

重组胶原蛋白、制备方法及其应用

技术领域

本发明涉及重组胶原蛋白、制备方法及其应用，尤其涉及重组表达的全长的胶原蛋白 $\alpha 1$ 链、制备方法及其应用，属于胶原蛋白表达技术领域。

背景技术

I、II 型胶原蛋白是人体内中典型的成纤维胶原蛋白，均由 3 条 α 肽链构成，每条 α 肽链均包含由氨基端肽区、特征性 (G-X-Y)_n 三联重复序列区、羧基端肽区三部分。

I 型胶原蛋白由两条 $\alpha 1$ 链和一条 $\alpha 2$ 链构成，是人体内所含的各种胶原蛋白中最丰富的一种，存在于肌肉、皮肤、动脉壁、纤维软骨中。II 型胶原蛋白由三条 $\alpha 1$ 链构成，主要分布在软骨组织、玻璃体、眼角膜中，占成人软骨基质胶原蛋白总量的 90% 以上，是软骨和骨型形成、骨骼生长和成熟软骨维持等所必需的成份。

作为一种重要的天然生物蛋白，胶原蛋白有良好的生物相容性、生物活性和可降解性等独特功能特征，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：加工过程降解严重，使其生物活性丧失；提取的胶原肽长度不等、性质不均、质量不稳定且有疯牛病、口蹄疫等病毒感染的安全隐患；同时动物源与人的胶原蛋白的氨基酸序列差别较大，属于异源性蛋白，会导致免疫排斥和过敏症状。

基因工程技术生产重组胶原蛋白则可有效避免这些缺陷。现有重组胶原蛋白的表达方法中，哺乳动物细胞表达系统、昆虫细胞（杆状病毒）表达系统、转基因动植物等表达系统成本高、产量低、周期长，多用于科研阶实验；大规模工业化生产中主要由原核（大肠杆菌）表达系统、毕赤酵母表达系统来表达人胶原蛋白。大肠杆菌中没有蛋白质的翻译后修饰，大规模表达为胞内表达，需裂解菌体，产生的大量杂质宿主蛋白和天然带有的（细胞壁成分）内毒素、肽聚糖，均需经复杂纯化工艺方能去除。但对于毕赤酵母来说，人胶原蛋白毕竟是一种外源蛋白，表达时会占用较多的细胞内资源（其所依赖的甲醇代谢途径最多可表达细胞可溶性蛋白 30% 的蛋白），细胞内会针对外源性的蛋白质进行相应调节，典型的状况是重组蛋白会出现严重的降解，人 I、II 型胶原蛋白的 $\alpha 1$ 链均为 1000 个氨基酸以上的长肽链，更易产生降解。

成熟的人 I、II 型胶原蛋白的 $\alpha 1$ 链序列均包含氨基端肽、三螺旋区域、羧基端肽三部分，人 I 型胶原蛋白 $\alpha 1$ 链（后文中按规范称之为 $\alpha 1(I)$ ）全长 1057AA，人 II 型胶原蛋白 $\alpha 1$ 链（后文中按规范称之为 $\alpha 1(II)$ ）全长 1060AA。以毕赤酵母表达人 $\alpha 1(I)$ 链的研究、专利较多，

表达人 $\alpha 1(\text{II})$ 链的研究、专利较少。现有毕赤酵母表达全长 $\alpha 1(\text{I})$ 链、全长 $\alpha 1(\text{II})$ 链的成果中，大部分研究中均只表达了 $\alpha 1(\text{I})$ 链的部分序列，而非成熟的全长 $\alpha 1(\text{I})$ 链序列。有一些公开的成果中虽然表达了全长的 $\alpha 1(\text{I})$ 链，但均会于表达时产生与目的产物占比基本相同的主降解产物，于 SDS-PAGE 电泳上表现为一条几乎与全长 $\alpha 1$ 链目的条带（目的产物）占比量基本相当的主降解条带（主降解产物）。这样的降解不但降低表达生产的全长 $\alpha 1$ 链的产量，而且因其与全长 $\alpha 1$ 链的相关性质接近，想获得高纯度的单一全长 $\alpha 1$ 链产品还需要进行两步双亲和纯化方能获得，增加了纯化复杂程度，相应的提高了纯化成本。所以保持全长肽链完整、减少降解，同时保持胶原蛋白生物学活性不发生改变是以毕赤酵母生产重组胶原蛋白面临的重要挑战。

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，提供毕赤酵母重组表达的全长的胶原蛋白 $\alpha 1$ 链、制备方法及其应用。本发明中的重组人 I 型胶原蛋白 $\alpha 1$ 链的变体（记为 $\alpha 1(\text{I})\text{M1}$ ）和重组人 II 型胶原蛋白 $\alpha 1$ 链的变体（记为 $\alpha 1(\text{II})\text{M6}$ ），相对于天然全长的 $\alpha 1(\text{I})$ 链与 $\alpha 1(\text{II})$ 链，在毕赤酵母中表达时均消除了几乎与全长 $\alpha 1$ 链目的条带（目的产物）占比量基本相同的主降解条带（主降解产物），提高了目的产物的产率；与毕赤酵母表达的天然全长的 $\alpha 1(\text{I})$ 链胶原蛋白与 $\alpha 1(\text{II})$ 链胶原蛋白及商品化人胶原蛋白有类似的理化特征和相同的生物学活性，均有应用于生物医学材料领域的价值。

为实现上述目的，本发明采用了以下技术方案：

本发明提供了重组胶原蛋白 $\alpha 1$ 链，所述重组胶原蛋白 $\alpha 1$ 链为 $\alpha 1(\text{I})\text{M1}$ 或 $\alpha 1(\text{II})\text{M6}$ ，所述 $\alpha 1(\text{I})\text{M1}$ 由人 I 型胶原蛋白 $\alpha 1$ 链的天然全长氨基酸序列经过氨基酸突变获得，所述 $\alpha 1(\text{II})\text{M6}$ 由人 II 型胶原蛋白 $\alpha 1$ 链天然全长氨基酸序列经过氨基酸突变获得。

优选的，所述 $\alpha 1(\text{I})\text{M1}$ 的氨基酸突变的位点数为 4；所述 $\alpha 1(\text{II})\text{M6}$ 的氨基酸突变位点数为 9。

优选的，所述人 I 型胶原蛋白 $\alpha 1$ 链为 SEQ.ID.NO.1 所示，所述氨基酸突变的位点为第 106 位的 M、109 位的 R、190 位的 M、193 位的 R，具体的，均改变为 P；

所述人 II 型胶原蛋白的 $\alpha 1$ 链为 SEQ.ID.NO.4 所示，所述氨基酸突变的位点为第 67 位的 V、68 位的 M、72 位的 M、75 位的 M、78 位的 R、108 位的 M、111 位的 R、162 位的 M、165 位的 R，具体的，均改变为 P。

进一步的，所述 $\alpha 1(\text{I})\text{M1}$ 的氨基酸序列如 SEQ.ID.NO.2 所示，所述 $\alpha 1(\text{II})\text{M6}$ 的氨基酸序列如 SEQ.ID.NO.5 所示。

本发明中，氨基酸序列突变的位点上相应氨基酸变更的种类发生一些变化，即突变的位

点相同，但变更的氨基酸种类不同，也会产生与本专利类似的技术效果；改变其中一处或几处进行突变处理，也有可能获得与本专利类似的技术效果。

本发明还提供编码所述重组胶原蛋白 $\alpha 1$ 链的核苷酸，所述编码重组胶原蛋白 $\alpha 1$ 链的核苷酸序列包括编码 $\alpha 1(I)M1$ 或 $\alpha 1(II)M6$ 的核苷酸序列。

进一步的，所述编码 $\alpha 1(I)M1$ 和核苷酸序列如 SEQ.ID.NO.3 所示，编码 $\alpha 1(II)M6$ 的核苷酸序列如 SEQ.ID.NO.6 所示。

本发明还提供了重组表达载体，含有编码所述重组胶原蛋白 $\alpha 1$ 链的核苷酸。

本发明还提供由上述重组表达载体构建的工程菌，所述工程菌含有所述重组表达载体或表达所述重组胶原蛋白 $\alpha 1$ 链。

所述工程菌的宿主菌优选为毕赤酵母，所述工程菌保藏日期为 2021 年 03 月 11 日，保藏编号为 CGMCC NO.21891 或 CGMCC NO. 21892，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。其中保藏编号为 CGMCC NO.21891 的工程菌表达的是重组 $\alpha 1(I)M1$ 胶原蛋白 $\alpha 1$ 链，保藏编号为 CGMCC NO. 21892 的工程菌表达的是重组 $\alpha 1(II)M6$ 胶原蛋白 $\alpha 1$ 链。

需要说明的是，本发明的宿主菌不限于毕赤酵母，只要根据本发明的方法，于毕赤酵母或其它种类的酵母菌中均可分泌表达，理论上均可获得与本专利类似的技术效果。

本发明还提供所述重组表达载体或所述的工程菌在表达所述重组胶原蛋白 $\alpha 1$ 链中的应用。

本发明还提供了所述重组胶原蛋白 $\alpha 1$ 链的制备方法，包括：

(1) 合成编码重组胶原蛋白 $\alpha 1$ 链的核苷酸序列：

将天然胶原蛋白 $\alpha 1$ 链序列将其氨基酸序列位点上相应氨基酸进行改变，I 型胶原蛋白的 $\alpha 1$ 链改变 4 个氨基酸得到 $\alpha 1(I)M1$ ，II 型胶原蛋白的 $\alpha 1$ 链改变 9 个氨基酸得到 $\alpha 1(II)M6$ ；在序列的氨基端和羧基端添加亲和纯化标签，合成编码 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 的 DNA 序列，使其含有双特异性亲和纯化标记，这样便于以两种标签序列为基础进行免疫学抗体检测。

经检测表明， $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 分别消除了全长 $\alpha 1(I)$ 链、全长 $\alpha 1(II)$ 链于毕赤酵母中表达时所产生的一条几乎与全长 $\alpha 1$ 链目的条带（目的产物）占比量基本相同的主降解条带（主降解产物）。

$\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 相对天然原始序列，其突变的氨基酸位于特征性的 (G-X-Y)_n 三联重复区域，但均是位于 X、Y 上的氨基酸，并没有改变胶原蛋白 (G-X-Y)_n 三联重复的氨基酸序列结构特征；且仍保持与原胶原蛋白类似的理化特征和生物学活性。

(2) 构建重组表达载体:

将合成的 DNA 连接入表达载体 pPIC9K 中, 分别构建表达重组 $\alpha 1(I)M1$ 胶原蛋白的 pPIC9K-COL1A1M1 和表达重组 $\alpha 1(II)M6$ 胶原蛋白的 pPIC9K-COL2A1M6 两种重组表达载体。

(3) 构建重组工程菌株、诱导表达和菌株筛选:

以 Sac I 线性化重组表达载体, 电转入毕赤酵母感受态细胞, 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; 筛选表达量高的工程菌株。

所述筛选到的表达量高的工程菌为巴斯德毕赤酵母 *Pichia pastoris*, 保藏编号分别为 CGMCC NO.21891、CGMCC NO. 21892。

(4) 高密度发酵培养:

将经过蛋白质表达鉴定表达量高的工程菌, 采用发酵罐进行高密度发酵培养。

(5) 蛋白质纯化:

使用一次阳离子交换层析对发酵上清液进行纯化, 以冷冻干燥的方法获得纯度高的 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 胶原蛋白。

本发明中将得到的重组胶原蛋白进行蛋白性质表征和体外实验, 对获取的毕赤酵母表达的 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 进行分析验证, 本发明得到的蛋白符合重组胶原蛋白结构特征, 具有细胞粘附活性, 与商品化的人胶原蛋白基本一致, 更重要的是 2 种变体蛋白与未突变前的胶原蛋白的结构特征、细胞黏附活性是类似的或相同的。

本发明还提供一种组合物, 所述组合物包含所述重组胶原蛋白 $\alpha 1$ 链或上述方法制备的胶原蛋白 $\alpha 1$ 链。

本发明还提供一种制品, 所述制品包含所述重组胶原蛋白 $\alpha 1$ 链或上述方法制备的胶原蛋白 $\alpha 1$ 链或上述组合物。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。

进一步的, 所述制品包括为细胞提供黏附、支撑、生长迁移空间的材料或作为输送营养物质与新陈代谢产物通道的材料。

进一步的, 所述制品为胶原蛋白水凝胶。

本发明还提供所述重组胶原蛋白 $\alpha 1$ 链、核苷酸、重组表达载体、工程菌、组合物等在制备成品中的用途, 包括但不限于药物、医疗器材、生物材料、组织工程产品、化妆品或保健品中的用途。

本发明还提供所述重组胶原蛋白 $\alpha 1$ 链、核苷酸、重组表达载体、工程菌或组合物在制备

促进创伤修复或组织再生的产品中的用途；进一步的，所述产品为胶原蛋白水凝胶。

本发明的有益效果：

(1) 本发明中胶原蛋白 $\alpha 1$ 链变体，改变的氨基酸位点于天然原始序列中占比极少（突变的氨基酸占比低于 1%，突变前后氨基酸序列同源性均高于 99%）、得到完整的重组 $\alpha 1$ 链胶原蛋白的同时并没有改变原始蛋白本身的性质（理化特性、生物学活性）、并且制备成相关产品的时候与天然序列的重组蛋白有相同的性质、生物学活性等。

$\alpha 1(I)M1$ 与 $\alpha 1(I)$ 相比、 $\alpha 1(II)M6$ 与 $\alpha 1(II)$ 相比，有类似的理化特征和生物学活性，均有应用于生物医学材料领域的价值。本发明通过进行 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 、 $\alpha 1(I)$ 、 $\alpha 1(II)$ 的细胞粘附实验发现， $\alpha 1(I)$ 与 $\alpha 1(I)M1$ 相比、 $\alpha 1(II)M6$ 与 $\alpha 1(II)$ 相比，细胞粘附活性无明显差别，且均与商品化的人胶原蛋白基本一致。使用毕赤酵母表达的 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 、 $\alpha 1(I)$ 、 $\alpha 1(II)$ 胶原蛋白制备胶原蛋白水凝胶，对水凝胶的液体力学特征进行检测， $\alpha 1(I)$ 与 $\alpha 1(I)M1$ 相比、 $\alpha 1(II)M6$ 与 $\alpha 1(II)$ 相比，其所制备的水凝胶在黏度、弹性模量、溶胀度上均无明显区别。使用扫描电镜扫描冻干后的胶原蛋白水凝胶，均为多孔网状结构，孔径范围集中于 100-200 μm ，有应用于生物医学材料领域的潜力。将四种胶原蛋白水凝胶与 NIH/3T3 细胞于体外共培养，加入钙黄绿素 AM 后，可检测到黏附、生长于水凝胶中的发绿色荧光的活细胞；加入 MTT 检测，均可观察到黏附、迁移生长至水凝胶内部的活细胞形成的蓝紫色结晶。

(2) 本发明通过 SDS-PAGE 电泳、Western Blot 对表达的蛋白质进行鉴定，结果表明，相对于全长的 $\alpha 1(I)$ 链与 $\alpha 1(II)$ 链， $\alpha 1(I)M1$ 与 $\alpha 1(II)M6$ 在毕赤酵母中表达时均消除了几乎与全长 $\alpha 1$ 链目的条带（目的产物）占比量基本相同的主降解条带（主降解产物），提高了目的产物的产率。并且，通过采用发酵罐进行高密度发酵实验，发酵产物经 SDS-PAGE 电泳检测发现， $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 可于高密度发酵条件下仍能保持目的条带的完整性，主降解条带依然不会产生，而同样高密度发酵条件下发酵生产的重组人 $\alpha 1(I)$ 、 $\alpha 1(II)$ 则会产生明显的主降解条带。

并且，本发明的重组胶原蛋白只需要使用一次阳离子交换层析对发酵上清液进行纯化，以冷冻干燥的方法获得纯度高的 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 胶原蛋白冻干海绵，经 SDS-PAGE 电泳检测，其主要是作为目的产物的单一条带，无主降解条带，获得了纯度高目的产物，降低了纯化的成本。而 $\alpha 1(I)$ 、 $\alpha 1(II)$ 同样以一步阳离子交换层析则只能获得目的条带（目的产物，全长 $\alpha 1$ 链）、主降解条带（主降解产物）混合的纯化蛋白，需要两步亲和层析方能获得纯度高的 $\alpha 1(I)$ 、 $\alpha 1(II)$ 胶原蛋白。

附图说明

图 1 为 $\alpha 1(I)M1$ 与 $\alpha 1(I)$ 氨基酸序列的差异图，图中灰色背景、加粗的氨基酸所示为差

异位点。

图 2 为 $\alpha 1(\text{II})\text{M6}$ 与 $\alpha 1(\text{II})$ 氨基酸序列的差异图，图中灰色背景、加粗的氨基酸所示为差异位点。

图 3 为 pPIC9K-COL1A1M1 载体图谱。

图 4 为 pPIC9K-COL2A1M6 载体图谱。

图 5 为 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白诱导表达 24h 的上清的 SDS-PAGE 图。

图 6 为 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白诱导表达 24h 的上清的 WB 图，图中，左边为抗 6×His Tag 抗体的 WB 图，右边为抗 Strep-Tag II 抗体的 WB 图。

图 7 为 $\alpha 1(\text{I})$ 、 $\alpha 1(\text{II})$ 胶原蛋白 SDS-PAGE 检测结果中目的条带、主降解条带质谱分析结果。

图 8 为 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白 SDS-PAGE 检测结果中目的条带质谱分析结果。

图 9 为 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 、 $\alpha 1(\text{I})$ 、 $\alpha 1(\text{II})$ 胶原蛋白诱导 48h 发酵上清的 SDS-PAGE 图。

图 10 为 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 、 $\alpha 1(\text{I})$ 、 $\alpha 1(\text{II})$ 纯化后冻干海绵的 SDS-PAGE 图。

图 11 为 $\alpha 1(\text{I})$ 、 $\alpha 1(\text{II})$ 胶原蛋白傅里叶变换红外光谱(FT-IR)分析图。

图 12 为 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白傅里叶变换红外光谱(FT-IR)分析图。

图 13 为 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 、 $\alpha 1(\text{I})$ 、 $\alpha 1(\text{II})$ 胶原蛋白细胞黏附活性检测。

图 14 为冻干 $\alpha 1(\text{I})$ 、 $\alpha 1(\text{I})\text{M1}$ 胶原蛋白水凝胶表面扫描电镜图（上两图 $\alpha 1(\text{I})$ 胶原蛋白水凝胶，下两图 $\alpha 1(\text{I})\text{M1}$ 胶原蛋白水凝胶）。

图 15 为冻干 $\alpha 1(\text{II})$ 、 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白水凝胶表面扫描电镜图（上两图 $\alpha 1(\text{II})$ 胶原蛋白水凝胶，下两图 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白水凝胶）。

图 16 为 $\alpha 1(\text{I})$ 、 $\alpha 1(\text{I})\text{M1}$ 胶原蛋白水凝胶中 NIH/3T3 细胞黏附生长结果图。

图中，上三图从左至右分别为： $\alpha 1(\text{I})$ 胶原蛋白水凝胶上黏附生长的 NIH/3T3 细胞（明场显微镜拍摄）、 $\alpha 1(\text{I})$ 胶原蛋白水凝胶上黏附生长的 NIH/3T3 细胞（钙黄绿素 AM 染色，明亮部分为发绿色荧光的细胞，荧光显微镜拍摄）、 $\alpha 1(\text{I})$ 胶原蛋白水凝胶中生长的 NIH/3T3 细胞形成的蓝紫色结晶（MTT 染色，黑色部分为结晶，明场显微镜拍摄）；

下三图从左至右分别为： $\alpha 1(\text{I})\text{M1}$ 胶原蛋白水凝胶上黏附生长的 NIH/3T3 细胞（明场显微镜拍摄）、 $\alpha 1(\text{I})\text{M1}$ 胶原蛋白水凝胶上黏附生长的 NIH/3T3 细胞（钙黄绿素 AM 染色，明亮部分为发绿色荧光的细胞，荧光显微镜拍摄）、 $\alpha 1(\text{I})\text{M1}$ 胶原蛋白水凝胶中生长的 NIH/3T3 细胞形成的蓝紫色结晶（MTT 染色，黑色部分为结晶，明场显微镜拍摄）。

图 17 为 $\alpha 1(\text{II})$ 、 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白水凝胶中 NIH/3T3 细胞黏附生长结果图。

图中, 上三图从左至右分别为: $\alpha 1(\text{II})$ 胶原蛋白水凝胶上黏附生长的 NIH/3T3 细胞 (明场显微镜拍摄)、 $\alpha 1(\text{II})$ 胶原蛋白水凝胶上黏附生长的 NIH/3T3 细胞 (钙黄绿素 AM 染色, 明亮部分为发绿色荧光的细胞, 荧光显微镜拍摄)、 $\alpha 1(\text{II})$ 胶原蛋白水凝胶中生长的 NIH/3T3 细胞形成的蓝紫色结晶 (MTT 染色, 黑色部分为结晶, 明场显微镜拍摄);

下三图从左至右分别为: $\alpha 1(\text{II})\text{M6}$ 胶原蛋白水凝胶上黏附生长的 NIH/3T3 细胞 (明场显微镜拍摄)、 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白水凝胶上黏附生长的 NIH/3T3 细胞 (钙黄绿素 AM 染色, 明亮部分为发绿色荧光的细胞, 荧光显微镜拍摄)、 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白水凝胶中生长的 NIH/3T3 细胞形成的蓝紫色结晶 (MTT 染色, 黑色部分为结晶, 明场显微镜拍摄)。

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. 氨基酸序列的设计和合成

人 I 型胶原蛋白 $\alpha 1$ 链 (记为 $\alpha 1(\text{I})$) 的氨基酸序列参考 Uniprot 数据库 P02452-1 (<https://www.uniprot.org/uniprot/P02452>) 序列中第 162 – 1218 部分 (PRO_0000005720), 是成熟形态人 I 型胶原蛋白 $\alpha 1$ 链氨基酸序列, 不含信号肽、C 端前肽、N 端前肽等 $\alpha 1(\text{I})$ 前体蛋白中会加工脱落的部分, 其序列如 SEQ.ID.NO.1 所示。

SEQ.ID.NO.1:

QLSYGYDEKSTGGISVPGPMGPSGRGLPGPPGAPGPQGFQGPPEPGEASGPMGPR
GPPGPPGKNGDDGEAGKPGRPGERGPQGPQGARGLPGLPMKGHRGFSGLDGAAGD
AGPAGPKGEPGSPGENGAPGQMGPRLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGP
AGPPGFPGAVGAKGEAGPQGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGA
NGAPGIAGAPGFPGARGPSGPQGPGGPPGPKGNSGEPGAPGSKGDTGAKGEPGPVGVQGP
GPAGEEGKRGARGEPGPTGLPGPPGERGGPSRGFPADGVAGPKGPAGERGSPGPAGPKG
SPGEAGRPGEAGLPAGKGLTGSPGSPGPDGKTGPPGPAGQDGRPGPPPGARGQAGVMG
FPGPKGAAGEPGKAGERGVPPPGAVGPAGKDGEAGAQQPPGPAGPAGERGEQGPAGSPG
FQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQPPGPAGPRGANG
APGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGADGSPGK
DGVRLTGPIGPPGPAGAPGDKGESGPSGPAGPTGARGAPGDRGEPGPPGPAGFAGPPGAD
GQPGAKGEPGDAGAKGDAGPPGPAGAPPPGPIGNVGAPGAKGARGSGAGPPGATGFPGAA

GRVGPPGPSGNAGPPGPPGAPAGKEGGKGRGETGPAGRPGEVGP GPPGPAGEKGSPGADG
PAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGKQGPPSGASGERGPPGPMGP
PGLAGPPGESGREGAPGAEGSPGRDGSPGAKGDRGETGPAGPPGAPGAPGAPGVPAGKS
GDRGETGPAGPAGVPVGARGPAGPQGPRGDKGETGEQGDRIKGHRGFSGLQGPPGPP
GSPGEQGPSGASGPAGPRGPPGSAGAPGKDGLNLPGPIGPPGPRGRTGDAGVP GPPGPPGP
PGPPGPPSAGFDFSFLPQPPQEKAHDGGRYRA

经过长期的实验研究，获得了重组人 I 型胶原蛋白 $\alpha 1$ 链的变体，记为 $\alpha 1(I)M1$ 。
 $\alpha 1(I)M1$ 相对于 $\alpha 1(I)$ 氨基酸序列中改变 4 个氨基酸，将这 4 个氨基酸突变为脯氨酸（Pro，
简称为 P），在 SEQ.ID.NO.1 所示的氨基酸序列第 106 位的 M、109 位的 R、190 位的 M、193
位的 R，均改变为 P；其余部分的氨基酸序列不变。 $\alpha 1(I)M1$ 与 $\alpha 1(I)$ 的同源性为 99.6%。

改变后的氨基酸序列（ $\alpha 1(I)M1$ ）全长 1057AA，序列如 SEQ.ID.NO.2 所示。

SEQ.ID.NO.2:

QLSYGYDEKSTGGISVPGPMGPSGRGLPGPPGAPGPQGFQGPPGEPGEPGASGPPGPP
GPPGPPGKNGDDGEAGKPGRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGA KGD
AGPAGPKGEPGSPGENGAPGQPGPPGLPGERGRPGAPGAPAGARGNDGATGAAGPPGPTGPA
GPPGFPGAVGAKGEAGPQGPRGSEGPQGV RGEPPGPAGAAGPAGNPGADGQPGAKGAN
GAPGIAGAPGFPGARGPSGPQPGGPPGPKGNSGEPGAPGSKGDTGAKGEPGPVG VQGPPG
PAGEEGKRARGEPPGTGLPGPPGERGGPGSRGFPGADGVAGPKGPAGERGSPGPAGPKGS
PGEAGRPGEAGLP GAKGLTGSPGSPGPDGKTGPPGPAGQDGRPGPPGPPGARGQAGVMGF
PGPKGAAGEPGKAGERGVPGPPGAVGPAGKDGEAGAQGPPGPAGPAGERGEQGPAGSPGF
QGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQGP GPPGAPPRGANGA
PGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLP GPKGDRGDAGPKGADGSPGKD
GVRGLTGPIGPPGPAGAPGDKGESGPSGPAGPTGARGAPGDRGEPGPPGPAGFAGPPGADG
QPGAKGEPGDAGAKGDAGPPGPAGPAGPPGPIGNVGAPGAKGARGSA GPPGATGFPGAAG
RVGPPGPSGNAGPPGPPGPAGKEGGKGRGETGPAGRPGEVGP GPPGPAGEKGSPGADGP
AGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGKQGPPSGASGERGPPGPMGPP
GLAGPPGESGREGAPGAEGSPGRDGSPGAKGDRGETGPAGPPGAPGAPGAPGVPAGKS
GDRGETGPAGPAGVPVGARGPAGPQGPRGDKGETGEQGDRIKGHRGFSGLQGPPGPP
GSPGEQGPSGASGPAGPRGPPGSAGAPGKDGLNLPGPIGPPGPRGRTGDAGVP GPPGPPGP
PGPPGPPSAGFDFSFLPQPPQEKAHDGGRYRA

$\alpha 1(I)M1$ 与 $\alpha 1(I)$ 氨基酸序列的差异如图 1 中灰色背景、加粗的氨基酸所示。

编码 SEQ.ID.NO.2 所示的 $\alpha 1(I)M1$ 的基因（记为 COL1A1M1）DNA 序列如 SEQ.ID.NO.3
所示：

SEQ.ID.NO.3:

CAACTTAGTTATGGATACGATGAAAAATCCACAGGTGGAATCAGTGTCCTGGACCT
ATGGGTCCATCAGGTCCAAGAGGTTTACCAGGACCTCCAGGTGCCCCAGGTCCCCAGGG
ATTTC AAGGTCCACCAGGAGAGCCTGGTGAGCCAGGAGCTTCTGGTCCACCTGGTCCCC
CTGGACCACCTGGTCCTCCAGGAAAGAATGGAGATGATGGTGAAGCTGGA AAAACCTGG
AAGACCTGGAGAAAGAGGACCACCAGGACCCCAGGGTGCCAGAGGACTGCCAGGTAC
CGCAGGTCTGCCTGGAATGAAAGGTCATAGAGGATTTTCAGGATTAGACGGTGCAAAGG
GAGACGCTGGACCTGCAGGACCAAAGGGTGAGCCAGGAAGTCCAGGAGAGAATGGTG
CACCAGGACAGCCAGGTCCACCTGGACTGCCCCGGTGAAAGAGGTAGACCCGGAGCACC
AGGACCAGCAGGTGCAAGAGGAAATGATGGAGCTACAGGTGCTGCAGGACCCCCAGGT
CCAACAGGACCAGCCGGTCCTCCCGGTTTCCAGGTGCCGTTGGAGCAAAGGTGAAG
CTGGTCCACAGGGTCCAAGAGGTTCTGAAGGTCCACAGGGAGTTAGAGGAGAACCAGG
ACCCCCTGGACCAGCTGGTGACGACGAGGACCAGCTGGTAACCTGGTGCTGACGGTCAG
CCAGGTGCTAAGGGAGCAAATGGAGCACCAGGAATAGCTGGTGCCCCAGGATTTCCCG
GTGCTAGAGGTCCAAGTGGTCCACAAGGACCAGGAGGTCCACCCGGTCCCAAAGGAAA
CAGTGGAGAACCAGGTGCACCCGGTTCAAAGGGAGATACAGGAGCTAAAGGAGAGCC
CGGTCCAGTGGGTGTTTCAGGGACCACCCGGACCTGCTGGAGAGGAAGGTAAAAGAGGT
GCAAGAGGTGAGCCAGGACCAACAGGTCTGCCTGGTCCCCCTGGTGAAAGAGGTGGTC
CAGGTAGTAGAGGATTTCCAGGAGCTGATGGTGTTGCAGGACCAAAGGGACCCGCAGG
TGAGAGAGGATCACCCGGTCCAGCCGGACCAAAAGGATCACCAAGGAGAAGCTGGTAGA
CCAGGAGAAGCTGGTCTGCCAGGTGCTAAAGGATTGACAGGATCACCCGGTTACCTG
GTCCTGATGGAAAGACAGGACCTCCAGGTCCCGCTGGTCAGGACGGTAGACCAGGACC
CCCAGGACCCCCAGGTGCAAGAGGTCAGGCAGGTGTAATGGGTTTCCCGGACCTAAA
GGAGCAGCTGGAGAACCTGGTAAAGCTGGAGAGAGAGGAGTGCCTGGACCCCCTGGA
GCTGTTGGTCCAGCAGGAAAGGATGGTGAGGCAGGTGCACAAGGTCCACCTGGACCCG
CTGGACCTGCAGGTGAGAGAGGAGAGCAAGGTCCCGCAGGTTCTCCAGGTTTTCAGGG
TTTGCCAGGTCCAGCCGGTCCTCCTGGAGAGGCAGGAAAGCCAGGAGAACAAAGGAGTT
CCAGGAGACCTGGGTGCACCAGGACCCTCTGGTGCAAGAGGAGAGAGAGGATTTCTCTG
GAGAAAGAGGTGTGCAGGGACCACCAGGTCCCGCCGGTCCAAGAGGAGCAAATGGAG
CCCCTGGAAATGACGGAGCTAAGGGTGACGCTGGTGACCAAGGAGCACCAGGTTCTCA
AGGTGCTCCCGGATTGCAGGGTATGCCTGGAGAGAGAGGTGCAGCTGGACTGCCAGGT
CCAAAAGGTGACAGAGGAGACGCCGGTCCTAAGGGAGCTGACGGTTCTCCTGGAAAG
GACGGTGTGAGAGGTTTGACAGGACCAATAGGTCCACCCGGTCCTGCTGGAGCCCCTG
GAGACAAAGGTGAATCAGGTCCTTCCGGTCCAGCCGGACCAACAGGAGCAAGAGGAG
CACCTGGAGACAGAGGAGAGCCAGGTCTCCAGGACCTGCAGGTTTCGCTGGTCCTCC
CGGAGCAGATGGACAGCCAGGAGCTAAGGGAGAACCCGGTGACGCTGGTGCTAAGGG

AGATGCAGGTCCACCAGGTCCTGCTGGTCTGCTGGACCTCCCGGACCAATAGGTAATG
TTGGAGCACCCGGAGCAAAAGGTGCCAGAGGTTCCGCAGGTCCTCCCGGAGCAACTGG
TTTTCCAGGAGCTGCCGGAAGAGTGGGTCCACCTGGTCCTTCTGGAAATGCAGGACCA
CCAGGTCCTCCTGGTCCAGCCGGAAGGAAGGTGGAAAGGGACCTAGAGGAGAAACA
GGTCCCGCAGGTAGACCCGGTGAGGTGGGTCCACCTGGTCCACCCGGTCCAGCTGGTG
AGAAAGGAAGTCCTGGAGCAGACGGACAGCTGGTGCCCCTGGTACACCAGGACCCCA
AGGAATAGCTGGTCAAAGAGGTGTTGTTGGTTTACCAGGTCAGAGAGGAGAAAGAGGT
TTTCCAGGATTACCAGGTCCCTCAGGTGAGCCCGGAAAACAGGGTCCCTCAGGAGCAA
GTGGTGAAAGAGGACCACCAGGACCAATGGGACCTCCAGGATTAGCTGGTCCACCAGG
AGAATCAGGAAGAGAGGGTGCTCCTGGAGCAGAAGGTTTACCAGGAAGAGACGGTTC
ACCCGGAGCCAAGGGAGACAGAGGTGAAACAGGTCCCGCAGGTCCACCAGGAGCACC
CGGAGCCCCTGGTGCTCCAGGACCTGTTCGACCAGCAGGAAAATCCGGTGACAGAGGT
GAGACTGGACCCGCAGGTCCTGCTGGTCCTGTTGGACCAGTGGGTGCAAGAGGACCAG
CAGGTCCACAAGGTCCAAGAGGTGACAAAGGTGAGACAGGTGAGCAGGGTGACAGAG
GAATTAAAGGTCACAGAGGATTTTCAGGACTGCAGGGACCACCCGGTCCTCCCGGTTCC
CCAGGAGAGCAAGGTCCATCCGGTGCATCCGGTCCAGCTGGACCCAGAGGACCACCTG
GTTCTGCTGGTGACACCAGGTAAAGATGGATTGAACGGTTTGCCTGGTCCAATAGGACCT
CCTGGTCCAAGAGGAAGAACTGGTGACGCCGGTCCCGTTCGGACCACCCGGTCCACCAG
GTCCCCCAGGTCCACCCGGACCACCATCCGCAGGATTTGATTCTCATTCTCCTCAAC
CTCCTCAAGAGAAAGCACATGATGGAGGTAGATACTATAGAGCC

人 II 型胶原蛋白 $\alpha 1$ 链（记为 $\alpha 1(\text{II})$ ）的氨基酸序列参 Uniprot 数据库 P02458（<https://www.uniprot.org/uniprot/P02458>）序列中第 182 – 1241 部分（PRO_0000005730）是成熟形态人 II 型胶原蛋白 $\alpha 1$ 链氨基酸序列，不含信号肽、C 端前肽、N 端前肽等 $\alpha 1(\text{II})$ 前体蛋白中加工脱落的部分，其序列如 SEQ.ID.NO.4 所示：

SEQ.ID.NO.4:

QMAGGFDEKAGGAQLGVMQGPMPGMPGRGPPGPAGAPGPQGFQGNPGEPEGVVS
GPMGPRGPPGPPGKPGDDGEAGKPGKAGERGPPGPQGARGFPPTPLPGVKHGRGYPLD
GAKGEAGAPGVKGESGSPGENSGSPGPMGPRGLPGERGRTGPAGAAGARGNDGQPGPAGPP
GPVGPAGGPGFPGAPGAKGEAGPTGARGPEGAQGPGEPTPGSPGPAGASGNPGTDGIPG
AKGSAGAPGIAGAPGFPGPRGPPGPQGATGPLGPKGQTGEPGIAGFKGEQGPKEPAGP
QGAPGPAGEEGKRGARGEPGGVGPPIGPPGERGAPGNRGFPQDGLAGPKGAPGERGPSGL
AGPKGANGDPGRPGEPGLPGARGLTGRPGDAGPQKVGPSGAPGEDGRPGPPGPQGARGQ
PGVMGFPGPKGANGEPGKAGEKGLPGAPGLRGLPGKDGETGAAGPPGPAGPAGERGEQG
APGPSGFQGLPGPPGPPGEGGKPGDQGVPEAGAPGLVGPRGERGFPGERGSPPAQGLQGP

改变后的序列 (α1(II)M6) 全长 1060AA, 序列如 SEQ.ID.NO.5 所示:

Q MAGGFDEKAGGAQLGPPQGGPPGPPGPPGPAGAPGPQGFQGNPGEPGEPGVSGPP
 PPGPPPGPPGKPGDDGEAGKPGKAGERGPPGPQGARGFPGTPGLPGVKGHRGYPLDGA
 GEAGAPGVKGESGSPGENSGPPGPPGLPGERGRTGPAGAAGARGNDGQGPAGPPGPVG
 PAGGPGFPGAPGAKGEAGPTGARGPEGAQGPGRGEPGTPGSPGPAGASGNPGTDGIPGAKGS
 AGAPGIAGAPGFFGPRGPPGPQGATGPLGPKGQTGEPGIAGFKGEQGPKEGPGAPGQAP
 GPAGEEGKRGARGEPGGVGPIGPPGERGAPGNRGFPQDGLAGPKGAPGERGPSGLAGPK
 GANGDPGRPGEPGLPGARGLTGRPGDAGPQGKVGPSGAPGEDGRPGPPGPQGARGQPGV
 MGFFGPKGANGEPPGKAGEKGLPGAPGLRGLPGKDGETGAAGPPGPAGPAGERGEQGAPGP
 SGFQGLPGPPGPPGEGGKPGDQGVPGEAGAPGLVGPRGERGFPGERGSPPGAQGLQGPRGLP
 GTPGTDGPKGASGPAGPPGAQGPPLQGMPGERGAAGIAGPKGDRGDVGEKGPEGAPGK
 DGGRGLTGPIGPPGPAGANGEKGEVGPMPGAPSAGARGAPGERGETGPPGPAGFAGPPGAD
 GQPGAKGEQGEAGQKGDAGAPGPQGSPGAPGPQGPTGVTGPKGARGAQGPPGATGFPGA
 AGRVGPPGSNGNPGPPGPPGPSGKDGPKGARGDSGPPGRAGEPGLQGPA GPPGEKGEPGDD
 GPSGAEGPPGPQGLAGQRGIVGLPGQRGERGFPLGPSGEPGKQGAPGASGDRGPPGPVG
 PPGLTGAPGEPGREGSPGADGPPGRDGAAGVKGDRGETGAVGAPGAPGPPGSPGAPPTGK
 QGDRGEAGAQQGPMGPSGPAGARGIQGPQGPRGDKGEAGEPGERGLKGHRGFTGLQGLPG
 PPGPSGDQGASGPAGPSGPRGPPGPVGPSGKDGANGIPGPIGPPGPRGRSGETGPAGPPGNPG
 PPGPPGPPGPGIDMSAFAGLGPREGKPPDLOYMRA

$\alpha 1$ (II)M6 与 $\alpha 1$ (II)氨基酸序列的差异如图 2 中灰色背景和加粗的氨基酸所示。

编码 SEQ.ID.NO.5 所示的 $\alpha 1$ (II)M6 的基因(记为 COL2A1M6)DNA 序列如 SEQ.ID.NO.6 所示:

SEQ.ID.NO.6:

```
CAAATGGCTGGTGGATTTCGATGAAAAGGCTGGTGGAGCCCAATTAGGTCCTCCACA
AGGTCCTCCCGGTCCACCTGGTCCTCCCGGTCCAGGTCCCGCCGGTGCTCCTGGAC
CACAGGGTTTCCAAGGAAACCCCGGTGAACCAGGTGAGCCTGGTGTTCAGGTCCTCC
CGGTCTCCAGGACCACCTGGACCACCAGGAAAGCCTGGTGACGACGGAGAAGCTGGT
AAACCAGGAAAGGCAGGAGAGAGAGGTCCACCTGGACCTCAGGGTGCCAGAGGTTTC
CCAGGTACCCCTGGTCTTCCTGGTGTCAAGGGTCATAGAGGTTACCCCGGTTTGGATGGT
GCCAAGGGTGAAGCCGGTGCCCCTGGTGTTAAGGGTGAATCAGGAAGTCCCGGTGAAA
ATGGAAGTCCCGGTCCACCCGGTCCACCTGGACTGCCAGGTGAGAGAGGAAGAACCGG
ACCAGCTGGTGCTGCAGGTGCTAGAGGAAATGACGGACAGCCCGGACCAGCCGGACCT
CCCGGTCTGTTGGGCCCCGACAGGTGGTCCTGGTTTCCCtgGTGCTCCTGGAGCCAAAGG
AGAAGCCGGACCCACCGGAGCCAGAGGTCCCGAGGGAGCACAGGGACCTAGAGGAGA
ACCAGGTACACCAGGTAGTCCCGGTCTGCTGGTGCATCAGGAAATCCCGGAACTGACG
GTATTCCAGGAGCAAAGGGATCTGCAGGAGCACACAGGAATAGCTGGTGCTCCTGGATTT
CCAGGTCCCAGAGGACCTCCCGGTCTCAAGGAGCAACAGGTCCTTTGGGACCAAAAG
GTCAAACAGGAGAACCAGGTATTGCTGGATTCAAAGGAGAGCAAGGTCCAAAGGGAG
AGCCCGGTCCCGCAGGTCCCCAAGGAGCCCCAGGACCAGCTGGTGAAGAAGGAAAAA
GAGGAGCCAGAGGTGAACCTGGAGGAGTAGGACCTATTGGTCCTCCTGGTGAGAGAGG
TGCTCCCGGAAACAGAGGTTTTCTGGTCAAGATGGTCTGGCTGGACCTAAAGGTGCTC
CAGGAGAGAGAGGACCTTCAGGACTTGCTGGTCCAAAAGGTGCTAACGGAGATCCAGG
AAGACCCGGTGAACCTGGTCTGCCTGGAGCTAGAGGATTAACAGGAAGACCAGGTGAC
GCAGGTCCCCAGGGTAAAGTGGGTCCAGTGGTGCCCCAGGTGAAGATGGAAGACCTG
GTCCTCCCGGACCCCAAGGTGCAAGAGGTGAGCCTGGAGTGATGGGATTTCTGGACC
CAAGGGTGCTAACGGAGAACCTGGAAAAGCTGGTGAGAAAGGACTGCCCGGTGCCCC
AGGTCTTAGAGGTTTGCCAGGTAAAGATGGAGAAACAGGAGCCGCAGGACCACCCGGT
CCAGCCGGACCAGCAGGAGAGAGAGGTGAACAAGGAGCACCTGGTCCAAGTGGTTTT
CAGGGTCTTCCAGGTCCCCCTGGTCCACCAGGAGAGGGAGGTAAACCAGGTGACCAAG
GTGTCCCTGGAGAAGCAGGTGCACCCGGTCTTGTGGGTCCAAGAGGTGAAAGAGGATT
CCCTGGTGAGAGAGGATCTCCCGGAGCCCAGGGACTTCAAGGTCCTAGAGGTCTGCCA
GGTACCCCTGGTACAGACGGACCAAAGGGAGCATCAGGACCCGCTGGACCTCCCGGAG
CCCAAGGTCTCCAGGTTTACAAGGTATGCCTGGTGAAAGAGGTGCTGCAGGTATAGCT
GGACCAAAAGGAGACAGAGGTGACGTTGGTGAGAAGGGTCCCGAAGGAGCCCCTGGA
AAAGATGGTGGAAGAGGATTAACAGGTCCTATAGGACCACCCGGTCCAGCCGGTGCTAA
TGGAGAAAAAGGAGAAGTAGGTCTCCAGGTCCAGCAGGATCTGCAGGTGCTAGAGGT
GCCCCCTGGAGAGAGAGGTGAAACAGGACCACCTGGTCCAGCTGGTTTTCGCTGGTCCCC
CAGGAGCTGATGGACAGCCCGGTGCAAAAGGTGAACAAGGAGAAGCCGGACAGAAGG
GAGATGCTGGAGCCCCCGGTCCACAAGGTCCCTCAGGAGCACCAGGTCTCAAGGTCC
AACTGGTGTGACCGGGCCAAAGGGTGCAAGAGGAGCACAGGGACCTCCAGGAGCAAC
AGGTTTCCCAGGAGCTGCTGGTAGAGTCGGTCCACCCGGATCTAATGGTAACCCCGGAC
CACCAGGACCACCTGGACCATCTGGAAAGGATGGACCCAAAGGAGCAAGAGGAGATTC
AGGACCACCCGGAAGAGCAGGAGAACCTGGATTACAGGGTCCCGCCGGTCCACCAGGA
```

GAGAAAGGAGAGCCCGGAGATGATGGTCCCTCAGGTGCAGAGGGACCCCCAGGACCC
CAAGGTCTGGCAGGTCAAAGAGGTATAGTGGGTCTTCCAGGTCAAAGAGGTGAAAGAG
GATTTCCAGGACTTCCAGGTCCTTCAGGTGAACCCGGTAAACAGGGAGCCCCCGGAGC
CTCAGGTGACAGAGGTCTCCAGGACCAGTAGGACCCCCAGGTTTAACCGGACCAGCA
GGTGAGCCAGGAAGAGAAGGTTCTCCTGGAGCCGATGGACCTCCAGGAAGAGACGGT
GCAGCTGGTGTTAAGGGTGACAGAGGTGAAACTGGAGCCGTAGGAGCCCCAGGTGCCC
CCGACCACCCGGATCACCCGGACCTGCAGGTCCTACTGGTAAACAAGGAGATAGAGG
AGAAGCCGGTGCCCAGGGTCCTATGGGTCTTCTGGTCCTGCAGGAGCAAGAGGTATAC
AAGGTCCACAGGGTCCCAGAGGTGACAAGGGTGAAGCAGGAGAACCCGGTGAGAGAG
GTCTGAAGGGTCATAGAGGATTACCCGGTTACAGGGTTTGCCAGGACCCCCTGGACCA
AGTGGTGACCAGGGTGATCCGGTCCAGCAGGTCCTTCTGGACCAAGAGGTCTCTCCG
GTCCAGTTGGTCCATCAGGTAAAGACGGAGCCAACGGTATCCCAGGTCCCATCGGTCTCT
CCAGGTCCTAGAGGAAGAAGTGGAGAGACTGGTCCTGCTGGACCTCCTGGAAACCCTG
GTCCTCCAGGACCTCCAGGTCCTCCAGGTCCCGGAATAGATATGTCCGCTTTCGCTGGAT
TGGGACCAAGAGAGAAAGGTCCTGACCCTCTTCAATATATGAGAGCA

在编码 $\alpha 1(I)M1$ 的 DNA 序列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, $\alpha 1(I)M1$ 最终表达获得的是含有标签的蛋白, 共 1071 个氨基酸, 如 SEQ.ID.NO.7 所示:

SEQ.ID.NO.7:

WSHPQFEKQLSYGYDEKSTGGISVPGPMGPSGRPLPGPPGAPGPQGFQPPGEPGEPG
ASGPPGPPGPPGPPGKNGDDGEAGKPRPGERGPPGPQGARGLPGTAGLPGMKGHRGFSG
LDGAKGDAGPAGPKGEPGSPGENGAPGQPGPPGLPGERGRPGAPGPAGARGNDGATGAAG
PPGPTGPAGPPGFPGAVGAKGEAGPQGPRGSEGPQGVVRGEPGPPGPAGAAGPAGNPGADGQ
PGAANGAPGIAGAPGFPGARGSPGPGGPPGPKGNSGEPGAPGSKGDTGAKGEPGP
VGVQGPMPGAGEEGKRGARGEPGPTGLPGPPGERGGPGSRGFPADGVAGPKGPAGERGSP
GPAGPKGSPGEAGRPGEAGLPAGKGLTGSPGSPGPDGKTGPPGPAGQDGRPGPPGPPGARG
QAGVMGFPGPKGAAGEPGKAGERGVPPGPAGVPAGKDGEAGAQQPPGPAGPAGERGEQ
GPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGERGVQGPMPGAG
PRGANGAPGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGA
DGSPGKDGVRLTGPIGPPGPAGAPGDKGESGPSGPAGPTGARGAPGDRGEPGPPGPAGFA
GPPGADGQPGAKGEPGDAGAKGDAGPPGPAGPAGPPGPIGNVGPAGKARGSGAGPPGAT
GFPGAAGRVGPPGPSGNAGPPGPPGPAGKEGKGPRGETGPAGRPGEVGPMPGPPGPAGEKG
SPGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGKQGSPSGASGERGP
PGPMGPPGLAGPPGESGREGAPGAEGSPGRDGSPPGAKGDRGETGPAGPPGAPGAPGAPGPV
GPAGKSGDRGETGPAGPAGVPVGPARGPAGPQGPGRGDKGETGEQGDRIKGHRGFSGLQ
GPPGPPGSPGEQGPSGASGPAGPRGPPGSAGAPGKDGLNGLPGPIGPPGPRGRTGDAGVPVP
GPPGPPGPPGPPSAGFDFSFLPQPPQEKAHDGGRYYRAHHHHHH

经过优化设计后，编码 SEQ.ID.NO.7 氨基酸序列 ($\alpha 1(I)M1$) 的基因 (记为 COL1A1M1) 的 DNA 序列如 SEQ.ID.NO.8 所示：

SEQ.ID.NO.8:

TGGTCTCATCCACAATTTGAAAAGCAACTTAGTTATGGATACGATGAAAAATCCACA
GGTGAATCAGTGTTTCCTGGACCTATGGGTCCATCAGGTCCAAGAGGTTTACCAGGACC
TCCAGGTGCCCCAGGTCCCCAGGGATTTCAAGGTCCACCAGGAGAGCCTGGTGAGCCA
GGAGCTTCTGGTCCACCTGGTCCCCCTGGACCACCTGGTCCCTCCAGGAAAGAATGGAGA
TGATGGTGAAGCTGGAAAACCTGGAAGACCTGGAGAAAGAGGACCACCAGGACCCCA
GGGTGCCAGAGGACTGCCAGGTACCGCAGGTCTGCCTGGAATGAAAGGTCATAGAGGA
TTTTCAGGATTAGACGGTGCAAAGGGAGACGCTGGACCTGCAGGACCAAAGGGTGAGC
CAGGAAGTCCAGGAGAGAATGGTGCACCAGGACAGCCAGGTCCACCTGGACTGCCCCG
GTGAAAGAGGTAGACCCGGAGCACCAGGACCAGCAGGTGCAAGAGGAAATGATGGAG
CTACAGGTGCTGCAGGACCCCCAGGTCCAACAGGACCAGCCGGTCTCCCGGTTTCCC
AGGTGCCGTTGGAGCAAAAGGTGAAGCTGGTCCACAGGGTCCAAGAGGTTCTGAAGGT
CCACAGGGAGTTAGAGGAGAACCAGGACCCCCCTGGACCAGCTGGTGCAGCAGGACCA
GCTGGTAACCCTGGTGCTGACGGTCAGCCAGGTGCTAAGGGAGCAAATGGAGCACCAG
GAATAGCTGGTGCCCCAGGATTTCCCGGTGCTAGAGGTCCAAGTGGTCCACAAGGACCA
GGAGGTCCACCCGGTCCCAAAGGAAACAGTGGAGAACCAGGTGCACCCGGTTCAAAG
GGAGATACAGGAGCTAAAGGAGAGCCCGGTCCAGTGGGTGTTTCAGGGACCACCCGGAC
CTGCTGGAGAGGAAGGTAAAAGAGGTGCAAGAGGTGAGCCAGGACCAACAGGTCTGC
CTGGTCCCCCTGGTGAAAGAGGTGGTCCAGGTAGTAGAGGATTTCCAGGAGCTGATGGT
GTTGCAGGACCAAAGGGACCCGCAGGTGAGAGAGGATCACCCGGTCCAGCCGGACCA
AAAGGATCACCAGGAGAAGCTGGTAGACCAGGAGAAGCTGGTCTGCCAGGTGCTAAAG
GATTGACAGGATCACCCGGTTCACCTGGTCTGATGGAAAGACAGGACCTCCAGGTCCC
GCTGGTCAGGACGGTAGACCAGGACCCCCAGGACCCCCAGGTGCAAGAGGTGAGGCA
GGTGTAATGGGTTTCCCCGGACCTAAAGGAGCAGCTGGAGAACCTGGTAAAGCTGGAG
AGAGAGGAGTGCCTGGACCCCCCTGGAGCTGTTGGTCCAGCAGGAAAGGATGGTGAGGC
AGGTGCACAAGGTCCACCTGGACCCGCTGGACCTGCAGGTGAGAGAGGAGAGCAAGG
TCCCGCAGGTTCTCCAGGTTTTTCAGGGTTTGCCAGGTCCAGCCGGTCTCCTGGAGAGG
CAGGAAAGCCAGGAGACAAGGAGTTCCAGGAGACCTGGGTGCACCAGGACCCTCTG
GTGCAAGAGGAGAGAGAGGATTTCTGGAGAAAGAGGTGTGCAGGGACCACCAGGTG
CCGCCGGTCCAAGAGGAGCAAATGGAGCCCCCTGGAAATGACGGAGCTAAGGGTGACGC
TGGTGCACCAGGAGCACCAGGTCTCAAGGTGCTCCCGGATTGCAGGGTATGCCTGGAG
AGAGAGGTGCAGCTGGACTGCCAGGTCCAAAAGGTGACAGAGGAGACGCCGGTCTTA
AGGGAGCTGACGGTTCTCCTGGAAAGGACGGTGTGAGAGGTTTGACAGGACCAATAGG

TCCACCCGGTCCTGCTGGAGCCCCTGGAGACAAAGGTGAATCAGGTCCTTCCGGTCCA
GCCGGACCAACAGGAGCAAGAGGAGCACCTGGAGACAGAGGAGAGCCAGGTCCTCCA
GGACCTGCAGGTTTCGCTGGTCCTCCCGGAGCAGATGGACAGCCAGGAGCTAAGGGAG
AACCCGGTGACGCTGGTGCTAAGGGAGATGCAGGTCCACCAGGTCCTGCTGGTCCTGCT
GGACCTCCCGGACCAATAGGTAATGTTGGAGCACCCGGAGCAAAAGGTGCCAGAGGTT
CCGCAGGTCCTCCCGGAGCAACTGGTTTTCCAGGAGCTGCCGGAAGAGTGGGTCCACC
TGGTCCTTCTGGAAATGCAGGACCACCAGGTCCTCCTGGTCCAGCCGGAAGGAAGGT
GGAAAGGGACCTAGAGGAGAAACAGGTCCCGCAGGTAGACCCGGTGAGGTGGGTCCA
CCTGGTCCACCCGGTCCAGCTGGTGAGAAAGGAAGTCCTGGAGCAGACGGACCAGCTG
GTGCCCCTGGTACACCAGGACCCCAAGGAATAGCTGGTCAAAGAGGTGTTGTTGGTTTA
CCAGGTCAGAGAGGAGAAAGAGGTTTTCCAGGATTACCAGGTCCCTCAGGTGAGCCCG
GAAAACAGGGTCCCTCAGGAGCAAGTGGTGAAAGAGGACCACCAGGACCAATGGGAC
CTCCAGGATTAGCTGGTCCACCAGGAGAATCAGGAAGAGAGGGTGCTCCTGGAGCAGA
AGGTTACACAGGAAGAGACGGTTCACCCGGAGCCAAGGGAGACAGAGGTGAAACAGG
TCCCGCAGGTCCACCAGGAGCACCCGGAGCCCCTGGTGCTCCAGGACCTGTCGGACCA
GCAGGAAAATCCGGTGACAGAGGTGAGACTGGACCCGCAGGTCCTGCTGGTCCTGTTG
GACCAGTGGGTGCAAGAGGACCAGCAGGTCCACAAGGTCCAAGAGGTGACAAAGGTG
AGACAGGTGAGCAGGGTGACAGAGGAATTAAGGTACACAGAGGATTTTCAGGACTGCA
GGGACCACCCGGTCCTCCCGGTTCCCCAGGAGAGCAAGGTCCATCCGGTGCATCCGGTC
CAGCTGGACCCAGAGGACCACCTGGTTCTGCTGGTGACACCAGGTAAAGATGGATTGAA
CGGTTTGCCTGGTCCAATAGGACCTCCTGGTCCAAGAGGAAGAACTGGTGACGCCGGT
CCCGTCGGACACCCGGTCCACCAGGTCCCCAGGTCCACCCGGACCACCATCCGCAG
GATTTGATTTCTCATTCCTTCCTCAACCTCCTCAAGAGAAAGCACATGATGGAGGTAGAT
ACTATAGAGCCCATCACCACCATCATCATTA

在编码 $\alpha 1$ (II)M6 的 DNA 序列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, $\alpha 1$ (II)M6 最终表达获得的是含有标签的蛋白, 共 1076 个氨基酸, 序列如 SEQ.ID.NO.9 所示:

SEQ.ID.NO.9:

EFWSHPQFEKQMAGGFDEKAGGAQLGPPQGPPGPPGPPGPPGAPGPQGFQGNPGE
PGEPGVSGPPGPPGPPGPPGKPGDDGEAGKPKGAGERGPPGPQGARGFPGTPLPGVKGHR
GYPLDGAKEAGAPGVKGESGSPGENGSPGPPGPPGLPGERGRTGPAGAAGARGNDGQP
GPAGPPGPVGPAGGPGFPGAPGAKGEAGPTGARGPEGAQGPREGPTPGSPGPAGASGNPG
TDGIPGAKGSAGAPGIAGAPGFPGRGPPGPQGATGPLGPKGQTGEPGIAGFKGEQGPKEP
GPAGPQGAPGPAGEEGKRGARGEPGGVGPIGPPGERGAPGNRGFPQDGLAGPKGAPGER
GPSGLAGPKGANGDPGRPGEPGLPGARGLTGRPGDAGPQGVGPSGAPGEDGRPGPPGPQ

GARGQPGVMGFPKPGKANGEPGKAGEKGLPGAPGLRGLPGKDGETGAAGPPGPAGPAGE
RGEQGAPGPSGFQGLPGPPGPPGEGGKPGDQGVPGGEAGAPGLVGPRGERGFPGERGSPGAQ
GLQGPRGLPGTPGTDGPKGASGPAGPPGAQGPGLQGMPGERGAAGIAGPKGDRGDVGEK
GPEGAPGKDGGRGLTGPIGPPGPAGANGEKGEVGPMPGAGSAGARGAPGERGETGPPGPAG
FAGPPGADGQPGAKGEQGEAGQKGDAGAPGPQGPSGAPGPQGPTGVTGPKGARGAQGPP
GATGFPGAAGRVPGPSNGNPGPPGPPGPSKDGPKGARGDSGPPGRAGEPGLQGPAGPPG
EKGEPGDDGPSGAEGPPGPQGLAGQRGIVGLPGQRGERGFPLGPSGEPGKQGAPGASGD
RGPPGPVGPGLTGPAEPGREGSPGADGPPGRDGAAGVKGDRGETGAVGAPGAPGPPGPS
GPAGPTGKQGDRGEAGAQQPMGPSGPAGARGIQGPQGPRGDKGEAGEPGERGLKGHRGF
TGLQGLPGPPGPSGDQGASGPAGPSGPRGPPGPVGPSKDGANGIPGPIGPPGPRGRSGETGP
AGPPGNPGPPGPPGPPGPGIDMSAFAGLGPKEKGPDPQLQYMRHHHHHH

经过优化设计后, 编码 SEQ.ID.NO.10 ($\alpha 1$ (II)M6) 氨基酸序列的基因 (记为 COL2A1M6) 的 DNA 序列如 SEQ.ID.NO.10 所示:

SEQ.ID.NO.10:

GAATTCTGGAGTCATCCTCAATTCGAAAAACAAATGGCTGGTGGATTCGATGAAAA
GGCTGGTGGAGCCCAATTAGGTCCTCCACAAGGTCCTCCCGGTCCACCTGGTCCTCCCG
GTCCTCCAGGTCCTCCCGCGGTGCTCCTGGACCACAGGGTTTCCAAGGAAACCCCGGTGA
ACCAGGTGAGCCTGGTGTTCAGGTCCTCCCGGTCTCCAGGACCACCTGGACCACCA
GGAAAGCCTGGTGACGACGGAGAAGCTGGTAAACCAGGAAAGGCAGGAGAGAGAGGT
CCACCTGGACCTCAGGGTGCCAGAGGTTTCCCAGGTACCCCTGGTCTTCTGGTGTCAA
GGGTCATAGAGGTTACCCCGGTTTGGATGGTGCCAAGGGTGAAGCCGGTGCCCCTGGTG
TTAAGGGTGAATCAGGAAGTCCCGGTGAAAATGGAAGTCCCGGTCCACCCGGTCCACC
TGGACTGCCAGGTGAGAGAGGAAGAACCGGACCAGCTGGTGCTGCAGGTGCTAGAGG
AAATGACGGACAGCCCGGACCAGCCGGACCTCCCGGTCTGTTGGGCCCCGAGGTGGT
CCTGGTTTCCCTgGTGCTCCTGGAGCCAAAGGAGAAGCCGGACCCACCGGAGCCAGAGG
TCCCGAGGGAGCACAGGGACCTAGAGGAGAACCAGGTACACCAGGTAGTCCCGGTCTCT
GCTGGTGCATCAGGAAATCCCGGAACTGACGGTATTCCAGGAGCAAAGGGATCTGCAG
GAGCACCAGGAATAGCTGGTGCTCCTGGATTTCCAGGTCCCAGAGGACCTCCCGGTCTCT
CAAGGAGCAACAGGTCCTTTGGGACCAAAAGGTCAAACAGGAGAACCAGGTATTGCTG
GATTCAAAGGAGAGCAAGGTCCAAAGGGAGAGCCCGGTCCCGCAGGTCCCCAAGGAG
CCCCAGGACCAGCTGGTGAAGAAGGAAAAAGAGGAGCCAGAGGTGAACCTGGAGGAG
TAGGACCTATTGGTCCTCCTGGTGAGAGAGGTGCTCCCGGAAACAGAGGTTTTCTGGT
CAAGATGGTCTGGCTGGACCTAAAGGTGCTCCAGGAGAGAGAGGACCTTCAGGACTTG
CTGGTCCAAAAGGTGCTAACGGAGATCCAGGAAGACCCGGTGAACCTGGTCTGCCTGG
AGCTAGAGGATTAACAGGAAGACCAGGTGACGCAGGTCCCAGGGTAAAGTGGGTCCC

AGTGGTGCCCCAGGTGAAGATGGAAGACCTGGTCCTCCCGGACCCCAAGGTGCAAGAG
GTCAGCCTGGAGTGATGGGATTTCTTGACCCAAGGGTGCTAACGGAGAACCTGGAAA
AGCTGGTGAGAAAGGACTGCCCCGGTGCCCCAGGTCTTAGAGGTTTGCCAGGTAAAGAT
GGAGAAACAGGAGCCGCAGGACCACCCGGTCCAGCCGGACCAGCAGGAGAGAGAGGT
GAACAAGGAGCACCTGGTCCAAGTGTTTTCAGGGTCTTCCAGGTCCCCCTGGTCCACC
AGGAGAGGGAGGTAAACCAGGTGACCAAGGTGTCCCTGGAGAAGCAGGTGCACCCGG
TCTTGTGGGTCCAAGAGGTGAAAGAGGATTCCCTGGTGAGAGAGGATCTCCCGGAGCC
CAGGGACTTCAAGGTCCTAGAGGTCTGCCAGGTACCCCTGGTACAGACGGACCAAAGG
GAGCATCAGGACCCGCTGGACCTCCCGGAGCCCAAGGTCCTCCAGGTTTACAAGGTATG
CCTGGTGAAAGAGGTGCTGCAGGTATAGCTGGACCAAAAGGAGACAGAGGTGACGTTG
GTGAGAAGGGTCCCGAAGGAGCCCCTGGAAAAGATGGTGGAAGAGGATTAACAGGTCC
TATAGGACCACCCGGTCCAGCCGGTGCTAATGGAGAAAAAGGAGAAGTAGGTCCTCCA
GGTCCAGCAGGATCTGCAGGTGCTAGAGGTGCCCCCTGGAGAGAGAGGTGAAACAGGAC
CACCTGGTCCAGCTGGTTTTCGCTGGTCCCCCAGGAGCTGATGGACAGCCCGGTGCAA
AGGTGAACAAGGAGAAGCCGGACAGAAGGGAGATGCTGGAGCCCCCGGTCCACAAGG
TCCCTCAGGAGCACCAGGTCCTCAAGGTCCAAGTGGTGACCGGGCCAAAGGGTGCA
AGAGGAGCACAGGGACCTCCAGGAGCAACAGGTTTCCCAGGAGCTGCTGGTAGAGTCG
GTCCACCCGGATCTAATGGTAACCCCGGACCACCAGGACCACCTGGACCATCTGGAAAG
GATGGACCCAAAGGAGCAAGAGGAGATTTCAGGACCACCCGGAAGAGCAGGAGAACCT
GGATTACAGGGTCCCGCCGGTCCACCAGGAGAGAAAGGAGAGCCCGGAGATGATGGTC
CCTCAGGTGCAGAGGGACCCCCAGGACCCCAAGGTCTGGCAGGTCAAAGAGGTATAGT
GGGTCTTCCAGGTCAAAGAGGTGAAAGAGGATTTCCAGGACTTCCAGGTCTTTCAGGT
GAACCCGGTAAACAGGGAGCCCCCGGAGCCTCAGGTGACAGAGGTCTCTCCAGGACCA
GTAGGACCCCCAGGTTTAACCGGACCAGCAGGTGAGCCAGGAAGAGAAGGTTCTCCTG
GAGCCGATGGACCTCCAGGAAGAGACGGTGCAGCTGGTGTTAAGGGTGACAGAGGTGA
AACTGGAGCCGTAGGAGCCCCAGGTGCCCCCGGACCACCCGGATCACCCGGACCTGCA
GGTCTTACTGGTAAACAAGGAGATAGAGGAGAAGCCGGTGCCCAGGGTCTATGGGTC
CTTCTGGTCTGTCAGGAGCAAGAGGTATACAAGGTCCACAGGGTCCCAGAGGTGACAA
GGGTGAAGCAGGAGAACCCGGTGAGAGAGGTCTGAAGGGTCATAGAGGATTACCCGG
GTTACAGGGTTTGCCAGGACCCCTGGACCAAGTGGTGACCAGGGTGCATCCGGTCCA
GCAGGTCTTCTGGACCAAGAGGTCTCCCGGTCCAGTTGGTCCATCAGGTAAAGACGG
AGCCAACGGTATCCCAGGTCCCATCGGTCTCCAGGTCTAGAGGAAGAAGTGAGAG
ACTGGTCTGCTGGACCTCTTGAAACCCTGGTCTCTCCAGGACCTCCAGGTCTCTCAGG
TCCCGGAATAGATATGTCCGCTTTCGCTGGATTGGGACCAAGAGAGAAAGGTCCTGACC
CTCTTCAATATATGAGAGCACACCATCACCATCATCACTAA

DNA 序列的合成委托南京金斯瑞生物科技股份有限公司完成，合成 SEQ.ID.NO.8、SEQ.ID.NO.10 两种基因的 DNA 片段。

实施例 2. 重组表达载体的构建、菌种筛选

(1) 重组表达载体的构建

将合成后的基因片段 SEQ.ID.NO.8、SEQ.ID.NO.10 重组至 pPIC9K 空载体（购自赛默飞世尔科技公司）中，使目的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，获得表达 $\alpha 1(\text{II})\text{M6}$ 的 pPIC9K-COL2A1M6 和表达 $\alpha 1(\text{I})\text{M1}$ 的 pPIC9K-COL1A1M1 两种重组表达载体质粒。

将 pPIC9K-COL2A1M6、pPIC9K-COL1A1M1 质粒转化进入感受态大肠杆菌 DH5 α （购自生工生物工程(上海)股份有限公司），在含有氨苄青霉素的 LB 抗性平板筛选阳性克隆，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），验证正确。pPIC9K-COL1A1M1、pPIC9K-COL2A1M6 的质粒图谱分别如图 3 和图 4 所示。

(2) 菌种筛选

将上述重组表达载体质粒 10 μg ，用 Sac I（购自大连 TaKaRa 公司，具体操作按试剂盒说明书进行）37 $^{\circ}\text{C}$ 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，使体积控制在 10 μL 左右。

将线性化质粒电转化入宿主菌种毕赤酵母 SMD1168（购自赛默飞世尔科技公司）感受态细胞中，将电转后的菌液涂布于 MD 平板上，每 100 μL ~200 μL 涂布一块平板，室温静置 10min，于 30 $^{\circ}\text{C}$ 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，并转移到 50mL 离心管中。以无菌双蒸水稀释菌悬液，10⁵ 个细胞涂布于含有 0.5mg/mL G418 的 YPD 平板上，倒置，30 $^{\circ}\text{C}$ 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（200 μL YPD/孔），混匀，于 30 $^{\circ}\text{C}$ 培养 48h；混匀孔中菌液，各取 10 μL 接入至一块新的无菌 96 孔板，于 30 $^{\circ}\text{C}$ 培养 24h 后再重复一次此操作；24h 后，从第三块 96 孔板中取出 1 μL 分别点在含有 1.0mg/mL 和 4mg/mL G418 的 YPD 平板上，于 30 $^{\circ}\text{C}$ 继续培养 96h~120h。毕赤酵母转化子若能在含高浓度（4mg/mL）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。

构建含 pPIC9K-COL1A1M1、pPIC9K-COL2A1M6 的两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保藏。

含重组表达载体 pPIC9K-COL1A1M1 的工程菌表达重组 $\alpha 1(\text{I})\text{M1}$ 胶原蛋白，保藏于中国

微生物菌种保藏管理委员会普通微生物中心，保藏编号是：CGMCC NO.21891，地址：北京市朝阳区北辰西路1号院3号；保藏日期：2021年03月11日；分类命名：巴斯德毕赤酵母 *Pichia pastoris*。

含重组表达载体 pPIC9K-COL2A1M6 的工程菌表达重组 $\alpha 1(\text{II})\text{M6}$ 胶原蛋白，保藏于中国微生物菌种保藏管理委员会普通微生物中心，保藏编号是：CGMCC NO. 21892。地址：北京市朝阳区北辰西路1号院3号；保藏日期：2021年03月11日；分类命名：巴斯德毕赤酵母 *Pichia pastoris*。

实施例 3.诱导表达与重组胶原蛋白的鉴定

分别取实施例 2 得到的表达 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 的重组工程菌，同时取已知专利中的表达全长 I 型胶原 $\alpha 1$ 链蛋白毕赤酵母工程菌株和表达全长 II 型胶原 $\alpha 1$ 链蛋白的毕赤酵母工程菌株作为对照，2 个对照工程菌株均为发明人团队前期研究成果，所表达的全长胶原 $\alpha 1$ 链同样于肽链氨基端添加 Strep-Tag II 标签、羧基端添加 6 $\times$ His Tag 标签），其分别来自于申请号 201911135958.0（名称：酵母重组人源 I 型胶原 $\alpha 1$ 链蛋白、合成方法及其应用，专利中表达全长 $\alpha 1(\text{I})$ 链的毕赤酵母工程菌种保藏于中国微生物菌种保藏管理委员会普通微生物中心，保藏编号：CGMCC NO.17150）、申请号 201911088025.0（名称：毕赤酵母生产重组人源 II 型胶原蛋白单链的方法，专利中表达全长 $\alpha 1(\text{II})$ 链的毕赤酵母工程菌种保藏于中国微生物菌种保藏管理委员会普通微生物中心，保藏编号：CGMCC NO.17149）的专利。将 4 种工程菌置于装有 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% 甲醇至培养基中甲醇的终浓度为 1.0%。加甲醇诱导 16h 以上，就可收取菌液样品，取样量为 1mL，置于 1.5mL EP 管中，4℃ 下以 12000g 离心 5min，收集表达上清，待检测样品于 -80℃ 保存备用。

收取的表达上清，加入 5 $\times$ 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-Tag II 标签，羧基端有 6 $\times$ His Tag 标签，可以抗 Srtep-Tag II、抗 6 $\times$ His Tag 的抗体（购自南京金斯瑞生物科技股份有限公司）进行 Western Blot 检测（具体操作参看说明书进行）。

表达上清的 SDS-PAGE 如下图 5 所示， $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 、 $\alpha 1(\text{I})$ 、 $\alpha 1(\text{II})$ 在诱导表达 24h 均可高效分泌表达于胞外的表达上清中， $\alpha 1(\text{I})$ 、 $\alpha 1(\text{II})$ 在预期的目的条带（>116kDa）下出现了明确的主降解条带（<116kDa），而本发明的 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 则只有预期的目

的条带 (>116kDa)。

使用 Image Lab 软件 (Bio-Rad Gel Doc XR+成像仪) 测算, 结果如下:

(1) $\alpha 1(I)$ M1 目的条带表观分子量 (116.3 kDa) 与 $\alpha 1(I)$ 目的条带表观分子量 (116.4 kDa)、 $\alpha 1(II)$ M6 目的条带表观分子量 (118.2kDa) 与 $\alpha 1(II)$ 目的条带表观分子量 (118.1kDa) 相比基本一致, 明显大于 $\alpha 1(I)$ 的主降解条带的表观分子量 (104.5kDa) 与 $\alpha 1(II)$ 的主降解条带的表观分子量 (106.9kDa)。

(2) $\alpha 1(I)$ 电泳结果中目的条带与主降解条带的比值为 51.5%: 48.3%; $\alpha 1(II)$ 电泳结果中目的条带与主降解条带的比值为 52.1%: 47.8%, 主降解产物与目的产物占比基本相同。

从图 6 的 ECL 化学发光显色结果(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 Srtp-Tag II 标签, 羧基端 6×His Tag 标签均可检测到, 且目的条带均与 SDS-PAGE 中表观分子量大小相同, 说明 $\alpha 1(I)$ M1、 $\alpha 1(II)$ M6 两种重组胶原蛋白成功的进行了全长序列的高效分泌表达, 目的条带的表达符合预期, 而 $\alpha 1(I)$ 、 $\alpha 1(II)$ 两种胶原蛋白的目的条带序列虽然是全长完整的, 但是主降解条带缺失了氨基端的序列, 只能检测到羧基端 6×His Tag 标签。

将 $\alpha 1(I)$ M1 和 $\alpha 1(II)$ M6 在 SDS-PAGE 上的目的条带、 $\alpha 1(I)$ 和 $\alpha 1(II)$ 在 SDS-PAGE 上的目的条带与主降解带条带切割下来, 用胰蛋白酶将其酶解, Nano-HPLC-MS/MS 质谱检测重组胶原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), 并将检测到肽段进行序列比对(Uniprot 数据库), 数据比对结果及鉴定肽段与天然序列比对覆盖图(底色为灰色部分: 条带中质谱鉴定到肽段与天然序列相完全相同的部分), 如图 7 和图 8 所示, 结果可见:

(1) $\alpha 1(I)$ M1、 $\alpha 1(I)$ 的目的条带及 $\alpha 1(I)$ 的主降解带被酶解后检测到的肽段均属于 I 型胶原 $\alpha 1$ 链上的序列。

(2) $\alpha 1(II)$ M6 和 $\alpha 1(II)$ 的目的条带及 $\alpha 1(II)$ 的主降解带被酶解后检测到的肽段均属于人 II 型胶原 $\alpha 1$ 链上的序列。

以上结果说明 $\alpha 1(I)$ M1、 $\alpha 1(II)$ M6 与 $\alpha 1(I)$ 、 $\alpha 1(II)$ 一样成功表达, 分别属于人 I 型胶原 $\alpha 1$ 链、人 II 型胶原 $\alpha 1$ 链的重组胶原蛋白, 但 $\alpha 1(I)$ 、 $\alpha 1(II)$ 表达时产生了降解, 主降解条带也属于相应种类胶原蛋白。

实施例 4. 高密度发酵与纯化

(1) 对基因工程菌进行高密度发酵

重组 $\alpha 1(I)$ M1、 $\alpha 1(II)$ M6 胶原蛋白规模化表达生产, 获取含有重组胶原蛋白的发酵液。

种子培养基 YPG (含酵母粉 10g/L、酵母蛋白胨 20g/L、无水甘油 10g/L); 发酵培养基 (含 $\text{NH}_4\text{H}_2\text{PO}_4$ 190.4 g/L、 KH_2PO_4 10.06 g/L、 $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ 1.18 g/L、 K_2SO_4 18.2 g/L、 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$

14.9 g/L、甘油 40 g/L)；补料培养基(含 50%W/V 甘油，每升加 12mL PTM1 微量元素)；诱导培养基(含 100%甲醇，每升加入 12mL PTM1 微量元素)；PTM1：用 0.22 μ m 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后加入 PTM1，用氨水调节 pH 值至 5.0。

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℃。

工程菌接入含种子培养基 YPG 的 1L 摇瓶，220rpm、30℃，培养 18-20h，至 OD600=2~10。使用 5L 发酵罐(保兴生物)，装液量 2L 发酵培养基，2%甘油分开灭菌，接种前调节转速为 300rpm，通气量 4L/min，温度 30℃，用浓氨水配制好的碱液调 pH，设置 pH 至 4.5。然后先接入 0.9mL PTM1，然后再将制备好的 200mL 种子液接入罐内(火焰圈接种)，然后点击溶氧电极校百，校百后开始发酵。待生长溶氧第一次掉至 30%时，采用溶氧串级转速功能，保持生长溶氧为 30%；等待甘油耗完，溶氧反弹、溶氧大于 70% (OD600 值约 20)，取消溶氧串级转速，调高搅拌 650rpm，甘油采用 30%联动补料，补料 80mL。停止补甘油，溶氧反弹至 70%以上后，设置 pH 值为 4、温度 29℃，以甲醇、甘油混合碳源(甲醇：50%甘油=7：3)进行诱导培养。手动补加 5mL，待溶氧反弹至 70%以上后，设定补料速度为 8mL/h，一小时后提高到为 10mL/h，一小时后再次提高设定到 20mL/h。待溶氧值低于 30%，停止补料，等待溶氧反弹，溶氧回升至 30%后联动补料。诱导 40~60h，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： $C(\text{mg/mL})=0.144*(A_{215}-A_{225})$ ， $A_{215}<1.5$ 。同时，分别同时取表达全长 $\alpha 1(\text{I})$ 链的毕赤酵母工程菌种(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：CGMCC NO.17150)、表达全长 $\alpha 1(\text{II})$ 链的毕赤酵母工程菌种(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：CGMCC NO.17149)进行高密度发酵。

结果如表 1 所示，在诱导 48 h 后， $\alpha 1(\text{I})$ 与 $\alpha 1(\text{I})\text{M1}$ 相比、 $\alpha 1(\text{II})\text{M6}$ 与 $\alpha 1(\text{II})$ 相比，菌浓度(OD600)、菌湿重、发酵液中表达蛋白浓度 UV 定量三个指标并无明显差异。但收集发酵上清进行 SDS-PAGE 电泳检测，结果如图 9 所示，可发现在高密度发酵的条件下， $\alpha 1(\text{I})$ 、 $\alpha 1(\text{II})$ 的主降解条带极为明显，与摇瓶诱导表达无区别；而 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 最主要的产物仍是其目的条带，主降解条带没有出现，这说明 $\alpha 1(\text{I})\text{M1}$ 、 $\alpha 1(\text{II})\text{M6}$ 在消除主降解条带(主降解产物)的效果在高密度发酵的条件下仍能有效保持。

表 1.发酵小试实验菌浓度、菌湿重、蛋白质表达量（UV 定量）

种类	OD600	菌湿重（g/L）	蛋白质（UV， g/L）
α1(Ⅱ)	189.0	260.0	17.8
α1(Ⅱ)M6	198.0	265.0	18.7
α1(Ⅰ)	215.0	310.0	18.1
α1(Ⅰ)M1	201.0	301.0	17.3

（2）胶原蛋白纯化

缓冲液 A： 20mM KH₂PO₄， pH 值为 4.0；
缓冲液 B： 20mM KH₂PO₄、 0.5M NaCl， pH 值为 4.0。

收集发酵液，2000g、30min、4℃离心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 装载于利穗科技产 GCC-50-400 层析柱，使用 GE AKTA Pure 蛋白质分离层析纯化系统）至 A215 吸光值和电导率值都保持不变后，设置 40us/cm 的流速上样，上样体积 0.5L/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，直至紫外和电导降至最低且不再变化。收集洗脱液，分别检测确定好组份后，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。取纯化后冻干海绵溶于超纯水，进行 SDS-PAGE 电泳，如图 10 所示。α1(Ⅰ)M1、α1(Ⅱ)M6 经一步离子交换纯化后，去除多数的杂蛋白和小降解条带等，即可获得纯度高的单一目的蛋白质（使用 Image Lab 软件测算，α1(Ⅰ)M1 纯度为 90.1%、α1(Ⅱ)M6 纯度为 88.3%）；而 α1(Ⅰ)、α1(Ⅱ)经相同的纯化步骤后，主降解带依旧出现、无法消除，主降解产物与目的产物大小相差不多，性质相近，很难使用一步纯化将二者分离，根据申请号 201911135958.0 和申请号 201911088025.0 的专利内容可知，想获得单一目的全长 α1(Ⅰ)链、全长 α1(Ⅱ)链产物，需要利用全长 α1(Ⅰ)链、全长 α1(Ⅱ)链氨基端有 Srtep-Tag II 标签，羧基端有 6×His 标签的性质，使用 Ni-NTA、Strep-Tactin 亲和层析介质进行双亲和纯化才能做到，这样的话在 SDS-PAGE 上表现为主降解条带的主降解产物则被舍弃，浪费了菌体的生物合成资源，增加了纯化步骤，目的产物得率降低。

实施例 5. 重组胶原蛋白的检测

（1）傅里叶变换红外光谱(FT-IR) 分析

试验取微量 α1(Ⅰ)M1、α1(Ⅱ)M6、α1(Ⅰ)、α1(Ⅱ)胶原蛋白纯化后冻干样品分别混合溴化钾(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 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 、 $\alpha 1(I)$ 、 $\alpha 1(II)$ 蛋白纯化样品的红外光谱扫描见图 11 和图 12, 可以看到其特征吸收均波数均符合重组胶原蛋白结构特征: 酰胺 A (3299cm^{-1} 左右)、酰胺 B (3081cm^{-1} 左右)、酰胺 I (1650cm^{-1} 左右)、酰胺 II ($1530\sim 1550\text{cm}^{-1}$ 左右)、酰胺 III (1240cm^{-1} 左右), 说明 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 中氨基酸的突变不影响胶原蛋白本身性质 (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.)。

(2) 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)。取重组 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 、 $\alpha 1(I)$ 、 $\alpha 1(II)$ 胶原蛋白冻干海绵、对照人胶原蛋白 (购买自 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 (用超纯水或 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 PBS (pH 7.4) 稀释至 0.5mg/mL 。向 96 孔细胞培养板中加入 $100\mu\text{L}$ 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 NIH/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 OD_{492nm} 的吸光度值 (具体操作参照说明书执行)。

OD_{492nm} 的吸光度相应的表征可以代表胶原蛋白样品的细胞粘附活性: OD_{492nm} 的吸光度越高, 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 13 所示, 重组 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 、 $\alpha 1(I)$ 、 $\alpha 1(II)$ 胶原蛋白均与商品化的天然人胶原蛋白有类似的细胞黏附活性, 且都显著高于对照组; $\alpha 1(I)M1$ 与 $\alpha 1(I)$ 相比、 $\alpha 1(II)M6$ 与 $\alpha 1(II)$ 相比, 细胞黏附活性基本一致, 无显著差异。

(3) 重组胶原蛋白水凝胶的制备与检测

取重组 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 、 $\alpha 1(I)$ 、 $\alpha 1(II)$ 胶原蛋白, 以 10% 浓度溶于注射用水, 调节

pH 值控制在 4~6 的范围内,0.22μm 无菌滤器过滤除菌;以每 1g 胶原干粉的比例加入 0.1g 10% (w/w) 无菌 N-羟基琥珀酰亚胺(NHS)溶液,混匀;然后以每 g 胶原干粉的比例加入 0.13g 50% (w/w) 无菌 1-(3-二甲基氨丙基)-3-乙基碳二亚胺盐酸盐(EDC)溶液;室温 (20-30℃) 静置反应 2-6h,形成水凝胶。将水凝胶置于无菌 PBS 溶液(NaCl 8.5g/L、Na₂HPO₄ 0.5 g/L、NaH₂PO₄ 0.15 g/L, pH 7.2)透析,凝胶: PBS 透析液=1: 6 (m/m),连续透析 120h,每 24h 彻底更换一次透析液,去除 NHS、EDC 残留。将透析后的水凝胶装入无菌容器中,室温放置。

将水凝胶冷冻干燥去除水分,取冻干后水凝胶称重,再置于无菌 PBS 溶液中放置 24h,待其彻底吸水溶胀,取出水凝胶,吸水纸将表面水分吸干后称重。参考文献中方法计算溶胀率: $Q_{\text{溶胀率}} = (W_{\text{吸水溶胀质量}} - W_{\text{干凝胶重量}}) / W_{\text{干凝胶重量}}$ 。流变仪 (Discovery HR-2) 检测水凝胶的弹性模量 (储能模量,小振幅频率扫描,25℃、应力 0.5%、0.1-100.0 rad/s)、动力粘度 (流动峰值保持,25℃、剪切速率 2.0 s⁻¹)。将冻干后水凝胶转入液氮中速冻,掰断,使用扫描电镜 (日立 TM3030PLUS) 对冻干后水凝胶表面进行扫描。弹性模量、动力粘度、溶胀率结果如表 2 所示,α1(I)M1 与 α1(I)相比、α1(II)M6 与 α1(II)相比,相同条件下所制备的水凝胶的流体力学性质基本一致,并无明显改变。

表 2. 四种水凝胶的弹性模量、动力粘度、溶胀率检测

凝胶种类	弹性模量(Pa)	粘度(Pa·S)	溶胀率(g/g 干凝胶)
α1(II)	104.63	82.87	14.18
α1(II)M6	102.02	77.80	14.01
α1(I)	234.67	190.67	12.53
α1(I)M1	292.68	170.97	12.22

如图 14、15 所示,由 α1(I)M1、α1(II)M6 所制备的水凝胶与由 α1(I)、α1(II)所制备的水凝胶一样,均为多孔网状结构,孔径范围集中于 100-200μm,具有良好的通透性,具备保持大量水分的空间结构基础,可作为细胞提供黏附、支撑、生长迁移的空间和输送营养物质与新陈代谢产物的通道,有应用于生物医学材料领域的潜力。

(4) 重组胶原蛋白水凝胶的细胞检测

无菌存放的水凝胶放置于 24 孔细胞培养板中。取正常培养 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), PBS 清洗、胰酶消化、加培养基吹打均匀并计数, 10⁵ 细胞/孔接入到含水凝胶的培养皿中,共培养 24-72 小时,观察细胞于水凝胶上的粘附与增值情况。

(1) 取一块 24 孔板用 DMSO 制备 1mM 的钙黄绿素 AM (购于碧云天生物技术研究 所),并用 D-PBS 将其稀释成 50 μM 的钙黄绿素工作液。吸去孔中培养基,用 PBS 清洗数 次,加入 1 mL 的含血清 DMEM 培养基,加入 100 μL 钙黄绿素 AM 溶液(培养基的 1/10),

孵育 30 min，对细胞进行染色，再更换生长培养基，培养 30 min，轻取出水凝胶，置入新的培养孔中，荧光显微镜下拍照（最大激发光波长为 494nm，最大发射光波长为 514nm）。

（2）另取一块 24 孔板中加入 200 μ L MTT 溶液（购于碧云天生物技术有限公司），培养 NIH/3T3 细胞 4 h 后，观察细胞中蓝紫色结晶形成情况，弃培养基，用 PBS 清洗水凝胶，纵切开水凝胶，置入新的培养孔中，显微镜下拍照。

该实施例实验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。

结果如图 16、17 所示，由 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 所制备的水凝胶与由 $\alpha 1(I)$ 、 $\alpha 1(II)$ 所制备的水凝胶一样，明场显微镜下观察 NIH/3T3 细胞形态正常，为典型的成纤维细胞形态；水凝胶上黏附生长的 NIH/3T3 细胞被钙黄绿素 AM 染色后可检测到绿色荧光（照片中明亮部分）；水凝胶中生长的 NIH/3T3 细胞在加入 MTT 后，可形成的蓝紫色结晶（照片中黑色部分）；绿色荧光与蓝紫色结晶只能由活细胞形成，说明 NIH/3T3 细胞均能正常黏附、生长、迁移于水凝胶中，由 $\alpha 1(I)M1$ 、 $\alpha 1(II)M6$ 制备的水凝胶与天然序列的 $\alpha 1(I)$ 、 $\alpha 1(II)$ 所制备的水凝胶有类似的生物学功能，可作为新型生物医疗器械应用于创伤修复、组织再生等领域。

权 利 要 求 书

1. 重组胶原蛋白 $\alpha 1$ 链, 其特征在于, 所述重组胶原蛋白 $\alpha 1$ 链为 $\alpha 1(I)M1$ 或 $\alpha 1(II)M6$, 所述 $\alpha 1(I)M1$ 由人 I 型胶原蛋白 $\alpha 1$ 链的天然全长氨基酸序列经过氨基酸突变获得, 所述 $\alpha 1(II)M6$ 由人 II 型胶原蛋白 $\alpha 1$ 链天然全长氨基酸序列经过氨基酸突变获得; 所述 $\alpha 1(I)M1$ 的序列如 SEQ.ID.NO.2 所示, 所述 $\alpha 1(II)M6$ 的序列如 SEQ.ID.NO.5 所示。
2. 编码权利要求 1 所述重组胶原蛋白 $\alpha 1$ 链的核苷酸。
3. 根据权利要求 2 所述的核苷酸, 其特征在于, 所述编码 $\alpha 1(I)M1$ 的核苷酸序列如 SEQ.ID.NO.3 所示; 所述编码 $\alpha 1(II)M6$ 的核苷酸序列如 SEQ.ID.NO.6 所示。
4. 重组表达载体, 含有权利要求 2 所述的核苷酸。
5. 工程菌, 所述工程菌含有权利要求 4 所述的重组表达载体或表达权利要求 1 所述重组胶原蛋白 $\alpha 1$ 链, 所述工程菌为巴斯德毕赤酵母 (*Pichia pastoris*)。
6. 根据权利要求 5 所述的工程菌, 其特征在于, 所述工程菌保藏于中国微生物菌种保藏管理委员会普通微生物中心, 保藏编号为 CGMCC NO.21891 或 CGMCC NO. 21892, 其中保藏编号为 CGMCC NO.21891 的工程菌表达的是重组 $\alpha 1(I)M1$ 胶原蛋白 $\alpha 1$ 链, 保藏编号为 CGMCC NO. 21892 的工程菌表达的是重组 $\alpha 1(II)M6$ 胶原蛋白 $\alpha 1$ 链。
7. 权利要求 4 所述的重组表达载体或权利要求 5-6 任一项所述的工程菌在制备所述重组胶原蛋白 $\alpha 1$ 链中的应用。
8. 权利要求 1 所述重组胶原蛋白 $\alpha 1$ 链的制备方法, 其特征在于, 包括如下步骤:
 - (1) 合成编码权利要求 1 所述重组胶原蛋白 $\alpha 1$ 链的核苷酸序列;
 - (2) 构建重组表达载体;
 - (3) 构建工程菌、诱导表达和菌株筛选, 得到表达量高的工程菌株;
 - (4) 将筛选得到的表达量高的工程菌, 高密度发酵培养;
 - (5) 发酵上清液纯化后, 冷冻干燥获得重组胶原蛋白 $\alpha 1$ 链。
9. 根据权利要求 8 所述的制备方法, 其特征在于, 所述表达量高的工程菌为巴斯德毕赤酵母 (*Pichia pastoris*), 保藏于中国微生物菌种保藏管理委员会普通微生物中心, 保藏编号为 CGMCC NO.21891 或 CGMCC NO. 21892。
10. 组合物, 其特征在于, 所述组合物包含权利要求 1 所述重组胶原蛋白 $\alpha 1$ 链或权利要求 8-9 任一项所述方法制备的重组胶原蛋白 $\alpha 1$ 链。
11. 制品, 其特征在于, 所述制品包含权利要求 1 所述重组胶原蛋白 $\alpha 1$ 链、权利要求 8-9 任一项所述方法制备的重组胶原蛋白 $\alpha 1$ 链或权利要求 10 所述的组合物, 所述制品为药物、医疗器械、生物材料、组织工程产品或化妆品。
12. 根据权利要求 11 所述的制品, 其特征在于, 所述制品为为细胞提供黏附、支撑、生长迁

移空间的材料或作为输送营养物质与新陈代谢产物通道的材料。

13. 根据权利要求 12 所述的制品，其特征在于，所述制品为胶原蛋白水凝胶。
14. 权利要求 1 所述重组胶原蛋白 $\alpha 1$ 链、权利要求 2-3 任一项所述的核苷酸、权利要求 4 所述的重组表达载体、权利要求 5-6 任一项所述的工程菌、权利要求 8-9 任一项所述方法制备的重组胶原蛋白 $\alpha 1$ 链或权利要求 10 所述的组合物在制备制成品中的用途，所述制成品为药物、医疗器械、生物材料、组织工程产品或化妆品。
15. 权利要求 1 所述重组胶原蛋白 $\alpha 1$ 链、权利要求 2-3 任一项所述的核苷酸、权利要求 4 所述的重组表达载体、权利要求 5-6 任一项所述的工程菌、权利要求 8-9 任一项所述方法制备的重组胶原蛋白 $\alpha 1$ 链或权利要求 10 所述的组合物在制备促进创伤修复或组织再生的产品中的用途。
16. 根据权利要求 14 所述的用途，其特征在于，所述制成品为胶原蛋白水凝胶。
17. 根据权利要求 15 所述的用途，其特征在于，所述产品为胶原蛋白水凝胶。