



# (12) 发明专利申请

(10) 申请公布号 CN 114934026 A

(43) 申请公布日 2022. 08. 23

(21) 申请号 202210557719.X

*C12N 15/70* (2006.01)

(22) 申请日 2022.05.19

*C12N 1/21* (2006.01)

(30) 优先权数据

*C12P 19/34* (2006.01)

17/699,354 2022.03.21 US

*C12Q 1/686* (2018.01)

*C12Q 1/6806* (2018.01)

(71) 申请人 武汉爱博泰克生物科技有限公司

*C12R 1/19* (2006.01)

地址 430000 湖北省武汉市东湖新技术开发区高新二路388号武汉光谷国际生物医药企业加速器1.1期7栋4层01室

(72) 发明人 朱振宇 孙大鹏

(74) 专利代理机构 北京市中伦律师事务所

11410

专利代理师 王奕勋

(51) Int. Cl.

*C12N 9/00* (2006.01)

*C12N 15/52* (2006.01)

权利要求书2页 说明书23页

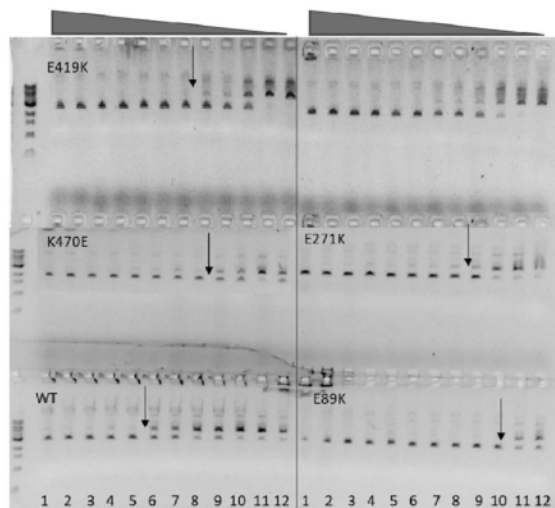
序列表33页 附图1页

(54) 发明名称

具有增加的连接效率的T4 DNA连接酶变体

(57) 摘要

本发明包括一种突变型T4DNA连接酶或其生物活性片段,其具有比野生型T4DNA连接酶更强的活性。如发明内容中更完全地描述的,所述突变型T4DNA连接酶或其生物活性片段具有不同于所述野生型的一个或多个取代。



1kb 梯

1. 一种突变型T4 DNA连接酶或其生物活性片段,其包含:下列氨基酸突变中的一个或多个,其中所述突变是在每个氨基酸序列中所指示位置处的取代,并且其中整个氨基酸序列以相邻偶数编号的序列标识号表示,但是其中相应突变型T4 DNA连接酶的每个所述氨基酸序列在每个所述偶数编号的序列标识号中不包括在其C末端处的6-元组氨酸标签(SEQ ID NO:26)和6个紧接在前的甘氨酸、丝氨酸和亮氨酸氨基酸:

E89K(SEQ ID NO:4)、E271K(SEQ ID NO:6)、D340R(SEQ ID NO:8)、D371Q(SEQ ID NO:10)、D371R(SEQ ID NO:12)、E419K(SEQ ID NO:14)、E438K(SEQ ID NO:16)、E440R(SEQ ID NO:18)、E440W(SEQ ID NO:20)、D452R(SEQ ID NO:22)、以及K470E(SEQ ID NO:24)。

2. 一种多核苷酸,其编码根据权利要求1所述的突变型T4 DNA连接酶中的一种的氨基酸序列。

3. 一种突变型T4 DNA连接酶或其生物活性片段,其包含:下列氨基酸突变中的一个或多个,其中所述突变是在每个氨基酸序列中所指示位置处的取代,并且其中整个氨基酸序列以相邻偶数编号的序列标识号表示:

E89K(SEQ ID NO:4)、E271K(SEQ ID NO:6)、D340R(SEQ ID NO:8)、D371Q(SEQ ID NO:10)、D371R(SEQ ID NO:12)、E419K(SEQ ID NO:14)、E438K(SEQ ID NO:16)、E440R(SEQ ID NO:18)、E440W(SEQ ID NO:20)、D452R(SEQ ID NO:22)、以及K470E(SEQ ID NO:24),但是其中相应突变型T4 DNA连接酶的每个所述氨基酸序列在每个所述偶数编号的序列标识号中不包括在其C末端处的6-元组氨酸标签(SEQ ID NO:26)和6个紧接在前的甘氨酸和丝氨酸氨基酸;并且

其中每个所述氨基酸序列对其某些氨基酸具有保守取代,仅达到与以所述相邻偶数编号的序列标识号表示的序列保持至少70%序列同一性的程度。

4. 根据权利要求3所述的突变型T4 DNA连接酶,其中每个所述氨基酸序列对其某些氨基酸具有保守取代,仅达到与以所述相邻偶数编号的序列标识号表示的序列保持至少80%、90%、95%、96%、97%、98%或99%序列同一性的程度。

5. 一种多核苷酸,其编码根据权利要求3所述的突变型T4 DNA连接酶中的一种的氨基酸序列。

6. 一种多核苷酸,其编码根据权利要求4所述的突变型T4 DNA连接酶中的一种的氨基酸序列。

7. 根据权利要求2所述的多核苷酸,其具有下列DNA序列中的一种,但不包括编码在其C末端处的6-元组氨酸标签(SEQ ID NO:26)和6个紧接在前的甘氨酸和丝氨酸氨基酸的DNA序列:

E89K(SEQ ID NO:3)、E271K(SEQ ID NO:5)、D340R(SEQ ID NO:7)、D371Q(SEQ ID NO:9)、D371R(SEQ ID NO:11)、E419K(SEQ ID NO:13)、E438K(SEQ ID NO:15)、E440R(SEQ ID NO:17)、E440W(SEQ ID NO:19)、D452R(SEQ ID NO:21)和K470E(SEQ ID NO:23)。

8. 一种载体,其掺入有根据权利要求4所述的多核苷酸。

9. 一种载体,其掺入有根据权利要求7所述的多核苷酸。

10. 一种细胞,其用根据权利要求8所述的多核苷酸转化并表达根据权利要求8所述的多核苷酸。

11. 一种细胞,其用根据权利要求9所述的多核苷酸转化并表达根据权利要求9所述的

多核苷酸。

12. 一种在不同多核苷酸之间或通过连接多核苷酸的5'末端和3'末端来进行多核苷酸连接以产生环状多核苷酸的方法,其中所述多核苷酸具有平滑末端或粘性末端,所述方法包括:

提供连接混合物,其包含待连接的多核苷酸和根据权利要求1所述的突变型T4 DNA连接酶或生物活性片段;以及

将所述连接混合物置于其中发生连接的温度下。

13. 一种在不同多核苷酸之间或通过连接多核苷酸的5'末端和3'末端来进行多核苷酸连接以产生环状多核苷酸的方法,其中所述多核苷酸具有平滑末端或粘性末端,所述方法包括:

提供连接反应混合物,其包含缓冲剂、待连接的多核苷酸和根据权利要求1所述的突变型T4 DNA连接酶或生物活性片段;以及

将所述连接反应混合物置于适于连接的温度条件下。

14. 根据权利要求13所述的方法,其中所述连接反应混合物包括Tris-HCl、MgCl<sub>2</sub>、ATP、二硫苏糖醇和水。

15. 一种在不同多核苷酸之间或通过连接多核苷酸的5'末端和3'末端来进行多核苷酸连接以产生环状多核苷酸的方法,其中所述多核苷酸具有平滑末端或粘性末端,所述方法包括:

提供连接反应混合物,其包含缓冲剂、待连接的多核苷酸和根据权利要求3所述的突变型T4 DNA连接酶或生物活性片段;和

将所述连接反应混合物置于适于连接的温度条件下。

16. 根据权利要求15所述的方法,其中所述连接反应混合物包括Tris-HCl、MgCl<sub>2</sub>、ATP、二硫苏糖醇和水。

## 具有增加的连接效率的T4 DNA连接酶变体

### [0001] 序列表

[0002] 本申请包含序列表,该序列表已以ASCII格式以电子方式提交,并通过引用在此整体并入。上述ASCII副本创建于2022年4月4日,名称为ABCL-T4HiAct\_SL.txt,大小为78749字节。

### 背景技术

[0003] 连接酶通常用于分子生物学,以用于在并置的5'磷酸和3'羟基末端的交叉处形成双螺旋核酸片段之间的磷酸二酯键。通过设计每个双链片段之间的互补突出端,连接可以是位置特异性连接和定向连接二者。这允许DNA或RNA片段特异性整合到更大的载体中,以满足分子生物学的需要。在可商购获得的连接酶中,T4 DNA连接酶是一种多功能的酶,其催化在粘性末端和平滑末端两者处连接双螺旋DNA或RNA的键,具有快速的连接速度,并且还经由连接修复带切口DNA中存在的错配。它可以连接DNA、RNA和RNA-DNA杂交体的粘性末端或平滑末端(但不连接单链核酸)。它还可以按比大肠杆菌DNA连接酶更高的效率连接平滑末端DNA。

[0004] 因为T4 DNA连接酶是许多分子生物学方案的骨架并且被持续需要,因此由于它在分子生物学试剂市场中的重要性而使它为许多生命科学产品公司创造了很大份额的收入。增加T4 DNA连接酶的活性有益于T4 DNA连接酶本身的制造,从而允许减少产生大量最终产品所需的生产量。取决于使用者的目标,增加T4 DNA连接酶的活性还通过减少酶或底物完全连接所需的总时间来增加酶的有用性。

### 发明内容

[0005] 本发明涉及与野生型连接酶相比表现出增强的连接活性的工程化T4 DNA连接酶突变体。以下T4 DNA连接酶突变体(在指示位置处具有取代,并且其中每个突变体的氨基酸序列是指定取代后的序列识别号,并且其中每个这种突变体的DNA序列是序列表中在每个突变体的氨基酸序列识别号之前的奇数序列识别号)被鉴定为具有这种增强的连接活性(其中每个都具有在C末端处添加的标签序列:GSGSSGHHHHH (SEQ ID NO:25)):E89K (SEQ ID NO:4)、E271K (SEQ ID NO:6)、D340R (SEQ ID NO:8)、D371Q (SEQ ID NO:10)、D371R (SEQ ID NO:12)、E419K (SEQ ID NO:14)、E438K (SEQ ID NO:16)、E440R (SEQ ID NO:18)、E440W (SEQ ID NO:20)、D452R (SEQ ID NO:22)和K470E (SEQ ID NO:24)。本发明还包括具有以上突变中的至少一种的T4 DNA连接酶突变体氨基酸序列,但是其中T4 DNA连接酶突变体氨基酸序列的剩余部分仅具有保守取代,使得所述分子与序列表中的相应T4 DNA连接酶突变体氨基酸序列具有至少70%、80%、90%、95%、96%、97%、98%或99%的同一性(在下文中被称为“变体序列”)。

[0006] 本发明还包括在以上突变体的每个氨基酸序列之前的DNA序列(即,分别为SEQ ID NO:),并且还包括前述DNA序列和其他简并核酸序列(统称为“简并核酸序列”),其编码(i)以上T4 DNA连接酶突变体中的每一个,和(ii)变体序列中的任一个的氨基酸序列。

[0007] 本发明还包括包含任何简并核酸序列的载体;以及用任何此类载体或简并核酸序列转化并能够表达以上T4 DNA连接酶突变体氨基酸序列或变体序列中的任一个的细胞。

[0008] 本发明还包括一种组合物或试剂盒,其包含以上T4 DNA连接酶突变体氨基酸序列或变体序列、简并核酸序列或包含此类简并核酸序列的载体中的任一个。本发明还包括一种扩增靶核酸的方法,其中在被设计为扩增靶核酸的反应混合物中采用以上T4 DNA连接酶突变体或变体序列中的任一个,并且将所述试剂混合物经受用于扩增所述靶核酸的条件。

[0009] 与野生型相比,以上突变型T4 DNA连接酶突变体在扩增靶DNA序列方面在更低浓度下具有更强的活性,并且变体序列也预期具有这种更强的活性。

## 附图说明

[0010] 图1A和1B示出了来自与T4 DNA连接酶突变体中的每一种相比野生型T4DNA连接酶(图1a的左下子图中的“WT”)的活性测定的一系列凝胶电泳结果。每个图1a和1b中三个最左侧子图的左边的1kb梯是凝胶带的参考组。每个凝胶中存在12个柱,使得从左到右,每个柱代表T4 DNA连接酶在1:2连续稀释后的浓度,并且其中在柱1中,700ng/ $\mu$ l的T4 DNA连接酶(或突变体)初始存在于溶液中。每个凝胶在稀释水平处都有一个箭头,其中显著量(上条带,右侧标记为“1”)的超螺旋质粒产物和限制性消化的线性化底物质粒(中条带,右侧标记为“2”)是明显的。最下面的条带(右侧标记为“3”)是由T4 DNA连接酶野生型或指示的突变体连接的开环质粒产物。

## 具体实施方式

[0011] 术语“生物活性片段”是指T4 DNA连接酶突变体或变体序列的任何片段、衍生物、同源物或类似物,其具有生物分子特有的体内或体外活性;包括例如连接酶活性,或经由连接修复带切口DNA中存在的错配。在一些实施方式中,突变型T4 DNA连接酶的生物活性片段、衍生物、同源物或类似物在任何体内或体外测定中均具有突变型T4 DNA连接酶的任何程度的生物活性。

[0012] 在一些实施方式中,生物活性片段可以任选地包括突变型T4 DNA连接酶的任何数量的连续氨基酸残基或变体序列。本发明还包括编码任何这种生物活性片段和/或简并核酸序列的多核苷酸。

[0013] 生物活性片段可以来自转录后加工或可替代剪接RNA的翻译,或者可替代地可以通过工程化、批量合成或其他合适的操作产生。生物活性片段包括在天然或内源细胞中表达的片段,以及在表达系统(例如像细菌、酵母、植物、昆虫或哺乳动物细胞)中产生的片段。

[0014] 如本文所用,短语“保守氨基酸取代”或“保守突变”是指一个氨基酸被另一个具有共同特性的氨基酸替换。定义单个氨基酸之间共同特性的功能性方法是分析同源生物体的相应蛋白质之间氨基酸变化的归一化频率(Schulz(1979)Principles of Protein Structure[蛋白质结构原理],Springer-Verlag)。根据此类分析,氨基酸组可以定义为组内的氨基酸优先相互交换,因此在它们对整个蛋白质结构的影响方面彼此最相似(前述Schulz(1979))。以这种方式定义的氨基酸组的实例可以包括:“带电荷/极性组”,包括Glu、Asp、Asn、Gln、Lys、Arg和His;“芳族或环状组”,包括Pro、Phe、Tyr和Trp;以及“脂族组”,包括Gly、Ala、Val、Leu、Ile、Met、Ser、Thr和Cys。在每个组内,还可以鉴定亚组。例如,带电荷/

极性氨基酸的组可以细分为多个亚组,包括:“带正电荷的亚组”,包括Lys、Arg和His;“带负电荷的亚组”,包括Glu和Asp;以及“极性亚组”,包括Asn和Gln。在另一个实例中,芳族或环状组可以细分为多个亚组,包括:“氮环亚组”,包括Pro、His和Trp;“苯基亚组”,包括Phe和Tyr。在另一个进一步的实例中,脂族组可以细分为多个亚组,包括:“大脂族非极性亚组”,包括Val、Leu和Ile;“脂族弱极性亚组”,包括Met、Ser、Thr、和Cys;以及“小残基亚组”,包括Gly和Ala。保守突变的实例包括上述亚组内氨基酸的氨基酸取代,诸如但不限于:Lys取代Arg,或反之,使得可以保持正电荷;Glu取代Asp,或反之,使得可以保持负电荷;Ser取代Thr,或反之,使得可以保持游离-OH;以及Gln取代Asn,或反之,使得可以保持游离-NH<sub>2</sub>。“保守变体”是包括一个或多个氨基酸的多肽,所述一个或多个氨基酸已被取代以用具有共同特性的氨基酸(例如,属于如上所述的相同氨基酸组或亚组)替换参考多肽(例如,其序列在出版物或序列数据库中公开或其序列已通过核酸测序确定的多肽)的一个或多个氨基酸。

[0015] 当提及基因时,“突变体”意指相对于天然或野生型基因,所述基因具有至少一个碱基(核苷酸)改变、缺失或插入。突变(一个或多个核苷酸的改变、缺失和/或插入)可以在基因的编码区中,或者可以在内含子、3'UTR、5'UTR或启动子区域中。作为非限制性实例,突变基因可以是在启动子区域内具有插入的基因,所述插入可以增加或减少所述基因的表达;可以是具有缺失的基因,导致无功能蛋白、截短蛋白、显性阴性蛋白或无蛋白的产生;或者,可以是一个或多个点突变的基因,导致所编码蛋白质的氨基酸改变或导致基因转录物的异常剪接。

[0016] 术语“本发明的突变型T4 DNA连接酶”和“突变型T4 DNA连接酶”在此具体实施方式部分中使用时,根据上下文,共同或单独地是指测试的并表现出增强的连接活性的突变型T4 DNA连接酶多肽,它们是:

[0017] E89K(SEQ ID NO:4)、E271K(SEQ ID NO:6)、D340R(SEQ ID NO:8)、D371Q(SEQ ID NO:10)、D371R(SEQ ID NO:12)、E419K(SEQ ID NO:14)、E438K(SEQ ID NO:16)、E440R(SEQ ID NO:18)、E440W(SEQ ID NO:20)、D452R(SEQ ID NO:22)和K470E(SEQ ID NO:24),其中每个具有标签序列:在C末端处添加的GSGSSGHHHHH(SEQ ID NO:25),和/或变体序列和/或简并核酸序列,如发明内容部分中所定义的那些术语。

[0018] “天然存在”或“野生型”是指在自然界中发现的形式。例如,天然存在或野生型多肽或多核苷酸序列是存在于生物体中的序列,其没有被人为操作有意修饰。

[0019] 关于核酸或多肽序列的术语“同一性百分比”或“同源性”定义为在排列序列以获得最大同一性百分比并引入缺口(如果需要)以实现最大同源性百分比之后,候选序列中与已知多肽相同的核苷酸或氨基酸残基的百分比。N-末端或C-末端插入或缺失不应被解释为影响同源性。核苷酸或氨基酸序列水平上的同源性或同一性可以通过BLAST(基本局部比对搜索工具(Basic Local Alignment Search Tool))分析来确定,所述分析使用由程序blastp、blastn、blastx、tblastn和tblastx采用的算法(Altschul(1997),Nucleic Acids Res[核酸研究].25,3389-3402和Karlin(1990),Proc.Natl.Acad.Sci.USA[美国国家科学院院刊]87,2264-2268),所述程序是为序列相似性搜索定制的。BLAST程序使用的方法是首先考虑查询序列和数据库序列之间的相似片段(有或没有缺口),然后评估被鉴定的所有匹配的统计显著性,最后仅总结满足预先选择的显著性阈值的那些匹配。关于序列数据库相似性搜索中基本问题的讨论,参见Altschul(1994),Nature Genetics[自然遗传学]6,119-

129.直方图、描述、比对、期望(即,用于报告与数据库序列匹配的统计显著性阈值)、截止、矩阵和过滤器(低复杂度)的搜索参数可以是默认设置。blastp、blastx、tblastn和tblastx使用的默认评分矩阵是BLOSUM62矩阵(Henikoff(1992),Proc.Natl.Acad.Sci.USA[美国国家科学院院刊]89,10915-10919),推荐用于长度超过85个单位(核苷酸碱基或氨基酸)的查询序列。

[0020] 在一些实施方式中,本发明涉及用于进行连接反应的方法(以及相关试剂盒、系统、设备和组合物),所述方法包括或由以下组成:在一种或多种核苷酸存在下,使突变型T4 DNA连接酶或其生物活性片段与核酸模板接触,并使用突变型T4 DNA连接酶或其生物活性片段连接所述一种或多种核苷酸中的至少一种。

[0021] 在一些实施方式中,所述方法可以包括将双链RNA或DNA多核苷酸链连接到环状分子中。在一些实施方式中,所述方法可以进一步包括通过使用传感器检测指示连接的信号。在一些实施方式中,传感器是ISFET。在一些实施方式中,传感器可以包括连接反应中的可检测标记或可检测试剂。

[0022] 在一些实施方式中,本发明涉及用于进行核酸的滚环扩增(参见美国专利号5,714,320,通过引用并入)的方法(以及相关的试剂盒、系统、设备和组合物),所述方法使用突变型T4 DNA连接酶作为扩增过程的连接步骤中的酶。扩增包括在溶液中扩增核酸,以及在固体支持物(诸如存在于固体支持物表面上的核酸珠、流动池、核酸阵列或孔)上克隆扩增核酸。

[0023] 制备突变型T4 DNA连接酶

[0024] 本发明的突变型T4 DNA连接酶可以在任何合适的宿主系统中表达,所述宿主系统包括细菌、酵母、真菌、杆状病毒、植物或哺乳动物宿主细胞。对于细菌宿主细胞,用于指导本公开的核酸构建体的转录的合适启动子包括从以下获得的启动子:大肠杆菌(E.coli)乳糖操纵子、天蓝色链霉菌琼脂糖酶基因(dagA)、枯草芽孢杆菌果聚糖蔗糖酶基因(sacB)、地衣芽孢杆菌 $\alpha$ -淀粉酶基因(amyL)、嗜热脂肪芽孢杆菌麦芽糖淀粉酶基因(amyM)、解淀粉芽孢杆菌 $\alpha$ -淀粉酶基因(amyQ)、地衣芽孢杆菌青霉素酶基因(penP)、枯草芽孢杆菌xylA和xylB基因和原核 $\beta$ -内酰胺酶基因(Villa-Kamaroff等人,1978,Proc.Natl Acad.Sci.USA[美国国家科学院院刊]75:3727-3731),以及tac启动子(DeBoer等人,1983,Proc.Natl Acad.Sci.USA[美国国家科学院院刊]80:21-25)。

[0025] 对于丝状真菌宿主细胞,用于指导本公开的核酸构建体的转录的合适启动子包括从以下酶的基因获得的启动子:米曲霉TAKA淀粉酶、米黑根毛霉天冬氨酸蛋白酶、黑曲霉中性 $\alpha$ -淀粉酶、黑曲霉酸稳定性 $\alpha$ -淀粉酶、黑曲霉或泡盛曲霉葡萄糖淀粉酶(glaA)、米黑根毛霉脂肪酶、米曲霉碱性蛋白酶、米曲霉磷酸丙糖异构酶、构巢曲霉乙酰胺酶和尖孢镰孢胰蛋白酶样蛋白酶(WO 96/00787),以及NA2-tpi启动子(来自黑曲霉中性 $\alpha$ -淀粉酶和米曲霉磷酸丙糖异构酶的基因的启动子的杂交体)及其突变体、截短和杂交启动子。

[0026] 在酵母宿主中,有用的启动子可以来自以下酶的基因:酿酒酵母(Saccharomyces cerevisiae)烯醇化酶(ENO-1)、酿酒酵母半乳糖激酶(GAL1)、酿酒酵母乙醇脱氢酶/甘油醛-3-磷酸脱氢酶(ADH2/GAP)、以及酿酒酵母3-磷酸甘油酸激酶。酵母宿主细胞的其他有用的启动子由Romanos等人,1992,Yeast[酵母]8:423-488描述。

[0027] 对于杆状病毒表达,来源于鳞翅目(蛾和蝴蝶)的昆虫细胞系,诸如草地贪夜蛾被

用作宿主。基因表达受强启动子(例如,pPolh)的控制。

[0028] 植物表达载体基于根癌土壤杆菌(*Agrobacterium tumefaciens*)的Ti质粒,或基于烟草花叶病毒(TMV)、马铃薯X病毒或豇豆花叶病毒。植物表达载体中常用的组成型启动子是花椰菜花叶病毒(CaMV) 35S启动子。

[0029] 对于哺乳动物表达,培养的哺乳动物细胞系诸如中国仓鼠卵巢(CHO)、COS,包括人细胞系诸如HEK和HeLa,可以用于产生突变型T4 DNA连接酶。哺乳动物表达载体的实例包括腺病毒载体、pSV和pCMV系列质粒载体、牛痘病毒和逆转录病毒载体以及杆状病毒。巨细胞病毒(CMV)和SV40的启动子通常用于哺乳动物表达载体中以驱动基因表达。非病毒启动子,诸如延伸因子(EF)-1启动子也是已知的。

[0030] 用于表达的控制序列也可以是合适的转录终止子序列,即,由宿主细胞识别以终止转录的序列。终止子序列可操作地连接至编码所述多肽的核酸序列的3'末端。可以使用在所选择宿主细胞中具有功能性的任何终止子。

[0031] 例如,用于丝状真菌宿主细胞的示例性转录终止子可以从以下酶的基因获得:米曲霉TAKA淀粉酶、黑曲霉葡糖淀粉酶、构巢曲霉邻氨基苯甲酸合酶、黑曲霉 $\alpha$ -葡糖苷酶和尖孢镰孢胰蛋白酶样蛋白酶。

[0032] 用于酵母宿主细胞的示例性终止子可以从以下酶的基因获得:酿酒酵母烯醇化酶、酿酒酵母细胞色素C(CYC1)、以及酿酒酵母甘油醛-3-磷酸脱氢酶。

[0033] 用于昆虫、植物和哺乳动物宿主细胞的终止子也是熟知的。

[0034] 控制序列也可以是合适的前导序列,即对由宿主细胞进行的翻译很重要的mRNA的非翻译区。前导序列可操作地连接至编码所述多肽的核酸序列的5'末端。可以使用在所选择宿主细胞中具有功能性的任何前导序列。用于丝状真菌宿主细胞的示例性前导序列从米曲霉TAKA淀粉酶和构巢曲霉磷酸丙糖异构酶的基因获得。用于酵母宿主细胞的合适的前导序列从以下酶的基因获得:酿酒酵母烯醇化酶(ENO-1)、酿酒酵母3-磷酸甘油酸激酶、酿酒酵母 $\alpha$ -因子、以及酿酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶(ADH2/GAP)。

[0035] 控制序列还可以是多腺苷酸化序列,一种可操作地连接至所述核酸序列的3'末端并且当转录时由宿主细胞识别为将多腺苷酸残基添加至所转录的mRNA的信号的序列。在所选择宿主细胞中具有功能性的任何多腺苷酸化序列均可以用于本发明中。用于丝状真菌宿主细胞的示例性多腺苷酸化序列可以来自以下酶的基因:米曲霉TAKA淀粉酶、黑曲霉葡糖淀粉酶、构巢曲霉邻氨基苯甲酸合酶、尖孢镰孢胰蛋白酶样蛋白酶和黑曲霉 $\alpha$ -葡糖苷酶。

[0036] 控制序列还可以是编码与多肽的氨基末端连接的氨基酸序列并指导所编码多肽进入细胞的分泌途径的信号肽编码区。核酸序列的编码序列的5'端本身可以含有在翻译阅读框中与编码分泌的多肽的编码区的区段天然连接的信号肽编码区。可替代地,编码序列的5'端可以含有对编码序列是外源的信号肽编码区。在编码序列天然地不含有信号肽编码区的情况下,可能需要外源信号肽编码区。

[0037] 可替代地,外源信号肽编码区可以单纯地替代天然信号肽编码区以便增强多肽的分泌。然而,可以使用指导已表达多肽进入所选择宿主细胞的分泌途径的任何信号肽编码区。

[0038] 用于细菌宿主细胞的有效信号肽编码区是从以下酶的基因获得的信号肽编码区:芽孢杆菌NCIB 11837麦芽糖淀粉酶、嗜热脂肪芽孢杆菌 $\alpha$ -淀粉酶、地衣芽孢杆菌枯草杆菌

蛋白酶、地衣芽孢杆菌 $\beta$ -内酰胺酶、嗜热脂肪芽孢杆菌中性蛋白酶(nprT、nprS、nprM)、以及枯草芽孢杆菌prsa。进一步的信号肽由Simonen和Palva,1993, Microbiol Rev[微生物评论]57:109-137描述。

[0039] 用于丝状真菌宿主细胞的有效信号肽编码区可以从以下酶的基因获得的信号肽编码区:米曲霉TAKA淀粉酶、黑曲霉中性淀粉酶、黑曲霉葡糖淀粉酶、米黑根毛霉天冬氨酸蛋白酶、特异腐质霉纤维素酶、以及疏棉状腐质霉脂肪酶。

[0040] 用于酵母宿主细胞的有用的信号肽可以来自酿酒酵母 $\alpha$ -因子和酿酒酵母转化酶的基因。用于其他宿主细胞系统的信号肽也是熟知的。

[0041] 控制序列还可以是编码位于多肽的氨基末端处的氨基酸序列的前肽编码区。所得的多肽被称为前体酶(proenzyme)或多肽原(或在一些情况下被称为酶原(zymogen))。多肽原通常是无活性的并且可以通过催化切割或自身催化切割来自多肽原的前肽而转化为成熟的活性多肽。前肽编码区可以从以下酶的基因获得:枯草芽孢杆菌碱性蛋白酶(aprE)、枯草芽孢杆菌中性蛋白酶(nprT)、酿酒酵母 $\alpha$ -因子、米黑根毛霉天冬氨酸蛋白酶、以及嗜热毁丝霉乳糖酶(WO 95/33836)。

[0042] 在信号肽和前肽区二者都存在于多肽的氨基末端处的情况下,所述前肽区位于紧邻多肽的氨基末端,并且所述信号肽区位于紧邻前肽区的氨基末端。

[0043] 还可期望的是添加调节序列,其允许相对于宿主细胞的生长调节突变型T4DNA连接酶的表达。调节系统的实例是响应于化学或物理刺激(包括调节化合物的存在)而引起基因表达被开启或关闭的那些。在原核宿主细胞中,合适的调节序列包括lac、tac和trp操纵子系统。在酵母宿主细胞中,合适的调节系统包括例如,ADH2系统或GAL1系统。在丝状真菌中,合适的调节序列包括TAKA $\alpha$ 淀粉酶启动子、黑曲霉葡糖淀粉酶启动子和米曲霉葡糖淀粉酶启动子。用于其他宿主细胞的调节系统也是熟知的。

[0044] 调节序列的其他实例是允许基因扩增的那些序列。在真核系统中,这些包括在甲氨蝶呤存在下扩增的二氢叶酸还原酶基因以及用重金属扩增的金属硫蛋白基因。在这些情况下,编码本发明KRED多肽的核酸序列将与调节序列可操作地连接。

[0045] 另一个实施方式包括重组表达载体,其包含编码工程化突变型T4 DNA连接酶或其变体的多核苷酸,和一个或多个表达调节区诸如启动子和终止子,以及复制起点,这取决于它们将被引入的宿主的类型。上述各种核酸和控制序列可以连接在一起以产生重组表达载体,所述重组表达载体可以包括一个或多个便利的限制位点以允许编码突变型T4 DNA连接酶的核酸序列在此类位点处的插入或取代。可替代地,可以通过将核酸序列或包含所述序列的核酸构建体插入用于表达的适当载体中来表达突变型T4 DNA连接酶的核酸序列。在产生表达载体时,编码序列位于载体中,使得编码序列与用于表达的适当控制序列可操作地连接。

[0046] 重组表达载体可以是方便地经受重组DNA程序并且可以引起突变型T4 DNA连接酶多核苷酸序列表达的任何载体(例如,质粒或病毒)。载体的选择将通常取决于载体与待引入载体的宿主细胞的相容性。载体可以是直链质粒或闭合环状质粒。

[0047] 表达载体可以是自主复制载体,即作为染色体外实体存在的载体,其复制不依赖于染色体复制,例如质粒、染色体外元件、微染色体或人工染色体。载体可以含有用于确保自我复制的任何手段。可替代地,载体可以是这样的载体,当它引入宿主细胞中时被整合到

基因组中并与其中已整合了它的一个或多个染色体一起被复制。此外,可以使用单独的载体或质粒或共同含有待引入宿主细胞基因组的总DNA的两个或更多个载体或质粒,或可以使用转座子。

[0048] 本文表达载体优选地含有一个或多个选择性标记,其允许容易地选择转化的细胞。选择性标记是一种基因,其产物提供了杀生物剂抗性或病毒抗性、对重金属抗性、对营养缺陷型的原营养等。细菌选择性标记的实例是来自枯草芽孢杆菌或地衣芽孢杆菌的 $dal$ 基因,或赋予抗生素抗性(诸如氨基青霉素、卡那霉素、氯霉素(实施例1)或四环素抗性)的标记。用于酵母宿主细胞的合适标记是 $ADE2$ 、 $HIS3$ 、 $LEU2$ 、 $LYS2$ 、 $MET3$ 、 $TRP1$ 、和 $URA3$ 。用于丝状真菌宿主细胞中的选择性标记包括但不限于 $amdS$ (乙酰胺酶)、 $argB$ (鸟氨酸氨甲酰基转移酶)、 $bar$ (草丁膦乙酰转移酶)、 $hph$ (潮霉素磷酸转移酶)、 $niaD$ (硝酸还原酶)、 $pyrG$ (乳清酸核苷-5'-磷酸脱羧酶)、 $sC$ (硫酸腺苷酰基转移酶)、和 $trpC$ (邻氨基苯甲酸合酶)、以及其等效物。用于曲霉属细胞的实施方式包括构巢曲霉或米曲霉的 $amdS$ 和 $pyrG$ 基因,以及吸水链霉菌的 $bar$ 基因。用于昆虫、植物和哺乳动物细胞的选择性标记也是熟知的。

[0049] 本发明的表达载体优选地含有允许载体整合到宿主细胞的基因组中或载体在细胞中不依赖于基因组的自主复制的一个或多个元件。对于整合到宿主细胞基因组中,载体可以依赖于编码所述多肽的核酸序列或用于通过同源或非同源重组将载体整合到基因组中的所述载体的任何其他元件。

[0050] 可替代地,表达载体可以含有另外的核酸序列,用于指导通过同源重组整合到宿主细胞的基因组中。另外的核酸序列使得载体能够整合到宿主细胞基因组中一个或多个染色体的一个或多个精确位置处。整合元件可以是与宿主细胞基因组内的靶序列同源的任何序列。此外,整合元件可以是非编码或编码核酸序列。另一方面,载体可以通过非同源重组整合到宿主细胞的基因组中。

[0051] 为了自主复制,载体可以进一步包含复制起点,该复制起点使得载体能够在讨论中的宿主细胞中自主复制。细菌复制起点的实例是 $P15A$  ori,或允许在大肠杆菌中复制的质粒 $pBR322$ 、 $pUC19$ 、 $pACYC177$ (所述质粒具有 $P15A$  ori)或 $pACYC184$ 的复制起点、以及允许在芽孢杆菌中复制的质粒 $pUB110$ 、 $pE194$ 、 $pTA1060$ 或 $pAM31$ 的复制起点。用于酵母宿主细胞中的复制起点的实例是2微米复制起点, $ARS1$ 、 $ARS4$ 、 $ARS1$ 与 $CEN3$ 的组合,以及 $ARS4$ 与 $CEN6$ 的组合。复制起点可以是具有使其在宿主细胞中的功能对温度敏感的突变的复制起点(参见例如,Ehrlich,1978,Proc Natl Acad Sci.USA[美国国家科学院院刊]75:1433)。

[0052] 可以将突变型T4 DNA连接酶的多于一个拷贝的核酸序列插入宿主细胞中,以增加基因产物的产生。通过将序列的至少一个另外的拷贝整合到宿主细胞基因组中或者通过包括与所述核酸序列一起的可扩增的选择性标记基因可以获得核酸序列的增加的拷贝数目,其中通过在适当的选择性试剂的存在下培养细胞可以选择含有选择性标记基因的经扩增的拷贝以及由此所述核酸序列的额外拷贝的细胞。

[0053] 用于突变型T4 DNA连接酶多核苷酸的表达载体可商购获得。合适的商业表达载体包括来自圣路易斯的西格玛奥德里奇化学公司(Sigma-Aldrich Chemicals,St.Louis Mo.)的 $p3xFLAGTM$ 表达载体,其包括用于在哺乳动物宿主细胞中表达的CMV启动子和hGH聚腺苷酸化位点,以及用于在大肠杆菌中扩增的 $pBR322$ 复制起点和氨基青霉素抗性标记物。合适的其他表达载体是可从加利福尼亚州拉荷亚(LaJolla Calif.)的Stratagene公司商

购获得的pBluescriptII SK(-)和pBK-CMV,以及来源于pBR322(Gibco BRL)、pUC(Gibco BRL)、pREP4、pCEP4(Invitrogen)或pPoly(Lathe等人,1987,Gene[基因]57:193-201)的质粒。

[0054] 用于表达编码突变型T4 DNA连接酶的多核苷酸的合适宿主细胞是本领域熟知的,并且包括但不限于:细菌细胞,诸如大肠杆菌、克菲尔乳杆菌(Lactobacillus kefir)、短乳杆菌、微小乳杆菌、链霉菌和鼠伤寒沙门氏菌细胞;真菌细胞,诸如酵母细胞(例如,酿酒酵母或巴斯德毕赤酵母(ATCC登记号201178));昆虫细胞,诸如果蝇S2和夜蛾Sf9细胞;动物细胞,诸如CHO、COS、BHK、293和鲍斯黑色素瘤细胞(Bowes melanoma cell);以及植物细胞。用于上述宿主细胞的适当培养基和生长条件是本领域熟知的。

[0055] 可以通过本领域已知的各种方法将用于表达突变型T4 DNA连接酶的多核苷酸引入细胞。这些技术包括电穿孔、生物射弹粒子轰击、脂质体介导的转染、氯化钙转染和原生质体融合。用于将多核苷酸引入细胞的各种方法是技术人员已知的。

[0056] 可以根据已知的合成方法,通过标准固相方法来制备编码突变型T4 DNA连接酶的多核苷酸。在一些实施方式中,可以单独合成多达约100个碱基的片段,然后连接(例如,通过酶促或化学连接方法,或聚合酶介导的方法)以形成任何所希望的连续序列。例如,可以使用例如由Beaucage等人,1981,Tet Lett[四面体快报]22:1859-69描述的经典亚磷酰胺方法,或由Matthes等人,1984,EMBO J.[欧洲分子生物学学报]3:801-05描述的方法(例如,如其通常应用于自动合成方法中)而通过化学合成来制备多核苷酸。根据亚磷酰胺方法,例如在自动DNA合成仪中合成寡核苷酸,将其纯化、退火、连接并克隆到合适的载体中。另外,基本上任何核酸都可以从各种商业来源获得,所述商业来源诸如德克萨斯州米德兰的Midland Certified Reagent Company公司、加利福尼亚州拉蒙纳的Great American Gene Company公司、伊利诺伊州芝加哥的ExpressGen公司和加利福尼亚州阿拉米达的Operon Technologies公司。

[0057] 使用任何一种或多种熟知的蛋白质纯化技术,包括溶菌酶处理、超声处理、过滤、盐析、超速离心和色谱法,可以从细胞和/或培养基中回收在宿主细胞中表达的工程化突变型T4 DNA连接酶。用于从细菌(诸如大肠杆菌)裂解和高效提取蛋白质的合适溶液可从圣路易斯的西格玛奥德里奇公司(Sigma-Aldrich)以商品名CellLytic B.TM.商购获得。

[0058] 用于分离突变型T4 DNA连接酶的色谱技术包括反相色谱、高效液相色谱、离子交换色谱、凝胶电泳和亲和色谱。纯化条件将部分取决于诸如净电荷、疏水性、亲水性、分子量、分子形状的因素,并且对于本领域技术人员将是显而易见的。

[0059] 在一些实施方式中,亲和技术可以用于分离突变型T4 DNA连接酶。对于亲和色谱纯化,可以使用特异性结合突变型T4 DNA连接酶的任何抗体。为了产生抗体,各种宿主动物(包括但不限于兔、小鼠、大鼠等)可以通过注射化合物进行免疫。所述化合物可以通过侧链官能团或附接至侧链官能团的接头附接至合适的载体,诸如BSA。各种佐剂可以用于增强免疫应答,这取决于宿主种类,包括但不限于弗氏(Freund's)(完全和不完全)、矿物凝胶诸如氢氧化铝、表面活性物质诸如溶血卵磷脂、复合多元醇(pluronic polyols)、聚阴离子、肽、油乳剂、钥孔咸血蓝蛋白(keyhole limpet hemocyanin)、二硝基酚,以及潜在有用的人佐剂诸如BCG(卡介苗)和短小棒状杆菌。

[0060] 制备T4 DNA连接酶突变体的实例

[0061] 通过常规PCR诱变生成T4 DNA连接酶突变体,其中引物被设计成包含所需的碱基取代,并且在PCR过程中,突变被并入到扩增子中,从而替换原始序列。所有的T4 DNA连接酶突变体和野生型具有添加的C末端6聚体His标签(SEQ ID NO:26),以便于纯化,前面是Ser和Gly残基的6聚体系列,如图所示。

[0062] PCR之后进行DpnI消化,其破坏甲基化模板(不包含取代),从而仅留下具有取代的未甲基化PCR扩增子。

[0063] 然后将PCR扩增子直接转化到化学感受态大肠杆菌宿主细胞中,其中细菌用化学物质预处理,以使它们能够吸收并入扩增子的质粒。参见ThermoFisher Scientific,化学感受态细胞网页(提供用于生成化学感受态细胞的试剂盒)。

[0064] 基于具有凝胶电泳的标准连接测定表征和选择由转化的大肠杆菌宿主细胞表达的突变型T4 DNA连接酶多肽。连接酶催化在双螺旋DNA的互补粘性末端或平滑末端的5'与3'末端之间形成磷酸二酯键,并且与不同的T4 DNA连接酶突变体的连接程度可以使用适当的DNA染料在琼脂糖凝胶上可视化。在这种情况下,使用GelRed(加利福尼亚州旧金山的Biotium公司)对凝胶进行染色,以便在UV光下进行可视化。基于其在降低的酶浓度下的连接活性检查每个T4 DNA连接酶突变体的性能,并且将所得的活性与类似稀释的野生型("WT")连接酶的活性进行比较,允许确定在相同的条件下与野生型相比哪些突变体显示增加的活性。

[0065] 如下制备用于表征由转化的大肠杆菌宿主细胞表达的突变型T4 DNA连接酶多肽的连接底物。

[0066] 所使用的DNA载体是pUC19(新英格兰生物实验室,目录号N3041S)。PUC19是2686个碱基对长的双链环。用具有偏移/不对称切割位点的BsaI-HF<sup>®</sup>v2(新英格兰生物实验室,目录号R3733S)消化PUC19,其中在识别序列添加一个随机(N1)核苷酸之后切割5'链,并且在互补识别序列添加五个额外的随机(N5)核苷酸之后在3'链上发生切割。5'识别序列是GGTCTC。通过BsaI-HF<sup>®</sup>v2对PUC19的切割位点指定为5'-GGTCTC(N1)/(N5)-3'。

[0067] 将5 $\mu$ l浓度为1mg/ml的puC19与2.5 $\mu$ l 20000个单位/ml的BsaI-HF<sup>®</sup>v2、5 $\mu$ l10X rCutSmart<sup>™</sup>缓冲液(新英格兰生物实验室,目录号B6004S)(50mM乙酸钾、20mM Tris-乙酸盐、10mM乙酸镁、100 $\mu$ g/ml重组白蛋白)和35 $\mu$ l水组合。在将所有组分组合之后,将组合物在37 $^{\circ}$ C下温育,以进行消化。在37 $^{\circ}$ C下1小时后,将反应在80 $^{\circ}$ C下温育20分钟,以使BsaI-HF<sup>®</sup>v2热失活。然后将混合物用水稀释至浓度为10ng/ $\mu$ l。

[0068] 实施例:连接测定和结果

[0069] 如下进行连接程序。将每种T4 DNA连接酶(无论是野生型还是变体)用酶稀释剂(50%甘油,10mM tris-HCl)在连续稀释下进行稀释,使得每个样品获得12个不同的浓度。每个连续稀释的起始浓度为700ng/ $\mu$ ,在每个下一次稀释中稀释两倍,使得最终浓度为稀释之前样品浓度的50%,并且依此类推,总共12个1:2稀释。然后将4 $\mu$ l每种酶或系列稀释液添加到PCR板中。

[0070] 对于酶,将16 $\mu$ l由2 $\mu$ l of 10X T4 DNA连接酶反应缓冲液(新英格兰生物实验室,目录号B0202A;包括50mM Tris-HCl、10mM MgCl<sub>2</sub>、1mM ATP、10mM DTT(二硫苏糖醇))、1 $\mu$ l 10ng/ $\mu$ l用BsaI-HF<sup>®</sup>v2消化的pUC19和13 $\mu$ l水组成的主混合物添加到每个反应中。这使得

总体积/反应达到20 $\mu$ l。将反应在16 $^{\circ}$ C下温育10分钟。在10分钟的温育期之后,反应接受2分钟的80 $^{\circ}$ C热休克以停止任何进一步的活动。将6 $\mu$ l终止溶液液(120mM EDTA、30%甘油、50mM Tris-HCl pH 8.0、0.0125%溴酚蓝、0.1%SDS和5x凝胶红核酸染色剂(加利福尼亚州弗里蒙特的Biotium公司))添加到每个反应中。

[0071] 使用具有1.2%琼脂糖凝胶的凝胶电泳来对连接反应产物进行可视化。每个凝胶具有野生型T4 DNA连接酶样品系列和7个变体T4 DNA连接酶样品系列。将每个凝胶在180V下运行35分钟。

[0072] 与野生型比较的结果在图1A和1B中示出。表现出增加的连接活性的鉴定的12个突变体如下:

[0073] E89K(SEQ ID NO:4)、E271K(SEQ ID NO:6)、D340R(SEQ ID NO:8)、D371Q(SEQ ID NO:10)、D371R(SEQ ID NO:12)、E419K(SEQ ID NO:14)、E438K(SEQ ID NO:16)、E440R(SEQ ID NO:18)、E440W(SEQ ID NO:20)、D452R(SEQ ID NO:22)和K470E(SEQ ID NO:24),其中每个具有标签序列:在C末端处添加的GSGSSGHHHHH(SEQ ID NO:25),如图1A、1B所示。

[0074] 比较活性结果在下表1中列出,显示每个T4 DNA连接酶突变体的泳道差异,其中每个大于WT的泳道差异被分配2倍的活性增加。例如,相对于WT的1个泳道活性改善被赋予2值(如通过在泳道中很少或没有明显的代表超螺旋质粒产物的上条带和很少或没有明显的代表限制性消化的线性底物质粒的中间条带),而相对于WT的2个泳道活性改善被赋予4值并依此类推。

[0075] 表1:具有比WT更强的活性的T4 DNA连接酶变体

[0076]

突变体	相对于WT的活性(倍数)
E89K	32
K271K	8
D340R	2
D371R	8
D371Q	4
E419K	4
E438K	8
E440R	4
E440W	4
K470E	8
D452R	2

[0077] SEQ ID NO:1 T4 DNA连接酶CH野生型DNA

	1	ATGATTCTTAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGGCCAGGTTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTAAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTTGATGCTTATCCTGAA	720
[0078]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTAAGTCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCCTGGTAGAAATATACAGTCTTCTGCATTTTCGTTTGAAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGTAATAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAATTATTATATTGGAAAAATCTAGAGTGCAGAA	1320
	1321	TGCAACGGTTGGTTAAAATCTGATGGCCGCACTGATTACGTTAAATTATTCTTCCGATT	1380
	1381	GCGATTGTTTACGTGAAGATAAACTAAAGCTAATACATTCCAAGATGTATTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0079]		SEQ ID NO:2 T4 DNA连接酶CH野生型蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKGIATQSF	60
	61	GMLTLTDMLDFIEFTLATRKLGTGNAIEELTGYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGLDLLKEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
[0080]	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVVNNLDEAKVIYKKYIDQGLEGIILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLIVGIYPHRKDPTKAGGFILESECGKIKVNAGSGLKDKAGVKSHEL	420
	421	DRTRIMENQNYIYGKILECECNGWLKSDGRDQVYKLFPLAIRLREDKTKANTFEDVFGD	480
[0081]	481	FHEVTGLGSGSSGHHHHHH*	499
[0082]		SEQ ID NO:3 T4 DNA连接酶CH E89K DNA	

	1	ATGATTCTTAAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAAAAATTAACTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAAGTTTTGCGTCGAGTGATGATGCGGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGGCCAGGTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTGTGCTTATCCTGAA	720
[0083]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGACTGCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCGTTGGTAGAAATATACAGTCTTCTGCATTTTCGTTTGAAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGAAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGTAATTCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAACCAAAATTATTATATTGAAAAATTCTAGAGTGCAGAA	1320
	1321	TGCAACGGTTGGTTAAAATCTGATGGCCGCACTGATTACGTAAATTTTCTTCCGATT	1380
	1381	GCGATTGTTTACGTGAAGATAAACTAAAGCTAATACATTCTGAAGATGATTTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0084]		SEQ ID NO:4 T4 DNA连接酶CH E89K蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKPIATQSF	60
	61	GMLTLTDMLDFIEFTLATRKLGTGNAAIEKLTGYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
[0085]	181	SRAGNEYLGDLLEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVVNNLDEAKVIYKKYIDQGLEGIILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLKIVGIYPHRKDPTKAGGFILESEC GKIKVNAGSLKDKAGVKSHEL	420
	421	DRTRIMENQNYIYGKILECECNGWLKSDGRTDYVKLFLPIAIRLREDKTKANTFEDVFGD	480
[0086]	481	FHEVTGLGSGSSGHHHHHH*	499
[0087]		SEQ ID NO:5 T4 DNA连接酶CH E271K DNA	

	1	ATGATTCTTAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGCCAGGTTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTTGATGCTTATCCTGAA	720
[0088]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTAAGTCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCT <u>AAAA</u> AGAAGCACAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCCTGGTAGAAATATACAGTCTTCCTGCATTTCGTTTGAAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGTAAAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAATTATTATATTGGAAAAATTCTAGAGTGCAGAA	1320
	1321	TGCAACGGTTGGTAAAATCTGATGGCCGCACTGATTACGTTAAATTATTTCTCCGATT	1380
	1381	GCGATTGTTTACGTGAAGATAAACTAAAGCTAATACATTGGAAGATGTATTTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0089]		SEQ ID NO:6 T4 DNA连接酶CH E271K蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKGIATQSF	60
	61	GMLTLTDMLDFIEFTLATRKLGTGNAIEELTYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
[0090]	181	SRAGNEYLGLDLLKEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
	241	NSKAKEFAEVAESRTASNGIANKSLKGTISKKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVNNLDEAKVIYKKYIDQGLEGIILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLKIVGIYPHRKDPTKAGGFILESEC GKIKVNAGSGLKDKAGVKSHEL	420
[0091]	421	DRTRIMENQNYIYGKILECECNGWLKSDGRD YVKLFLPIAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLGSGSSGHHHHHHH*	499
[0092]		SEQ ID NO:7 T4 DNA连接酶CH D340R DNA	

	1	ATGATTCTTAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGCCAGGTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTTGATGCTTATCCTGAA	720
[0093]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTAAGTCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCGTTGGTAGAAATATACAGTCTTCCTGCATTTCGTTTGAAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTCGT	1020
	1021	CAAGGTCTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGAAAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAATTATTATATTGGAAAAATTCTAGAGTGCAGAA	1320
	1321	TGCAACGGTTGGTAAAAATCTGATGGCCGCACTGATTACGTTAAATTATTTCTCCGATT	1380
	1381	GCGATTGTTTTACGTGAAGATAAACTAAAGCTAATACATTCCAAGATGTATTTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0094]		SEQ ID NO:8 T4 DNA连接酶CH D340R蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKGIATQSF	60
	61	GMLTLTDMLDFIEFTLATRKLGTGNAIEELTYITDGGKDDVEVLRVMMRDLECGASVS	120
[0095]	121	IANKVWPLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGLDLLKEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVNNLDEAKVIYKKYIRQGLEGILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLIVGIYPHRKDPTKAGGFILESECGKIKVNAGSGLKDKAGVKSHEL	420
[0096]	421	DRTRIMENQNYIYGKILECECNGWLKSDGRDQVYKFLPLAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLGSGSSGHHHHHH*	499
[0097]		SEQ ID NO:9 T4 DNA连接酶CH D371Q DNA	

	1	ATGATTCTTAAATTCGAAACGAAATAGCATCTATTGGTTCAACTAAACAGAAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAACCTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAAACAAAGTTTGGCCAGGTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATTAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTGGATGCTTATCCTGAA	720
[0098]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTACTGCTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCGTTGGTAGAAATATACAGTCTTCCTGCATTTTCGTTTGAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTC <u>CA</u> ATTGATTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGATAAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAATTATTATATTGGAAAAATTCTAGAGTGCGAA	1320
	1321	TGCAACGGTTGGTTAAAATCTGATGGCCGCACTGATTACGTAAATTTCTTCCGATT	1380
	1381	GCGATTGTTTTACGTGAAGATAAACTAAAGCTAATACATTCCAAGATGTATTTGGTGAT	1440
	1441	TTTCATGAGGTAACCTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0099]	SEQ ID NO:10	T4 DNA连接酶CH D371Q蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLRVYRLTYSRGLQYYIKKWPKGIATQSF	60
[0100]	61	GMLTLTDMLDFIEFTLATRKLGNAAIEELTYITDGGKKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGDLLEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVVNNLDEAKVIYKKYIDQGLEGILKNIDGLWENARS	360
[0101]	361	KNLYKFKEVIQVDLKIVGIYPHRKDPTKAGGFILESEC GKIKVNAGSGLKDKAGVKSHEL	420
	421	DRTRIMENQNYIYGKILECECNGWLKSDGRD YVKLFLPIAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLSGSGSHHHHHH*	499
[0102]	SEQ ID NO:11	T4 DNA连接酶CH D371R DNA	

	1	ATGATTCTTAAAATCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAAACAAAGTTTGGCCAGGTTAATTCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATTAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTGTGCTTATCCTGAA	720
[0103]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTACTGCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCGTTGGTAGAAATATACAGTCTTCTGCATTCGTTTGAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTCGTGTTGATTTAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGTAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGAAAACCAAATTTATTATATTGAAAAATTCTAGAGTGCAGAA	1320
	1321	TGCAACGGTTGGTTAAAATCTGATGGCCGCACTGATTACGTAAATTTTCTCCGATT	1380
	1381	GCGATTCGTTTACGTGAAGATAAACTAAAGCTAATACATTCGAAGATGATTTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0104]	SEQ ID NO:12	T4 DNA连接酶CH D371R蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKPGIATQSF	60
[0105]	61	GMLTLTDMLEDFIEFTLATRKLGTGNAIEELTYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGDLLEELIKMTAEARQIHPEGVLDGELVYHEQVKKEPEGLDFLFDAYPE	240
	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVNNLDEAKVIYKKYIDQGLEGILKNIDGLWENARS	360
[0106]	361	KNLYKFKEVIRVDLKIVGIYPHRKDPTKAGGFILESEC GKIKVNAGSGLKDKAGVKSHEL	420
	421	DRTRIMENQNYIYGKILECECNGWLKSDGRTDYVKLFLPIAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLGSGSGSHHHHHH*	499
[0107]	SEQ ID NO:13	T4 DNA连接酶CH E419K DNA	

	1	ATGATTCTTAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGCCAGGTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTTGATGCTTATCCTGAA	720
[0108]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTAAGTCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCCTGGTAGAAATATACAGTCTTCCTGCATTTCGTTTGAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGTAAAATCGCATAAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAATTATTATATTGGAAAAATTCTAGAGTGCAGAA	1320
	1321	TGCAACGGTTGGTAAAATCTGATGGCCGCACTGATTACGTTAAATTATTTCTCCGATT	1380
	1381	GCGATTGTTTACGTGAAGATAAACTAAAGCTAATACATTGGAAGATGTATTTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0109]		SEQ ID NO:14 T4 DNA连接酶CH E419K蛋白质	
[0110]	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKGIATQSF	60
	61	GMLTLTDMLEDFIEFTLATRKLGTNAIEELTYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGLDLLKEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
[0111]	301	DVRFSKLEQMTSGYDKVILIENQVNNLDEAKVIYKKYIDQGLEHILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLIVGIYPHRKDPTKAGGFILESECGKIKVNAGSGLKDKAGVKSHKL	420
	421	DRTRIMENQNYIYGKILECECNGWLKSDGRDVTYVVKLFLPIAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLSGSSGHHHHHHH*	499
[0112]		SEQ ID NO:15 T4 DNA连接酶CH E438K DNA	

	1	ATGATTCTTAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGCCAGGTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTTGATGCTTATCCTGAA	720
[0113]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTAAGTCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCGTTGGTAGAAATATACAGTCTTCCTGCATTTCGTTTGAAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGTAAAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAATTATTATATTGGAAAAATTCTAAAATGCGAA	1320
	1321	TGCAACGGTTGGTAAAAATCTGATGGCCGCACTGATTACGTTAAATTATTTCTCCGATT	1380
	1381	GCGATTGTTTTACGTGAAGATAAACTAAAGCTAATACATTGGAAGATGATTTTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0114]		SEQ ID NO:16 T4 DNA连接酶CH E438K蛋白质	
[0115]	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKPGIATQSF	60
	61	GMLTLTDMLDFIEFTLATRKLGTGNAIEELTYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGLDLLKEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
[0116]	301	DVRFSKLEQMTSGYDKVILIENQVNNLDEAKVIYKKYIDQGLEGIILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLKIVGIYPHRKDPTKAGGFILESEC GKIKVNAGSGLKDKAGVKSHEL	420
	421	DRTRIMENQNYIYGKILKCECNGWLKSDGRTDYVKLFLPIAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLGSGSSGHHHHHH*	499
[0117]		SEQ ID NO:17 T4 DNA连接酶CH E440R DNA	

	1	ATGATTCTTAAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGGCCAGGTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGCTTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTTGATGCTTATCCTGAA	720
[0118]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTAAGTCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCGTTGGTAGAAATATACAGTCTTCTGCATTTCGTTTGAAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGTAAAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAAATTATTATATTGGAAAAATTCTAGAGTGCCGT	1320
	1321	TGCAACGGTTGGTAAAAATCTGATGGCCGCACTGATTACGTAAATTATTTCTCCGATT	1380
	1381	GCGATTGTTTACGTGAAGATAAACTAAAGCTAATACATTGGAAGATGTATTTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0119]		SEQ ID NO:18 T4 DNA连接酶CH E440R蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKGIATQSF	60
	61	GMLTLTDMLDFIEFTLATRKLGTGNAIEELTGYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGLDLLKEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
[0120]	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVNNLDEAKVIYKKYIDQGLEGIILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLKIVGIYPHRKDPTKAGGFILESECGKIKVNAGSGLKDKAGVKSHEL	420
	421	DRTRIMENQNYIYGKILECR_CNGWLKSDGRD_YVKLFLPIAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLSGSGSHHHHHH*	499
[0121]		SEQ ID NO:19 T4 DNA连接酶CH E440W DNA	

	1	ATGATTCTTAAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCCTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGGCCAGGTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTTGATGCTTATCCTGAA	720
[0122]	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTAAGTCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCCTGGTAGAAATATACAGTCTTCTGCATTTCGTTTGAAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTAAAAGATAAAGCCGGTGTAAAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAATTATTATATTGGAAAAATTCTAGAGTGCTGG	1320
	1321	TGCAACGGTTGGTAAAAATCTGATGGCCGCACTGATTACGTTAAATTATTTCTCCGATT	1380
	1381	GCGATTGTTTACGTGAAGATAAACTAAAGCTAATACATTGGAAGATGTATTTGGTGAT	1440
	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACCACCATCATCACTAA	1500
[0123]		SEQ ID NO:20 T4 DNA连接酶CH E440W蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKPGIATQSF	60
	61	GMLTLTDMLDFIEFTLATRKLGTGNAIEELTYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGDLLEELIKMTAEARQIHPEGVLDGELVYHEQVKKEPEGLDFLFDAYPE	240
[0124]	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCМКFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVNNLDEAKVIYKKYIDQGLEGIILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLKIVGIYPHRKDPTKAGGFILESEC GKIKVNAGSGLKDKAGVKSHEL	420
	421	DRTRIMENQNYIYGKILECWCNGWLKSDGRTDYVKLFLPIAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLSGSGSGHHHHHHH*	499
[0125]		SEQ ID NO:21 T4 DNA连接酶CH D452R DNA	

	1	ATGATTCTTAAATTCTGAACGAAATAGCATCTATTGGTTCAACTAAACAGAAGCAAGCA	60
	61	ATTCTTGAAAAGAATAAAGATAATGAATTGCTTAAACGAGTATATCGTCTGACTTATTCT	120
	121	CGTGGGTTACAGTATTATATCAAGAAATGGCCTAAACCTGGTATTGCTACCCAGAGTTTT	180
	181	GGAATGTTGACTCTTACCGATATGCTTGACTTCATTGAATTCACATTAGCTACTCGGAAA	240
	241	TTGACTGGAAATGCAGCAATTGAGGAATTAAGTGGATATATCACCGATGGTAAAAAAGAT	300
	301	GATGTTGAAGTTTTGCGTCGAGTGATGATGCGAGACCTTGAATGTGGTGCTTCAGTATCT	360
	361	ATTGCAAACAAAGTTTGCCAGGTTAATTCCTGAACAACCTCAAATGCTCGCAAGTTCT	420
	421	TATGATGAAAAAGGCATTAATAAGAATATCAAATTTCCAGCCTTTGCTCAGTTAAAAGCT	480
	481	GATGGAGCTCGGTGTTTTGCTGAAGTTAGAGGTGATGAATTAGATGATGTTTCGTCTTTTA	540
	541	TCACGAGCTGGTAATGAATATCTAGGATTAGATCTTCTTAAGGAAGAGTTAATAAAATG	600
	601	ACCGCTGAAGCCCGCCAGATTCATCCAGAAGGTGTGTTGATTGATGGCGAATTGGTATAC	660
[0126]	661	CATGAGCAAGTTAAAAAGGAGCCAGAAGGCCTAGATTTTCTTTTTGATGCTTATCTGAA	720
	721	AACAGTAAAGCTAAAGAATTCGCCGAAGTAGCTGAATCACGTAAGTCTTCTAATGGAATC	780
	781	GCCAATAAATCTTTAAAGGGAACCATTTCTGAAAAAGAAGCACAAATGCATGAAGTTTCAG	840
	841	GTCTGGGATTATGTCCCGTTGGTAGAAATATACAGTCTTCCTGCATTTCGTTTGAATAT	900
	901	GATGTACGTTTTTCTAAACTAGAACAAATGACATCTGGATATGATAAAGTAATTTAATT	960
	961	GAAAACCAGGTAGTAAATAACCTAGATGAAGCTAAGGTAATTTATAAAAAGTATATTGAC	1020
	1021	CAAGGTCTTGAAGGTATTATTCTCAAAAATATCGATGGATTATGGGAAAATGCTCGTTCA	1080
	1081	AAAAATCTTTATAAATTTAAAGAAGTAATTGATGTTGATTTAAAAATTGTAGGAATTTAT	1140
	1141	CCTCACCGTAAAGACCCCTACTAAAGCGGGTGGATTTATTCTTGAGTCAGAGTGTGGAAAA	1200
	1201	ATTAAGGTAAATGCTGGTTCAGGCTTAAAAGATAAAGCCGGTGTAAAATCGCATGAACTT	1260
	1261	GACCGTACTCGCATTATGGAAAACCAAATTATTATATTGGAAAAATTCTAGAGTGCAGAA	1320
	1321	TGCAACGGTTGGTAAAAATCTGATGGCCGCACTCGTTACGTTAAATTATTTCTCCGATT	1380
	1381	GCGATTGTTTACGTGAAGATAAACTAAAGCTAATACATTGGAAGATGTATTTGGTGAT	1440
[0127]	1441	TTTCATGAGGTAAGTGGTCTAGGTTCTGGCAGTTCAGGTCATCACACCATCATCACTAA	1500
[0128]		SEQ ID NO:22 T4 DNA连接酶CH D452R蛋白质	
	1	MILKILNEIASIGSTKQKQAILEKNKDNELLKRVYRLTYSRGLQYYIKKWPKGIATQSF	60
	61	GMLTLTDMLDFIEFTLATRKLGTGNAIEELTYITDGGKDDVEVLRVMMRDLECGASVS	120
	121	IANKVWPGLIPEQPQMLASSYDEKGINKNIKFPFAQLKADGARCFAEVRGDELDDVRL	180
	181	SRAGNEYLGLDLLKEELIKMTAEARQIHPEGVLIDGELVYHEQVKKEPEGLDFLFDAYPE	240
[0129]	241	NSKAKEFAEVAESRTASNGIANKSLKGTISEKEAQCMKFQVWDYVPLVEIYSLPAFRLKY	300
	301	DVRFSKLEQMTSGYDKVILIENQVNNLDEAKVIYKKYIDQGLEGIILKNIDGLWENARS	360
	361	KNLYKFKEVIDVDLKIVGIYPHRKDPTKAGGFILESEC GKIKVNAGSGLKDKAGVKSHL	420
	421	DRTRIMENQNYIYGKILECECNGWLKSDGRTRYVKLFLPIAIRLREDKTKANTFEDVFGD	480
	481	FHEVTGLGSGSSGHHHHHH*	499
[0130]		SEQ ID NO:23 T4 DNA连接酶CH K470E DNA	



并且包含在由权利要求书的范围所限定的本发明的精神内。对于本领域技术人员来说,将显而易见的是,在不脱离本发明的范围和精神的情况下,可以对本文公开的发明进行各种替换和修改。可以在不存在本文没有明确公开为必要的任何一种或多种要素、或者任何一种或多种限制的情况下适当地实践本文说明性描述的本发明。因此,例如,在本文的每种情况下,在本发明的实施方式或实例中,术语“包含”、“包括”、“含有”等中的任一项都应该被广泛而无限制地阅读。本文中示例性描述的方法和过程可以不同的步骤顺序适当地实施,并且它们不必限于本文或权利要求中指示的步骤顺序。还应注意,除非上下文另外清楚地指出,否则如在本文以及在所附权利要求所用的,单数形式“一个/一种”以及“所述”包括复数指示物,并且复数包括单数形式。在任何情况下,都不能将本专利申请解释为限于本文具体公开的具体实例或实施方式或方法。在任何情况下,本专利申请都不应被解释为受专利商标局的任何审查员或任何其他官员或雇员所做的任何声明的限制,除非此声明由申请人在响应性书面材料中明确且无条件或专门保留地采用。

[0136] 已经在本文中宽泛且概括地对本发明进行了描述。落在整个公开文本之内的更窄的种类和亚属的分类中的每一个也形成了本发明的一部分。已经采用的术语和表达被用作描述的术语且没有限制性,并且并非旨在使用此类术语和表达而将所示出和描述的特征的任何等效物或其部分排除在外,但是将认识到的是提出要求保护的本发明范围内的不同修改是可能的。因此,应当理解,虽然已经通过优选实施方式和任选特征具体地公开了本发明,但本领域技术人员可以采用本文所公开概念的修改和变更,包括但不限于变体序列,并且认为此类修改和变更在由所附权利要求限定的本发明范围内。

## 序列表

- <110> 武汉爱博泰克生物科技有限公司  
 <120> 具有增加的连接效率的T4 DNA连接酶变体  
 <130> ABCL-T4HIACT  
 <140> 17/699,354  
 <141> 2022-03-21  
 <160> 26  
 <170> PatentIn version 3.5  
 <210> 1  
 <211> 1500  
 <212> DNA  
 <213> 未知的  
 <220>  
 <223> 未知的描述：

## 噬菌体T4序列

<400> 1  
 atgattctta aaattctgaa cgaaatagca tctattggtt caactaaaca gaagcaagca 60  
 attccttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120  
 cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt 180  
 ggaatgttga ctcttaccga tatgcttgac ttcattgaat tcacattagc tactcggaaa 240  
 ttgactggaa atgcagcaat tgaggaatta actggatata tcaccgatgg taaaaaagat 300  
 gatgttgaag ttttgcgtcg agtgatgatg cgagacctg aatgtggtgc ttcagtatct 360  
 attgcaaaca aagtttggcc aggtttaatt cctgaacaac ctcaaagct cgcaagttct 420  
 tatgatgaaa aaggcattaa taagaatata aaatttcag cctttgctca gttaaaagct 480  
 gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta 540  
 tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaaatg 600  
 accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attgggtatac 660  
 catgagcaag ttaaaaagga gccagaaggc ctagattttc tttttgatgc ttatcctgaa 720  
 aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780  
 gccataaat ctttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag 840  
 gtctgggatt atgtcccgtt ggtagaaata tacagtcttc ctgcatttcg tttgaaatat 900  
 gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aatttttaatt 960  
 gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020  
 caaggtcttg aaggtattat tetcaaaaat atcgatggat tatgggaaaa tgctcgttca 1080  
 aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaaattgt aggaatttat 1140  
 cctcaccgta aagaccctac taaagcgggt ggattttatc ttgagtcaga gtgtggaaaa 1200  
 attaaggtaa atgctggttc aggettaaaa gataaagccg gtgtaaaatc gcatgaactt 1260  
 gaccgtactc gcattatgga aaacaaaaat tatttatattg gaaaaattct agagtgcgaa 1320

tgcaacgggtt ggttaaaatc tgatggccgc actgattacg ttaaattatt tcttccgatt 1380  
 gcgattcggtt tacgtgaaga taaaactaaa gctaatacat tcgaagatgt atttgggtgat 1440  
 tttcatgagg taactgggtct aggttctggc agttcaggtc atcaccacca tcatcactaa 1500

<210> 2

<211> 499

<212> PRT

<213> 未知的

<220>

<223> 未知的描述:

噬菌体T4序列

<400> 2

Met	Ile	Leu	Lys	Ile	Leu	Asn	Glu	Ile	Ala	Ser	Ile	Gly	Ser	Thr	Lys	1	5	10	15
Gln	Lys	Gln	Ala	Ile	Leu	Glu	Lys	Asn	Lys	Asp	Asn	Glu	Leu	Leu	Lys	20	25	30	
Arg	Val	Tyr	Arg	Leu	Thr	Tyr	Ser	Arg	Gly	Leu	Gln	Tyr	Tyr	Ile	Lys	35	40	45	
Lys	Trp	Pro	Lys	Pro	Gly	Ile	Ala	Thr	Gln	Ser	Phe	Gly	Met	Leu	Thr	50	55	60	
Leu	Thr	Asp	Met	Leu	Asp	Phe	Ile	Glu	Phe	Thr	Leu	Ala	Thr	Arg	Lys	65	70	75	80
Leu	Thr	Gly	Asn	Ala	Ala	Ile	Glu	Glu	Leu	Thr	Gly	Tyr	Ile	Thr	Asp	85	90	95	
Gly	Lys	Lys	Asp	Asp	Val	Glu	Val	Leu	Arg	Arg	Val	Met	Met	Arg	Asp	100	105	110	
Leu	Glu	Cys	Gly	Ala	Ser	Val	Ser	Ile	Ala	Asn	Lys	Val	Trp	Pro	Gly	115	120	125	
Leu	Ile	Pro	Glu	Gln	Pro	Gln	Met	Leu	Ala	Ser	Ser	Tyr	Asp	Glu	Lys	130	135	140	
Gly	Ile	Asn	Lys	Asn	Ile	Lys	Phe	Pro	Ala	Phe	Ala	Gln	Leu	Lys	Ala	145	150	155	160
Asp	Gly	Ala	Arg	Cys	Phe	Ala	Glu	Val	Arg	Gly	Asp	Glu	Leu	Asp	Asp	165	170	175	
Val	Arg	Leu	Leu	Ser	Arg	Ala	Gly	Asn	Glu	Tyr	Leu	Gly	Leu	Asp	Leu	180	185	190	
Leu	Lys	Glu	Glu	Leu	Ile	Lys	Met	Thr	Ala	Glu	Ala	Arg	Gln	Ile	His	195	200	205	
Pro	Glu	Gly	Val	Leu	Ile	Asp	Gly	Glu	Leu	Val	Tyr	His	Glu	Gln	Val	210	215	220	



<220>

<223> 人工序列的描述: 合成的多核苷酸

<400> 3

```

atgattctta aaattctgaa cgaaatagca tctattggtt caactaaaca gaagcaagca 60
attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120
cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt 180
ggaatgttga ctcttaccga tatgcttgac ttatttgaat tcacattagc tactcggaaa 240
ttgactggaa atgcagcaat tgaaaaatta actggatata tcaccgatgg taaaaaagat 300
gatgttgaag ttttgcgtcg agtgatgatg cgagacctg aatgtggtgc ttcagtatct 360
attgcaaaca aagtttggcc aggtttaatt cctgaacaac ctcaaagct cgcaagttct 420
tatgatgaaa aaggcattaa taagaatata aaatttccag cctttgctca gttaaaagct 480
gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta 540
tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaatg 600
accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attggtatac 660
catgagcaag ttaaaaagga gccagaaggc ctagattttc tttttgatgc ttatcctgaa 720
aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780
gccaataaat ctttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag 840
gtctgggatt atgtcccgtt ggtagaaata tacagtcttc ctgcatttcg tttgaaatat 900
gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aattttaatt 960
gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020
caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaaa tgctcgttca 1080
aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaattgt aggaatttat 1140
cctcaccgta aagaccctac taaagcgggt ggatttattc ttgagtcaga gtgtggaaaa 1200
attaaggtaa atgctggttc aggcttaaaa gataaagccg gtgtaaaatc gcatgaactt 1260
gaccgtactc gcattatgga aaaccaaact tattatattg gaaaattct agagtgcgaa 1320
tgcaacgggt ggttaaaatc tgatggccgc actgattacg ttaaattatt tcttccgatt 1380
gcgattcgtt tacgtgaaga taaaactaaa gctaatacat tcgaagatgt atttggtgat 1440
tttcatgagg taactggtct aggttctggc agttcaggtc atcaccacca tcatcactaa 1500

```

<210> 4

<211> 499

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述: 合成的多肽

<400> 4

```

Met Ile Leu Lys Ile Leu Asn Glu Ile Ala Ser Ile Gly Ser Thr Lys
1           5           10           15
Gln Lys Gln Ala Ile Leu Glu Lys Asn Lys Asp Asn Glu Leu Leu Lys
                20           25           30

```

Arg Val Tyr Arg Leu Thr Tyr Ser Arg Gly Leu Gln Tyr Tyr Ile Lys																			
	35						40						45						
Lys Trp Pro Lys Pro Gly Ile Ala Thr Gln Ser Phe Gly Met Leu Thr																			
	50						55						60						
Leu Thr Asp Met Leu Asp Phe Ile Glu Phe Thr Leu Ala Thr Arg Lys																			
65					70								75						80
Leu Thr Gly Asn Ala Ala Ile Glu Lys Leu Thr Gly Tyr Ile Thr Asp																			
					85								90						95
Gly Lys Lys Asp Asp Val Glu Val Leu Arg Arg Val Met Met Arg Asp																			
					100								105						110
Leu Glu Cys Gly Ala Ser Val Ser Ile Ala Asn Lys Val Trp Pro Gly																			
	115																		125
Leu Ile Pro Glu Gln Pro Gln Met Leu Ala Ser Ser Tyr Asp Glu Lys																			
	130																		140
Gly Ile Asn Lys Asn Ile Lys Phe Pro Ala Phe Ala Gln Leu Lys Ala																			
145																			160
Asp Gly Ala Arg Cys Phe Ala Glu Val Arg Gly Asp Glu Leu Asp Asp																			
					165														175
Val Arg Leu Leu Ser Arg Ala Gly Asn Glu Tyr Leu Gly Leu Asp Leu																			
	180																		190
Leu Lys Glu Glu Leu Ile Lys Met Thr Ala Glu Ala Arg Gln Ile His																			
	195																		205
Pro Glu Gly Val Leu Ile Asp Gly Glu Leu Val Tyr His Glu Gln Val																			
	210																		220
Lys Lys Glu Pro Glu Gly Leu Asp Phe Leu Phe Asp Ala Tyr Pro Glu																			
225																			240
Asn Ser Lys Ala Lys Glu Phe Ala Glu Val Ala Glu Ser Arg Thr Ala																			
					245														255
Ser Asn Gly Ile Ala Asn Lys Ser Leu Lys Gly Thr Ile Ser Glu Lys																			
	260																		270
Glu Ala Gln Cys Met Lys Phe Gln Val Trp Asp Tyr Val Pro Leu Val																			
	275																		285
Glu Ile Tyr Ser Leu Pro Ala Phe Arg Leu Lys Tyr Asp Val Arg Phe																			
	290																		300
Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile																			
305																			320
Glu Asn Gln Val Val Asn Asn Leu Asp Glu Ala Lys Val Ile Tyr Lys																			
					325														335
Lys Tyr Ile Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Ile Asp																			

	340		345		350										
Gly	Leu	Trp	Glu	Asn	Ala	Arg	Ser	Lys	Asn	Leu	Tyr	Lys	Phe	Lys	Glu
	355		360		365										
Val	Ile	Asp	Val	Asp	Leu	Lys	Ile	Val	Gly	Ile	Tyr	Pro	His	Arg	Lys
	370		375		380										
Asp	Pro	Thr	Lys	Ala	Gly	Gly	Phe	Ile	Leu	Glu	Ser	Glu	Cys	Gly	Lys
385				390					395						400
Ile	Lys	Val	Asn	Ala	Gly	Ser	Gly	Leu	Lys	Asp	Lys	Ala	Gly	Val	Lys
			405		410				415						
Ser	His	Glu	Leu	Asp	Arg	Thr	Arg	Ile	Met	Glu	Asn	Gln	Asn	Tyr	Tyr
	420		425		430										
Ile	Gly	Lys	Ile	Leu	Glu	Cys	Glu	Cys	Asn	Gly	Trp	Leu	Lys	Ser	Asp
	435		440		445										
Gly	Arg	Thr	Asp	Tyr	Val	Lys	Leu	Phe	Leu	Pro	Ile	Ala	Ile	Arg	Leu
	450		455		460										
Arg	Glu	Asp	Lys	Thr	Lys	Ala	Asn	Thr	Phe	Glu	Asp	Val	Phe	Gly	Asp
465			470		475										480
Phe	His	Glu	Val	Thr	Gly	Leu	Gly	Ser	Gly	Ser	Ser	Gly	His	His	His
			485		490										495

His His His

<210> 5

<211> 1500

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 合成的多核苷酸

<400> 5

```

atgattctta aaattctgaa cgaatagca tctattggtt caactaaaca gaagcaagca 60
attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120
cgtgggttac agtattatat caagaaatgg ctaaactg gtattgctac ccagagtttt 180
ggaatgttga ctctaccga tatgcttgac tcattgaat tcacattagc tactcggaaa 240
ttgactggaa atgcagcaat tgaggaatta actggatata tcaccgatgg taaaaaagat 300
gatgttgaag ttttgcgtcg agtgatgatg cgagacctg aatgtggtgc ttcagtatct 360
attgcaaaca aagtttgcc aggtttaatt cctgaacaac ctcaaatgct cgcaagtctt 420
tatgatgaaa aaggcattaa taagaatata aaatttccag cctttgctca gttaaaagct 480
gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta 540
tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaaatg 600
accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attggtatac 660
catgagcaag ttaaaaagga gccagaagc ctagattttc tttttgatgc ttatcctgaa 720

```

aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780  
 gccaataaat ctttaaagg aaccatttct aaaaaagaag cacaatgcat gaagtttcag 840  
 gtctgggatt atgtcccgtt gtagaaata tacagtcttc ctgcatttcg tttgaaatat 900  
 gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aattttaatt 960  
 gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020  
 caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaaa tgctcgttca 1080  
 aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaattgt aggaatttat 1140  
 cctcaccgta aagaccctac taaagcgggt ggatttattc ttgagtcaga gtgtggaaaa 1200  
 attaaggtaa atgctgggtc aggcttaaaa gataaagccg gtgtaaaatc gcatgaactt 1260  
 gaccgtactc gcattatgga aaacaaaat tattatattg gaaaaattct agagtgcgaa 1320  
 tgcaacggtt ggtaaaaatc tgatggccgc actgattacg ttaaattatt tcttccgatt 1380  
 gcgattcgtt tacgtgaaga taaaactaaa gctaatacat tcgaagatgt atttggtgat 1440  
 tttcatgagg taactggtct aggttctggc agttcagtc atcaccacca tcatacactaa 1500

<210> 6

<211> 499

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：合成的多肽

<400> 6

Met	Ile	Leu	Lys	Ile	Leu	Asn	Glu	Ile	Ala	Ser	Ile	Gly	Ser	Thr	Lys
1			5						10					15	
Gln	Lys	Gln	Ala	Ile	Leu	Glu	Lys	Asn	Lys	Asp	Asn	Glu	Leu	Leu	Lys
			20					25					30		
Arg	Val	Tyr	Arg	Leu	Thr	Tyr	Ser	Arg	Gly	Leu	Gln	Tyr	Tyr	Ile	Lys
		35					40					45			
Lys	Trp	Pro	Lys	Pro	Gly	Ile	Ala	Thr	Gln	Ser	Phe	Gly	Met	Leu	Thr
		50				55					60				
Leu	Thr	Asp	Met	Leu	Asp	Phe	Ile	Glu	Phe	Thr	Leu	Ala	Thr	Arg	Lys
65					70				75					80	
Leu	Thr	Gly	Asn	Ala	Ala	Ile	Glu	Glu	Leu	Thr	Gly	Tyr	Ile	Thr	Asp
				85					90					95	
Gly	Lys	Lys	Asp	Asp	Val	Glu	Val	Leu	Arg	Arg	Val	Met	Met	Arg	Asp
			100					105				110			
Leu	Glu	Cys	Gly	Ala	Ser	Val	Ser	Ile	Ala	Asn	Lys	Val	Trp	Pro	Gly
		115					120					125			
Leu	Ile	Pro	Glu	Gln	Pro	Gln	Met	Leu	Ala	Ser	Ser	Tyr	Asp	Glu	Lys
		130				135						140			
Gly	Ile	Asn	Lys	Asn	Ile	Lys	Phe	Pro	Ala	Phe	Ala	Gln	Leu	Lys	Ala

145	150	155	160
Asp Gly Ala Arg Cys Phe Ala Glu Val Arg Gly Asp Glu Leu Asp Asp			
	165	170	175
Val Arg Leu Leu Ser Arg Ala Gly Asn Glu Tyr Leu Gly Leu Asp Leu			
	180	185	190
Leu Lys Glu Glu Leu Ile Lys Met Thr Ala Glu Ala Arg Gln Ile His			
	195	200	205
Pro Glu Gly Val Leu Ile Asp Gly Glu Leu Val Tyr His Glu Gln Val			
	210	215	220
Lys Lys Glu Pro Glu Gly Leu Asp Phe Leu Phe Asp Ala Tyr Pro Glu			
225	230	235	240
Asn Ser Lys Ala Lys Glu Phe Ala Glu Val Ala Glu Ser Arg Thr Ala			
	245	250	255
Ser Asn Gly Ile Ala Asn Lys Ser Leu Lys Gly Thr Ile Ser Lys Lys			
	260	265	270
Glu Ala Gln Cys Met Lys Phe Gln Val Trp Asp Tyr Val Pro Leu Val			
	275	280	285
Glu Ile Tyr Ser Leu Pro Ala Phe Arg Leu Lys Tyr Asp Val Arg Phe			
	290	295	300
Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile			
305	310	315	320
Glu Asn Gln Val Val Asn Asn Leu Asp Glu Ala Lys Val Ile Tyr Lys			
	325	330	335
Lys Tyr Ile Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Ile Asp			
	340	345	350
Gly Leu Trp Glu Asn Ala Arg Ser Lys Asn Leu Tyr Lys Phe Lys Glu			
	355	360	365
Val Ile Asp Val Asp Leu Lys Ile Val Gly Ile Tyr Pro His Arg Lys			
	370	375	380
Asp Pro Thr Lys Ala Gly Gly Phe Ile Leu Glu Ser Glu Cys Gly Lys			
385	390	395	400
Ile Lys Val Asn Ala Gly Ser Gly Leu Lys Asp Lys Ala Gly Val Lys			
	405	410	415
Ser His Glu Leu Asp Arg Thr Arg Ile Met Glu Asn Gln Asn Tyr Tyr			
	420	425	430
Ile Gly Lys Ile Leu Glu Cys Glu Cys Asn Gly Trp Leu Lys Ser Asp			
	435	440	445
Gly Arg Thr Asp Tyr Val Lys Leu Phe Leu Pro Ile Ala Ile Arg Leu			
	450	455	460



<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述: 合成的多肽

<400> 8

```

Met Ile Leu Lys Ile Leu Asn Glu Ile Ala Ser Ile Gly Ser Thr Lys
1           5           10           15
Gln Lys Gln Ala Ile Leu Glu Lys Asn Lys Asp Asn Glu Leu Leu Lys
           20           25           30
Arg Val Tyr Arg Leu Thr Tyr Ser Arg Gly Leu Gln Tyr Tyr Ile Lys
           35           40           45
Lys Trp Pro Lys Pro Gly Ile Ala Thr Gln Ser Phe Gly Met Leu Thr
           50           55           60
Leu Thr Asp Met Leu Asp Phe Ile Glu Phe Thr Leu Ala Thr Arg Lys
65           70           75           80
Leu Thr Gly Asn Ala Ala Ile Glu Glu Leu Thr Gly Tyr Ile Thr Asp
           85           90           95
Gly Lys Lys Asp Asp Val Glu Val Leu Arg Arg Val Met Met Arg Asp
           100          105          110
Leu Glu Cys Gly Ala Ser Val Ser Ile Ala Asn Lys Val Trp Pro Gly
           115          120          125
Leu Ile Pro Glu Gln Pro Gln Met Leu Ala Ser Ser Tyr Asp Glu Lys
           130          135          140
Gly Ile Asn Lys Asn Ile Lys Phe Pro Ala Phe Ala Gln Leu Lys Ala
145          150          155          160
Asp Gly Ala Arg Cys Phe Ala Glu Val Arg Gly Asp Glu Leu Asp Asp
           165          170          175
Val Arg Leu Leu Ser Arg Ala Gly Asn Glu Tyr Leu Gly Leu Asp Leu
           180          185          190
Leu Lys Glu Glu Leu Ile Lys Met Thr Ala Glu Ala Arg Gln Ile His
           195          200          205
Pro Glu Gly Val Leu Ile Asp Gly Glu Leu Val Tyr His Glu Gln Val
           210          215          220
Lys Lys Glu Pro Glu Gly Leu Asp Phe Leu Phe Asp Ala Tyr Pro Glu
225          230          235          240
Asn Ser Lys Ala Lys Glu Phe Ala Glu Val Ala Glu Ser Arg Thr Ala
           245          250          255
Ser Asn Gly Ile Ala Asn Lys Ser Leu Lys Gly Thr Ile Ser Glu Lys
           260          265          270

```

Glu Ala Gln Cys Met Lys Phe Gln Val Trp Asp Tyr Val Pro Leu Val  
 275 280 285  
 Glu Ile Tyr Ser Leu Pro Ala Phe Arg Leu Lys Tyr Asp Val Arg Phe  
 290 295 300  
 Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile  
 305 310 315 320  
 Glu Asn Gln Val Val Asn Asn Leu Asp Glu Ala Lys Val Ile Tyr Lys  
 325 330 335  
 Lys Tyr Ile Arg Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Ile Asp  
 340 345 350  
 Gly Leu Trp Glu Asn Ala Arg Ser Lys Asn Leu Tyr Lys Phe Lys Glu  
 355 360 365  
 Val Ile Asp Val Asp Leu Lys Ile Val Gly Ile Tyr Pro His Arg Lys  
 370 375 380  
 Asp Pro Thr Lys Ala Gly Gly Phe Ile Leu Glu Ser Glu Cys Gly Lys  
 385 390 395 400  
 Ile Lys Val Asn Ala Gly Ser Gly Leu Lys Asp Lys Ala Gly Val Lys  
 405 410 415  
 Ser His Glu Leu Asp Arg Thr Arg Ile Met Glu Asn Gln Asn Tyr Tyr  
 420 425 430  
 Ile Gly Lys Ile Leu Glu Cys Glu Cys Asn Gly Trp Leu Lys Ser Asp  
 435 440 445  
 Gly Arg Thr Asp Tyr Val Lys Leu Phe Leu Pro Ile Ala Ile Arg Leu  
 450 455 460  
 Arg Glu Asp Lys Thr Lys Ala Asn Thr Phe Glu Asp Val Phe Gly Asp  
 465 470 475 480  
 Phe His Glu Val Thr Gly Leu Gly Ser Gly Ser Ser Gly His His His  
 485 490 495

His His His

<210> 9

<211> 1499

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成的多核苷酸

<400> 9

atgattctta aaattctgaa cgaaatagca tctattggtt caactaaca gaagcaagca 60  
 attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120  
 cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt 180





385	390	395	400
Ile Lys Val Asn Ala Gly Ser Gly Leu Lys Asp Lys Ala Gly Val Lys			
	405	410	415
Ser His Glu Leu Asp Arg Thr Arg Ile Met Glu Asn Gln Asn Tyr Tyr			
	420	425	430
Ile Gly Lys Ile Leu Glu Cys Glu Cys Asn Gly Trp Leu Lys Ser Asp			
	435	440	445
Gly Arg Thr Asp Tyr Val Lys Leu Phe Leu Pro Ile Ala Ile Arg Leu			
	450	455	460
Arg Glu Asp Lys Thr Lys Ala Asn Thr Phe Glu Asp Val Phe Gly Asp			
465	470	475	480
Phe His Glu Val Thr Gly Leu Gly Ser Gly Ser Ser Gly His His His			
	485	490	495

His His His

<210> 11

<211> 1500

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成的多核苷酸

<400> 11

```

atgattctta aaattctgaa cgaaatagca tctattggtt caactaaaca gaagcaagca 60
attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120
cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt 180
ggaatgttga ctcttaccga tatgcttgac ttattgaat tcacattagc tactcggaaa 240
ttgactggaa atgcagcaat tgaggaatta actggatata tcaccgatgg taaaaaagat 300
gatgttgaag ttttgcgtcg agtgatgatg cgagacctg aatgtggtgc ttcagtatct 360
attgcaaaca aagtttgcc aggtttaatt cctgaacaac ctcaaatgct cgcaagttct 420
tatgatgaaa aaggcattaa taagaatata aaatttccag ctttgctca gttaaaagct 480
gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta 540
tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaaatg 600
accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attggtatac 660
catgagcaag ttaaaaagga gccagaaggc ctagatttct tttttgatgc ttatcctgaa 720
aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780
gccaataaat cttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag 840
gtctgggatt atgtcccgtt ggtagaaata tacagtcttc ctgcatttcg tttgaaatat 900
gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aattttaatt 960
gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020
caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaaa tgctcgttca 1080

```



195	200	205
Pro Glu Gly Val Leu Ile Asp	Gly Glu Leu Val Tyr His Glu Gln Val	
210	215	220
Lys Lys Glu Pro Glu Gly Leu Asp Phe Leu Phe Asp Ala Tyr Pro Glu		
225	230	235
Asn Ser Lys Ala Lys Glu Phe Ala Glu Val Ala Glu Ser Arg Thr Ala		
245	250	255
Ser Asn Gly Ile Ala Asn Lys Ser Leu Lys Gly Thr Ile Ser Glu Lys		
260	265	270
Glu Ala Gln Cys Met Lys Phe Gln Val Trp Asp Tyr Val Pro Leu Val		
275	280	285
Glu Ile Tyr Ser Leu Pro Ala Phe Arg Leu Lys Tyr Asp Val Arg Phe		
290	295	300
Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile		
305	310	315
Glu Asn Gln Val Val Asn Asn Leu Asp Glu Ala Lys Val Ile Tyr Lys		
325	330	335
Lys Tyr Ile Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Ile Asp		
340	345	350
Gly Leu Trp Glu Asn Ala Arg Ser Lys Asn Leu Tyr Lys Phe Lys Glu		
355	360	365
Val Ile Arg Val Asp Leu Lys Ile Val Gly Ile Tyr Pro His Arg Lys		
370	375	380
Asp Pro Thr Lys Ala Gly Gly Phe Ile Leu Glu Ser Glu Cys Gly Lys		
385	390	395
Ile Lys Val Asn Ala Gly Ser Gly Leu Lys Asp Lys Ala Gly Val Lys		
405	410	415
Ser His Glu Leu Asp Arg Thr Arg Ile Met Glu Asn Gln Asn Tyr Tyr		
420	425	430
Ile Gly Lys Ile Leu Glu Cys Glu Cys Asn Gly Trp Leu Lys Ser Asp		
435	440	445
Gly Arg Thr Asp Tyr Val Lys Leu Phe Leu Pro Ile Ala Ile Arg Leu		
450	455	460
Arg Glu Asp Lys Thr Lys Ala Asn Thr Phe Glu Asp Val Phe Gly Asp		
465	470	475
Phe His Glu Val Thr Gly Leu Gly Ser Gly Ser Ser Gly His His His		
485	490	495
His His His		

&lt;210&gt; 13

<211> 1500

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成的多核苷酸

<400> 13

```

atgattctta aaattctgaa cgaaatagca tctattggtt caactaaaca gaagcaagca 60
attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120
cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt 180
ggaatgttga ctcttaccga tatgcttgac tcattgaat tcacattagc tactcggaaa 240
ttgactggaa atgcagcaat tgaggaatta actggatata tcaccgatgg taaaaaagat 300
gatgttgaag ttttgcgtcg agtgatgatg cgagacctg aatgtgggtgc ttcagtatct 360
attgcaaaca aagtttgccc aggtttaatt cctgaacaac ctcaaagtct cgcaagttct 420
tatgatgaaa aaggcattaa taagaatata aaatttccag cctttgctca gttaaaagct 480
gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta 540
tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaaatg 600
accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attgggtatac 660
catgagcaag ttaaaaagga gccagaaggc ctagattttc tttttgatgc ttatcctgaa 720
aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780
gccaataaat ctttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag 840
gtctgggatt atgtcccgtt ggtagaaata tacagtcttc ctgcatttcg tttgaaatat 900
gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aattttaatt 960
gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020
caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaa tgctcgttca 1080
aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaattgt aggaatttat 1140
cctcaccgta aagaccctac taaagcgggt ggatttattc ttgagtcaga gtgtggaaaa 1200
attaaggtaa atgctggttc aggcttaaaa gataaagccg gtgtaaaatc gcataaactt 1260
gaccgtactc gcattatgga aaaccaaact tattatattg gaaaaattct agagtgcgaa 1320
tgcaacgggt ggtaaaaatc tgatggccgc actgattacg ttaaattatt tcttccgatt 1380
gcgattcgtt tacgtgaaga taaaactaaa gctaatacat tcgaagatgt atttggtgat 1440
tttcatgagg taactggtct aggttctggc agttcaggtc atcaccacca tcatcactaa 1500

```

<210> 14

<211> 499

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：合成的多肽

<400> 14

Met Ile Leu Lys Ile Leu Asn Glu Ile Ala Ser Ile Gly Ser Thr Lys

1	5	10	15
Gln Lys Gln Ala Ile Leu Glu Lys Asn Lys Asp Asn Glu Leu Leu Lys			
	20	25	30
Arg Val Tyr Arg Leu Thr Tyr Ser Arg Gly Leu Gln Tyr Tyr Ile Lys			
	35	40	45
Lys Trp Pro Lys Pro Gly Ile Ala Thr Gln Ser Phe Gly Met Leu Thr			
	50	55	60
Leu Thr Asp Met Leu Asp Phe Ile Glu Phe Thr Leu Ala Thr Arg Lys			
65	70	75	80
Leu Thr Gly Asn Ala Ala Ile Glu Glu Leu Thr Gly Tyr Ile Thr Asp			
	85	90	95
Gly Lys Lys Asp Asp Val Glu Val Leu Arg Arg Val Met Met Arg Asp			
	100	105	110
Leu Glu Cys Gly Ala Ser Val Ser Ile Ala Asn Lys Val Trp Pro Gly			
	115	120	125
Leu Ile Pro Glu Gln Pro Gln Met Leu Ala Ser Ser Tyr Asp Glu Lys			
	130	135	140
Gly Ile Asn Lys Asn Ile Lys Phe Pro Ala Phe Ala Gln Leu Lys Ala			
145	150	155	160
Asp Gly Ala Arg Cys Phe Ala Glu Val Arg Gly Asp Glu Leu Asp Asp			
	165	170	175
Val Arg Leu Leu Ser Arg Ala Gly Asn Glu Tyr Leu Gly Leu Asp Leu			
	180	185	190
Leu Lys Glu Glu Leu Ile Lys Met Thr Ala Glu Ala Arg Gln Ile His			
	195	200	205
Pro Glu Gly Val Leu Ile Asp Gly Glu Leu Val Tyr His Glu Gln Val			
	210	215	220
Lys Lys Glu Pro Glu Gly Leu Asp Phe Leu Phe Asp Ala Tyr Pro Glu			
225	230	235	240
Asn Ser Lys Ala Lys Glu Phe Ala Glu Val Ala Glu Ser Arg Thr Ala			
	245	250	255
Ser Asn Gly Ile Ala Asn Lys Ser Leu Lys Gly Thr Ile Ser Glu Lys			
	260	265	270
Glu Ala Gln Cys Met Lys Phe Gln Val Trp Asp Tyr Val Pro Leu Val			
	275	280	285
Glu Ile Tyr Ser Leu Pro Ala Phe Arg Leu Lys Tyr Asp Val Arg Phe			
	290	295	300
Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile			
305	310	315	320



tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaatg 600  
accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attggtatac 660  
catgagcaag ttaaaaagga gccagaaggc ctagattttc tttttgatgc ttatcctgaa 720  
aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780  
gccaataaat ctttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag 840  
gtctgggatt atgtcccgtt gtagaaata tacagtcttc ctgcatttcg tttgaaatat 900  
gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aattttaatt 960  
gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020  
caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaaa tgctcgttca 1080  
aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaaattgt aggaatttat 1140  
cctcaccgta aagaccctac taaagcgggt ggattttatc ttgagtcaga gtgtggaaaa 1200  
attaaggtaa atgctggttc aggettaaaa gataaagccg gtgtaaaatc gcatgaactt 1260  
gaccgtactc gcattatgga aaaccaaatt tattatattg gaaaaattct aaaatgcgaa 1320  
tgcaacggtt ggttaaaatc tgatggccgc actgattacg ttaaattatt tcttccgatt 1380  
gcgattcgtt tacgtgaaga taaaactaaa gctaatacat tcgaagatgt atttggtgat 1440  
tttcatgagg taactggtct aggttctggc agttcaggtc atcaccacca tcatcactaa 1500

<210> 16

<211> 499

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述: 合成的多肽

<400> 16

Met	Ile	Leu	Lys	Ile	Leu	Asn	Glu	Ile	Ala	Ser	Ile	Gly	Ser	Thr	Lys
1			5						10					15	
Gln	Lys	Gln	Ala	Ile	Leu	Glu	Lys	Asn	Lys	Asp	Asn	Glu	Leu	Leu	Lys
			20						25					30	
Arg	Val	Tyr	Arg	Leu	Thr	Tyr	Ser	Arg	Gly	Leu	Gln	Tyr	Tyr	Ile	Lys
			35						40					45	
Lys	Trp	Pro	Lys	Pro	Gly	Ile	Ala	Thr	Gln	Ser	Phe	Gly	Met	Leu	Thr
			50						55					60	
Leu	Thr	Asp	Met	Leu	Asp	Phe	Ile	Glu	Phe	Thr	Leu	Ala	Thr	Arg	Lys
65									70					75	
Leu	Thr	Gly	Asn	Ala	Ala	Ile	Glu	Glu	Leu	Thr	Gly	Tyr	Ile	Thr	Asp
									85					90	
Gly	Lys	Lys	Asp	Asp	Val	Glu	Val	Leu	Arg	Arg	Val	Met	Met	Arg	Asp
									100					105	
Leu	Glu	Cys	Gly	Ala	Ser	Val	Ser	Ile	Ala	Asn	Lys	Val	Trp	Pro	Gly
									115					120	
															125

Leu Ile Pro Glu Gln Pro Gln Met Leu Ala Ser Ser Tyr Asp Glu Lys  
 130 135 140  
 Gly Ile Asn Lys Asn Ile Lys Phe Pro Ala Phe Ala Gln Leu Lys Ala  
 145 150 155 160  
 Asp Gly Ala Arg Cys Phe Ala Glu Val Arg Gly Asp Glu Leu Asp Asp  
 165 170 175  
 Val Arg Leu Leu Ser Arg Ala Gly Asn Glu Tyr Leu Gly Leu Asp Leu  
 180 185 190  
 Leu Lys Glu Glu Leu Ile Lys Met Thr Ala Glu Ala Arg Gln Ile His  
 195 200 205  
 Pro Glu Gly Val Leu Ile Asp Gly Glu Leu Val Tyr His Glu Gln Val  
 210 215 220  
 Lys Lys Glu Pro Glu Gly Leu Asp Phe Leu Phe Asp Ala Tyr Pro Glu  
 225 230 235 240  
 Asn Ser Lys Ala Lys Glu Phe Ala Glu Val Ala Glu Ser Arg Thr Ala  
 245 250 255  
 Ser Asn Gly Ile Ala Asn Lys Ser Leu Lys Gly Thr Ile Ser Glu Lys  
 260 265 270  
 Glu Ala Gln Cys Met Lys Phe Gln Val Trp Asp Tyr Val Pro Leu Val  
 275 280 285  
 Glu Ile Tyr Ser Leu Pro Ala Phe Arg Leu Lys Tyr Asp Val Arg Phe  
 290 295 300  
 Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile  
 305 310 315 320  
 Glu Asn Gln Val Val Asn Asn Leu Asp Glu Ala Lys Val Ile Tyr Lys  
 325 330 335  
 Lys Tyr Ile Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Ile Asp  
 340 345 350  
 Gly Leu Trp Glu Asn Ala Arg Ser Lys Asn Leu Tyr Lys Phe Lys Glu  
 355 360 365  
 Val Ile Asp Val Asp Leu Lys Ile Val Gly Ile Tyr Pro His Arg Lys  
 370 375 380  
 Asp Pro Thr Lys Ala Gly Gly Phe Ile Leu Glu Ser Glu Cys Gly Lys  
 385 390 395 400  
 Ile Lys Val Asn Ala Gly Ser Gly Leu Lys Asp Lys Ala Gly Val Lys  
 405 410 415  
 Ser His Glu Leu Asp Arg Thr Arg Ile Met Glu Asn Gln Asn Tyr Tyr  
 420 425 430  
 Ile Gly Lys Ile Leu Lys Cys Glu Cys Asn Gly Trp Leu Lys Ser Asp

435	440	445
Gly Arg Thr Asp Tyr Val Lys Leu Phe Leu Pro Ile Ala Ile Arg Leu		
450	455	460
Arg Glu Asp Lys Thr Lys Ala Asn Thr Phe Glu Asp Val Phe Gly Asp		
465	470	475
Phe His Glu Val Thr Gly Leu Gly Ser Gly Ser Ser Gly His His His		
485	490	495
His His His		

<210> 17

<211> 1500

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成的多核苷酸

<400> 17

```

atgattctta aaattctgaa cgaaatagca tctattggtt caactaaaca gaagcaagca 60
attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120
cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt 180
ggaatgttga ctcttaccga tatgcttgac ttattgaat tcacattagc tactcggaaa 240
ttgactggaa atgcagcaat tgaggaatta actggatata tcaccgatgg taaaaaagat 300
gatgttgaag ttttgcgtcg agtgatgatg cgagacctg aatgtggtgc ttcagtatct 360
attgcaaaca aagtttgcc aggtttaatt cctgaacaac ctcaaatgct cgcaagttct 420
tatgatgaaa aaggcattaa taagaatata aaatttcag ctttgctca gttaaaagct 480
gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta 540
tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaatg 600
accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attggtatac 660
catgagcaag ttaaaaagga gccagaaggc ctagattttc tttttgatgc ttatcctgaa 720
aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780
gccaataaat ctttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag 840
gtctgggatt atgtcccgtt ggtagaaata tacagtcttc ctgcatttcg tttgaaatat 900
gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aattttaatt 960
gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020
caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaaa tgctcgttca 1080
aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaattgt aggaatttat 1140
cctcaccgta aagaccctac taaagcgggt ggatttattc ttgagtcaga gtgtggaaaa 1200
attaaggtaa atgctggttc aggettaaaa gataaagccg gtgtaaaatc gcatgaactt 1260
gaccgtactc gcattatgga aaacaaaaat tattatattg gaaaaattct agagtgccgt 1320
tgcaacgggt ggtaaaaatc tgatggccgc actgattacg ttaaattatt tcttccgatt 1380
gcgattcgtt tacgtgaaga taaaactaaa gctaatacat tcgaagatgt atttggtgat 1440

```



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

His His His

&lt;210&gt; 19

&lt;211&gt; 1500

&lt;212&gt; DNA

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述：合成的多核苷酸

&lt;400&gt; 19

atgattctta aaattctgaa cgaaatagca tctattggtt caactaaaca gaagcaagca 60  
 attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120  
 cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt 180  
 ggaatgttga ctcttaccga tatgcttgac ttattgaat tcacattagc tactcggaaa 240  
 ttgactggaa atgcagcaat tgaggaatta actggatata tcaccgatgg taaaaaagat 300  
 gatgttgaag ttttgcgtcg agtgaatgag cgagacctg aatgtggtgc ttcagtatct 360  
 attgcaaaca aagtttgcc aggtttaatt cctgaacaac ctcaaatgct cgcaagttct 420  
 tatgatgaaa aaggcattaa taagaatata aaatttccag ctttgctca gttaaaagct 480  
 gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta 540  
 tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaaatg 600  
 accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attgggtatac 660  
 catgagcaag ttaaaaagga gccagaaggc ctagatttct tttttgatgc ttatcctgaa 720  
 aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780  
 gccataaat cttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag 840  
 gtctgggatt atgtcccgtt ggtagaaata tacagtcttc ctgcatttcg tttgaaatat 900  
 gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aattttaatt 960  
 gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020  
 caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaaa tgctcgttca 1080  
 aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaattgt aggaatttat 1140  
 cctcaccgta aagaccctac taaagcgggt ggatttattc ttgagtcaga gtgtggaaaa 1200  
 attaaggtaa atgctggttc aggcttaaaa gataaagccg gtgtaaaatc gcatgaactt 1260  
 gaccgtactc gcattatgga aaacaaaat tattatattg gaaaaattct agagtgtctgg 1320  
 tgcaacgggt ggtaaaaatc tgatggccgc actgattacg ttaaattatt tcttccgatt 1380  
 gcgattcgtt tacgtgaaga taaaactaaa gctaatacat tcgaagatgt atttggtgat 1440  
 tttcatgagg taactggtct aggttctggc agttcagtc atcaccacca tcatcactaa 1500  
 <210> 20  
 <211> 499  
 <212> PRT  
 <213> 人工序列  
 <220>  
 <223> 人工序列的描述：合成的多肽  
 <400> 20  
 Met Ile Leu Lys Ile Leu Asn Glu Ile Ala Ser Ile Gly Ser Thr Lys  
 1                   5                   10                   15  
 Gln Lys Gln Ala Ile Leu Glu Lys Asn Lys Asp Asn Glu Leu Leu Lys  
                   20                   25                   30  
 Arg Val Tyr Arg Leu Thr Tyr Ser Arg Gly Leu Gln Tyr Tyr Ile Lys  
                   35                   40                   45  
 Lys Trp Pro Lys Pro Gly Ile Ala Thr Gln Ser Phe Gly Met Leu Thr

50	55	60
Leu Thr Asp Met Leu Asp Phe Ile Glu Phe Thr Leu Ala Thr Arg Lys		
65	70	75
Leu Thr Gly Asn Ala Ala Ile Glu Glu Leu Thr Gly Tyr Ile Thr Asp		80
	85	90
Gly Lys Lys Asp Asp Val Glu Val Leu Arg Arg Val Met Met Arg Asp		95
	100	105
Leu Glu Cys Gly Ala Ser Val Ser Ile Ala Asn Lys Val Trp Pro Gly		110
	115	120
Leu Ile Pro Glu Gln Pro Gln Met Leu Ala Ser Ser Tyr Asp Glu Lys		125
	130	135
Gly Ile Asn Lys Asn Ile Lys Phe Pro Ala Phe Ala Gln Leu Lys Ala		140
145	150	155
Asp Gly Ala Arg Cys Phe Ala Glu Val Arg Gly Asp Glu Leu Asp Asp		160
	165	170
Val Arg Leu Leu Ser Arg Ala Gly Asn Glu Tyr Leu Gly Leu Asp Leu		175
	180	185
Leu Lys Glu Glu Leu Ile Lys Met Thr Ala Glu Ala Arg Gln Ile His		190
	195	200
Pro Glu Gly Val Leu Ile Asp Gly Glu Leu Val Tyr His Glu Gln Val		205
	210	215
Lys Lys Glu Pro Glu Gly Leu Asp Phe Leu Phe Asp Ala Tyr Pro Glu		220
225	230	235
Asn Ser Lys Ala Lys Glu Phe Ala Glu Val Ala Glu Ser Arg Thr Ala		240
	245	250
Ser Asn Gly Ile Ala Asn Lys Ser Leu Lys Gly Thr Ile Ser Glu Lys		255
	260	265
Glu Ala Gln Cys Met Lys Phe Gln Val Trp Asp Tyr Val Pro Leu Val		270
	275	280
Glu Ile Tyr Ser Leu Pro Ala Phe Arg Leu Lys Tyr Asp Val Arg Phe		285
	290	295
Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile		300
305	310	315
Glu Asn Gln Val Val Asn Asn Leu Asp Glu Ala Lys Val Ile Tyr Lys		320
	325	330
Lys Tyr Ile Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Ile Asp		335
	340	345
Gly Leu Trp Glu Asn Ala Arg Ser Lys Asn Leu Tyr Lys Phe Lys Glu		350
	355	360
		365

Val Ile Asp Val Asp Leu Lys Ile Val Gly Ile Tyr Pro His Arg Lys  
 370 375 380  
 Asp Pro Thr Lys Ala Gly Gly Phe Ile Leu Glu Ser Glu Cys Gly Lys  
 385 390 395 400  
 Ile Lys Val Asn Ala Gly Ser Gly Leu Lys Asp Lys Ala Gly Val Lys  
 405 410 415  
 Ser His Glu Leu Asp Arg Thr Arg Ile Met Glu Asn Gln Asn Tyr Tyr  
 420 425 430  
 Ile Gly Lys Ile Leu Glu Cys Trp Cys Asn Gly Trp Leu Lys Ser Asp  
 435 440 445  
 Gly Arg Thr Asp Tyr Val Lys Leu Phe Leu Pro Ile Ala Ile Arg Leu  
 450 455 460  
 Arg Glu Asp Lys Thr Lys Ala Asn Thr Phe Glu Asp Val Phe Gly Asp  
 465 470 475 480  
 Phe His Glu Val Thr Gly Leu Gly Ser Gly Ser Ser Gly His His His  
 485 490 495

His His His

<210> 21

<211> 1500

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 合成的多核苷酸

<400> 21

atgattctta aaattctgaa cgaaatagca tctattggtt caactaaaca gaagcaagca 60  
 attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct 120  
 cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt 180  
 ggaatgttga ctcttaccga tatgcttgac ttattgaat tcacattagc tactcggaaa 240  
 ttgactggaa atgcagcaat tgaggaatta actggatata tcaccgatgg taaaaaagat 300  
 gatgttgaag ttttgcgtcg agtgatgatg cgagacctg aatgtggtgc ttcagtatct 360  
 attgcaaaca aagtttgcc aggtttaatt cctgaacaac ctcaaagct cgcaagttct 420  
 tatgatgaaa aaggcattaa taagaatata aaatttcag cctttgctca gttaaaagct 480  
 gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta 540  
 tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaatg 600  
 accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attggtatac 660  
 catgagcaag ttaaaaagga gccagaaggc ctagattttc tttttgatgc ttatcctgaa 720  
 aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc 780  
 gccataaat ctttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag 840  
 gtctgggatt atgtcccgtt ggtagaaata tacagtcttc ctgcatttcg tttgaaatat 900

gatgtacgtt tttctaaact agaacaaatg acatctggat atgataaagt aattttaatt 960  
 gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac 1020  
 caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaaa tgctcgttca 1080  
 aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaattgt aggaatttat 1140  
 cctcaccgta aagaccctac taaagcgggt ggatttattc ttgagtcaga gtgtggaaaa 1200  
 attaaggtaa atgctggttc aggcttaaaa gataaagccg gtgtaaaatc gcatgaactt 1260  
 gaccgtactc gcattatgga aaaccaaatt tattatattg gaaaaattct agagtgcgaa 1320  
 tgcaacgggt ggtaaaaatc tgatggccgc actcgttacg ttaaattatt tcttccgatt 1380  
 gcgattcgtt tacgtgaaga taaaactaaa gctaatacat tcgaagatgt atttggatgat 1440  
 tttcatgagg taactggctt aggttctggc agttcaggtc atcaccacca tcatcactaa 1500

<210> 22

<211> 499

<212> PRT

<213> 人工序列

<220>

<223> 人工序列的描述：合成的多肽

<400> 22

Met	Ile	Leu	Lys	Ile	Leu	Asn	Glu	Ile	Ala	Ser	Ile	Gly	Ser	Thr	Lys
1			5						10					15	
Gln	Lys	Gln	Ala	Ile	Leu	Glu	Lys	Asn	Lys	Asp	Asn	Glu	Leu	Leu	Lys
			20						25					30	
Arg	Val	Tyr	Arg	Leu	Thr	Tyr	Ser	Arg	Gly	Leu	Gln	Tyr	Tyr	Ile	Lys
			35					40						45	
Lys	Trp	Pro	Lys	Pro	Gly	Ile	Ala	Thr	Gln	Ser	Phe	Gly	Met	Leu	Thr
			50				55					60			
Leu	Thr	Asp	Met	Leu	Asp	Phe	Ile	Glu	Phe	Thr	Leu	Ala	Thr	Arg	Lys
65					70					75					80
Leu	Thr	Gly	Asn	Ala	Ala	Ile	Glu	Glu	Leu	Thr	Gly	Tyr	Ile	Thr	Asp
					85					90					95
Gly	Lys	Lys	Asp	Asp	Val	Glu	Val	Leu	Arg	Arg	Val	Met	Met	Arg	Asp
					100					105				110	
Leu	Glu	Cys	Gly	Ala	Ser	Val	Ser	Ile	Ala	Asn	Lys	Val	Trp	Pro	Gly
					115					120				125	
Leu	Ile	Pro	Glu	Gln	Pro	Gln	Met	Leu	Ala	Ser	Ser	Tyr	Asp	Glu	Lys
					130					135				140	
Gly	Ile	Asn	Lys	Asn	Ile	Lys	Phe	Pro	Ala	Phe	Ala	Gln	Leu	Lys	Ala
145						150					155				160
Asp	Gly	Ala	Arg	Cys	Phe	Ala	Glu	Val	Arg	Gly	Asp	Glu	Leu	Asp	Asp
					165						170				175

Val Arg Leu Leu Ser Arg Ala Gly Asn Glu Tyr Leu Gly Leu Asp Leu  
 180 185 190  
 Leu Lys Glu Glu Leu Ile Lys Met Thr Ala Glu Ala Arg Gln Ile His  
 195 200 205  
 Pro Glu Gly Val Leu Ile Asp Gly Glu Leu Val Tyr His Glu Gln Val  
 210 215 220  
 Lys Lys Glu Pro Glu Gly Leu Asp Phe Leu Phe Asp Ala Tyr Pro Glu  
 225 230 235 240  
 Asn Ser Lys Ala Lys Glu Phe Ala Glu Val Ala Glu Ser Arg Thr Ala  
 245 250 255  
 Ser Asn Gly Ile Ala Asn Lys Ser Leu Lys Gly Thr Ile Ser Glu Lys  
 260 265 270  
 Glu Ala Gln Cys Met Lys Phe Gln Val Trp Asp Tyr Val Pro Leu Val  
 275 280 285  
 Glu Ile Tyr Ser Leu Pro Ala Phe Arg Leu Lys Tyr Asp Val Arg Phe  
 290 295 300  
 Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile  
 305 310 315 320  
 Glu Asn Gln Val Val Asn Asn Leu Asp Glu Ala Lys Val Ile Tyr Lys  
 325 330 335  
 Lys Tyr Ile Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Ile Asp  
 340 345 350  
 Gly Leu Trp Glu Asn Ala Arg Ser Lys Asn Leu Tyr Lys Phe Lys Glu  
 355 360 365  
 Val Ile Asp Val Asp Leu Lys Ile Val Gly Ile Tyr Pro His Arg Lys  
 370 375 380  
 Asp Pro Thr Lys Ala Gly Gly Phe Ile Leu Glu Ser Glu Cys Gly Lys  
 385 390 395 400  
 Ile Lys Val Asn Ala Gly Ser Gly Leu Lys Asp Lys Ala Gly Val Lys  
 405 410 415  
 Ser His Glu Leu Asp Arg Thr Arg Ile Met Glu Asn Gln Asn Tyr Tyr  
 420 425 430  
 Ile Gly Lys Ile Leu Glu Cys Glu Cys Asn Gly Trp Leu Lys Ser Asp  
 435 440 445  
 Gly Arg Thr Arg Tyr Val Lys Leu Phe Leu Pro Ile Ala Ile Arg Leu  
 450 455 460  
 Arg Glu Asp Lys Thr Lys Ala Asn Thr Phe Glu Asp Val Phe Gly Asp  
 465 470 475 480  
 Phe His Glu Val Thr Gly Leu Gly Ser Gly Ser Ser Gly His His His

	485	490	495
His His His			
<210> 23			
<211> 1500			
<212> DNA			
<213> 人工序列			
<220>			
<223> 人工序列的描述: 合成的多核苷酸			
<400> 23			
atgattctta aaattctgaa cgaaatagca tctattggtt caactaaaca gaagcaagca	60		
attcttgaaa agaataaaga taatgaattg cttaaacgag tatatcgtct gacttattct	120		
cgtgggttac agtattatat caagaaatgg cctaaacctg gtattgctac ccagagtttt	180		
ggaatgttga ctcttaccga tatgcttgac tcattgaat tcacattagc tactcggaaa	240		
ttgactggaa atgcagcaat tgaggaatta actggatata tcaccgatgg taaaaagat	300		
gatgttgaag ttttgcgtcg agtgatgat cgagacctg aatgtggtgc ttcagtatct	360		
attgcaaaca aagtttgcc aggtttaatt cctgaacaac ctcaaatgct cgcaagttct	420		
tatgatgaaa aaggcattaa taagaatata aaatttcag cctttgctca gttaaaagct	480		
gatggagctc ggtgttttgc tgaagttaga ggtgatgaat tagatgatgt tcgtctttta	540		
tcacgagctg gtaatgaata tctaggatta gatcttctta aggaagagtt aattaaatg	600		
accgctgaag cccgccagat tcatccagaa ggtgtgttga ttgatggcga attggtatac	660		
catgagcaag ttaaaaagga gccagaaggc ctagatttcc tttttgatgc ttatcctgaa	720		
aacagtaaag ctaaagaatt cgccgaagta gctgaatcac gtactgcttc taatggaatc	780		
gccaataaat ctttaaaggg aaccatttct gaaaaagaag cacaatgcat gaagtttcag	840		
gtctgggatt atgtcccgtt ggtagaaata tacagtcttc ctgcatttcg tttgaaatat	900		
gatgtacgtt tttctaact agaacaaatg acatctggat atgataaagt aattttaatt	960		
gaaaaccagg tagtaaataa cctagatgaa gctaaggtaa tttataaaaa gtatattgac	1020		
caaggtcttg aaggtattat tctcaaaaat atcgatggat tatgggaaaa tgctcgttca	1080		
aaaaatcttt ataaatttaa agaagtaatt gatgttgatt taaaattgt aggaatttat	1140		
cctcaccgta aagaccctac taaagcgggt ggatttattc ttgagtcaga gtgtggaaaa	1200		
attaaggtaa atgctggttc aggttaaaa gataaagccg gtgtaaaatc gcatgaactt	1260		
gaccgtactc gcattatgga aaacaaaaat tattatattg gaaaaattct agagtgcgaa	1320		
tgcaacgggt ggtaaaaatc tgatggccgc actgattacg ttaaattatt tcttccgatt	1380		
gcgattcgtt tacgtgaaga taaaactgaa gctaatacat tcgaagatgt atttggtgat	1440		
tttcatgagg taactggctc aggttctggc agttcagtc atcaccacca tcatcactaa	1500		
<210> 24			
<211> 499			
<212> PRT			
<213> 人工序列			
<220>			

<223> 人工序列的描述：合成的多肽

<400> 24

Met	Ile	Leu	Lys	Ile	Leu	Asn	Glu	Ile	Ala	Ser	Ile	Gly	Ser	Thr	Lys
1				5					10					15	
Gln	Lys	Gln	Ala	Ile	Leu	Glu	Lys	Asn	Lys	Asp	Asn	Glu	Leu	Leu	Lys
			20					25					30		
Arg	Val	Tyr	Arg	Leu	Thr	Tyr	Ser	Arg	Gly	Leu	Gln	Tyr	Tyr	Ile	Lys
		35					40					45			
Lys	Trp	Pro	Lys	Pro	Gly	Ile	Ala	Thr	Gln	Ser	Phe	Gly	Met	Leu	Thr
	50					55					60				
Leu	Thr	Asp	Met	Leu	Asp	Phe	Ile	Glu	Phe	Thr	Leu	Ala	Thr	Arg	Lys
65					70				75						80
Leu	Thr	Gly	Asn	Ala	Ala	Ile	Glu	Glu	Leu	Thr	Gly	Tyr	Ile	Thr	Asp
				85					90						95
Gly	Lys	Lys	Asp	Asp	Val	Glu	Val	Leu	Arg	Arg	Val	Met	Met	Arg	Asp
			100						105					110	
Leu	Glu	Cys	Gly	Ala	Ser	Val	Ser	Ile	Ala	Asn	Lys	Val	Trp	Pro	Gly
		115						120					125		
Leu	Ile	Pro	Glu	Gln	Pro	Gln	Met	Leu	Ala	Ser	Ser	Tyr	Asp	Glu	Lys
	130						135						140		
Gly	Ile	Asn	Lys	Asn	Ile	Lys	Phe	Pro	Ala	Phe	Ala	Gln	Leu	Lys	Ala
145					150					155					160
Asp	Gly	Ala	Arg	Cys	Phe	Ala	Glu	Val	Arg	Gly	Asp	Glu	Leu	Asp	Asp
				165					170					175	
Val	Arg	Leu	Leu	Ser	Arg	Ala	Gly	Asn	Glu	Tyr	Leu	Gly	Leu	Asp	Leu
		180							185					190	
Leu	Lys	Glu	Glu	Leu	Ile	Lys	Met	Thr	Ala	Glu	Ala	Arg	Gln	Ile	His
		195					200					205			
Pro	Glu	Gly	Val	Leu	Ile	Asp	Gly	Glu	Leu	Val	Tyr	His	Glu	Gln	Val
	210						215					220			
Lys	Lys	Glu	Pro	Glu	Gly	Leu	Asp	Phe	Leu	Phe	Asp	Ala	Tyr	Pro	Glu
225					230					235					240
Asn	Ser	Lys	Ala	Lys	Glu	Phe	Ala	Glu	Val	Ala	Glu	Ser	Arg	Thr	Ala
				245					250					255	
Ser	Asn	Gly	Ile	Ala	Asn	Lys	Ser	Leu	Lys	Gly	Thr	Ile	Ser	Glu	Lys
		260							265					270	
Glu	Ala	Gln	Cys	Met	Lys	Phe	Gln	Val	Trp	Asp	Tyr	Val	Pro	Leu	Val
		275					280						285		
Glu	Ile	Tyr	Ser	Leu	Pro	Ala	Phe	Arg	Leu	Lys	Tyr	Asp	Val	Arg	Phe

290	295	300
Ser Lys Leu Glu Gln Met Thr Ser Gly Tyr Asp Lys Val Ile Leu Ile		
305	310	315
Glu Asn Gln Val Val Asn Asn Leu Asp Glu Ala Lys Val Ile Tyr Lys		
	325	330
Lys Tyr Ile Asp Gln Gly Leu Glu Gly Ile Ile Leu Lys Asn Ile Asp		
	340	345
Gly Leu Trp Glu Asn Ala Arg Ser Lys Asn Leu Tyr Lys Phe Lys Glu		
	355	360
Val Ile Asp Val Asp Leu Lys Ile Val Gly Ile Tyr Pro His Arg Lys		
370	375	380
Asp Pro Thr Lys Ala Gly Gly Phe Ile Leu Glu Ser Glu Cys Gly Lys		
385	390	395
Ile Lys Val Asn Ala Gly Ser Gly Leu Lys Asp Lys Ala Gly Val Lys		
	405	410
Ser His Glu Leu Asp Arg Thr Arg Ile Met Glu Asn Gln Asn Tyr Tyr		
	420	425
Ile Gly Lys Ile Leu Glu Cys Glu Cys Asn Gly Trp Leu Lys Ser Asp		
435	440	445
Gly Arg Thr Asp Tyr Val Lys Leu Phe Leu Pro Ile Ala Ile Arg Leu		
450	455	460
Arg Glu Asp Lys Thr Glu Ala Asn Thr Phe Glu Asp Val Phe Gly Asp		
465	470	475
Phe His Glu Val Thr Gly Leu Gly Ser Gly Ser Ser Gly His His His		
	485	490
		495

His His His

&lt;210&gt; 25

&lt;211&gt; 12

&lt;212&gt; PRT

&lt;213&gt; 人工序列

&lt;220&gt;

&lt;223&gt; 人工序列的描述：合成的肽

&lt;400&gt; 25

Gly Ser Gly Ser Ser Gly His His His His His His

1

5

10

&lt;210&gt; 26

&lt;211&gt; 6

&lt;212&gt; PRT

&lt;213&gt; 人工序列

<220>

<223> 人工序列的描述：合成的6x组氨酸标签

<400> 26

His His His His His His

1

5

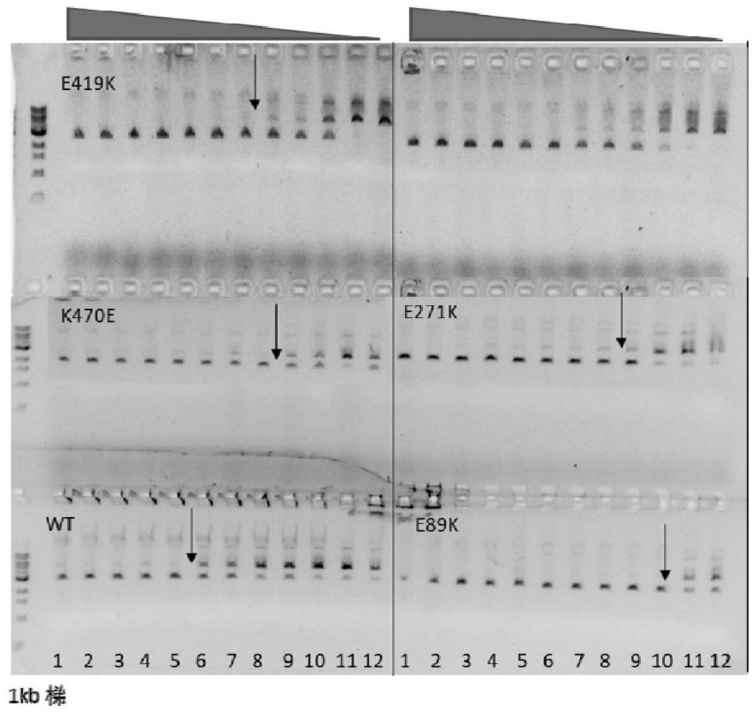


图1A

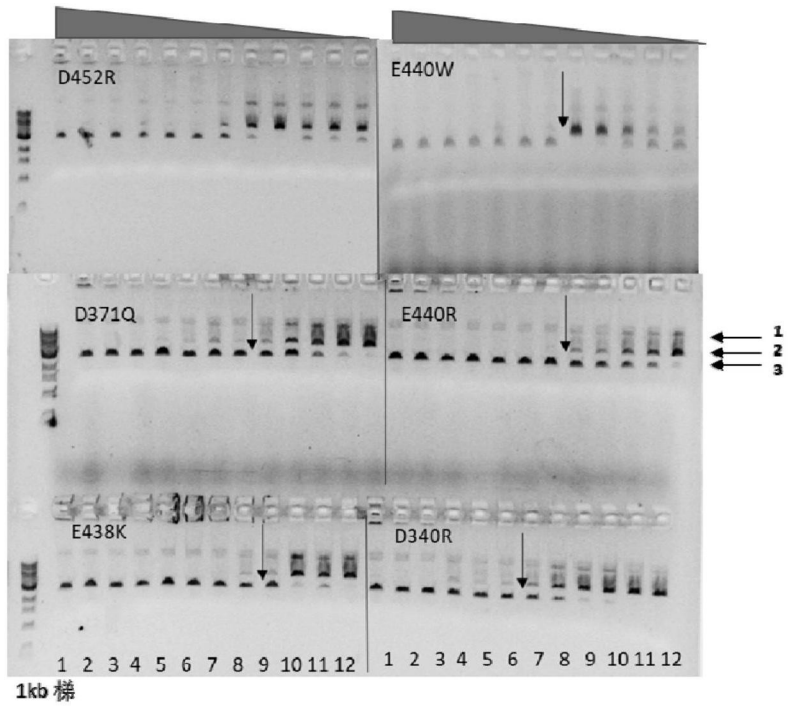


图1B