

大规模制备164.88 α 三螺旋结构胶原蛋白生物方法

本申请要求申请日：2023年9月28日，申请号：202311272338.8，发明名称为“大规模制备164.88 α 三螺旋结构胶原蛋白生物方法”的中国发明专利申请的优先权权益。

技术领域

本发明涉及利用大规模生物发酵实现具有164.88 α 三螺旋结构的胶原蛋白制备方法属于合成生物学技术领域。

背景技术

人有28种不同的胶原蛋白，胶原蛋白是人和动物体内含量最多的蛋白，占体内蛋白总量的25-30%，约占体重的6%。胶原蛋白是皮肤、骨骼、牙齿、角膜、肌腱、韧带以及各种组织器官中存在的重要的结构蛋白，起支撑和保护机体组织的功能。胶原蛋白是细胞外基质的主要组分，其生物学活性主要表现在促进细胞粘附、生长分化、细胞信号通路的应答、血小板凝集、细胞和组织器官功能的维护和损伤修复等功能。胶原蛋白在空间结构上呈现出三股螺旋缠绕的结构，三条独立的胶原 α 肽链主要以氢键的形式缠绕形成螺旋结构。在体内胶原蛋白以三螺旋的结构与组织和细胞结合以发挥其生物学效应，因此胶原蛋白的生物学活性很大程度上依赖其天然三螺旋结构。胶原蛋白优良的生物学活性使其在生物医学材料，医疗美容、化妆品等领域拥有巨大的应用价值和潜力。

现在市场上主要的胶原蛋白产品来源于动物组织提取，这种方法在提取的过程中破坏了天然三螺旋结构，影响了胶原生物学活性的发挥。而且使用动物来源提取工艺得到的胶原蛋白其肽段分子量不均一，可加工性较差。由于动物胶原蛋白与人胶原蛋白在氨基酸序列上存在差异，因此可能会在临床上造成潜在的免疫原性，安全性较差。而且动物组织受其来源的影响需要解决动物病毒的安全隐患才能进入市场。随着市场对胶原蛋白产品安全性、有效性要求的提高，传统的从动物组织中提取胶原蛋白的方式已经无法满足现有市场的需求。而且，以胶原蛋白作为生物材料应用于医疗场景的产品需求量也在日益增大，需要通过别的途径来解决目前市场上存在的问题。通

过基因工程等分子生物学技术来生产具有天然三螺旋结构的重组胶原蛋白在解决以上所述问题方面具有极大的潜力。

市场上已经有企业通过基因重组的方式制备III型胶原蛋白的产品，但是难以实现三螺旋结构，利用合成生物发酵具有正确三螺旋结构的胶原蛋白具有较高技术壁垒。

本领域中需要合成生物发酵具有正确三螺旋结构的胶原蛋白的方法。

发明内容

发明人发现了一种人III型胶原蛋白，结构研究证实，人III型胶原蛋白核心功能区具有164.88°三螺旋结构（CN109593126A）。发明人已经通过基因重组的方式，利用大肠杆菌表达系统成功制备了重组III型人源化胶原蛋白。然而，发明人发现大肠杆菌表达系统存在蛋白表达量有限，工艺放大受限等缺点。因此，发明人尝试以真核表达系统表达人III型胶原蛋白。经过大量前期筛选工作，发明人发现当以毕赤酵母表达人III型胶原蛋白时可以体外合成大量具有164.88°三螺旋结构胶原蛋白。发明人在先前的申请中还发现人III型胶原蛋白在通过大肠杆菌表达时存在细菌内毒素，这增加了胶原蛋白纯化的复杂性和难度。发明人通过在酵母表达系统中表达胶原蛋白消除了该技术问题。

本发明通过改造“底盘细菌”获得了大规模生产，具有164.88°三螺旋结构的重组人源化胶原蛋白，属于A型，通过大规模生产能够满足市场需求。

在一方面，提供了一种构建表达胶原蛋白的酵母细胞的方法，其包括将重组表达载体导入酵母细胞中，其中重组表达载体包含编码胶原蛋白的多核苷酸，所述胶原蛋白包含以SEQ ID No. 1所示的序列的重复单元，重复单元的数目是2-32，重复单元是直接连接的。在一个实施方案中，胶原蛋白是重组III型人源化胶原蛋白。在一个实施方案中，胶原蛋白是164.88°三螺旋结构胶原蛋白。

在一个实施方案中，胶原蛋白包含SEQ ID No. 2或3所示的氨基酸序列。在一个实施方案中，多核苷酸包含SEQ ID No. 4或5所示的核苷酸序列。本发明的多核苷酸可以在酵母宿主细胞中有效表达胶原蛋白。

在一个实施方案中，重组表达载体是酵母系统表达载体。在一个实施方案中，酵母系统表达载体为pPICZα A、pPIC9、pPIC9K、pHIL-S1、或pYAM75P

载体。

在一个实施方案中，酵母是巴斯德毕赤酵母。在一个实施方案中，巴斯德毕赤酵母是巴斯德毕赤酵母X33、GS115、SMD1168、KM71或KM71H菌株。

5 在一个实施方案中，方法包括以下一个或多个步骤：

(1) 将多核苷酸插入表达载体，优选pPicza α A表达载体的NotI和XhoI酶切位点之间，以得到重组表达质粒；

(2) 对重组表达质粒进行酶切，优选用sacI进行酶切；

(3) 对酶切后的重组表达质粒进行纯化；

10 (4) 通过电转化，将酶切后的重组表达质粒导入酵母细胞中。

在另一个方面，提供了表达胶原蛋白的酵母细胞，其中所述胶原蛋白包含以SEQ ID No .1所示的序列的重复单元，重复单元的数目是2-32，重复单元是直接连接的；酵母细胞是巴斯德毕赤酵母。

在一个实施方案中，巴斯德毕赤酵母是巴斯德毕赤酵母X33、GS115、
15 SMD1168、KM71或KM71H菌株。在一个实施方案中，胶原蛋白包含SEQ ID No .2或3所示的氨基酸序列。在一个实施方案中，胶原蛋白是重组III型人源化胶原蛋白。在一个实施方案中，胶原蛋白是164.88°三螺旋结构胶原蛋白。

在一个实施方案中，酵母细胞包含酵母系统表达载体。在一个实施方案中，酵母系统表达载体为pPICZ α A、pPIC9、pPIC9K、pHIL-S1、或pYAM75P
20 载体。在一个实施方案中，表达载体包含多核苷酸，该多核苷酸包含SEQ ID No. 2或3所示的核苷酸序列。

在另一个方面，提供了本文所述的酵母细胞用于制备胶原蛋白的用途。

在一个实施方案中，胶原蛋白包含以SEQ ID No .1所示的序列的重复单元，重复单元的数目是2-32，重复单元是直接连接的。在一个实施方案中，
25 胶原蛋白包含SEQ ID No .2或3所示的氨基酸序列。在一个实施方案中，胶原蛋白是重组III型人源化胶原蛋白。在一个实施方案中，胶原蛋白是164.88°三螺旋结构胶原蛋白。

在另一个方面，提供了通过发酵生产胶原蛋白的方法，其包括在合适的条件下培养本文所述的酵母细胞，对酵母细胞添加甲醇以进行诱导表达，并
30 且收获培养上清液以得到胶原蛋白。

在一个实施方案中，培养基为补充有无氨基酵母氮源和生物素的BMMY

培养基。

在一个实施方案中，胶原蛋白包含以SEQ ID No .1所示的序列的重复单元，重复单元的数目是2-32，重复单元是直接连接的。在一个实施方案中，胶原蛋白包含SEQ ID No .2或3所示的氨基酸序列。在一个实施方案中，胶原蛋白是重组III型人源化胶原蛋白，优选164.88°三螺旋结构胶原蛋白。

本发明的优点包括：

1. 发明人发现了适合于表达重组III型人源化胶原蛋白的酵母细胞、其构建方法和生产方法；

2. 本发明的酵母细胞是巴斯德毕赤酵母相比于原核表达系统，诸如大肠杆菌具有表达量大、工艺可放大的优点；

3. 本发明的方法体外合成具有164.88°三螺旋结构胶原蛋白，实现吨级大规模量产，具有重要产业意义；

4. 本发明可以不需要消除细胞内毒素的步骤。

附图说明

图1显示了高分子量P-012诱导表达电泳图。

图2显示了中分子量P-C3T8诱导表达电泳图。

图3显示了高分子量P-012细胞粘附活性。

图4显示了中分子量P-C3T8细胞粘附活性。

图5显示了高分子量P-012圆二色谱结果。

图6显示了中分子量P-C3T8圆二色谱结果。

具体实施方式

为使本发明的目的、技术方案和优点更加清楚，下面将结合本发明的实施例，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

如本文使用的，重组胶原蛋白是利用前沿的结构生物学、基因工程等技
术，以人特定型别胶原蛋白功能区基因编码为模板，进行筛选、制备得到与人胶原蛋白氨基酸序列相同或类似的一类新型生物材料。

如本文中所用,“人重组III型胶原蛋白”指由或基本上由源自人III型胶原蛋白的序列组成的重组蛋白。在本文中,人重组III型胶原蛋白可以由或基本上由人III型胶原蛋白衍生的片段或片段的多个重复组成。如本文使用的,“酵母”包括产子囊酵母(*ascosporogenous yeast*)(内孢霉目(*Endomycetales*))、产担子酵母(*basidiosporogenous yeast*)和属于半知菌纲(*Fungi Imperfecti*) (芽孢纲(*Blastomycetes*))的酵母。

如本文使用的,酵母宿主细胞意指易于用包含本发明的多核苷酸的核酸构建体或重组表达载体进行转化、转染、转导等的任何酵母宿主细胞。酵母宿主细胞可以是假丝酵母属(*Candida*)、汉逊酵母属(*Hansenula*)、克鲁弗酵母属(*Kluyveromyces*)、毕赤酵母属(*Pichia*)、酵母属(*Saccharomyces*)、裂殖酵母属(*Schizosaccharomyces*)或耶氏酵母属(*Yarrowia*)细胞,如乳酸克鲁弗酵母(*Kluyveromyces lactis*)、卡尔酵母(*Saccharomyces carlsbergensis*)、酿酒酵母(*Saccharomyces cerevisiae*)、糖化酵母(*Saccharomyces diastaticus*)、道格拉氏酵母(*Saccharomyces douglasii*)、克鲁弗酵母(*Saccharomyces kluyveri*)、诺地酵母(*Saccharomyces norbensis*)、卵形酵母(*Saccharomyces oviformis*)或解脂耶氏酵母(*Yarrowia lipolytica*)细胞。在本文中,优选的酵母细胞是巴斯德毕赤酵母。更优选地,巴斯德毕赤酵母X33、GS115、SMD1168、KM71或KM71H菌株。

如本文使用的,术语“表达”包括涉及多肽产生的任何步骤,包括但不限于:转录、转录后修饰、翻译、翻译后修饰、以及分泌。

如本文使用的,术语“表达载体”意指直链或环状DNA分子,其包含编码多肽的多核苷酸并且可操作地连接至提供用于其表达的控制序列。在本文中,表达载体是酵母系统表达载体,优选为pPICZα A、pPIC9、pPIC9K、pHIL-S1、或pYAM75P载体。

如本文使用的,术语“重组表达载体”意指单链或双链的核酸分子,该核酸分子是从天然存在的基因中分离的,或以本来不存在于自然界中的方式被修饰以含有核酸区段,或是合成的,该核酸分子包含一个或多个控制序列。

如本文使用的,术语“控制序列”意指表达编码本发明的成熟多肽的多核苷酸所必需的核酸序列。每个控制序列对于编码该多肽的多核苷酸来说可以是天然的(即,来自相同基因)或外源的(即,来自不同基因),或相对于彼此是天然的或外源的。此类控制序列包括但不限于前导序列、多腺苷酸化序列、

前肽序列、启动子、信号肽序列、以及转录终止子。最少，控制序列包括启动子、以及转录和翻译终止信号。出于引入有利于将控制序列与编码多肽的多核苷酸的编码区连接的特异性限制位点的目的，这些控制序列可以提供有多个接头。

5 胶原蛋白

本发明提供了胶原蛋白。胶原蛋白可以包含以SEQ ID No .1所示的序列或其该序列具有80-100%（例如80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%）序列同一性的序列的重复单元。重复单元的数目是2-32，例如2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31或32。各个重复单元可以是直接连接的。胶原蛋白是重组III型人源化胶原蛋白，优选是164.88°三螺旋结构胶原蛋白。胶原蛋白可以包含SEQ ID No .2或3所示的氨基酸序列或其该序列具有80-100%（例如80%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%）序列同一性的氨基酸序列。本文所述的胶原蛋白可以具有三螺旋结构或三条相同的链（即为三聚体形式）。每条链的序列可以为本文所述的序列。

重组表达载体

本发明还涉及重组表达载体，其包含可操作地连接至一个或多个控制序列的本发明所述的核酸，在与控制序列相容的条件下，该一个或多个控制序列指导编码序列在合适的宿主细胞中的表达。载体可以包含重组表达载体。本发明所述的核酸可以包含SEQ ID NO. 4或5的核苷酸序列。

可用多种方式操纵核酸以提供胶原蛋白的表达。取决于表达载体，在核酸插入载体之前对其进行操纵可能是理想的或必需的。用于利用重组DNA方法修饰核酸的技术是本领域熟知的。

控制序列可为启动子，即，被宿主细胞识别用于表达编码本发明的胶原蛋白或多肽的多核苷酸。启动子包含介导胶原蛋白或多肽的表达的转录控制序列。启动子可以是在宿主细胞中显示出转录活性的任何核酸，包括变体、截短型及杂合型启动子，并且可以从编码与该宿主细胞同源或异源的细胞外或细胞内胶原蛋白或多肽的基因获得。

用于指导本发明的载体或重组表达载体在酵母宿主细胞中的转录的合

适启动子的实例没有特别限制。在酵母宿主中，有用的启动子从以下的基因中获得：酿酒酵母烯醇酶（ENO-1）、酿酒酵母半乳糖激酶（GAL1）、酿酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶（ADH1、ADH2/GAP）、酿酒酵母丙糖磷酸异构酶（TPI）、酿酒酵母金属硫蛋白（CUP1）、以及酿酒酵母3-磷酸甘油酸激酶。

控制序列也可为由宿主细胞识别以终止转录的转录终止子。终止子可操作地连接至编码胶原蛋白或多肽的多核苷酸的3'末端。在宿主细胞中有功能的任何终止子可用于本发明中。酵母宿主细胞的优选终止子从以下的基因中获得：酿酒酵母烯醇酶、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。控制序列还可以是启动子下游和基因的编码序列上游的mRNA稳定子区，其提高该基因的表达。

合适的mRNA稳定子区的实例从以下基因中获得：苏云金芽孢杆菌 *cryIIIa* 基因（WO 94/25612）和枯草芽孢杆菌 SP82 基因（Hue 等人，1995, *Journal of Bacteriology* [细菌学杂志] 177: 3465-3471）。

控制序列也可以是前导序列，即对宿主细胞翻译很重要的mRNA的非翻译区域。前导序列可操作地连接至编码胶原蛋白或多肽的多核苷酸的5'末端。可以使用在宿主细胞中有功能的任何前导序列。酵母宿主细胞的合适的前导序列从以下的基因中获得：酿酒酵母烯醇酶（ENO-1）、酿酒酵母3-磷酸甘油酸激酶、酿酒酵母 α -因子和酿酒酵母醇脱氢酶/甘油醛-3-磷酸脱氢酶（ADH2/GAP）。

控制序列还可以是多腺苷酸化序列，一种可操作地连接至该多核苷酸的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。可以使用在宿主细胞中有功能的任何多腺苷酸化序列。酵母宿主细胞的有用的多腺苷酸化序列由Guo和Sherman, 1995, *Mol. Cellular Biol.* [分子细胞生物学] 15: 5983-5990描述。

控制序列还可以是编码与胶原蛋白的N-末端连接的信号肽并指导胶原蛋白进入细胞的分泌途径的信号肽编码区。多核苷酸的编码序列的5'-端本身可以含有在翻译阅读框中天然与编码胶原蛋白的编码序列区段相连接的信号肽编码序列。可替代地，编码序列的5'-端可以含有对编码序列而言是外源的信号肽编码序列。在编码序列不天然地含有信号肽编码序列的情况下，可能需要外源信号肽编码序列。可替代地，外源信号肽编码序列可以单纯地替

代天然信号肽编码序列以便增强胶原蛋白或多肽的分泌。然而，可以使用指导已表达胶原蛋白进入宿主细胞的分泌途径的任何信号肽编码序列。酵母宿主细胞的有用的信号肽从以下的基因中获得：酿酒酵母 α -因子和酿酒酵母转化酶。酵母细胞可以应用于用于制备本文的胶原蛋白。

5 宿主细胞

本发明还涉及重组宿主细胞，这些重组宿主细胞包含可操作地连接至一个或多个控制序列的本发明的多核苷酸，该一个或多个控制序列指导本发明的胶原蛋白的生产。将包含多核苷酸的构建体引入宿主细胞中，这样使得该构建体作为染色体整合体或作为自主复制的染色体外载体维持。术语“宿主细胞”涵盖由于复制期间出现的突变而与亲本细胞不相同的任何亲本细胞子代。宿主细胞的选择将在很大程度上取决于编码胶原蛋白或多肽的基因及其来源。

在本文中，宿主细胞可以是酵母细胞，包括产子囊酵母(ascosporogenous yeast)(内孢霉目(*Endomycetales*))、产担子酵母(basidiosporogenous yeast)和属于半知菌纲(*Fungi Imperfecti*)(芽孢纲(*Blastomycetes*))的酵母。酵母宿主细胞可以是假丝酵母属(*Candida*)、汉逊酵母属(*Hansenula*)、克鲁维酵母菌属(*Kluyveromyces*)、毕赤酵母属(*Pichia*)、酵母菌属(*Saccharomyces*)、裂殖酵母属(*Schizosaccharomyces*)或耶氏酵母属(*Yarrowia*)细胞，如乳酸克鲁维酵母(*Kluyveromyces lactis*)、卡尔酵母(*Saccharomyces carlsbergensis*)、酿酒酵母、糖化酵母(*Saccharomyces diastaticus*)、道格拉氏酵母(*Saccharomyces douglasii*)、克鲁弗酵母(*Saccharomyces kluyveri*)、诺地酵母(*Saccharomyces norbensis*)、卵形酵母(*Saccharomyces oviformis*)或解脂耶氏酵母(*Yarrowia lipolytica*)细胞。特别的是，本发明人发现巴斯德毕赤酵母适合于生产本文所述的胶原蛋白。酵母细胞可以表达本文所述的胶原蛋白。优选的是胶原蛋白是重组III型人源化胶原蛋白，优选164.88°三螺旋结构胶原蛋白。酵母细胞是巴斯德毕赤酵母；优选是巴斯德毕赤酵母X33、GS115、SMD1168、KM71或KM71H菌株。酵母细胞可以包含酵母系统表达载体，优选为pPICZ α A、pPIC9、pPIC9K、pHIL-S1、或pYAM75P载体。表达载体可以包含多核苷酸，该多核苷酸包含 SEQ ID No. 4或5所示的核苷酸序列。

构建方法

本发明提供了构建表达胶原蛋白的酵母细胞的方法，其包括将重组表达载体导入酵母细胞中，其中重组表达载体包含编码本文的胶原蛋白的多核苷酸。重组表达载体可以包含多核苷酸包含SEQ ID No. 4或5所示的核苷酸序列。重组表达载体可以是酵母系统表达载体，优选为pPICZ α A、pPIC9、pPIC9K、pHIL-S1、或pYAM75P载体。酵母可以是巴斯德毕赤酵母，优选是巴斯德毕赤酵母X33、GS115、SMD1168、KM71或KM71H菌株。

方法可以包括以下一个或多个步骤：

- (1) 将多核苷酸插入表达载体，优选pPicza α A表达载体的NotI和XhoI酶切位点之间，以得到重组表达质粒；
- (2) 对重组表达质粒进行酶切，优选用sacI进行酶切；
- (3) 对酶切后的重组表达质粒进行纯化；
- (4) 通过电转化，将酶切后的重组表达质粒导入酵母细胞中。

发酵方法

本文提供了通过发酵生产胶原蛋白的方法。方法可以包括在合适的条件下培养根据本文所述的酵母细胞，对酵母细胞添加甲醇以进行诱导表达，并且收获培养上清液以得到胶原蛋白。培养酵母细胞的条件和培养基是本领域技术人员已知的，并且没有特别的限定。例如，培养基可以为补充有无氨基酵母氮源和生物素的BMMY培养基。

发酵方法可以包括(1)种子液培养，所述种子液为本文的酵母细胞的种子液；(2)发酵培养：将种子液接种到发酵培养基中，发酵罐培养，在适合的生长阶段开始流加甲醇，诱导表达；(3)发酵上清的制备：培养至适合的时间结束培养，取发酵液去除菌体，过滤，获得发酵上清液；

发酵方法还可以包括对胶原蛋白进行纯化的步骤。纯化的步骤是本领域技术人员已知的，并且没有特别限定。例如，纯化方法可以采用CN109593126A中描述的纯化步骤。

通过以下实施例进一步阐述本发明。应当理解本发明的范围应当以所附权利要求书为准，实施例不应解释为限制本发明的范围。

实施例

本发明通过下述实施例进一步阐明，但任何实施例或其组合不应理解为对本发明的范围或实施方式的限制。本发明的范围由所附权利要求书限定，结合本说明书和本领域一般常识，本领域普通技术人员可以清楚地明白

权利要求书所限定的范围。在不偏离本发明的精神和范围的前提下，本领域技术人员可以对本发明的技术方案进行任何修改或改变，这种修改和改变也包含在本发明的范围内。

- 核苷酸的PCR、克隆、连接等的一般方法是本领域技术人员熟知的，并且可以例如在以下文献中发现：“Molecular cloning: A laboratory manual[分子克隆：实验室手册]”，Sambrook等人(1989)，Cold Spring Harbor lab.[冷泉港实验室]，冷泉港，纽约州；Ausubel, F. M.等人(编辑)；“Current protocols in Molecular Biology[分子生物学实验指南]”，John Wiley and Sons(1995)；Harwood, C. R.和Cutting, S. M.(编辑)；“DNA Cloning: A Practical Approach, Volumes I and II [DNA克隆：实用方法，第I卷和第II卷]”，D. N. Glover编辑(1985)；“Oligonucleotide Synthesis[寡核苷酸合成]”，M. J. Gait编辑(1984)；“Nucleic Acid Hybridization[核酸杂交]”，B. D. Hames & S. J. Higgins编辑(1985)；“A Practical Guide To Molecular Cloning[分子克隆实用指南]”，B. Perbal (1984)。

15 实施例1：蛋白片段的构建、表达和筛选

1) 高分子量重组III型人源化胶原蛋白氨基酸序列：

- gergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergapgergapgfrgpagpngipgekpgagergap (SEQ ID No .2; 重复单位为gergapgfrgpagpngipgekpgagergap; SEQ ID No .1)。

25 高分子量重组III型人源化胶原蛋白蛋白核苷酸序列：

- aggtgaaagaggtgcc ccagggttta gaggtccagc aggtccaaat ggtattccag gagagaaagg tccagcagga
gaaagaggtg ccccggtga gagaggtgct cccggattca gaggtcctgc cgggtccaaat ggtattcccg gtgaaaagg
acctgcagga gagagaggtg cccagggaga gagaggagca ccaggattca gaggtccagc tgggtccaaat ggtatccctg
gagagaaggg acccgctgga gaaagaggtg cccctggaga aagaggtgcc cccggttta gaggacctgc tggacccaac
ggaatcccag gagagaaggg tccgcaggt gagagaggtg ctctggaga gagaggagca cctggattca gaggacctgc
aggacccaac ggaatacctg gtgaaaaagg tctgccggt gaaagaggtg ctccaggtga aagaggagca ccagggttta
gaggaccagc tggctctaac ggtatccctg gtgaaaaagg acccgctggt gaaagaggag cccaggtga gagaggtgct
cccggttta gaggtccagc aggtccaaac ggaatacccg gtgaaaaagg acctgctgga gagagaggtg ctctggaga
aagaggtgct cctggttca gaggtccagc aggacccaat ggaatcccag gagaaaaagg acctgcaggt gaaagaggag
ccctggtga aagaggagca cctggttta gaggaccagc aggtcctaata ggtattccag gagagaaggg acccgccgga

gaaagaggag caccgggtga aagaggagca ccaggtttca gaggaccagc tggaccaaac ggtattcccg gtgaaaaagg
accagctgga gagagagggtg caccaggaga aagagggtgt cccggtttca gaggtcacgc cggaccaaat ggtatacttg
gagaaaaagg tccagcagga gaaagagggtg caccgggtga aagagggtgca ccaggtttta gaggtcctgc cgttccaaat
ggaatccctg gtgaaaaagg acccgttgtt gaaagagggtg ctccaggaga gagaggagcc cccggtttta gaggtcctgc
tggtccaaat ggaatacctg gagagaaggg tctgtctgtt gagagagggtg cacctggaga gagagggtgcc ccaggtttca
gaggacctgc cgttcccaat ggaatacccg gaaaaaaagg tcccgagga gagagagggtg cccaggtgga aagagggtgca
cccgttttca gaggaccgc cgttctaat ggtatacttg gaaaaaaagg accagccgtt gagagagggtg ctccgggaga
gagaggagcc cctggtttca gaggtcacgc aggaccaaac ggtattccag gaaaaaagg accagcagga gagagaggag
ccccatg (SEQ ID No .4)

10 2) 中分子量重组III型人源化胶原蛋白氨基酸序列:

Gergapgfrpagpngipgekpgagergapgergapgfrpagpngipgekpgagergapgergapgfr
gpagpngipgekpgagergapgergapgfrpagpngipgekpgagergapgergapgfrpagpngipge
kpgagergapgergapgfrpagpngipgekpgagergapgergapgfrpagpngipgekpgagergap
ergapgfrpagpngipgekpgagergap (SEQ ID No .3)

15 中分子量重组III型人源化胶原蛋白蛋白核苷酸序列:

ggcgagagagggggcaccgggtttcagaggtcctgccggaccaatgggtattccaggggagaaaggccccgctggggaaacggggag
cgcttgagagcgaggtgccccggggtttcgtggtcggcgaggtccaaatggcatccctggcgaaaaggaccagccggcgagag
gggtgcaccgggagaaaggggtgtcctggggttcggcgaccgccgggtccgaacggaattctggggagaaggggccagcggg
cgaaagaggtgtcctggggagagggggcgccaggattcgggggtctgtgtggccgaacggcataccggagagaaaggccca
gcaggggagcgaggagcaccggcgaaagggggggcgccaggatttcgggaccgccggccccaacggtatccgggggaaaa
agggcctgcgggtgaagagggggccccgggcgaacgtggggcaccgggtttccgaggaccagcgggccctaacggaatacctgtg
gagaaggggtccagctggggagcggggtgcgcggggggaagaggtgcaccggcttcctgtggccagccggccctaattggtatcc
ccggcgaaaaagggcctgcaggtgaacgtgggggtcctggggagcgcggcgctcccgatttcggggccccggctggaccaaatgg
aattctggagagaaaggccccggcgggtgaacgcggagcgcgg (SEQ ID No .5)

25

上述编码核苷酸序列为商业合成。将上述编码核苷酸序列插入pPic2alαA表达载体的NotI和XhoI酶切位点之间，且目的基因在signal cleavage后端，得到重组表达质粒P-012（高分子量）、P-C3T8（中分子量）。

实施例2: 重组III型人源化胶原蛋白酵母菌改造

30 1 质粒酶切

取构建成功的表达质粒P-012和P-C3T8各20-40ug，使用sacI进行酶切，根据限制性酶切比例进行37℃酶切2-3h，酶切完成后进行核酸胶电泳检测质粒酶切前和酶切后，验证是否全部酶切，且酶切后条带位置大于酶切前。

2 质粒纯化

35 对酶切后的质粒使用质粒纯化试剂盒进行纯化去除离子等杂质，使用20-25ul灭菌的纯化水进行溶解，检测质粒浓度，将质粒放入4℃待用。

3 酵母菌改造

取出酵母感受态细胞 (X33), 在生物安全柜下进行如下操作:

- a、取0.2cm的电转杯，使用75%的酒精浸泡10min消毒，使用无菌水进行反复冲洗3次并晾干；
- b、将电转杯进行冰浴20min，开启电转仪；
- c、在灭菌的1.5ml的EP管中加入80ul的感受态细胞和5-15ug的酶切后的质粒，使用移液枪反复混匀；
- 5 d、将100ul的混合液转移至电转杯中，冰浴5min后，擦拭电转杯外表面，将电转杯转移至电转仪槽中，点击开始，进行电击；
- e、电转结束后，在电转杯中加入1ml的1M的无菌山梨醇溶液，使用封口膜密封，30℃进行静置孵育2h；
- 10 f、将电转杯中的混合液转移至灭菌的10ml EP管中，然后在EP管中加入1ml的YPD液体培养基（酵母蛋白胨0.2g/L，酵母浸粉0.1g/L）进行30℃，300rpm培养1h。
- g、铺平板：在生物安全柜下，将转化并孵育后的菌液分别取200ul加入倒好的1-3号平板中，4号平板作为对照，使用无菌涂布棒进行均匀涂布平板，
- 15 完成后进行30℃平板培养2-5天左右。
- h、菌种培养：平板培养至长出菌斑，将10ml无水葡萄糖溶液加入40ml的YPD液体培养基中，待培养基温度降至室温时；将50ml培养基按5ml/管分装于10个50ml生物反应管中，按比例2000mg/L加入博来霉素溶液，挑取平板上10个单克隆分别接种于10个小管中进行30℃，300rpm过夜培养；准备
- 20 BMGY培养基200ml，灭菌完成后，在摇瓶中加入20ml 10*YNB储液和200ul 500*生物素储液，混合均匀；将200ml培养基进行分装置10个50ml的离心管中，20ml/管；然后将培养好的菌液，分别取4ml接种于10个20ml/管的离心管中，过夜培养至菌液OD值在3-10之间。
- i、诱导表达：准备BMMY培养基500ml，灭菌完成后，在摇瓶中加入50ml
- 25 10*YNB储液，500ul 500*生物素储液，混合均匀；将500ml培养基进行分装置10个250ml的摇瓶中，50ml/瓶；将培养好的菌液进行1500rpm离心10min，弃掉上清液；分别用10个摇瓶中的培养基将10个离心管中的菌体沉淀进行悬浮，并转移至各自的摇瓶中，30℃，300rpm进行摇瓶培养；每隔24h，分别对10个摇瓶进行取样1ml至1.5ml EP管中，-20℃保存；并向摇瓶中添加1%的
- 30 甲醇（使用0.22um滤芯进行过滤）进行诱导表达，直至第5天结束；将1.5ml EP管中的样品，进行12000rpm离心5min，取上清进行电泳检测，分别观察

10个克隆的表达情况。

4 电泳检测

具体过程为：取样品液40 μ l，加入10 μ l 5 $\times$ 的蛋白上样缓冲液(250mM的Tris-HCl(pH:6.8)，10 % SDS , 0.5 % 溴酚蓝，50 % 甘油，5 % β -巯基乙醇)，置于100 $^{\circ}$ C沸水中煮10 min，然后每孔10 μ l加入SDS-PAGE蛋白胶中，电压80V跑2h后，用考马斯亮蓝染色液(0.1 % 考马斯亮蓝R-250，25 % 异丙醇，10 % 冰醋酸)进行蛋白染色20 min，再利用蛋白脱色液(10 % 醋酸，5 % 乙醇)进行脱色。电泳检测结果见图1、2：P-012目的蛋白理论分子量为44.75KD，表观分子量为50KD；P-C3T8目的蛋白理论分子量为22.39KD，表观分子量为28KD。

10 电泳检测结果显示酵母菌改造成功。

实施例3：重组III型人源化胶原蛋白的质谱检测

1 实验仪器：

- 1) 高分辨质谱仪: XevoG2-XS QToF (Waters)
- 2) 超高效液相色谱:UPLC(Acquity UPLC I-Class) (Waters)

15 2 材料和试剂

- 1) Guanidine HCl(Sigma)
- 2) Urea (Bio-Rad)
- 3) Tris-base (Bio-Rad)
- 4) DTT (Bio-Rad)
- 20 5) IAM (Sigma)
- 6) Zeba Spin column (Pierce)
- 7) ACQUITY UPLC Peptide BEH C18 Column, 300 $\text{\AA}$, 1.7 μ m, 2.1 mm X 150 mm(Waters)
- 8) UNIFI(Waters)
- 25 9) Trypsin(Promega)
- 10) Chymotrypsin(Sigma)
- 11) Glu-C(Wako)
- 12) LysC(Wako)

3 实验方法

30 1) Trypsin、Chymotrypsin、Glu-C酶解：取适量供试品经适当前处理后加入Trypsin、Chymotrypsin、Glu-C，37 $^{\circ}$ C酶切20小时。

2) 高效液相色谱: 供试品经酶解处理后采超高效液相系统Acquity UPLC I-Class进行分离。液相A液为0.1% FA水溶液, B液为0.1% FA 乙腈溶液。供试品由自动进样器上样到柱, 再经色谱柱分离, 柱温为55℃, 流速为300μl/min, TUV检测器波长为214nm。相关液相梯度如下:

	Time/min	A/%	B%
1	3	100	0
2	63	60	40
3	63.1	2	98
4	66	2	98
5	66.1	100	0
6	75	100	0

5 3) 质谱鉴定: 供试品经超高效液相色谱脱盐及分离后用XevoG2-XS QToF质谱仪(Waters)进行质谱检测分析。分析时长:63min, 检测方式:正离子, MS, 扫描范围(m/z):300-2000。

4) 质谱数据处理:原始使用UNIFI(1.8.2,Waters)软件查库,主要参数如(表1):

Item	Value
Enzyme	Trypsin、Chymotrypsin、Glu-C
modifications	Deamidated(NQ) 、 Oxidation(M) 、 Carbamidomethyl(C)
M/Z tolerance	±15ppm
Fragment tolerance	Fragment tolerance
Database	20200604366.fasta
Filter	Minimum fragment ions:3

10 4 实验结果和分析

在使用Trypsin、Chymotrypsin、Glu-C分别对供试品进行溶液内酶解后, 得到的肽段样品,经过LC-MS/MS设备的分析,得到的原始数据经过UNIFI软件进行查库。

1) 质谱检测出P-012分子量及对应多肽

观察值	Mr (预期值)	肽
1171.2805	1170.1811	GERGAPGFRGPA

1660.8642	1659.8569	GPAGPNGIPGEKGPAGER
1678.8994	1677.8921	GAPGFRGPAGPNGIPGEK
2244.0007	2242.9934	GPAGPNGIPGEKGPAGERGAPGER
2244.3176	2243.3104	GPAGPNGIPGEKGPAGERGAPGER
2246.2154	2245.2081	GAPGFRGPAGPNGIPGEKGPAGER

检出多肽片段的覆盖率为100%。

2) 质谱检测出P-C3T8分子量及对应多肽

观察值	Mr (预期值)	肽
1706.8815	1705.5321	GERGAPGFRGPAGPNGIP
1093.5637	1092.5564	GPAGPNGIPGEK
1660.8636	1659.8563	GPAGPNGIPGEKGPAGER
1678.8907	1677.8834	GAPGFRGPAGPNGIPGEK
2244.3176	2243.3104	GPAGPNGIPGEKGPAGERGAPGER
2246.1276	2245.1203	GAPGFRGPAGPNGIPGEKGPAGER

检出多肽片段的覆盖率为100%。

酶解后的样品经过LC-MS/MS分析后查库结果进行整合，最终得到样品的肽段覆盖率为100%，检测结果非常可信。

实施例4: 重组III型人源化胶原蛋白的细胞粘附活性检测

胶原蛋白的活性检测方法可以参考文献Juming Yao, Satoshi Yanagisawa, Tetsuo Asakura, Design, Expression and Characterization of Collagen-Like Proteins Based on the Cell Adhesive and Crosslinking Sequences Derived from Native Collagens, J Biochem. 136, 643-649(2004)。具体实施方法如下:

(1) 利用紫外吸收法检测待测蛋白样品的浓度, 包括牛I型胶原蛋白(中国食品药品检定研究院, 编号: 380002)、本发明提供的重组人源化蛋白。具体为分别测定样品在215nm和225nm下的紫外光吸收, 利用经验公式 $C(\mu\text{g/mL})=144\times(A_{215}-A_{225})$ 计算蛋白质浓度, 注意需在 $A_{215}<1.5$ 的情况下检测。该方法的原理是: 测定肽键在远紫外光下的特征吸收, 不受生色团含量的影响, 干扰物质少, 操作简便, 适合检测考马斯亮蓝不显色的人胶原蛋白及其类似物。(参考文献为Walker JM. The Protein Protocols Handbook, second edition. HumanaPress. 43-45.)。检测完蛋白浓度后, 用PBS将所有待测蛋白浓度调整到0.5mg/mL。

(2) 向96孔板中加入100μL各种蛋白溶液和空白PBS溶液对照，室温静置60min。

(3) 每孔中加入105个培养状态良好的NIH/3T3细胞，37℃孵育60min。

(4) 每孔用PBS清洗4次。

5 (5) 用LDH检测试剂盒(Roche, 04744926001)检测 OD492nm的吸光度。根据空白对照的数值，可以计算出细胞的贴壁率。计算公式如下：细胞贴壁率= $\frac{\text{OD}_{492\text{nm}} - \text{OD}_{492\text{nm}}^{\text{blank}}}{\text{OD}_{492\text{nm}}^{\text{max}}} \times 100\%$ 。细胞的贴壁率即可以反应胶原蛋白的活性。蛋白的活性越高，越能在短时间给细胞提供优质的外环境，帮助细胞贴壁。检测结果见下图3、4。

10 实施例5: 重组III型人源化胶原蛋白圆二色谱紫外扫描分析
实验方法

(1) 仪器参数设定

带宽(Band width): 1.0nm

步(Step) 1.0nm

15 测量范围: 190-260nm(远-UV区扫描)/250-340nm(近-UV区扫描)

每个点的时间(Time-per-point) 0.5s

重复(Repeats) 3次

池长度(Cell Length) 10mm|0.5mm

温度 室温

20 (2) 标准品远近紫外扫描

设定扫描波长180-340nm进行背景测试、空白缓冲液测试，然后采集1mg/mL CSA 标准品溶液在180-340nm范围的圆二色远近紫外吸收。

(3) 样品处理

25 取P-012、P-C3T8蛋白样品，用10KD的超滤浓缩管（密理博）将蛋白分别浓缩至蛋白浓度为1mg/ml。

(4) 样品远紫外扫描

将比色皿用2M HNO3浸泡过夜，去离子水冲洗干净后晾干，先采集背景，再采集空白缓冲液，然后在比色皿中加适量供试品按照上述参数进行190-260nm的远紫外扫描并采集数据。

30 (5) 样品近紫外扫描

将比色皿用2M HNO3浸泡过夜，去离子水冲洗干净后晾干，先采集背

景，再采集空白缓冲液，然后在比色皿中加入适量供试品按照上述参数进行250-340nm的近紫外扫描并采集数据。

(6) 扫描图谱处理

- 5 对扫描后的所有图谱用软件Pro-Data Viewer进行subtract baseline、smoothing处理。

实验结果和分析

结果显示：P-012、P-C3T8在221nm处均有正峰，表明该蛋白均具有三螺旋结构，结果见图5、图6。

权利要求书

1. 一种构建表达胶原蛋白的酵母细胞的方法,其包括将重组表达载体导入酵母细胞中,其中重组表达载体包含编码胶原蛋白的多核苷酸,所述胶原蛋白包含以SEQ ID No .1所示的序列的重复单元,重复单元的数目是2-32,重复单元是直接连接的;优选地,胶原蛋白是重组III型人源化胶原蛋白,优选164.88°三螺旋结构胶原蛋白。
2. 根据权利要求1所述的方法,其中胶原蛋白包含SEQ ID No .2或3所示的氨基酸序列;优选地,多核苷酸包含SEQ ID No. 4或5所示的核苷酸序列。
3. 根据权利要求1或2所述的方法,其中重组表达载体是酵母系统表达载体,优选为pPICZ α A、pPIC9、pPIC9K、pHIL-S1、或pYAM75P载体。
4. 根据权利要求1-3中任一项所述的方法,其中酵母是巴斯德毕赤酵母;优选是巴斯德毕赤酵母X33、GS115、SMD1168、KM71或KM71H菌株。
5. 根据权利要求1-4中任一项所述的方法,其包括以下一个或多个步骤:
- (1) 将多核苷酸插入表达载体,优选pPicZ α A表达载体的NotI和XhoI酶切位点之间,以得到重组表达质粒;
- (2) 对重组表达质粒进行酶切,优选用sacI进行酶切;
- (3) 对酶切后的重组表达质粒进行纯化;
- (4) 通过电转化,将酶切后的重组表达质粒导入酵母细胞中。
6. 表达胶原蛋白的酵母细胞,其中所述胶原蛋白包含以SEQ ID No .1所示的序列的重复单元,重复单元的数目是2-32,重复单元是直接连接的;酵母细胞是巴斯德毕赤酵母;优选是巴斯德毕赤酵母X33、GS115、SMD1168、KM71或KM71H菌株;优选地,胶原蛋白包含SEQ ID No .2或3所示的氨基酸序列;优选地,胶原蛋白是重组III型人源化胶原蛋白,优选164.88°三螺旋结构胶原蛋白。
7. 根据权利要求6所述的酵母细胞,其包含酵母系统表达载体,优选为pPICZ α A、pPIC9、pPIC9K、pHIL-S1、或pYAM75P载体;表达载体包含多核苷酸,该多核苷酸包含SEQ ID No. 4或5所示的核苷酸序列。
8. 根据权利要求6或7所述的酵母细胞用于制备胶原蛋白的用途,优选地,胶原蛋白包含以SEQ ID No .1所示的序列的重复单元,重复单元的数目是2-32,重复单元是直接连接的;优选地,胶原蛋白包含SEQ ID No .2或3所示的氨基酸序列;优选地,胶原蛋白是重组III型人源化胶原蛋白,优选164.88°

三螺旋结构胶原蛋白。

9. 通过发酵生产胶原蛋白的方法,其包括在合适的条件下培养根据权利要求6或7所述的酵母细胞,对酵母细胞添加甲醇以进行诱导表达,并且收获培养上清液以得到胶原蛋白;优选地,培养基为补充有无氨基酵母氮源和生物素的BMMY培养基。

10. 根据权利要求9所述的方法,其包括:

(1)种子液培养,所述种子液为权利要求6或7的酵母细胞的种子液;

(2)发酵培养:将种子液接种到发酵培养基中,发酵罐培养,在适合的生长阶段开始流加甲醇,诱导表达;

- 10 (3)发酵上清的制备:培养至适合的时间结束培养,取发酵液去除菌体,过滤,获得发酵上清液;

优选地,胶原蛋白包含以SEQ ID No .1所示的序列的重复单元,重复单元的数目是2-32,重复单元是直接连接的;优选地,胶原蛋白包含SEQ ID No .2或3所示的氨基酸序列;优选地,胶原蛋白是重组III型人源化胶原蛋白,

- 15 优选164.88°三螺旋结构胶原蛋白。