

微生物学报 *Acta Microbiologica Sinica*
54(9):998-1009; 4 September 2014
ISSN 0001-6209; CN 11-1995/Q
http://journals.im.ac.cn/actamicroen
doi: 10.13343/j.cnki.wsxb.2014.09.005

基于 PHA 颗粒-PhaP 标签和内含肽元件的极端嗜盐古菌蛋白的表达纯化系统

伍锦花^{1,2}, 蔡双凤^{1,2}, 刘海龙¹, 侯靖^{1,2}, 韩静¹, 周坚¹, 向华^{1*}

¹ 中国科学院微生物研究所, 微生物资源前期开发国家重点实验室, 北京 100101

² 中国科学院大学, 北京 100049

摘要:【目的】建立一个基于 PHA 颗粒-PhaP 标签的简便实用的极端嗜盐古菌蛋白表达纯化系统, 并探讨嗜盐古菌内含肽在该系统优化中的应用。【方法】以嗜盐古菌-大肠杆菌穿梭载体 pWL502 为基本骨架, 构建带有嗜盐古菌强启动子和 PhaP 融合标签的表达载体; 在地中海富盐菌 *phaP* 缺陷株 ($\Delta phaP$) 中融合表达目的基因, 通过蔗糖密度梯度离心分离纯化结合于 PHA 颗粒上的 PhaP 融合蛋白; 在 *phaP* 基因与多克隆位点之间引入特定的嗜盐古菌内含肽元件, 尝试通过定点突变改变该内含肽的剪切活性。【结果】成功构建了以 PhaP 作为 N 端融合标签的表达载体 pPM 以及作为 C 端融合标签的载体 pIP; 在 *phaP* 基因簇强启动子控制下, 二者均实现了目标蛋白的高效表达; 通过 PHA 颗粒介导的蛋白分离纯化策略, 实现了以 PhaP 为融合标签的目标蛋白的分离纯化; 发现内含肽序列 Hbt21 在地中海富盐菌中保持了高效的剪接活性, 通过定点突变其 C 端末位氨基酸天冬酰胺 (N182) 及邻位的丝氨酸 (S183) 失活了该内含肽的 C 端剪接活性。【结论】首次建立了一个基于 PHA 颗粒-PhaP 标签的简便节约的极端嗜盐古菌蛋白表达纯化系统, 并确定了嗜盐古菌型内含肽 C 端剪接的活性位点, 为该内含肽将来应用于 PhaP 融合蛋白的标签去除奠定了基础。

关键词: 极端嗜盐古菌, 表达载体, 聚羟基脂肪酸酯颗粒, PhaP 标签, 内含肽

中图分类号: Q933 **文章编号:** 0001-6209(2014)09-0998-12

嗜盐古菌是一类生活在各种高盐环境中的极端微生物。它们主要通过胞内积累高浓度的无机盐离子 (主要为 K^+ 离子) 来保持渗透压平衡, 因此极端嗜盐古菌蛋白一般需要在高盐条件下才能正确折叠和行使功能。由于在胞内盐浓度低的异源表达系统如大肠杆菌等中表达嗜盐古菌蛋白时, 其蛋白通常不能正确折叠而聚集形成包涵体, 不仅需要经过变

性、复性等繁琐的步骤, 而且复性成功率极低, 限制了嗜盐古菌蛋白结构与功能的研究, 因此有必要建立一个简便实用的以极端嗜盐古菌为宿主的蛋白表达纯化体系。

聚羟基脂肪酸酯 (Polyhydroxyalkanoate, PHA) 是由许多不同种属的细菌或嗜盐古菌在碳源过量而某些其它营养元素 (如氮、磷等) 受限的条件下, 在

基金项目: 国家自然科学基金重点项目 (31330001)

* 通信作者。Tel/Fax: +86-10-64807472; E-mail: xiangh@im.ac.cn

作者简介: 伍锦花 (1988-), 女, 湖南娄底人, 硕士研究生, 主要从事极端嗜盐古菌蛋白表达纯化系统的研究。E-mail: wujinhua2468@163.com

收稿日期: 2014-03-18; **修回日期:** 2014-04-08

胞内积累的一类作为碳源和能源储存的可降解的生物聚酯^[1]。PHA 在胞内以颗粒形式存在,其表面结合有多种参与 PHA 合成、降解及调节的蛋白,称为 PHA 颗粒结合蛋白^[2],其中 PhaP 为 PHA 颗粒的主要结构蛋白,其含量最高可达到细胞蛋白总量的 5% 和 PHA 颗粒蛋白总量的 90% 以上^[3]。PHA 颗粒可通过简单的离心与其他胞质蛋白分离,因此可作为一种天然的分离介质;而 PHA 颗粒蛋白能够有效结合于 PHA 颗粒上,因此可以作为蛋白融合表达标签,介导目标蛋白的固定与表达纯化^[4]。内含肽元件则是位于某些前体蛋白编码序列中间的一段插入序列,在 mRNA 翻译成前体蛋白质后,内含肽能够自我剪切,并介导其两端的外显肽序列以肽键相连接,从而形成成熟的蛋白质^[5]。内含肽主要通过四步亲核置换反应实现蛋白质的剪接,通过定点突变的方法,可以实现内含肽某一端的可控断裂^[6],从而应用于蛋白质融合表达系统中标签肽段的清除。

地中海富盐菌 (*Haloferax mediterranei*) 是研究嗜盐古菌的重要模式菌株,具有生长速度快、可遗传操作等特点,并可在胞内大量积累 PHA 颗粒。此前,本实验室 Cai 等人^[3]首次在地中海富盐菌中鉴定了

PHA 颗粒结构蛋白 PhaP 的功能,但该 PhaP 的应用还尚未有报道。同样地,虽然 Sakakibara 等人曾在 *Halobacterium* sp. NRC-1 证实了 Hsp-NRC1 CDC21 内含肽(以下简称为 Hbt21)在其天然宿主蛋白(CDC21)中能够发生自我剪接反应^[7],但该内含肽的剪接活性位点及应用方面的研究亦未见报道。本研究将利用地中海富盐菌 PHA 颗粒及其 PhaP 蛋白,建立一套以嗜盐古菌为表达宿主、基于 PHA 颗粒和 PhaP 标签的蛋白表达纯化系统;更进一步地,为了能方便地将 PhaP 标签从纯化后的融合蛋白上清除,我们还在该系统中引入了一个内含肽元件,分析嗜盐古菌内含肽 Hbt21 的 C 端剪接活性位点,探索其用于嗜盐古菌蛋白标签去除的可能性。

1 材料和方法

1.1 材料

1.1.1 菌株和质粒: 本研究中所用的菌株和质粒见表 1。

1.1.2 引物: 本研究中所用的引物见表 2。

表 1. 本文中所使用的菌株和质粒

Table 1. Strains and plasmids used in this study

strains or plasmids	description	sources/references
<i>H. mediterranei</i> Δ <i>phaP</i>	<i>phaP</i> -deleted mutant of <i>H. mediterranei</i> DF50	[3]
<i>Halobacterium</i> sp. NRC-1	Wild-type strain	[8]
<i>E. coli</i> JM109	<i>recA1 supE44 endA1 hsdR17 gyrA96 relA1 thi</i>	[9]
<i>E. coli</i> JM110	<i>dam- dcm-</i> strain of <i>E. coli</i> JM109	[10]
Plasmids		
pUC57	Cloning vector, Amp ^r	Sangon
pUC57M	2.8 kb, derivative of pUC57 by insertion of a 142-bp fragment containing multiple cloning sites and the synthesized terminator T _{syn} at the <i>Xba</i> I and <i>Hind</i> III sites	this study
pWL502	7.9 kb, derivative of pWL102 by replacing Mev ^r with <i>pyrF</i> marker	[3]
pWLP	8.6 kb; pWL502-based expression vector containing <i>phaP</i> and promoter of <i>gap12</i>	[3]
pPM	8.7 kb; derivative of pWL502 by insertion of P _{5219a} promoter, <i>phaP</i> and a 137-bp fragment containing multiple cloning sites, T _{syn} terminator at the <i>Xba</i> I and <i>Nco</i> I sites	this study
pPE	9.3 kb, derivative of pPM by insertion of <i>gvpE</i> at the <i>Nde</i> I and <i>Eco</i> RV sites	this study
pIP	9.1 kb, derivative of pWL502 by insertion of P _{5219a} promoter, multiple cloning sites, the intein fragment (<i>hbt21</i>) of NRC-1, and <i>phaP</i>	this study
pUCm-T	Cloning vector, Amp ^r	Sangon
pTIP	3.9 kb, derivative of pUCm-T by insertion of P _{5219a} promoter, multiple cloning sites, the intein fragment (<i>hbt21</i>) of NRC-1, and <i>phaP</i>	this study
pN182A	9.1 kb, pIP with mutation of N182 of <i>hbt21</i> to A	this study
pS183A	9.1 kb, pIP with mutation of S183 of <i>hbt21</i> to A	this study
pBIP	9.5 kb, pN182A with insertion of 357-bp <i>bktBβ</i> fragment at the <i>Not</i> I and <i>Eco</i> RV sites	this study

表 2. 本文中所使用的引物

Table 2. The primers used in this study

primers	sequences (5'→3') ^a
PaF	CGGGGTACCCTTATAAGGTCGTTACTGCGGCT
PaR	GCAGGATCCCTCCTAACTCGGTGTTGT
phaPF	CGTGGATCCATGAGTGAACAAGCCAA
phaPR	GCTCTAGACTCCGGCGTGTCTGGTGTG
gvpEF	GGAATTCCATATGGAACGGCTCATCGAGGAAC
gvpER	CGGATATCTTATCGCTCATTTGGATTCACT
PbF	CGGGATCCAATGGTGTGGAAGGGAAC
PbR	ACTTCCCGTTCCGATATCGAGCTCGCGGCCCCATATGGTCTGTCATCTCTCTA
IF	GATGACGAACGACCATATGGCGCGCCGAGCTCGATATCGGAACCGGGAAGTGCCTG
IR	TGGGTTGGCTGTTCACCTCGCTGCGGCTGCCGCTGCGGAGATCATCTGGGACTTGTGC
phaPF2	GCACAACCTCCAGATGATCTCCGACGCGCAGCCGACGAGTGAACAAGCCAACCCA
phaPR2	GGGGTACC GGCGGGCTAAAATAGTT
N182AF	TGGTCTCGCACGCGTCCAGATG
N182AR	CATCTGGGACGCGTGCAGACCA
S183AF	TCGCACAACGCCAGATGATCTCCG
S183AR	CGGAGATCATCTGGGCGTTGTGCGA
BF	ATAAGAATCGGGCCGACCCTGGAAGCTGGCAAG
BR	GCGATATCCACAGGTCCAGCGATA

^a Restriction sites are underlined.

1.1.3 培养基和培养条件:大肠杆菌采用 LB 培养基于 37℃ 培养,必要时添加终浓度为 100 mg/L 的氨苄青霉素。*H. mediterranei* Δ *phaP* 采用 AS-168 培养基(添加终浓度为 50 mg/L 的尿嘧啶)于 37℃ 培养。带有重组质粒的 *H. mediterranei* Δ *phaP* 通常用 AS-168 SY 培养基 37℃ 培养^[11];为了实现 *phaP* 基因簇启动子的高表达,并使菌体尽可能多地积累 PHA 颗粒,则采用发酵培养基 MGN(每升含葡萄糖 10 g, NaCl 200 g, MgSO₄·7H₂O 20 g, KCl 2 g, NH₄Cl 2 g, 谷氨酸钠 1 g, KH₂PO₄ 37.5 mg, 柠檬酸铁铵 8 mg, 微量元素溶液 SL-6 1 mL, PIPES 15.1 g, pH 7.0)于 37℃ 摇瓶培养^[12]。

1.1.4 主要试剂和仪器:KOD Plus 聚合酶购自 Toyobo 公司, KAPA HiFi DNA 聚合酶购自 KAPA Biosystems 公司,其他常用限制性内切酶、T4 连接酶、*DpnI* 购自 New England Biolabs (NEB) 公司。胰蛋白酶购自 Roche 公司,蛋白质分子量标准品 Blue plus 购自全式金公司。pUCm-T 载体购自上海生工。所用仪器:超声破碎仪(JY92-II D),超速离心机(Optima L-100 XP),MALDI TOF/TOF 型质谱仪(AB SCIEX 5800)。

1.2 内含肽的选取

为了寻找到最合适的嗜盐古菌型内含肽作为研究对象,参考内含肽的专业数据库 InBase^[13]以及 *H.*

mediterranei 基因组数据,选取剪接位点相对保守的内含肽元件作为载体插入片段。同时,为了不影响蛋白的折叠以及质粒的稳定性,尽量选择不含归巢核酸内切酶结构域(DOD domain)、长度较短且在 *H. mediterranei* 基因组中无同源基因的微型内含肽。利用在线工具 EMBL clustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>)及 GeneDoc 软件进行多序列比对分析。

1.3 PCR 扩增及定点突变

本文中所使用的 PCR 扩增引物详见表 2。

常规的 PCR 采用 KOD Plus 高保真 DNA 聚合酶进行扩增,反应体系(25 μL)如下:模板 0.2 μL、KOD Plus buffer 2.5 μL、KOD Plus dNTP(2 mmol/L) 2.5 μL、KOD Plus 0.3 μL(1 U/L)、正向引物(10 μmol/L)1 μL、反向引物(10 μmol/L)1 μL、ddH₂O 17.5 μL。

重叠延伸 PCR(overlap PCR)的方法主要参考文献[14]:首先采用 KAPA Hifi DNA 聚合酶分别扩增强启动子 P_{5219b}、*phaP* 基因以及内含肽元件 *Hbt21* 相对应的 DNA 序列 *hbt21*, 3 个片段再通过两轮重叠延伸 PCR 反应融合在一起。第一轮 PCR 反应:P_{5219b} 与 *hbt21* 片段的融合,反应体系(25 μL)如下:P_{5219b} 与 *hbt21* 的 PCR 扩增产物各 1 μL、5 × KAPA HiFi High GC buffer 5 μL、dNTP(10 mmol/L)0.75 μL、正向引

物(10 $\mu\text{mol/L}$) 0.75 μL 、反向引物(10 $\mu\text{mol/L}$) 0.75 μL 、KAPA Hifi DNA 聚合酶 0.5 μL (1U/L)、ddH₂O 15.25 μL ; 反应程序如下: 95 $^{\circ}\text{C}$ 2 min; 98 $^{\circ}\text{C}$ 20 s, 68 $^{\circ}\text{C}$ 15 s, 72 $^{\circ}\text{C}$ 30 s, 30 个循环; 72 $^{\circ}\text{C}$ 5 min。反应完成后, 扩增产物 Fsl 再与 *phaP* 进行第二轮 PCR 反应。

PCR 定点突变的方法主要参考文献[15]: 考虑到 pIP (~9.1 kb) 载体较大, PCR 扩增时可能引入新的突变, 因此以中间载体 pTIP (~3.9 kb) 为模板, 分别设计引物 N182AF/R、S183AF/R, 反应体系与上述常规的 PCR 反应相同, 反应程序如下: 95 $^{\circ}\text{C}$ 5 min; 95 $^{\circ}\text{C}$ 30 s, 54 $^{\circ}\text{C}$ 30 s, 68 $^{\circ}\text{C}$ 4 min, 30 个循环; 68 $^{\circ}\text{C}$ 8 min。PCR 反应后, 直接在反应体系中加入 1 μL *DpnI*, 并于 37 $^{\circ}\text{C}$ 水浴锅中酶切消化 1–2 h, 取 5 μL 消化产物转化 *E. coli* JM109, 过夜培养后分别提质粒, 测序验证是否含有正确的突变。

1.4 质粒转化嗜盐古菌及其验证

各质粒均在 *E. coli* JM109 中构建完成。为了提高转化效率, 所有构建好的质粒均经穿梭 *E. coli* JM110 后^[16], 再通过聚乙二醇(PEG)介导的转化方法^[17]导入到宿主菌 *H. mediterranei* ΔphaP 中, 通过 *pyrF* 选择标记筛选转化子, PCR 验证转化子中是否含有正确的重组质粒。

1.5 带有 PhaP 标签的融合蛋白的表达以及 PHA 颗粒的分离与纯化

从平板上挑取正确的转化子, 接至 10 mL AS-168SY 培养基中培养 4–5 d, 待菌株生长至稳定期后以 1:10 比例转接 1 次, 待菌体生长至对数期(约 2–3 d)再以 1:10 比例转接至 200 mL 发酵培养基 MGN(培养基成分见上文)摇瓶培养, 此过程中菌体开始大量积累 PHA 颗粒并表达 PhaP 融合蛋白。菌体在发酵培养基中生长 4–5 d 后, 离心收集菌体, 再用适量体积的 TBSW 缓冲液(NaCl 60 g/L, MgSO₄·7H₂O 2 g/L, KCl 8 g/L, Tris-HCl 40 mmol/L, pH 7.0)悬浮菌体, 超声破碎后, 低速离心, 上清即为含有 PHA 颗粒的细胞粗提液。

PHA 颗粒的分离主要参考文献[3], 采用蔗糖密度梯度离心的方法, 同时为了尽可能的保证融合蛋白的活性, 蔗糖梯度采用 TBSW 缓冲液代替水溶液来进行配制, 具体过程如下(为了避免蛋白的降解, 以下步骤均应在 4 $^{\circ}\text{C}$ 进行): 用移液枪吸取 3 mL 细胞粗提液于 12.5 mL 大小的超速离心管中, 用

带长针头的注射器于离心管底部依次加入 2 mL 1.0 mol/L, 1.3 mol/L, 1.6 mol/L 的蔗糖溶液(用 TBSW 配制), 超速离心(165000 $\times g$, 2.5 h, 4 $^{\circ}\text{C}$)后, 用注射器小心吸取颗粒层至 Eppendorf 离心管, 用 TBSW 稀释后, 离心(16000 $\times g$, 1 h, 4 $^{\circ}\text{C}$), 弃上清, 离心管底部白色沉淀物即为 PHA 颗粒。

1.6 SDS-PAGE、蛋白胶内酶切以及质谱鉴定

用 1 \times SDS-PAGE 上样缓冲液悬浮 PHA 颗粒, 离心后上清即为颗粒蛋白。SDS-PAGE 检测 PhaP 融合蛋白的表达及纯化情况。SDS-PAGE (12% 浓缩胶, 5% 分离胶)的电泳条件为 100 V, 2.5 h。SDS-PAGE 胶用考马斯亮蓝染色过夜, 脱色 3 h, 从凝胶上依次切下目标条带, 用胰蛋白酶酶切后, 进行 MALDI TOF/TOF 质谱分析。胶内酶切步骤流程主要参考文献[18], 质谱数据分析所用搜索引擎为 MASCOT 2.0(Matrix Science, London, U. K.), 序列比对所使用的数据库为 Swiss-Prot 数据库及本地数据库。

2 结果

2.1 以 PhaP 作为蛋白 N 端融合标签的表达载体的构建

以 *H. mediterranei* 基因组为模板, 分别以 *phaPF*/R, *PaF*/R 为引物, PCR 扩增得到 *phaP* 基因及包含其自身启动子序列的 DNA 片段 P_{5219_a} (*phaP* 基因簇强启动子, 长度为 188 bp)。为了实现目的基因的高效表达与转录终止, 人工合成多克隆位点(MCS)及含有 *phaP* 基因自身终止序列的片段 T_{syn}; 同时, 为了尽量不影响 PhaP 蛋白的折叠, 在片段中引入 5 \times GS-linker 序列, 并在多克隆位点之后添加 6 \times His-tag, 构成片段 *ins1*。该 *ins1* 片段的合成由上海生工完成, 通过 *XbaI* 和 *HindIII* 酶切位点连入 pUC57, 得到重组质粒 pUC57M。重组质粒经 *XbaI*、*NcoI* 双酶切, 回收得到 *ins* 片段; *phaP* 基因经 *KpnI*、*XbaI* 双酶切, P_{5219_a} 经 *KpnI*、*BamHI* 酶切, 3 个片段经 *KpnI* 和 *NcoI* 酶切位点连入载体 pWL502, 得到重组表达载体 pPM(图 1-A)。测序及酶切结果(图 1-B)表明该载体已构建成功。

2.2 PhaP 标签融合蛋白的表达与纯化

为了验证 PhaP 作为融合标签用于分离纯化蛋白的可行性以及表达载体 pPM 中多克隆位点的实用性, 以引物 *gvpEF*, *gvpER* 扩增得到嗜盐古菌中气

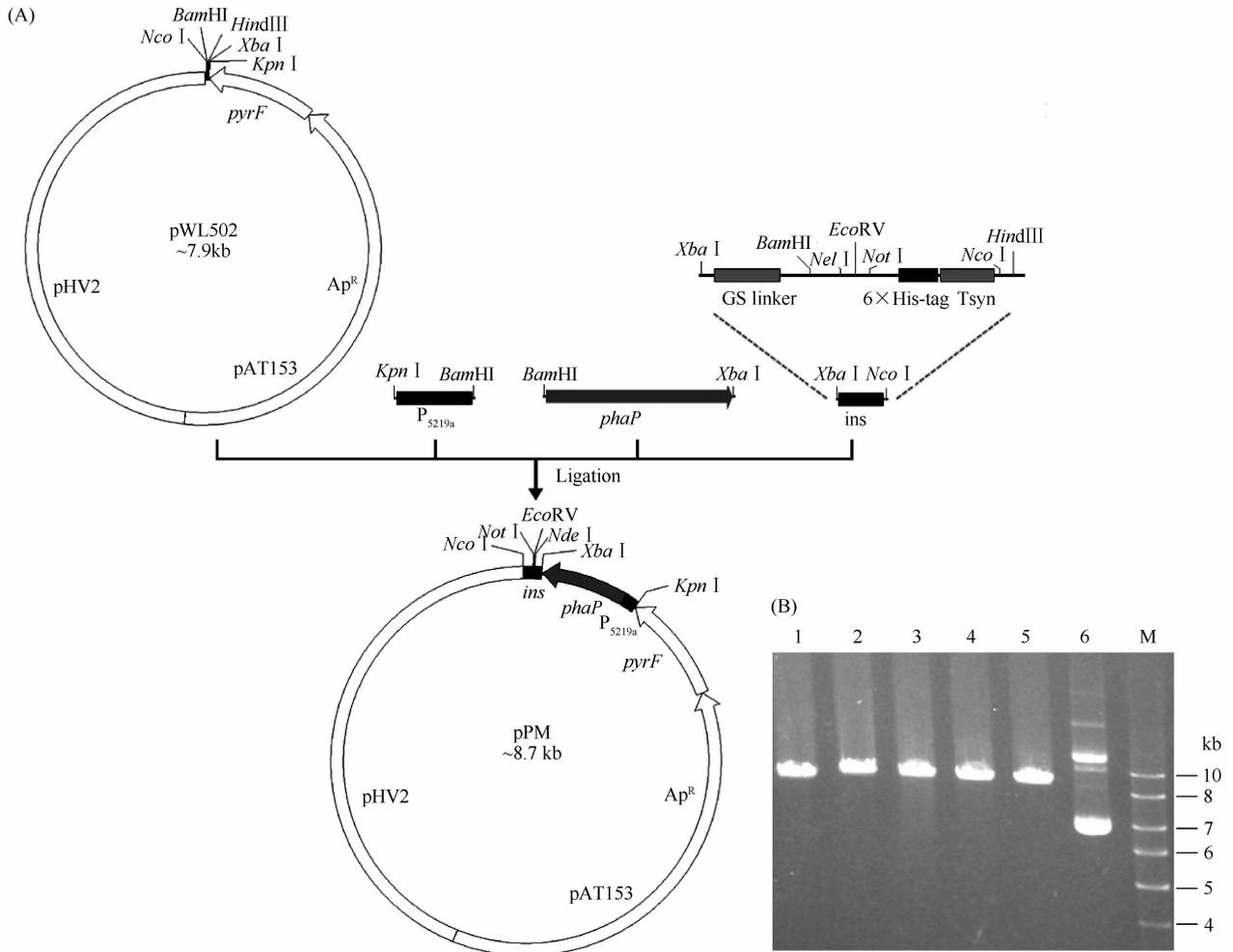


图 1. PhaP 作为蛋白 N 端亲和和标签的表达载体的构建与验证

Figure 1. Construction and restriction digests of the expression vector (pPM) with PhaP as an N-terminal fusion tag. A: Construction of pPM. B: pPM digested by *NcoI* (lane 1), *NotI* (lane 2), *EcoRV* (lane 3), *NdeI* (lane 4), *XbaI* (lane 5), respectively; lane 6, untreated pPM as the control; M: 1 kb DNA ladder.

泡合成的调控基因 *gvpE*, 经 *NdeI*、*EcoRV* 双酶切连入 pPM 得到重组质粒 pPE。GvpE-PhaP 融合蛋白经表达后, 通过 PhaP 标签结合于 PHA 颗粒上。蔗糖密度梯度离心结果表明, PHA 颗粒层在蔗糖溶液中呈现为单一白色条带, 密度介于 1.0 - 1.3 mol/L 蔗糖梯度之间(图 2-A)。SDS-PAGE 结果表明, PhaP 为主要的 PHA 颗粒结合蛋白, 作为融合表达标签, 可以有效结合于 PHA 颗粒, 从而介导目标蛋白的分离与纯化(图 2-B)。

2.3 内含肽的结构特征及序列分析

鉴于 PhaP 作为蛋白融合表达标签, 已取得良好纯化效果, 为了能简便有效地从纯化的目标蛋白上去除 PhaP 标签, 我们拟在表达载体中引入能够

自我剪接的内含肽这一特殊的蛋白元件。结合内含肽数据库 InBase 以及 *H. mediterranei* 基因组数据, 通过比较分析发现, 嗜盐菌内含肽大小为 180 - 642 个氨基酸不等, 分布于 DNA 聚合酶、核糖核苷酸还原酶以及细胞分裂控制蛋白 21 (CDC21 或 MCM) 等蛋白中。在这些可能的嗜盐古菌内含肽中, Hsp-NRC1 CDC21 (以下简称为 Hbt21) 大小仅为 182 个氨基酸, 属于微型内含肽; 其 N、C 端为剪接结构域 (splicing domain), 含有 4 个保守的基序, 依次为 Block A, B, F, G, 两个剪接结构域之间以柔性区域 (Linker) 连接 (图 3-A)。内含肽的序列比对结果表明, 极端嗜盐古菌内含肽 Hbt21 与其他古菌以及细菌的内含肽一样, 含有 4 个保守的剪接位

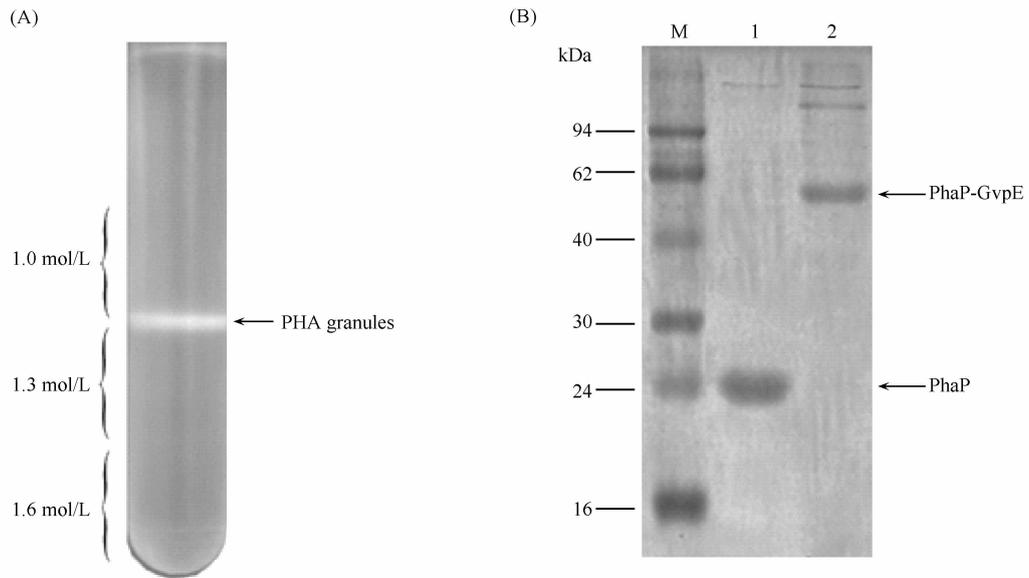


图 2. 含有 PhaP 标签的融合蛋白分离纯化及 SDS-PAGE 分析

Figure 2. Purification and SDS-PAGE analysis of the expressed PhaP-tagged protein. A: Sucrose gradient centrifugation of cell lysates of *H. mediterranei* Δ *phaP* harboring plasmid pWLP or pPM, arrow indicates the layer of PHA granules. B: SDS-PAGE analysis of PHA-granule associated proteins. Lane 1-2, PHA-granule associated proteins of *H. mediterranei* Δ *phaP* harboring plasmid pWLP and pPM, respectively. M, Marker.

点(图 3-B):N 端起始氨基酸为半胱氨酸(C),C 端末两位氨基酸分别为组氨酸(H)及天冬酰胺(N),C 端外显肽第一位氨基酸为丝氨酸(S),暗示着嗜盐古菌可能具有与细菌或者其他古菌相同或相似的剪接机制。

2.4 基于内含肽-PhaP 的表达载体的构建

根据 overlap PCR 引物设计原则,分别设计引物 PbF, PbR,并在反向引物中添加多克隆位点(包括 *NdeI*、*NotI*、*SacI*、*EcoRV* 四个限制性内切酶位点),以 *H. mediterranei* 基因组为模板扩增得到启动子片段 P_{5219b} (*phaP* 基因簇强启动子,包含多克隆位点,总长度为 99 bp)。由于内含肽的剪接活性在异源蛋白中会受到不同程度的影响,因此分别保留了 Hbt21 内含肽上游 4 个氨基酸(GTGK)及其下游 5 个氨基酸(SQMIS),构成片段 *hbt21*。 P_{5219b} 与 *hbt21* 片段通过 overlap PCR 的方法融合在一起,构成片段 Fs1。为了便于后续蛋白纯化及内含肽的定点突变实验,同时验证 PhaP 作为蛋白 C 端亲和标签的可行性,将片段 Fs1 融合于 *phaP* 基因的 N 端构成片段 Fs2,再经 *Bam*HI 以及 *Kpn*I 限制性酶切位点连入 pWL502,构成表达载体 pIP(图 4),插入片段序列 Fs2 经测序验证为正确。

2.5 内含肽剪接活性的验证及应用于蛋白表达纯化的探索

表达载体 pIP 中嗜盐古菌型内含肽 Hbt21(图 5 中简称为 1)的 N 端外显肽为多克隆位点 MCS 编码肽(简称为 M),C 端外显肽为 PhaP(简称为 P)。为了验证该内含肽在外源蛋白中是否仍具有高效的剪接活性,从 PHA 颗粒上分离得到融合蛋白后,采用 SDS-PAGE 方法进行检测。结果表明:Hbt21-PhaP 融合蛋白表达终产物大小约为 24 kDa,与 PhaP 大小相近,说明该内含肽在异源宿主蛋白(MIP)中发生了高效的自我剪接反应,终产物为 MP(图 5-B,泳道 2),比不含 MCS 编码肽的野生型 PhaP 蛋白(图 5-B,泳道 1)略大。

为了确定内含肽 Hbt21 的 C 端剪接活性位点,并阻断其 C 端的剪接反应,使得内含肽稳定连接于 PhaP 标签上(可用于未来将内含肽与 PhaP 标签通过内含肽 N 端的可控剪切从目标蛋白上去除),分别将内含肽 C 端末位氨基酸天冬酰胺突变成丙氨酸(N182A),C 端外显肽第一位氨基酸丝氨酸突变成丙氨酸(S183A)。SDS-PAGE 结果表明,保守的天冬酰胺位点经突变后,内含肽无法再进行自我环化,从而阻断了内含肽 C 端断裂反应,表达终产物为 MIP(图 5-B,泳道 3)。值得一提的是,图 5-B 中 MIP 所

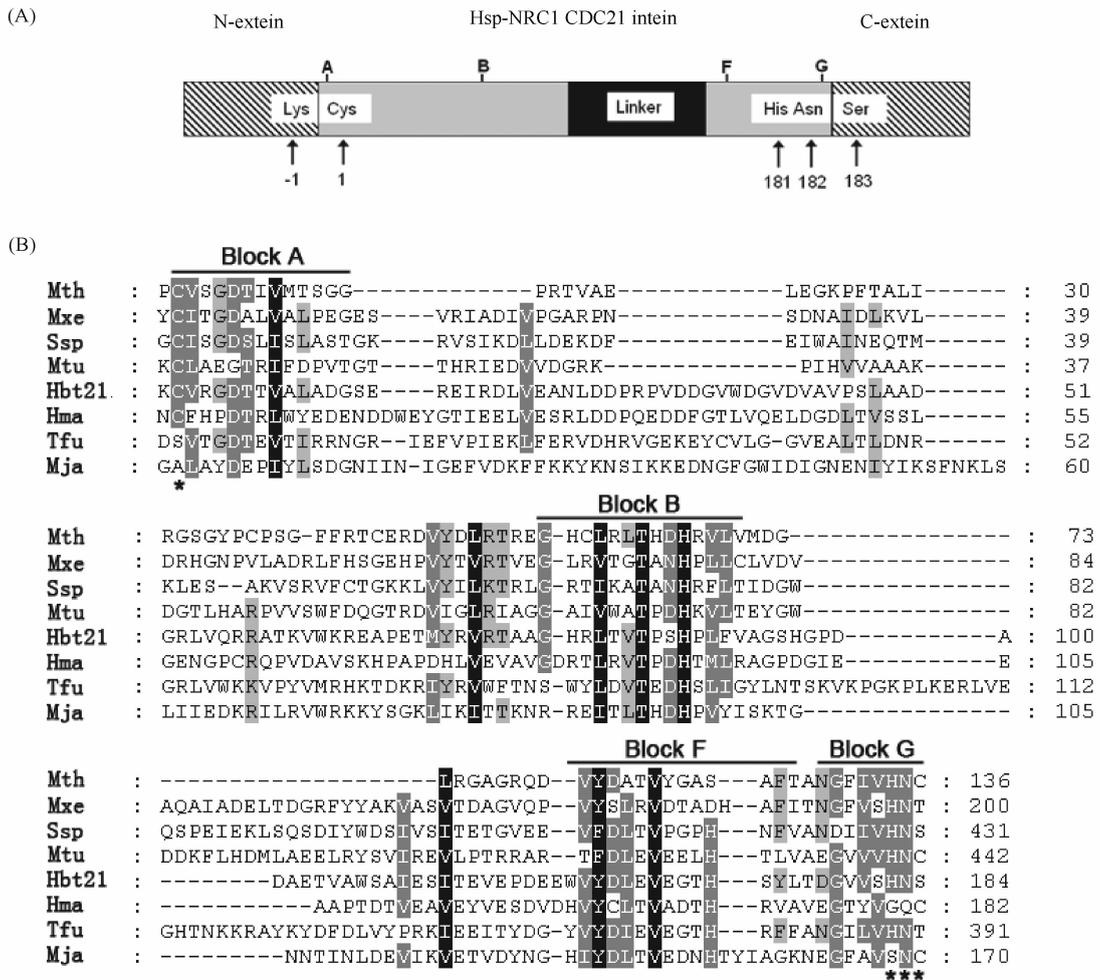


图 3. 极端嗜盐古菌型内含肽 Hbt21 序列特征分析

Figure 3. Sequence analysis of the haloarchaeal-type intein Hbt21. A: splicing motifs of Hbt21. The N-, C-terminal splicing domains (grey boxes), the linker domain (dark box) of the intein, and the four splicing motifs consisting of block A, B, F, G are shown; the N-extein and C-extein are shown in striped boxes; the arrows and numbers below indicate the amino acid positions. B: alignment of Hbt21 and (partial) intein sequences from other archaea and bacteria (containing -1 and +1 extein residues). Asterisks indicate the four conserved splice junction residues. The sequences corresponding to the four conserved motifs of block A, B, F, G in figure A are shown. Abbreviations: Mth, Mth RIR1 intein; Mxe, Mxe GyrA intein; Ssp, Ssp DnaB intein; Mtu, Mtu-H37Rv RecA intein; Hbt21, Hsp-NRC1 CDC21; Hma, Hma Pol-II intein; Tfu, Tfu Pol-2 intein; Mja, Mja KlbA intein.

示条带中可能存在 IP 剪切产物(由内含肽 N 端剪切副反应产生,与 MIP 分子量相近),但不影响我们对内含肽(I)与 PhaP-标签(P)已稳定连接的判断。类似地,突变丝氨酸这一位点,阻断了内含肽剪接过程中的第二步亲核置换反应,从而降低了内含肽自我剪接的效率,但 C 末端位天冬酰胺仍可进行自我环化并发生 C 端断裂反应,因此通过 PHA 颗粒纯化得到的蛋白为 MIP 及 P 的混合物(图 5-B,泳道 4)。上述结果表明,内含肽 C 端的剪接反应关键氨基酸 N182A 突变,即可成功实现内含肽稳定连接于 PhaP

标签上的目的。

为了进一步探索该内含肽-PhaP 标签(IP)在极端嗜盐古菌蛋白表达纯化中的应用,还需进一步了解内含肽 N 端剪切活性是否可控,以便表达纯化得到融合蛋白 XIP 后(图 5-A),通过在特定理化条件下激活内含肽的 N 端剪切活性,将目标蛋白 X 从固定于 PHA 颗粒上的 IP 上释放出来,达到表达纯化不带标签的天然嗜盐古菌蛋白的目的。本实验采用嗜盐古菌特异的 β -酮硫解酶 β 亚基作为报告蛋白 BktB β ,通过 PCR 扩增 *bktB β* 基因,经 *NotI*、*EcoRV* 双

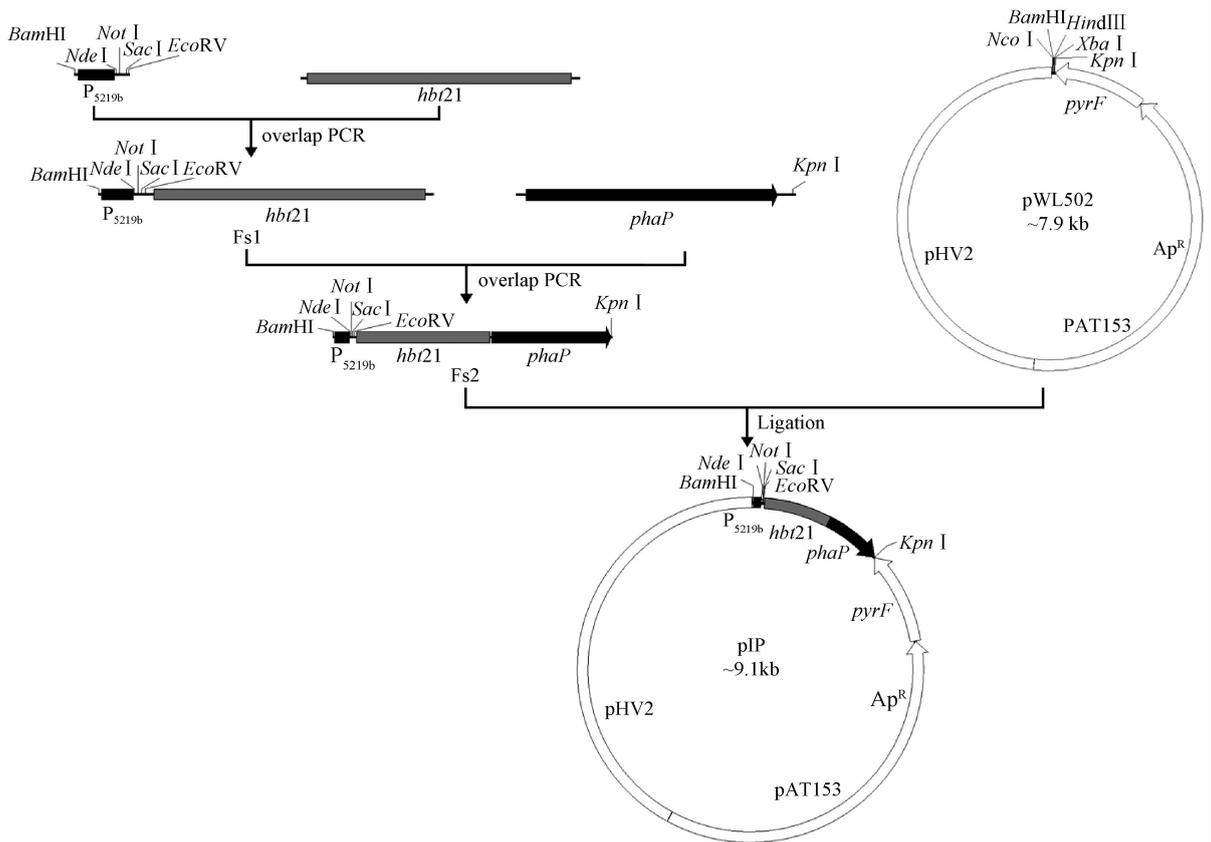


图 4. 基于内含肽-PhaP 的表达载体 pIP 的构建

Figure 4. Construction of the expression vector (pIP) containing intein and *phaP*.

酶切后连入质粒 pN182A 中, 构成重组质粒 pBIP。BktB β -Hbt21-PhaP (BIP) 融合蛋白经表达纯化后, SDS-PAGE 检测结果表明, 图中出现了 3 条较为明显的蛋白条带(条带 1, 2, 3), 其中条带 1 丰度最高, 大小与内含肽-PhaP 融合蛋白表达产物(IP)相近, 约为 35 kDa; 另两条蛋白条带丰度相近, 在 SDS-PAGE 胶上位于 62 kDa 和 94 kDa 之间(图 5-C)。分别从凝胶上切下这 3 条蛋白条带, 胶内酶切后, 进行 MALDI TOF/TOF 质谱鉴定分析。酶切肽段的序列匹配结果表明, 条带 1 为 Hbt21-PhaP 融合蛋白(IP), 条带 2 和 3 经鉴定均为 BIP 融合蛋白(表 3)。以上实验结果说明, 利用该载体系统可以成功实现目标蛋白-内含肽-PhaP 融合蛋白的表达与纯化, 为内含肽-PhaP 标签综合应用于嗜盐古菌蛋白表达奠定了基础。但由于内含肽发生大量 N 端断裂的副反应, 使得大部分目标蛋白(B)在分离纯化融合蛋白前已丢失(即主要余下 IP), 将影响目标蛋白的回收率。未来还需进一步研究该内含肽 N 端断裂的可控机制, 以实现从目标蛋白-内含肽-PhaP 融合蛋白中可控地获得目标蛋白。

3 讨论

由于嗜盐古菌蛋白具有其特殊的理化性质, 该领域一直缺乏一个较为方便实用的蛋白表达纯化体系, 从而在一定程度上限制了嗜盐古菌蛋白结构及生化性质的研究。pWL102^[19] 是较为常用的嗜盐古菌-大肠杆菌穿梭载体, 但由于该载体选用的是来源于嗜盐菌自身基因的 *Mev*^r 选择标记, 易于与染色体 DNA 发生同源重组, 使得筛选正确转化子时出现很高的假阳性, 大大增加了工作量。pWL502 是以 pWL102 为基本骨架, 将该载体的选择性标记 *Mev*^r 替换为 *pyrF* 改造而成。pWL502 不仅同样具有稳定性好、宿主范围较为广泛等特点, 同时, 选用 *pyrF* 作为选择性标记, 有效避免了 pWL102 载体上 *Mev*^r 假阳性率高的缺点^[11]。本研究以 pWL502 作为载体基本骨架, 并首次引入嗜盐古菌 PHA 颗粒结构蛋白 PhaP 作为蛋白表达标签, 同时添加多克隆位点、启动子(P_{5219a} 或 P_{5219b})和终止子等元件, 分别成功构建

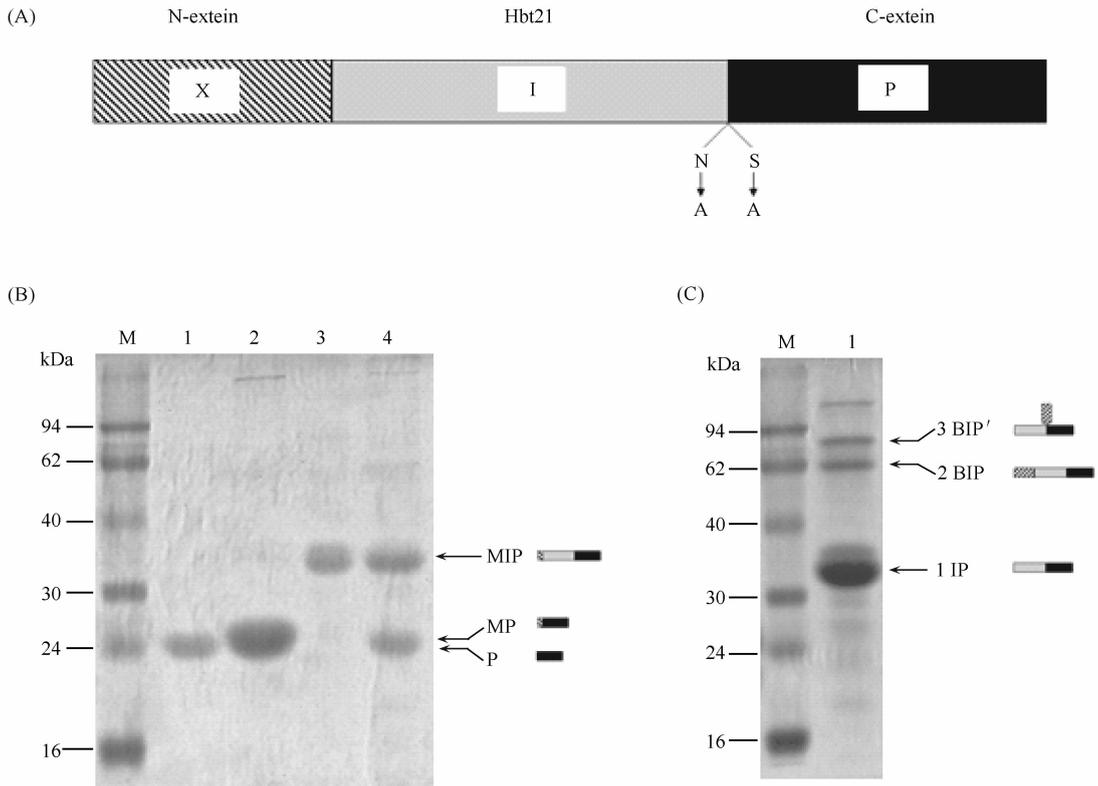


图 5. 含有内含肽-PhaP 标签的表达载体所表达的融合蛋白的 SDS-PAGE 分析

Figure 5. SDS-PAGE analysis of the expressed intein-PhaP-tagged proteins. A: Schematic diagram of the chimeric XIP fusion protein consisting of target protein (X), the Hbt21 intein-containing fragment (I), and the phasin protein PhaP (P). B: SDS-PAGE analysis of PHA-granule associated proteins of *H. mediterranei* Δ phaP harboring plasmid pIP or its mutant derivatives. Lane 1, pWLP; lane 2, pIP; lane 3, pN182A; lane 4, pS183A. The predicted unspliced precursor and various spliced intermediates and products are shown on the right as follows; MIP, unspliced products; MP, spliced products; P, the wild type PhaP or product of cleavage at C-terminal junction of MIP. C: SDS-PAGE analysis of PHA-granule associated proteins of *H. mediterranei* Δ phaP harboring plasmid pBIP. The predicted unspliced precursor and spliced products are shown on the right as follows; BIP*, branched intermediate; BIP, unspliced precursor; IP, product of cleavage at N-terminal junction of BIP. M, marker.

表达载体 pPM 和 pIP。载体经转化到宿主菌 *H. mediterranei* Δ phaP 后,在菌体大量积累 PHA 颗粒的同时,就能实现目标蛋白的高效表达。而 PhaP 作为表达标签,一方面可以直接以胞内积累的 PHA 颗粒作为分离基质,大大降低了成本,另一方面通过蔗糖密度梯度离心就能分离得到纯度很高的融合蛋白,简化了蛋白分离的步骤。本文实验结果表明,PhaP 可分别作为蛋白 N 端或 C 端亲和和标签,蛋白表达量高,分离效果良好。

为了探索能够有效地从目标蛋白上去除 PhaP 标签的方法,本文还首次在嗜盐古菌表达载体中引入内含肽这一特殊的元件。Perler 等人的研究表明,通过对内含肽的保守剪接位点进行改造,可以阻断或降低内含肽 N 端或 C 端的剪接反应^[20],分离纯化

得到前体蛋白后,再通过添加巯基试剂、改变温度和 pH 等方法,便可以实现内含肽介导的融合标签的体外去除。Banki 等人^[21]以能够积累聚羟基丁酸酯 (polyhydroxybutyrate, PHB, 为 PHA 的一种) 的重组大肠杆菌为宿主,利用 PhaP 结合于 PHB 颗粒的特性,结合内含肽介导的 C 端断裂反应,建立了一个经济而又简便的新型蛋白表达纯化系统,但是该系统并不适用于极端嗜盐古菌蛋白的表达与纯化。基于上述原理,同时又由于嗜盐古菌宿主的特殊性,本文首次探讨了嗜盐古菌型内含肽 Hbt21 用于相关蛋白表达纯化载体构建的可能性。本文实验结果表明,来自 *Halobacterium* sp. NRC-1 的野生型内含肽在异源蛋白中仍具有很高的剪接活性。而当内含肽 C 端保守剪接位点 N182 或者 S183 经突变成丙氨酸后,

完全阻断或者减弱了 C 端的剪接反应,说明这两个保守位点为内含肽正常行使剪接功能所必需。

表 3. 图 5-C 中蛋白条带 1-3 的质谱鉴定分析

Table 3. Mass spectrum analysis of protein bands 1-3 in Figure 5-C

band number	MASCOT score	matched proteins	Coverage/%	matched peptides
1	608	PhaP	28	[RSIEQSQR] [SMHQGI EFQK] [GTFARTAVK] [SYFEMMDSSVPGDNR]
		Hbt21	22	[GDTTVALADGSER] [EAPETMYR] [LTVTTPSHPLFVAGSHGPD AVR]
2	618	BktB β	82	[TLEAGKCPNGHVSYPHPR] [VVTWTHSTATPPGVR] [QPNTMAIVEFEVDGQAVR] [ALGQVTDDIETGDVVEPVYV] [EELRDPEVGIK] [APESQSWDGYR]
		PhaP	21	[RSIEQSQR] [SMHQGI EFQK] [SYFEMMDSSVPGDNR]
3	706	Hbt21	13	[GTGK] [LTVTTPSHPLFVAGSHGPD AVR]
		BktB β	82	[TLEAGKCPNGHVSYPHPR] [VVTWTHSTATPPGVR] [QPNTMAIVEFEVDGQAVR] [ALGQVTDDIETGDVVEPVYV] [EELRDPEVGIK] [APESQSWDGYR]
3	706	PhaP	21	[RSIEQSQR] [SMHQGI EFQK] [SYFEMMDSSVPGDNR]
		Hbt21	24	[GTGK] [GDTTVALADGSER] [EAPETMYR] [LTVTTPSHPLFVAGSHGPD AVR]

在质粒 pN182A 的多克隆位点连入 *bktB β* 基因后,表达产物经 SDS-PAGE 电泳检测,泳道中出现 3 条蛋白条带,其中蛋白条带 2 和 3 经质谱鉴定为融合蛋白 BIP。BIP 分别位于泳道中不同位置,推测是由于前体蛋白在内含肽的作用下形成了不同构型的中间酯(蛋白条带 2 为线性酯 BIP,蛋白条带 3 为分支酯 BIP^{*});而泳道中丰度最高的条带 1 经质谱鉴定为 Hbt21-PhaP 融合蛋白,说明内含肽在 N 端发生了断裂的副反应,推测可能的原因如下:(1)内含肽

C 端剪接反应被阻断,导致酯类中间物的积累,因酯的水解反应而发生 N-端断裂;(2)内含肽位于异源的宿主蛋白(N 端为 BktB β ,C 端为 PhaP)中,未能正确折叠形成稳定构象,导致内含肽 N 端副反应增多;(3)存在其他内源性的亲核试剂,对酯键发动亲核攻击而使酯键断裂。以上实验结果表明,基于内含肽-PhaP 的载体系统应用于嗜盐古菌蛋白的表达纯化具有一定的可行性,但要实现内含肽介导的可控的嗜盐古菌蛋白表达纯化,该载体系统还需进一

步的摸索以及优化:如增加内含肽两端野生型外显肽的长度而降低外源环境对其剪接活性的影响,通过改变培养条件(如培养时间、温度等)、优化剪接位点附近的氨基酸残基等方式来尽可能减少内含肽 N 端副反应的发生。另外值得一提的是,嗜盐菌内含肽 C 端保守剪接位点具有一定多态性,末两位氨基酸分别有 HN、GQ、GN、KQ、KN 多种组合形式,暗示着可能存在新型的剪接机制,因此该载体系统的建立也可以为进一步研究嗜盐古菌内含肽的剪接反应机制提供重要的工具。

致谢:感谢美国马里兰大学 DasSarma 教授提供 *Halobacterium* sp. NRC-1 菌株,感谢 NEB 公司 Perler 教授对嗜盐古菌内含肽使用的参考建议,感谢中国科学院微生物研究所微生物资源前期开发国家重点实验室罗元明博士和王前老师在质谱分析方面的技术支持。

参考文献

[1] Steinbüchel A, Fuchtenbusch B. Bacterial and other biological systems for polyester production. *Trends in Biotechnology*, 1998, 16(10): 419-427.

[2] Pötter M, Steinbüchel A. Poly (3-hydroxybutyrate) granule-associated proteins; impacts on poly (3-hydroxybutyrate) synthesis and degradation. *Biomacromolecules*, 2005, 6(2): 552-560.

[3] Cai SF, Cai L, Liu HL, Liu XQ, Han J, Zhou J, Xiang H. Identification of the haloarchaeal phasin (PhaP) that functions in polyhydroxyalkanoate accumulation and granule formation in *Haloferax mediterranei*. *Applied and Environmental Microbiology*, 2012, 78(6): 1946-1952.

[4] Peters V, Rehm BHA. In vivo enzyme immobilization by use of engineered polyhydroxyalkanoate synthase. *Applied and Environmental Microbiology*, 2006, 72(3): 1777-1783.

[5] Perler FB, Davis EO, Dean GE, Gimble FS, Jack WE, Neff N, Noren CJ, Thormer J, Belfort M. Protein splicing elements-inteins and exteins-a definition of terms and recommended nomenclature. *Nucleic Acids Research*, 1994, 22(7): 1125-1127.

[6] Xu MQ, Perler FB. The mechanism of protein splicing and its modulation by mutation. *EMBO Journal*, 1996, 15(19): 5146-5153.

[7] Sakakibara N, Han M, Rollor CR, Gilson RC, Busch C,

Heo G, Kelman Z. Cloning, purification, and partial characterization of the *Halobacterium* sp. NRC-1 minichromosome maintenance (MCM) helicase. *Open Microbiol Journal*, 2008, 2: 13-17.

[8] Ng WV, Kennedy SP, Mahairas GG, Berquist B, Pan M, Shukla HD, Lasky SR, Baliga NS, Thorsson V, Sbrogna J, Swartzell S, Weir D, Hall J, Dahl TA, Welti R, Goo YA, Leithauser B, Keller K, Cruz R, Danson MJ, Hough DW, Maddocks DG, Jablonski PE, Krebs MP, Angevine CM, Dale H, Isenbarger TA, Peck RF, Pohlschroder M, Spudich JL, Jung KH, Alam M, Freitas T, Hou SB, Daniels CJ, Dennis PP, Omer AD, Ebhardt H, Lowe TM, Liang R, Riley M, Hood L, DasSarma S. Genome sequence of *Halobacterium* species NRC-1. *Proceedings of the National Academy of Sciences of the United States of America*, 2000, 97(22): 12176-12181.

[9] Joseph S, David W. *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory Press, 2001.

[10] Palmer BR, Marinus MG. The *dam* and *dcm* strains of *Escherichia coli* - a review. *Gene*, 1994, 143(1): 1-12.

[11] Liu HL, Han J, Liu XQ, Zhou J, Xiang H. Development of *pyrF*-based gene knockout systems for genome-wide manipulation of the archaea *Haloferax mediterranei* and *Haloarcula hispanica*. *Journal of Genetics and Genomics*, 2011, 38(6): 261-269.

[12] Zhao D, Cai L, Wu J, Li M, Liu H, Han J, Zhou J, Xiang H. Improving polyhydroxyalkanoate production by knocking out the genes involved in exopolysaccharide biosynthesis in *Haloferax mediterranei*. *Applied Microbiology Biotechnology*, 2013, 97(7): 3027-3036.

[13] Perler FB. Inbase: The intein database. *Nucleic Acids Research*, 2002, 30(1): 383-384.

[14] Bryksin AV, Matsumura I. Overlap extension PCR cloning: A simple and reliable way to create recombinant plasmids. *Biotechniques*, 2010, 48(6): 463-465.

[15] Li SL, Wilkinson MF. Site-directed mutagenesis: A two-step method using PCR and *DpnI*. *Biotechniques*, 1997, 23(4): 588-90.

[16] Holmes ML, Nuttall SD, Dyallsmith ML. Construction and use of halobacterial shuttle vectors and further studies on *Haloferax* DNA gyrase. *Journal of Bacteriology*, 1991, 173(12): 3807-3813.

[17] Cline SW, Lam WL, Charlebois RL, Schalkwyk LC, Doolittle WF. Transformation-methods for halophilic archaeobacteria. *Canadian Journal of Microbiology*, 1989,

- 35(1): 148-152.
- [18] Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Analytical Chemistry*, 1996, 68(5): 850-858.
- [19] Nieuwlandt DT, Daniels CJ. An expression vector for the archaeobacterium *Haloferax volcanii*. *Journal of Bacteriology*, 1990, 172(12): 7104-7110.
- [20] Chong SR, Shao Y, Paulus H, Benner J, Perler FB, Xu MQ. Protein splicing involving the *Saccharomyces cerevisiae* VMA intein - the steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an *in vitro* splicing system. *Journal of Biological Chemistry*, 1996, 271(36): 22159-22168.
- [21] Banki MR, Gerngross TU, Wood DW. Novel and economical purification of recombinant proteins: Intein-mediated protein purification using *in vivo* polyhydroxybutyrate (PHB) matrix association. *Protein Science*, 2005, 14(6): 1387-1395.

Development of PhaP-tagged protein expression and purification systems for extremely halophilic archaea

Jinhua Wu^{1,2}, Shuangfeng Cai^{1,2}, Hailong Liu¹, Jing Hou^{1,2}, Jing Han¹, Jian Zhou¹, Hua Xiang^{1*}

¹ State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

² University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: [**Objective**] To establish a convenient halophilic protein expression and purification system based on the haloarchaeal-type PhaP and polyhydroxyalkanoate (PHA) granule. [**Methods**] We cloned a strong haloarchaeal promoter and the *phaP*-tag into the haloarchaea-*Escherichia coli* shuttle vector pWL502, and then used the constructed vector to express the PhaP-tagged haloarchaeal proteins in the *phaP*-deleted strain *Haloferax mediterranei* Δ *phaP*. We purified the PhaP-fusion proteins, which were associated with PHA granules, by sucrose density gradient centrifugation. We also inserted a haloarchaeal intein-containing fragment between *phaP* and multiple cloning sites, and modulated the intein splicing activity by site-directed mutagenesis. [**Results**] We successfully constructed two expression vectors, pPM and pIP, in which PhaP was used as N-terminal and C-terminal fusion tag, respectively. The haloarchaeal proteins were effectively expressed by both vectors. The PhaP-tagged proteins were easily purified through the strategy of PHA granule-mediated protein purification. In addition, we found that the intein-containing fragment Hbt21 from *Halobacterium* sp. NRC-1 had maintained splicing activity in *H. mediterranei*, and its C-terminal cleavage could be blocked or attenuated by mutating the conserved asparagine (N182) or serine (S183), respectively. [**Conclusion**] We have established a convenient and economical halophilic protein expression and purification system. We have also identified the splicing active sites of a haloarchaeal intein, which showed potential for removing the PhaP-tag from the purified proteins.

Keywords: extremely halophilic archaea, expression vector, polyhydroxyalkanoate granules, PhaP-tag, intein

(本文责编:张晓丽)