

Man₅GlcNAc₂哺乳动物甘露糖型糖蛋白的毕赤酵母表达系统构建

杨晓鹏¹, 刘波¹, 宋森^{1,2}, 巩新¹, 唱韶红¹, 薛奎晶^{1,2}, 吴军¹

1 军事医学科学院生物工程研究所, 北京 100071

2 沈阳药科大学生命科学与生物制药学院, 沈阳 110016

摘要: 蛋白的糖基化对蛋白的活性、高级结构及功能都有重要的影响。酵母表达的糖蛋白不同于哺乳动物表达的杂合型或复杂型糖蛋白, 而是高甘露糖型或过度甘露糖化糖蛋白。在前期成功敲除毕赤酵母 α -1,6-甘露糖转移酶(Och1p) 基因、阻断毕赤酵母过度糖基化, 获得毕赤酵母过度糖基化缺陷菌株 GJK01 (*ura3*, *och1*) 的基础上, 通过表达不同物种来源的 α -1,2-甘露糖苷酶 I (MDSI) 的活性区与酵母自身定位信号的融合蛋白, 并通过 DSA-FACE (基于 DNA 测序仪的荧光辅助糖电泳) 分析筛选报告蛋白 HSA/GM-CSF (人血清白蛋白与粒细胞-巨噬细胞集落刺激因子融合蛋白) 的糖基结构, 发现当编码酿酒酵母 α -1,2-甘露糖苷酶 (MnsI) 基因的内质网定位信号与带有完整 C-端催化区的拟南芥 MDSI 基因融合表达时, 毕赤酵母工程菌株能够合成 Man₅GlcNAc₂ 哺乳动物甘露糖型糖蛋白。这在酵母体内合成类似于哺乳动物杂合型或复杂型糖基化修饰的糖蛋白奠定了基础。

关键词: α -1,2-甘露糖苷酶 I, Man₅GlcNAc₂, 糖基化, 毕赤酵母, 甘露糖型

Construction of yeast *Pichia pastoris* to produce Man₅GlcNAc₂ mammalian mannose-type glycoprotein

Xiaopeng Yang¹, Bo Liu¹, Miao Song^{1,2}, Xin Gong¹, Shaohong Chang¹, Kuijing Xue^{1,2}, and Jun Wu¹

1 Institute of Biotechnology, Academy of Military Medical Sciences, Beijing 100071, China

2 School of Life and Biopharmaceutics, Shenyang Pharmaceutical University, Shenyang 110016, China

Abstract: Glycosylation is vital for activity, higher structure and function of protein. Glycoproteins derived from yeast contain N-glycan of high mannose type and are usually hyperglycosylated, while those from mammalian cells contain N-glycan of hybrid or complex type. We introduced the α -1,2-mannosidase I (MDSI) into yeast cells, which catalyzed an essential proceeding of

Received: March 31, 2010; **Accepted:** July 16, 2010

Supported by: National High Technology Research and Development Program of China (863 Program) (No. 2007A A02Z103), National Natural Science Foundation of China (No. 30870050), National Key Technology R&D Program (No. 2008BAI66B03), Scientific and Technological Major Special Project-Significant Creation of New Drugs (No. 2009ZX09031-002).

Corresponding author: Jun Wu. Tel: +86-10-66948825; Fax: +86-10-63833521; E-mail: junwu1969@163.com

国家高技术研究发展计划 (863 计划) (No.2007AA02Z103), 国家自然科学基金 (No. 30870050), 国家科技支撑计划 (No. 2008BAI66B03), 国家科技重大新药创制 (No. 2009ZX09031-002) 资助。

N-glycan structures from Man₈GlcNAc₂ to Man₅GlcNAc₂. The plasmids contained MDSI genes from *Homo sapiens* [HMDSI(Δ 185)] or *Arabidopsis thaliana* [ATMDSI(Δ 48)], and three ER-signals were used to be transformed a mutant *Pichia pastoris* GJK01, respectively. The reporter protein HSA/GM-CSF (human serum albumin and granulocyte-macrophage colony stimulating factor fusion protein) was expressed and its N-glycans were analyzed by DSA-FACE (DNA sequencer assisted fluorophore-assisted carbohydrate electrophoresis). The plasmid contained ER-ScMnsI-ATMDSI(Δ 48) was expressed in *Pichia pastoris*, the Man₅GlcNAc₂ N-glycan on secreted glycoprotein HSA/GM-CSF was observed. The research reported here provided basic substrate to obtain the hybrid- and complex-type glycans in mammalian cell.

Keywords: α -1,2-mannosidase I, Man₅GlcNAc₂, glycosylation, *Pichia pastoris*, mannose-type

近年来, 药用蛋白市场的需求越来越大, 在这些蛋白中糖蛋白占绝大多数, 例如各种治疗性抗体、细胞因子等。糖蛋白是糖基与蛋白共价相连构成的结合蛋白, 糖蛋白的糖基与蛋白的功能、稳定性均有着密切的联系^[1]。糖蛋白根据糖基和蛋白质的连接方式不同, 可分为 2 大类, 即 O-连接和 N-连接糖蛋白。其中, 对 N-连接糖蛋白的糖基化修饰研究得比较透彻, 其修饰序列极端保守, 即在 Asn-X-Thr/Ser (X 为除 Pro 外的任意氨基酸) 的 Asn 残基上。

目前重组糖蛋白药物的生产主要是哺乳动物细胞表达系统。但该系统培养成本高、生产周期长^[2]。因此, 利用酵母细胞来替代哺乳动物细胞生产重组蛋白药物已日益受到人们的重视。酵母表达系统具有原核细胞系统生长速度快、便于基因操作和可工业化大规模培养等优点, 同时又具有真核细胞大部分的翻译后加工修饰能力, 因而广泛用于各种蛋白的表达^[3]。巴斯德毕赤酵母属于甲基营养型酵母, 是第 2 代酵母表达系统的代表, 作为成熟的表达系统, 它还具有表达量高、较稳定、可分泌、易纯化及成本低等优点^[4-7], 但目前主要应用于非糖蛋白的生产, 这是因为当毕赤酵母表达系统用于多数重组糖蛋白药物的生产时, 还存在一些问题, 其中最关键的是毕赤酵母对糖蛋白的过度糖基化修饰, 即形成不同于哺乳动物细胞的杂合型和复杂型糖基结构, 而是高甘露糖型糖基或过度糖基化结构^[8-9], 这种高甘露糖型糖基结构的糖蛋白在人体中容易被清除、半衰期较短、且易产生较高的免疫原性; 同时对某些蛋白过度的糖基化修饰也会造成产物的分子量不均一, 或者一些活性位点的遮蔽^[10], 影响其功

能和活性。这些都限制了毕赤酵母在糖蛋白类药物生产方面的应用。

酵母和哺乳动物 N-糖基化修饰途径在内质网中的起始步骤都是相同的, 在内质网中, 在寡糖转移酶作用下, Glc₃Man₉GlcNAc₂ 被连接到新生肽链的专一序列上, 即 Asn-X-Thr/Ser 的 Asn 残基上, 随后在葡萄糖苷酶 I、II 和内质网甘露糖苷酶 I (MnsI) 的作用下, 蛋白的糖基最终被加工成 Man₈GlcNAc₂ 糖基, 并转运至高尔基体。在高尔基体内, 酵母和哺乳动物糖基修饰过程明显不同, 在酵母高尔基体中, 由于存在一个关键起始酶 α -1,6-甘露糖转移酶 (Och1p), Man₈GlcNAc₂ 糖基在它的作用下接受第一个 α -1,6-甘露糖, 形成 Man₉GlcNAc₂ 糖基, 该糖基是 α -1,2-、 α -1,3 甘露糖转移酶及一些磷酸甘露糖转移酶的底物, 在这些甘露糖转移酶的作用下, 蛋白的糖基可以被加至数十至上百个甘露糖, 形成高甘露糖, 发生过度糖基化修饰^[10]; 而在哺乳动物高尔基体内, 蛋白的糖基在 α -1,2 甘露糖苷酶 I (MDSI) 的作用下, Man₈GlcNAc₂ 糖基会被剪切去除 3 个 α -1,2 连接的甘露糖, 而形成 Man₅GlcNAc₂ 糖基, 这一糖基是哺乳动物细胞合成杂合型和复杂型糖基的前体, 也是哺乳动物表达的甘露糖型糖蛋白的主要糖型, 如牛核糖核酸酶 B (Ribonuclease B)。而这种糖基在酵母中并不存在, 所以这一结构也是酵母和哺乳动物 N-糖基化修饰过程的“分界线”。

我们前期已经敲除了 Och1 这一形成高甘露糖型糖基结构的关键酶基因, 获得了过度糖基化缺失的毕赤酵母菌株 GJK01 (*ura3 och1*)^[11]。本研究拟通过选择不同物种来源 MDSI 及其定位信号, 并实

现其在 GJK01 (*ura3*、*och1*) 中的融合表达, 以获得从酵母糖基向哺乳动物糖基改造的第一个关键糖型, 即 $\text{Man}_5\text{GlcNAc}_2$ 结构。

1 材料与amp;方法

1.1 材料

1.1.1 质粒和菌株

α -1,6-甘露糖转移酶 (*Och1p*) 基因缺陷的毕赤酵母 GJK01 (*ura3*、*och1*) 菌株由本室构建。HSA/GM-CSF 表达载体 pHIL-D2/HSA/GM-CSF、PGE1203-URA3 均由本室构建保存^[11]。载体 pPICZ α A、pIB2、毕赤酵母 GS115 均购自 Invitrogen 公司。HepG2 细胞由本室保存。

1.1.2 试剂和仪器

培养基和试剂: 酵母抽提物、蛋白胨均为 Oxoid 公司产品; 无氨基酸酵母氮源 (Yeast nitrogen base without amino acids, YNB) 为 Difco 公司产品; 所用限制性内切酶、T4 DNA 连接酶和去磷酸化酶 (CIAP) 均为大连宝生物工程有限公司产品; 牛核糖核酸酶 B (RNase B)、氢化硼氰化钠 (NaBH_3CN) 均为 Sigma 产品; Sephadex-G10 柱色谱填料 (AB); DNA 3100 测序仪 (ABI); 8-氨基苝基-1,3,6-三磺酸 (8-aminopyrene-1,3,6-trisulfonate, APTS) (Molecular probes); 碳黑柱 (Alltech associates)。

1.2 α -1,2-甘露糖苷酶 (MDSI) 表达载体的构建

1.2.1 PGE-URA3-GAP1 载体构建

GAP 启动子基因的获取: 以 pGAP01 和 pGAP02 为引物, 以毕赤酵母表达载体 pIB2 为模板, PCR 扩增 500 bp GAP 启动子基因片段; CY CTT 终止子基因的获取: 以 pCYCTT01 和 pCYCTT02 为引物, 以毕赤酵母表达载体 pPICZ α A 为模板, PCR 扩增 300 bp 的 CYCTT 终止子基因片段; 然后通过引物 pGAP01 和 pCYCTT02, 利用 PCR 方法融合 GAP 和 CYCTT, 获得 GAP 表达盒, 片段大小约 800 bp。利用 *Kpn* I /*Xba* I 酶切该片段, 并与经过同样酶切的 pGE1203-URA3 进行连接, 获得含 GAP 表达盒的载

体 pGEURA3-GAP。利用一对互补配对的寡核苷酸引物 1203MSC5、1203MSC3, 退火形成双链, 构建至 pGEURA3-GAP 载体中, 获得含有多克隆位点 *Ssp* I-*Sse*838 I-*Swa* I 的载体, 命名为 pGEURA3-GAP1 载体。

1.2.2 载体 PGE-URA3-GAP1-signal-MDSI 的构建

定位信号的克隆: 以酿酒酵母基因组为模板, 以 pMnsI5、pMnsI3 为引物, PCR 扩增 90 bp 的 MnsI 定位信号, 以 pSCsec12-5、pSCsec12-3 为引物, PCR 扩增 300 bp 的 Sec12 定位信号; 以乳酸克鲁维酵母基因组为模板, 以 pKLsec12-5、pKLsec12-3 为引物, PCR 扩增 300 bp 的 Sec12 定位信号。将扩增得到的 3 种定位信号基因片段分别用 *Sse*8387 I /*Ssp* I 酶切, 构建至同样酶切的 pGE-URA3-GAP1 载体中, 获得 3 种 pGE-URA3-GAP1-signal 载体。

MDSI 基因的克隆: 从人的肝癌细胞 HepG2 中提取总 RNA, 以 pHmsdi(Δ 185)、pHMDSI3 为引物, RT-PCR 扩增 1 800 bp 人源的编码缺失 N 端 185 个氨基酸但包含完整 C 端催化区的 MDSI 基因片段, 记为 HMDSI(Δ 185); 从植物拟南芥叶片中提取总 RNA, 以 pATMDSI5(Δ 48)、pATMDSI3 为引物, RT-PCR 扩增 1 700 bp 拟南芥的编码缺失 N 端 48 个氨基酸但包括完整 C 端催化区的 MDSI 基因片段, 记为 ATMDSI(Δ 48)。利用 *Sse*8387 I /*Swa* I 酶切位点将 2 种基因片段分别构建至 3 种 pGE-URA3-GAP1-signal 载体中, 获得 6 种载体 PGE-URA3-GAP1-signal-MDSI。

1.2.3 表达载体 PGE-URA3-PNOI-GAP-signal-MDSI 的构建

为了实现 MDSI 表达载体在酵母基因组中的定点整合, 本研究选取了磷酸甘露糖转移酶 PNO1 基因开放读码框 ORF 上游同源序列作为整合位点。因此, 以毕赤酵母 GS115 基因组为模板, 以 PNOI5-5、Pno 为引物, PCR 扩增 PNO1 基因 ORF 上游序列, 片段大小为 1 000 bp 左右, 电泳回收, 然后, 再以回收产物与 PNOI3-3 为引物, 以毕赤酵母 GS115 基

表 1 α -1,2-甘露糖苷酶 (MDSI) 表达载体的构建引物Table 1 Primer used for constructing α -1,2-mannosidase I (MDSI) expression plasmid

Gene name	Primer name	Primer sequence (5'-3')
CYCTT terminator	pCYCTT01	GAACAACATAT AATATTC ACGTCCGACGGCGGCCCA
	pCYCTT02	CATCTAGAGC GCGGCCG CAGCTTGCAAATTAAGCCTTCGA
GAP promoter	pGAP01	ATCGGTACCGCGGCCGAGATCCTTTTTTTGTAGAAATGTC
	pGAP02	GTG AATATT TATAGTTGTTCAATTGATTGA
SCMnsI signal	pMnsI5	AAACGATGAAGAACTCTGTCTGGTAT
	pMnsI3	AATCCTGCAGGTCTCTCAAAGTGTTCGTACC
SCSec12 signal	pSCsec12-5	AAACGATGGCCAACACTATCCACAT
	pSCsec12-3	AATCCTGCAGGCACGCCATGCACTTTTATGG
KL Sec12 signal	pKLsec12-5	AAACGATGGCAAACACCGTTCACAT
	pKLsec12-3	AATCCTGCAGGGAAAGAAATAGTGGTTGATTCTGAGTGGCC
HMDSI	pHmsi(Δ 185)	AATCCTGCAGGCGAGCCC GCCGACGCCGCCAT
	pHMDSI3	GGC ATTTAAAT TCTATTATTCCTCTCTGATTTCAAC
ATMDSI	pATMDSI5(Δ 48)	AATCCTGCAGGCCTTGCCCGAGAACATGAGGTTGAA
	pATMDSI3	GGC ATTTAAAT TACTAAACGTTAATCTGATGACC
Multiple cloning sites	1203MSC5	ATTACGCCTGCAGGCTGA ATTTAAATC
	1203MSC3	G ATTTAAAT TTCAGCCTGCAGGCGTAAT
PNOI	PNOI5-5	ATTCTAGAAGCCAGAATTCTAATATGTAACACTAC
	Pno	CATGAGCCAACCAAGTGGGAGAGCGGCCGCTTAATAGTTACGTACGAGCAATGTT
	PNOI3-3	ACGGTACCTCTTGAAGTAGATTTGGAGATTTTG

基因组为模板, PCR 扩增 2 000 bp 左右基因片段, 电泳回收, 回收产物利用 *Kpn* I /*Xba* I 酶切位点克隆至 pGE-URA3-GAP 载体中, 获得 pGE-URA3-PNOI 载体; *Not* I 酶切 6 个基因载体 PGE-URA3-GAP-signal-MDSI, 将获得的带有 MDSI 基因的表达盒克隆到限制性内切酶 *Not* I 酶切并 CIAP 去磷酸化的 pGE-URA3-PNOI 载体中, 获得表达载体 PGE-URA3-PNOI-GAP1-signal-MDSI。

1.3 HSA/GM-CSF (His₆-tag) 在过度糖基化缺陷菌 GJK01 (*ura3*、*och1*) 中的表达及鉴定

制备 GJK01 (*ura3*、*och1*) 的电转化感受态细胞, 将 pHIL-D2/HSA/GM-CSF(His₆-tag) 以限制性内切酶 *Not* I 线性化后电转入感受态细胞中, 将电击后的菌液涂布于含有尿嘧啶和精氨酸的 MD 培养基 (YNB 1.34%, 生物素 4×10^{-5} %, 葡萄糖 2%, 琼脂 1.5%, 精氨酸 100 μ g/mL, 尿嘧啶 100 μ g/mL) 上。5~8 d 后, 随机挑取克隆接种到 2 mL YPD 培养基中, 25 $^{\circ}$ C 培养 48 h 后, 以 5% 的接种量接种到 BMGY

培养基中, 24 h 后加入 0.5% 甲醇进行诱导表达, 每 12 h 补加 1 次甲醇, 诱导 72 h 后离心取上清, 上清用于融合蛋白的表达分析, 筛选 HSA/GM-CSF(His₆-tag) 在缺陷菌 GJK01 (*ura3*、*och1*) 中的表达株。

1.4 MDSI 在 GJK01(*ura3*、*och1*、pHIL-D2/HSA/GM-CSF(His₆-tag) 中表达及鉴定

制备 HSA/GM-CSF(His₆-tag) 在 GJK01 (*ura3*、*och1*) 中的表达株的感受态细胞, 将 6 种 PGE-URA3-PNOI-GAP1-signal-MDSI 分别以位于 PNOI 基因 ORF 上游手臂上的限制性内切酶 *Bam*H I 酶切, 线性化后分别电转入感受态细胞中, 将电击后的菌液涂布于含有精氨酸的 MD 培养基 (YNB 1.34%, 生物素 4×10^{-5} %, 葡萄糖 2%, 琼脂 1.5%, 精氨酸 100 μ g/mL) 上。5~8 d 后, 对长出的 URA3⁺ 菌进行基因组鉴定, 鉴定引物为 MDSI 钓取引物。将鉴定得到的 MDSI 转入的阳性菌接种到 2 mL YPD 培养基中, 25 $^{\circ}$ C 培养 48 h 后, 以 5% 的接种量接种到 BMGY 培养基中, 24 h 后加入 0.5% 甲醇进行诱

导表达, 每 12 h 补加一次甲醇, 诱导 72 h 后离心取上清, 上清备用。

1.5 HSA/GM-CSF(His₆-tag) 的纯化

收集培养上清, 利用镍亲和层析方法纯化末端带有 6 个 His 标签的报告蛋白 HSA/GM-CSF。流动相为 20 mmol/L Tris-HCl, 0.5 mol/L NaCl, pH 7.5, 洗脱液另含 0.5 mol/L 咪唑。在 280 nm 处紫外吸收峰值收集纯化产物, 10% 的 SDS-PAGE 鉴定表达产物。

1.6 HSA/GM-CSF(His₆-tag) 的糖基分析

1.6.1 HSA/GM-CSF(His₆-tag) 糖基的释放

蛋白在 0.5% SDS 和 1% β-巯基乙醇变性缓冲液中 100 °C 变性 10 min, 然后在含有 50 mmol/L 磷酸钠 (pH 7.5)、1% NP-40、N-糖苷酶 F (PNGase F) 的溶液中 37 °C 酶切 16 h, 使 HSA/GM-CSF(His₆-tag) 释放糖基。

1.6.2 HSA/GM-CSF(His₆-tag) 糖基的纯化

用 80% 乙腈、0.1% 三氟醋酸活化碳黑柱, 20% 乙腈、0.05% 三氟醋酸洗脱, 所得糖链真空冷冻抽干备用。

1.6.3 APTS 标记糖链

取糖链样品, 加 0.02 mol/L 的 APTS 溶液 1 μL [溶于 1.2 mol/L 柠檬酸], 1 mol/L 的 NaBH₃CN 溶液 1 μL (溶于 DMSO)。混匀, 封管, 置 37 °C 水浴反应 18 h。

1.6.4 APTS 标记糖链 Sephadex G10 纯化

据文献[12]报道该方法简单, 可除去约 90% 的单体 APTS, 同时能保留约 70% 以上的标记物复合物。样品经 G10 两次纯化后真空抽干。

1.6.5 3100 DNA 测序仪分析 APTS 标记糖链^[17]

5 μL 纯水溶解 G10 纯化且冷冻抽干的 APTS 标记糖链。最终待测样品中含 1 μL 样品, 8 μL 去离子甲酰胺和 1 μL ROX 修饰内标混合物, 其中 ROX 内标混合物是指由罗丹明修饰的 6-、30-寡核苷酸 (由 5'-TAC-3' 重复序列组成, 由 Invitrogen 公司 PAGE 纯化得到)。上样后利用标准的 36 cm 毛细管及 POP-6 胶 (ABI) 分离, 并利用 GeneScan 3.7 软件进行数据分析。

2 结果

2.1 HSA/GM-CSF(His₆-tag) 在 *P. pastoris* GJK01 (*ura3*、*och1*) 中表达

为了分析外源的 MDSI 在 GJK01 (*ura3*、*och1*) 中的活性, 本研究以 HSA/GM-CSF(His₆-tag) 作为报告蛋白。为此, 必须获得 HSA/GM-CSF(His₆-tag) 在 GJK01 (*ura3*、*och1*) 中的表达菌株。将表达质粒 pHIL-D2/HSA/GM-CSF(His₆-tag) 转入毕赤酵母 GJK01 (*ura3*、*och1*) 后, 随机挑取克隆, 接种至 BMGY 进行诱导表达, 取培养上清进行 SDS-PAGE 分析, 1、2、4、5、7 号克隆在分子量接近 97 kDa 处出现一条蛋白条带 (图 1), 而非糖基化的 HSA/GM-CSF(His₆-tag) 分子量为 85 kDa, 提示该条带为糖基化的 HSA/GM-CSF(His₆-tag)。



图 1 HSA/GM-CSF(His₆-tag) 在 *P. pastoris* GJK01 (*ura3*、*och1*) 中表达的 SDS-PAGE 图谱

Fig. 1 The SDS-PAGE for identification of HSA/GM-CSF expressed in *P. pastoris* GJK01 (*ura3*, *och1*). M: molecular weight marker; 1, 2, 4, 5, 7: positive transformants; 3, 6, 8, 9: negative transformants.

2.2 α-1,2-甘露糖苷酶 (MDSI) 表达载体的构建及其转化 GJK01 (*ura3*、*och1*、*pHIL-D2/HSA/GM-CSF(His₆-tag)*) 基因组鉴定

构建的 α-1,2-甘露糖苷酶 (MDSI) 表达载体 (图 2), 包含 PNOI 基因 ORF 上下游各 1 000 bp 的同源臂、GAP 启动子、目的基因及其定位信号、CYCTT 终止子。其中, 两种不同来源的 MDSI 基因: 人源的编码缺失 N 端 185 个氨基酸但包含完整 C 端催化区的 MDSI 基因 [HMDSI(Δ185)] 和拟南芥的编码缺失 N 端 48 个氨基酸但包含完整 C 端催化区的 MDSI

基因[ATMDSI(Δ 48)]; 3 种不同的定位信号: 酿酒酵母 MnsI (ScMnsI) 的内质网定位信号、酿酒酵母 Sec12 (Scsec12) 的内质网定位信号以及乳酸克鲁维酵母 Sec12 (KLSec12) 的内质网定位信号, 分别组合, 构成 6 种表达载体 (表 2)。

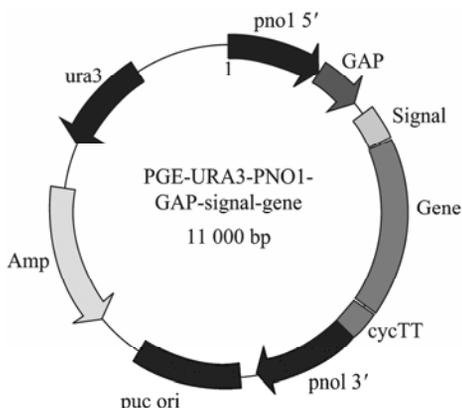


图 2 PGE-GAP-URA3-PNOI-signal-MDSI 的构建
Fig. 2 Construction of the plasmid PGE-GAP-URA3-PNOI-signal-MDSI.

表 2 MDSI 表达载体

Table 2 MDSI expression plasmid

	Gene (MDSI)	Signal	Promoter	Vector
1	HMDSI(Δ 185)	ScmnsI	GAP	PGE-URA3-PNOI
2	HMDSI(Δ 185)	Scsec12	GAP	PGE-URA3-PNOI
3	HMDSI(Δ 185)	KLsec12	GAP	PGE-URA3-PNOI
4	ATMDSI(Δ 48)	ScmnsI	GAP	PGE-URA3-PNOI
5	ATMDSI(Δ 48)	Scsec12	GAP	PGE-URA3-PNOI
6	ATMDSI(Δ 48)	KLsec12	GAP	PGE-URA3-PNOI

定位信号扩增的片段, 琼脂糖核酸电泳如图 3A 所示, 在 90 bp 左右、300 bp 左右有特异的扩增产物片段, 分别与 ScMnsI、Scsec12、KLSec12 的定位信号大小一致; MDSI 基因克隆的片段, 琼脂糖核酸电泳如图 3B、图 3C 所示, 在 1 800 bp 左右、1 700 bp 左右有特异的扩增片段, 分别与 HMDSI (Δ 185)、ATMDSI(Δ 48)大小一致。最终获得的表达载体 PGE-URA3-PNOI-GAP1-signal-MDSI 经 *Not* I 酶切分析, 发现有约 2.5~3.0 kb 的基因表达盒片段

与理论值一致 (图 4)。并且经过测序分析显示表达载体 PGE-URA3-PNOI-GAP1-signal-MDSI 构建成功。

将 3 种定位信号 (ScMnsI 的定位信号、Scsec12 的定位信号、KLSec12 的定位信号) 与 2 种来源的 MDSI (人源的及拟南芥来源) 表达载体电转化 GJK01(*ura3*、*och1*、*pHIL-D2/HSA/GM-CSF(His₆-tag)*) 菌, 对转化菌基因组进行 PCR 鉴定, 筛选 MDSI 基因整合到酵母染色体上的重组菌, 1 800 bp 左右的条带与 MDSI 大小一致 (图 5)。

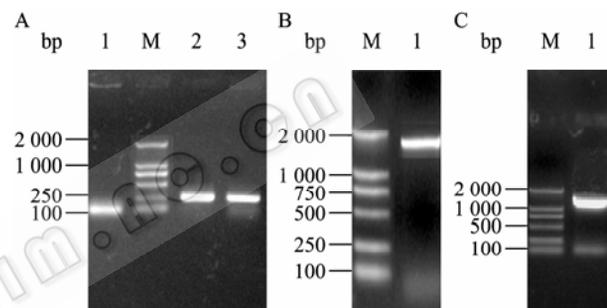


图 3 定位信号及 MDSI 基因 PCR 扩增分析

Fig. 3 PCR analysis of signal and MDSI. M: DL2000 DNA marker. (A) 1: ScmnsI; 2: Scsec12; 3: KLsec12. (B) HMDSI(Δ 185). (C) ATMDSI(Δ 48).

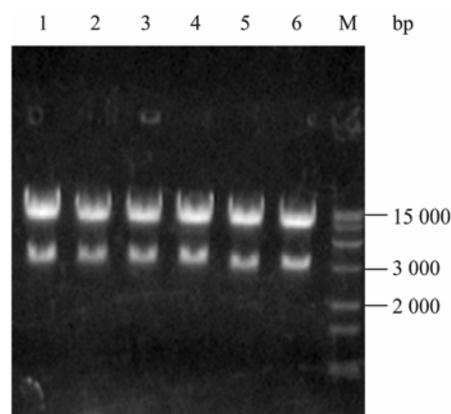


图 4 PGE-GAP-URA3-PNOI-signal-MDSI 载体质粒的 *Not* I 酶切图谱

Fig. 4 Characterization of PGE-GAP-URA3-PNOI-signal-MDSI digestion by *Not* I. M:15 kb DNA marker; 1: Scsec12-ATMDSI(Δ 48); 2: KLsec12-ATMDSI(Δ 48); 3: Scsec12-HMDSI(Δ 185); 4: KLsec12-HMDSI(Δ 185); 5: ScMnsI-ATMDSI(Δ 48); 6: ScMnsI-HMDSI(Δ 185).

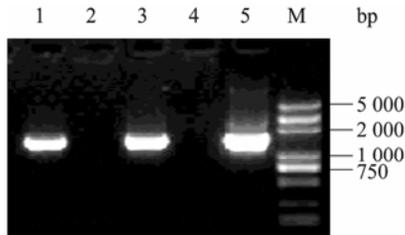


图 5 MDSI 转化 GJK01 (*ura3*、*och1*、*pHIL-D2/HSA/GM-CSF(His₆-tag)*) 的 PCR 鉴定

Fig. 5 PCR identification of MDSI transformed in GJK01(*ura3*, *och1*, *pHIL-D2/HSA/GM-CSF(His₆-tag)*). 1-4: GJK01(*ura3*, *och1*, *pHIL-D2/HSA/GM-CSF(His₆-tag)*) transformed with MDSI; 5: positive control; M: 2000 plus DNA marker.

2.3 HSA/GM-CSF(His₆-tag) 糖链的制备与分析

利用 PNGaseF 酶对 MDSI 转化 GJK01/ (*ura3*、*och1*、*pHIL-D2/HSA/GM-CSF(His₆-tag)*) 基因组鉴定正确的克隆中表达的报告蛋白 HSA/GM-CSF(His₆-tag) 进行酶切, 利用 SDS-PAGE 对酶切的报告蛋白分析, 发现 HSA/GM-CSF 的糖基经 PNGaseF 酶切后, 分子量减小, 说明其 N-糖链被 PNGaseF 酶切除 (图 6)。

PNGaseF 酶切报告蛋白 HSA/GM-CSF(His₆-tag) 后, 样品经碳黑柱纯化获得糖链, APTS 标记糖链。由于单体 APTS 的存在会影响糖链分析, 所以再用

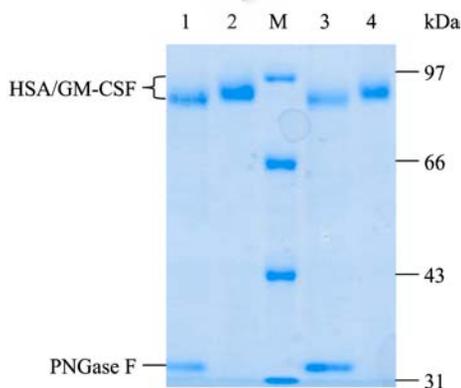


图 6 PNGaseF 酶切报告蛋白 HSA/GM-CSF (His₆-tag) 的 SDS-PAGE 图谱

Fig. 6 The SDS-PAGE for characterization of HSA/GM-CSF digested by PNGaseF. 1, 3: HSA/GM-CSF expressed in GJK01 (*ura3*, *och1*, *ScmnsI*, *ATMDSI*, *pHIL-D2/HSA/GM-CSF*) digestion by PNGaseF; 2, 4: HSA/GM-CSF expressed in GJK01 (*ura3*, *och1*, *ScmnsI*, *ATMDSI*, *pHIL-D2/HSA/GM-CSF*); M: molecular weight marker.

Sephadex G10 柱纯化标记糖链, 以去除单体 APTS。然后利用 3100 DNA 测序仪分析 APTS 标记、纯化的报告蛋白糖链, 并利用 GeneScan 3.7 软件进行数据分析, 分析结果如图 7 所示。

从图 7A 中可以看出, α -1,6-甘露糖转移酶 (Och1p) 基因缺陷的毕赤酵母 GJK01 (*ura3*、*och1*) 表达的报告蛋白 HSA/GM-CSF(His₆-tag) 的糖基是 Man₁₂GlcNAc₂~Man₁₆GlcNAc₂ 高甘露糖型糖基, 其中 Man₁₄GlcNAc₂、Man₁₅GlcNAc₂ 占绝大部分。而将酿酒酵母 *MnsI* 基因内质网定位信号与编码拟南芥 MDSI(Δ 48) 基因进行融合, 并将这一融合基因整合到 GJK01 基因组诱导表达后, 糖基分析发现拟南芥 MDSI 能在酵母 GJK01 (*ura3*、*och1*) 宿主菌中发挥了较好的生物活性, 能够将酵母 GJK01 中表达的报告蛋白的糖基 Man₁₂GlcNAc₂~Man₁₆GlcNAc₂ 绝大部分剪切成 Man₅GlcNAc₂~Man₉GlcNAc₂ 结构 (图 7B), 这与哺乳动物来源的牛核糖核酸酶 B 的 Man₅GlcNAc₂~Man₉GlcNAc₂ 甘露糖型糖基一致 (图 7C)。通过本研究, 我们利用酿酒酵母 *MnsI* 内质网定位信号与外源拟南芥 MDSI(Δ 48) 基因的融合表达, 在 GJK01 (*ura3*、*och1*) 基础上, 构建了具有合成 Man₅GlcNAc₂ 哺乳动物甘露糖型糖蛋白能力的毕赤酵母表达系统。这为在酵母体内合成类似于哺乳动物杂合型或复杂型糖基奠定了基础。

3 讨论

酵母糖基工程改造是在酵母中模拟哺乳动物的糖基修饰过程, 即阻断酵母形成高甘露糖型糖基, 引入外源的参与糖基修饰的酶基因, 并将其定位在适当位置, 使其发挥活性对糖基进行修饰, 从而获得类似于哺乳动物杂合型或复杂型糖基的过程。哺乳动物中的糖苷酶或糖基转移酶对糖基修饰是逐个顺序完成的。Man₅GlcNAc₂ 是杂合型糖基修饰酶 N-乙酰葡萄糖胺转移酶 (GnTI) 的底物, 在 GnTI 的作用下, Man₅GlcNAc₂ 上添加一个 N-乙酰葡萄糖胺, 形成 GlcNAcMan₅GlcNAc₂ 杂合型糖基。因而获得

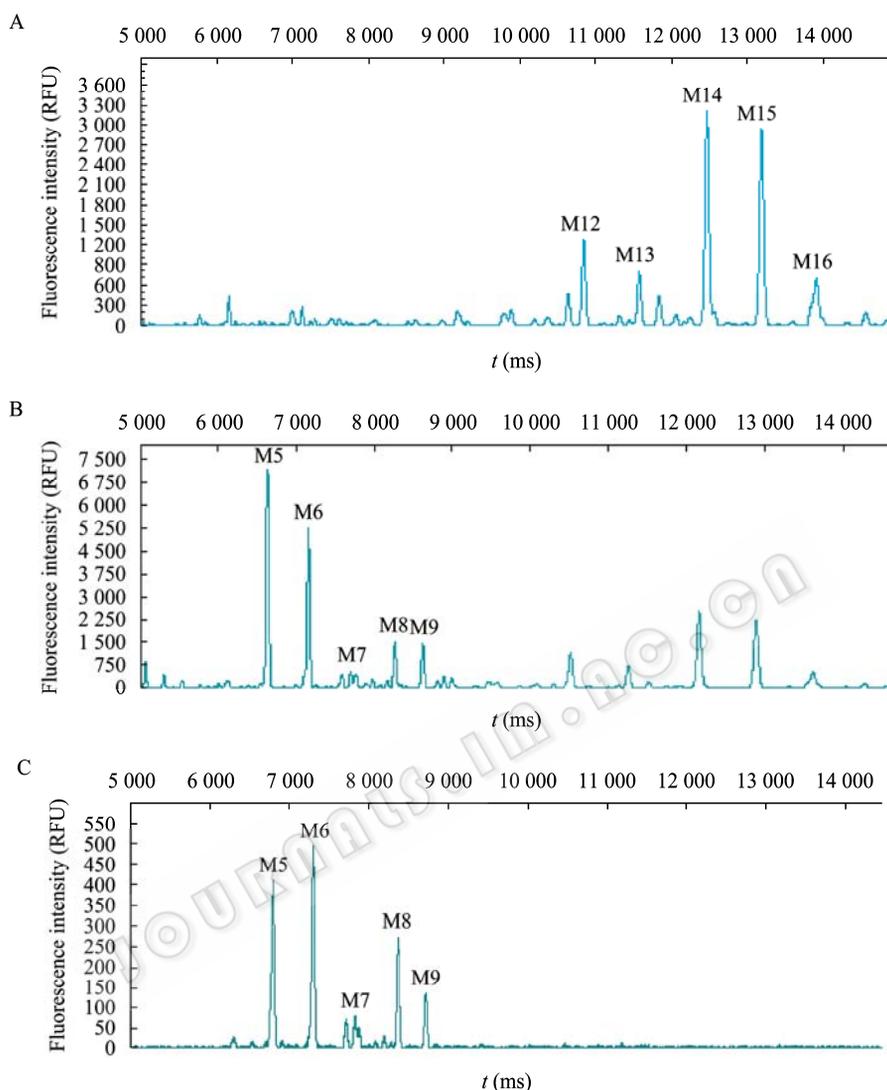


图 7 DSA-FACE 分析报告蛋白 HSA/GM-CSF(His₆-tag) 的 N-糖基结构

Fig. 7 Analysis of N-Glycan on HSA/GM-CSF(His₆-tag) by DSA-FACE. (A) N-Glycan of GJK01(*ura3, och1, pHIL-D2/HSA/GM-CSF(His₆-tag)*). (B) N-Glycan of GJK01(*ura3, och1, ScmnsI-ATmSI pHIL-D2/HSA/GM-CSF(His₆-tag)*). (C) N-Glycan of RNaseB M5–M9: $\text{Man}_5\text{GlcNAc}_2$ – $\text{Man}_9\text{GlcNAc}_2$; M12–M16: $\text{Man}_{12}\text{GlcNAc}_2$ – $\text{Man}_{16}\text{GlcNAc}_2$.

$\text{Man}_5\text{GlcNAc}_2$ 糖基是酵母糖基化改造中的关键一步, 没有 $\text{Man}_5\text{GlcNAc}_2$ 结构, GnTI 就很难发挥作用进而会影响到后续酶的进一步加工。酵母糖基工程改造是一个复杂的系统工程, 这其中涉及到诸多因素。

例如要实现外源基因的高效表达, 需解决以下 3 方面的问题。第一, 基因来源的选择, 作为异源表达必须考虑外源基因在酵母中的活性问题, 即

酵母内环境、pH 值、最适温度等等。我们优先选择了哺乳动物自身来源的基因以获得哺乳动物的糖基结构, 如本研究中的 MDSI 基因。但由于人体温度在 37 °C 左右, 与酵母生长温度 25 °C~30 °C 存在差异, 因此又选择了植物来源的基因, 其生存环境温度范围很广, 包含了酵母生长温度。参与糖基化修饰的酶都是 II 型膜蛋白 (如 MDSI), II 型膜蛋白的共同特点是, 具有 N 端的氨基端细胞质尾、穿膜锚

定区、茎区以及 C 端的催化区^[13]。本研究中我们保留了不同长度的茎区和完整的 C 端催化区。其次，外源基因的准确定位问题。糖蛋白的糖基是蛋白经过内质网，内侧、中间和外侧高尔基体过程中逐个顺序添加上去的，因此，模拟哺乳动物的糖基修饰过程，将酶定位在准确位置是必要的。为了将 MDSI 定位于糖基化修饰的初始环境中，即内质网或内侧高尔基体，本研究选择酿酒酵母 MnsI、Sec12 的定位信号和乳酸克鲁维酵母 Sec12 的定位信号。第三，启动子问题，我们选择了 GAP 启动子。一般用于毕赤酵母表达的醇氧化酶 (AOX) 启动子属于强启动子，它需要甲醇诱导，存在一定的毒性隐患，不利于操作；GAP 启动子不需要甲醇诱导，是一种中等强度的组成型启动子，是一种成熟的用于酵母表达的启动子，例如人壳多糖酶基因在 GAP 调控下连续表达超过 1 个月，表达量稳定在 300 mg/L^[14]。同时，GAP 启动子在 AOX 启动报告蛋白 (HSA/GM-CSF) 之前，就启动 MDSI 基因的表达，使得 MDSI 能更好地修饰随后表达的报告蛋白。

本室在构建酵母缺陷菌时，利用了 URA3 作为营养缺陷型筛选标记，URA3 是一种常用的筛选标记，可以用来进行正负筛选。缺陷菌 GJK01 为 URA3 缺陷型^[11]，本研究构建的表达载体以 URA3 为筛选标记，当 MDSI 表达载体被转入工程菌 GJK01 中时，在目的基因 MDSI 被整合至基因组的同时，URA3 也被整合至基因组，重组酵母菌可以在无尿嘧啶的培养基上生长，得到 URA3⁺MDSI 菌株；进一步可以利用 5-氟乳清酸 (5-FOA) 使 URA3⁺菌株发生第 2 次重组，即 URA3 基因重新被剔除出基因组，从而再次获得 URA3 的缺陷菌株，该筛选标记将运用于下一次基因的敲除。如本实验前期利用 URA3 基因已经成功敲除了毕赤酵母的 Ocl1p^[11]和乳酸克鲁维酵母的 Ocl1p 和 Mnn1p (α -1,3-甘露糖转移酶)^[15]。

HSA/GM-CSF(His₆-tag) 是本实验室构建的 HSA 与 GM-CSF 的融合蛋白，GM-CSF 即粒细胞-巨噬细胞集落刺激因子，由 127 个氨基酸组成，有 2

个 N-糖基化位点 (Asn27、Asn37)，天然的 GM-CSF 约有 34% 的糖基化^[16]。为了对糖基进行有效分析，本研究采用了基于 DNA 测序仪荧光辅助糖电泳 (DNA sequencer assisted FACE, DSA-FACE)^[14]的方法，现在用于糖基分析的方法大都是荧光辅助糖电泳 (FACE)、质谱 (MS)、核磁共振 (NMR)，这些方法需要大量的样品，而且很难做到高通量。DSA-FACE 是借助于 DNA 测序仪的一种新型的荧光毛细管电泳方法，在常规的 FACE 的基础上建立起来的一种高灵敏度的糖分析方法。它用毛细管电泳代替了传统的 PAGE 胶，用荧光检测器代替了紫外灯下的荧光观察，从而在分离能力、灵敏度上都有了很大提高。DSA-FACE 方法借助 DNA 测序仪，如 ABI 3100，可以实现了 96 个样品的同步分析，因而可以满足糖工程和糖组学高通量筛选的要求。

REFERENCES

- [1] Helenius A, Aebi M. Intracellular functions of N-linked glycans. *Science*, 2001, 291(5512): 2364-2369.
- [2] Hamilton SR, Bobrowicz P, Brobrowicz B, et al. Production of complex human glycoproteins in yeast. *Science*, 2003, 301(5637): 1244-1246.
- [3] Cregg JM, Cereghino JL, Shi JY, et al. Recombinant protein expression in *Pichia pastoris*. *Mol Biotechnol*, 2000, 16(1): 23-52.
- [4] Hitzeman RA, Hagie FE, Levine HL, et al. Expression of a human gene for interferon in yeast. *Nature*, 1981, 293(5835): 717-722.
- [5] OuYang LM, Zhang HZ, Zhang ST, et al. Advances in the studies of *Pichia pastoris* as a heterologous gene expression system. *Prog Biochem Biophys*, 2000, 27(2): 151-154.
欧阳立明, 张惠展, 张嗣良, 等. 巴斯德毕赤酵母的基因表达系统研究进展. *生物化学与生物物理进展*, 2000, 27(2): 151-154.
- [6] Peng Y, Yang XC, Kang LY. Factors affecting expression of heterologous protein in methylotrophic yeast. *Biotechnol Inf*, 2000, 4: 33-36.
彭毅, 杨希才, 康良仪. 影响甲醇酵母中外源蛋白表达的因素. *生物技术通报*, 2000, 4: 33-36.

- [7] Brethauer RK, Castellino FJ. Glycosylation of *Pichia pastoris*-derived proteins. *Biotechnol Appl Biochem*, 1999, 30: 193–200.
- [8] Hubbard SC, Ivatt RJ. Synthesis and processing of asparagine-linked oligosaccharides. *Anna Rev Biochem*, 1981, 50 (7): 555–583.
- [9] Wildt S, Gerngross TU. The humanization of *N*-glycosylation pathways in yeast. *Nature Microbiol*, 2005, 3(2): 119–128.
- [10] Walsh G. Biopharmaceutical benchmarks-2003. *Nature Biotechnol*, 2003, 21(8): 865–870.
- [11] Wang Y, Liu B, Wu J, et al. A *Pichia pastoris* with α -1,6-mannosyltransferase deletion and its use in expression of HSA/GM-CSF Chimera. *Chin J Biotech*, 2007, 23(5): 907–914.
王越, 刘波, 吴军, 等. α -1,6-甘露糖转移酶基因敲除的毕赤酵母菌株构建及其用于融合蛋白 HSA/GM-CSF 表达的研究. *生物工程学报*, 2007, 23(5): 907–914.
- [12] Laroy W, Contreras R, Callewaert N. Glycome mapping on DNA sequencing equipment. *Nature Protocols*, 2006, 1(1): 397–405.
- [13] Paulson JC, Colley KJ. Glycosyltransferases. *Biological Chemistry*, 1989, 264: 17615.
- [14] Dorr R T. Clinical properties of yeast-derived versus *Escherichia coil*-derived granulocyte-macrophage colony-stimulating factor. *Clin Ther*, 1993, 15(1): 19–29.
- [15] Liu B, Gong X, Wu J, et al. Disruption of the *OCHI* and *MNN1* genes decrease *N*-glycosylation on glycoprotein expressed in *Kluyxeromyces lactis*. *J Biotechnol*, 2009, 143(2): 95–102.
- [16] Goodrick JC, Xu M, Finnegan R, et al. High-level expression and stabilization of recombinant human chitinase produced in a continuous constitutive *Pichia pastoris* expression system. *Biotechnol Bioeng*, 2001, 74(6): 492–497.
- [17] Liu B, Song M, Wu J, et al. A Novel method for the Oligosaccharides from glycoproteins using DSA-FACE. *Lett Biotechnol*, 2008, 19(6): 885–888.
刘波, 宋森, 吴军, 等. 一种利用 DSA-FACE 分析寡糖链的方法. *生物技术通讯*, 2008, 19(6): 885–888.



本 期 广 告 索 引

企 业	版 位	企 业	版 位
GE Healthcare 公司	封 底	生物谷网站	内 页
宝生物工程 (大连) 有限公司	封 二	艾本德 (上海) 国际贸易有限公司	内 页
北京鼎国昌盛生物技术有限责任公司	扉 一	镇江东方生物工程公司	内 页
安琪酵母股份有限公司	内 页		