

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(43) 国际公布日:

2004年12月9日(09.12.2004)

PCT

(10) 国际公布号:

WO 2004/106525 A1

(51) 国际分类号⁷: C12N 15/62, 15/17, 15/12, 15/51, C12P 21/02

新区张江高科技园区李时珍路288号, Shanghai 201203 (CN)。

(21) 国际申请号: PCT/CN2003/000426

(74) 代理人: 上海专利商标事务所(SHANGHAI PATENT & TRADEMARK LAW OFFICE); 中国上海市桂平路435号徐迅, Shanghai 200233 (CN)。

(22) 国际申请日: 2003年6月3日(03.06.2003)

(25) 申请语言: 中文

(81) 指定国(国家): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(26) 公布语言: 中文

(71) 申请人(对除美国以外的所有指定国): 上海新药研究开发中心(SHANGHAI CENTRE OF RESEARCH & DEVELOPMENT OF NEW DRUGS) [CN/CN]; 中国上海市浦东新区张江高科技园区李时珍路288号, Shanghai 201203 (CN)。上海新生源医药研究有限公司(SHANGHAI NEWSUMMIT BIOPHARMA CO., LTD) [CN/CN]; 中国上海市田林路398号4楼A座, Shanghai 200233 (CN)。上海益众生物技术有限公司(SHANGHAI YIZHONG BIOTECHNOLOGY CO., LTD) [CN/CN]; 中国上海市浦东新区张江高科技园区李时珍路288号, Shanghai 201203 (CN)。

(84) 指定国(地区): ARIPO专利(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI专利(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

(72) 发明人:及

(75) 发明人/申请人(仅对美国): 陆怡(LU, Yi) [CN/CN]; 高鑫(GAO, Xin) [CN/CN]; 崔大敷(CUI, Dafu) [CN/CN]; 张友尚(ZHANG, Youshang) [CN/CN]; 中国上海市浦东新区张江高科技园区李时珍路288号, Shanghai 201203 (CN)。黄阳滨(HUANG, Yangbin) [CN/CN]; 中国上海市田林路398号4楼A座, Shanghai 200233 (CN)。孙九如(SUN, Jiuru) [CN/CN]; 李江(LI, Jiang) [CN/CN]; 费俭(FEI, Jian) [CN/CN]; 中国上海市浦东

本国际公布:

— 包括国际检索报告。

所引用双字母代码和其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: A FUSION PROTEIN SUITABLE TO BE EXPRESSED HIGH EFFECTIVELY AND THE PRODUCTION METHOD THEREOF

(54) 发明名称: 一种适合于高效表达的融合蛋白及其生产方法

(57) Abstract: The invention provides a novel fusion protein suitable to be expressed high effectively and the production method thereof. The structure form of the fusion protein is: A-C1-K-(B-C2-K)_nD, wherein A is the upstream peptide; C1 and C2 are each independently a short peptide of 20-40 amino acids of length and contain no Lys in sequence; K is Lys; B is linker; D is downstream peptide; n is integer of 3-30. The fusion protein is able to express high effectively and steadily in host cells. The purification and enzymic cut technology of the expression product is simple, which is able to produce a short peptide high effectively, and has very high value of industrialization.

(57) 摘要

本发明提供了一种适合于高效表达的新型融合蛋白及其生产方法。该融合蛋白结构式是: A-C1-K-(B-C2-K)_n-D, 其中A为上游肽, C1和C2各自独立地是长度为20-40氨基酸且序列中不含Lys的短肽; K是Lys; B是接头; D是下游肽; n为3-30的整数。该融合蛋白能在宿主细胞内高效、稳定地表达。表达产品的纯化、酶切工艺简便, 可以高效地生产短肽, 具有很高的产业化价值。

一种适合于高效表达的融合蛋白及其生产方法

技术领域

本发明涉及了基因工程领域。更具体地，本发明涉及一种适合于高效表达的融合蛋白及其生产方法。该融合蛋白特别适合高效生产不含赖氨酸的短肽。

背景技术

毕赤酵母和大肠杆菌这两种表达系统的适合表达的产物的分子量均在5到200KD左右。对于分子量在1到5KD的短多肽，或者不能表达，或者虽能表达，但通常情况下都存在表达量偏低的问题。

此外，低分子量的短肽表达产物还存在不易进行检测的问题。对于常规的SDS-PAGE电泳来说，常用分子量标准的分子量就在14到65KD。短肽的分子量在检测范围之外(SDS-PAGE电泳凝胶的浓度最高为20%左右，此浓度下检测的分子量线性范围在10到40KD)，因此容易跑出凝胶而流失，无法得到准确的检验结果。

许多常见病可用短肽进行治疗。例如，糖尿病肾病(DN)是糖尿病患者常见的慢性并发症之一，并已成为终末期肾病(ESRD)的主要原因。约35%的1型糖尿病患者和15%的2型糖尿病患者最终将发展为糖尿病肾病。胰岛素原C肽是胰岛β细胞分泌的由31个氨基酸组成的多肽。研究已表明，胰岛素原C肽具有改善糖尿病远期并发症(包括糖尿病肾病)的作用。此外，研究还表明，C肽单独应用或与胰岛素联合应用，可以减少肾小球细胞外基质积聚导致系膜扩张。C肽的这一作用是其治疗糖尿病微血管病变包括糖尿病肾病的主要机制。因此在胰岛素中等控制糖代谢情况下，联合C肽治疗可以更明显改善糖尿病肾病的发生和发展。鉴于糖尿病患者人数众多，因此C肽等短肽药物的需求是很大。

目前国际市场上短肽类药物多为化学合成，价格昂贵。因此，迫切需要工艺简单、成本低廉的大规模生产短肽的基因工程方法。

多拷贝是一种提高短肽表达水平的有效手段。构建多拷贝基因可以表达较大分子量的融合蛋白，达到高效表达和易检测的目的。

对于高效表达的多拷贝融合蛋白或其编码基因，必须处理好以下问题：(a)目的基因(或目的蛋白)的选择；(b).长度(即拷贝数)；(c).连接接头，所述的连接包括融合蛋白与表达质粒的连接、目的蛋白单体之间的连接、信号肽选择等；(d).表达后的酶切等。不恰当的选择都会影响表达水平，甚至使融合蛋白不能表达。

一种常见的方法是将多拷贝的短肽与分子量较大的肽(约10-20KD)融合，从而使表达的融合蛋白能稳定地在宿主细胞中存在。这种方法的缺点是，所需的短肽在融合蛋白中所占比例较小，因此效率不够高。此外，与GST等形成的

融合蛋白在纯化时多一步酶切和纯化，工艺较复杂。

另一种理想的方法是在多拷贝的短肽的上游和下游添加长度较短的上游序列和下游序列，这样可以显著提高短肽在融合蛋白中所占比例。但是目前用这种方法的所构建的融合蛋白大多不稳定，难以获得大量的表达产物。

5 因此，本领域迫切需要开发一种高效、稳定表达多拷贝短肽的方法。

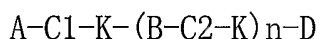
发明内容

本发明的目的就是提供一种高效、稳定表达多拷贝短肽的方法。

本发明的另一目的是提供有关的融合蛋白、载体和宿主细胞。

10

在本发明的第一方面，提供了一种融合蛋白，其结构式是：



其中A为MHHHHHHRK(SEQ ID NO:4)，

C1和C2各自独立地是长度为20-40氨基酸且序列中不含Lys的短肽；

15

K是Lys；

B是AGSK(SEQ ID NO:5)；

D是长度为3-15个氨基酸且前3个氨基酸为AGS的下游肽；n为3-30的整数。

在另一优选例中，所述的C1和C2是相同的，更佳地C1和C2都是胰岛素原C肽。

在另一优选例中，所述的C1和C2是不同的，且两者的长度相差8-10个氨基酸。

20

在另一优选例中，所述的C1和C2选自：胰岛素原C肽、 α -人心房钠尿肽、乙肝前S1抗原肽。

在另一优选例中，所述融合蛋白具有SEQ ID NO:2或3所示的氨基酸序列。

在本发明第二方面，提供了一种DNA分子，它编码本发明上述的融合蛋白。

在本发明第三方面，提供了一种表达载体，它含有本发明上述的DNA分子。

25

在本发明的第四方面，提供了一种宿主细胞，它含有本发明所述的表达载体，或者在其基因组中整合有本发明所述的DNA分子。较佳地，所述的宿主细胞是毕赤酵母或大肠杆菌。

在本发明的第五方面，提供了一种融合蛋白的生产方法，包括步骤：

(a)培养本发明所述的宿主细胞，从而表达所述的融合蛋白；

30

(b)分离融合蛋白。

在本发明的第六方面，提供了一种生产短肽的方法，包括步骤：

(c)用胰蛋白酶和羧肽酶B酶切本发明所述的融合蛋白，产生短肽C1和C2；

(d)分离短肽C1和C2。

较佳地，在所述的方法，所述的C1和C2是相同的，都是胰岛素原C肽。

35

附图说明

图1. 毕赤酵母表达人胰岛素原C肽融合蛋白质粒构建图。

图2. 毕赤酵母表达人胰岛素原C肽融合蛋白SDS-PAGE电泳图。泳道a: 分子量标准(从上到下分子量为96KD、66KD、43KD、31KD、20.1KD、14.4KD); b: 空白对照; c: 培液上清(C肽融合蛋白)。

图3. 大肠杆菌表达人胰岛素原C肽融合蛋白SDS-PAGE电泳图(摇瓶鉴定)。泳道a、b、c、d分别表示诱导时间为0小时、1小时、2小时和3小时(其中目的蛋白含量分别是:0%, 11.47%, 11.88%, 12.36%); e: 分子量标准(从上到下分子量为96KD、66KD、43KD、31KD、20.1KD、14.4KD);

图4. 毕赤酵母发酵表达人胰岛素原C肽融合蛋白SDS-PAGE电泳图。泳道a: 分子量标准(从上到下分子量为96KD、66KD、43KD、31KD、20.1KD、14.4KD); b. 诱导前; c-g: 诱导8、16、24、32、48hr。

图5. 大肠杆菌发酵表达人胰岛素原C肽融合蛋白SDS-PAGE电泳图。泳道a、b、c、d分别表示诱导时间为0小时、1小时、2小时和3小时(其中目的蛋白含量分别是:0%, 24.87%, 27.89%, 30.33%); e: 分子量标准(从上到下分子量为96KD、66KD、43KD、31KD、20.1KD、14.4KD)。

图6显示了酶切后C肽的HPLC分析结果。

图7显示了纯化后C肽的HPLC分析结果。

具体实施方式

本发明人经过广泛而深入的研究发现, 在多拷贝短肽上游添加上游肽MHSHHHHRSK (SEQ ID NO:4) 所形成的A-C1-K-(B-C2-K)_n-D的结构结合了多种优势, 可以使融合蛋白在毕赤酵母或大肠杆菌中高效表达。在此基础上完成了本发明。

融合蛋白的结构

本发明融合蛋白为非天然的形式, 整个融合蛋白中各元件线性连接且结构上一般不存在二硫键, 因此融合蛋白呈线形结构, 没有空间折叠, 易被各种蛋白酶酶切。

(i) 上游肽

在本发明通式为A-C1-K-(B-C2-K)_n-D的融合蛋白中, 对稳定高效表达起关键作用的元件是上游肽A, 即MHSHHHHRSK (SEQ ID NO:4)。它使得含多拷贝元件C1-K-(B-C2-K)_n的融合蛋白不仅能够表达, 而且能够稳定存在于宿主细胞中。

此外, 由于上游元件的很短, 因此融合蛋白中目的蛋白的比例很高。

(ii)短肽

在本发明融合蛋白中，C1和C2各自独立地是长度为10-40氨基酸的短肽。由于在将融合蛋白切割为短肽时，使用针对碱性氨基酸如赖氨酸(Lys)的酶(如胰蛋白酶、羧肽酶B等)，因此C1和C2的序列中不可含有碱性氨基酸如Lys和Arg。合适的例子包括胰岛素原C肽(31aa)、 α -人心房钠尿肽(28aa)、乙肝前S1抗原肽(28aa)等。

C1和C2可相同，也可不同。当两者相同时，融合蛋白被切割后就形成长度为20-40氨基酸的所需的短肽，以及长度为4-15氨基酸的接头、上游肽和下游肽。由于长度不同，因此可以方便地纯化出短肽。一种优选情况是C1和C2都是胰岛素原C肽。

当C1和C2不同时，两者的长度宜相差8-10个氨基酸，以便在切割后将C1和C2分离开。当然，如果C1和C2是可以联用的两种短肽，那么就不一定需要将两者分开，此时C1和C2可以没有长度差异。

在本发明融合蛋白中，n为3-30的整数，较佳地，n为5-25的整数，更佳地为8-15的整数。此时，融合蛋白的分子量约为6-60KD，适合毕赤酵母和大肠杆菌表达系统。当然，对于其他适合表达系统，n的上限可以超过20，例如25，30，35，40甚至更高。

(iii)接头

在本发明融合蛋白中，B为接头。接头的长度通常为4-10个氨基酸。长度越短，则短肽在融合蛋白中的比例越高，而且短肽与接头的长度差越大从而更有利于分离，因此短接头更适用于本发明。此外，接头的羧基端应为Lys或Arg。一种优选的接头是AGSK(SEQ ID NO:5)。

(iv)下游肽

在本发明融合蛋白中，D是下游肽，其作用是与下游肽之前的Lys一起提供羧肽酶的酶切位点。有时，下游肽的编码序列还可提供用于克隆操作的限制性位点。

下游肽是可有可无的。当B为AGSK(SEQ ID NO:5)时，D宜为长度3-15个氨基酸，且前3个氨基酸为AGS的下游肽。合适的例子包括(但不限于)：AGSLNSLGRPRINS(SEQ ID NO:6)、AGSLNSP(SEQ ID NO:7)。

融合基因、载体和宿主细胞的构建

在本发明中，一旦确定了融合蛋白的序列，则融合基因、载体和宿主细胞的构建都可按常规方法进行。适用于本发明的载体和宿主细胞没有特别限制。

通常，先根据本发明融合蛋白的氨基酸序列设计出编码DNA。较佳地，可

根据选用的宿主细胞如毕赤酵母或大肠杆菌，进行密码子优化。

然后，选择表达载体，如毕赤酵母系统的pPIC9或pPIC9K，或大肠杆菌系统的pET30(a)等，并且根据选定的载体，设计酶切位点。

接着，全合成所设计的融合蛋白编码序列，酶切后克隆入表达载体。

- 5 最后，用常规方法将表达载体转染毕赤酵母或转化大肠杆菌，经过抗性筛选等方法，就可获得高表达的宿主细胞。

融合蛋白的生产和纯化

- 10 对于转化的表达融合蛋白的宿主细胞，可用常规方法和条件进行培养，从而表达出融合蛋白，并用常规方法(如盐析方法、离心、分子筛层析、吸附层析、离子交换层析、HPLC等)进行分离纯化。

以毕赤酵母培养与表达为例，合适的培养基包括(但不限于)以下培养基：

YPD平板	酵母提取物	1%
	蛋白胨	2%
	葡萄糖	2%
	琼脂	1%
BMGY培养基	酵母提取物	1%
	蛋白胨	2%
	磷酸钾缓冲液, pH 6.0	100mM
	YNB	1.34%
	生物素	$(4 \times 10^{-5})\%$
	甘油	1%
MM培养液	YNB	1.34%
	生物素	$(4 \times 10^{-5})\%$
	甲醇	0.5%

- 15 一种优选的摇瓶培养条件是：挑取YPD平板上的单克隆，以BMGY为摇瓶培养液，培养至OD₆₀₀2-20，加入1%浓度甲醇诱导表达，诱导24小时后，培液上清的目的蛋白浓度可达到50-200mg/L。

为了中试和大规模生产，需要对P.Pastoris工程细胞的中试表达与表达条件进行优化。中试发酵指的是发酵罐发酵。优选的生产条件如下：

- 20 1.对于培养基的选择而言，发酵罐发酵时可选择低盐培养基，也可在低盐培养基基础上进行一定更改，但所含离子成份与低盐培养基相似。

低盐发酵培养基	H ₃ PO ₄ (85%储存浓度)	26.7ml/L
	CaSO ₄ •2H ₂ O	0.93g/L
	K ₂ SO ₄	18.2g/L
	MgSO ₄	7.27g/L
	KOH(固体)	4.13g/L
	柠檬酸钠.2H ₂ O	1.47g/L
	PTM ₁	2ml/L
	甘油	4%(w/v)
PTM ₁	CuSO ₄ .5H ₂ O	6g/L
	KI	0.08g/L
	MnSO ₄ .H ₂ O	3g/L
	钼酸钠	0.2g/L
	硼酸	0.02g/L
	CaSO ₄ .2H ₂ O	0.5g/L
	ZnCl ₂	20g/L
	氯化钴	0.5g/L
	FeSO ₄	65g/L
	生物素	0.2g/L
	H ₂ SO ₄	5ml/L

2. 就温度的控制而言, 发酵及诱导温度保持在25-31℃。

3. 就诱导期的pH值而言, 诱导期pH控制在4-8, 较佳地为pH5-7。

5 4. 就溶氧(DO)的控制而言 DO控制在20-90%。溶氧的维持可以用氧气/空气混合气体的通入来解决。

5. 就补料的流加而言, 补料种类宜包括甘油、甲醇、葡萄糖等糖原, 可单独补料或混合补料。

10 6. 就PTM₁(混合微量元素)的量: 起始培养阶段PTM₁的量为1-4ml/L培液, 在补料阶段加入量为2-20ml/L补料。

7.就诱导期甲醇浓度而言, 常规诱导浓度都可用于本发明, 通常甲醇浓度控制在0.5-2%。

8.就诱导时间而言, 通常为10-100小时, 较佳地为20-60小时。

15 短肽单体的制备

本发明的融合蛋白是制备短肽单体的优良原料。经过酶切和分离, 就可得

到短肽单体。

通常，所用的酶是针对碱性氨基酸的酶，如胰蛋白酶、羧肽酶B等。酶切的方法可用二步法，即先酶切形成带K的短肽(如C1K或C2K)，再将K切除，得到短肽单体。也可使用单步法，即用双酶混合酶切，直接获得短肽单体。

- 5 酶切的条件没有特别限制，可采用常规条件。例如
融合蛋白浓度是0.1-100mg/ml，较佳地1-50mg/ml；
酶与蛋白比例是1：10-1：5000，较佳地，1：50-1：500；
酶切温度是15-35℃，较佳地20-30℃；
酶切pH是5.0-10.0,较佳地6.0-9.0；

- 10 缓冲体系是磷酸盐缓冲液、醋酸盐缓冲液、Tris、或碳酸盐缓冲液，缓冲液浓度10-200mM；
酶切时间是1-50小时，较佳地1-10小时。

- 15 对于酶切混合物，可用常规技术(如阳离子交换层析、阴离子交换层析、凝胶过滤层析、疏水层析、反相层析、亲和层析等)纯化短肽单体。优选方法是酶切后的样品进行离子交换层析、疏水层析和反相层析。

- 对于制备的短肽单体，可用常规方法将其制成相应的制剂。例如当短肽是胰岛素原C肽时，可在纯化后的原液加入适当辅料，如1-10%甘露醇(或蔗糖、乳糖等)稳定剂，同时还可添加表面活性剂、抗氧化剂等保护剂，及适当缓冲液，
20 分装，制成C肽水溶液剂，或者冷冻抽干，制成C肽的注射用粉针剂。

本发明的主要优点在于：

- i. 表达量高且稳定；
ii. 短肽(如人胰岛素原C肽)的产量高；
25 iii. 制备方法简单，易纯化所以成本低；
iv. 目的蛋白若用真核细胞表达系统-毕赤酵母进行表达，与哺乳动物有较强同源性，产生的毒副作用较少。

- 下面结合具体实施例，进一步阐述本发明。应理解，这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件，例如Sambrook等人，分子克隆：实验室手册(New York：Cold Spring Harbor Laboratory Press, 1989)中所述的条件，或按照制造厂商所建议的条件。
30

实施例1 毕赤酵母表达人胰岛素原C肽融合蛋白-工程细胞的构建方法

- 35 在本实施例中，选择C1=C2=胰岛素原C肽，n=8。

经过筛选,也得到高效表达株(经检验,其基因拷贝整合数 ≥ 20 个)。BMGY培养液摇瓶试验表明,诱导24小时后,培养上清目的蛋白占总蛋白50%以上,表达水平100ug/ml以上。根据其消耗甲醇的速度,判断此工程细胞株为Mut⁺。

5 实施例3 毕赤酵母表达人胰岛素原C肽融合蛋白-工程细胞的构建方法

在本实施例中,比较了毕赤酵母表达人胰岛素原C肽融合蛋白的几种方案。

方案1是实施例1。

方案2中融合蛋白的结构是:(His)₆-Lys-C肽-Lys-C肽-Lys。

方案3中融合蛋白的结构与实施例1基本相同,不同点仅在于n=4。

10 结果表明,方案2表达目的产物得到的量很低,难以检测出来。这表明,6His的上游肽以及连接肽的设计形成一个有利于高表达和分泌的构象。方案3得到了所需的表达蛋白。

15 实施例4. 大肠杆菌表达人胰岛素原C肽融合蛋白-工程细胞的构建方法

选择合适的表达载体和宿主菌系统: 表达载体质粒pET-30a(+)来自Novagen公司,受体菌为大肠杆菌BL21(DE3)。

实施例1中的融合蛋白基因序列经查证没有大肠杆菌特别稀有的密码子,将该基因克隆到pUC19中测序验证后,克隆入表达载体pET-30a(+)中。测序确认没有表达框改变后,转化入大肠杆菌BL21(DE3),获得工程菌。

20 接种确证的工程菌株于10mL LB(卡那霉素浓度为100μg/mL)培养液中,在37℃,250rpm条件下培养至OD₆₀₀=0.5-1.0,此时细胞处于对数生长期;无菌条件下加入IPTG至终浓度为1mM,继续在37℃,250rpm条件下诱导表达,3小时后取样,SDS-PAGE检测,在41KD处有目的蛋白表达(图3)。

25 实施例5. 毕赤酵母表达人胰岛素原C肽融合蛋白的中试发酵

实施例1所得的毕赤酵母工程细胞用以下方法发酵:

NBS BioFlo 3000发酵罐(5L):

1. 种子液: BMGY培养基

2. 发酵培养基: 低盐发酵培养基

30 3. 发酵培养阶段:

温度: 30 °C

pH: 5.0

DO: $\geq 30\%$

甲醇诱导前湿细胞重: 88g/L

35 4. 发酵诱导阶段:

温度: 30 °C

pH: 6.5

DO: $\geq 20\%$

甲醇补料中PTM₁量10ml/L

5 甲醇诱导结束湿细胞重: 257g/L

在培养阶段后期, 甘油被耗尽后流加甲醇进行诱导, 整个诱导过程甲醇浓度控制在0.5-1%, 诱导时间20-50 小时。

诱导48小时后, 人胰岛素原C肽融合蛋白表达量占发酵上清液总蛋白47%, 表达量达1mg/ml (用改良Lowry法测) (图4)。

10

实施例 6大肠杆菌表达人胰岛素原C肽融合蛋白的中试发酵

实施例4所得的大肠杆菌工程细胞用以下方法发酵:

NBS BioFlo 3000发酵罐(5L):

1. 种子液: LB培养基

15 2. 发酵培养基: M9培养基(3.5L)

3. 发酵培养阶段:

温度: 37 °C

pH: 7.0

DO: 80%

20 诱导前OD₆₀₀: 1.52

4. 发酵诱导阶段:

温度: 35 °C

pH: 6.5

DO: 50%

25 葡萄糖补料中微量元素母液(PTM)的量: 10ml/L

IPTG诱导结束OD₆₀₀: 5.72

在培养阶段后期, OD₆₀₀上升到1.5左右, 一次性加入IPTG进行诱导。IPTG终浓度为1mM。整个诱导过程通过流加葡萄糖补料来控制在pH在6.9到7.2范围内, 诱导时间3 小时。

30 诱导3小时后, 离心, 得到25.7克菌体。人胰岛素原C肽融合蛋白表达量占发酵菌体总蛋白35%, 蛋白含量为1.02mg/ml (用改良Lowry法测)。目的蛋白总量为1.25克, 表达量达0.357g/L发酵液(图5)。

实施例7. 毕赤酵母表达人胰岛素原C肽融合蛋白—酶切

35 实施例5所得的毕赤酵母工程细胞表达产物上清经初步纯化, 进行单步法

酶切，酶切条件如下：

所用的酶：TPCK-胰蛋白酶(Sigma)、羧肽酶 B(Sigma公司)

缓冲体系 10mM PB(pH7.4)

C肽融合蛋白与TPCK-胰蛋白酶比例 200: 1(W/W)

5 C肽融合蛋白与羧肽酶 B 比例 200: 1(W/W)

酶切温度：25℃

酶切时间：6小时。

酶切得到的C肽经HPLC分析，HPLC鉴定条件如下：

10 C18分析柱 Φ 4.6*250mm

λ : 230nm

缓冲液 A: 0.1%TFA 缓冲液 B: 0.08%TFA 70%乙腈

梯度:	0-5min	0% 缓冲液B
	5-10min	30% 缓冲液B
	10-30min	60% 缓冲液B
	30-31min	100% 缓冲液B

15 结果如图6所示，酶切的主要产物的保留时间为 27 ± 1 min，与C肽单体的保留时间相符，这表明酶切后产生了C肽单体。

实施例8. 酶切产物纯化

实施例7所得的酶切产物经过纯化，具体条件如下：

层析介质：Q-Sepharose FF (Pharmacia)

20 缓冲液 A: 10mM PB 缓冲液 B: 10mMPB + 1M NaCl

梯度：0-100% 缓冲液B 20柱体积

流速：60cm/h

25 将实施例7所得的酶切产物超滤或是透析以脱盐和去少部分杂质。然后，将此样品上样，上样完毕后用缓冲液 A冲洗至UV基线平稳，而后使用线形梯度，用20CV使B%从0-100。在20%B浓度左右出现样品，分管收集。样品回收率达到60%以上，纯度达到95%以上。

由于此样品盐浓度较高，将收集的样品使用G-50 Sephadex脱盐，使用缓冲液A为缓冲液，样品回收率在90%以上，HPLC纯度在98%以上(图7)。

30 样品经鉴定，其序列与天然人胰岛素原C肽完全相同，证实获得的是人胰岛素原C肽。

实施例9 纯化产物的制剂

对实施例8中获得的纯化后原液加入10%甘露醇，分装，冷冻抽干，制成C肽的注射用粉针剂。经稳定性试验表明，该粉针剂稳定。经活性测定表明，本发明方法制得的人胰岛素原C肽的活性与标准品相同。

5

工业实用性

综上所述，通过毕赤酵母等表达系统来表达本发明的融合蛋白，其优点在于：高表达、高稳定，而且制得的融合蛋白经简单酶切和分离后即可C肽等短肽单体。本发明方法使工艺复杂程度大大降低，产率得到较大提高，特别适合大规模生产C肽等短肽。

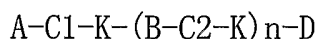
10

在本发明提及的所有文献都在本申请中引用作为参考，就如同每一篇文献被单独引用作为参考那样。此外应理解，在阅读了本发明的上述讲授内容之后，本领域技术人员可以对本发明作各种改动或修改，这些等价形式同样落于本申请所附权利要求书所限定的范围。

15

权 利 要 求

1. 一种融合蛋白，其特征在于，其结构式是：



5 其中A为MHHHHHRSK (SEQ ID NO:4)，

C1和C2各自独立地是长度为20-40氨基酸且序列中不含Lys的短肽；

K是Lys；

B是AGSK (SEQ ID NO:5)；

D是长度为3-15个氨基酸且前3个氨基酸为AGS的下游肽；

10 n为3-30的整数。

2. 如权利要求1所述的蛋白，其特征在于，所述的C1和C2是相同的。

3. 如权利要求1所述的蛋白，其特征在于，所述的C1和C2是不同的，且两者的长度相差8-10个氨基酸。

4. 如权利要求1所述的蛋白，其特征在于，所述的C1和C2选自：胰岛素原C肽、
15 α -人心房钠尿肽、乙肝前S1抗原肽。

5. 如权利要求1所述的蛋白，其特征在于，具有SEQ ID NO:2或3所示的氨基酸序列。

6. 一种DNA分子，其特征在于，它编码权利要求1所述的融合蛋白。

7. 一种表达载体，其特征在于，它含有权利要求6所述的DNA分子。

20 8. 一种宿主细胞，其特征在于，它含有权利要求7所述的表达载体，或者在其基因组中整合有权利要求6所述的DNA分子。

9. 如权利要求8所述的宿主细胞，其特征在于，所述的宿主细胞是毕赤酵母或大肠杆菌。

10. 一种融合蛋白的生产方法，其特征在于，包括步骤：

25 (a)培养权利要求6所述的宿主细胞，从而表达所述的融合蛋白；

(b)分离融合蛋白。

11. 一种生产短肽的方法，其特征在于，包括步骤：

(c)用胰蛋白酶和羧肽酶B酶切权利要求1所述的融合蛋白，产生短肽C1和
C2；

30 (d)分离短肽C1和C2。

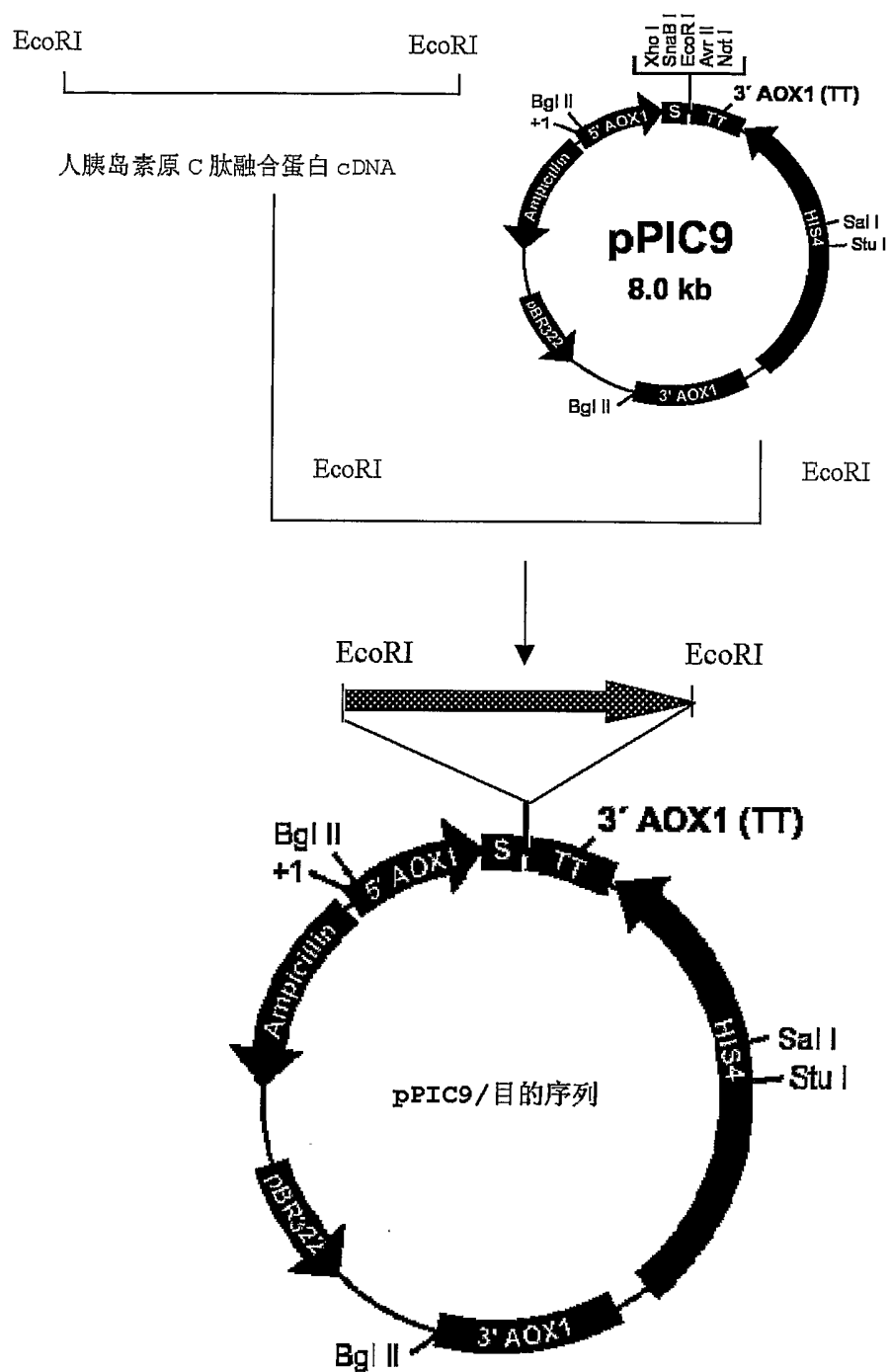


图 1

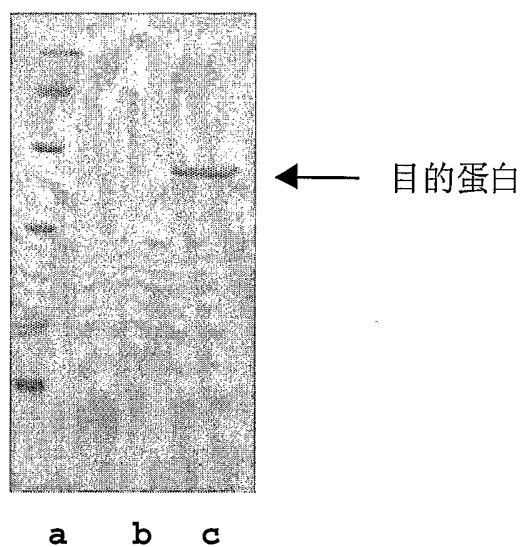


图 2

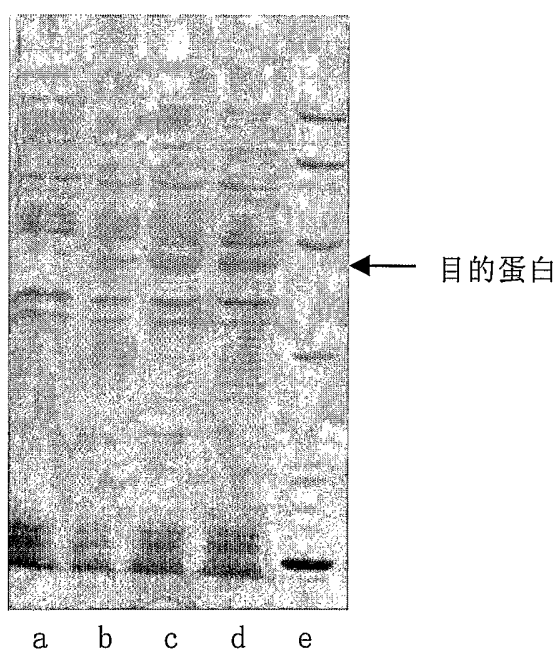


图 3

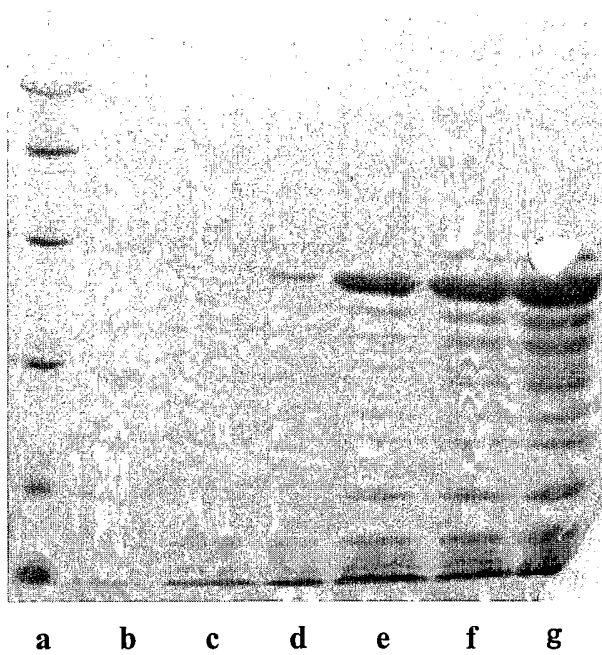


图 4

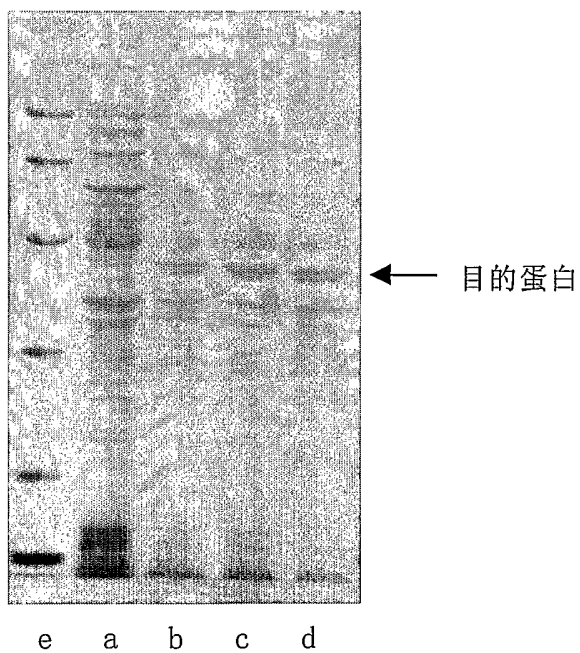


图 5

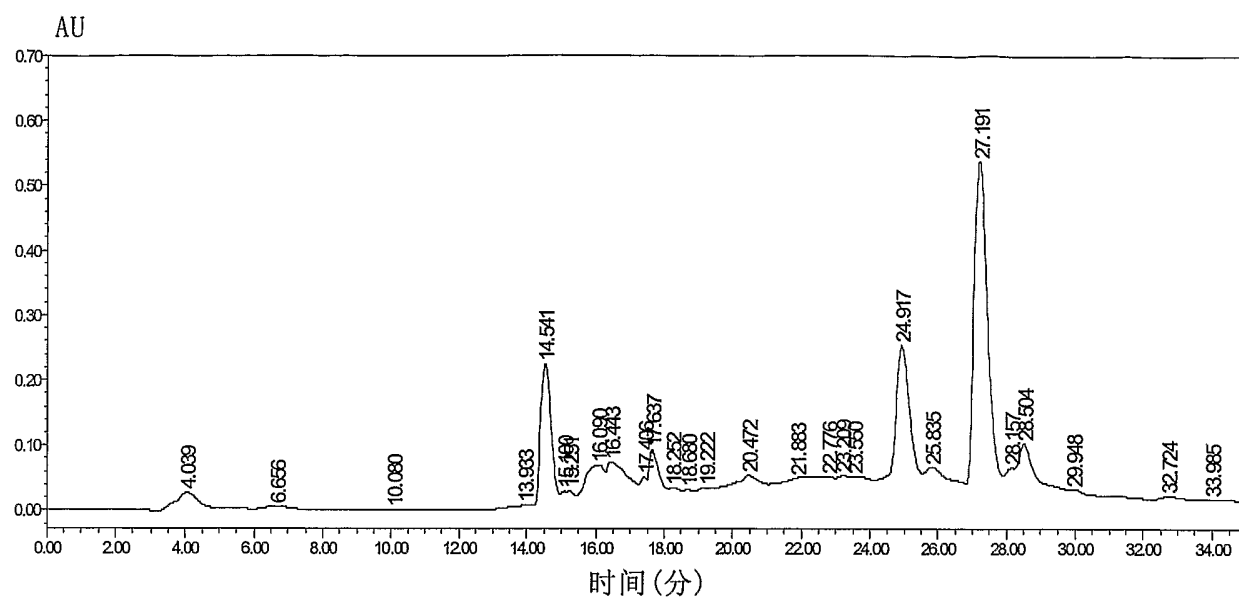


图 6

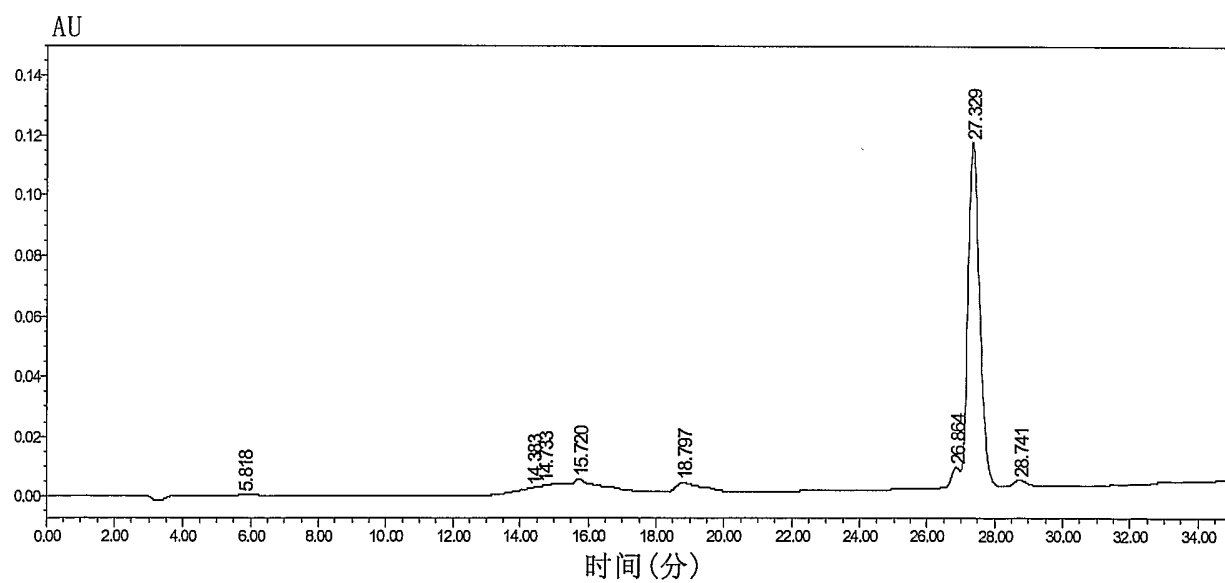


图 7

序列表

<110> 上海新药研究开发中心
上海新生源医药研究有限公司
上海益众生物技术有限公司

<120> 一种适合于高效表达的融合蛋白及其生产方法

<130> 030255

<160> 7

<170> PatentIn version 3.1

<210> 1

<211> 1056

<212> DNA

<213> 人工序列

<400> 1

gaggctgaag cttacgtaga attcatgcat catcatcatc atcatagatc taaggaagct	60
gaagatttgc aagttggtca agttgaattg ggtggtgggc ccggtgctgg ttctttgcaa	120
ccattggcctt tggaagggttc ttgcaaaaag gctggatcta aggaagctga agatttgcaa	180
gttgggtcaag ttgaattggg tgggtggccc ggtgctggtt ctttgcaacc attggctttg	240
gaaggttctt tgcaaaaggc tggatctaag gaagctgaag atttgcaagt tggatcaagtt	300
gaattgggtg gtgggcccgg tgctggttct ttgcaaccat tggctttgga aggttctttg	360
caaaaggctg gatctaagga agctgaagat ttgcaagttg gtcaagttga attgggtggt	420
gggcccgggtg ctggttcttt gcaaccattg gctttggaag gttctttgca aaaggctgga	480
tctaaggaag ctgaagattt gcaagttggt caagttgaat tgggtggtgg gcccggtgct	540
ggttctttgc aaccattggc ttggaaggt tctttgcaa aggctggatc taaggaagct	600
gaagatttgc aagttggtca agttgaattg ggtggtgggc ccggtgctgg ttctttgcaa	660
ccattggcctt tggaagggttc ttgcaaaaag gctggatcta aggaagctga agatttgcaa	720
gttgggtcaag ttgaattggg tgggtggccc ggtgctggtt ctttgcaacc attggctttg	780
gaaggttctt tgcaaaaggc tggatctaag gaagctgaag atttgcaagt tggatcaagtt	840
gaattgggtg gtgggcccgg tgctggttct ttgcaaccat tggctttgga aggttctttg	900
caaaaggctg gatctaagga agctgaagat ttgcaagttg gtcaagttga attgggtggt	960
gggcccgggtg ctggttcttt gcaaccattg gctttggaag gttctttgca aaaggctgga	1020
tccctgaatt ccctagggcg gccgcgaatt aattcg	1056

<210> 2
 <211> 344
 <212> PRT
 <213> 人工序列

<400> 2

Met His His His His His His Arg Ser Lys Glu Ala Glu Asp Leu Gln
 1 5 10 15

Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln
 20 25 30

Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala
 35 40 45

Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala
 50 55 60

Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly
 65 70 75 80

Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly
 85 90 95

Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu
 100 105 110

Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val
 115 120 125

Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu
 130 135 140

Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln
 145 150 155 160

Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln
 165 170 175

Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala
 180 185 190

Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala
 195 200 205

Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly
 210 215 220

Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly
 225 230 235 240

Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu
 245 250 255

Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val
 260 265 270

Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu
 275 280 285

Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln
 290 295 300

Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln
 305 310 315 320

Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly Ser Leu Asn Ser
 325 330 335

Leu Gly Arg Pro Arg Ile Asn Ser
 340

<210> 3
 <211> 445
 <212> PRT
 <213> 人工序列

<400> 3

Met His His His His His His Arg Ser Lys Glu Ala Glu Asp Leu Gln
 1 5 10 15

Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln
 20 25 30

Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala
 35 40 45

Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala
 50 55 60

Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly
 65 70 75 80

Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly
 85 90 95

Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu
 100 105 110

Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val
 115 120 125

Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu
 130 135 140

Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln
 145 150 155 160

Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln
 165 170 175

Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala
 180 185 190

Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala
 195 200 205

Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly
 210 215 220

Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly
 225 230 235 240

Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu
 245 250 255

Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val
 260 265 270

Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu

275 280 285
 Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln
 290 295 300
 Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln
 305 310 315 320
 Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly Ser Lys Glu Ala
 325 330 335
 Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly Gly Pro Gly Ala
 340 345 350
 Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu Gln Lys Ala Gly
 355 360 365
 Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val Glu Leu Gly Gly
 370 375 380
 Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu Glu Gly Ser Leu
 385 390 395 400
 Gln Lys Ala Gly Ser Lys Glu Ala Glu Asp Leu Gln Val Gly Gln Val
 405 410 415
 Glu Leu Gly Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu Ala Leu
 420 425 430
 Glu Gly Ser Leu Gln Lys Ala Gly Ser Leu Asn Ser Pro
 435 440 445

<210> 4
 <211> 10
 <212> PRT
 <213> 人工序列

<220>
 <221> MISC_FEATURE
 <222> (1)..(10)
 <223> 上游肽

<400> 4

Met His His His His His His Arg Ser Lys

1 5 10

<210> 5
<211> 4
<212> PRT
<213> 人工序列

<220>
<221> MISC_FEATURE
<222> (1).. (4)
<223> 接头

<400> 5

Ala Gly Ser Lys
1

<210> 6
<211> 14
<212> PRT
<213> 人工序列

<220>
<221> MISC_FEATURE
<222> (1).. (14)
<223> 下游肽

<400> 6

Ala Gly Ser Leu Asn Ser Leu Gly Arg Pro Arg Ile Asn Ser
1 5 10

<210> 7
<211> 7
<212> PRT
<213> 人工序列

<220>
<221> MISC_FEATURE
<222> (1).. (7)
<223> 下游肽

<400> 7

Ala Gly Ser Leu Asn Ser Pro
1 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN03/00426

A. CLASSIFICATION OF SUBJECT MATTER

IPC⁷ : C12N15/62,15/17,15/12,15/51,C12P21/02

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched(classification system followed by classification symbols)

IPC⁷ : C12N15/62,15/17,15/12,15/51,C12P21/02

Documentation searched other than minimum documentation to the extent that such documents are included in the field searched

Electronic data base consulted during the international search(name of data base and, where practicable, search terms used)

WPI,EPODOC,PAJ,CNPAT

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant claim No.
A	WO0017336 A1; LEE S,ET AL, C12N15/00,C12N15/62, 2000/03/30 See the abstract.	1-11
A	WO9711186 A1; BOEHRINGER MANNHEIM GMBH,ET AL, C12N15/62, 1997/03/27 ; See the abstract.	1-11
A	WO9310152 A1; SMIK,ET AL, C07K15/00,C12N15/62, 1993/05/27 See the abstract.	1-11
A	DE4012818 A; HOECHST AG, C12N15/62, 1991/10/24 See the abstract.	1-11
A	EP0281418 A; SUNTORY LTD, C12N15/62, 1988/09/07 See the abstract.	1-11

☐ Further documents are listed in the continuation of Box C.
 ☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X"document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document but published on or after the international filing date	"Y"document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason(as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 24 June 2003 (24.06. 03)	Date of mailing of the international search report 10 JUL 2003 (10.07.03)
Name and mailing address of the ISA/ The Chinese Patent Office 6, Xitucheng Road, Haidian District, Beijing, 100088, China Facsimile No. 86-010-62019451	Authorized officer ZENG,Fanhui Telephone No. 62093733

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN03/00426

Patent document cited in search report	Publication date	Patent family members	Publication date
WO0017336 A1	2000/03/30	KR2000023309 A AU5762699 A	2000/04/25 2000/04/10
WO9711186 A1	1997/03/27	JP11511333 T DE19535445 A1 AU7130096 A EP0851930 A1	1999/10/05 1997/03/27 1997/04/09 1998/07/08
WO9310152 A1	1993/05/27	KR251505 B1 AU2927892 A EP0614465 A1 JP7501213 T US5928902 A	2000/05/01 1993/06/15 1994/09/14 1995/02/09 1999/07/27
DE4012818 A	1991/10/24	CA2040810 C EP0453969 A AU7515491 A BR9101587 A CN1055952 A US5426036 A	2001/07/24 1991/10/30 1991/10/24 1991/12/10 1991/11/06 1995/06/20
EP0281418 A	1988/09/07	CA1340830 C AU1272488 A JP1010999 A US4987070 A HU210625 B	1999/11/30 1988/09/15 1989/01/13 1991/01/22 1995/06/28

国际检索报告

国际申请号

PCT/CN03/00426

A. 主题的分类

Int.Cl⁷: C12N15/62,15/17,15/12,15/51,C12P21/02

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类体系和分类号)

Int.Cl⁷: C12N15/62,15/17,15/12,15/51,C12P21/02

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词)

WPI,EPODOC,PAJ,CNPAT

C. 相关文件

类 型*	引用文件, 必要时, 包括相关段落的说明	相关的权利要求编号
A	WO0017336 A1; LEE S 等, C12N15/00,C12N15/62, 2000/03/30 见摘要	1—11
A	WO9711186 A1; BOEHRINGER MANNHEIM GMBH 等, C12N15/62, 1997/03/27; 见摘要	1—11
A	WO9310152 A1; SMIK 等, C07K15/00,C12N15/62, 1993/05/27 见摘要	1—11
A	DE4012818 A; HOECHST AG, C12N15/62, 1991/10/24 见摘要	1—11
A	EP0281418 A; SUNTORY LTD, C12N15/62, 1988/09/07 见摘要	1—11

☐ 其余文件在 C 栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的专用类型:

“A” 明确表示了一般现有技术、不认为是特别相关的文件
“E” 在先文件, 但是在国际申请日的同一日或之后公布的
“L” 对优先权要求可能产生怀疑或者用来确定另一篇引用
文件的公布日期或其它特殊理由而引用的文件(如详
细说明)

“O” 涉及口头公开、使用、展览或其他手段的文件

“P” 在国际申请日之前但迟于所要求的优先权日公布的文件

“T” 在国际申请日或优先权日之后公布的在后文件, 它与申请不
相抵触, 但是引用它是为了解构成发明基础的理论或原理
“X” 特别相关的文件; 当该文件被单独使用时, 要求保护的发明
不能认为是新颖的或不能认为具有创造性

“Y” 特别相关的文件; 当该文件与其他一篇或多篇这类文件结
合在一起, 这种结合对本领域技术人员是显而易见的, 要
求保护的发明不能认为具有创造性

“&” 同族专利成员的文件

国际检索实际完成的日期

24. 6 月 2003(24.06.03)

国际检索报告邮寄日期

10. 7 月 2003(10.07.03)

国际检索单位名称和邮寄地址

中国国家知识产权局专利局
中国北京市海淀区西土城路 6 号(100088)

传真号: 86-010-62019451

受权官员

曾繁辉

电话号码: 010-62093733

国际检索报告
同族专利成员的情报

国际申请号

PCT/CN03/00426

检索报告中引用的 专利文件	公布日期	同族专利成员	公布日期
WO0017336 A1	2000/03/30	KR2000023309 A AU5762699 A	2000/04/25 2000/04/10
WO9711186 A1	1997/03/27	JP11511333 T DE19535445 A1 AU7130096 A EP0851930 A1	1999/10/05 1997/03/27 1997/04/09 1998/07/08
WO9310152 A1	1993/05/27	KR251505 B1 AU2927892 A EP0614465 A1 JP7501213 T US5928902 A	2000/05/01 1993/06/15 1994/09/14 1995/02/09 1999/07/27
DE4012818 A	1991/10/24	CA2040810 C EP0453969 A AU7515491 A BR9101587 A CN1055952 A US5426036 A	2001/07/24 1991/10/30 1991/10/24 1991/12/10 1991/11/06 1995/06/20
EP0281418 A	1988/09/07	CA1340830 C AU1272488 A JP1010999 A US4987070 A HU210625 B	1999/11/30 1988/09/15 1989/01/13 1991/01/22 1995/06/28