例,基于本申请的实施例,本领域的技术人员在不付出创造性劳动的前提下所获得的其他实施例,均属于本申请保护的范围。

[0050] 下面详细描述本申请的实施例,描述的实施例是示例性的,仅用于解释本申请,而不能理解为对本申请的限制。

[0051] 在本申请的描述中,若干的含义是一个以上,多个的含义是两个以上,大于、小于、超过等理解为不包括本数,以上、以下、以内等理解为包括本数。如果有描述到第一、第二只是用于区分技术特征为目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的

技术特征的数量或者隐含指明所指示的技术特征的先后关系。

[0052] 本申请的描述中,参考术语"一个实施例"、"一些实施例"、"示意性实施例"、"示例"、"具体示例"、或"一些示例"等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本申请的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任何的一个或多个实施例或示例中以合适的方式结合。

[0053] 下述实施例中使用的材料说明如下:

[0054] Phi 29-EL: 将野生型Phi 29聚合酶引入坎德勒氏甲烷嗜热菌拓扑异构酶V的HhH基序的E~L基序得到的嵌合体聚合酶(具体参考中国专利CN110747191A)。

[0055] pET28a-Phi29EL:在pET28a质粒载体上插入Phi29-EL的基因片段得到的重组载体 (具体参考中国专利CN110747191A)。

[0056] 定点突变试剂盒 Q5[®] Site-Directed Mutagenesis Kit:购自NEB公司,货号 E0552。

[0057] 大肠杆菌表达菌株BL21:购自Transgene公司。

[0058] 普通DNA产物纯化试剂盒:购自天根公司,货号DP204-02。

[0059] KpnI限制性内切酶:购自NEB公司,货号R0142S。

[0060] BstBI限制性内切酶:购自NEB公司,货号R0519S。

[0061] 连接溶液Solution I:购自Takara公司,货号6022Q。

[0062] 质粒提取试剂盒 Wolact® Plasmid DNA Purification Kit: 购自威邦生命科学(香港)有限公司,货号NP11001.50/200。

[0063] 实施例1

[0064] 本实施例提供聚合酶突变体,该聚合酶突变体的制备过程如下:

[0065] 1实验步骤

[0066] 1.1聚合酶突变体Phi29EL-14B基因片段合成

[0067] 以Phi29-EL的核苷酸序列为基础,通过全基因合成获得包含6个突变位点(T368F、T372L、E375S、K478Y、A484E、K512Y)的基因片段(记为pGH-14B,长度858bp),pGH-14B片段的5′和3′端分别包含BstBI和KpnI酶切位点,并连接到pGH载体上。

[0068] 1.2点突变生成pGH-18B基因片段

[0069] 以pGH-14B为模板,设计引物,利用定点突变试剂盒Q5[®]Site-Directed Mutagenesis Kit进行S375K定点突变,得到相对于Phi29-EL包含6个突变位点(T368F、T372L、E375K、K478Y、A484E、K512Y)的基因片段(记为pGH-18B)。准备PCR反应体系,如表1所示:

[0070] 表1.点突变反应体系

	试剂	用量
	Q5 Hot Start High-Fidelity 2× Master Mix	12.5 μL
[0071]	模板(10 ng/μL)	1 μL
	正向引物 (10 μM)	1.25 μL
	反向引物 (10 μM)	1.25 μL
	无核酸酶纯水	加至总体积 50μL

[0072] 突变正向引物的序列为: AAAGGAGCGATCAAGCAAC (SEQ ID No.7);

[0073] 突变反向引物的序列为: TGATGTCAACTTGATGTAAAAC (SEQ ID No.8)。

[0074] PCR扩增的条件为98℃变性30s,循环周期为98℃,10s,55℃,22s,72℃,2min,30个循环;72℃,2min。

[0075] 取PCR产物,进行Kinase-Ligase-DpnI(KLD)处理去除模板、连接片段。反应体系如表2所示:

[0076] 表2.PCR产物处理体系

•

[0078] 将反应体系混匀,室温孵育5min。转化大肠杆菌表达菌株BL21,将测序验证正确的载体pGH-18B,用于下一步实验。

[0079] 1.3载体酶切

[0080] 选取pET28a-Phi29EL、pGH-14B和pGH-18B的质粒进行双酶切。准备酶切反应体系,进行分步酶切。第一步酶切如下:

[0081] 表3.Kpn I酶切反应体系

	试剂	pET28a-Phi29EL 质粒	pGH-14B(或 pGH-18B)质粒
	10×Buffer 2.1	8 μL	8 μL
[0082]	质粒	5μg	5μg
	KpnI ($10U/\mu L$)	$4\mu L$	$4\mu L$
	ddH_2O	加至总体积 80 μL	加至总体积 80 μL

[0083] 酶切反应条件为37℃,16h。选取酶切产物,准备第二步酶切体系,如表4:

[0084] 表4.BstBI酶切反应体系

	试剂	pET28a-Phi29EL 质粒	pGH-14B (或 pGH-18B) 质粒
[0085]	KpnI 酶切产物	80 μL	80 μL
	BstBI (10 $U/\mu L$)	$4~\mu L$	$4~\mu L$

[0086] 酶切反应条件为:65℃,4h。

[0087] 酶切产物用琼脂糖凝胶电泳检测,目标条带选取凝胶用普通DNA产物纯化试剂盒进行回收。

[0088] 1.4片段连接

[0089] 连接Phi29EL片段、pGH-14B片段和pGH-18B,用1.3中pGH-14B(或pGH-18B)质粒酶 切回收产物片段替换野生型Phi29EL片段中对应片段,得到Phi29EL-14B(或Phi29EL-18B)。准备连接反应体系,如表5:

[0090] 表5.Phi29EL-14B(或Phi29EL-18B)连接反应体系

	试剂	用量
	Solution I	5 μL
[0091]	Phi29EL 片段酶切回收产物	2.5 μL
	14B(或 18B) 片段酶切回收产物	2.5 μL

[0092] 连接溶液为Solution I。连接反应条件为:16℃,30min。

[0093] 取连接产物转化大肠杆菌Top10感受态细胞,挑取阳性克隆。选用质粒提取试剂盒 Wolact[®]Plasmid DNA Purification Kit提取质粒,进行测序,获得序列正确的突变型 Phi29DNA聚合酶质粒pET28a-Phi29EL-14B(或pET28a-Phi29EL-18B)。

[0094] 1.5蛋白表达纯化

[0095] 取质粒pGEX6P-1-Phi29、pET28a-Phi29EL、pET28a-Phi29EL-14B和pET28a-Phi29EL-18B,转化大肠杆菌表达菌株BL21,表达野生型Phi29 DNA聚合酶、Phi29-EL DNA聚合酶、Phi29EL-14B DNA聚合酶和Phi29EL-18B DNA聚合酶。蛋白诱导表达选择低温诱导,诱导条件为16℃,16h。

[0096] 蛋白纯化步骤包括:

[0097] ①收集菌体,超声破碎细胞。功率50~100W,超声5~15s,间隔10~30s。共持续20~40min。

[0098] ②离心去除细胞碎片,取上清。分别采用采用Glutathione Sepharose 4B和Ni Sepharose亲和层析柱纯化步骤裂解粗产物。

[0099] ③用12%SDS-PAGE检测蛋白纯度和浓度。

[0100] 2.实验结果

[0101] 2.1pGH-18B点突变结果

[0102] 图1是pGH-18B的PCR产物的电泳结果,扩增产物约为4kb。从图中可以看出,该次PCR已扩增出目的条带。

[0103] 2.2酶切结果

[0104] 图2是质粒双酶切结果。其中,(A) 为pET28a-Phi29EL,(B) 为pGH-14B,(C) 为pGH-18B; (A)、(B)、(C) 中左侧为marker,右侧为对应质粒的两个重复的条带。结合图2,pET28a-Phi29EL、pGH-14B和pGH-18B经质粒双酶切分别回收图中所示的7000bp左右和855bp左右的片段。

[0105] 2.3pET28a-Phi29EL-14B载体酶切验证

[0106] 图3是表达载体双酶切结果。其中,(A) 为pET28a-Phi29EL-14B,(B) 为pET28a-Phi29EL-18B; (A)、(B) 中左侧为marker,右侧为对应表达载体的条带。其中,pET28a-Phi29EL-14B用KpnI和BstBI双酶切,pET28a-Phi29EL-18B用EcoRI和NotI双酶切,并对正确的克隆菌株测序验证。

[0107] 2.4Phi29EL-14B和Phi29EL-18B蛋白检测

[0108] 图4是聚合酶突变体的12%SDS-PAGE电泳结果,(A)为Phi29EL-14B聚合酶的结果,(B)为Phi29EL-18B聚合酶的结果。其中,M为蛋白分子量Marker;泳道1和泳道2为突变型聚合酶的两个重复。结合图4,上述聚合酶突变体的分子量均为116.6kD。

[0109] 其中,Phi29EL-14B的氨基酸序列如SEQ ID No.2所示,核苷酸序列如SEQ ID No.5 所示;Phi29EL-18B的氨基酸序列如SEQ ID No.3所示,核苷酸序列如SEQ ID No.6所示。

[0110] 实施例2

[0111] 聚合酶突变体底物结合能力检测

[0112] 天然核苷酸底物由碱基、核糖和三个磷酸组成,每个核苷酸的整合,伴随着DNA链的延伸和焦磷酸的释放。在DNA聚合酶辅助的纳米孔测序技术中,单个DNA聚合酶分子共价结合到α-溶血素七聚物纳米孔上,DNA聚合酶结合模板引物复合物后,以携带标记的核苷酸为底物,进行边合成边测序。纳米孔依次捕获和检测携带标记的核苷酸,从而产生核苷酸特异的电流信号。根据这一原理,本实施例中设计合成了包含六个磷酸基团、聚乙二醇(PEG)侧链和香豆素(Coumarin)的核苷酸类似物,记为Coumarin-PEGm-dN6P(CPm-dN6P),其中,m表示核苷酸类似物中聚乙二醇单元的重复数量,N表示A、T、C、G中的任一种。

[0113] 具体如下: Coumarin-PEG12-dA6P (CP12-dA6P)、Coumarin-PEG24-dC6P (CP24-dC6P)、Coumarin-PEG24-dA6P (CP24-dA6P)、Coumarin-PEG24-dT6P (CP24-dT6P)、Coumarin-PEG24-dG6P (CP24-dG6P)、Coumarin-PEG36-dT6P (CP36-dT6P) 和Coumarin-PEG44-dG6P (CP44-dG6P),部分核苷酸类似物的结构式如图5中的 $(A) \sim (D)$ 所示。

[0114] 取等量的Phi29ELDNA聚合酶、Phi29EL-14B DNA聚合酶和Phi29EL-18B DNA聚合酶,以核苷酸类似物为底物,进行发夹延伸实验,检测突变体底物结合能力。发夹模板如图6 所示,长度为53nt,延伸产物长度为74nt。延伸反应结束后,模板链含量减少,生成产物链。根据有无产物链以及产物链的含量,即可判断DNA聚合酶对特异性底物的结合能力。

[0115] 将200nM实施例1中制得的DNA聚合酶、500nM的发夹模板及反应缓冲液混合,得到反应体系,另外以Phi29EL DNA聚合酶作为对照。反应体系包括:dNTP(-KC1)组、dNTP(+KC1)组、CP-dN6P(-KC1)组、CP-dN6P(+KC1)组,具体组成如表6所示:

[0116] 其中,反应缓冲液由50mM的Tris-HCl、10mM的MgCl $_2$ 、10mM的 (NH $_4$) $_2$ SO $_4$ 、4mM的DTT及 200 μ M dNTPs组成,反应缓冲液在25℃条件下的pH为7.5。将反应体系在温度为30℃的条件下反应10min,得到反应产物。

[0117] 表6发夹延伸反应体系

	21 ZI	dNTP	dNTP	CP-dN6P	CP-dN6P
	试剂	(-KCl)	(+KCl)	(-KC1)	(+KCl)
	10×Buffer	2 μL	2 μL	2 μL	2 μL
	牛血清白蛋白(20 mg/ml)	0.1 μL	$0.1~\mu L$	0.1 μL	$0.1~\mu L$
[0118]	dNTP (1 mM)	$0.4~\mu L$	$0.4~\mu L$	-	-
	dNTP 类似物(1mM)	-	-	$0.8~\mu L$	$0.8~\mu L$
	酶(2 μΜ)	$1 \mu L$	1 μL	1 μL	1 μL
	KCI (2 M)	-	$3~\mu L$	-	$3~\mu L$
	H_2O	到 20 μL	到 20 μL	到 20 μL	到 20 μL

[0119] 将反应产物加入6×DNA loading后,上样到20%TBE-Urea聚丙烯酰胺凝胶中;电泳条件为180V,3h。电泳结束后,将PAGE胶浸泡在SYBR Gold染液中,缓慢水平震荡约20min。然后利用Azure Biosystems成像检测模板链和产物链的亮度。通过模板链和产物链的亮度确定实施例1制备的DNA聚合酶的活性。

[0120] 实验结果如图7所示,从图中可以看出,实施例1制备的DNA聚合酶的具有DNA酶的活性,且Phi29EL-14B和Phi29EL-18B分别能够以CP24-dC6P、CP24-dA6P、CP24-dT6P和CP24-dG6P的混合物为底物,还能够以CP12-dA6P、CP24-dC6P、CP36-dT6P和CP44-dG6P混合物为底物,合成全长的延伸产物。而作为对照的Phi29EL DNA聚合酶以两类核苷酸类似物为底物时,全部或者大部分产物为不完全延伸产物。

[0121] 比较不同KC1浓度对反应的影响,从图中可以看出,添加0.3M KC1时,三种DNA聚合酶的活性均有一定程度的下降,但仍然能够利用核苷酸类似物合成产物链。其中,Phi29EL-14B和Phi29EL-18B这两种DNA聚合酶能够利用天然底物和修饰底物,合成少量全长产物;而Phi29EL聚合酶仅能生成少量的不完全延伸产物。证实Phi29EL-14B和Phi29EL-18B这两种DNA聚合酶在低盐和高盐浓度下,均能有效利用包含PEG侧链的核苷酸类似物为底物,进行DNA复制和延伸。

[0122] 上面结合实施例对本申请作了详细说明,但是本申请不限于上述实施例,在所属技术领域普通技术人员所具备的知识范围内,还可以在不脱离本申请宗旨的前提下作出各种变化。此外,在不冲突的情况下,本申请的实施例及实施例中的特征可以相互组合。

```
SEQUENCE LISTING
<110> 深圳清华大学研究院
安序源生物科技(深圳)有限公司
〈120〉聚合酶突变体及其应用
<130> 1
<160> 8
<170> PatentIn version 3.5
<210> 1
<211> 1028
<212> PRT
〈213〉人工序列
<400> 1
Met Lys His Met Pro Arg Lys Met Tyr Ser Cys Asp Phe Glu Thr Thr
                5
                                    10
Thr Lys Val Glu Asp Cys Arg Val Trp Ala Tyr Gly Tyr Met Asn Ile
                                25
Glu Asp His Ser Glu Tyr Lys Ile Gly Asn Ser Leu Asp Glu Phe Met
                            40
                                                45
        35
Ala Trp Val Leu Lys Val Gln Ala Asp Leu Tyr Phe His Asn Leu Lys
                        55
                                            60
Phe Asp Gly Ala Phe Ile Ile Asn Trp Leu Glu Arg Asn Gly Phe Lys
                    70
                                                            80
65
Trp Ser Ala Asp Gly Leu Pro Asn Thr Tyr Asn Thr Ile Ile Ser Arg
                85
                                    90
                                                        95
Met Gly Gln Trp Tyr Met Ile Asp Ile Cys Leu Gly Tyr Lys Gly Lys
            100
                                105
Arg Lys Ile His Thr Val Ile Tyr Asp Ser Leu Lys Lys Leu Pro Phe
        115
                            120
                                                125
Pro Val Lys Lys Ile Ala Lys Asp Phe Lys Leu Thr Val Leu Lys Gly
                        135
                                            140
Asp Ile Asp Tyr His Lys Glu Arg Pro Val Gly Tyr Lys Ile Thr Pro
                    150
145
                                        155
                                                            160
Glu Glu Tyr Ala Tyr Ile Lys Asn Asp Ile Gln Ile Ile Ala Glu Ala
                165
                                    170
Leu Leu Ile Gln Phe Lys Gln Gly Leu Asp Arg Met Thr Ala Gly Ser
            180
                                185
                                                    190
Asp Ser Leu Lys Gly Phe Lys Asp Ile Ile Thr Thr Lys Lys Phe Lys
        195
                            200
                                                205
```

Lys	Val 210	Phe	Pro	Thr	Leu	Ser 215	Leu	Gly	Leu	Asp	Lys 220	Glu	Val	Arg	Tyr
Ala 225		Arg	G1y	Gly	Phe 230		Trp	Leu	Asn	Asp 235		Phe	Lys	G1u	Lys 240
Glu	Ile	Gly	G1u	Gly 245	Met	Val	Phe	Asp	Val 250	Asn	Ser	Leu	Tyr	Pro 255	Ala
Gln	Met	Tyr	Ser 260	Arg	Leu	Leu	Pro	Tyr 265	Gly	Glu	Pro	Ile	Val 270	Phe	Glu
Gly	Lys	Tyr 275	Val	Trp	Asp	Glu	Asp 280	Tyr	Pro	Leu	His	I1e 285	G1n	His	Ile
Arg	Cys 290	Glu	Phe	Glu	Leu	Lys 295	Glu	Gly	Tyr	Ile	Pro 300	Thr	Ile	G1n	Ile
Lys 305	Arg	Ser	Arg	Phe	Tyr 310	Lys	Gly	Asn	Glu	Tyr 315	Leu	Lys	Ser	Ser	Gly 320
Gly	G1u	Ile	Ala	Asp 325	Leu	Trp	Leu	Ser	Asn 330	Val	Asp	Leu	G1u	Leu 335	Met
Lys	Glu	His	Tyr 340	Asp	Leu	Tyr	Asn	Val 345	Glu	Tyr	Ile	Ser	Gly 350	Leu	Lys
Phe	Lys	Ala 355	Thr	Thr	Gly	Leu	Phe 360	Lys	Asp	Phe	Ile	Asp 365	Lys	Trp	Thr
Tyr	I1e 370	Lys	Thr	Thr	Ser	G1u 375	Gly	Ala	Ile	Lys	G1n 380	Leu	Ala	Lys	Leu
Met 385	Leu	Asn	Ser	Leu	Tyr 390	Gly	Lys	Phe	Ala	Ser 395	Asn	Pro	Asp	Val	Thr 400
Gly	Lys	Val	Pro	Tyr 405	Leu	Lys	Glu	Asn	Gly 410	Ala	Leu	Gly	Phe	Arg 415	Leu
Gly	Glu	Glu	G1u 420	Thr	Lys	Asp	Pro	Val 425	Tyr	Thr	Pro	Met	Gly 430	Val	Phe
Ile	Thr	Ala 435	Trp	Ala	Arg	Tyr	Thr 440	Thr	Ile	Thr	Ala	Ala 445	Gln	Ala	Cys
Tyr	Asp 450	Arg	Ile	Ile	Tyr	Cys 455	Asp	Thr	Asp	Ser	Ile 460	His	Leu	Thr	Gly
Thr 465	Glu	Ile	Pro	Asp	Val 470	Ile	Lys	Asp	Ile	Val 475	Asp	Pro	Lys	Lys	Leu 480
Gly	Tyr	Trp	Ala	His 485	Glu	Ser	Thr	Phe	Lys 490	Arg	Ala	Lys	Tyr	Leu 495	Arg
Gln	Lys	Thr	Tyr 500	Ile	G1n	Asp	Ile	Tyr 505	Met	Lys	G1u	Val	Asp 510	Gly	Lys
Leu	Val	Glu	Gly	Ser	Pro	Asp	Asp	Tyr	Thr	Asp	Ile	Lys	Phe	Ser	Val

	515					520					525			
Lys Cys	Ala	G1y	Met	Thr	Asp	Lys	Ile	Lys	Lys	G1u	Val	Thr	Phe	Glu
530					535					540				
Asn Phe	Lys	Val	Gly	Phe	Ser	Arg	Lys	Met	Lys	Pro	Lys	Pro	Val	Gln
545				550					555					560
Val Pro	Gly	Gly		Val	Leu	Val	Asp		Thr	Phe	Thr	Ile		Gly
_, _,	_		565	_	_		_	570	_				575	_
Thr Gly	Ser		Ala	Leu	Lys	Thr		Glu	Ser	He	Val		Asp	Leu
C1 I	۸.1	580	01	T	т	٨	585	т	01	C	A 1	590	۸ 1	1 7 1
Glu Lys		Asp	GIU	Leu	Lys		Lys	lyr	Gly	Ser		Ser	Ala	vai
Arg Arg	595	Dno	Vo 1	C1,,	C1,,	600	Ana	C1,,	Lou	C1,,,	605	Son	Aan	Aan
610	Leu	110	vai	GIU	615	Leu	лгg	GIU	Leu	620	rne	261	лър	лѕр
Glu Ile	Ala	G111	T1e	Lvs		T1e	Pro	Lvs	Lvs		Arg	G111	Ala.	Phe
625		0 2 0		630	01)			-, -	635		0	0 2 0		640
Asp Leu	Glu	Thr	Ala	Ala	Glu	Leu	Tyr	Glu	Arg	Tyr	Gly	Ser	Leu	Lys
			645					650					655	
Glu Ile	Gly	Arg	Arg	Leu	Ser	Tyr	Asp	Asp	Leu	Leu	G1u	Leu	G1y	Ala
		660					665					670		
Thr Pro	Lys	Ala	Ala	Ala	Glu	Ile	Lys	G1y	Pro	Glu	Phe	Lys	Phe	Leu
	675					680					685			
Leu Asn	Ile	G1u	G1y	Val		Pro	Lys	Leu	Ala		Arg	Ile	Leu	Glu
690				_	695		_		_	700		_		
Ala Val	Asp	Tyr	Asp		Glu	Arg	Leu	Ala		Leu	Asn	Pro	Glu	
705	01	т	17 1	710	01	T	01	01	715	T	A 1	01	Λ	720
Leu Ala	GIU	Lys	va1 725	GIU	GIY	Leu	GIY	730	GIU	Leu	Ala	GIU	Arg 735	vaı
Val Tyr	Δ1a	Δ1a		G111	Ara	Va1	C111		Ara	Ara	Lvc	Sor		Δνα
vai iyi	пια	740	mg	oru	nı g	vai	745	261	mg	nı g	Lys	750	Uly	mg
Gln Glu	Arg		G1u	G1u	G1u	Trp		Glu	Trp	Leu	G1u		Lvs	Val
	755					760	J				765	O	J	
Gly Glu	Gly	Arg	Ala	Arg	Arg	Leu	Ile	Glu	Tyr	Phe	G1y	Ser	Ala	Gly
770					775					780				
Glu Val	Gly	Lys	Leu	Val	Glu	Asn	Ala	Glu	Val	Ser	Lys	Leu	Leu	Glu
785				790					795					800
Val Pro	Gly	Ile	Gly	Asp	Glu	Ala	Val	Ala	Arg	Leu	Val	Pro	Gly	Tyr
			805					810					815	
Lys Thr	Leu		Asp	Ala	G1y	Leu		Pro	Ala	G1u	Ala	G1u	Arg	Val
		820					825					830		

Leu Lys Arg 835	Tyr Gl	Ser	Val	Ser 840	Lys	Val	Gln	Glu	Gly 845	Ala	Thr	Pro
Asp Glu Leu 850	Arg Glu	ı Leu	Gly 855	Leu	Gly	Asp	Ala	Lys 860	Ile	Ala	Arg	Ile
Leu Gly Leu 865	Arg Sei	E Leu 870	Val	Asn	Lys	Arg	Leu 875	Asp	Val	Asp	Thr	Ala 880
Tyr Glu Leu	Lys Arg	_	Tyr	Gly	Ser	Val 890	Ser	Ala	Val	Arg	Lys 895	Ala
Pro Val Lys	Glu Let 900	ı Arg	Glu	Leu	Gly 905	Leu	Ser	Asp	Arg	Lys 910	Ile	Ala
Arg Ile Lys 915	Gly Ile	e Pro	Glu	Thr 920	Met	Leu	Gln	Val	Arg 925	Gly	Met	Ser
Val Glu Lys 930	Ala Gli	ı Arg	Leu 935	Leu	Glu	Arg	Phe	Asp 940	Thr	Trp	Thr	Lys
Val Lys Glu 945	Ala Pro	950 Val	Ser	Glu	Leu	Val	Arg 955	Val	Pro	Gly	Val	Gly 960
Leu Ser Leu	Val Lys 965		Ile	Lys	Ala	Gln 970	Val	Asp	Pro	Ala	Trp 975	Lys
Ala Leu Leu	Asp Va.	l Lys	G1y	Val	Ser 985	Pro	Glu	Leu	Ala	Asp 990	Arg	Leu
Val Glu Glu 995	Leu Gl	ser Ser	Pro	Tyr		Val	Leu	Thr	Ala 1008		Lys	Ser
Asp Leu Met 1010	Arg Va	l Glu	Arg 101		Gly	Pro	Lys	Leu 1020		Glu	Arg	
Ile Arg Ala 1025	Ala Gl	7										
<210> 2												
<211> 1028												
<212> PRT <213> 人工序	≅ <i>万</i> 山											
<400> 2	70											
Met Lys His	Met Pro	Arg	Lys	Met	Tyr	Ser	Cys	Asp	Phe	Glu	Thr	Thr
1	5					10					15	
Thr Lys Val	Glu Ası 20	Cys	Arg	Val	Trp 25	Ala	Tyr	G1y	Tyr	Met 30	Asn	Ile
Glu Asp His 35	Ser Glu	ı Tyr	Lys	Ile 40	Gly	Asn	Ser	Leu	Asp 45	Glu	Phe	Met
Ala Trp Val 50	Leu Lys	s Val	Gln 55	Ala	Asp	Leu	Tyr	Phe 60	His	Asn	Leu	Lys

CN 113	1225	517 A	L					序	列	表	<u>=</u>						5/16 页
Р	he	Asp	Gly	Ala	Phe	Ile	Ile	Asn	Trp	Leu	Glu	Arg	Asn	Gly	Phe	Lys	
6	5					70					75					80	
Т	rp	Ser	Ala	Asp	Gly 85	Leu	Pro	Asn	Thr	Tyr 90	Asn	Thr	Ile	Ile	Ser 95	Arg	
M	let	G1y	G1n	Trp 100	Tyr	Met	Ile	Asp	Ile 105	Cys	Leu	Gly	Tyr	Lys 110	Gly	Lys	
A	.rg	Lys	Ile 115	His	Thr	Val	Ile	Tyr 120	Asp	Ser	Leu	Lys	Lys 125	Leu	Pro	Phe	
Р	ro	Val 130	Lys	Lys	Ile	Ala	Lys 135	Asp	Phe	Lys	Leu	Thr 140	Val	Leu	Lys	G1y	
	sp 45	Ile	Asp	Tyr	His	Lys 150	Glu	Arg	Pro	Val	Gly 155	Tyr	Lys	Ile	Thr	Pro 160	
		Glu	Tyr	Ala	Tyr 165		Lys	Asn	Asp	Ile 170		Ile	Ile	Ala	Glu 175		
L	eu	Leu	Ile	Gln 180	Phe	Lys	Gln	Gly	Leu 185	Asp	Arg	Met	Thr	Ala 190	Gly	Ser	
A	sp	Ser	Leu 195	Lys	Gly	Phe	Lys	Asp 200	Ile	Ile	Thr	Thr	Lys 205	Lys	Phe	Lys	
L	ys	Val 210	Phe	Pro	Thr	Leu	Ser 215	Leu	Gly	Leu	Asp	Lys 220	Glu	Va1	Arg	Tyr	
	1a 25	Tyr	Arg	Gly	Gly	Phe 230	Thr	Trp	Leu	Asn	Asp 235	Arg	Phe	Lys	Glu	Lys 240	
G	lu	Ile	Gly	Glu	Gly 245	Met	Val	Phe	Asp	Val 250	Asn	Ser	Leu	Tyr	Pro 255	Ala	
G	ln	Met	Tyr	Ser 260	Arg	Leu	Leu	Pro	Tyr 265	Gly	Glu	Pro	Ile	Val 270	Phe	Glu	
G	1y	Lys	Tyr 275	Val	Trp	Asp	Glu	Asp 280	Tyr	Pro	Leu	His	Ile 285	Gln	His	Ile	
A	rg	Cys 290	Glu	Phe	Glu	Leu	Lys 295	Glu	Gly	Tyr	Ile	Pro 300	Thr	Ile	Gln	Ile	
	ys 05	Arg	Ser	Arg	Phe	Tyr 310	Lys	Gly	Asn	Glu	Tyr 315	Leu	Lys	Ser	Ser	G1y 320	
G	1y	Glu	Ile	Ala	Asp 325	Leu	Trp	Leu	Ser	Asn 330	Val	Asp	Leu	Glu	Leu 335	Met	
L	ys	Glu	His	Tyr 340	Asp	Leu	Tyr	Asn	Val 345	Glu	Tyr	Ile	Ser	Gly 350	Leu	Lys	
Р	he	Lys	Ala 355		Thr	Gly	Leu	Phe 360	Lys	Asp	Phe	Ile	Asp 365	Lys	Trp	Phe	
Т	yr	Ile	Lys	Leu	Thr	Ser	Ser	Gly	Ala	Ile	Lys	Gln	Leu	Ala	Lys	Leu	

	370					375					380				
Met	Leu	Asn	Ser	Leu	Tyr	G1y	Lys	Phe	Ala	Ser	Asn	Pro	Asp	Val	Thr
385					390					395					400
Gly	Lys	Val	Pro	Tyr	Leu	Lys	Glu	Asn	Gly	Ala	Leu	Gly	Phe	Arg	Leu
				405					410					415	
Gly	Glu	Glu	Glu	Thr	Lys	Asp	Pro	Val	Tyr	Thr	Pro	Met	Gly	Val	Phe
			420					425					430		
Ile	Thr		Trp	Ala	Arg	Tyr		Thr	Ile	Thr	Ala		Gln	Ala	Cys
_		435			_	_	440			_		445	_		
Tyr		Arg	Ile	Ile	Tyr		Asp	Thr	Asp	Ser		His	Leu	Thr	Gly
m1	450	T-1	ъ		** 1	455			T 1	** 1	460				
	Glu	He	Pro	Asp		He	Lys	Asp	He		Asp	Pro	Tyr	Lys	
465	т	т	01	11:	470	C	TT1	DI	T	475	۸ 1	т	т	T	480
GIY	1yr	rp	GIU	His	GIU	ser	Inr	Pne		Arg	Ala	Lys	1yr		Arg
C1n	Lvc	Thr	Тик	485 Ile	C1n	Aan	T10	Тик	490 Mot	Lvc	C111	Vo 1	Aan	495	Тиг
GIII	LyS	1111	500	116	GIII	лър	116	505	Met	LyS	GIU	vai	510	СТУ	1 y 1
Ι Δ11	Va1	G111		Ser	Pro	Asn	Asn		Thr	Asn	T16	Ive		Sor	Va1
Lea	vai	515	Oly	501	110	пър	520	1 9 1	1111	пър	110	525	1110	501	vai
Lvs	Cvs		G1 v	Met	Thr	Asp		I1e	Lvs	Lvs	G111		Thr	Phe	G111
_5 _	530		3			535	-3-		-3-	-3-	540				
Asn	Phe	Lys	Val	Gly	Phe	Ser	Arg	Lys	Met	Lys	Pro	Lys	Pro	Val	Gln
545					550					555					560
Val	Pro	G1y	Gly	Val	Val	Leu	Val	Asp	Asp	Thr	Phe	Thr	Ile	Lys	Gly
				565					570					575	
Thr	Gly	Ser	Gly	Ala	Leu	Lys	Thr	Leu	Glu	Ser	Ile	Val	Gly	Asp	Leu
			580					585					590		
Glu	Lys		Asp	Glu	Leu	Lys	_	Lys	Tyr	Gly	Ser		Ser	Ala	Val
		595	_				600			_		605	_		
Arg		Leu	Pro	Val	Glu		Leu	Arg	Glu	Leu		Phe	Ser	Asp	Asp
0.1	610	4 1	0.1	т1		615	т1	Б	.	.	620	4	0.1	4.1	D1
	He	Ala	Glu	Ile		Gly	He	Pro	Lys		Leu	Arg	Glu	Ala	
625	Ι	C1	тъ	۸1.	630	C1	Ι	Т	C1	635	Т	C1	C	Ι	640
Asp	Leu	GIU	Inr	Ala 645	Ala	GIU	Leu	lyr	650	Arg	lyr	GIY	ser	655	Lys
G1 ₁₁	T10	C1 ₃₇	Ara	Arg	Lou	Sor	Тик	Acn		I 011	I 011	G111	I 011		Δ1a
oru	116	оту	660	ni g	ren	PET	тут	665	ush	ren	ren	oru	670	оту	ша
Thr	Pro	Lvs		Ala	Ala	G111	T1e		G1v	Pro	G111	Phe		Phe	Len
1111	110	675	111U	1110	1114	514	680	2,5	013	110	514	685	2,5	1110	104
												233			

Leu	Asn 690	Ile	Glu	Gly	Val	Gly 695	Pro	Lys	Leu	Ala	Glu 700	Arg	Ile	Leu	Glu
Ala	Val	Asp	Tyr	Asp	Leu	Glu	Arg	Leu	Ala	Ser	Leu	Asn	Pro	Glu	Glu
705					710					715					720
Leu	Ala	Glu	Lys	Val	Glu	Gly	Leu	Gly	Glu	Glu	Leu	Ala	Glu	Arg	Val
				725					730					735	
Val	Tyr	Ala	Ala	Arg	Glu	Arg	Val	Glu	Ser	Arg	Arg	Lys	Ser	Gly	Arg
			740					745					750		
Gln	Glu	Arg	Ser	Glu	Glu	Glu	Trp	Lys	Glu	Trp	Leu	Glu	Arg	Lys	Val
		755					760					765			
Gly	Glu	Gly	Arg	Ala	Arg	Arg	Leu	Ile	Glu	Tyr	Phe	Gly	Ser	Ala	G1y
	770					775					780				
Glu	Val	Gly	Lys	Leu	Val	Glu	Asn	Ala	Glu	Val	Ser	Lys	Leu	Leu	G1u
785					790					795					800
Val	Pro	Gly	Ile		Asp	Glu	Ala	Val	Ala	Arg	Leu	Val	Pro	G1y	Tyr
				805					810					815	
Lys	Thr	Leu		Asp	Ala	Gly	Leu		Pro	Ala	Glu	Ala		Arg	Val
			820		_		_	825					830		
Leu	Lys		Tyr	Gly	Ser	Val		Lys	Val	Gln	Glu		Ala	Thr	Pro
		835					840				_	845			
Asp	Glu	Leu	Arg	Glu	Leu		Leu	Gly	Asp	Ala		Ile	Ala	Arg	Ile
_	850	_		~	_	855		_		_	860			m1	
	Gly	Leu	Arg	Ser		Val	Asn	Lys	Arg		Asp	Val	Asp	Thr	
865	0.1				870	m.	0.1	0	17 1	875	4.1	77 1		.	880
Tyr	Glu	Leu	Lys		Arg	Tyr	Gly	Ser		Ser	Ala	Val	Arg		Ala
D	17 1	T	01	885	٨	01	T	01	890	C	Λ	Λ	T	895	۸ 1
Pro	Val	Lys		Leu	Arg	GIU	Leu	-	Leu	ser	Asp	Arg	-	11e	Ala
Λ	т1.	T	900	т1.	D	C1	т1	905	Ι	C1	W - 1	۸	910	M - 4	C
Arg	Ile		GIY	11e	Pro	GIU		мет	Leu	GIN	vai		GIY	мет	ser
Vo 1	C1	915	110	C1,,	Amor	Lou	920	C1,,	120	Dho	Aan	925	Tan	Tha	L
vai	Glu 930	LyS	Ala	GIU	Arg	935	Leu	GIU	Arg	rne	940	1111	rrp	1111	LyS
Vo 1	Lys	C1,,	11 ₀	Dno	Vo 1		C1,,	Lou	Vo 1	Ana		Dno	C1,,	Vo 1	C1 ₁₇
945	LyS	GIU	пта	110	950	ser	GIU	Leu	Val	955	Val	110	GIY	Val	960
	Ser	Ι Δ11	Va1	Lvc		T1 ₀	Lvc	Δ1a	G1n		Acn	Pro	Δ1a	Trn	
Leu	DEI	Leu	vai	965	oru	116	Lys	пια	970	vai	пър	110	пια	975	Lys
Ala	Leu	I 611	Asn		Lve	G1v	Va1	Ser		G111	<u>ا م</u> ا ا	Ala	Asn		Len
ıııa	LCU	LCU	980	, a 1	шуз	оту	, a 1	985	110	oru	LCu	111a	990	111 8	LCu
Va1	Glu	G111		G1 v	Ser	Pro	Tvr		Va1	[_. e ₁₁	Thr	Ala		Lvs	Ser
, aı	O L U	O L U	Lou	O I y	501	0	- J -	8	, а т	Lou			- 1 5	_, 5	501

Asp Leu Met Arg Val Glu Arg Val Gly Pro Lys Leu Ala Glu Arg Ile Arg Ala Ala Gly <210> 3 <211> 1028 <212> PRT 〈213〉人工序列 <400> 3 Met Lys His Met Pro Arg Lys Met Tyr Ser Cys Asp Phe Glu Thr Thr Thr Lys Val Glu Asp Cys Arg Val Trp Ala Tyr Gly Tyr Met Asn Ile Glu Asp His Ser Glu Tyr Lys Ile Gly Asn Ser Leu Asp Glu Phe Met Ala Trp Val Leu Lys Val Gln Ala Asp Leu Tyr Phe His Asn Leu Lys Phe Asp Gly Ala Phe Ile Ile Asn Trp Leu Glu Arg Asn Gly Phe Lys Trp Ser Ala Asp Gly Leu Pro Asn Thr Tyr Asn Thr Ile Ile Ser Arg Met Gly Gln Trp Tyr Met Ile Asp Ile Cys Leu Gly Tyr Lys Gly Lys Arg Lys Ile His Thr Val Ile Tyr Asp Ser Leu Lys Lys Leu Pro Phe Pro Val Lys Lys Ile Ala Lys Asp Phe Lys Leu Thr Val Leu Lys Gly Asp Ile Asp Tyr His Lys Glu Arg Pro Val Gly Tyr Lys Ile Thr Pro Glu Glu Tyr Ala Tyr Ile Lys Asn Asp Ile Gln Ile Ile Ala Glu Ala Leu Leu Ile Gln Phe Lys Gln Gly Leu Asp Arg Met Thr Ala Gly Ser Asp Ser Leu Lys Gly Phe Lys Asp Ile Ile Thr Thr Lys Lys Phe Lys Lys Val Phe Pro Thr Leu Ser Leu Gly Leu Asp Lys Glu Val Arg Tyr Ala Tyr Arg Gly Gly Phe Thr Trp Leu Asn Asp Arg Phe Lys Glu Lys

225					230					235					240
Glu	Ile	G1y	Glu	G1y 245	Met	Val	Phe	Asp	Val 250	Asn	Ser	Leu	Tyr	Pro 255	Ala
G1n	Met	Tyr			Leu	Leu	Pro		Gly	G1u	Pro	Ile			Glu
	_		260	_				265	_	_			270		
Gly	Lys	Tyr 275	Val	Trp	Asp	Glu	Asp 280	Tyr	Pro	Leu	His	11e 285	Gln	His	He
Arg	Cys 290	Glu	Phe	Glu	Leu	Lys 295	Glu	Gly	Tyr	Ile	Pro 300	Thr	Ile	G1n	Ile
Lvs		Ser	Arg	Phe	Tvr		G1v	Asn	Glu	Tvr		Lvs	Ser	Ser	G1v
305	0	202	0		310	_, _	02)		0 = 0	315		_, _	202	202	320
G1y	Glu	Ile	Ala	Asp 325	Leu	Trp	Leu	Ser	Asn 330	Val	Asp	Leu	G1u	Leu 335	Met
Lvs	Glu	His	Tvr		Leu	Tvr	Asn	Val	Glu	Tvr	Ile	Ser	G1v		Lvs
, i			340	-				345					350		
Phe	Lys	Ala	Thr	Thr	Gly	Leu	Phe	Lys	Asp	Phe	Ile	Asp	Lys	Trp	Phe
		355					360					365			
Tyr	Ile	Lys	Leu	Thr	Ser	Lys	Gly	Ala	Ile	Lys	G1n	Leu	Ala	Lys	Leu
	370					375					380				
Met	Leu	Asn	Ser	Leu	Tyr	Gly	Lys	Phe	Ala	Ser	Asn	Pro	Asp	Val	Thr
385					390					395					400
Gly	Lys	Val	Pro	Tyr	Leu	Lys	Glu	Asn	Gly	Ala	Leu	G1y	Phe	Arg	Leu
				405					410					415	
Gly	Glu	Glu	Glu	Thr	Lys	Asp	Pro		Tyr	Thr	Pro	Met	Gly	Val	Phe
			420					425					430		
Ile	Thr		Trp	Ala	Arg	Tyr		Thr	Ile	Thr	Ala		G1n	Ala	Cys
_		435				~	440	m.t		~		445	_		0.1
Tyr		Arg	He	He	Tyr		Asp	Thr	Asp	Ser		His	Leu	Thr	Gly
TT1	450	т 1	D	٨	17 1	455	ī	Λ	т1	17 1	460	D	т	т	T
	Glu	He	Pro	Asp		He	Lys	Asp	Ile		Asp	Pro	lyr	Lys	
465	т	т	01	11:	470	C	TT1	DI	т	475	A 1	т	т	T	480
Gly	lyr	Irp	Glu	H1S 485	Glu	Ser	Inr	Phe	Lys 490	Arg	Ala	Lys	lyr	Leu 495	Arg
G1n	Lys	Thr	Tyr 500	Ile	G1n	Asp	Ile	Tyr 505	Met	Lys	Glu	Val	Asp 510	G1y	Tyr
Leu	Val	Glu	G1y	Ser	Pro	Asp	Asp	Tyr	Thr	Asp	Ile	Lys	Phe	Ser	Val
		515					520					525			
Lys	Cys	Ala	G1y	Met	Thr	Asp	Lys	Ile	Lys	Lys	Glu	Val	Thr	Phe	Glu
	530					535					540				

CN 113122	517 A	L					序	列	表	<u>=</u>						10/16 页
Asn	Phe	Lys	Va1	Gly	Phe	Ser	Arg	Lys	Met	Lys	Pro	Lys	Pro	Va1	Gln	
545					550					555					560	
Val	Pro	Gly	Gly	Va1	Val	Leu	Val	Asp	Asp	Thr	Phe	Thr	Ile	Lys	Gly	
				565					570					575		
Thr	Gly	Ser	Gly 580	Ala	Leu	Lys	Thr	Leu 585	Glu	Ser	Ile	Val	Gly 590	Asp	Leu	
Glu	Lys	Ala		Glu	Leu	Lvs	Arg		Tvr	Glv	Ser	Ala		Ala	Val	
	J	595	1			5	600	J	J	J		605				
Arg	Arg	Leu	Pro	Va1	Glu	Glu	Leu	Arg	Glu	Leu	Gly	Phe	Ser	Asp	Asp	
	610					615					620					
Glu	Ile	Ala	Glu	Ile	Lys	Gly	Ile	Pro	Lys	Lys	Leu	Arg	Glu	Ala	Phe	
625					630					635	_		_	_	640	
Asp	Leu	Glu	Thr		Ala	Glu	Leu	Tyr		Arg	Tyr	Gly	Ser		Lys	
C1	T1.	C1	A 72 CT	645	Lou	Con	Т	Aan	650	Lou	Lou	C1.,	Lou	655	11 ₀	
GIU	Ile	GIY	660	Arg	Leu	ser	lyr	665	ASP	Leu	Leu	GIU	670	GIY	Ala	
Thr	Pro	Lys 675	Ala	Ala	Ala	Glu	Ile 680	Lys	Gly	Pro	Glu	Phe 685	Lys	Phe	Leu	
Leu	Asn 690	Ile	Glu	Gly	Val	Gly 695	Pro	Lys	Leu	Ala	Glu 700	Arg	Ile	Leu	Glu	
Ala	Val	Asp	Tyr	Asp	Leu		Arg	Leu	Ala	Ser		Asn	Pro	Glu	Glu	
705		_	-	_	710					715					720	
Leu	Ala	Glu	Lys	Val 725	Glu	Gly	Leu	Gly	G1u 730	Glu	Leu	Ala	Glu	Arg 735	Val	
Va1	Tyr	Ala	Ala		Glu	Arg	Va1	Glu		Arg	Arg	Lvs	Ser		Arg	
	J		740	Ü		Ü		745		Ü	Ü	J	750	J	J	
Gln	Glu	Arg	Ser	Glu	Glu	Glu	Trp	Lys	Glu	Trp	Leu	Glu	Arg	Lys	Val	
		755					760					765				
Gly	Glu 770	Gly	Arg	Ala	Arg	Arg 775	Leu	Ile	Glu	Tyr	Phe 780	Gly	Ser	Ala	Gly	
Glu	Va1	Gly	Lys	Leu	Va1	Glu	Asn	Ala	Glu	Val	Ser	Lys	Leu	Leu	Glu	
785					790					795					800	
Val	Pro	Gly	Ile		Asp	Glu	Ala	Val		Arg	Leu	Val	Pro		Tyr	
Ī	Т1	Ι	۸	805	۸1.	C1	Τ	ть	810	۸1.	C1	۸1.	C1	815	W - 1	
Lys	Thr	Leu	Arg 820	Asp	Ala	Gly	Leu	1nr 825	Pro	Ala	GIU	Ala	830	Arg	vaı	
Leu	Lys	Arg 835	Tyr	Gly	Ser	Val	Ser 840	Lys	Val	Gln	Glu	Gly 845	Ala	Thr	Pro	
Asp	Glu		Arg	Glu	Leu	Gly		G1y	Asp	Ala	Lys		Ala	Arg	Ile	

850			855					860				
Leu Gly Leu	Arg Ser	Leu	Val	Asn	Lys	Arg	Leu	Asp	Val	Asp	Thr	Ala
865		870					875					880
Tyr Glu Leu	Lys Arg	Arg	Tyr	Gly	Ser	Val	Ser	Ala	Val	Arg	Lys	Ala
	885)				890					895	
Pro Val Lys		Arg	Glu	Leu	Gly	Leu	Ser	Asp	Arg	Lys	Ile	Ala
	900				905					910		
Arg Ile Lys	Gly Ile	Pro	Glu		Met	Leu	Gln	Val		Gly	Met	Ser
915	.1 01			920	0.1		D1		925	m	mı.	
Val Glu Lys	Ala Glu	Arg		Leu	Glu	Arg	Phe		Thr	Trp	Thr	Lys
930	41 D	37 1	935	0.1	т	17 1	Δ.	940	D	0.1	77 1	01
Val Lys Glu	Ala Pro		Ser	Glu	Leu	Val	_	Val	Pro	Gly	Val	-
945	Vol I v	950	T1.	I *** a	110	C1n	955 Vo.1	Aan	Dao	110	Тип	960
Leu Ser Leu	var Lys 965		11e	Lys	Ala	970	vai	ASP	Pro	Ala	975	Lys
Ala Leu Leu			G1 _v	Va1	Sar		G111	I 011	Δ1a	Acn		I 011
Mia Led Led	980	Lys	Uly	vai	985	110	Ulu	Leu	пια	990	mg	Leu
Val Glu Glu		Ser	Pro	Tvr		Va1	Len	Thr	A1a		Lvs	Ser
995	Lea dij	501	110	1000		rai	Doa	1111	1005		Ljo	501
Asp Leu Met	Ara Vol	C1,,	Ara			Dno	Luc	Lou			Δνα	
MSD LCU MCC	лід уаі	. utu	$\Lambda I \aleph$	vaı	$\Omega T \lambda$	L L O	LVS	Leu	пта	arange	$\Pi \times$	
1010	nig vai	GIU	1015		Gly	110	LyS	1020		Glu	лгg	
					Gly	110	LyS			Giu	ni g	
1010					Gly	110	LyS			GIU	ΝΊĞ	
1010 Ile Arg Ala					Gly	710	LyS			Giù	MIG	
1010 Ile Arg Ala 1025					Gly	710	LyS			Giù	AIG	
1010 Ile Arg Ala 1025 <210> 4					Gly	110	LyS			Giu	AIG	
1010 Ile Arg Ala 1025 <210> 4 <211> 447	Ala Gly				Gly	110	LyS			Giu	AIG	
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT	Ala Gly				Gly	710	Lys			Giu	AIG	
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序	Ala Gly ^等 列		1015	5				1020)			Glu
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4	Ala Gly ^等 列		1015	5				1020)			Glu
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4 Leu Lys Thr	Ala Gly S列 Leu Glu 5 Lys Tyr	Ser	1018	Val	Gly Ser	Asp 10	Leu	1020 Glu) Lys	Ala Leu	Asp 15	
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4 Leu Lys Thr 1 Leu Lys Arg	Ala Gly F列 Leu Glu 5 Lys Tyr 20	Ser Gly	Ile Ser	Val Ala	Gly Ser 25	Asp 10 Ala	Leu Val	1020 Glu Arg	Lys Arg	Ala Leu 30	Asp 15 Pro	Val
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4 Leu Lys Thr 1 Leu Lys Arg	Ala Gly F列 Leu Glu 5 Lys Tyr 20	Ser Gly	Ile Ser	Val Ala Phe	Gly Ser 25	Asp 10 Ala	Leu Val	1020 Glu Arg	Lys Arg Ile	Ala Leu 30	Asp 15 Pro	Val
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4 Leu Lys Thr 1 Leu Lys Arg Glu Glu Leu 35	Ala Gly Eeu Glu 5 Lys Tyr 20 Arg Glu	Ser Gly Leu	Ile Ser Gly	Val Ala Phe 40	Gly Ser 25 Ser	Asp 10 Ala Asp	Leu Val Asp	Glu Glu Glu	Lys Arg Ile 45	Ala Leu 30 Ala	Asp 15 Pro Glu	Val Ile
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4 Leu Lys Thr 1 Leu Lys Arg Glu Glu Leu 35 Lys Gly Ile	Ala Gly Eeu Glu 5 Lys Tyr 20 Arg Glu	Ser Gly Leu	Ile Ser Gly Leu	Val Ala Phe 40	Gly Ser 25 Ser	Asp 10 Ala Asp	Leu Val Asp	Glu Arg Glu Asp	Lys Arg Ile 45	Ala Leu 30 Ala	Asp 15 Pro Glu	Val Ile
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4 Leu Lys Thr 1 Leu Lys Arg Glu Glu Leu 35 Lys Gly Ile 50	Ala Gly Leu Glu 5 Lys Tyr 20 Arg Glu Pro Lys	Ser Gly Leu	Ile Ser Gly Leu 55	Val Ala Phe 40 Arg	Gly Ser 25 Ser Glu	Asp 10 Ala Asp	Leu Val Asp Phe	Glu Arg Glu Asp 60	Lys Arg Ile 45 Leu	Ala Leu 30 Ala Glu	Asp 15 Pro Glu Thr	Val Ile Ala
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4 Leu Lys Thr 1 Leu Lys Arg Glu Glu Leu 35 Lys Gly Ile 50 Ala Glu Leu	Ala Gly Leu Glu 5 Lys Tyr 20 Arg Glu Pro Lys	Ser Gly Leu Lys	Ile Ser Gly Leu 55	Val Ala Phe 40 Arg	Gly Ser 25 Ser Glu	Asp 10 Ala Asp	Leu Val Asp Phe Lys	Glu Arg Glu Asp 60	Lys Arg Ile 45 Leu	Ala Leu 30 Ala Glu	Asp 15 Pro Glu Thr	Val Ile Ala Arg
1010 Ile Arg Ala 1025 <210> 4 <211> 447 <212> PRT <213> 人工序 <400> 4 Leu Lys Thr 1 Leu Lys Arg Glu Glu Leu 35 Lys Gly Ile 50	Ala Gly Folia Folia Arg Glu Pro Lys Tyr Glu	Ser Gly Leu Lys Arg 70	Ile Ser Gly Leu 55 Tyr	Val Ala Phe 40 Arg	Gly Ser 25 Ser Glu Ser	Asp 10 Ala Asp Ala Leu	Leu Val Asp Phe Lys 75	Glu Arg Glu Asp 60 Glu	Lys Arg Ile 45 Leu Ile	Ala Leu 30 Ala Glu Gly	Asp 15 Pro Glu Thr	Val Ile Ala Arg 80

				85					90					95	
Ala	Glu	Ile	Lys	Gly	Pro	Glu	Phe	Lys	Phe	Leu	Leu	Asn	Ile	Glu	Gly
			100					105					110		
Val	Gly	Pro	Lys	Leu	Ala	Glu	Arg	Ile	Leu	Glu	Ala	Val	Asp	Tyr	Asp
		115					120					125			
Leu	Glu	Arg	Leu	Ala	Ser	Leu	Asn	Pro	Glu	Glu	Leu	Ala	Glu	Lys	Val
	130					135					140				
Glu	Gly	Leu	Gly	Glu	Glu	Leu	Ala	Glu	Arg	Val	Val	Tyr	Ala	Ala	Arg
145					150					155					160
Glu	Arg	Val	Glu	Ser	Arg	Arg	Lys	Ser	Gly	Arg	G1n	Glu	Arg	Ser	Glu
				165					170					175	
Glu	Glu	Trp		Glu	Trp	Leu	Glu		Lys	Val	Gly	Glu		Arg	Ala
		_	180	0.1		D .	0.1	185		0.1	0.1		190	_	_
Arg	Arg		He	Glu	Tyr	Phe		Ser	Ala	Gly	Glu		Gly	Lys	Leu
17 1	0.1	195	4.1	0.1	37 1	0	200	T	T	0.1	1 7 1	205 D	0.1	т1	0.1
Val	Glu	Asn	Ala	Glu	Val		Lys	Leu	Leu	Glu		Pro	Gly	He	Gly
٨	210	A 1	17 1	۸.1	٨	215	17 1	D	01	T	220	TT1	T	Λ	Λ
	Glu	Ala	vai	Ala		Leu	vai	Pro	GIY		Lys	Inr	Leu	Arg	
225	C1,,,	Lou	Thm	Dno	230	C1,,	11a	C1,,	Ana	235	Lou	Lvo	Ana	Тиги	240
на	Gly	Leu	1111	245	пта	GIU	на	GIU	250	vai	Leu	LyS	AIg	255	GIY
Sor	Val	Sor	Lvc		G1n	G111	G1 _v	Δ1a		Pro	Acn	G111	I 011		G111
261	vai	261	260	vai	OIII	Ulu	Uly	265	1111	110	пър	Ulu	270	nı g	UIU
Len	G1y	Len		Asn	A1a	Lvs	T1e		Arg	T1e	Len	G1v		Arg	Ser
200	01)	275	01)	пор	1110	2,5	280	1110	111 0	110	200	285	200		501
Leu	Val		Lvs	Arg	Leu	Asp		Asp	Thr	Ala	Tvr		Leu	Lvs	Arg
	290		J	Ü		295					300			J	O
Arg	Tyr	G1y	Ser	Va1	Ser	Ala	Val	Arg	Lys	Ala	Pro	Val	Lys	Glu	Leu
305					310					315					320
Arg	Glu	Leu	Gly	Leu	Ser	Asp	Arg	Lys	Ile	Ala	Arg	Ile	Lys	G1y	Ile
				325					330					335	
Pro	Glu	Thr	Met	Leu	Gln	Val	Arg	Gly	Met	Ser	Val	Glu	Lys	Ala	Glu
			340					345					350		
Arg	Leu	Leu	Glu	Arg	Phe	Asp	Thr	Trp	Thr	Lys	Val	Lys	Glu	Ala	Pro
		355					360					365			
Val	Ser	Glu	Leu	Val	Arg	Val	Pro	Gly	Val	Gly	Leu	Ser	Leu	Val	Lys
	370					375					380				
Glu	Ile	Lys	Ala	Gln	Val	Asp	Pro	Ala	Trp	Lys	Ala	Leu	Leu	Asp	Val
385					390					395					400

<211> 3093

<212> DNA

<213> 人工序列

<400> 5

atgccgagaa agatgtatag ttgtgacttt gagacaacta ctaaagtgga agactgtagg 60 gtatgggcgt atggttatat gaatatagaa gatcacagtg agtacaaaat aggtaatagc 120 ctggatgagt ttatggcgtg ggtgttgaag gtacaagctg atctatattt ccataacctc 180 aaatttgacg gagcttttat cattaactgg ttggaacgta atggttttaa gtggtcggct 240 gacggattgc caaacacata taatacgatc atatctcgca tgggacaatg gtacatgatt 300 gatatatgtt taggctacaa agggaaacgt aagatacata cagtgatata tgacagctta 360 aagaaactac cgtttcctgt taagaagata gctaaagact ttaaactaac tgttcttaaa 420 ggtgatattg attaccacaa agaaagacca gtcggctata agataacacc cgaagaatac 480 gcctatatta aaaacgatat tcagattatt gcggaagctc tgttaattca gtttaagcaa 540 ggtttagacc ggatgacagc aggcagtgac agtctaaaag gtttcaagga tattataacc 600 actaagaaat tcaaaaaggt gtttcctaca ttgagtcttg gactcgataa ggaagtgaga 660 tacgcctata gaggtggttt tacatggtta aatgataggt tcaaagaaaa agaaatcgga 720 gaaggcatgg tettegatgt taatagteta tateetgeac agatgtatag cegteteett 780 ccatatggtg aacctatagt attcgagggt aaatacgttt gggacgaaga ttacccacta 840 cacatacagc atatcagatg tgagttcgaa ttgaaagagg gctatatacc cactatacag 900 ataaaaagaa gtaggtttta taaaggtaat gagtacctaa aaagtagcgg cggggagata 960 gccgacctct ggttgtcaaa tgtagaccta gaattaatga aagaacacta cgatttatat 1020 aacgttgaat atatcagcgg cttaaaattt aaagcaacta caggtttgtt taaagatttt 1080 atagataaat ggttttacat caagttgaca tcaagcggag cgatcaagca actagcaaaa 1140 ctgatgttaa acagtctata cggtaaattc gctagtaacc ctgatgttac agggaaagtc 1200 ccttatttaa aagagaatgg ggcgctaggt ttcagacttg gagaagagga aacaaaagac 1260 cctgtttata cacctatggg cgttttcatc actgcatggg ctagatacac gacaattaca 1320 gcggcacagg cttgttatga tcggataata tactgtgata ctgacagcat acatttaacg 1380 ggtacagaga tacctgatgt aataaaagat atagttgacc cttacaaatt gggatactgg 1440 gaacatgaaa gtacattcaa aagagctaaa tatctgagac agaagaccta tatacaagac 1500 atctatatga aagaagtaga tggttattta gtagaaggta gtccagatga ttacactgat 1560 ataaaattta gtgttaaatg tgcgggaatg actgacaaga ttaagaaaga ggttacgttt 1620 gagaatttca aagtcggatt cagtcggaaa atgaagccta agcctgtgca agtgccgggc 1680

```
ggggtggttc tggttgatga cacattcaca atcaaaggta ccggctctgg cgccctgaag 1740
accetggaga geatagtagg ggatetggag aaggeegaeg agetgaageg gaagtaegga 1800
teegegteeg eggttegaeg tetgeeegta gaggagetae gegaaetegg gtteteegae 1860
gatgagatcg ccgagatcaa ggggatacct aagaagctcc gggaggcctt cgaccttgag 1920
accgccgcgg aactctacga gcggtacggt tcgctgaaag agatcggtcg ccgactctct 1980
tacgacgate tactegaget eggtgegaet eegaaggeeg eggeegagat eaaggggeeg 2040
gagttcaagt teeteetgaa categaaggg gteggaeega aactegetga geggataete 2100
gaggccgtgg attatgacct cgagcgactg gcttccctga atcccgagga acttgcggag 2160
aaggtggaag gactgggcga agagctcgcg gagcgcgtcg tgtacgctgc tagggagcgc 2220
gtagaaagtc gcaggaagtc cggccgccag gagcggtcgg aggaagaatg gaaggagtgg 2280
ctcgagcgta aggtcggcga ggggagggct cgccggttga ttgagtattt cggctccgcg 2340
ggtgaagtag gaaagctggt cgagaacgcc gaggtgtcga agctactgga ggtcccgggt 2400
ataggcgacg aggccgtcgc taggctcgta ccgggctaca agaccctacg agacgccggt 2460
ctcacgccgg ccgaagcgga gcgcgtgctg aaacggtacg gctcggtctc caaagtgcag 2520
gaaggagcca ctccggacga gttacgcgag ctcggcctcg gcgacgccaa gatcgcgagg 2580
atcctgggcc tgcgcagcct ggtgaacaag aggctggacg tggacaccgc gtacgagctc 2640
aagcgtagat acggttccgt ctccgccgtc cggaaggccc cggtgaaaga actgcgcgag 2700
ctcggcctct ccgatcggaa gatcgcacgt atcaagggca tcccggagac gatgcttcag 2760
gtccgaggga tgagcgtgga gaaagcggag cggctgctgg agcgtttcga tacctggacc 2820
aaggtgaagg aagctcccgt ctcggagctg gtgagagtcc cgggtgtcgg attgagtttg 2880
gtgaaggaga tcaaggctca ggtggatccg gcctggaagg cacttctgga tgtcaaaggg 2940
gtcagtccgg agctggccga ccggctcgtc gaggagctcg gcagcccgta tcgggtgctg 3000
acggccaaga aatccgacct gatgagagtc gagagagtcg gaccgaagct cgccgagcga 3060
atccgggccg cgggctaagc ggccgcatcg tga 3093
<210> 6
<211> 3093
<212> DNA
〈213〉人工序列
<400> 6
atgccgagaa agatgtatag ttgtgacttt gagacaacta ctaaagtgga agactgtagg 60
gtatgggcgt atggttatat gaatatagaa gatcacagtg agtacaaaat aggtaatagc 120
ctggatgagt ttatggcgtg ggtgttgaag gtacaagctg atctatattt ccataacctc 180
aaatttgacg gagcttttat cattaactgg ttggaacgta atggttttaa gtggtcggct 240
gacggattgc caaacacata taatacgatc atatctcgca tgggacaatg gtacatgatt 300
gatatatgtt taggctacaa agggaaacgt aagatacata cagtgatata tgacagctta 360
aagaaactac cgtttcctgt taagaagata gctaaagact ttaaactaac tgttcttaaa 420
ggtgatattg attaccacaa agaaagacca gtcggctata agataacacc cgaagaatac 480
gcctatatta aaaacgatat tcagattatt gcggaagctc tgttaattca gtttaagcaa 540
ggtttagacc ggatgacagc aggcagtgac agtctaaaag gtttcaagga tattataacc 600
```

tcaaaaaggt	gtttcctaca	ttgagtcttg	gactcgataa	ggaagtgaga	660
gaggtggttt	tacatggtta	$a at {\tt gat aggt}$	tcaaagaaaa	agaaatcgga	720
tcttcgatgt	taatagtcta	tatcctgcac	agatgtatag	${\tt ccgtctcctt}$	780
aacctatagt	${\tt attcgagggt}$	aaatacgttt	gggacgaaga	ttacccacta	840
atatcagatg	tgagttcgaa	ttgaaagagg	gctatatacc	cactatacag	900
gtaggtttta	taaaggtaat	gagtacctaa	aaagtagcgg	cggggagata	960
ggttgtcaaa	tgtagaccta	gaattaatga	aagaacacta	cgatttatat	1020
atatcagcgg	$\tt cttaaaattt$	aaagcaacta	${\tt caggtttgtt}$	taaagatttt	1080
ggttttacat	caagttgaca	tcaaaaggag	cgatcaagca	actagcaaaa	1140
acagtctata	cggtaaattc	gctagtaacc	${\tt ctgatgttac}$	agggaaagtc	1200
aagagaatgg	ggcgctaggt	ttcagacttg	gagaagagga	aacaaaagac	1260
cacctatggg	${\tt cgttttcatc}$	actgcatggg	ctagatacac	gacaattaca	1320
cttgttatga	tcggataata	tactgtgata	ctgacagcat	a catttaacg	1380
tacctgatgt	aataaaagat	atagttgacc	$\verb cttacaaatt \\$	gggatactgg	1440
gtacattcaa	aagagctaaa	tatctgagac	agaagaccta	tatacaagac	1500
aagaagtaga	tggttattta	gtagaaggta	gtccagatga	ttacactgat	1560
gtgttaaatg	tgcgggaatg	actgacaaga	ttaagaaaga	ggttacgttt	1620
aagtcggatt	cagtcggaaa	atgaagccta	agcctgtgca	agtgccgggc	1680
tggttgatga	cacattcaca	atcaaaggta	ccggctctgg	cgccctgaag	1740
gcatagtagg	ggatctggag	aaggccgacg	agctgaagcg	gaagtacgga	1800
cggttcgacg	tctgcccgta	gaggagctac	gcgaactcgg	gttctccgac	1860
ccgagatcaa	ggggatacct	aagaagctcc	gggaggcctt	${\tt cgaccttgag}$	1920
aactctacga	gcggtacggt	tcgctgaaag	agatcggtcg	ccgactctct	1980
tactcgagct	cggtgcgact	ccgaaggccg	cggccgagat	caaggggccg	2040
tcctcctgaa	catcgaaggg	gtcggaccga	aactcgctga	gcggatactc	2100
attatgacct	cgagcgactg	gcttccctga	atcccgagga	acttgcggag	2160
gactgggcga	agagctcgcg	gagcgcgtcg	tgtacgctgc	tagggagcgc	2220
gcaggaagtc	cggccgccag	gagcggtcgg	aggaagaatg	gaaggagtgg	2280
aggtcggcga	${\tt ggggagggct}$	${\tt cgccggttga}$	${\tt ttgagtattt}$	${\tt cggctccgcg}$	2340
gaaagctggt	cgagaacgcc	gaggtgtcga	agctactgga	ggtcccgggt	2400
aggccgtcgc	taggctcgta	ccgggctaca	agaccctacg	agacgccggt	2460
ccgaagcgga	gcgcgtgctg	aaacggtacg	gctcggtctc	${\tt caaagtgcag}$	2520
ctccggacga	gttacgcgag	ctcggcctcg	gcgacgccaa	gatcgcgagg	2580
tgcgcagcct	ggtgaacaag	aggctggacg	tggacaccgc	gtacgagctc	2640
acggttccgt	$\verb ctccgccgtc $	cggaaggccc	cggtgaaaga	actgcgcgag	2700
ccgatcggaa	gatcgcacgt	atcaagggca	tcccggagac	gatgcttcag	2760
tgagcgtgga	gaaagcggag	cggctgctgg	agcgtttcga	tacctggacc	2820
aagctcccgt	ctcggagctg	gtgagagtcc	cgggtgtcgg	attgagtttg	2880
tcaaggctca	ggtggatccg	gcctggaagg	cacttctgga	tgtcaaaggg	2940
	gaggtggttt tcttcgatgt aacctatagt atatcagatg gtaggtttta ggttgtcaaa atatcagcgg ggttttacat aagagaatgg cacctatggg cacctatggg cactatgtg gtacattcaa aagaagtaga gtgttaaatg aagtcgatt tggtgatga gcatagtagg ccgattcgacg ccgagatcaa aactctacga tactcaga tagacgtga aagctccct acgatccga aagctccct	gaggtggttt tacategtta tcttcgatgt taatagtcta aacctatagt attcgagggt atatcagatg tgagttcgaa gtaggtttta taaaggtaat ggttgtcaaa tgtagaccta atatcagcgg cttaaaatt ggttttacat caagttgaca acagtctata caggtaaatc aagagaatgg ggtaaatt cacctatggg cgttttcatc cttgttatga tcggataata tacctgatgt aataaaagat gtacattcaa aagagctaaa aagagatga tggttattta gtacattcaa aagagctaaa aagtcggatt cagtcggaatg aagttgatag cacattcaca gcatagtaga gggtacggaa cggttcgacg tctgccgta ccgagatcag gcggtacgg tactcagagc cggtgcgcact tactcagagc cggtgcgcact gactgggaaggc cggtgcgcact gactgggaaggc cggcgcacc gacggaagc cggcgcacg gaagctgga ggggaacag g	gaggtggttt tacatggtta aatgataggt tettegatgt taatagteta tateetgeae aacetatagt gagttegaa ttgaaagagg gtaggttta taaaggtaat gagtaectaa ggttgteaa tgaagecta gaattaatga atateagegg ettagaea tgaagecta gagtttaea eaggttgaea teaagggaggttta eaaggtagea teaaggagg gegttaea gggttgaea eaggtaecta eaggttgaea eaggtegae eaggtegae eaggtegae eaggtegae eaggegae eaggega	gaggtggttttacatggttaaatgataggttcaaagaaaatcttcgatgttaatagtctatatcctgcacagatgtatagaactatagtattcgagggtaaatacgtttggaccgaagaatatcagatgtgagttcgaattgaaagagggctatataccgtaggttttataaaggtaatgagtacctaaaaagaacactaatatcagcggcttaaaatttaaagcaactacaggtttgtgttttacatcaagttgacatcaaaaggagcgatcaagcaacagtctatacggtaaattcgctagtaccctgaagaaggacactatgggcgttttcatcactgatgggctgaatacaccatgttatgatccaaatgggctgaatacacctgaagaaggacacttatggcgttttcatcactgatgggctgaatacaccttgttatgatccgataatatactggatgactgacaaggacacttatggcgttttcatcactgaatggctgacaaggacacttatggcgttttattatactggaaggactcaaattgtacattcaaaagaggataaatactgaaggaagaagacctaaagaagtagatggttattatgtagaaggtaagcctgtggagtgttaaatgtgcgggaatgactgaaaggaagcctgtggagtgttgatgacacattcacaatcaaaggtaagcctgtggagcatagtagggagttcggaaaaggecgacaagctgaaggcggttcgacgtctgcccgtagaggagctcggggaggcttggcggatcaggcggtgacgggggaggctggggaggctgccgaggatcacggtggaagtcggaaggccgaactcgagggcgatcggaagtccggagaggcggggagggggggagagggggcatgggaagtccggaggggggggagggggggagagggggca	aagaagtaga tggttattta gtagaaggta gtccagatga ttacactgat gtgttaaatg tgcgggaatg actgacaaga ttaagaaaga ggttacgttt aagtcggatt cagtcggaaa atgaagccta agcctgtgca agtgccgggc tggttgatga cacattcaca atcaaaggta ccggctctgg cgccctgaag gcatagtagg ggatctggag aaggccgacg agctgaagcg gaagtacgga

gtcagtccgg agctggccga ccggctcgtc gaggagctcg gcagcccgta tcgggtgctg 3000 acggccaaga aatccgacct gatgagagtc gagagagtcg gaccgaagct cgccgagcga 3060 atccgggccg cgggctaagc ggccgcatcg tga 3093

<210> 7

<211> 19

<212> DNA

〈213〉人工序列

⟨400⟩ 7

aaaggagcga tcaagcaac 19

<210> 8

<211> 22

<212> DNA

〈213〉人工序列

⟨400⟩ 8

tgatgtcaac ttgatgtaaa ac 22

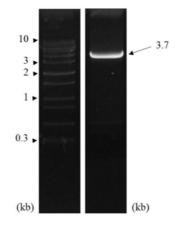


图1

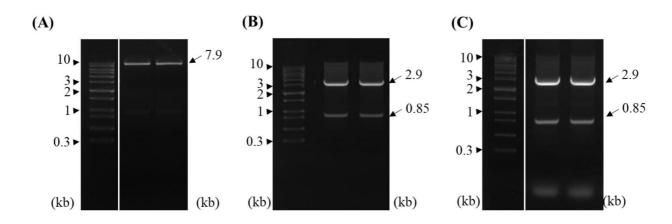


图2

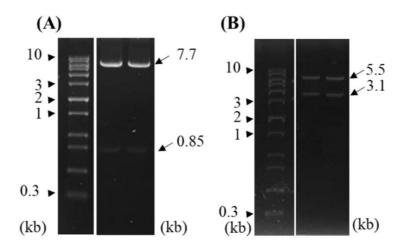


图3

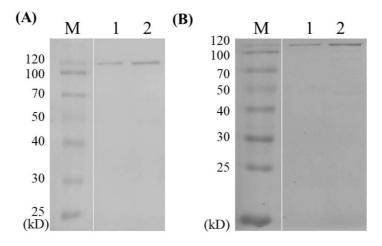


图4

(A) CP12-dA6P: Coumarin-PEG12-dA6P

(B) CP24-dC6P: Coumarin-PEG24-dC6P

(C) CP36-dT6P: Coumarin-PEG36-dT6P

(D) CP44-dG6P:Coumarin-PEG44-dG6P

图5

3'-AC GAG GGC GCC GGC CGT || || || || || 6' 5'- TAG CGA AGG ATG TGA ACC TAA T<u>CC C</u>TG CTC CCG CGG CCGATC+

图6

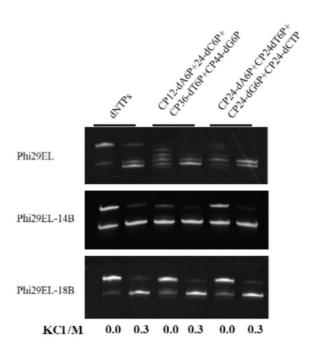


图7