

# 分子串联重复策略用于制备超短肽

赵晨, 李端华, 李进军, 王轲

成都大学药学院 四川抗菌素工业研究所 抗生素研究与再评价四川省重点实验室, 四川 成都 610106

赵晨, 李端华, 李进军, 王轲. 分子串联重复策略用于制备超短肽. 生物工程学报, 2022, 38(12): 4587-4600.

ZHAO C, LI DH, LI JJ, WANG L. Molecular tandem repeat strategy for production of ultrashort peptides. Chin J Biotech, 2022, 38(12): 4587-4600.

**摘要:** 超短肽具有更好的稳定性、组织渗透性、生物相容性以及更低的免疫原性。GHK (glycyl-L-histidyl-L-lysine) 和 GQPR (glycyl-L-glutamyl-L-prolyl-L-arginine) 具有刺激胶原蛋白产生、减缓胶原蛋白降解的作用, 作为抗皱成分广泛应用于化妆品。超短肽一般都是通过固相合成方法制备, 其缺陷是制备过程中大量使用有机化学试剂而造成环境负担, 故本文探讨了一种设计和制备超短肽的新方法。因序列短而无法直接重组表达, 文中首先构建了适用于融合表达的载体骨架 pET28a-Trxm。以 GHK 和 GQPR 串联重复基因作为滚环扩增的基本单元 (tandem repeat of short peptides, TRSP), 反应时随机掺入 5-甲基胞嘧啶获得长基因片段, 然后经 *Acc65 I* 和 *Apa I* 消化产生随机长度的基因。胶回收 500 bp 到 1 500 bp 的 DNA 片段, 克隆得到表达载体 pET28a-Trxm-(TRSP)<sub>n</sub> 并转化获得重组菌。双酶切及测序结果表明, 成功构建获得串联重复数  $n=1$ 、2、3、4、6、7、8、9 的阳性克隆。蛋白表达结果显示, 当串联重复数  $n=1$ 、2、3、4、8、9 时均有相应融合蛋白表达, 表达水平随着重复数增加而降低。Trxm-(TRSP)<sub>1</sub> 表达水平最高, 达总蛋白的 50%, 而 Trxm-(TRSP)<sub>2</sub> 表达水平为总蛋白的 30%。进一步地, 含 Trxm-(TRSP)<sub>1</sub> 的清液先后经肠激酶和胰蛋白酶切割后, HPLC 分析结果表明, 成功获得超短肽 GHK 和 GQPR。该结果对于超短肽重组制备的工业化应用具有重要价值。

**关键词:** 串联重复基因; 滚环扩增; 超短肽; 融合蛋白; 重组制备

**Received:** May 15, 2022; **Accepted:** August 19, 2022; **Published online:** September 28, 2022

**Supported by:** International Science and Technology Innovation Cooperation Project of Sichuan Province, China (2021YFH0016, 2020YFH0013)

**Corresponding author:** WANG Lu. E-mail: wanglu@cdu.edu.com

**基金项目:** 四川省国际科技创新合作项目 (2021YFH0016, 2020YFH0013)

# Molecular tandem repeat strategy for production of ultrashort peptides

ZHAO Chen, LI Duanhua, LI Jinjun, WANG Lu

Antibiotics Research and Re-evaluation Key Laboratory of Sichuan Province, Sichuan Industrial Institute of Antibiotics, School of Pharmacy, Chengdu University, Chengdu 610106, Sichuan, China

**Abstract:** Ultrashort peptides have higher stability, tissue penetrability, biocompatibility, and less immunogenicity, and are widely applied in biology and medicine. GHK (glycyl-L-histidyl-L-lysine) and GQPR (glycyl-L-glutamyl-L-prolyl-L-arginine) can stimulate collagen renewal and inhibit collagen degradation. GHK and GQPR have been used in cosmetic anti-wrinkle skincare and make-up products. The most common approach for ultrashort peptide production is the solid-phase synthesis, which is eco-unfriendly due to heavy usage of organic chemical reagents during the manufacturing process. Here we report a new approach to the production of ultrashort peptides. Recombinant expression of ultrashort peptides is usually unfeasible because of the short amino acid sequences. A vector pET28a-Trxm harboring the thioredoxin gene was first constructed for subsequent fusion expression. The tandem repeats of GHK and GQPR genes were used as the templates for rolling circle amplification (RCA). The RCA reaction was tuned to incorporate noncanonical nucleotides 5-methylcytosine to obtain long DNA fragments. Gene sequences with various lengths were generated through double digestion of *Acc65* I and *Apa* I. The resulting digestion products were gel recovered by size (from 500 bp to 1 500 bp) and cloned into pET28a-Trxm to obtain the recombinant vector pET28a-Trxm-(TRSP)<sub>n</sub>. The pET28a-Trxm-(TRSP)<sub>n</sub> was introduced into *E. coli* BL21(DE3) to generate a library of Trxm-(TRSP)<sub>n</sub> sequences with a controlled distribution of lengths. Through double digestion and sequencing, positive clones with tandem repeats  $n=1, 2, 3, 4, 6, 7, 8, 9$  were obtained. Protein expression results showed protein bands with corresponding molecular weight, and the protein expression level decreased as the tandem repeats increased. The expression level of Trxm-(TRSP)<sub>1</sub> achieved 50% of the total protein, while the expression level of Trxm-(TRSP)<sub>2</sub> was 30% of the total protein. The crude extracts from cell pellets were further treated with enterokinase cleavage, and the supernatants containing (TRSP)<sub>1</sub> were collected after ultrafiltration and then subjected to trypsin cleavage. HPLC analysis indicated that the ultrashort peptides GHK and GQPR were successfully obtained through two-step cleavage. This study may facilitate the commercial production of ultrashort peptides.

**Keywords:** tandem repeat gene; rolling circle amplification; ultrashort peptide; fusion protein; recombinant production

短肽包含生命分子的基本信息, 由于其独特的特性和巨大的应用前景, 在生物学、医学、化学等领域引起了广泛的关注。短肽具有结构多样和构象灵活的特点, 易于调控与受体结合的相互

作用, 同时显示出靶标结合的高选择性<sup>[1]</sup>。超短肽 (ultrashort peptide) 被精确地定义为肽链长度不超过 7 个氨基酸的肽<sup>[2-5]</sup>。与更长的肽链相比, 超短肽具有非常突出的特性, 其合

成更为经济且具有更好的可调节性、结构稳定性、组织渗透性、生物相容性及细胞相容性以及更低的免疫原性<sup>[6-7]</sup>。化妆品中重要的功能性成分之一就是超短肽。许多超短肽在细胞外基质合成、色素沉着、先天免疫及炎症等发挥作用,具有刺激胶原蛋白生成、促进伤口愈合、抗皱以及抗氧化等效果<sup>[8]</sup>。三肽 GHK (glycyl-L-histidyl-L-lysine) 具有刺激胶原蛋白产生的特性,四肽 GQPR (glycyl-L-glutamyl-L-prolyl-L-arginine) 具有抗炎以减缓胶原蛋白降解的特性<sup>[9-10]</sup>。超短肽进行棕榈酰基修饰后更利于皮肤吸收,棕榈酰三肽-1 (Pal-GHK) 和棕榈酰四肽-7 (Pal-GQPR) 的混合物在局部应用于皮肤时具有抗皱的特性,被广泛应用于化妆品中<sup>[11-12]</sup>。目前,短肽主要通过固相合成法 (solid phase peptide synthesis, SPPS) 制备, SPPS 中肽的纯化是通过简单洗涤附着在聚合物上的肽来实现的,这使得合成过程自动化。由于短肽的分子量较小,一般无法直接进行重组表达,超短肽更是如此。

滚环扩增 (rolling circle amplification, RCA) 是一种等温酶促反应,它以 1 个非常小的环状 DNA 作为模板,用特定的 DNA 或 RNA 聚合酶扩增短 DNA 或 RNA 引物以形成长链 DNA 或 RNA,其产物是包含数十个到数百个与环状模板互补的串联重复序列的多联体。RCA 技术的强大和多功能性使其成为生物医学研究和纳米生物技术的有吸引力的工具<sup>[13-16]</sup>。

RCA 反应时加入非典型核苷酸,这些非典型核苷酸会在反应时随机掺入在产物中且不能被特定的限制性内切酶消化,故消化后的产物可按照 DNA 片段大小分离获得,此方法可制备特定长度范围的串联重复 DNA 片段混合物<sup>[17]</sup>。为了开发一种筛选超短肽基因串联重复数的方法,同时开发一种制备超短肽 GHK 和 GQPR 混合物的方法,本研究以 GHK 和 GQPR 的串联重复基因为模板,利用 RCA 来生成不同长度的串联重复基因片段,随后克隆至以硫氧还蛋白作为融合蛋白的载体骨架 pET28a-Trxm,转化获得大肠杆菌工程菌。通过对比蛋白表达水平确定基因串联重复数,同时建立了一种制备超短肽 GHK 和 GQPR 混合物的方法。

## 1 材料与方法

### 1.1 菌株和质粒

大肠杆菌 (*Escherichia coli*) BL21(DE3) 及 pET28a 购自 Merck 公司; pET28a-Trx27 为本实验室前期构建和保存;具体信息见表 1。

### 1.2 培养基及试剂

Luria-Bertani (LB) 培养基 (g/L): 酵母提取物 5, 胰蛋白胨 10, 氯化钠 10。

Nt. *Bbv*C I, 5-Methylcytosine (5mC), phi29 DNA Polymerase 购自 New England Biolabs 公司。小量质粒提取试剂盒、胶回收试剂盒购自 Biotek 公司。FastDigest *Dpn* I, FastDigest *Acc65* I, FastDigest *Apa* I, FastDigest *Nco* I,

表 1 文中所用菌株和质粒

Table 1 Strains and plasmids used in this study

Strains/plasmids	Characteristics/relevant genotype	Source
<i>E. coli</i> BL21 (DE3)	F <sup>-</sup> <i>ompT hsdS<sub>B</sub> (r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm</i> (DE3)	Merck
<i>E. coli</i> Top10	F <sup>-</sup> <i>mcrA Δ(mrr-hsdRMS-mcrBC)φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu) 7697 galE15 galK16 rpsL (Str<sup>R</sup>) endA1 nupG λ-</i>	Our lab
pET28a	T7/lac promoter, kanamycin resistance	Merck
pET28a-Trx27	pET28a harboring thioredoxin gene	Our lab

FastDigest *EcoR* I, FastDigest *Sac* I, dNTPs, Phusion High-Fidelity DNA Polymerase 购自 Thermofisher 公司。超滤离心管购自 Merck 公司。

### 1.3 载体骨架 pET28a-Trxm 构建

#### 1.3.1 pET28a: 酶切位点 *Apa* I 破坏

根据 pET28a 载体上需要破坏“失活”的酶切位点 *Apa* I (1 334), 参考 QuickChange 定点突变方法<sup>[18-19]</sup>, 设计引物对 g1331a\_c1334a (表 2) 并以 pET28a 为模板进行 PCR。PCR 反应总体积为 50  $\mu$ L, 具体为: 10  $\mu$ mol/L 正向引物 1.5  $\mu$ L, 10  $\mu$ mol/L 反向引物 1.5  $\mu$ L, 模板 pET28a 10–30 ng, 2 $\times$ Phusion PCR mix 25  $\mu$ L, 补水至 50  $\mu$ L。PCR 程序为: 98  $^{\circ}$ C 预变性 10 s; 98  $^{\circ}$ C 10 s; 72  $^{\circ}$ C 4 min 30 s, 30 个循环; 72  $^{\circ}$ C 延伸 10 min。所得 PCR 产物用 *Dpn* I 于 37  $^{\circ}$ C 消化 1 h 后用乙醇沉淀法回收目的质粒 DNA<sup>[20]</sup>。随后将 DNA 沉淀重溶于 5  $\mu$ L 超纯水, 热激法转化 *E. coli* Top10 感受态细胞, 37  $^{\circ}$ C 过夜培养得到转化子。挑取若干转化子扩增过夜, 离心收集菌体用试剂盒提取质粒后, 委托生工生物工程 (上海) 股份有限公司测序并获得含目的突变的质粒 pET28a(g1331a/c1334a)。

#### 1.3.2 pET28a: *Apa* I 以及新终止密码子 TGA 引入

为了实现 RCA 策略, 需要将 pET28a 上的多克隆位点 (multiple cloning site, MCS)

引入 *Apa* I (*Not* I 突变为 *Apa* I), 同时插入得到新终止密码子 TGA。此突变可通过同一引物对 c153g\_g155c\_c160t 实现 (表 2), 模板为 pET28a(g1331a/c1334a), 具体突变方法同 1.3.1, 获得含目的突变的质粒 pET28a(g1331a/c1334a/c153g/g155c/c160t)。

#### 1.3.3 pET28a-Trx27: 5'端酶切位点 *Nco* I 引入

因短肽串联表达本身分子量较小, 故将其 5'端构建引入硫氧还蛋白 (thioredoxin, Trx) 作为融合蛋白, 并通过肠激酶酶切位点连接至短肽, 其结构式为: 硫氧还蛋白-酶切位点-(短肽) $n$ ,  $n$  为  $\geq 1$  的自然数。pET28a-Trx27 为本实验室前期构建的用于表达多肽“27”的重组质粒, 其融合蛋白为 Trx, 拟通过双酶切获得 DNA 片段。因其 5'端为无缝克隆, 故通过 QuickChange 方法将 5'端引入酶切位点 *Nco* I, 便于后续实验。此步突变模板为 pET28a-Trx27, 引物对 TA\_ins3nt\_62 (表 2), 具体突变方法同 1.3.1, 获得目的质粒 pET28a-Trx27(62ggt)。

#### 1.3.4 载体骨架 pET28a-Trxm 的获得

分别将质粒 pET28a(g1331a/c1334a/c153g/g155c/c160t) 和 pET28a-Trx27(62ggt) 进行 *Nco* I 和 *EcoR* I 双酶切, 37  $^{\circ}$ C 酶切 30 min。酶切结束后将产物进行 1% 琼脂糖凝胶电泳, 电泳条件为 100 V, 40 min。电泳结束后分别切胶回收, 并获得线性化载体和目的 DNA 片段。将

表 2 文中所用引物

Table 2 Primers used in this study

Primer name	Primer sequence (5'→3')	Purification method
g1331a_c1334a_f	CAAATCGCGCTGTTAGCGAGCACATTAAGTTCTGTCTCGG	PAGE
g1331a_c1334a_r	CCGAGACAGAACTTAATGTGCTCGCTAACAGCGCGATTTG	PAGE
c153g_g155c_c160t_f	GACAAGCTTGC <sup>GGGCC</sup> CACTTGAGCACCACCACC	PAGE
c153g_g155c_c160t_r	GGTGGTGGTGCTCAAGTGGGCCCGCAAGCTTGTC	PAGE
TA_ins3nt_62_f	AAGAAGGAGATATACCATGGGTATCATCGAATACGATGGCG	PAGE
TA_ins3nt_62_r	CGCCATCGTATTCGATGATACCCATGGTATATCTCCTTCTT	PAGE

The mutagenesis sites were indicated as bold letters.

载体和目的 DNA 片段于 16 °C 连接 30 min, 连接产物热激法转化 *E. coli* Top10, 37 °C 培养过夜得到转化子。转化子过夜扩增并提取质粒后, 将质粒送出测序并比对, 保存正确的克隆。

#### 1.4 短肽串联重复 (tandem repeat of short peptides, TRSP) 基本单元的设计及模板构建

1 个 TRSP 基本单元的基因片段长度为 96 bp, 委托生工生物工程 (上海) 股份有限公司进行基因合成并构建至 pUC57。此基因片段包含 5 个酶切位点 (图 1), 以实现串联重复拷贝数的随机化。具体为: 其 5' 端和 3' 端各具有 1 个 *Sac* I 位点, 用于片段自身环化。*Acc65* I 和 *Apa* I 位点为 TRSP 基本单元的克隆位点, (TRSP)<sub>n</sub> 通过这两个位点构建至表达载体上。因 *Acc65* I 和 *Apa* I 均为 Dcm 甲基化敏感, 当片段中引入 5mC 代替胞嘧啶时切割就会被阻断。*Acc65* I 位点紧邻 5' 端 *Sac* I 之后, *Apa* I 位点之后紧接着 3' 端 *Sac* I。*Bbv*C I 位点紧接 *Acc65* I 之后, nt. *Bbv*C I 酶切此位点可产生“切口”, 从而作为 Phi29 DNA 聚合酶的“引物”位点。紧接 *Bbv*C I 位点之后 (下划线双实线部分) 为编码短肽 GHK 和 GQPR 的串联重复短肽基因序列。

质粒 pUC57-TRSP 经 *Sac* I 酶切后, 胶回收 TRSP 片段, 然后用 DNA 连接试剂盒中的

试剂 I 进行连接环化, 之后 65 °C 水浴 10 min 热灭活酶, 然后使用 nt. *Bbv*C I 将环状 DNA 切割产生“切口”, 之后 80 °C 水浴 20 min 热灭活酶。

#### 1.5 RCA 反应

随后进行 RCA 反应, 50 μL 反应体系具体如下: 带有“切口”的环状 TRSP 片段 10 μL, 1 mmol/L dATP, 1 mmol/L dGTP, 1 mmol/L dTTP, 0.5 mmol/L dCTP, 0.5 mmol/L 5mC, 0.25 U/μL phi29 DNA 聚合酶, 0.1 μg/μL 牛血清白蛋白 (bovine serum albumin, BSA)。30 °C 孵育 19 h, 孵育结束后 65 °C 水浴 10 min 热灭活酶。

#### 1.6 含不同长度 TRSP 的表达质粒及重组工程菌构建

RCA 反应产物经 *Acc65* I 和 *Apa* I 消化后, 因 5mC 的掺入而产生不同长度的 TRSP 片段 (TRSP)<sub>n</sub>, 进行 1% 琼脂糖凝胶电泳后, 用胶回收试剂盒纯化获得长度约 500 bp 到 1 500 bp 的 TRSP 片段混合物。载体 pET28a-Trxm 经 *Acc65* I 和 *Apa* I 消化后, 胶回收线性化载体。将回收获得的不同长度的 (TRSP)<sub>n</sub> 片段与线性化载体 pET28a-Trxm 用 Solution I 于 16 °C 连接 30 min, 随后热激法转化感受态细胞 *E. coli* BL21(DE3), 获得含不同长度插入片段的单克隆 *E. coli* BL21(DE3)/pET28a-Trxm-(TRSP)<sub>n</sub>。

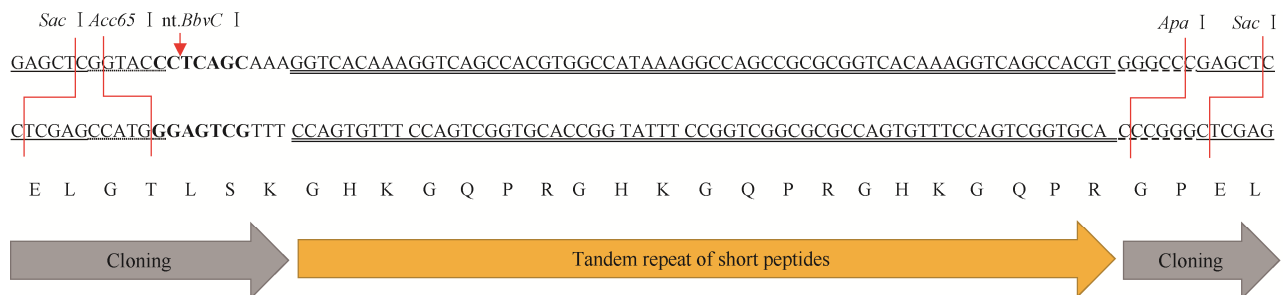


图 1 TRSP 的基本单元 ( $n=1$ )

Figure 1 TRSP unit ( $n=1$ ). *Sac* I was denoted as “\_\_\_”. *Acc65* I was denoted as “.....”. The tandem repeat sequences encoding peptide GHK and GQPR were denoted as “\_\_\_”. *Apa* I was denoted as “\_ \_ \_”.

随机挑取若干 *E. coli* 单克隆并分别接种于含硫酸卡那霉素的 2 mL LB 培养基, 37 °C、700 r/min 扩增过夜。培养液 12 000 r/min 离心 3 min 收集菌体, 然后用质粒提取试剂盒提取质粒后, 取少量质粒用 *Nco* I、*Apa* I 于 37 °C 酶切 20 min, 酶切产物进行 1% 琼脂糖凝胶电泳验证。将验证正确的质粒送出测序。

### 1.7 重组融合蛋白 Trxm-(TRSP)<sub>n</sub> 表达

将序列正确的单克隆接种于 2 mL 含硫酸卡那霉素的 LB 培养基, 37 °C、700 r/min 振荡培养过夜。将培养过夜的新鲜种子以 1% (*V/V*) 接种量接种于 2 mL 的 LB 培养基, 37 °C、700 r/min 振荡培养至 *OD*<sub>600</sub> 约为 0.6, 再加入 IPTG 至 0.4 mmol/L, 于 37 °C、700 r/min 诱导 6 h。诱导结束后将培养液离心 (12 000 r/min) 3 min 收集菌体, 用去离子水洗涤一遍, 再加入一定体积的 5×上样缓冲液混匀, 100 °C 煮沸 5 min 制备样品。所得样品进行 SDS-PAGE 分析。

### 1.8 酶切重组融合蛋白 Trxm-(TRSP)<sub>1</sub> 获得目的短肽

将工程菌 *E. coli* BL21(DE3)/pET28a-Trxm-(TRSP)<sub>1</sub> 接种于 500 mL LB 培养基并诱导表达, 具体方法同 1.7。培养结束后离心收集菌体, 并将其重悬于 50 mL 的提取缓冲液 (20 mmol/L Tris-HCl, 200 mmol/L NaCl, pH 8.0) 并进行超声破碎, 超声条件为: 功率 200 W, 工作 3 s, 间隔 7 s, 重复 60 次。超声后悬液离心 (14 000 r/min) 10 min 收集上清。超声上清液用分子量 30 kDa 的超滤离心管进行超滤浓缩后, 超滤内液中加入 20 mmol/L 的 Tris-HCl 缓冲液 (pH 8.0) 离心洗涤 3 遍, 得到 15 mL 含重组融合蛋白 Trxm-(TRSP)<sub>1</sub> 的上清液。在此上清液中加入 16 U 胰激酶, 于 37 °C 水浴 4 h 进行酶切反应, 酶切结束后反应液用分子量 5 kDa

的超滤离心管进行超滤, 并用相同的 Tris-HCl 缓冲液洗涤后, 超滤内液加入 32 U 的胰蛋白酶, 于 25 °C 条件下酶切 3 h 后将反应液离心, 取上清进行 HPLC 分析。

## 2 结果与分析

### 2.1 载体骨架 pET28a-Trxm 的构建

为了实现不同拷贝数的短肽 GHK 和 GQPR 基因串联重复策略, 需要在目的基因 (TRSP)<sub>n</sub> 片段两端引入对于 *Dcm* 甲基化敏感的酶切位点 *Acc65* I 和 *Apa* I, 通过这两个位点构建至表达载体, 当基因片段中引入 5mC 代替胞嘧啶时切割就会被阻断。本研究以 pET28a 为出发质粒, 通过定点突变将 *Apa* I (1 334) 破坏掉, 随后在多克隆位点 (multiple cloning site, MCS) 引入 *Apa* I, 获得了突变后的 pET28a(g1331a/c1334a/c153g/g155c/c160t)。不同拷贝数的短肽 GHK 和 GQPR 基因在 *E. coli* 中串联表达时仍然存在分子量较小而不稳定的问题, 故本研究采取与硫氧还蛋白融合表达的方式进行表达。以 pET28a-Trx27 为出发质粒, 先在 Trx27 的 5' 引入 *Nco* I 位点, 获得 pET28a-Trx27(62ggt)。pET28a(g1331a/c1334a/c153g/g155c/c160t) 和 pET28a-Trx27(62ggt) 分别用 *Nco* I 和 *Eco*R I 双酶切, 胶回收获得线性化载体和 Trxm 片段, 随后连接并转化感受态细胞, 获得序列正确的载体骨架 pET28a-Trxm。构建过程示意图见图 2。

### 2.2 RCA 反应获得不同长度的 (TRSP)<sub>n</sub> 片段

图 3 阐释了串联重复片段的构建流程。具体为: (1) pUC57-TRSP 经 *Sac* I 酶切得到线性的 TRSP 片段; (2) 线性 TRSP 片段再经 Solution I 连接得到环状 TRSP 分子; (3) 环状 TRSP 经 *nt.Bbv*C I 切割产生“切口”; (4) 带“切

口”的 TRSP 作为模板, 使用 5mC 代替 1/2 的 dCTP 进行滚环扩增, 扩增时 5mC 随机掺入至序列; (5) 滚环扩增反应产物经 *Acc65 I* 和 *Apa I* 双酶切产生不同长度的  $(\text{TRSP})_n$  片段混合物;

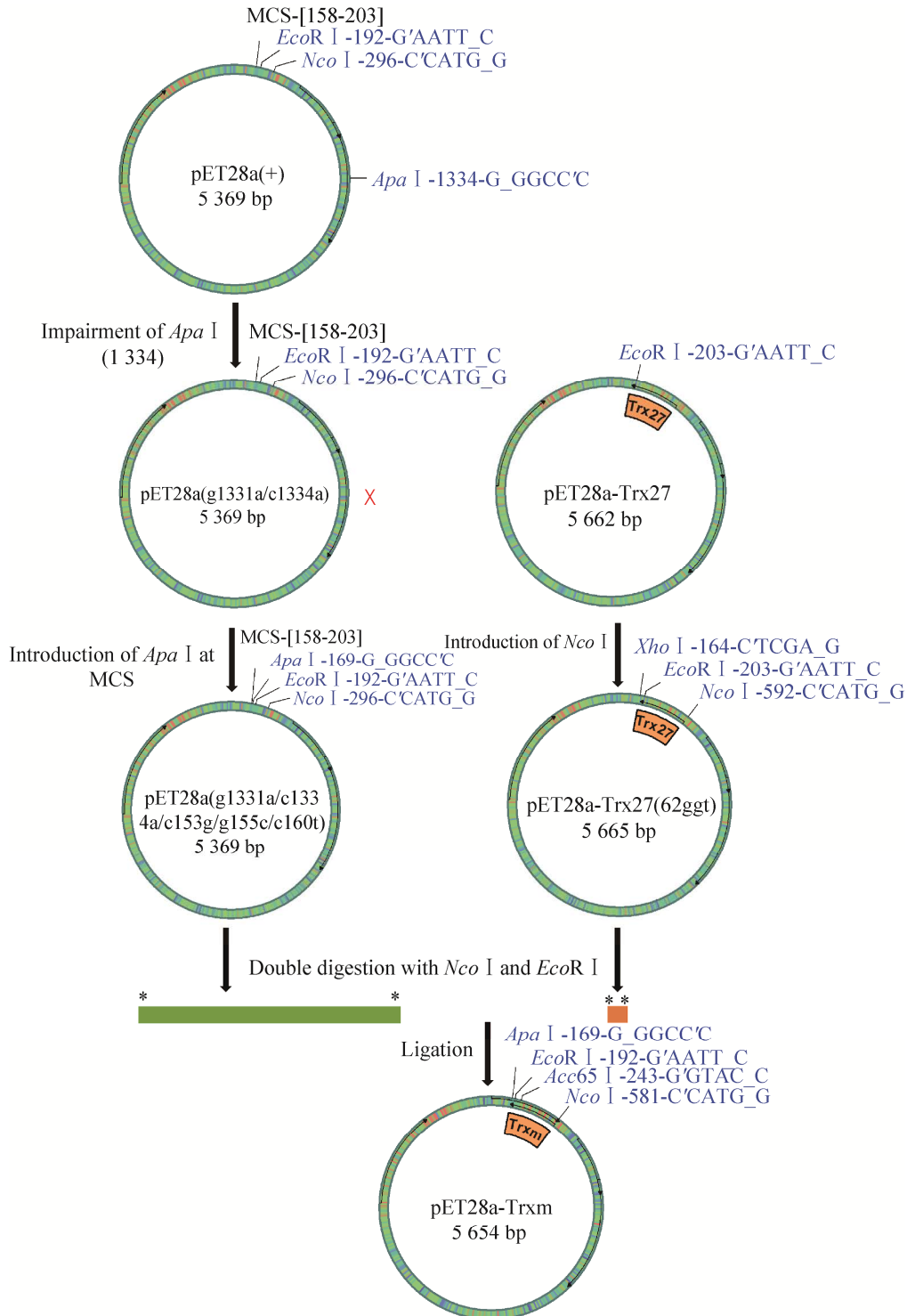


图 2 载体骨架 pET28a-Trxm 构建示意图

Figure 2 Illustration of pET28a-Trxm construction.

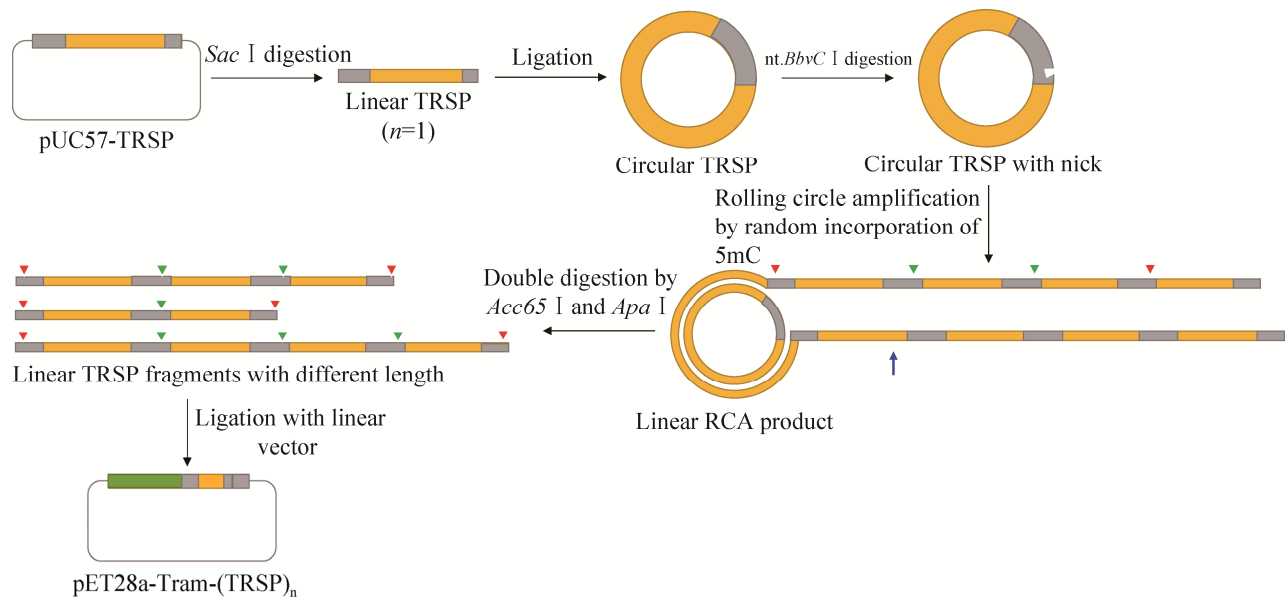


图3 pET28a-Tram-(TRSP)<sub>n</sub> 的构建流程

Figure 3 Procedures of constructing pET28a-Tram-(TRSP)<sub>n</sub>. The tandem repeat sequences encoding peptide GHK and GQPR were denoted as the yellow box. Cloning regions were denoted as the grey box. Locations where dCTP and 5mC were involved were denoted as red and green triangle, respectively. Linear RCA product was denoted as the blue arrow.

(6) (TRSP)<sub>n</sub> 片段混合物与线性 pET28a-Trxm 连接得到重组表达载体 pET28a-Trxm-(TRSP)<sub>n</sub>。经热激转化获得含不同长度插入片段的单克隆 *E. coli* BL21(DE3)/pET28a-Trxm-(TRSP)<sub>n</sub>。

RCA 产物经 *Acc65* I 和 *Apa* I 消化后进行 1% 琼脂糖凝胶电泳, 结果显示无清晰条带出现且整个泳道呈弥散状态 (图 4), 可能是由于 RCA 反应时 5mC 掺入序列的随机性而导致条带分子量的随机性。切胶回收分子量约 500 bp 到 1 500 bp 的 DNA 片段并克隆至表达载体, 相应的蛋白质分子量约为 18 kDa 到 55 kDa。

### 2.3 (TRSP)<sub>n</sub> 基因序列鉴定

将随机挑取的若干 *E. coli* 单克隆扩增并提取质粒后, 质粒经 *Nco* I、*Apa* I 双酶切结果如图 5 所示。所有泳道均出现目的条带 Trxm-(TRSP)<sub>n</sub>, 证明其成功构建至表达载体。此片段分子量大部分位于 750 bp 以下, 其中泳道 21 所对应的 Trxm-(TRSP)<sub>n</sub> 片段分子量约为

3 500 bp。此外, 经双酶切后, 部分质粒出现了 2 个除线性化质粒外的小分子量条带, 这是由于 TRSP 基本单元本身包含 1 个 *Apa* I 酶切位点 GGGCCC, 当滚环扩增过程产生  $n$  个 ( $n > 1$ ) 串联重复时, 此序列就相应地包含  $n$  个 *Apa* I 位点。测序结果显示 6#、7# 编码区长

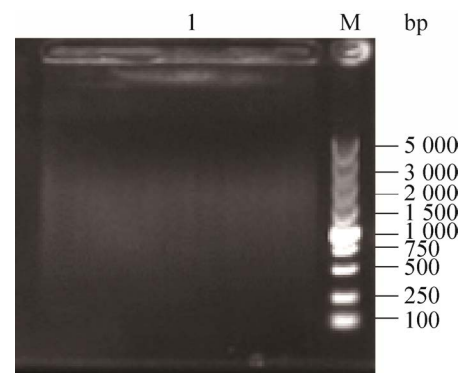


图4 RCA 产物经 *Acc65* I 和 *Apa* I 双酶切后电泳图

Figure 4 Double digestion of RCA product by *Acc65* I and *Apa* I.



度为 873 bp, 12#长度为 1 053 bp, 这也与双酶切结果相符。但是, 也有部分单克隆在融合蛋白 Trxm 的基因序列正确的情况下, 在滚环扩增反应时出现错误, 导致串联重复短肽 (TRSP)<sub>n</sub> 部分的序列与预期有较大差异。在 TRSP 基本单元 ( $n=1$ ) 中 (图 1), 已经包含 GHKGQPR 基因序列的 3 个串联重复, 故在滚环扩增过程时, 每增加 1 个 TRSP 的拷贝数, 相当于增加了 3 个 GHKGQPR 的串联重复。而串联重复基因序列在细胞中复制和维持时本身就具有不稳定性<sup>[4]</sup>, 这可能也是滚环扩增时

出错的原因之一。

将单克隆的测序结果进行分析, 结果显示, 通过滚环扩增反应时随机掺入 5mC 成功实现了短肽基因序列的串联重复, 获得了不同串联重复数的重组融合蛋白编码基因,  $n=1、2、3、4、6、7、8、9$ 。随着串联重复数  $n$  的增加, Trxm-(TRSP)<sub>n</sub> 基因序列长度从 423 bp 逐渐增加至 1 143 bp, 相应的氨基酸个数从 141 增加至 381, 理论蛋白分子量从 15.94 kDa 增加至 41.27 kDa (表 3)。每增加 1 个拷贝数 ( $n=1$ ), 基因序列就增加 90 bp, 相应的氨基酸个数增加 30 个。

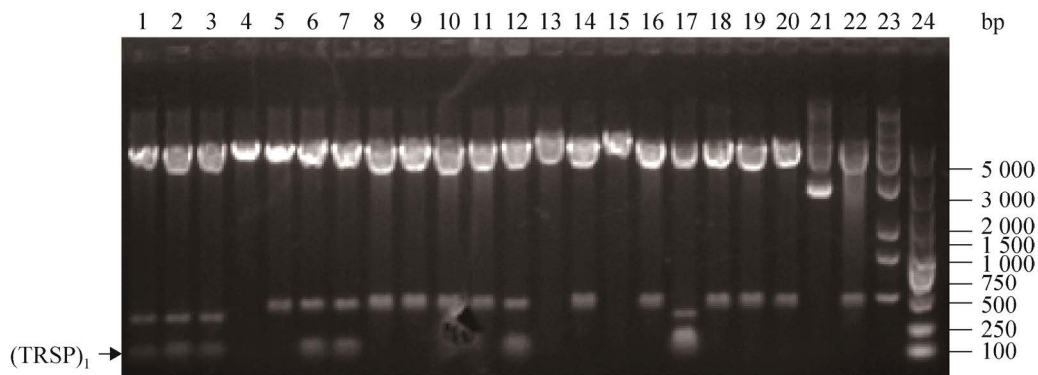


图 5 质粒 pET28a-Trxm-(TRSP)<sub>n</sub> 双酶切验证

Figure 5 Double digestion of pET28a-Trxm-(TRSP)<sub>n</sub>. Lane 1–22: single clones; lane 24: marker DL5000; TRSP units generated by *Apa* I digestion were denoted as the arrow, which suggests  $n>1$  of the tandem repeat number.

表 3 基因序列分析

Table 3 Analysis of gene sequence

Protein	Gene length (bp)	Number of amino acids	Theoretical molecular weight (kDa)	TRSP unit number ( $n$ )
Trxm-(TRSP) <sub>1</sub>	423	141	15.94	1
Trxm-(TRSP) <sub>2</sub>	513	171	19.11	2
Trxm-(TRSP) <sub>3</sub>	603	201	22.27	3
Trxm-(TRSP) <sub>4</sub>	693	231	25.44	4
Trxm-(TRSP) <sub>6</sub>	873	291	31.77	6
Trxm-(TRSP) <sub>7</sub>	963	321	34.93	7
Trxm-(TRSP) <sub>8</sub>	1 053	351	38.10	8
Trxm-(TRSP) <sub>9</sub>	1 143	381	41.27	9

## 2.4 重组融合蛋白 Trxm-(TRSP)<sub>n</sub> 的表达

将含不同串联重复数的单克隆 *E. coli* BL21(DE3)/pET28a-Trxm-(TRSP)<sub>n</sub> 分别培养并诱导表达, 诱导结束菌体总蛋白经 SDS-PAGE

分析结果如下图所示。当 (TRSP)<sub>n</sub> 串联重复数  $n=1$ 、 $n=2$  时, 分别在分子量 16 kDa 和 19 kDa 处有高表达条带, 与预期相符合 (图 6A, 红色箭头)。Trxm-(TRSP)<sub>1</sub> 表达水平约为 *E. coli* 总蛋

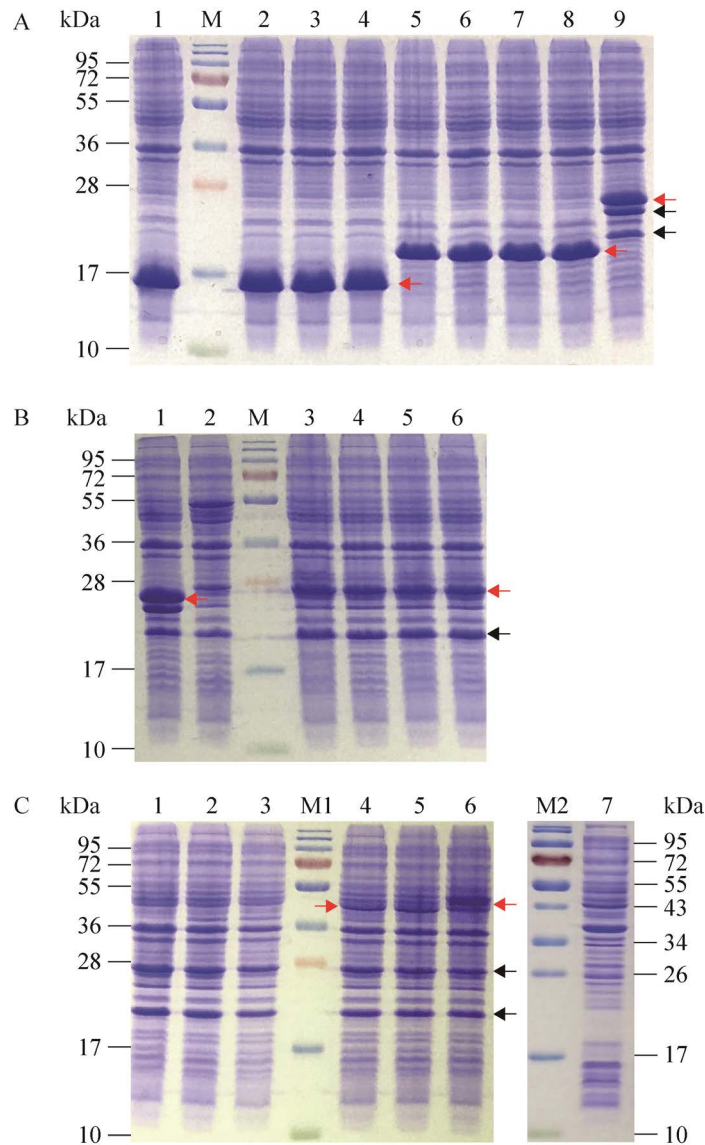


图 6 *E. coli* 菌体总蛋白 SDS-PAGE 分析

Figure 6 SDS-PAGE analysis of total protein from *E. coli* cells. (A) Lane 1–4: Trxm-(TRSP)<sub>1</sub>,  $n=1$ ; lane 5–8: Trxm-(TRSP)<sub>2</sub>,  $n=2$ ; lane 9: Trxm-(TRSP)<sub>3</sub>,  $n=3$ ; lane M: marker 26619. (B) Lane 1: Trxm-(TRSP)<sub>3</sub>,  $n=3$ ; lane 2–6: Trxm-(TRSP)<sub>4</sub>,  $n=4$ ; lane M: marker 26619. (C) Lane 1–2: Trxm-(TRSP)<sub>6</sub>,  $n=6$ ; lane 3: Trxm-(TRSP)<sub>7</sub>,  $n=7$ ; lane 4–5: Trxm-(TRSP)<sub>8</sub>,  $n=8$ ; lane 6: Trxm-(TRSP)<sub>9</sub>,  $n=9$ ; lane 7: *E. coli* BL21(DE3)/pET28a (negative control); lane M1: marker 26619; lane M2: marker 26616. Target bands were indicated as the red arrows. Unexpected overexpression bands were indicated as the black arrows.

白的 40%–60%，Trxm-(TRSP)<sub>2</sub> 表达水平约为总蛋白 30%–40%。当 (TRSP)<sub>n</sub> 串联重复数  $n=3$ 、 $n=4$  时，在预测分子量 22 kDa 和 25 kDa 处有目的条带表达，除此之外， $n=3$  时在 19 kDa 到 22 kDa 之间另有两条非预期分子量的过表达条带 (图 6A 和 6B，黑色箭头)， $n=4$  时在 19 kDa 到 22 kDa 之间另有 1 条过表达条带 (图 6B，黑色箭头)；当  $n=6$ 、 $n=7$  时，在预测分子量 32 kDa 和 35 kDa 处并无目的蛋白表达，当  $n=8$ 、 $n=9$  时，分别在预期分子量 38 kDa 和 41 kDa 处有较低水平的蛋白过表达，推测可能为目的蛋白表达 (图 6C，红色箭头)。 $n=6$ 、 $n=7$ 、 $n=8$ 、 $n=9$  时均有杂蛋白条带出现 (图 6C，黑色箭头)。

在表达 Trxm-(TRSP)<sub>n</sub> 时，当串联重复数  $n>2$  时，均会产生非预期蛋白的过表达，且均有 1 条相同分子量的蛋白条带出现。这可能是在表达串联重复数较高的基因序列时，大肠杆菌的一种自我应对机制，也可能是由于融合蛋白 Trxm-(TRSP)<sub>n</sub> 的胞内降解。从上述结果可看出，当  $n=1$  和  $n=2$  时，目的蛋白 Trxm-(TRSP)<sub>1</sub> 和 Trxm-(TRSP)<sub>2</sub> 表达水平较高，而随着基因串联重复数的增加，Trxm-(TRSP)<sub>n</sub> 的表达水平逐渐降低。 $n=6$  和  $n=7$  时未见目的蛋白表达，同时在  $n=8$  和  $n=9$  时分别在预期分子量处有较低水平的蛋白过表达。考虑到融合蛋白表达水平以及其他杂质蛋白的影响，故选取  $n=1$  进行后续实验。

## 2.5 目的短肽混合物获得

将诱导表达后的工程菌 ( $n=1$ ) 菌体经过超声破碎、超滤浓缩、顶洗等步骤获得含重组融合蛋白 Trxm-(TRSP)<sub>1</sub> 的上清液。将此上清液经肠激酶酶切 4 h 后进行 SDS-PAGE 分析。图 7 中泳道 2 可看出，酶反应 0 h 时，反应液中主

要条带为融合蛋白 Trxm-(TRSP)<sub>1</sub> (16 kDa)。泳道 1 为酶切 4 h 的反应液，可看出在酶切 4 h 后，反应液中已无融合蛋白 Trxm-(TRSP)<sub>1</sub> 条带，而相应地出现了一个分子量为 12 kDa 的新条带，与 Trxm 理论分子量一致，从而推测 Trxm-(TRSP)<sub>1</sub> 被完全切割为硫氧还蛋白 Trxm 和短肽片段 (TRSP)<sub>1</sub>。此外，并未观察到 (TRSP)<sub>1</sub> 的相应条带 (理论分子量为 3 kDa)，这可能是由于小于 5 kDa 的蛋白条带在 SDS-PAGE 时极易弥散。肠激酶酶切结束后，反应液再经超滤浓缩等步骤获得含 (TRSP)<sub>1</sub> 的清液。随后将此清液进行胰蛋白酶切割反应，同时对反应液进行 HPLC 分析。通过与市售 GHK 及 GQPR 对照品的 HPLC 图谱相对比，显示经酶切后的反应液中出现了新峰，且与对照品的保留时间一致 (图 8)。此结果表明，反应液中初步生成了目的短肽 GHK 和 GQPR。

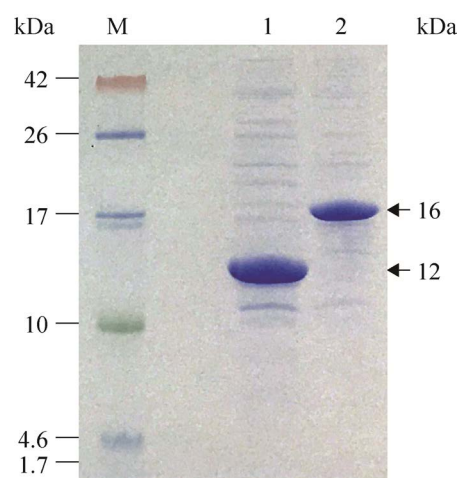


图 7 重组 Trxm-(TRSP)<sub>1</sub> 经肠激酶酶切 SDS-PAGE 结果

Figure 7 Recombinant Trxm-(TRSP)<sub>1</sub> cleavage by enterokinase. Lane M: marker 26628; lane 1: reaction mixture after 4 hours of cleavage; lane 2: reaction mixture at the beginning. Target bands were indicated as the arrows.

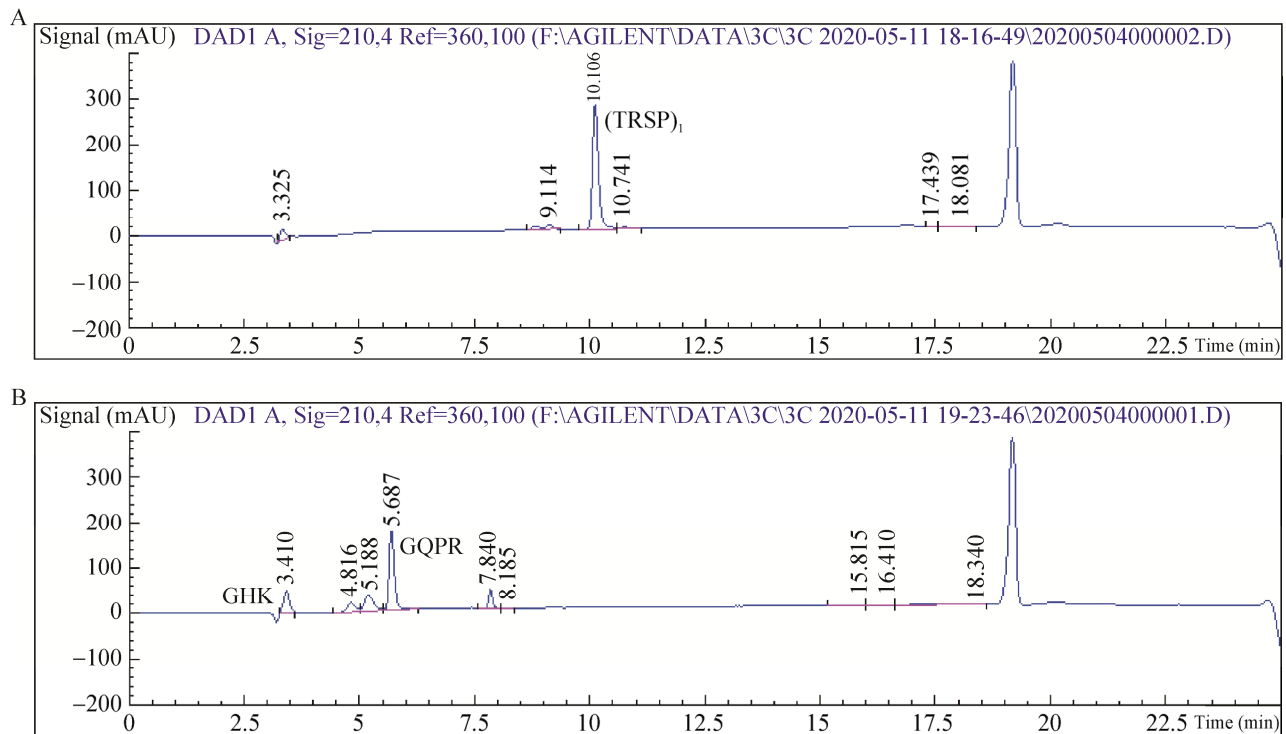


图8 (TRSP)<sub>1</sub>酶切反应HPLC分析

Figure 8 HPLC analysis of (TRSP)<sub>1</sub> enzymatic cleavage. Reaction mixture at 0 h (A) and 3 h (B).

### 3 讨论

随着社会发展,人们对于健康的要求越来越高,包括拥有健康且年轻的皮肤。近年来化妆品市场的快速增长清楚地印证了这点,含活性成分的化妆品愈发具有开发前景,而应用最为广泛的活性成分之一就是超短肽<sup>[21]</sup>。超短肽GHK和GQPR均为信号肽,GHK能够增强胶原蛋白的产生且具有较强的促伤口愈合和组织修复作用,GQPR具有抗炎以减缓胶原蛋白降解的作用。法国Sederma公司开发出Pal-GHK和Pal-GQPR的混合物作为抗衰老明星成分Matrixyl™ 3000,可刺激胶原蛋白产生的同时不会引起皮肤刺激性<sup>[22-23]</sup>。本研究通过RCA法实现了GHK和GQPR基因串联重复数的确定。我们通过在RCA过程中随机掺入非典型核苷酸,从而获得不同长度的串联重复基因,并与

硫氧还蛋白融合表达,通过对比蛋白表达水平确定基因的串联重复数,从而获得高表达水平的工程菌。当重复数 $n=1$ 和 $n=2$ 时,(TRSP)<sub>1</sub>和(TRSP)<sub>2</sub>的表达水平分别达到菌体总蛋白的50%和30%。随着基因串联重复数的增加,重组融合蛋白的表达水平逐渐降低,且会产生杂质蛋白的过表达。有报道通过RCA法构建并生产特定分子量范围的蛋白,用以研究蛋白序列-结构-功能间的关系<sup>[17]</sup>,而目前尚未见其他报道用此方法来进行超短肽基因串联重复数的筛选。总之,应用本方法可实现超短肽串联表达时基因重复数的筛选。

短肽通常是用固相合成法来制备,但此方法的缺点是制备过程中大量使用有机化学试剂而带来一定的环境负担。因序列较短,故一般无法直接用重组策略来生产短肽<sup>[1,21]</sup>。本研究实现了重组法制备超短肽GHK和GQPR。工程菌

*E. coli* BL21(DE3)/pET28a-Trxm-(TRSP)<sub>1</sub> 经培养获得菌体, 菌体经超声破碎获得重组融合蛋白 Trxm-(TRSP)<sub>1</sub> 粗提液, 再经超滤浓缩等步骤获得含重组融合蛋白的上清液。此上清液先后经肠激酶和胰蛋白酶酶切后, 获得超短肽 GHK 和 GQPR 混合物。此混合物可作为活性成分直接添加于化妆品中, 也可经过后续的纯化步骤分别得到 GHK 和 GQPR。我们进一步探讨了工程菌 *E. coli* BL21(DE3)/pET28a-Trxm-(TRSP)<sub>1</sub> 30 L 发酵罐的放大实验, 通过发酵、提取、酶切及分离纯化等步骤, 最后分别获得 3.6 g 纯化的 GHK 和 4.9 g 纯化的 GQPR。目前尚未发现有其他研究用重组策略制备 GHK 和 GQPR 混合物。本研究成功建立了重组策略制备超短肽的方法, 与固相合成法相比大大减少了化学试剂使用, 更为环境友好, 且初步具备工业化特征。然而, 本研究仅是初步探讨了超短肽的串联表达及分离纯化方法, 后续可通过研究 *E. coli* 高密度发酵、破胞提取方式、膜过滤及离子交换色谱等深入探讨大规模制备 GHK 和 GQPR 的方法。此外, 也可尝试将此方法应用于其他生物技术领域的超短肽, 探索其重组制备策略。

## REFERENCES

- [1] Apostolopoulos V, Bojarska J, Chai TT, Elnagdy S, Kaczmarek K, Matsoukas J, New R, Parang K, Lopez OP, Parhiz H, et al. A global review on short peptides: frontiers and perspectives *Molecules*, 2021, 26(2): E430.
- [2] Hauser CAE, Deng RS, Mishra A, Loo Y, Khoe U, Zhuang FR, Cheong DW, Accardo A, Sullivan MB, Riekel C, et al. Natural tri- to hexapeptides self-assemble in water to amyloid beta-type fiber aggregates by unexpected alpha-helical intermediate structures. *PNAS*, 2011, 108(4): 1361-1366.
- [3] Seow WY, Hauser CAE. Short to ultrashort peptide hydrogels for biomedical uses. *Materials Today*, 2014, 17(8): 381-388.
- [4] Makovitzki A, Avrahami D, Shai Y. Ultrashort antibacterial and antifungal lipopeptides. *PNAS*, 2006, 103(43): 15997-16002.
- [5] Ni M, Zhuo SM. Applications of self-assembling ultrashort peptides in bionanotechnology. *RSC Adv*, 2019, 9(2): 844-852.
- [6] Haggag YA. Peptides as drug candidates: limitations and recent development perspectives. *Biomed J*, 2018, 8(4): 6659-6662.
- [7] Chan KH, Xue B, Robinson RC, Hauser CAE. Systematic moiety variations of ultrashort peptides produce profound effects on self-assembly, nanostructure formation, hydrogelation, and phase transition. *Sci Rep*, 2017, 7: 12897.
- [8] Rahnamaeian M, Vilcinskis A. Short antimicrobial peptides as cosmetic ingredients to deter dermatological pathogens. *Appl Microbiol Biotechnol*, 2015, 99(21): 8847-8855.
- [9] Pickart L. The human tri-peptide GHK and tissue remodeling. *J Biomater Sci Polym Ed*, 2008, 19(8): 969-988.
- [10] Watson RE, Ogden S, Cotterell LF, Bowden JJ, Bastrilles JY, Long SP, Griffiths CE. Effects of a cosmetic 'anti-ageing' product improves photoaged skin corrected. *Br J Dermatol*, 2009, 161(2): 419-426.
- [11] Schagen S. Topical peptide treatments with effective anti-aging results. *Cosmetics*, 2017, 4(2): 16.
- [12] Mondon P, Hillion M, Peschard O, Andre N, Marchand T, Doridot E, Feuilloley MG, Pionneau C, Chardonnet S. Evaluation of dermal extracellular matrix and epidermal-dermal junction modifications using matrix-assisted laser desorption/ionization mass spectrometric imaging, *in vivo* reflectance confocal microscopy, echography, and histology: effect of age and peptide applications. *J Cosmet Dermatol*, 2015, 14(2): 152-160.
- [13] Mohsen MG, Kool ET. The discovery of rolling circle amplification and rolling circle transcription. *Acc Chem Res*, 2016, 49(11): 2540-2550.
- [14] Ali MM, Li F, Zhang ZQ, Zhang KX, Kang DK, Ankrum JA, Le XC, Zhao WA. Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. *Chem Soc Rev*, 2014, 43(10): 3324-3341.

- [15] Yue SZ, Li YW, Qiao ZJ, Song WL, Bi S. Rolling circle replication for biosensing, bioimaging, and biomedicine. *Trends Biotechnol*, 2021, 39(11): 1160-1172.
- [16] 刘淑贞, 张志庆, 王芳, 周亭, 王秀凤, 张国栋, 刘婷婷, 张洪芝. DNA 的滚环扩增合成研究. *物理化学学报*, 2017, 33(10): 2052-2057.  
Liu SZ, Zhang ZQ, Wang F, Zhou T, Wang XF, Zhang Guodong, Liu TT, Zhang HZ. Study on the synthesis of DNA via rolling circle amplification. *Acta Phys -Chim Sin*, 2017, 33(10): 2052-2057 (in Chinese).
- [17] Jung H, Pena-Francesch A, Saadat A, Sebastian A, Kim DH, Hamilton RF, Albert I, Allen BD, Demirel MC. Molecular tandem repeat strategy for elucidating mechanical properties of high-strength proteins. *PNAS*, 2016, 113(23): 6478-6483.
- [18] Liu HT, Naismith JH. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol*, 2008, 8: 91.
- [19] Novoradovsky A, Zhang V, Ghosh M, et al. Computational principles of primer design for site directed mutagenesis. 2005 NSTI Nanotechnol Conf Trade Show NSTI Nanotech 2005 Tech Proc, 2005: 532-535.
- [20] Green MR, Sambrook J. *Molecular Cloning: A Laboratory Manual*. 4th ed. New York: Cold Spring Harbor Laboratory Press, 2012: 17-20.
- [21] Errante F, Ledwoń P, Latajka R, Rovero P, Papini AM. Cosmeceutical peptides in the framework of sustainable wellness economy. *Front Chem*, 2020, 8: 572923.
- [22] Johnson W Jr, Bergfeld WF, Belsito DV, Hill RA, Klaassen CD, Liebler DC, Marks JG Jr, Shank RC, Slaga TJ, Snyder PW, et al. Safety assessment of tripeptide-1, hexapeptide-12, their metal salts and fatty acyl derivatives, and palmitoyl tetrapeptide-7 as used in cosmetics. *Int J Toxicol*, 2018, 37(3\_suppl): 90S-102S.
- [23] Dou Y, Lee A, Zhu LD, Morton J, Ladiges W. The potential of GHK as an anti-aging peptide. *Aging Pathobiol Ther*, 2020, 2(1): 58-61.

(本文责编 陈宏宇)