

自交联重组人源化胶原蛋白高分子生物材料及其制备方法

相关申请的引用

5 本公开要求于2023年08月04日提交中国专利局、申请号为202310983181.3、发明名称为“自交联重组人源化胶原蛋白高分子生物材料及其制备方法”的发明专利申请的优先权，通过引用将其全部内容结合在本公开中。

技术领域

10 本公开属于合成生物技术领域，具体涉及自交联重组人源化胶原蛋白高分子生物材料及其制备方法；更具体涉及一种促进重组III型人源化胶原蛋白自交联形成三螺旋结构的胶原蛋白自组装元件及相应的具有三螺旋结构的重组III型人源化胶原蛋白。

背景技术

15 胶原蛋白是广泛分布于人体结缔组织的一类蛋白质，也是人体含量最多的蛋白质，可以占到蛋白质总量的25%~35%。胶原蛋白的主要功能体现在维护细胞外环境，维持组织器官的正常生理功能，以及修复肌体损伤等方面。胶原蛋白是天然的生物资源，具有其它高分子材料所望尘莫及的生物组织相容性，对细胞的支撑弹性和可降解性。因此，胶原蛋白可以广泛的应用在医药和化妆品等行业中。

20 天然胶原分子可形成一种特殊的超螺旋结构，这种结构是以3个氨基酸残基为基本重复体的左手螺旋，这三个氨基酸残基通常为Gly-X-Pro。Gly对胶原蛋白中氢键的形成是必须的，其本身没有侧链，从而使胶原蛋白紧密堆积。在更高级的结构层次上，胶原超螺旋会进一步缔合成为胶原原纤维。在生物体中，胶原蛋白的合成和修饰从原胶原开始，经历了羟基化、糖基化、相互交联等诸多化学变化，受到了多种生物酶的复杂调控。原胶原除了含有胶原链之外，还含有球状的头部和尾部。没有这些头部和尾部，胶原链就不会折叠成为正确的三螺旋，从而缺乏胶原蛋白的生物学活性。因此，按照原始基因序列制备的胶原蛋白在体外很难自发的组织形成正确的空间结构。这样的困难严重阻碍了人胶原蛋白的研发和生产。

25 人们对此进行了适当的探索研究，例如引用文献1公开可以在表达多肽的C末端加入铰链区氨基酸GPPGPCCGGG（SEQ ID NO.15）等序列帮助成胶。然而引用文献2公开，如果在其提供的多肽序列的C末端加入上述如SEQ ID NO.15所示的铰链区氨基酸时，会使得该多肽在发酵培养的时候形成胶状沉淀导致无法溶解提纯，得率大大降低。

30 因此，目前亟需一种能够促进重组III型人源化胶原蛋白交联形成三螺旋结构的方法，使得重组III型人源化胶原蛋白能够通过生物合成的方法大量生产，且能够具有三螺旋结构并发挥其特定的功能。

35 此外，圆二色光谱技术是用来检测蛋白质二级结构的一种方法。其检测天然胶原蛋白时，在合适的检测条件下会出现221nm附近的吸收峰。这个吸收峰与其三螺旋结构存在一定关系，但同时受到脯氨酸、羟脯氨酸含量等的影响（多聚脯氨酸本身就有强烈的221nm峰，但是不形成三螺旋）。因此圆二色光谱不能作为三螺旋结构的直接证据。但是由于圆二色光谱实验检测简单，因此可用于辅助表征胶原蛋白高级结构。针对不同的胶原蛋白，圆二色光谱实验的条件和结果差异很大，主要原因是重组胶原蛋白的序列长度和氨基酸组成（尤其是脯氨酸、羟脯氨酸的含量）与天然全长胶原蛋白不同，因此不同的胶原蛋白的圆二色光谱检测条件（温度、浓度、溶液配方等）无法保持一致。为了表征胶原蛋白能否形成三螺旋（比如生产中变性的胶原蛋白或者设计有缺陷的胶原蛋白完全无法形成三螺旋结构），可通过圆二色光谱在不同条件下对胶原蛋白进行表征。

40 参考文献：

引用文献1： Yao J, Yanagisawa S, Asakura T. Design, expression and characterization of collagen-like proteins based on the cell adhesive and crosslinking sequences derived from native collagens. J Biochem. 2004 Nov;136(5):643-9.

45 引用文献2： CN 109593126 B。

发明内容

发明要解决的问题

目前，现有技术已经对例如按照原始基因序列制备的胶原蛋白在体外很难自发的组织形成正确的空间结构，导致其无法有效发挥生理功能的问题进行了适当的研究，但其仍然不能说是充分的。对此，本公开提供了一种胶原蛋白自组装元件，其能够促进重组Ⅲ型人源化胶原蛋白自交联形成三螺旋结构，通过使重组Ⅲ型人源化胶原蛋白的氨基酸序列的末端含有人Ⅲ型胶原蛋白铰链区的氨基酸序列，帮助重组Ⅲ型人源化胶原蛋白自交联形成三螺旋结构，且该具有三螺旋结构的重组Ⅲ型人源化胶原蛋白生物材料能够通过生物合成的方法顺利制备。

用于解决问题的方案

本公开第一方面提供了一种胶原蛋白自组装元件，所述胶原蛋白自组装元件包含如下(i)~(iii)中任一项所示的序列：

(i) 如SEQ ID NO.1所示的氨基酸序列的C端的至少21个连续的氨基酸；

(ii) 如(i)所述的氨基酸序列中取代、缺失或添加一个或多个氨基酸，且保留促进重组Ⅲ型人源化胶原蛋白自交联形成三螺旋结构的功能的氨基酸序列；

(iii) 与(i)所述的氨基酸序列具有至少90%、92%、95%、96%、97%、98%或99%同一性，且保留促进重组Ⅲ型人源化胶原蛋白自交联形成三螺旋结构的功能的氨基酸序列。

在一些实施方案中，所述(i)所示的序列为如SEQ ID NO.1、SEQ ID NO.2、SEQ ID NO.3或SEQ ID NO.4所示的氨基酸序列。

本公开第二方面提供了根据本公开第一方面所述的胶原蛋白自组装元件在用于促进重组Ⅲ型人源化胶原蛋白自交联形成三螺旋结构中的用途。

在一些实施方案中，所述胶原蛋白自组装元件通过存在于所述重组Ⅲ型人源化胶原蛋白的氨基酸序列的末端发挥作用；

优选地，所述重组Ⅲ型人源化胶原蛋白的氨基酸序列的末端为C末端。

本公开第三方面提供了一种重组Ⅲ型人源化胶原蛋白，所述重组Ⅲ型人源化胶原蛋白的末端包含根据权利要求1或2所述的胶原蛋白自组装元件；

优选地，所述重组Ⅲ型人源化胶原蛋白的末端为C末端。

在一些实施方案中，所述重组Ⅲ型人源化胶原蛋白包含如下(a)~(c)中任一项所示的序列：

(a) 如SEQ ID NO.6、SEQ ID NO.7、SEQ ID NO.8或SEQ ID NO.9所示的氨基酸序列；

(b) 如(i)所述的氨基酸序列中取代、缺失或添加一个或多个氨基酸，且保留自交联形成三螺旋结构能力的氨基酸序列；

(c) 与(i)所述的氨基酸序列具有至少90%、92%、95%、96%、97%、98%或99%同一性，且保留自交联形成三螺旋结构能力的氨基酸序列。

本公开第四方面提供了多核苷酸，所述多核苷酸编码根据本公开第一方面所述的胶原蛋白自组装元件，或编码根据本公开第三方面所述的重组Ⅲ型人源化胶原蛋白。

在一些实施方案中，所述多核苷酸包含如下(z₁)或(z₂)所示的序列：

(z₁) 如SEQ ID NO.11、SEQ ID NO.12、SEQ ID NO.13或SEQ ID NO.14所示的核苷酸序列；

(z₂) 在严格条件下与(z₁)所示的核苷酸序列杂交且编码具有自交联形成三螺旋结构能力的蛋白的核苷酸序列，所述严格条件是中等严格条件，中-高严格条件，高严格条件或非常高严格条件。

本公开第五方面提供了重组表达载体，所述重组表达载体包含根据本公开第四方面所述的多核苷酸。

本公开第六方面提供了重组宿主细胞，所述重组宿主细胞包含根据本公开第五方面所述的重组表达载体。

本公开第七方面提供了根据本公开第三方面所述的重组Ⅲ型人源化胶原蛋白的制备方法，所述制备方法包括如下步骤：

S1: 构建包含根据第四方面所述的编码根据第三方面所述的重组Ⅲ型人源化胶原蛋白的多核苷酸的重组表达载体, 并通过转化构建重组宿主细胞;

S2: 在培养基中培养S1获得的所述重组宿主细胞并生产蛋白;

S3: 收获并纯化所述蛋白, 优选用Ni柱和/或阴离子交换层析纯化所述蛋白; 和

S4: 任选地对所述蛋白进行酶切, 优选用TEV蛋白酶酶切蛋白。

发明的效果

通过上述技术方案的实施, 本公开取得了如下技术效果:

首先, 本公开提供了一种胶原蛋白自组装元件, 通过使重组Ⅲ型人源化胶原蛋白的氨基酸序列的末端含有该胶原蛋白自组装元件, 也即含有人Ⅲ型胶原蛋白铰链区的氨基酸序列, 可以有效的帮助重组Ⅲ型人源化胶原蛋白自交联形成三螺旋结构, 大大提高了通过生物合成方法制备的重组Ⅲ型人源化胶原蛋白的可利用性。

同时, 本公开提供的末端包含胶原蛋白自组装元件, 也即包含人Ⅲ型胶原蛋白铰链区的氨基酸序列的重组Ⅲ型人源化胶原蛋白可以形成良好的三螺旋结构, 该结果得到了圆二色谱检测结果的支持。并且该重组Ⅲ型人源化胶原蛋白能够通过生物合成的方法大量制备。

附图说明

图1为重组Ⅲ型人源化胶原蛋白TE16c制备过程中的电泳检测结果。

图2为重组Ⅲ型人源化胶原蛋白C3T16H1制备过程中的电泳检测结果。

图3为重组Ⅲ型人源化胶原蛋白C3T16H2制备过程中的电泳检测结果。

图4为重组Ⅲ型人源化胶原蛋白C3T16H3制备过程中的电泳检测结果。

图5为重组Ⅲ型人源化胶原蛋白C3T16H4制备过程中的电泳检测结果。

图6为重组Ⅲ型人源化胶原蛋白C3T16H1、C3T16H2、C3T16H3、C3T16H4和TE16c的圆二色谱紫外扫描分析结果, 检测到221 nm附近的正峰。

具体实施方式

以下对本公开的实施方式进行说明, 但本公开不限于此。本公开不限于以下说明的各构成, 在公开请求保护的范围内可以进行各种变更, 而适当组合不同实施方式、实施例中各自公开的技术手段而得到的实施方式、实施例也包含在本公开的技术范围中。

在本公开中, 使用“数值A~数值B”或“数值A-数值B”表示的数值范围是指包含端点数值A、B的范围。

在本公开中, 使用“以上”或“以下”表示的数值范围是指包含本数的数值范围。

在本公开中, “任选的”或“任选地”是指接下来描述的事件或情况可发生或可不发生, 并且该描述包括该事件发生的情况和该事件不发生的情况。

在本公开中, 术语“一(a)”或“一(an)”或“一(the)”可以指“一个”, 也可以指“一个或多个”、“至少一个”以及“一个或多于一个”。

在本公开中, 术语“包含”、“具有”、“包括”或“含有”可以指包括在内的或开放式的, 并不排除额外的、未引述的元件或方法步骤。与此同时, “包含”、“具有”、“包括”或“含有”也可以表示封闭式的, 排除额外的、未引述的元件或方法步骤。

在本公开中, 术语“约”可以表示: 一个值包括测定该值所使用的装置或方法的误差的标准偏差。用以界定本公开的数值范围与参数皆是约略的数值, 此处已尽可能精确地呈现具体实施例中的相关数值。然而, 任何数值本质上不可避免地含有因前述测试装置或方法所致的标准偏差。因此, 除非另有明确的说明, 应当理解本公开所用的所有范围、数量、数值与百分比均经过“约”的修饰。在此处, “约”通常是指实际数值在某一特定数值或范围的±10%、±5%、±3%、±1%或±0.5%之内。

在本公开中, 术语“多肽”、“蛋白”可互换地指通过共价键(例如肽键)相互连接的一串至少两个氨基酸残基, 可以是重组多肽、天然多肽或合成多肽。多肽可以是线形或分支的, 它可以包含修

饰的氨基酸，并且它可以由非氨基酸隔断。该术语也包括已经被修饰（例如，二硫键形成、糖基化、脂质化、乙酰化、磷酸化或任何其他操作，如以标记组分缀合）的氨基酸聚合物。

在本公开中，术语“氨基酸”可以包括天然氨基酸、非天然氨基酸、氨基酸类似物以及所有它们的D和L立体异构体。

5 在本公开中，氨基酸修饰可以包括对天然序列的修饰，例如官能团的修饰、分子内共价键合（例如，侧链之间成环）、甲基化、酰基化、泛素化、磷酸化、氨基己烷化、生物素化等。

在本公开中，氨基酸缺失可指从氨基酸序列中删除1、2或3个以上氨基酸，只要改变后的序列完全或部分保留原有氨基酸序列的活性。

10 在本公开中，氨基酸添加可指在氨基酸序列的C端、N端或C端与N端中间的任意位置处添加1、2或3个以上氨基酸，只要改变后的序列完全或部分保留原有氨基酸序列的活性。

在本公开中，氨基酸取代可指在氨基酸序列的某个位置的氨基酸被其他氨基酸替代，只要改变后的序列完全或部分保留原有氨基酸序列的活性。氨基酸取代可以是保守氨基酸取代，指与原有氨基酸序列相比，若干个氨基酸被性质相似或相近的氨基酸所替换而形成肽（保守性变异肽）。示例性的，这些保守性变异肽可以根据下列氨基酸替换而产生：Val、Leu或Ile对Ala的取代，Lys、Gln、Asn或His对Arg的取代，Gln、His、Lys或Arg对Asn的取代，Glu或Asn对Asp的取代，Ser或Ala对Cys的取代，Asn或Glu对Gln的取代，Asp或Gln对Glu的取代，Ala对Gly的取代，Asn、Lys、Gln或Arg对His的取代，Leu、Met、Ala、Val、Phe或正亮氨酸对Ile的取代，Ile、Met、Ala、Val、Phe或正亮氨酸对Leu的取代，Asn、Gln或Arg对Lys的取代，Ile、Leu或Phe对Met的取代，Leu、Val、Ile、Ala或Tyr对Phe的取代，Ala对Pro的取代，Thr对Ser的取代，Ser或Val对Thr的取代，Phe或Tyr对Trp的取代，Trp、Phe、Thr或Ser对Tyr的取代，及Phe、Ala、Met、Ile、Leu或正亮氨酸对Val的取代。氨基酸取代也可以是非保守氨基酸取代。

15 His对Arg的取代，Gln、His、Lys或Arg对Asn的取代，Glu或Asn对Asp的取代，Ser或Ala对Cys的取代，Asn或Glu对Gln的取代，Asp或Gln对Glu的取代，Ala对Gly的取代，Asn、Lys、Gln或Arg对His的取代，Leu、Met、Ala、Val、Phe或正亮氨酸对Ile的取代，Ile、Met、Ala、Val、Phe或正亮氨酸对Leu的取代，Asn、Gln或Arg对Lys的取代，Ile、Leu或Phe对Met的取代，Leu、Val、Ile、Ala或Tyr对Phe的取代，Ala对Pro的取代，Thr对Ser的取代，Ser或Val对Thr的取代，Phe或Tyr对Trp的取代，Trp、Phe、Thr或Ser对Tyr的取代，及Phe、Ala、Met、Ile、Leu或正亮氨酸对Val的取代。氨基酸取代也可以是非保守氨基酸取代。

20 Ser对Tyr的取代，及Phe、Ala、Met、Ile、Leu或正亮氨酸对Val的取代。氨基酸取代也可以是非保守氨基酸取代。

在本公开中，术语“序列同一性”和“同一性百分比”指两个或更多个多核苷酸或多肽之间相同（即同一）的核苷酸或氨基酸的百分比。两个或更多个多核苷酸或多肽之间的序列同一性可通过以下方法测定：将多核苷酸或多肽的核苷酸或氨基酸序列对准且对经对准的多核苷酸或多肽中含有相同核苷酸或氨基酸残基的位置数目进行评分，且将其与经对准的多核苷酸或多肽中含有不同核苷酸或氨基酸残基的位置数目进行比较。多核苷酸可例如通过含有不同核苷酸（即取代或突变）或缺失核苷酸（即一个或两个多核苷酸中的核苷酸插入或核苷酸缺失）而在一个位置处不同。多肽可例如通过含有不同氨基酸（即取代或突变）或缺失氨基酸（即一个或两个多肽中的氨基酸插入或氨基酸缺失）而在一个位置处不同。序列同一性可通过用含有相同核苷酸或氨基酸残基的位置数目除以多核苷酸或多肽中氨基酸残基的总数来计算。举例而言，可通过用含有相同核苷酸或氨基酸残基的位置数目除以多核苷酸或多肽中核苷酸或氨基酸残基的总数且乘以100来计算同一性百分比。

25 将多核苷酸或多肽的核苷酸或氨基酸序列对准且对经对准的多核苷酸或多肽中含有相同核苷酸或氨基酸残基的位置数目进行评分，且将其与经对准的多核苷酸或多肽中含有不同核苷酸或氨基酸残基的位置数目进行比较。多核苷酸可例如通过含有不同核苷酸（即取代或突变）或缺失核苷酸（即一个或两个多核苷酸中的核苷酸插入或核苷酸缺失）而在一个位置处不同。多肽可例如通过含有不同氨基酸（即取代或突变）或缺失氨基酸（即一个或两个多肽中的氨基酸插入或氨基酸缺失）而在一个位置处不同。序列同一性可通过用含有相同核苷酸或氨基酸残基的位置数目除以多核苷酸或多肽中氨基酸残基的总数来计算。举例而言，可通过用含有相同核苷酸或氨基酸残基的位置数目除以多核苷酸或多肽中核苷酸或氨基酸残基的总数且乘以100来计算同一性百分比。

30 序列同一性可通过用含有相同核苷酸或氨基酸残基的位置数目除以多核苷酸或多肽中氨基酸残基的总数来计算。举例而言，可通过用含有相同核苷酸或氨基酸残基的位置数目除以多核苷酸或多肽中核苷酸或氨基酸残基的总数且乘以100来计算同一性百分比。

在本公开中，术语“中等严格条件”，“中-高严格条件”，“高严格条件”或“非常高严格条件”，其描述了核酸杂交和洗涤的条件。进行杂交反应的指导参见Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989)，6.3.1-6.3.6，其通过引用并入本文。在该文献中描述了含水的和非含水的方法，且可以使用任一种。例如，具体的杂交条件如下：（1）低严格性杂交条件在6×氯化钠/柠檬酸钠(SSC)中，在约45℃，然后在至少50℃，在0.2×SSC，0.1%SDS中洗涤2次（对于低严格性条件，可以将洗涤温度升高到55℃）；（2）中等严格性杂交条件在6×SSC，在约45℃，然后在60℃，在0.2×SSC，0.1%SDS中洗涤1次或多次；（3）高严格性杂交条件在6×SSC，在约45℃，然后在65℃，在0.2×SSC，0.1%SDS中洗涤1次或多次且优选；（4）非常高的严格性杂交条件是0.5M磷酸钠，7%SDS，在65℃，然后在65℃，在0.2×SSC，1%SDS中洗涤1次或多次。

35 低严格性杂交条件在6×氯化钠/柠檬酸钠(SSC)中，在约45℃，然后在至少50℃，在0.2×SSC，0.1%SDS中洗涤2次（对于低严格性条件，可以将洗涤温度升高到55℃）；（2）中等严格性杂交条件在6×SSC，在约45℃，然后在60℃，在0.2×SSC，0.1%SDS中洗涤1次或多次；（3）高严格性杂交条件在6×SSC，在约45℃，然后在65℃，在0.2×SSC，0.1%SDS中洗涤1次或多次且优选；（4）非常高的严格性杂交条件是0.5M磷酸钠，7%SDS，在65℃，然后在65℃，在0.2×SSC，1%SDS中洗涤1次或多次。

40 非常高的严格性杂交条件是0.5M磷酸钠，7%SDS，在65℃，然后在65℃，在0.2×SSC，1%SDS中洗涤1次或多次。

在本公开中，术语“多核苷酸”指任何长度的核苷酸的聚合形态，不论是脱氧核糖核苷酸或核糖核苷酸，或其类似物。多核苷酸可具有任何三维结构，可实施任何已知或未知的功能。

在本公开中，术语“重组表达载体”指用于表达例如编码所需多肽的多核苷酸的DNA结构。重组表达载体可包括，例如包含 i) 对基因表达具有调控作用的遗传元素的集合，例如启动子和增强子；ii) 转录成mRNA并翻译成蛋白质的结构或编码序列；以及iii) 适当的转录和翻译起始和终止序列的

45 ii) 转录成mRNA并翻译成蛋白质的结构或编码序列；以及iii) 适当的转录和翻译起始和终止序列的

转录亚单位。重组表达载体以任何合适的方式构建。载体的性质并不重要，并可以使用任何载体，包括质粒、病毒、噬菌体和转座子。用于本公开的可能载体包括但不限于染色体、非染色体和合成DNA序列，例如病毒质粒、细菌质粒、噬菌体DNA、酵母质粒以及从质粒和噬菌体DNA的组合中衍生的载体，来自如慢病毒、逆转录病毒、牛痘、腺病毒、鸡痘、杆状病毒、SV40和伪狂犬病等病毒的DNA。

5 在本公开中，术语“重组宿主细胞”是指已向其中引入了表达载体的细胞。宿主细胞可包括细菌、微生物、植物或动物细胞。易于转化的细菌包括肠杆菌科的成员，例如大肠杆菌或沙门氏菌的菌株；芽孢杆菌科例如枯草芽孢杆菌；肺炎球菌；链球菌和流感嗜血菌。适当的微生物包括酿酒酵母和毕赤酵母。适当的动物宿主细胞系包括CHO(中国仓鼠卵巢细胞系)和NS0细胞。

在本公开中，术语“密码子优化”是指编码多肽的核苷酸序列已被配置为包含宿主细胞或生物体
10 优选的密码子，以改善宿主细胞或生物体中的基因表达并提高翻译效率。

除非另有定义，本公开所用的其他技术和科学术语具有与本公开所属技术领域中的普通技术人员所通常理解的相同含义。

以下对本公开提供的技术方案进行更具体的描述：

胶原蛋白自组装元件

15 本公开提供了一种胶原蛋白自组装元件，其能够有效促进重组III型人源化胶原蛋白自交联形成三螺旋结构。

在一些实施方案中，所述胶原蛋白自组装元件包含如下(i)~(iii)中任一项所示的序列：

(i) 如SEQ ID NO.1所示的氨基酸序列的C端的至少21个连续的氨基酸；

20 (ii) 如(i)所述的氨基酸序列中取代、缺失或添加一个或多个氨基酸，且保留促进重组III型人源化胶原蛋白自交联形成三螺旋结构的功能的氨基酸序列；

(iii) 与(i)所述的氨基酸序列具有至少90%、92%、95%、96%、97%、98%或99%同一性，且保留促进重组III型人源化胶原蛋白自交联形成三螺旋结构的功能的氨基酸序列。

SEQ ID NO.1: GERGSEGPSHPGQPGPPGPPGAPGPCCGG，其是人III型胶原蛋白肽段。

本公开意外发现，不同长度的胶原蛋白自组装元件促进重组III型人源化胶原蛋白自交联形成三螺旋结构的能力是不同的。
25

在一些具体的实施方案中，所述胶原蛋白自组装元件的氨基酸序列包含如SEQ ID NO.1、SEQ ID NO.2、SEQ ID NO.3或SEQ ID NO.4所示的序列。

在一些更具体的实施方案中，所述胶原蛋白自组装元件的氨基酸序列为如SEQ ID NO.1、SEQ ID NO.2、SEQ ID NO.3或SEQ ID NO.4所示的序列。

30 SEQ ID NO.2: GSEGPSHPGQPGPPGPPGAPGPCCGG。

SEQ ID NO.3: GSPGHPGQPGPPGPPGAPGPCCGG。

SEQ ID NO.4: GHPGQPGPPGPPGAPGPCCGG。

35 为了有效发挥所述胶原蛋白自组装元件促进重组III型人源化胶原蛋白自交联形成三螺旋结构的功能，在一些实施方案中，所述胶原蛋白自组装元件位于所述重组III型人源化胶原蛋白的末端，优选C末端。

本公开所述的重组III型人源化胶原蛋白是指由DNA重组技术制备的III型人胶原蛋白基因编码的全长或部分氨基酸序列片段，或是含III型人胶原蛋白功能片段的组合。

40 在一些具体的实施方案中，所述重组III型人源化胶原蛋白的氨基酸序列的末端含有胶原蛋白自组装元件，可以通过如下方式实现：设计末端含有胶原蛋白自组装元件的重组III型人源化胶原蛋白的氨基酸序列及其基因序列，并构建含有该基因序列的表达载体，将其转化进宿主细胞，后续可对该宿主细胞进行发酵培养，并对重组III型人源化胶原蛋白进行诱导表达即可。

重组III型人源化胶原蛋白

45 本公开提供了一种重组III型人源化胶原蛋白，其末端包含胶原蛋白自组装元件，所述胶原蛋白自组装元件如上述本公开提供的重组III型人源化胶原蛋白由于含有胶原蛋白自组装元件，其在表达后可自交联形成三螺旋结构，且该蛋白可以由生物合成的方法大量制备。

为了更好的形成三螺旋结构，所述胶原蛋白自组装元件位于所述重组Ⅲ型人源化胶原蛋白的C末端。

在一些具体的实施方案中，所述重组Ⅲ型人源化胶原蛋白包含如下（a）～（c）中任一项所示的序列：

- 5 （a）如SEQ ID NO.6、SEQ ID NO.7、SEQ ID NO.8或SEQ ID NO.9所示的氨基酸序列；
 （b）如（i）所述的氨基酸序列中取代、缺失或添加一个或多个氨基酸，且保留自交联形成三螺旋结构能力的氨基酸序列；
 （c）与（i）所述的氨基酸序列具有至少90%、92%、95%、96%、97%、98%或99%同一性，且保留自交联形成三螺旋结构能力的氨基酸序列。

- 10 在一些优选的实施方案中，所述重组Ⅲ型人源化胶原蛋白的氨基酸序列为如SEQ ID NO.6、SEQ ID NO.7、SEQ ID NO.8或SEQ ID NO.9所示的序列。

多核苷酸、重组表达载体、重组宿主细胞

本公开提供了编码上述胶原蛋白自组装元件及上述重组Ⅲ型人源化胶原蛋白的多核苷酸。

- 15 在一些实施方案中，编码所述重组Ⅲ型人源化胶原蛋白的多核苷酸包含如下（z₁）或（z₂）所示的序列：

 （z₁）如SEQ ID NO.11、SEQ ID NO.12、SEQ ID NO.13或SEQ ID NO.14所示的核苷酸序列；
 （z₂）在严格条件下与（z₁）所示的核苷酸序列杂交且编码具有自交联形成三螺旋结构能力的蛋白的核苷酸序列，所述严格条件是中等严格条件，中-高严格条件，高严格条件或非常高严格条件。

本公开提供了包含上述多核苷酸的重组表达载体。

- 20 本公开选择的表达载体可以在原核或真核细胞的各种宿主中稳定存在并可自主复制，如本领域中的常规质粒（pET系列）、穿梭载体PNV 18.1、噬菌体或病毒载体等。在载体上通过酶切、连接等分子生物学操作，将本公开上述多核苷酸的序列克隆到载体中，构建得到重组表达载体。

本公开提供了包含上述重组表达载体的重组宿主细胞。

- 25 本公开所述的宿主细胞选自原核细胞、酵母或真核细胞，进一步可以选自大肠杆菌、赤红球菌、枯草芽孢杆菌、酵母菌。将本公开上述的重组表达载体转化至宿主细胞中，获得相应的基因工程菌。

重组Ⅲ型人源化胶原蛋白的制备方法

本公开提供了上述重组Ⅲ型人源化胶原蛋白的制备方法，所述制备方法包括如下步骤：

- S1：构建包含编码所述重组Ⅲ型人源化胶原蛋白的多核苷酸的重组表达载体，并通过转化构建重组宿主细胞；
30 S2：在培养基中培养所述重组宿主细胞并生产蛋白；
 S3：收获并纯化所述蛋白，优选用Ni柱和/或阴离子交换层析纯化所述蛋白；和
 S4：任选地对所述蛋白进行酶切，优选用胶原工具酶酶切蛋白。

实施例

- 35 下面将结合实施例对本公开的实施方案进行详细描述，但是本领域技术人员将会理解，下列实施例仅用于说明本公开，而不应视为限定本公开的范围。实施例中未注明具体条件者，按照常规条件或制造商建议的条件进行。所用材料或仪器除非特别说明，均为可以使用通过市购获得的常规产品。

实施例1：重组Ⅲ型人源化胶原蛋白的构建与表达

1. 将人Ⅲ型胶原蛋白铰链区不同长度的氨基酸序列与重组Ⅲ型人源化胶原蛋白（TE16c，由本
40 申请申请人的在先专利ZL201811438582.6所记载）的氨基酸序列组合，得到含不同长度的铰链区的重组Ⅲ型人源化胶原蛋白。

（1）氨基酸序列

- 重组Ⅲ型人源化胶原蛋白TE16c的氨基酸序列：GERGAPGFRGPAGPNGIPGEKGPAGERGAPGE
RGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGF
45 RGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGP

NGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEK
GPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGER
GAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGER
GAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFR
5 GPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAP (SEQ ID NO.5)。

重组III型人源化胶原蛋白C3T16H1氨基酸序列: GERGAPGFRGPAGPNGIPGEKGPAGERGAPGE
RGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGF
RGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGP
NGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEK
10 GPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGER
GAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGER
GAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFR
GPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGSEGS₂PGHPGQPG
PPGPPGAPGPCCGG (SEQ ID NO.6)。

重组III型人源化胶原蛋白C3T16H2的氨基酸序列: GERGAPGFRGPAGPNGIPGEKGPAGERGAP
GERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAP
GFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPA
GPNIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPG
EKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAG
20 ERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPG
ERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPG
FRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGSEGS₂PGHPGQPGGP
PGPPGAPGPCCGG (SEQ ID NO.7)。

重组III型人源化胶原蛋白C3T16H3的氨基酸序列: GERGAPGFRGPAGPNGIPGEKGPAGERGAP
25 GERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAP
GFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPA
GPNIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPG
EKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAG
ERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPG
30 ERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPG
FRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGSPGHPGQPGPPGP
PGAPGPCCGG (SEQ ID NO.8)。

重组III型人源化胶原蛋白C3T16H4的氨基酸序列: GERGAPGFRGPAGPNGIPGEKGPAGERGAP
GERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAP
35 GFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPA
GPNIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPG
EKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAG
ERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPG
ERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGERGAPG
40 FRGPAGPNGIPGEKGPAGERGAPGERGAPGFRGPAGPNGIPGEKGPAGERGAPGGHPGQPGPPGPPGA
PGPCCGG (SEQ ID NO.9)。

在上述序列中,下划线部分为人III型胶原蛋白铰链区不同长度的氨基酸序列,也即不同长度的胶
原蛋白自组装元件。

(2) 核苷酸序列

45 重组III型人源化胶原蛋白 TE16c 的核苷酸序列: ggtgaacgtggtgcaccaggttttcgtggtccggcaggtccgaatgg

aattccgggtgagaaaggaccggctggtagcgtggtgcgccgggtgaacgtggagcgcctgggtttctgtggccagcaggtccgaacggtattcctggtg
aaaaaggtccggcgaggagcgtggtgcaccgggtgaacgcgggtgcaccgggattcgtggtccagcaggaccgaatggtatccctggtgaaaaaggac
cggcaggtgagcgtggagcgcacgtgaacgtggcgcaccgggtttctgtggaccggcagggccgaatggtattccgggtgaaaaaggcccgaggt
gaacgtggtgccccgggtgaacgtggtgcctggtttctgtggccggcagggaccgaacggtatccctggagaaaaaggctcgtcaggtgagcgcgggt
5 gcccggcgagcgtggtgccccgtttctgtggctccggcagggccctaatggtattcctggagaaaaaggccctgcaggtgaacggcagcaccgggt
gagcgtggcgcacctggtttctgtgctcgtcagggccgaacggtattccggcgaaaaagggtccagcaggtgaacgtggtgctccgggtgaacgtggtgc
acctggtattcgtggctcgtggtccgaatggtattccaggtgaaaaagggtccggcagggagagcgtggagcaccgggagaaacgtggtgcaccgggtttc
gtgtccggcggtcctaacgtatccaggtgaaaaagggtccggcgagcgtggcgccctggtgagcgtggtgctcctggtttctgtgtccggct
gttccgaacggaattcgtggtgagaaagggtccggctggcgaacgtggtgcaccgggtgaacgtggtgcaccgggtttctgtgtccggcggtcctaag
10 gtatccgggtgaaaaagggtccggcaggtgaacgtggtgcaccgggtgaacgtggtgcaccgggtttctgtggaccggcagggacctaatggtattccggg
agaaaaaggacctgcgggtgaacgtggtgcaccgggtgaacgtggtgcaccgggtttctgtgtccggcaggtcctaaggaattcgtgagagaaaggga
cctgcaggtgaacgtggtgcaccgggtgaacgtggtgcaccgggtttctgtgtccggcaggtcctaaggtattccgggtgaaaaagggtccggcaggtga
acgtggtgcaccgggtgaacgtggtgcaccgggtttctgtgtccggcaggtccgaatggcattcctggtgaaaaagggtccggcaggtgaacgtggtgcac
cgggtgaacgtggtgcaccgggtttctgtgtccggcaggtccgaatggtattccgggtgaaaaagggtccggcaggtgaacgtggtgcaccg (SEQ I
15 D NO.10)。

重组III型人源化胶原蛋白 C3T16H1 核苷酸序列: ggagaaagggggcgccaggtttcgtggtccggcgccggccga
acggcatccgggtgaaaaaggaccggcggtgaacgtggtgcgccgggagaacgtggagcgcgggctttctgtgtccggcggtcccaacggcatt
ccgggtgagaagggcccgccgggtgaacgtggtgctccggcgagcgtggtgccccgggcttctgtggcccccggtccgaacggcatcccggtga
gaaagggtccggtggcgagcgtggcgacatggcgagcgcggcgccaggtttctgtgtccggcggtcctaaggtattccggcgaaaaagggtcc
20 ggacggcgaaacgtggtgctccgggtgagcgtggtgctccgggtttctgtgtccggcggtcccaacggcatccggcgagaaagggtccggcggtga
gcgggtgcaccaggtgagcgcggcgccctggttttagaggtccggcggtccgaacggtattccgggtgagaagggtccggcagggcgaacggcgag
cgccggcggaacgtggggccccgggtattcgtggaccggcaggtccgaacggcattccggcgaaaaaggccccggcggtgagcgcggcgccctg
gcgaacgtggcgctccgggtttctgtggccccggcgccgaatggtatccgggtgaaaaaggccccagcggtgaacgcgtgaccggcggaacgt
ggtgcccccggtattccggtcctggtggtccgaacggcattccggcgaaaaagggtccgctggtgaacgagcgcaccggcgagcgcggggtcc
25 gggcttctgtgaccagcggcgccgaatggtattcctggcgaaaaagggtccggcaggtgagcggggcgaccgggtgaacgtggcgaccgggttcc
gcggaccggcgggccgaatggtatccgggtgaaaaagggtccggcggtgagcgtggtgcgggggtgagcgtggcgaccaggtttctgtgtccg
gcagggccgaatggtatccgggtgagaagggccagcgggtgagcgtggtgcgggggtgaacgcggggccctggttccgtggtccggcggtcc
gaacggtatccgggtgaaaaaggccccagcagcgagcgcgggtgctccgggtgagcgtggagcgtccgggtttctgtggccccggcggtccgaacggca
ttccgggtgaaaaagggtccggcggtgagcgtggtgcccaggtgagcgtggtgctccgggtttctgtgtccggcggtccgaacggcatcccggtg
30 agaaagggtccggcagggcgaacgtggcgccagcggaacgtggggcgccaggtttccgggtccagcaggtccgaatggcatccgggtgagaagg
ccggcggtgagcgcgggtgacccgggtgagcgcggtagcgaaggttccccgggtcaccgggtcagcggcgccgggtccggcgccggcgcccc
gggtccgtgttcggtgt (SEQ ID NO.11)。

重组III型人源化胶原蛋白 C3T16H2 的核苷酸序列: ggagaaagggggcgccctggctttctgtggtccggctggtccg
aacggtattccgggtgaaaaagggtccggcgggcgagcgcgggtgcaccgggtgagcgtggtgccccgtgtttctgtggccagctggccgaacggcatc
35 ccgggtgagaagggccagcggcgagcgtggtgccccgggtgagcgtggagcgcacaggtattccgtggccccgggtgccccgaacggcatcccggtg
aaaaaggccccggcgccgaaggggcgccagggagaagaggtgcggcggtccggctggcggtccggcggaacggcatcccggtgaaaaag
gtccggcaggtgaaaggcgccggcgagcgtggtgcaccgggttctgtgtccggtgccccgaacggcatcccggtgagaagggccccggc
gggtgaacgtggcgccaggtgaacgtggtgcccgttccgtgtccggcggtccgaatggcatcccggggagaaggtccggcaggtgagc
gcgggtgcaccgggtgagcgtggtgcaccgggtattccggcgccctgcgggccgaacggcatcccggtgaaaaagggtccggcaggtgaacgcgtgc
40 ccagggcgaacgtggcgccctggtttctgtgtcctgcccgtcctaaggtattccggcgaaaaagggtccggcggtgagcgtggcgaccgggggaacgtggcg
acgtggtgccccgggtttctggaccggcggttcgaatggtatccggcgaaaaagggtccggcggtgagcgtggcgaccgggggaacgtggcg
cgccaggttccggcgcccgagggccgaacggcatcccgggcgagaagggtccggcggtgagcgtggcgctccgggtgagcgcgggtgccccgg
gatttccgggtccggcggtccgaatggcatcccgggcgagaagggtccggcggtgagcgcgggtgctccgggtgaacgtggtgccccgggttccgt
ggtcctccgggtccgaacgtattcctggtgagaagggccagcggtgaacgcgggtgcaccgggtgaaagaggtgcaccaggtttctgtggccccggcg
45 ggtcctaaggtattccggcgaaaaaggaccggcggtgagcgtggtgcggcgagcgcgggtgcaccgggttccgtggccccggcggtccgaa

tggcattccgggtgaaaagggggcggctggcgagcgtggtgctccgggtgaacgtggcgaccggggttcgcggtccgcgggccccaacggcatcc
cgggtgagaagggccggcggtgagcgcggcgccaggcgaacgtggagcgccggattccgcggccggcccccacgggtattccgggtg
agaaggggtccggctggagagcggggcgccgggtagcgaaggttccccggggccaccgggtcagccggccctccagggtccgctggtgctgctgg
cccttgttcgggtgt (SEQ ID NO.12)。

- 5 重组III型人源化胶原蛋白 C3T16H3 的核苷酸序列: ggagaaagggggcgccagggttcgcggtccggcggttc
gaacgggtattccggcgagaaggggtccggcggtgagcgtggagcgccgggtgagagggcgccggggttcggtgcccagcgggaccgaacgg
tattccagggtgaaaaggtccggctggcgagcgtggtgctgcccgggtgaacgtggggcaccggggttcggtgcccggcgggccccaacgggtattccgggt
gaaaagggccggcggtgagcgtggtgctggtgagcgcggagcgccgggttccggtgcccagcgggaccgaatggcatccgggtgaaaag
gtccggcgggcgagcgtggtgctccggcggaacgtggcgaccgggtttctggtccctgcggggtccgaatggcatccgggtgaaaagggccagca
10 ggtgaacgtggtgctccgggtgagcgggtgctgcccgggttcggtgcccgggttcctaatggcatccgggtgagaaggggtccagccggtgaacgt
ggcgaccggggcgagcgggtgccccgggttctggtggtggtgcccgaatggtattccgggtgaaaaggtccgggtggtgacgtggcgacc
gggtgagcgtggtgctccgggttccgggtccggcgggaccgaatggtattccgggtgagaaggtccggcggtgaacgtggtgccccggcgaga
gaggtgccccgtggttccgggtccggcaggtccgaacgggtatccgggtgagaaggtccgggtgagcgtggcgccaggtgagcgtggtgctgc
ctggtttccggcgcccgaggtcctaacggtattctggcgaaaaggtctcggtgagcgtggagccctggcggaacgtggtgctgcccaggttccgt
15 ggtccgggtgccccaacggcatccggcgagaaaggtccggcggtgagcgtggcgctccggggagcgcgccgccccgggttccaggggtccg
gcgggtccgaacggcatccgggtgagaagggccggcgggcgagcgggtgccccagcgagcgcggtccgggtttctggtgctgcccaggtcc
ccaacggcatccggcgaaaaggtccgggtgctggaacgcgggtgctccgggtgaacgtggtgacccgggttccgtggtgcccagcgggcggaatggc
attccggcgaaaaggtccggcggtgagcgggtgacccgggggagcgtggtgacgttctggtgctccgggtcccaacggcatcccggtg
cgagaaggtccggcggtggaacgcgtgccccggcggaacgtggtgccccaggtttctggtgcccggcggtccgaacggcattccggcgaaaag
20 ggccctgcggtggaacggggcgaccggcgagccccggggcaccgggtcagccggggccaccgggtccgctgggtgctccgggtccgtgttgcggtg
gt (SEQ ID NO.13)。

- 重组III型人源化胶原蛋白 C3T16H4 的核苷酸序列: ggagaaagggggcgccgggttccggtgctgctggtccg
aacggcatcccggtgaaaagggccggcgggggagcgtggagccccgggagaacgtggtgcccaggttccgtgcccagcggttccgaacggtat
ttccgggggagaaaagggccggcggtgacccagcggaagggcgccgaggttccagagggccgggtccgaatggcatcccg
25 gcgagaaggtccggcggtggaacgtggtgctgcccgggtgaacgcgtgctccgggtttctggtgctgcccgaacggcatcccggtgagaaa
ggccccggcggtgaacgtggaagcaggtgagcggcgccgggttccgtgcccagcgggccgaacgggtattccgggtgaaaaggttccgg
cgggtgaacgtggtgctgcccgggtgaacgtggtgctgctggtttagggcccgagctgcccgaatggtattccgggtgagaaggtccagctggagagc
gaggtgccccggcggaacgtggtgccccgttctggtgctgcccgggtccgaatggcatcccgggcgaaaaggtccgggtggtgagcgtggggcac
cggcggaacgggtgctgcccgggttccgtgcccggcagggccccaacggtatccgggtgagaaggggtccggcggtgagcgggtgacccggcg
30 aacgggtgctccgggtttctgcccggcggttccgaatggtattccggcgagaagggcccgagcgagcgtggtgctccaggtgaacgcgtg
ccccgggttctggtgctgcccgggttcccaacggtatccgggtgagaaggggtccagccggagagcgtggtgacccagagagcgggtgctgcccggc
ttccgggtgctgcccagggccgaatggcatcccggtgagaaaggtccggcgagcgtggtgcccagagagcgtggtgctgctggttccgtg
gcccaggttcccaacggcattccgggagaaaaggtccggcggtggaacgtggtgcccgggtgagcgtggtgccccgggttccgtggtgccc
ggtccgaacggcatcccggtgagaaaggtctgctggtgagcgtggtgcccgggtgaacgtggtgacccgggttcccggtgcccagggccccaat
35 ggtattccgggtgagaaggggtccagcggtgaacgtggtgctgcccggcggaacgcgtgccccgggtttctggtgctggtgcccgaacggcattccgg
gcgaaaagggccggcggtgaacgcgtgcccgggtgaacgcgtgctggtgctgctggttccgaacgggtattccggcgaaaag
ggccccggtggtgagcgtggtgctgcccgggacacccgggtcagccgggtcctccgggtccgctggtgctgcccggcgccctgttgcggtgt (SEQ I
D NO.14)。

2. 将重组III型人源化胶原蛋白C3T16H1、C3T16H2、C3T16H3、C3T16H4的基因序列进行合成,
40 将上述每种编码核苷酸序列 (在5'端添加胶原工具酶酶切位点, 胶原工具酶酶切位点的氨基酸序列
为ENLYFQ (SEQ ID NO.16), 核苷酸序列为gaaacctgtattccag (SEQ ID NO.17)) 插入pET-28a-Trx-His
表达载体的KpnI和XhoI酶切位点之间。TE16c重组表达质粒为现有质粒, 由本申请申请人的在先专利
ZL201811438582.6所记载。

3. 将构建成功的表达质粒转化大肠杆菌感受态细胞BL21(DE3)。具体过程为: (1)在超低温冰箱
45 中取出大肠杆菌感受态细胞BL21(DE3)置于冰上, 待半融时取2μl待转化的质粒加入大肠杆菌感受态细

胞BL21(DE3)中，稍微混匀2-3次。(2)将混合物置于冰上冰浴30min，然后于42℃水浴热激45-90s，取出后置于冰上冰浴2min。(3)转移至生物安全柜中，并加入700μl液体LB培养基，然后于37℃、220rpm条件下培养60min。(4)取200μl的菌液均匀涂布在含有氨苄西林钠的LB平板上。(5)将平板在37℃的培养箱中培养15-17h，待其长出大小均匀的菌落。

5 4. 从转化好的LB平板中挑取5-6个单菌落于含有抗生素储液（转化C3T16H1、C3T16H2、C3T16H3、C3T16H4重组表达质粒的单菌落对应添加至含硫酸卡那霉素的抗生素储液，转化TE16c重组表达质粒的单菌落对应添加至含氨苄西林钠的抗生素储液）的摇瓶中，在220 rpm，37℃恒温摇床中7h。再将培养后的摇瓶降温至16℃，添加IPTG诱导表达一段时间后，将菌液分装于离心瓶中，于8000 rpm、4℃离心10 min，收集菌体，并记录菌体重量，取样进行电泳检测。

10 5. 将收集的菌体用平衡工作液重悬（200mM氯化钠、25mM Tris、20mM咪唑，pH8.0），将菌液降温至≤15℃，进行均质，高压均质两次（取均质后的样品标注：均质），完成后收集菌液。将均质后的菌液分装至离心瓶中，于17000 rpm、4℃离心30 min，收集上清液，取上清（标注：上清）和沉淀（标注：沉淀）进行电泳检测。

6. 将收集的上清液进行纯化和酶切，具体过程是：(1)粗纯：a.水洗柱材（Ni6FF，Cytiva），5个CV。b.平衡液（200mM氯化钠、25mM Tris、20mM咪唑，pH8.0）平衡柱材，5个CV。c.上样：将收集的上清液加入柱材中，直至液体流完后，取流穿进行电泳检验（标注：流穿）。d.清洗杂蛋白：添加洗杂液25 mL(200 mM氯化钠，25 mM Tris，20mM咪唑，pH8.0)至液体流完，并取洗杂流穿进行电泳检验（标注：洗杂）。e.收集目的蛋白：添加20 mL洗脱液（200mM氯化钠、25mM Tris、250mM咪唑，pH8.0），并收集流穿液（标注：洗脱），采用紫外可见分光光度法检测蛋白浓度，并进行电泳检测。f.用1M咪唑工作液清洗柱材（标注：1M洗）。g.纯化水清洗柱材。(2)酶切：按蛋白质量与胶原工具酶质量比为20：1，添加胶原工具酶，16℃酶切4 h，取样（标注：酶切）进行电泳检测。将酶切后的蛋白液放入透析袋，透析液为A液（20 mM Tris，20 mM氯化钠，pH8.0），于4℃透析2 h，再转移至新的透析液中4℃过夜透析，取样（标注：换液）进行电泳检测。(3)精纯：a.平衡柱材（Capto Q，Cytiva）：使用A液（20 mM Tris，20 mM氯化钠，pH8.0）将柱材平衡，流速为10 ml/min。b.上样：流速为5 ml/min，上样并收集流穿（标注：Q流穿），并进行电泳检测。c.二元流动相体系梯度洗脱：分别设置0%（V/V）-15%（V/V） B液（20 mM Tris，1M氯化钠，pH8.0）和100%（V/V）-85%（V/V） A液2 min然后保持3个CV、15%（V/V）-30%（V/V） B液和85%（V/V）-70%（V/V） A液2 min然后保持3个CV、30%（V/V）-50%（V/V） B液和70%（V/V）-50%（V/V） A液2 min然后保持3个CV、50%（V/V）-100%（V/V） B液和50%（V/V）-0%（V/V） A液2 min然后保持3个CV，出峰收集（标注：B洗）并进行电泳检测。d.清洗柱材。e.检测目的蛋白含量并计算蛋白收率，将蛋白储存在4℃环境中。

7. 浓度检测

(1) 粗纯蛋白浓度检测

35 精确量取适量样品，用洗脱液稀释10-50倍，用玻璃棒充分搅拌均匀。使用紫外可见分光光度计在280nm处测定吸光度，根据公式 $C(\text{mg/ml}) = A_{280} \times \text{吸光系数} \times \text{稀释倍数}$ ，计算蛋白浓度（注：吸光度值需在0.1~1）。

浓度检测结果如下：

质粒	TE16c	C3T16H1	C3T16H2	C3T16H3	C3T16H4
吸光系数	3.48	4.18	4.15	4.13	4.13
稀释倍数	1	1	1	1	1
A280	0.624	0.720	0.612	0.932	0.766
浓度	2.17mg/ml	3.01mg/ml	2.54mg/ml	3.85mg/ml	3.16mg/ml
洗脱蛋白体积	20ml	20ml	20ml	20ml	20ml

蛋白量	43.40mg	60.20mg	50.80mg	76.98mg	63.27mg
-----	---------	---------	---------	---------	---------

蛋白表达量C3T16H3>C3T16H4>C3T16H1>C3T16H2>TE16c。

(2) 精纯蛋白浓度检测

精确量取适量样品，用洗脱液稀释10-50倍，用玻璃棒充分搅拌均匀。使用紫外可见分光光度计在215和225nm处测定吸光度，根据公式C (mg/ml) = (A215-A225) × 144 × 稀释倍数 ÷ 1000，计算蛋白浓度（注：吸光度值需在0.1~1）。

浓度检测结果如下：

质粒	TE16c	C3T16H1	C3T16H2	C3T16H3	C3T16H4
稀释倍数	10	10	10	10	10
A215	0.301	0.505	0.513	0.584	0.717
A225	0.121	0.167	0.156	0.252	0.322
浓度	0.26mg/ml	0.49mg/ml	0.51mg/ml	0.48mg/ml	0.57mg/ml
收集蛋白体积	51ml	53ml	48ml	51ml	60ml
蛋白量	13.26mg	25.80mg	24.68mg	24.38mg	34.13mg

精纯后蛋白量C3T16H4>C3T16H1>C3T16H2>C3T16H3>TE16c。

8. 电泳检测

具体过程为：取样品液40μl，加入10μl 5×的蛋白上样缓冲液(250mM的Tris-HCl(pH 6.8), 10% (V/V) SDS, 0.5% (V/V) 溴酚蓝, 50% (V/V) 甘油, 5% (V/V) β-巯基乙醇)，置于100℃沸水中煮10min，然后每孔10μl加入SDS-PAGE蛋白胶中，电压80V电泳2h后，用考马斯亮蓝染色液(0.1% (V/V) 考马斯亮蓝R-250, 25% (V/V) 异丙醇, 10% (V/V) 冰醋酸)进行蛋白染色20min，再利用蛋白脱色液(10% (V/V) 醋酸, 5% (V/V) 乙醇)进行脱色。

电泳检测结果如图1~5所示，C3T16H1、C3T16H2、C3T16H3、C3T16H4、TE16c蛋白表达量均较好，纯化后胶原蛋白纯度较好，无明显杂带。

实施例2：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

频带宽度 (Band width) : 1.0nm;

步进 (Step) : 1.0nm;

测量范围 (Measurement range) : 190-260nm (远紫外区域扫描 (far-UV region scan)) / 250-340nm (近紫外区域扫描 (near-UV region scan)) ;

每点时间 (Time-per-point) : 0.5s;

重复 (Repeats) : 3次;

比色皿长度 (Cell Length) : 10mm|0.5mm;

温度 (Temperature) : 室温 (Room Temperature) 。

(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 (buffer) 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

取精纯后的C3T16H1、C3T16H2、C3T16H3、C3T16H4、TE16c蛋白样品，用10kDa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。

(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

- 10 实验结果如图6所示，C3T16H1、C3T16H2、C3T16H3、C3T16H4、TE16c在221nm处均有正峰，且正峰值C3T16H2>C3T16H4≥C3T16H3>C3T16H1≥TE16c。表明C3T16H1、C3T16H2、C3T16H3、C3T16H4、TE16c在相同蛋白浓度条件下均具有三螺旋结构，人III型胶原蛋白铰链区的氨基酸序列有助于重组III型人源化胶原蛋白三螺旋结构的形成，不同长度的铰链区的氨基酸序列促进重组III型人源化胶原蛋白形成三螺旋结构的能力有所不同。

权 利 要 求 书

- 1、一种胶原蛋白自组装元件，其特征在于，所述胶原蛋白自组装元件包含如下(i)~(iii)中任一项所示的序列：
- (i) 如SEQ ID NO.1所示的氨基酸序列的C端的至少21个连续的氨基酸；
- (ii) 如(i)所述的氨基酸序列中取代、缺失或添加一个或多个氨基酸，且保留促进重组III型人源化胶原蛋白自交联形成三螺旋结构的功能的氨基酸序列；
- (iii) 与(i)所述的氨基酸序列具有至少90%、92%、95%、96%、97%、98%或99%同一性，且保留促进重组III型人源化胶原蛋白自交联形成三螺旋结构的功能的氨基酸序列。
- 2、根据权利要求1所述的胶原蛋白自组装元件，其特征在于，所述(i)所示的序列为如SEQ ID NO.1、SEQ ID NO.2、SEQ ID NO.3或SEQ ID NO.4所示的氨基酸序列。
- 3、根据权利要求1或2所述的胶原蛋白自组装元件在用于促进重组III型人源化胶原蛋白自交联形成三螺旋结构中的用途。
- 4、根据权利要求3所述的用途，其特征在于，所述胶原蛋白自组装元件通过存在于所述重组III型人源化胶原蛋白的氨基酸序列的末端发挥作用；
- 优选地，所述重组III型人源化胶原蛋白的氨基酸序列的末端为C末端。
- 5、一种重组III型人源化胶原蛋白，其特征在于，所述重组III型人源化胶原蛋白的末端包含根据权利要求1或2所述的胶原蛋白自组装元件；
- 优选地，所述重组III型人源化胶原蛋白的末端为C末端。
- 6、根据权利要求5所述的重组III型人源化胶原蛋白，其特征在于，所述重组III型人源化胶原蛋白包含如下(a)~(c)中任一项所示的序列：
- (a) 如SEQ ID NO.6、SEQ ID NO.7、SEQ ID NO.8或SEQ ID NO.9所示的氨基酸序列；
- (b) 如(i)所述的氨基酸序列中取代、缺失或添加一个或多个氨基酸，且保留自交联形成三螺旋结构能力的氨基酸序列；
- (c) 与(i)所述的氨基酸序列具有至少90%、92%、95%、96%、97%、98%或99%同一性，且保留自交联形成三螺旋结构能力的氨基酸序列。
- 7、多核苷酸，其特征在于，所述多核苷酸编码根据权利要求1或2所述的胶原蛋白自组装元件，或编码根据权利要求5或6所述的重组III型人源化胶原蛋白；
- 优选地，所述多核苷酸包含如下(z₁)或(z₂)所示的序列：
- (z₁) 如SEQ ID NO.11、SEQ ID NO.12、SEQ ID NO.13或SEQ ID NO.14所示的核苷酸序列；
- (z₂) 在严格条件下与(z₁)所示的核苷酸序列杂交且编码具有自交联形成三螺旋结构能力的蛋白的核苷酸序列，所述严格条件是中等严格条件，中-高严格条件，高严格条件或非常高严格条件。
- 8、重组表达载体，其特征在于，所述重组表达载体包含根据权利要求7所述的多核苷酸。
- 9、重组宿主细胞，其特征在于，所述重组宿主细胞包含根据权利要求8所述的重组表达载体。
- 10、根据权利要求5或6所述的重组III型人源化胶原蛋白的制备方法，其特征在于，所述制备方法包括如下步骤：
- S1：构建包含根据权利要求7所述的编码根据权利要求5或6所述的重组III型人源化胶原蛋白的多核苷酸的重组表达载体，并通过转化构建重组宿主细胞；
- S2：在培养基中培养S1获得的所述重组宿主细胞并生产蛋白；
- S3：收获并纯化所述蛋白，优选Ni柱和/或阴离子交换层析纯化所述蛋白；和
- S4：任选地对所述蛋白进行酶切，优选TEV蛋白酶酶切蛋白。