



(12)发明专利申请

(10)申请公布号 CN 110835366 A

(43)申请公布日 2020.02.25

(21)申请号 201911023604.7

C12R 1/19(2006.01)

(22)申请日 2019.10.25

(71)申请人 中国科学院生物物理研究所  
地址 100101 北京市朝阳区大屯路15号

(72)发明人 欧先金 李雪梅

(74)专利代理机构 北京植德律师事务所 11780  
代理人 唐华东

(51)Int.Cl.

C07K 7/08(2006.01)

C07K 14/00(2006.01)

C07K 19/00(2006.01)

C12N 15/11(2006.01)

C12N 15/62(2006.01)

C12N 15/70(2006.01)

C12N 1/21(2006.01)

权利要求书1页 说明书5页 附图1页

(54)发明名称

促进蛋白质可溶性表达的标签多肽及其用途

(57)摘要

本发明涉及一种促进蛋白质可溶性表达的标签多肽及其用途。其中,所述标签多肽具有选自以下的氨基酸序列:1)SEQ ID NO.:1所示氨基酸序列;2)SEQ ID NO.:1所示氨基酸序列经过取代、缺失或添加一个或多个氨基酸而得到的氨基酸序列;或3)与SEQ ID NO.:1具有90%以上序列同一性的氨基酸序列。利用本发明标签多肽可提高目标蛋白质的可溶性表达,避免外源融合蛋白质(比如GST、MBP)带来免疫干扰和蛋白质酶切割繁琐操作的不利因素,能够低成本、简便、高效地生产目标蛋白质。

1. 一种促进蛋白质可溶性表达的标签多肽,所述标签多肽具有选自以下的氨基酸序列:

1) SEQ ID NO.:1所示氨基酸序列;

2) SEQ ID NO.:1所示氨基酸序列经过取代、缺失或添加一个或多个氨基酸而得到的氨基酸序列;或

3) 与SEQ ID NO.:1具有90%以上序列同一性的氨基酸序列。

2. 一种融合蛋白质,所述融合蛋白质包含权利要求1所述的标签多肽以及目标蛋白质,其中,所述融合蛋白质从N端到C端具有以下通式表示的氨基酸序列:

A-B或B-A

其中,A是标签多肽,B是目标蛋白质;

任选地,标签多肽和目标蛋白质之间可直接连接或通过几个到几十个氨基酸残基连接;

优选地,所述目的多肽为鼠抵抗素;更优选地,所述鼠抵抗素的氨基酸序列如SEQ ID NO.:3所示。

3. 一种编码权利要求1所述标签多肽的多核苷酸;优选地,所述标签多肽的编码多核苷酸序列如SEQ ID NO.:2所示。

4. 一种编码权利要求2所述融合蛋白质的多核苷酸;优选地,所述鼠抵抗素的编码多核苷酸序列如SEQ ID NO.:4所示。

5. 包含权利要求3或4所述多核苷酸的重组表达载体;

优选地,所述表达载体为原核表达载体;更优选地,所述表达载体为质粒,优选为pET-28a或者pET-22a。

6. 包含权利要求3或4所述多核苷酸或权利要求5所述重组表达载体的宿主细胞;

优选地,所述宿主细胞选自真核细胞或原核细胞;优选地,所述真核细胞是酵母菌;优选地,所述原核细胞是大肠杆菌细胞,如大肠杆菌T7 express或BL21 (DE3)。

7. 一种用于重组蛋白质高效可溶性表达的重组表达载体的构建方法,其包括:

(1) 制备权利要求3所述的标签多肽的多核苷酸;

(2) 制备目标蛋白质的多核苷酸;优选地,所述目标蛋白质是鼠抵抗素;更优选地,所述鼠抵抗素的编码核酸如SEQ ID NO.:4所示;

(3) 连接标签多肽的核酸与目标蛋白质的核酸,获得重组表达载体;

优选地,连接方法通过核苷酸引物的重叠设计,通过重叠PCR扩增得到标签蛋白质-目标蛋白质的融合核苷酸序列。

8. 一种用于重组蛋白质高效可溶性表达的重组表达菌体的制备方法,其方法包括:

将权利要求8构建的所述重组表达载体转化宿主细胞,获得重组表达菌体。

9. 一种重组蛋白质高效可溶性表达的方法,其包括:

蛋白质表达:37°C培养培养重组表达菌体,在菌浓OD<sub>600nm</sub>达到0.6-0.8范围时,添加0.1-1mM终浓度的IPTG进行诱导蛋白质表达。

10. 权利要求1所述的标签多肽、权利要求2所述的融合蛋白质、权利要求3或4所述的多核苷酸、权利要求5所述的重组表达载体、权利要求6所述的宿主细胞在生产可溶性蛋白质中的用途。

## 促进蛋白质可溶性表达的标签多肽及其用途

### 技术领域

[0001] 本发明属于基因工程领域。更具体地,本发明涉及促进蛋白质可溶性表达的标签多肽及其用途。

### 背景技术

[0002] 蛋白质是生物医学中诸多研究的研究对象和材料,也是我们重要的医疗产业产品,比如胰岛素和干扰素等都是蛋白质,所以蛋白质的表达生产是我们科学研究和社会产业中重要的环节。

[0003] 从原始生物体中提取目标蛋白质,比如从猪胰腺提取胰岛素,成本高居不下导致无法普及推广。现代分子生物学的核苷酸重组技术可以低成本、大规模地重组表达目标蛋白质,是近半个世纪伟大的技术突破。但是,重组蛋白质高水平的、快速的表达会导致目标蛋白质的不正确折叠,在宿主细胞内就形成不可溶性的蛋白质沉淀,即包涵体。包涵体蛋白质仅仅具有正确的一级氨基酸序列,不能形成正确的空间三维结构,也就说包涵体蛋白质不具备它应有的生物活性。所以,如何让目标蛋白质由包涵体变成可溶性的表达,是蛋白质大规模低成本表达的重要难题。

[0004] 目前所有重组表达外源蛋白质的系统中,大肠杆菌表达系统(简称E.coli,下同)是技术背景研究最广泛最清楚的,该系统易于培养操作,繁殖快、大规模生产成本低且表达水平高。蛋白质的快速高水平表达是其优点,同时也带来该系统最大缺点,就是易导致蛋白质的不正确折叠,导致形成包涵体。

[0005] 针对E.coli表达系统中蛋白质表达易形成包涵体的问题,目前已经有很多公开的技术方法都能有效改善,比如降低诱导表达温度和诱导剂浓度,或者增加分子伴侣的表达。其中方法之一就是在目标蛋白质的上游增加促进目标蛋白质可溶性表达的其他蛋白质,比如常用谷胱甘肽S转移酶(简称GST,下同),最终表达“GST-目标蛋白质”可溶性蛋白质。或是麦芽糖结合蛋白质(简称MBP,下同),最终表达“MBP-目标蛋白质”可溶性蛋白质。

[0006] GST/MBP确实有促进目标蛋白质的可溶性表达的实例报道,但GST约27kDa,MBP约42kDa,促溶蛋白质分子量较大,会带来免疫干扰反应。同时,也会限制目标蛋白质的分子量上限。如果想得到不含有促溶蛋白质(GST和MBP),需要用蛋白质酶切割(比如TEV蛋白酶),最后得到目标蛋白质,但是这样会增加操作流程和成本。

### 发明内容

[0007] 为了解决现有技术中存在的上述缺陷,本发明的目的是提供一种短肽来替代GST/MBP,该短肽既能够促进蛋白质可溶性表达,又能够避免GST/MBP因分子量过大带来的不足,并能够低成本、简便、高效地生产多肽。

[0008] 为了实现上述目的,本发明提供了一种促进蛋白质可溶性表达的标签多肽,所述标签多肽具有选自以下的氨基酸序列:

[0009] 1) SEQ ID NO.:1所示氨基酸序列;

[0010] 2) SEQ ID NO.:1所示氨基酸序列经过取代、缺失或添加一个或多个氨基酸而得到的氨基酸序列;或

[0011] 3) 与SEQ ID NO.:1具有90%以上序列同一性的氨基酸序列。

[0012] 其中,SEQ ID NO.:1具体序列如下:

[0013] Glu-Glu-Glu-Glu-Asp-Tyr-Lys-Asp Asp-Asp-Asp-Lys,即EEEEYKDDDDK。

[0014] 优选地,所述标签多肽的编码核酸如SEQ ID NO.:2所示,即具有以下36个核苷酸残基的核酸:

[0015] GAAGAAGAAGAAGACTACAAAGACGACGACGACAAA

[0016] 本发明还提供了一种融合蛋白质,所述融合蛋白质包含所述促进蛋白质可溶性表达的标签多肽以及目标蛋白质,其中,从N端到C端具有以下通式表示的氨基酸序列:

[0017] A-B或B-A

[0018] 其中,A是标签多肽,B是目标蛋白质;

[0019] 任选地,标签多肽和目标蛋白质之间可直接连接或通过几个到几十个氨基酸残基连接。

[0020] 优选地,所述目的多肽为鼠抵抗素。

[0021] 优选地,所述鼠抵抗素的氨基酸序列如SEQ ID NO.:3所示,即具有以下94个氨基酸残基的多肽:

[0022] SSMPLCPIDEAIDKKIKQDFNSLFPNAIKNIGLNCWTVSSRGKLASCPEGT

[0023] AVLSCSCGSACGSWDIREEKVCHCQCARIDWTAARCKLQVAS

[0024] 优选地,所述鼠抵抗素的编码核酸如SEQ ID NO.:4所示,即具有以下282个核苷酸残基的核酸:

[0025] TCCAGCATGCCACTGTGTCCCATCGATGAAGCCATCGACAAGAAGATC

[0026] AAACAAGACTTCAACTCCCTGTTTCCAAATGCAATAAAGAACATTGGC

[0027] TTAAATTGCTGGACAGTCTCCTCCAGAGGGAAGTTGGCCTCCTGCCCA

[0028] GAAGGCACAGCAGTCTTGAGCTGCTCCTGTGGCTCTGCCTGTGGCTCG

[0029] TGGGACATTTCGTGAAGAAAAAGTGTGTCAGTCCAGTGTGCAAGGAT

[0030] AGACTGGACAGCAGCCCGCTGCTGTAAGCTGCAGGTCGCTTCC

[0031] 本发明还提供了编码所述标签多肽或融合蛋白质的多核苷酸、包含所述多核苷酸的重组表达载体以及包含所述载体的宿主细胞。

[0032] 优选地,所述表达载体选自为原核表达载体。优选地,所述表达载体为质粒,优选为pET-28a或者pET-22a。

[0033] 所述宿主细胞选自真核细胞或原核细胞;优选地,所述真核细胞是酵母菌;优选地,所述原核细胞是大肠杆菌细胞,如大肠杆菌T7 express或BL21 (DE3)。

[0034] 本发明还提供了一种用于重组蛋白质高效可溶性表达的重组表达载体构建方法,其包括:

[0035] (1) 制备标签多肽的多核苷酸;优选地,所述标签多肽的编码核酸如SEQ ID NO.:2所示

[0036] (2) 制备目标蛋白质的多核苷酸;优选地,所述目标蛋白质是鼠抵抗素;更优选地,所述鼠抵抗素的编码核酸如SEQ ID NO.:4所示;

[0037] (3) 连接标签多肽的核酸与目标蛋白质的核酸,获得重组表达载体;优选地,连接方法通过核苷酸引物的重叠设计,通过重叠PCR扩增得到标签蛋白质-目标蛋白质的融合核苷酸序列。

[0038] 本发明还提供了一种用于重组蛋白质高效可溶性表达的重组表达菌体的制备方法,其包括:将所述重组表达载体转化如宿主细胞,获得重组表达菌体。

[0039] 其中,可以将标签多肽-目标蛋白质的融合多肽视为一条多肽链,进行常规分子生物学的重组连接操作,构建在E.coli常规表达载体(比如pET系列表达载体)的多克隆位点,得到“标签肽-目标蛋白质”的重组表达载体。把重组表达载体转化进入相应的宿主细胞,比如BL21 (De3)。

[0040] 本发明还提供了一种重组蛋白质高效可溶性表达的方法,所述方法包括:

[0041] 蛋白质表达:37℃培养培养重组细胞,在菌浓OD<sub>600nm</sub>达到0.6-0.8范围时,添加0.1-1mM终浓度的IPTG进行诱导蛋白质表达。

[0042] 优选地,所述方法还包括对重组菌体进行常规的蛋白质表达分析,或利用LB摇瓶培养基进行蛋白质可溶性表达的分析。

[0043] 更优选地,蛋白质表达的可溶性分析包括:菌体离心收集,5000g离心力10分钟,丢弃离心液上清。用缓冲液进行菌体重悬,重悬后菌体浓度控制在80-1000D<sub>600nm</sub>范围。之后冰浴超声波破碎,200w,4秒破碎4秒间隔休息,总破碎时间20分钟。200微升破碎液高速离心,16000g离心力,20分钟,离心后的上清约190微升作为“表达上清”,离心沉淀物用500微升缓冲液轻轻重洗2次,之后用180微升缓冲液重悬沉淀物,得到样品视为“表达沉淀”。然后把“表达上清”和“表达沉淀”样品进行SDS-page蛋白质电泳,根据跑胶结果,由目标蛋白质的分布情况判断目标蛋白质可溶性表达情况。

[0044] 本发明的有益效果:

[0045] 利用本发明标签多肽可提高目标蛋白质的可溶性表达,添加标签多肽之后的融合RESISTIN的可溶性表达得到80%比例,上清样品经过检查具有生物活性。同时,采用本发明的标签多肽可以避免外源融合蛋白质(比如GST、MBP)带来免疫干扰和蛋白质酶切割繁琐操作的不利因素,能够低成本、简便、高效地生产目标蛋白质。

## 附图说明

[0046] 图1为融合RESISTIN可溶性表达的蛋白质电泳胶。从左到右分别是:诱导后全菌、表达上清、表达沉淀和蛋白质Marker。

## 具体实施方式

[0047] 为使本发明更加容易理解,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用于限制本发明的范围。本领域技术人员在不偏离本发明的精神和范围下对本发明技术方案的细节和形式进行的任何修改和替换,均落入本发明的保护范围内。下列实施例中未提及的具体实验方法,通常按照常规实验方法进行。

[0048] 实施例

[0049] 实施例1鼠抵抗素蛋白质的可溶性表达

[0050] 鼠抵抗素蛋白质(RESISTIN)序列含有114个氨基酸残基,起始20个氨基酸序列是

引导蛋白质质跨越细胞膜定位的信号肽,在重组表达时通常除去信号肽,余下如SEQ ID NO.:3所示的94氨基酸残基,其具体氨基酸序列如下:

[0051] SSMPLCPIDEAIDKKIKQDFNSLFPNAIKNIGLNCWTVSSRGKLASCPEGT

[0052] AVLSCSCGSACGSWDIREEKVCHCQCARIDWTAARCCKLQVAS)。

[0053] 所对应的核苷酸序列如SEQ ID NO.:4所示,共计282碱基对,其具体核苷酸序列如下:

[0054] TCCAGCATGCCACTGTGTCCCATCGATGAAGCCATCGACAAGAAGATC

[0055] AAACAAGACTTCAACTCCCTGTTTCCAAATGCAATAAAGAACATTGGC

[0056] TTAAATTGCTGGACAGTCTCCTCCAGAGGGAAGTTGGCCTCCTGCCCA

[0057] GAAGGCACAGCAGTCTTGAGCTGCTCCTGTGGCTCTGCCTGTGGCTCG

[0058] TGGGACATTCGTGAAGAAAAAGTGTGTCACTGCCAGTGTGCAAGGAT

[0059] AGACTGGACAGCAGCCCCTGCTGTAAGCTGCAGGTCGCTTCC

[0060] 鼠抵抗素蛋白质含有5对内部二硫键,通常含有3对以上内部二硫键的蛋白质在E.coli系统表达结果就是全部包涵体,本实施例鼠抵抗素在E.coli原核的常规表达也是完全不可溶性的包涵体。

[0061] 本实施例中,将本发明的12个氨基酸残基的标签多肽的对应核苷酸序列直接连接在RESISTIN的上游(-端),得到融合蛋白质的具体核苷酸序列(含有终止密码子TGA的总长度:321碱基对,简称融合RESISTIN),该融合蛋白质的氨基酸序列如SEQ ID NO.:5所示,即具有以下106个氨基酸残基的多肽:

[0062] EEEEDYKDDDDKSSMPLCPIDEAIDKKIKQDFNSLFPNAIKNIGLNCWTV

[0063] SSRGKLASCPEGTAVLSCSCGSACGSWDIREEKVCHCQCARIDWTAARC

[0064] CKLQVAS

[0065] 所述融合RESISTIN的编码核酸序列如SEQ ID NO.:6所示,即具有以下321个核苷酸残基的核酸:

[0066] GAAGAAGAAGAAGACTACAAAGACGACGACGACAAAATCCAGCATGC

[0067] CACTGTGTCCCATCGATGAAGCCATCGACAAGAAGATCAAAACAAGACT

[0068] TCAACTCCCTGTTTCCAAATGCAATAAAGAACATTGGCTTAAATTGCTG

[0069] GACAGTCTCCTCCAGAGGGAAGTTGGCCTCCTGCCAGAAGGCACAG

[0070] CAGTCTTGAGCTGCTCCTGTGGCTCTGCCTGTGGCTCGTGGGACATTC

[0071] GTGAAGAAAAAGTGTGTCACTGCCAGTGTGCAAGGATAGACTGGACA

[0072] GCAGCCCCTGCTGTAAGCTGCAGGTCGCTTCTCTGA

[0073] 融合RESISTIN的核苷酸序列通过核酸酶EcoR I和Xho I双酶切,再重组连接构建在pET-28(a)表达载体的EcoR I和Xho I之间,

[0074] 得到的重组载体经过核苷酸测序正确后再转回进入BL21(DE3)感受态细胞。再次鉴定正确的重组BL21(de3)菌落视为融合RESISTIN的重组表达菌。

[0075] 对于融合RESISTIN的重组表达菌,菌体离心收集,5000g离心力10分钟,丢弃离心液上清。用缓冲液进行菌体重悬,重悬后菌体浓度控制在80-1000D<sub>600nm</sub>范围。之后冰浴超声波破碎,200w,4秒破碎4秒间隔休息,总破碎时间20分钟。200微升破碎液高速离心,16000g离心力,20分钟,离心后的上清约190微升作为“表达上清”,离心沉淀物用500微升缓冲液轻

轻重洗2次,之后用180微升缓冲液重悬沉淀物,得到样品视为“表达沉淀”。然后把“表达上清”和“表达沉淀”样品进行SDS-page蛋白质电泳,根据跑胶结果,由目标蛋白质的分布情况判断目标蛋白质可溶性表达情况。实验得到的蛋白质可溶性分析结果如图1所示。由图1可见,添加标签多肽之后的融合RESISTIN的可溶性比例超过一半。

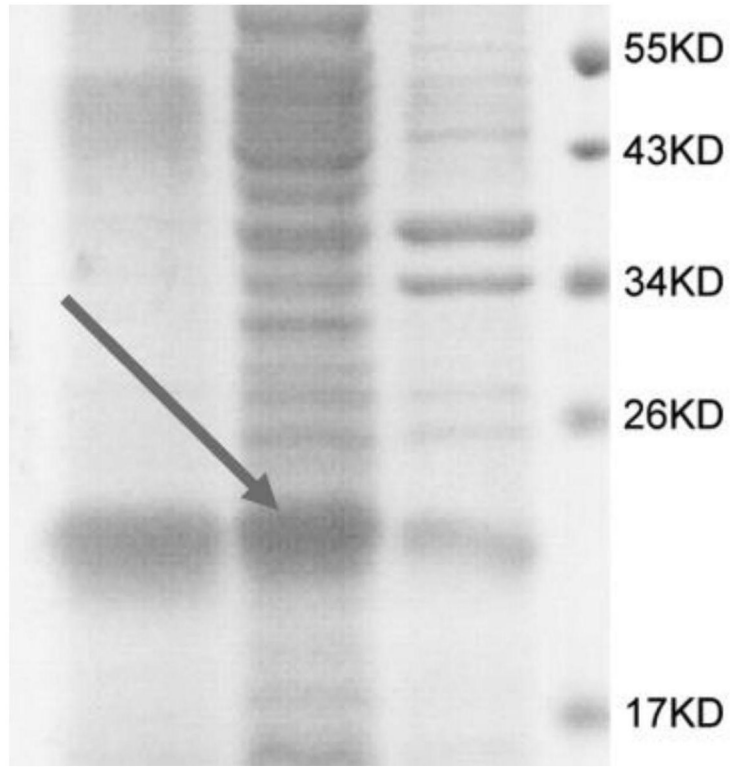


图1