

一种重组人 XVII 型胶原蛋白、制备方法和应用

技术领域

本发明涉及一种重组人 XVII 型胶原蛋白、制备方法和应用，属于基因工程技术领域。

背景技术

胶原蛋白作为重要的天然生物蛋白，可广泛应用于化工、医药、食品、化妆品等众多领域，尤其适合制备多种生物器械，是最为理想的生物材料来源，具有广阔的应用前景。市场上销售胶原蛋白主要是利用酸、碱、酶解法处理动物组织获得的胶原提取物。这类提取物在加工过程中，基本破坏了胶原的天然结构，降解严重，使其生物活性丧失；提取的胶原肽长度不等、性质不均、质量不稳定且有病毒感染的风险；同时动物源与人的胶原蛋白的氨基酸序列差别较大，会导致免疫排斥和过敏症状。

人 XVII 型胶原蛋白是一种跨膜型非成纤维胶原蛋白，其为三条相同 $\alpha 1$ (XVII) 链构成的均相三聚体，分为胞内、跨膜、胞外三大类结构域。人 XVII 型胶原蛋白是细胞中半桥粒的一种组成成分，在上皮细胞-基底膜的作用中发挥重要作用，可调控上皮细胞的黏附、分离和发育分化，对角化细胞的分化和再生有重要作用。人 XVII 型胶原蛋白在人体中含量极少，且 XVII 型胶原蛋白在动物体内含量同样极为稀少，提取难度也非常大，无法量产，同时动物源性胶原蛋白产品不可避免的有免疫原性和潜在的病毒、疫病等生物安全性隐患，没有大量应用。限于法理伦理的约束，人源 XVII 型胶原蛋白只能应用于科研层面，这些也就导致现在市场上没有人或其它动物来源的商品化 XVII 型胶原蛋白产品，这也进一步限制了人 XVII 型胶原蛋白的研究和应用。现在解决此类问题的主要方式则是通过基因工程等生物技术来完成胶原蛋白。

现有重组蛋白主要由原核（大肠杆菌）表达系统，毕赤酵母表达系统、哺乳动物细胞表达系统、昆虫细胞（杆状病毒）表达系统四大类表达系统进行生产。关于哺乳动物细胞表达人 XVII 型胶原蛋白的研究较少，且均处于实验研究阶段；最早使用原核表达（大肠杆菌、pGEX 表达载体）的是非螺旋区 NC16 序列的融合蛋白，其余表达系统暂无成功表达人 XVII 型胶原蛋白的报道，尤其是毕赤酵母表达系统。

其中哺乳动物细胞表达成本高、产量低，多用于高价值的药用蛋白表达生产。哺乳动物细胞表达的人 XVII 型胶原蛋白，无论是瞬时转染表达还是稳定转染细胞系表达，蛋白质表达量极低，同时其所使用的培养基昂贵，只能满足科学研究中微量的使用。昆虫细胞（杆状病毒）表达系统除成本高、成量低外，其翻译后与人细胞差异巨大，胶原蛋白的大量生产应用一般不采用这两种方式。原核表达（大肠杆菌表达系统）的胶原蛋白没有对蛋白质的翻译后修饰，有大规模放大前景的原核表达只能是胞内表达，需要进行菌体的裂解，会有大量杂质

蛋白与目的蛋白混合，对纯化工艺都有极高的要求，且原核表达体系天然带有（细菌细胞壁成分）内毒素、肽聚糖，需要复杂的纯化工艺方能去除。

以上的几个表达系统中，均存在序列的单一性与生物学活性的问题，序列选取比较单一、胶原蛋白肽均是分别独立进行表达等，相对于人 XVII 型胶原蛋白胞外域接近 1000 个氨基酸（分 15 个三螺旋区、15 个非三螺旋区）的序列丰富性而言，这两者是欠缺的，这也直接影响了生物学活性，胶原的生物学活性中细胞粘附活性只是最基础的一项，其余的生物学活性均未涉及。并且，无论是以上的哺乳动物细胞表达，还是大肠杆菌表达，均只停留于小体积实验室级别的培养、生产，均没有进行规模化生产。但重组胶原蛋白的应用前提是能够进行规模化、高密度、高表达的发酵生产和纯化，一般来经过说 500L 体积的中试发酵实验的验证方有工业级别、规模化放大生产的可能性。

人 XVII 型胶原蛋白表达存在一些难点，人 XVII 型胶原蛋白属于跨膜类的胶原蛋白，有胞内区、跨膜区、胞外区。一般来说跨膜蛋白于真核细胞中表达时多数时候并不会分泌于胞外，而是会固定于细胞膜上，而且人 XVII 型胶原蛋白的氨基酸序列很长（1497 个氨基酸）、蛋白质分子量较大（180kDa），理论上很难有效分泌于胞外且易于降解，要成功表达就需要进行相关序列的选取。同时，现有人 XVII 型胶原蛋白的研究对其氨基酸序列、结构、功能等方面的研究知之甚少。所以，如何选取氨基酸序列使其能在保持毕赤酵母表达体系优点（尤其是分泌表达）的前提下实现更多生物学功能是突出的难点，当然，这一点在其它表达系统中（包括大肠杆菌）表达人 XVII 型胶原蛋白也是难点。

发明内容

本发明的目的在于，克服现有技术中存在的一些技术问题，主要是克服关于重组人 XVII 型胶原蛋白序列（非单一性）的优化选取及能否在毕赤酵母高效分泌表达、现有重组人 XVII 型胶原蛋白仅有细胞黏附活性、现有重组人 XVII 型胶原蛋白未能规模化制备的生产的技术问题。

为此，本发明提供一种胶原蛋白、编码该胶原蛋白的多核苷酸、含有编码该胶原蛋白的多核苷酸的重组载体、工程菌、制备该胶原蛋白的方法、包含该胶原蛋白的组合物、该胶原蛋白的应用。本发明中已经证实该胶原蛋白不但能够实现人 XVII 型胶原蛋白于毕赤酵母这种真核宿主细胞的高效分泌可溶性表达，而且还具有比商品化的天然人胶原蛋白更为优异的细胞粘附活性、促细胞迁移活性以及优异组织再生和促进毛囊修复再生的生物学活性，并且可工业化应用。

本发明提供了一种胶原蛋白，所述胶原蛋白具有细胞粘附活性、促细胞迁移活性、促进组织再生和促进毛囊修复再生的生物学活性。

所述胶原蛋白，由 $(A)_n$ 所示的氨基酸序列组成或包含 $(A)_n$ 所示的氨基酸序列，其中：A为SEQ ID NO.2所示的氨基酸序列或在SEQ ID NO.2基础上进行一定程度上的氨基酸替换、插入、取代、添加、缺失等修改后的氨基酸序列或与SEQ ID NO.2氨基酸序列具有大于80%同源性的氨基酸序列；其中n为大于或等于1的整数；其中每个A相同或不同并且以A为基础单元串联重复，每个相同或不同的A之间直接通过肽键相连接。

进一步的，根据本发明的一个实施例，所述胶原蛋白由 $(A)_n$ 所示的氨基酸序列组成或包含 $(A)_n$ 所示的氨基酸序列：其中A为SEQ ID NO.2所示的氨基酸序列；n为大于或等于1的整数；其中每个A相同且以A为基础单元串联重复，每个A之间直接通过肽键相连接。

进一步的，根据本发明的一个实施例，所述胶原蛋白由 $(A)_n$ 所示的氨基酸序列组成，当n为1时，所述胶原蛋白的氨基酸序列如SEQ ID NO.2所示，本发明中记为170801；当n为2时，所述胶原蛋白的氨基酸序列如SEQ ID NO.3所示，本发明记为170802。

进一步的，根据本发明的记载，本发明所述胶原蛋白由 $(A)_n$ 所示的氨基酸序列组成或包含 $(A)_n$ 所示的氨基酸序列，其中A为在SEQ ID NO.2基础上进行一定程度上的氨基酸插入、取代、添加、缺失等修改后的氨基酸序列或与SEQ ID NO.2氨基酸序列具有大于80%同源性的氨基酸序列；其中n为大于或等于1的整数；其中每个A相同或不同并且以A为基础单元串联重复，每个相同或不同的A之间直接通过肽键相连接。

本发明还提供编码上述胶原蛋白的多核苷酸，具体为编码上述 $(A)_n$ 代表的胶原蛋白的多核苷酸序列；优选的，以毕赤酵母为宿主细胞时，所述多核苷酸为包含SEQ ID NO.6或SEQ ID NO.7所示的核苷酸序列或由SEQ ID NO.6、SEQ ID NO.7所示的核苷酸序列组成。

根据本发明的一个实施例，所述A为SEQ ID NO.2，n为1时的多核苷酸为由SEQ ID NO.6所示的核苷酸序列；根据本发明的一个实施例，所述A为SEQ ID NO.2，当n为2时，编码上述 $(A)_n$ 的多核苷酸为由SEQ ID NO.7所示的核苷酸序列。

本发明的胶原蛋白，是从人XVII型胶原蛋白的第15螺旋区、羧基端区、中间区域等多个螺旋区序列中优化筛选后选取的组合序列，与人XVII型胶原蛋白氨基酸序列相应部分100%相同，不但能够实现人XVII型胶原蛋白于毕赤酵母这种真核宿主细胞的高效分泌可溶性表达，而且还具有比商品化天然人胶原蛋白更为优异的细胞粘附活性、促细胞迁移活性以及促进组织再生和促进毛囊修复再生的生物学活性；这样的组合序使各区域序列的功能得以集成，并不是分别单独表达而是作为一个整体进行表达，避免了重组胶原蛋白的序列单一性。

本发明还提供了含有编码上述胶原蛋白的多核苷酸的重组表达载体。

根据本发明的一个实施例，本发明所述的重组表达载体构建过程中，首先对编码上述 $(A)_n$ 代表的胶原蛋白的多核苷酸进行修饰，分别于胶原蛋白氨基酸序列的氨基端添加编码

Strep-Tag II 标签的 DNA 序列、羧基端添加编码 6×His Tag 标签的 DNA 序列，使其含有双特异性亲和纯化标记，这样可进行亲和层析纯化，也便于以两种标签序列为基础进行免疫学抗体检测。将标记后的序列连接入表达载体 pPIC9K 中，构建重组表达载体（实施例中表示为 pPIC9K-170801、pPIC9K-170802）。

本发明还提供由上述重组表达载体构建的工程菌，所述工程菌的宿主细胞优选为毕赤酵母。所述工程菌保藏日期为 2020 年 09 月 09 日，保藏编号分别为 CGMCC NO.20626、CGMCC NO. 20627，分类命名为巴斯德毕赤酵母 *Pichia pastoris*，保藏单位为中国微生物菌种保藏管理委员会普通微生物中心，地址为北京市朝阳区北辰西路 1 号院 3 号。

需要注意的是，宿主细胞可以是真核细胞，例如真菌、酵母、植物细胞等；原核细胞，例如肠杆菌科细菌，如大肠杆菌；也可以是动物细胞，例如哺乳动物细胞 CHO 系列细胞、HEK293 系列细胞或昆虫细胞。应当理解，本领域技术人员可以通过用其它表达宿主代替上述毕赤酵母作为表达宿主表达出相同氨基酸序列的胶原蛋白。

本发明还提供上述胶原蛋白的制备方法，包括如下步骤：

（1）工程菌的培养和筛选

将构建的工程菌经过筛选后，接入 BMGY 培养基诱导表达筛选表达高的工程菌。所述筛选的表达量高的工程菌为巴斯德毕赤酵母（*Pichia pastoris*），保藏编号为 CGMCC No.20626 或 CGMCC No. 20627。

（2）规模化高密度发酵、培养和蛋白质纯化

采用 50L-500L 发酵罐进行联动分批补料发酵，对发酵上清液进行纯化，获取高纯度胶原蛋白。

根据本发明的一个实施例，本发明以 Sac I 线性化重组表达载体，电转入毕赤酵母感受态细胞，构建工程菌，将工程菌转涂至 MD 平板初筛后，再经过含有不同浓度 G418 的 YPD 平板筛选，挑取菌落接入 BMGY 培养基中，再以 BMMY 培养基诱导表达；挑取多株菌种，选择表达量高的工程菌进行后续实验。本发明中，以 SDS-PAGE 电泳、Western Blot 对表达的蛋白质进行初步鉴定，判断其有效分泌表达的能力，结果证明了本发明的重组人 XVII 型胶原蛋白能够高效分泌表达，而全长的人 XVII 型胶原蛋白主要集中于胞内无法有效分泌于胞外；将电泳获得的目标条带进行 Nano-HPLC-MS/MS 质谱检测，证明其中肽段来源于人 XVII 型胶原蛋白序列中优化选取的区域，本发明的重组人 XVII 型胶原蛋白 170801、170802 成功表达。

采用 50L-500L 发酵罐诱导表达 48h 时，本发明的胶原蛋白产量（UV 定量）可达到 12g/L、11g/L。

本发明中，使用 MALDI-TOF/TOF 进行分子量的测定，170801 胶原蛋白的分子量为 23811.6797Da（理论值为 23800Da）、170802 胶原蛋白的分子量为 45689.5781Da（理论值为 45400Da），毕赤酵母表达后存在翻译后修饰及分子量实验测定存在的系统误差，实际表达的胶原蛋白分子量与理论值一致。使用对胶原蛋白红外光谱扫描，其酰胺 A、酰胺 B、酰胺 I、酰胺 II、酰胺 III 的波数均符合重组胶原蛋白结构特征。电镜扫描纯化的胶原蛋白冻干海绵，有片层状结构，有应用于生物医学材料领域的潜力。

根据本发明的一个实施例，对得到的胶原蛋白的生物学活性检测，体外细胞实验验证其细胞粘附活性、促细胞迁移活性。对得到的胶原蛋白生物学活性的动物实验，兔耳瘢痕模型实验验证其能够有效的促进兔耳瘢痕的愈合、组织再生，尤其是有极为优异的促进毛囊修复再生的活性。

基于此，本发明还提供上述胶原蛋白在促进组织再生、促进毛囊修复方面的用途。

本发明还提供一种组合物，所述组合物包含上述胶原蛋白或上述方法制备的胶原蛋白。

本发明还提供一种制品，所述制品包含上述胶原蛋白或上述方法制备的胶原蛋白或上述组合物。所述制品包括但不限于药物、药物组合物、医疗器材、生物材料、组织工程产品、化妆品或保健品等。

本发明还提供所述胶原蛋白、核苷酸、重组表达载体、工程菌、组合物等在制备制成品中的用途，包括但不限于药物、药物组合物、医疗器材、生物材料、组织工程产品、化妆品、保健品中的用途。

本发明还提供所述胶原蛋白、核苷酸、重组表达载体、工程菌、或组合物在制备促进瘢痕愈合、组织修复或促进毛囊修复的产品中的用途。根据本发明的一个实施例，所述产品为外用制剂，优选为外用涂抹制剂，优选的为凝胶剂。

本发明的有益效果：

（1）本发明选择了毕赤酵母表达系统，毕赤酵母是真核微生物，拥有真核生物完整的细胞器，能对翻译的蛋白质进行一定的翻译后修饰（尤其是糖基化修饰），有力支撑蛋白质生物学功能的实现。以其建立表达系统同时拥有微生物表达系统可高密度发酵生产、极低的培养生成成本、短周期、高表达等规模化工业生产的优点；其可分泌于胞外的特性，可完全菌体裂解带来的杂质蛋白；细胞壁成分中不含内毒素、肽聚糖。

（2）本发明中重组人 XVII 型胶原蛋白选取的序列为经过筛选优化的序列组合而成，其氨基酸序列与天然胶原蛋白氨基酸序列相应部分同源为 100%，无免疫原性，可以广泛应用于医药、医疗器械、生物材料、组织工程、化妆品等领域。

（3）本发明中重组人 XVII 型胶原蛋白经过实验验证，具有与商品化的天然人胶原相同

或更优的细胞粘附活性、促细胞迁移活性、组织修复和促进毛囊修复再生等多种生物学活性，能达到真正产品应用的目的。

(4) 本发明中重组人 XVII 型胶原蛋白可于毕赤酵母表达系统中高效分泌表达，易于纯化获得高纯度的胶原蛋白；编码氨基酸序列的基因 DNA 序列进行了优化设计，使得构建的工程菌株蛋白表达量很高，可于高密度发酵表达时获得很高的产量；毕赤酵母易于进行高密度发酵表达，本发明筛选的菌种已能实现 500L 级中试规模的高密度高表达量发酵，已具备工业化广大生产的条件。

(5) 本发明中的宿主细胞毕赤酵母对表达的外源蛋白有一定的翻译后修饰，例如糖基化等。

附图说明

图 1 为本发明构建的 pPIC9K-170801 载体图谱。

图 2 为本发明构建的 pPIC9K-170802 载体图谱。

图 3 为本发明构建的 pPIC9K-col17a1 载体图谱。

图 4 为重组全长人 XVII 型胶原蛋白 $\alpha 1$ 链表达的 SDS-PAGE 结果图，其表观分子量约为 180kDa。

图 5 为重组全长人 XVII 型胶原蛋白 $\alpha 1$ 链使用抗 His 标签抗体进行 Western Blot 检测结果。

图 6 为诱导 24h 后 170801、170802 胶原蛋白的工程菌表达上清的 SDS-PAGE 结果。

图 7 为诱导 24h 后 170801、170802 胶原蛋白的工程菌表达上清使用抗 His 标签抗体进行 Western Blot 检测结果。

图 8 为诱导 24h 后 170801、170802 胶原蛋白的工程菌表达上清使用抗 Strep-Tag II 标签抗体进行 Western Blot 检测结果。

图 9 为 170801 胶原蛋白的 SDS-PAGE 条带的质谱分析与天然人 XVII 型胶原蛋白序列比对结果。

图 10 为 170802 胶原蛋白的 SDS-PAGE 条带的质谱分析与天然人 XVII 型胶原蛋白序列比对结果。

图 11 为 171801 胶原蛋白的工程菌在 500L 发酵罐中试发酵菌生长曲线和胶原蛋白表达曲线。

图 12 为 171802 胶原蛋白的工程菌在 500L 发酵罐中试发酵菌生长曲线和胶原蛋白表达曲线。

图 13 为纯化后的 170801 胶原蛋白 HPLC 图谱。

图 14 为纯化后的 170802 胶原蛋白 HPLC 图谱。

图 15 为 170801 胶原蛋白分子量的 MALDI-TOF MS 检测结果。

图 16 为 170802 胶原蛋白分子量的 MALDI-TOF MS 检测结果。

图 17 为纯化后的 170801、170802 胶原蛋白的红外图谱。

图 18 为 170801（左）、170802（右）胶原蛋白冻干海绵样品的 SEM 图（200×）。

图 19 为 170801、170802 胶原蛋白的细胞粘附活性检测结果。

图 20 为 170801、170802 胶原蛋白与天然人胶原蛋白的细胞迁移状态实际对比结果。

图 21 为细胞迁移 24h、48h 后，170801、170802 胶原蛋白与天然人胶原蛋白的细胞迁移率。

图 22 为 170801、170802 胶原蛋白对兔耳瘢痕干预 6 周的瘢痕大体照片和瘢痕切片图。

具体实施方式

为了使本领域技术人员更好的理解本发明的技术方案，下面对本发明的较佳实施例进行详细的阐述，但是如下实施例并不限制本发明的保护范围。

本发明的实施例中，没有多作说明的都是采用常规分子生物学实验方法完成，实施例中所涉及 PCR、酶切、连接、密码子优化等过程都是本领域技术人员根据产品说明书或本领域基础知识可以理解并且容易实现的，因此不再详细描述。

实施例 1. 重组人 XVII 型胶原蛋白的构建、表达和鉴定

(1) 重组人 XVII 型胶原蛋白氨基酸序列的设计

本发明中，基于人 XVII 型胶原蛋白序列优化，所述人 XVII 型胶原蛋白具体序列参考：Uniprot Q9UMD9-1 序列（<https://www.uniprot.org/uniprot/Q9UMD9>）和 NCBI 参照序列 Q9UMD9.3(<https://www.ncbi.nlm.nih.gov/protein/Q9UMD9.3>)，二者序列相同，具体如 SEQ ID NO.1 所示。

SEQ ID NO.1:

MDVTKKNKRDGTEVTERIVTETVTTRLTSLPPKGGTSNGYAKTASLGGGSRLKQSLTHGS
SGYINSTGSTRGHASTSSYRRAHSPASTLPNSPGSTFERKTHVTRHAYEGSSSGNSSPEYPRK
EFASSSTRGRSQTRESEIRVRLQSASPSTRWTELDDVKRLLKGSRASVSPTRNSSNTLPIPKK
GTVETKIVTASSQSVSGTYDATILDANLPSHVWSSTLPAGSSMGTYHNNMTTQSSSLLNTNA
YSAGSVFGVPNNMASCSPTLHPGLSTSSSVFGMQNNLAPSLTTLSHGTTTTSTAYGVKKNM
PQSPAAVNTGVSTSAACTTSVQSDDLLHKDCKFLILEKDNTPAKKEMELLIMTKDSGKVFT
ASPASIAATSFSEDTLKKEKQAAYNADSGLKAEANGDLKTVSTKGKTTTADIHSYGSSGGG
GSGGGGGVGGAGGGPWGPAPAWCPGSCCSWWKLLGLLLTWLLLLGLLFLIALAEEV
RKLKARVDELERIRRSILPYGDSMDRIEKDRLQGMAPAAGADLDKIGLHSDSQEELWMFVR

KKLMMEQENGNLRGSPGPKGDMGSPGPKGDRGFPGTPGIPGPLGHPGPQGPKGQKG
SVGDPMMEGPMMGQRGREGPMGPRGEAGPPGSGEKGERGAAGEPGPHGPPGVPGSVGP
KGSSGSPGPQGPPGPVGLQGLRGEVGLPGVKGDKGPMGPPGPKGDOGEKGPRGLTG
EPGMRGLPGAVGEPGAKGAMGPAGPDGHQGPGEQGLTGMPGIRGPPGPSGDPGKPGLTG
PQGPQGLPGTPGRPGIKGEPGAPGKIVTSEGSSMLTVPGPPGPPGAMGPPGPPGAPGAPGA
GLPGHQEVLNLQGP GPPGPRGPPGPSIPGPPGPRGPPGEGLPGPPGPPGSFLSNSETFLSGPP
GPPGPPGPKGDOGPPGPRGHQGEQGLPGFSTSGSSSFGLNLQGP GPPGPPGPQGPKGDKG
DPGVPGALGIPSGPSEGGSSSTMYVSGPPGPPGPPGPPGSISSSGQEIQYISEYMQSDSIRSY
LSGVQGP GPPGPPGPPGPPVTITGETFDYSELASHVVSYLRTSGYGVSLFSSSISSEDILAVLQRD
DVRQYLRQYLMGPRGPPGPPGASGDGSLSLDYAELSSRILSYMSSSGISIGLP GPPGPPGLP
GTSYEELLSLLRGSEFRGIVGPPGPPGPPGIPGNVWSSISVEDLSSYLHTAGLSFIPGPPGPPGP
PGPRGPPGVSGALATYAAENSDFRSELISYLTSPDVRSFIVGPPGPPGPQGPPGDSRLLSTDA
SHSRGSSSSSHSSSVRRGSSSYSSSMSTGGGGAGSLGAGGAFGEAAGDRGPYGTDIGPGGGY
GAAAEAGMYAGNGGLLGADFAGDLDYNELAVRVSESMQRQGLLQGMAYTVQGPPGQPG
PQGP GPGISKVFSAYSNVTADLMDFFQTYGAIQGP GPKGEMGTPGPKGDRGPAGPPGHP
GPPGPRGHKGEKGDKGQVYAGR RRRRRSIAVKP

上述序列 SEQ ID NO.1 中加粗有下划线部分即为本发明所选取的序列，共 233 个氨基酸。本发明所选取的序列是从人 XVII 型胶原蛋白的第 15 螺旋区、羧基端区、中间区域等多个螺旋区序列中优化筛选后选取的组合序列，不但能够实现人 XVII 型胶原蛋白在毕赤酵母这种真核宿主细胞的高效分泌可溶性表达，而且还具有比商品化的天然人胶原蛋白更为优异的细胞粘附活性、细胞迁移活性以及优异组织再生和促进毛囊修复再生的生物学活性；这样的组合使各区域序列的功能得以集成，并不是分别单独表达而是作为一个整体进行表达，避免了重组胶原蛋白的序列单一性，具有极强的创新性，本发明中将其命名为 170801，是人 XVII 型胶原蛋白的优选序列，具体序列如 SEQ ID NO.2 所示。

SEQ ID NO.2:

GSPGPKGDMGSPGPKGDRGFPGTPGIPGPLGHPGPQGPKGQKGSVGDPMMEGPMGEKGE
RGAAGEPGPHGPPGVPGSVGPKGSSGSPGPQGPPGPVGLQGLRGEVGLPGVKGDKGPMGP
PGPKGDQGEKGPPGPPGPPGPKGDQPPGPRGHQGEQGLPGFSGPPGPPGPQGPKGDKGDP
GVPGALGIPGPPGQKGEMGTPGPKGDRGPAGPPGHPGPPGPRGHKGEKGDKGQ

编码 170801 的 DNA 序列如 SEQ ID NO.6 所示：

SEQ ID NO.6:

GGTTCTCCTGGTCCAAAAGGTGACATGGGTTCTCCAGGTCCTAAAGGTGACAGAGG
TTTTCCTGGTACTCCAGGTATTCCTGGTCCATTGGGTCATCCAGGTCCTCAAGGTCCAAA
AGGTCAAAAAGGGTTCTGTTGGTGACCCTGGTATGGAAGGTCCAGGTGAAAAGGGTGAA

AGAGGTGCTGCTGGTGAACCTGGTCCACACGGTCCACCTGGTGTTCCAGGTTCTGTTGG
TCCTAAGGGTTCTTCTGGTTCTCCTGGTCCACAAGGTCCTCCTGGTCCAGTTGGTTTGCA
AGGTTTGAGAGGTGAAGTTGGTTTGCCAGGTGTTAAGGGTGACAAGGGTCCAATGGGT
CCTCCAGGTCCAAAAGGTGACCAAGGTGAAAAGGGTCCACCTGGTCCTCCTGGTCCAC
CTGGTCCAAAGGGTGACCAAGGTCCTCCAGGTCCTAGAGGTCATCAAGGTGAACAAGG
TTTGCCTGGTTTTTCTGGTCCACCAGGTCCTCCAGGTCCACAAGGTCCTAAAGGTGACA
AAGGTGACCCAGGTGTTCTGGTGCTTTGGGTATTCCAGGTCCACCTGGTCAAAAAGGT
GAAATGGGTACTCCAGGTCCTAAGGGTGACAGAGGTCCAGCTGGTCCACCTGGTCATCC
TGGTCCACCAGGTCCAAGAGGTCATAAGGGTGAAAAAGGTGACAAGGGTGACCAA

本发明中的胶原蛋白，还包括以 170801 为基本单位，进行 n 次串联重复，得到的氨基酸序列组成，其中 n 为大于或等于 1 的整数。

当以 170801 为基本单位，当 n 为 2 时，170801 进行 2 次串联重复，得到的氨基酸序列，共 466 个氨基酸，将其命名记为 170802，其氨基酸序列如 SEQ ID NO.3 所示。

SEQ ID NO.3:

GSPGPKGDMGSPGPKGDRGFPGTPGIPGPLGHPGPQGPKGQKGSVGDPMEGPGEKGE
RGAAGEPGPHGPPGVPVPGSVGPKGSSGSPGPQGPVGLQGLRGEVGLPGVKGDKGPMGP
PGPKGDQGEKGPPGPPGPPGPKGDQPPGPRGHQGEQGLPGFSGPPGPPGPQGPKGDKGDP
GVPGALGIPGPPGQKGEMGTPGPKGDRGPAGPPGHPGPPGPRGHKGEKGDKGDPGSPGPK
GDMGSPGPKGDRGFPGTPGIPGPLGHPGPQGPKGQKGSVGDPMEGPGEKGERGAAGEPG
PHGPPGVPVPGSVGPKGSSGSPGPQGPVGLQGLRGEVGLPGVKGDKGPMGPPGPKGDQGE
KGPVPGPPGPPGPKGDQPPGPRGHQGEQGLPGFSGPPGPPGPQGPKGDKGDPGVPGALGI
PGPPGQKGEMGTPGPKGDRGPAGPPGHPGPPGPRGHKGEKGDKGDP

编码 170802 的 DNA 序列如 SEQ ID NO.7 所示。

SEQ ID NO.7:

GGTTCTCCAGGTCCTAAAGGTGACATGGGTTCTCCAGGTCCTAAGGGTGACAGAGG
TTTTCCAGGTACTCCAGGTATTCCAGGTCCTTTGGGTCATCCAGGTCCTCAAGGTCCTAA
AGGTCAAAAAGGTTCTGTTGGTGACCCTGGTATGGAAGGTCCTGGTGAAAAAGGTGAA
AGAGGTGCTGCTGGTGAACCTGGTCCACACGGTCCTCCAGGTGTTCTGTTTCTGTTGG
TCCAAAAGGTTCTTCTGGTTCTCCTGGTCCACAAGGTCCTCCAGGTCCTGTTGGTTTGC
AAGGTTTGAGAGGTGAAGTTGGTTTGCCAGGTGTTAAAGGTGACAAAGGTCCAATGGG
TCCTCCAGGTCCAAAGGGTGACCAAGGTGAAAAAGGTCCACCTGGTCCTCCTGGTCCA
CCAGGTCCAAAAGGTGACCAAGGTCCACCTGGTCCAAGAGGTACCAAGGTGAACAA
GGTTTGCCTGGTTTTTCTGGTCCAGGTCCTCCTGGTCCCTCAAGGTCCAAAGGGTGA

CAAGGGTGACCCTGGTGTTCAGGTGCTTTGGGTATTCCTGGTCCTCCAGGTCAAAAGG
GTGAGATGGGTACTCCTGGTCCTAAGGGTGACAGAGGTCCAGCTGGTCCTCCTGGTCAC
CCAGGTCCCTCCTGGTCCTAGAGGTCATAAAGGTGAAAAAGGTGACAAGGGTGACCAAG
GTTCTCCAGGTCCAAAGGGTGACATGGGTTCCTGGTCCAAAAGGTGACAGAGGTTTC
CCTGGTACTCCAGGTATTCCTGGTCCATTGGGTACCCAGGTCCACAAGGTCCAAAAGG
TCAAAAAGGTTCTGTTGGTGACCCAGGTATGGAAGGTCCAGGTGAAAAGGGTGAAAGA
GGTGCTGCTGGTGAACCAGGTCCCTCATGGTCCACCAGGTGTTCCAGGTTCTGTTGGTCC
AAAGGGTTCTTCTGGTTCTCCAGGTCCACAAGGTCCTCCAGGTCCAGTTGGTTTGCAAG
GTTTGAGAGGTGAAGTTGGTTTGCTGGTGTTAAGGGTGACAAAGGTCCTATGGGTCCCT
CCTGGTCCTAAAGGTGACCAAGGTGAAAAGGGTCCACCAGGTCCTCCAGGTCCACCTG
GTCCAAAAGGTGACCAAGGTCCACCAGGTCCCTAGAGGTCATCAAGGTGAACAAGGTTT
GCCAGGTTTTTCTGGTCCACCAGGTCCACCAGGTCCTCAAGGTCCTAAGGGTGACAAAG
GTGACCCAGGTGTTCTCCTGGTGCTTTGGGTATTCCTGGTCCACCTGGTCAAAAGGGTGAA
ATGGGTACTCCTGGTCCTAAAGGTGACAGAGGTCCTGCTGGTCCACCTGGTCATCCAGG
TCCACCTGGTCCAAGAGGTCACAAAGGTGAAAAGGGTGACAAGGGTGACCAA

蛋白质的功能需要其氨基酸残基有序的排列方能实现，本发明中重组人 **XVII 型** 胶原蛋白所有创新性生物学活性建立在优化筛选选取的氨基酸序列的基础之上，有限的序列修改或一定比例的同源性（大于 80%）仍可能实现相同或类似的生物学活性。

所以本发明中，所述的胶原蛋白，还包括在 SEQ ID NO.2 基础上进行一定程度上的氨基酸插入、取代、添加、缺失等修改后的氨基酸序列或与 SEQ ID NO.2 氨基酸序列具有大于 80% 同源性的氨基酸序列，以该氨基酸序列为基本单元；进行 n 次串联重复得到的氨基酸序列。其中 n 为大于或等于 1 的整数，每个基本单元相同或不同。

其中，本发明所述的同源性指的是序列的重合率，可以是两序列直接的数量关系，如部分相同、相似的百分比；还包括两序列在同一位点核苷酸或氨基酸残基完全相同的序列比例；也可以是从某个共同祖先经趋异进化形成的不同的序列。可以通过常规生物信息学方法比对序列信息得到。

所述的氨基酸插入指在氨基酸序列，例如 SEQ ID NO .2 或 SEQ ID NO .3 序列的适当位置插入氨基酸残基，插入的氨基酸残基也可以全部或部分彼此相邻，或插入的氨基酸之间都不彼此相邻。只要所述蛋白具有与 170801、170802 胶原蛋白有类似的生物学活性，包括细胞粘附活性、促细胞迁移活性、促进组织再生和促进毛囊修复再生的活性。

所述的氨基酸取代指在氨基酸序列，例如 SEQ ID NO .2 或 SEQ ID NO .3 序列某个位置的某个或若干个氨基酸残基（包括连续若干个氨基酸残基或非连续的若干个氨基酸残基）被其他氨基酸残基替代，只要所述蛋白具有与 170801、170802 胶原蛋白有类似的生物学活性，

包括细胞粘附活性、促细胞迁移活性、促进组织再生和促进毛囊修复再生的活性。

所述的氨基酸添加指在氨基酸序列，例如 SEQ ID NO. 2 或 SEQ ID NO. 3 序列的 C 端或 N 端添加氨基酸，只要所述蛋白具有与 170801、170802 胶原蛋白有类似的生物学活性，包括细胞粘附活性、促细胞迁移活性、促进组织再生和促进毛囊修复再生的活性。

所述的氨基酸缺失指可以从氨基酸序列，例如 SEQ ID NO. 2 或 SEQ ID NO. 3 序列中删除 1、2 或 3 个及 3 个以上氨基酸，只要所述蛋白具有与 170801、170802 胶原蛋白有类似的生物学活性，包括细胞粘附活性、促细胞迁移活性、促进组织再生和促进毛囊修复再生的活性。

本发明中所述的氨基酸优选为 L 型氨基酸；同时其中的一个或多个(如 2-5 个、2-4 个或 2-3 个)氨基酸也可以用构象为 D 型的氨基酸、人工修饰的氨基酸、自然界存在的稀有氨基酸等进行替换，以提高多肽的生物利用度、稳定性和/或抗病毒活性。其中 D 型氨基酸是指与组成蛋白质的 L 型氨基酸相对应的氨基酸；人工修饰的氨基酸指经过甲基化、磷酸化等修饰的组成蛋白质的常见 L 型氨基酸；自然界存在的稀有氨基酸包括组成蛋白质的不常见氨基酸和不组成蛋白质的氨基酸，例如 5-羟基赖氨酸、甲基组氨酸、 γ 氨基丁酸、高丝氨酸等。

(2) 重组人 XVII 型胶原蛋白氨基酸序列和 DNA 序列优化

本发明中，对 170801、170802 的 DNA 序列进行修饰，分别于二者氨基端添加编码 Strep-Tag II 标签的 DNA 序列、羧基端添加编码 6 $\times$ His Tag 标签的 DNA 序列，使其含有双特异性亲和纯化标记，这样可进行亲和层析纯化，也便于以两种标签序列为基础进行免疫学抗体检测。具体的是以 170801、170802 的氨基酸序列及其基因序列为基础，进行 DNA 序列的密码子偏好性及其转录、翻译过程中相关优化参数进行计算、拼接重组和优化，合成更适宜在毕赤酵母中高效表达的 DNA 序列。

本发明中，在 170801 两端分别修饰，氨基端添加编码 Strep-Tag II 标签的 DNA 序列、羧基端添加编码 6 $\times$ His Tag 标签的 DNA 序列后，170801 最终表达获得的是含有标签的蛋白，共 249 个氨基酸，优化后的 170801 的氨基酸序列如 SEQ ID NO.4 所示。在 170801 两端分别修饰，氨基端添加编码 Strep-Tag II 标签的 DNA 序列、羧基端添加编码 6 $\times$ His Tag 标签的 DNA 序列、终止密码子 (TAA) 及 Not I 的酶切位点的 DNA 序列如 SEQ ID NO.8 所示。

SEQ ID NO.4:

YVWSHPQFEKGSPGPKGDMGSPGPKGDRGFPGTPGIPGPLGHPGPQGPKGQKGSVGDP
GMEGPGEKGERGAAGEPGPHGPPGVPGSVGPKGSSGSPGPQGPVGLQGLRGEVGLPG
VKGDKGPMGPPGPKGDQGEKGPPGPPGPKGDQGPGRGHQGEQGLPGFSGPPGPPG
PQGPKGDKGDGVPVPGALGIPGPPGQKGEMGTPGPKGDRGPAGPPGHPGPPGPRGHKGEKG
DKGDQHHHHH

TACGTATGGTCTCATCCACAATTCGAGAAGGGTTCTCCTGGTCCAAAAGGTGACATG
GGTTCTCCAGGTCCTAAAGGTGACAGAGGTTTTCTGGTACTCCAGGTATTCTGGTCCA
TTGGGTCATCCAGGTCCTCAAGGTCCAAAAGGTCAAAAGGGTTCTGTTGGTGACCCTGG
TATGGAAGGTCCAGGTGAAAAGGGTGAAAGAGGTGCTGCTGGTGAACCTGGTCCACAC
GGTCCACCTGGTGTTCAGGTTCTGTTGGTCCTAAGGGTTCTTCTGGTTCTCCTGGTCCA
CAAGGTCCTCCTGGTCCAGTTGGTTTGCAAGGTTTGAGAGGTGAAGTTGGTTTGCCAGG
TGTTAAGGGTGACAAGGGTCCAATGGGTCCTCCAGGTCCAAAAGGTGACCAAGGTGAA
AAGGGTCCACCTGGTCCCTCCTGGTCCACCTGGTCCAAAGGGTGACCAAGGTCTCCAG
GTCCTAGAGGTCATCAAGGTGAACAAGGTTTGCCTGGTTTTCTGGTCCACCAGGTCCT
CCAGGTCCACAAGGTCCTAAAGGTGACAAAGGTGACCCAGGTGTTCTCCTGGTGCTTTGG
GTATTCCAGGTCCACCTGGTCAAAAAGGTGAAATGGGTACTCCAGGTCCTAAGGGTGAC
AGAGGTCCAGCTGGTCCACCTGGTCATCCTGGTCCACCAGGTCCAAGAGGTCATAAGGG
TGAAAAAGGTGACAAGGGTGACCAACACCATCATCACCATCATTAAGCGGCCGC

同样的，在 170802 两端分别修饰添加氨基端添加编码 Strep-Tag II 标签的 DNA 序列、羧基端添加编码 6×His Tag 标签的 DNA 序列后，170802 最终表达获得的是含有标签的蛋白，共 482 个氨基酸，优化后的 170802 的氨基酸序列如 **SEQ ID NO.5** 所示。在 170802 两端分别修饰添加氨基端添加编码 Strep-Tag II 标签的 DNA 序列、羧基端添加编码 6×His Tag 标签的 DNA 序列和终止密码子（TAA）及 Not I 的酶切位点的 DNA 序列如 **SEQ ID NO.9** 所示。

SEQ ID NO.5:

YVWSHPQFEKGSPGPKGDMGSPGPKGDRGFPGTPGIPGPLGHPGPQGPKGQKGSVGD
PMEGPGEKGERGAAGEPGPHGPPGVPGSVGPKGSSGSPGPQGPVGLQGLRGEVGLPG
VKGDKGPMGPPGPKGDQGEKGPPGPPGPPGPKGDQGPGRGHQGEQGLPGFSGPPGPPG
PQGPKGDKGDPGVPGALGIPGPPGQKGEMGTPGPKGDRGPAGPPGHPGPPGPRGHKGEKG
DKGDQGSPPGPKGDMGSPGPKGDRGFPGTPGIPGPLGHPGPQGPKGQKGSVGDPMEGPGE
KGERGAAGEPGPHGPPGVPGSVGPKGSSGSPGPQGPVGLQGLRGEVGLPGVKGDKG
PMGPPGPKGDQGEKGPPGPPGPPGPKGDQGPGRGHQGEQGLPGFSGPPGPPGPPGPKGD
KGDPGVPGALGIPGPPGQKGEMGTPGPKGDRGPAGPPGHPGPPGPRGHKGEKGDKGDQHH
HHHH

SEQ ID NO.9:

TACGTATGGTCTCATCCTCAATTTGAAAAGGGTTCTCCAGGTCCTAAAGGTGACATG
GGTTCTCCAGGTCCTAAGGGTGACAGAGGTTTTCCAGGTACTCCAGGTATTCCAGGTCC
TTTGGGTCATCCAGGTCCTCAAGGTCCCTAAAGGTCAAAAAGGTTCTGTTGGTGACCCTG

GTATGGAAGGTCCTGGTGAAAAAGGTGAAAGAGGTGCTGCTGGTGAACCTGGTCCACA
CGGTCCTCCAGGTGTTCTCTGGTTCTGTTGGTCCAAAAGGTTCTTCTGGTTCTCCTGGTCC
ACAAGGTCCTCCAGGTCTGTTGGTTTGCAAGGTTTGAGAGGTGAAGTTGGTTTGCCAG
GTGTTAAAGGTGACAAAGGTCCAATGGGTCCTCCAGGTCCAAAGGGTGACCAAGGTGA
AAAAGGTCCACCTGGTCCTCCTGGTCCACCAGGTCCAAAAGGTGACCAAGGTCCACCT
GGTCCAAGAGGTACCAAGGTGAACAAGGTTTGCCTGGTTTTTCTGGTCCTCCAGGTCC
TCCTGGTCCTCAAGGTCCAAAGGGTGACAAGGGTGACCCTGGTGTTCAGGTGCTTTG
GGTATTCCTGGTCCTCCAGGTCAAAAGGGTGAGATGGGTACTCCTGGTCCTAAGGGTGA
CAGAGGTCCAGCTGGTCCTCCTGGTCACCCAGGTCCCTCCTGGTCCTAGAGGTGATAAAG
GTGAAAAAGGTGACAAGGGTGACCAAGGTTCTCCAGGTCCAAAGGGTGACATGGGTTC
TCCTGGTCCAAAAGGTGACAGAGGTTTCCCTGGTACTCCAGGTATTCCTGGTCCATTGG
GTCACCCAGGTCCACAAGGTCCAAAAGGTCAAAAAGGTTCTGTTGGTGACCCAGGTAT
GGAAGGTCCAGGTGAAAAGGGTGAAAGAGGTGCTGCTGGTGAACCAGGTCCCTCATGGT
CCACCAGGTGTTCCAGGTTCTGTTGGTCCAAAGGGTTCTTCTGGTTCTCCAGGTCCACA
AGGTCCCTCCAGGTCCAGTTGGTTTGCAAGGTTTGAGAGGTGAAGTTGGTTTGCTGGTG
TTAAGGGTGACAAAGGTCCTATGGGTCCTCCTGGTCCTAAAGGTGACCAAGGTGAAAA
GGGTCCACCAGGTCCCTCCAGGTCCACCTGGTCCAAAAGGTGACCAAGGTCCACCAGGT
CCTAGAGGTCATCAAGGTGAACAAGGTTTGCCAGGTTTTTCTGGTCCACCAGGTCCACC
AGGTCCCTCAAGGTCCTAAGGGTGACAAAGGTGACCCAGGTGTTCTCCTGGTGCTTTGGGTA
TTCCTGGTCCACCTGGTCAAAAGGGTGAAATGGGTACTCCTGGTCCTAAAGGTGACAGA
GGTCCTGCTGGTCCACCTGGTCATCCAGGTCCACCTGGTCCAAGAGGTGACAAAGGTG
AAAAGGGTGACAAGGGTGACCAACACCATCACCATCATCATTAAGCGGCCGC

本发明中，编码全长人 XVII 型胶原蛋白氨基酸序列（即人 XVII 型胶原蛋白 $\alpha 1$ 链）的 DNA 序列（氨基端添加 EcoR I 的酶切位点、Strep-Tag II 标签的 DNA 序列、羧基端添加编码 6xHis Tag 标签的 DNA 序列和终止密码子（TAA）及 Not I 的酶切位点）如 **SEQ ID NO.10** 所示。

SEQ ID NO.10:

GAATTCTGGAGTCATCCTCAATTCGAAAAAATGGATGTCACTAAAAAGAACAAGAG
AGACGGAAGTGAAGTCACTGAGAGAATCGTTACCGAAAGTGTACCACAAGACTTACT
TCATTACCTCCAAAGGGTGGAAGTTCTAATGGTTACGCAAAGACAGCATCATTAGGAGG
TGGTTCAAGACTTGAGAAACAATCCCTTACTCATGGTTCAAGTGGATACATAAATTCAAC
TGGTTCAACAAGAGGACATGCAAGTACTTCTTCTTATAGAAGAGCACATAGTCCAGCAT
CAACTTTGCCCAACTCTCCTGGTTCAACATTTGAGAGAAAACTCATGTAACCAGACAT
GCCTATGAGGGTTCTAGTTCTGGTAATTCATCTCCAGAATACCCTAGAAAAGAGTTTGCA

TCATCCTCAACTAGAGGTAGATCACAGACTAGAGAATCTGAAATCAGAGTCAGATTACA
ATCAGCATCTCCTTCAACTAGATGGACTGAGTTAGACGACGTGAAAAGATTATTAAAGG
GATCAAGAAGTGCAAGTGTTTCCCTTACAAGAAATCTTCCAACACCCTTCCCATTCTTA
AGAAAGGAACCGTTGAAACTAAAATCGTTACAGCATCATCTCAGTCTGTATCCGGAACC
TACGACGCTACCATTTCTGGACGCCAACTTACCTTCTCATGTCTGGTCTTCAACTTTACCC
GCAGGTTCTCAATGGGAACTTACCACAATAACATGACTACTCAATCAAGTTCTCTTCTG
AATACCAATGCATACTCAGCCGGTTCCGTTTTTGGAGTCCCTAACAATATGGCCTCTTGCT
CCCCAACTCTTCATCCAGGTCTTTCAACCTCATCAAGTGTATTTGGTATGCAGAACAACT
TAGCCCCCTTCTTGACAACCCTGTCTCATGGTACTACTACCACTAGTACAGCATATGGAG
TCAAGAAGAACATGCCACAGTCACCTGCTGCCGTTAACACTGGAGTTTCAACATCTGCT
GCCTGCACTACATCTGTTCAATCAGATGATCTTCTGCATAAGGACTGCAAGTTTTTAATTT
TAGAGAAAGACAATACCCCTGCCAAAAAGGAAATGGAGTTACTTATAATGACCAAAGAT
TCTGGTAAAGTATTCACTGCTTCTCCCGCTAGTATCGCCGCAACTTCATTCTCTGAAGATA
CTTTAAAGAAGGAAAAACAAGCCGCATACAACGCAGACTCAGGTTTAAAAGCAGAAGC
AAACGGTGACCTTAAAACAGTGTCAACTAAGGGAAAGACTACAACCGCAGACATCCAT
TCATATGGTTCTTCAGGTGGAGGAGGATCTGGAGGAGGTGGTGGAGTGGGAGGTGCTG
GAGGAGGTCCATGGGGTCTGCACCTGCATGGTGGCCTTGCGGTTCTTGCTGCAGTTGG
TGGAAGTGGCTTCTGGGTTTATTATTAACCTTGGCTGCTTTTACTGGGTCTTTTATTCGGTC
TTATCGCATTAGCAGAAGAGGTGAGAAAACCTAAGGCCAGAGTTGATGAGTTAGAGAGA
ATCAGAAGATCAATCCTTCCCTATGGAGACTCCATGGACAGAATCGAAAAGGATAGACT
TCAGGGTATGGCCCTGCAGCAGGAGCTGATCTGGATAAAATCGGACTTCATTGAGATTC
TCAAGAGGAATTATGGATGTTTGTAGAAAAGAACTTATGATGGAGCAGGAGAACGGTA
ACCTGAGAGGTTCTCCTGGTCCAAAAGGAGATATGGGTTTACCCGGTCCCAAAGGAGAT
AGAGGATTCCCTGGTACTCCAGGTATCCCCGGTCCCCTGGGTACCCCTGGACCTCAAGG
TCCTAAAGGTCAAAAGGGTTCTGTAGGAGATCCAGGTATGGAGGGTCCCATGGGTGAGA
GAGGTAGAGAAGGTCCCATGGGACCAAGAGGTGAAGCTGGACCTCCCGGAAGTGGTG
AAAAAGGAGAAAGAGGAGCAGCAGGAGAACCTGGACCCCATGGACCTCCAGGAGTTC
CTGGATCAGTCGGACCCAAAGGTTTCATCCGGTTCTCCTGGACCTCAAGGTCCACCAGGA
CCCGTCGGATTGCAAGGATTGAGAGGAGAAGTTGGACTTCCCGGAGTTAAGGGTGACA
AGGGTCCTATGGGTCTCCTGGTCCAAAAGGGAGATCAGGGTGAAAAGGGTCCTAGAGG
TCTGACTGGTGAACCAGGAATGAGAGGACTTCCCGGTGCCGTGGGTGAACCCGGTGCA
AAAGGAGCAATGGGTCTGCCGGTCTGATGGACACCAGGGACCCAGAGGAGAGCAG
GGATTAACAGGAATGCCTGGTATCAGAGGTCCCCCAGGTCCCTCAGGAGACCCAGGAA
AGCCAGGACTTACTGGTCCCCAGGGTCTCAAGGTCTGCCTGGAACCTCCCGGAAGACC
CGGAATCAAAGGTGAACCAGGAGCCCCAGGAAAAATCGTTACTAGTGAAGGATCTTCA

ATGCTGACTGTGCCAGGTCCACCTGGTCCTCCTGGAGCTATGGGTCTCCCGGTCCCCCT
GGAGCACCCGGTCCTGCAGGTCTGCGGATTACCTGGACATCAGGAAGTCTTAAACCT
GCAAGGTCCACCAGGTCTCCTGGTCCAAGAGGACCCCCTGGTCCTTCAATCCCTGGTC
CACCTGGACCTAGAGGTCCACCCGGAGAGGGACTTCCCGGACCACCAGGACCTCCTGG
ATCATTTCTGTCTAATTCTGAGACATTTCTTTTCAGGACCTCCTGGTCCTCCCGGACCACCT
GGACCAAAGGAGATCAGGGACCACCTGGTCCCAGAGGACATCAAGGAGAGCAGGGT
CTTCCAGGTTTCTCTACTTCTGGATCATCATCATTTGGTCTTAATCTTCAAGGTCCTCCTG
GACCCCCAGGACCACAGGGACCCAAGGGTGACAAGGGAGACCCTGGAGTCCCAGGTG
CACTGGGAATCCCTTCAGGTCTTTCAGAAGGTGGTTCATCTTCAACCATGTATGTGTCTG
GACCCCCAGGACCTCCCGGACCTCCAGGTCTCCTGGTTCAATCTCTTCTTCTGGTCAA
GAAATTCAGCAGTATATTTCTGAGTACATGCAATCTGACTCTATTAGAAGTTATTTGTCTG
GTGTGCAGGGTCCACCAGGTCTCCAGGTCCCCCTGGACCAGTCACTACTATCACTGGA
GAGACATTTGATTATAGTGAATTAGCATCACACGTCGTTTCTTATCTGAGAACTTCAGGTT
ATGGTGTTTCATTATTTTCCTCCTCAATCTCTTCAGAAGACATCTTAGCCGTACTGCAGAG
AGATGATGTAAGACAGTACCTTAGACAATACTTGATGGGTCCAAGAGGACCACCAGGTC
CACCCGGTGCATCAGGAGACGGATCATTATTATCTTTGGATTACGCTGAATTATCATCAAG
AATCCTTAGTTACATGTCTCTTCTGGTATCTCCATAGGTCTGCCCGGTCTCCTGGTCCT
CCCGGACTTCCTGGTACTTCTTACGAAGAGCTTCTGTCACTTTTAAGAGGATCTGAGTTT
AGAGGTATAGTTGGTCCCCCAGGACCTCCAGGTCTCCTCAGGTATCCCAGGTAACGTGTG
GTCATCTATCTCAGTTGAGGATCTTTCATCTTATCTTCACACCGCCGGATTGTCCTTTATAC
CTGGTCTCCTCAGGACCCCCTGGACCTCCTGGACCCAGAGGTCTCCTCAGGAGTATCCGGT
GCTTTAGCAACTTATGCAGCAGAAAACCTCCGATTCTTTTAGATCAGAATTGATCTCCTAC
CTTACTTCTCCTGATGTTAGATCATTCATCGTCGGACCTCCAGGACCACCTGGACCTCAA
GGACCTCCTGGAGATTCAAGATTACTGAGTACAGATGCATCACATTCAAGAGGTTCAAG
TTCTTCTTCTCACTCCTCTTCTGTTAGAAGAGGATCATCATATTCATCAAGTATGTCAACT
GGTGGAGGAGGAGCTGGTTCATTGGGTGCCGGAGGTGCATTTGGTGAGGCCGCTGGTG
ACAGAGGACCTTATGGTACTGACATTGGACCTGGTGGAGGTTATGGTGCAGCAGCAGAA
GGTGGAATGTATGCTGGTAATGGAGGTTTACTGGGAGCCGACTTTGCCGGAGACCTGGA
CTACAACGAGTTAGCTGTTAGAGTGTGAGAGTCAATGCAGAGACAAGGATTATTACAGG
GTATGGCTTATACAGTTTCAGGGACCCCCAGGTGAGCCAGGTCTCAAGGACCCCCTGGT
ATTTCAAAGTCTTCTCCGCTTATTCTAACGTTACTGCAGACCTGATGGATTTCTTCCAGA
CTTACGGTGCCATCCAAGGTCTCCTCAGGACAGAAGGGTGAAATGGGAACTCCCGGTCC
CAAGGGTGACAGAGGTCTGACAGGACCTCCTGGTCAACCCTGGACCACCTGGACCTAGA
GGACACAAAGGTGAGAAGGGTGACAAGGGAGACCAGGTTTACGCAGGAAGAAGAAGA
AGAAGAAGTATTGCCGTCAAGCCTCACCATCACCATCATCACTAAGCGGCCGC

(3) DNA 序列的合成与重组表达载体的构建

本发明中，优化后的序列 SEQ ID NO.8、SEQ ID NO.9 和 SEQ ID NO.10 委托南京金斯瑞生物科技股分有限公司合成，三个基因片段记为 pUC57-170801、pUC57-170802、pUC57-col17a1。将合成后的基因片段 SEQ ID NO.8、SEQ ID NO.9 使用 SnaB I 和 Not I 双酶切，将酶切后得到的目的基因片段连接入同样双酶切后的 pPIC9K 空载体（购自赛默飞世尔科技公司）。同时将合成后的基因片段 SEQ ID NO.10 使用 EcoR I 和 Not I 双酶切，将酶切后得到的目的片段连接入同样双酶切后 pPIC9K 空载体，作为对比。

步骤 a. 扩增目的基因、表达载体：以质粒小提试剂盒（购自生工生物工程(上海)股份有限公司）分别提取 pUC57-170801、pUC57-170802、pUC57-col17a1、pPIC9K 空载体的质粒，具体操作按试剂盒说明书进行。

步骤 b. 将步骤 a 得到的目的基因片段与 pPIC9K 载体的质粒在 37℃ 下双酶切 30 min，酶切反应体系如下（所用 QuickCut™限制性酶购自大连 TaKaRa 公司，酶切过程按其说明书进行）：

质粒	3μg
10×H 缓冲液	5μL
EcoRI/SnaB I	10U
NotI	10U
无菌水	至 50μL

将所得到的目的基因片段和载体片段，采用 PCR 产物纯化试剂盒进行纯化（购自大连 TaKaRa 公司，具体操作按试剂盒说明书进行）。回收经双酶切处理后的目的基因片段和载体 pPIC9K 线性化片段，用 SolutionI 连接试剂（连接试剂盒购自大连 TaKaRa 公司）进行连接（16℃ 连接 30min），将目的基因片段准确插入到含有分泌信号 α-因子的分泌型载体读码框内，连接反应体系如下（按连接试剂盒说明书进行）：

pPIC9K 线性化片段	2μL
目的基因片段	3μL
SolutionI 连接试剂	5μL

将连接产物转化进入感受态大肠杆菌 DH5α（购自生工生物工程(上海)股份有限公司），在含有 Amp 的 LB 抗性平板筛选阳性克隆，进行菌落 PCR 验证，提取重组质粒进行测序鉴定（交由生工生物工程(上海)股份有限公司完成），测序验证正确的证明本实施例成功构建了重组表达载体 pPIC9K-170801、pPIC9K-170802、pPIC9K-col17a1，相关质粒图谱如图 1、图 2 和图

3 所示。

(4) 重组工程菌株的构建、菌种筛选

分别将步骤(3)得到的重组表达载体质粒 10 μ g 用 SacI (酶切试剂盒购自大连 TaKaRa 公司, 具体操作按试剂盒说明书进行) 37℃ 酶切消化过夜, 使其线性化, 再使用以 PCR 产物纯化试剂盒 (购自生工生物工程(上海)股份有限公司), 回收线性化质粒, 回收的体积控制在 10 μ L 左右。

将线性化质粒电转化入空宿主菌种毕赤酵母 GS115 (购自中国工业微生物菌种保藏管理中心) 感受态细胞中, 将电转后的菌液涂布于 MD 平板上, 每 100 μ L~200 μ L 涂布一块平板, 室温静置 10min, 于 30℃ 倒置培养 2-5 天, 直至有单菌落 (阳性转化子) 出现。

向 MD 平板表面加入 2mL 无菌双蒸水, 然后用无菌三角涂布器轻轻刮下平板表面的 His⁺ 转化子 (阳性转化子), 并转移到 50mL 离心管中, 以无菌双蒸水稀释成菌悬液, 含 10⁵ 个细胞的菌悬液涂布于含有 0.5mg/mL G418 的 YPD 平板上, 倒置, 30℃ 培养 3~4d 后至单菌落出现。从 YPD 平板上挑取菌落至无菌 96 孔板中 (每孔含 200 μ L YPD), 混匀, 于 30℃ 培养 48h; 混匀孔中菌液, 各取 10 μ L 接入至一块新的无菌 96 孔板, 于 30℃ 培养 24h 后再重复一次该步骤; 24h 后, 从第三块 96 孔板中取出 1 μ L 分别点在含有 1.0mg/mL 和 4mg/mL G418 的 YPD 平板上, 于 30℃ 继续培养 96h~120h。毕赤酵母转化子若能在含高浓度 G418 的平板上生长, 说明该转化子含有多拷贝的目的基因, 即有多个重组片段进入了酵母体内并通过同源重组整合到酵母的染色体上。经过这一步筛选可得到的高拷贝、可高效表达的重组酵母工程菌种。

构建分别含 pPIC9K-col17a1、pPIC9K-170801、pPIC9K-170802 三种工程菌样本均送至中国微生物菌种保藏管理委员会普通微生物中心, 菌种保藏编号分别是: CGMCC No.18659、CGMCC No.20626、CGMCC No. 20627。地址: 北京市朝阳区北辰西路 1 号院 3 号; 保藏日期: CGMCC No.18659 为 2019 年 10 月 10 日、CGMCC No.20626、CGMCC No. 20627 为 2020 年 09 月 09 日。分类命名: 巴斯德毕赤酵母 *Pichia pastoris*。

(5) 诱导表达与重组胶原蛋白的鉴定

在步骤(4)中含有 1.0mg/mL 和 4mg/mL G418 的 YPD 平板上分别挑选单菌落, 置于装有 10mL BMGY 培养基的 100mL 三角瓶中, 于 28-30℃、220rpm 培养至 OD₆₀₀ 为 2~6(16-18 h)。室温下 1500~3000g 离心 5min, 收集菌体, 用 BMMY 培养基重悬菌体, 使 OD₆₀₀ 为 2 左右, 放置于 28-30℃、220 rpm 的摇床上继续生长 3 天, 每 24h 向培养基中添加 100% 甲醇至终浓度为 1.0%。按时间点 ($\geq$ 24h, 必要时 24h、48h、72h、96h 均取样) 分别取菌液样品, 取样量为 1mL, 置于 1.5mL EP 管中, 最大转速离心 2~3min, 收集表达上清和菌体。将菌体以玻璃珠 (购自生工生物 (上海) 有限公司) 破碎法裂解, 裂解缓冲液 (1mM PMSF、10mM Tris,

pH7.4), 4℃下以 12000g 离心 5min, 取胞内裂解上清, 再加入 8M 尿素振荡溶解沉淀, 分析目的蛋白的表达量和菌液最佳收获时间。待检测样品于-80℃保存备用。

收取的表达上清、胞内裂解上清、胞内沉淀溶解液, 加入 5×上样缓冲液(250mM Tris-HCl、pH 值为 6.8, 10%SDS, 0.5%溴酚蓝, 50%甘油, 5%β-巯基乙醇), 置于 100℃金属浴加热 10min, 进行 SDS-PAGE 检测。因表达的目的蛋白氨基端有 Srtep-TagII 标签, 羧基端有 6×His 标签, 可以抗 Srtep-Tag II、抗 6×His Tag 的抗体(购自南京金斯瑞生物科技股份有限公司), 采用 Western Blot 进行检测。

重组全长人 XVII 型胶原蛋白 α1 链的表达与鉴定:

如图 4 的 SDS-PAGE 结果图所示, 表达上清中并没有可见预期的目的条带(其表观分子量约为 180kDa)。

使用抗 6×His Tag 抗体进行 Western Blot 检测, 结果如图 5 所示, 发现表达上清中只能检测到极少量降解后的目的蛋白(表观分子量小于预期, 约为 120kDa), 并且是在诱导表达很长时间后方能看到; 胞内裂解上清中有较多的降解后的目的蛋白(表观分子量小于预期), 其表观分子量约为 120kDa(与人 XVII 型胶原蛋白胞外域的表观分子量一致); 只有胞内沉淀溶解部分才有预期表观分子量大小(180kDa)的目的蛋白存在。实验结果说明: 作为一种跨膜类型的胶原蛋白, 其应该被固定于胞内的膜系统之上, 只有降解后才能脱落溶解于胞内, 而表达上清中检测到的降解后目的蛋白极有可能是在培养长时间中已死亡的细胞裂解后释放到表达上清液中的, 但无论哪种可能, 全长的重组人 XVII 型胶原蛋白是无法有效分泌于毕赤酵母胞外的。

170801、170802 胶原蛋白的表达与鉴定:

图 6 为诱导 24h 后, 170801、170802 胶原蛋白的工程菌表达上清的 SDS-PAGE 结果, 如图 6 所示, 170801、170802 胶原蛋白在诱导表达 24h 即可高效分泌表达于胞外的培养上清中, 170801 的表观分子量约为 30kDa, 170802 的表观分子量约为 52kDa。

图 7 为诱导 24h 后, 170801、170802 胶原蛋白的工程菌表达上清的 Western Blot 结果(Anti-His, 抗体购自生工生物工程(上海)股份有限公司, 小鼠抗 His 单克隆抗体, 货号 D199987), 图 8 为诱导 24h 后, 170801、170802 胶原蛋白的工程菌表达上清的 Western Blot 结果(Anti-Strep-Tag II, 抗体购自南京金斯瑞生物科技股份有限公司, 兔抗 Strep-Tag 多克隆抗体, 货号 A00626)。从图 7 和图 8 的结果(ECL 化学发光显色, 全自动化学发光图像分析系统 Tanon 5200 将蛋白质分子质量标准合成于图像)中可以看到, 氨基端 Srtep-TagII 标签, 羧基端 6×His 标签均可检测到, 且目的条带大小均与 SDS-PAGE 中表观分子量大小相同, 说明重组人 XVII 型胶原蛋白 170801、170802 成功进行了高效分泌表达。

回收切割诱导表达的蛋白条带，用胰蛋白酶将其酶解，Nano-HPLC-MS/MS 质谱检测重组胶原蛋白的胰蛋白酶解后肽段（委托苏州普泰生物技术有限公司完成），并将检测到的肽段进行序列比对（Uniprot 数据库）。结果如图 9、图 10 所示：170801、170802 胶原蛋白被酶解后检测到的肽段均属于氨基酸序列选取设计时选择的人 XVII 型胶原蛋白序列的相关区域，说明 170801、170802 胶原蛋白成功表达。

实施例 2. 基因工程菌的中试发酵和蛋白纯化

（1）中试发酵

对构建的分别含 pPIC9K-170801、pPIC9K-170802 的工程菌进行 50L-500L 联动中试发酵，获取含有重组人 XVII 型胶原蛋白的发酵液，实现重组人 XVII 型胶原蛋白 170801、170802 的规模化表达生产。

种子培养基 YPG（配方为：酵母粉 10g/L、酵母蛋白胨 FP102 20g/L、无水甘油 10g/L）；发酵培养基（配方为：NH₄H₂PO₄ 190.4 g/L、KH₂PO₄ 10.06 g/L、CaSO₄·2H₂O 1.18 g/L、K₂SO₄ 18.2 g/L、MgSO₄·7H₂O 14.9 g/L、甘油 40 g/L）；补料培养基（50%W/V 甘油，每升加 12mL PTM₁ 微量元素）；诱导培养基（100%甲醇，每升加入 12mL PTM₁ 微量元素）；其中 PTM₁ 用 0.22μm 的滤膜过滤除菌，4℃保存。发酵培养基高温灭菌后待温度降至室温加入 PTM₁，用氨水调节 pH 值至 5.0。

构建的工程菌株分批培养条件和诱导表达条件为：

采用分批补料培养方法，培养温度 30℃。将构建的工程菌接入装有种子培养基 YPG 的摇瓶中，药瓶容量为 2L，220rpm、30℃培养 24h；以 10%接种量接种到含发酵培养基的种子发酵罐中（使其接种后 OD₆₀₀ 为 5），发酵罐容量为 50L，扩大培养，调节搅拌转速为 200r/min-600r/min，空气通量为 2VVM，DO₂≥30%，培养至碳源耗尽，移种至含发酵培养基的发酵罐中开始发酵培养，发酵罐容量为 500L。发酵培养条件为，搅拌转速 150r/min-500r/min，罐压为 0-0.05MPa，空气通量为 1VVM，DO₂≥30%，溶氧反弹流加补料培养基，流加速率维持 DO₂≥30%，至 OD₆₀₀ 生长到 190-210，停止补料。待溶氧反弹 DO₂≥70% 后，流加诱导培养基进行诱导表达，调节转速、通气量、罐压和流加速度使 DO₂≥30%。诱导 40~72h，UV 测量蛋白浓度增长幅度不明显或下降即可放罐。

UV 蛋白定量公式：C(mg/mL)=0.144*(A₂₁₅-A₂₂₅)，A₂₁₅<1.5。

经过对构建的工程菌进行高密度中试发酵，结果发现，诱导 48h 时，菌浓度（OD₆₀₀）均可达到 250 以上，菌湿重可达到 300g/L 左右，发酵液中蛋白浓度均大于 10g/L（UV 定量）。结果见图 11 和图 12，图 11 可见，171801 的工程菌生长良好，诱导 48h 时菌浓度（OD₆₀₀）可达到 270，菌湿重可达到 355g/L，发酵液中蛋白浓度可达到 12g/L（UV 定量）。图 12 可见，

171802 的工程菌生长良好，诱导 48h 时菌浓度 (OD600) 可达到 250，菌湿重可达到 293g/L，发酵液中蛋白浓度可达到 11g/L (UV 定量)。

(2) 纯化

本实施例中用到的缓冲液如下：

缓冲液 A：20mM KH_2PO_4 、pH 值为 4.0；

缓冲液 B：20mM KH_2PO_4 、1M NaCl，pH 值为 4.0。

收集得到的工程菌的发酵液，使用陶瓷膜分离系统（江苏久吾高科技股份有限公司，JWCMP-9）分离菌体和发酵上清。以缓冲液 A 平衡阳离子交换介质（层析填料为苏州纳微产 UniGel-80sp，装载于江苏汉邦科技产全自动层析系统）至 A215 吸光值和电导率值都保持不变后，设置上样流速为 100us/cm，上样体积 20L/次，检测紫外 A215 吸光值，当其上升时，开始接样。待上样结束后，关闭接样，再以缓冲液 A 平衡阳离子层析介质，当 A215 吸光值下降时，打开接样，直至紫外和电导降至最低且不再变化，停止接样。分别用 500mM 和 1M 的氯化钠洗脱，收集洗脱液，分别检测 500mM 洗脱液和 1M 洗脱液组分中纯化蛋白状况，确定好组份后，进行透析（透析液为超纯水），随后浓缩、冷冻干燥，收集冻干胶原蛋白海绵。

取 170801、170802 胶原蛋白冻干海绵用超纯水溶解至 2mg/mL，0.22 μm 滤膜过滤，进样 10 μL （Sepax Bio-C18 色谱柱，高效液相色谱仪为 Waters2695 或 Agilent LC1260），以分析纯化后胶原蛋白 170801、170802 的纯度。

纯化结果如图 13、图 14 所示，图中可见，纯化的 170801、170802 胶原蛋白，单峰明显，纯度高（91.2%、90%），说明本纯化方法可靠有效，得到的胶原蛋白纯度高。

除本实施例所用纯化方法外，因 170801、170802 胶原蛋白氨基端有 Srtep-TagII 标签，羧基端有 6xHis 标签，均可使用 Ni-NTA、Strep-Tactin 等亲和介质进行单种或双亲和纯化，也能达到与本实例类似的效果。

(3) 重组胶原蛋白的性质表征

分子量大小测定：

得到的 170801、170802 蛋白理论预测的分子量分别为 23.8kDa、45.4kDa，而 SDS-PAGE 上检测的表现分子量比理论预测值要大。对重组人 XVII 型胶原蛋白 170801、170802 蛋白纯化后冻干品进行基质辅助激光解吸电离飞行时间质谱 (AB SCIEX 5800 MALDI-TOF/TOF) 检测其相对分子量，判断其蛋白的实际分子量，委托生工生物工程（上海）股份有限公司检测。结果如图 15、图 16 所示，MALDI-TOF MS（电荷数+1）检测到的 170801 胶原蛋白的分子量为 23811.6797Da（理论值为 23800Da）、170802 胶原蛋白的分子量为 45689.5781Da（理

论值为 45400Da)，由于毕赤酵母表达后存在翻译后修饰及分子量实验测定存在的系统误差，实际表达的胶原蛋白分子量与理论值一致。

傅里叶变换红外光谱(FT-IR) 分析:

胶原蛋白基团的特征吸收峰可以被红外光谱分析检测，试验取微量 170801、170802 胶原蛋白纯化后冻干样品分别 KBr 研磨成粉后压片，室温下，在 4000~ 400 cm^{-1} 范围内扫描 (Thermo Scientific, Nicolet™ iS™ 10 FT-IR 光谱仪)，方法及结果分析参照 (Jeong, H., J. Venkatesan and S. Kim, Isolation and characterization of collagen from marine fish (Thunnus obesus). Biotechnology and Bioprocess Engineering, 2013. 18(6): p. 1185-1191.)

从 170801、170802 胶原蛋白纯化样品的红外光谱扫描图 17 可以看到其酰胺 A、酰胺 B、酰胺 I、酰胺 II、酰胺 III 的波数均符合重组胶原蛋白结构特征 (参见文献[1].陈静涛等, 重组胶原蛋白与牛源I型胶原蛋白红外光谱研究. 材料导报, 2008(03): 第 119-121 页. [2].Doyle, B.B., E.G. Bendit and E.R. Blout, Infrared spectroscopy of collagen and collagen-like polypeptides. Biopolymers, 1975. 14(5): p. 937-957. [3].周爱梅等, 重组人源胶原蛋白的分离纯化及其结构表征. 食品与发酵工业, 2015(03): 第 46-52 页.)。

胶原蛋白冻干海绵样品扫描电镜扫描分析:

使用扫描电镜 (日立 TM3030PLUS) 对 170801、170802 胶原蛋白冻干海绵样品表面进行扫描，结果如图 18 所示，图中可见，重组人 XVII 型胶原蛋白 170801、170802 均具有明显片层状结构，与人体中人 XVII 型胶原蛋白在基底膜上层状分布类似，这样的结构预示其具有应用于生物医学材料领域的潜力。

实施例 3. 重组胶原蛋白细胞黏附活性检测

重组胶原蛋白的细胞黏附活检测方法参考文献 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)。委托常州大学药学院功能纳米材料与生物医学检测实验室完成。

具体实施方法:

正常培养 NIH/3T3 细胞 (购自中国科学院细胞库，货号 GNM6，培养、传代方法参照细胞说明书执行)。取 170801、170802 蛋白纯化冻干品、对照品为天然人胶原蛋白 (购自 Sigma，货号 C7774) 及牛血清白蛋白 (BSA，购自生工生物 (上海) 股份有限公司) 溶解 (超纯水或 1M HCl 溶液)，以 UV 蛋白定量经验公式： $C(\text{mg/mL}) = 0.144 * (A_{215} - A_{225})$ 测定蛋白浓度，再以 PBS (pH 7.4) 稀释至 0.5mg/mL。向 96 孔细胞培养板中加入 100 μL 各种蛋白溶液和空白

PBS 溶液对照，室温静置 60min；再向每孔中加入 10^5 个培养状态良好的 3T3 细胞， 37°C 、5% CO_2 孵育 60min。以 PBS 清洗 4 次孔中细胞。使用 LDH 检测试剂盒(Roche, 04744926001)检测 OD492nm 的吸光度值（具体操作参照说明书执行）。

OD492nm 的吸光度可相应的表征出胶原蛋白样品的细胞粘附活性：吸光度值越高，说明蛋白粘附的细胞越多，黏附活性越高，胶原蛋白越能在短时间内帮助细胞贴壁或粘附于细胞外基质之上，更利于构建更佳的细胞外环境。结果如图 19 结果所示，重组人 XVII 型胶原蛋白 170801、170802 与商品化的天然人胶原蛋白相比，具有更好的细胞粘附活性。

实施例 4. 划痕法检测重组胶原蛋白细胞迁移活性

重组胶原蛋白的细胞迁移活性检测及分析方法参考文献 Bobadilla, A., et al., In vitro cell migration quantification method for scratch assays. J R Soc Interface, 2019. 16(151): p. 20180709. 委托常州大学药学院功能纳米材料与生物医学检测实验室完成。

具体实施方法：

取 170801、170802 胶原蛋白纯化冻干品、对照品为天然人胶原蛋白（购自 Sigma，货号 C7774）及牛血清白蛋白（BSA，购自生工生物（上海）股份有限公司）溶解（超纯水或 1M HCl 溶液），以 UV 蛋白定量经验公式： $C(\text{mg/mL}) = 0.144 \times (A_{215} - A_{225})$ 测定蛋白浓度，再以 DMEM 无血清培养（GIBCO, 货号 12800017, pH 7.4）稀释至 0.5mg/mL（稀释后，调节其 pH 稳定于 7.0~7.4）。正常培养、传代 NIH/3T3 细胞（购自中国科学院细胞库，货号 GNM6，培养、传代方法参照细胞说明书执行）。将状态良好的细胞接入 6 孔板，每孔按照 2 万细胞/mL 密度接种 2mL 细胞悬液，培养 36h。用 200 μL 枪头制备划痕，用 PBS 清洗细胞 3 次，去除划下的细胞。于孔中加入 DMEM 无血清培养基稀释的蛋白溶液，继续放入 37°C 、5% CO_2 培养箱培养，0h、24h、48h 取样、拍照。使用 Image J 软件对细胞迁移的图片进行处理，获得初始划痕面积和无细胞空白区域面积数据，计算迁移率，迁移率 = $(1 - \text{无细胞空白区域面积} / \text{初始划痕面积}) \times 100\%$ 。

体外细胞迁移实验在一定程度上模拟了体内细胞迁移的过程，直接反映了细胞与胞外基质及基质影响下细胞之间的相互作用。细胞迁移活性是更有效表征胶原蛋白生物学活性的指标，迁移率越高，速度越快，说明胶原蛋白的生物学活性越佳。结果如图 20 所示，不同时间下拍摄的细胞迁移实际对比图（两黑色线内为初始及细胞迁移后划痕伤口区域）、及图 21 所示的计算出的细胞迁移率（Image J 计算无细胞空白区域）的比较可知，重组人 XVII 型胶原蛋白 170801、170802 均有比商品化的天然人胶原蛋白更优的细胞迁移活性。

实施例 5. 170801、170802 重组胶原蛋白凝胶制备与兔耳瘢痕模型实验

取羧甲基纤维素钠、分别取 170801、170802 胶原蛋白纯化冻干品溶解于超纯水中，得到

浓度为 0.01%羧甲基纤维素钠、5%胶原蛋白，将 0.01%羧甲基纤维素钠和 5%胶原蛋白混匀后配制为胶原蛋白凝胶，不含胶原蛋白的羧甲基纤维素钠凝胶作为阴性对照。

新西兰白化兔（中国人民解放军海军军医大学实验动物中心购进），体重为 2-3kg。新西兰白化兔在通风良好的房间内饲养，室温 21℃，12 小时昼夜循环，单笼饲养，在实验前饲养 2 周适应环境。随后进行瘢痕模型构建：实验动物接受 30g/L 戊巴比妥钠 1.0mL/kg (30mg/kg) 作耳缘静脉麻醉。确定麻醉后，以酒精擦洗消毒皮肤 3 遍。在兔耳内侧以直径 1cm 皮肤活检器划出整齐皮肤切缘，眼科剪剪出约直径 1cm 的全层皮肤缺损创面，暴露至耳软骨骨面。以无菌凡士林油纱覆盖创面，外面覆盖无菌纱布，包扎固定。每 2-3 天以碘伏对创面消毒清创，并更换创面敷料。

第 21 天创面愈合后分组进行瘢痕干预处理，分别使用 170801 胶原蛋白凝胶、170802 胶原蛋白凝胶涂抹创面，配合含碘抗菌手术贴膜（Drape Antimicrob®, Minnesota Mining and Manufacturing Company, USA）、自粘性硅酮胶片敷料（Cica-Care®, Smith & Nephew, England）进行瘢痕干预处理。

使用数码照相机及皮肤镜记录 1-6 周瘢痕干预实验过程中的瘢痕大体，第 6 周予麻醉后予耳缘静脉注射过饱和氯化钾溶液处死，取耳部瘢痕样本，浸泡于固定液，予 HE 染色、MASSON 染色，显微镜观察、拍照，具体方法可参照：[1]刘彤. 重组人胶原蛋白水凝胶和人工真皮分别作为微粒皮移植覆盖物修复全层皮肤缺损的实验研究[D].中国人民解放军海军军医大学,2019.[2] 李继承主编. 组织学与胚胎学实验指导 [M]. 北京：人民卫生出版社, 2018.09.。

实验结果如图 22 所示，170801、170802 胶原蛋白凝胶干预的兔耳皮肤伤口愈合后瘢痕恢复好，兔耳内侧有明显新生毛发长出，表明有明显的具有生物学功能的毛囊形成并有毛发生长，从组织切片的 Masson 染色、HE 染色结果看，其组织中胶原蛋白纤维排列有序，已开始形成真皮组织，尤其是存在新生毛囊结构，并明确的毛发的生长（图中箭头指的圈内），这是有效真皮组织恢复和具有生物学功能（生成毛发）的毛囊新生的明确标志；而对照组中，预后瘢痕仍为大量规则致密胶原沉积，类似于瘢痕结构，没有出现毛囊结构，无真皮组织形成，也没有任何恢复生长的兔毛出现。实验结果表明，重组人 XVII 型胶原蛋白 170801、170802 有优异组织修复和促进毛囊修复再生的生物活性。

权 利 要 求 书

1. 一种重组人 XVII 型胶原蛋白，所述胶原蛋白的氨基酸序列选自 SEQ ID NO. 2 或 SEQ ID NO. 3 所示的氨基酸序列。
2. 根据权利要求 1 所述的胶原蛋白，其特征在于，所述胶原蛋白的肽链氨基端与羧基端设计添加蛋白质标签，所述氨基端添加 Strep-Tag II 标签、羧基端添加 6×His Tag 标签。
3. 根据权利要求 1 所述的胶原蛋白，其特征在于，所述胶原蛋白具有细胞粘附活性、促细胞迁移活性、促进组织再生和促进毛囊修复再生的活性。
4. 编码权利要求 1 所述胶原蛋白的多核苷酸。
5. 根据权利要求 4 所述的多核苷酸，其特征在于，所述多核苷酸的核苷酸序列选自 SEQ ID NO. 6 或 SEQ ID NO. 7 所示的核苷酸序列。
6. 一种重组表达载体，含有权利要求 4 或 5 所述的多核苷酸。
7. 一种工程菌，所述工程菌含有权利要求 6 所述的重组表达载体。
8. 根据权利要求 7 所述的工程菌，其特征在于，所述工程菌为毕赤酵母。
9. 一种制备权利要求 1 所述胶原蛋白的方法，其特征在于，包括如下步骤：
 - (1) 工程菌的培养和筛选：将构建的工程菌经过筛选后，筛选表达量高的工程菌；
 - (2) 规模化高密度发酵、培养和蛋白质纯化：采用发酵罐进行发酵，对发酵上清液进行纯化，获取高纯度胶原蛋白。
10. 根据权利要求 9 所述的方法，其特征在于，步骤 (1) 所述筛选的表达量高的工程菌为巴斯德毕赤酵母 (*Pichia pastoris*)，保藏编号为 CGMCC No. 20626 或 CGMCC No. 20627。
11. 一种组合物，所述组合物包含权利要求 1 所述胶原蛋白或权利要求 9 所述方法制备的胶原蛋白。
12. 一种制品，所述制品包含权利要求 1 所述胶原蛋白或权利要求 9 所述方法制备的胶原蛋白或权利要求 11 所述组合物，所述制品包括药物、医疗器材、生物材料、组织工程产品、化妆品或保健品。
13. 权利要求 1 所述胶原蛋白、权利要求 4 或 5 所述多核苷酸、权利要求 6 所述重组表达载体、权利要求 7 所述工程菌或权利要求 11 所述组合物在制备制成品中的用途，所述制成品包括药物、医疗器材、生物材料、组织工程产品、化妆品

或保健品。

14. 权利要求 1 所述胶原蛋白、权利要求 4 或 5 所述多核苷酸、权利要求 6 所述重组表达载体、权利要求 7 所述工程菌或权利要求 11 所述组合物在制备促进瘢痕愈合、组织修复或促进毛囊修复的产品中的用途。

15. 根据权利要求 14 所述的用途，其特征在于，所述产品为外用制剂。

16. 根据权利要求 15 所述的用途，其特征在于，所述产品为外用涂抹制剂。

17. 根据权利要求 16 所述的用途，其特征在于，所述产品为外用涂抹凝胶剂。

18. 一种细胞，所述细胞含有权利要求 6 所述的重组表达载体。

19. 根据权利要求 18 所述的细胞，其特征在于，所述细胞为毕赤酵母。