

PaP3 噬菌体 57kD 蛋白编码基因的确定

朱军民 金晓琳 饶贤才 申晓东 胡福泉*

(第三军医大学基础医学部微生物学教研室 重庆 400038)

摘 要 :分离鉴定了 PaP3 噬菌体 57 kD 蛋白的编码基因并对其功能进行了初步探讨。用 PEG 沉淀结合 CsCl 梯度密度离心分离、纯化噬菌体颗粒,通过 SDS-PAGE 分析该噬菌体的衣壳蛋白,转印 PVDF 膜后,对 57 kD 蛋白用 Edman 降解法进行 N-端氨基酸测序,进而从 PaP3 全基因组的 256 个 ORFs 中确认该蛋白质的编码基因及其对应的氨基酸序列。结果显示噬菌体 PaP3 有 9 种结构蛋白分子,其中 57 kD 蛋白是由 ORF₃₄₇₉₃ 编码的。57kD 蛋白编码基因全长 1542bp, G + C 百分含量为 49.16%, 编码 514 个氨基酸。该蛋白分子量为 57.4kD, 等电点为 5.879。实验表明该蛋白是一种结构性蛋白,很可能是噬菌体衣壳蛋白中的一种壳微粒。

关键词 噬菌体 PaP3 结构性蛋白 编码基因

中图分类号 :Q755 文献标识码 :A 文章编号 :0001-6209(2004)01-0050-04

噬菌体(Bacteriophage)是寄生于细菌等微生物体内的一类病毒,与宿主菌关系密切。宿主菌可通过噬菌体介导的基因转座而导致耐药性基因、致病性基因在细菌间横向转移。展开对噬菌体基因组学、功能基因组学以及噬菌体与其宿主菌相互作用的研究,对于探讨生物多样性,深入了解生命起源、遗传变异、生化代谢等许多重要生命现象都有着极为重要的意义,因此噬菌体的研究一直是一个热点领域。本研究室自行分离到一株铜绿假单胞菌噬菌体,命名为 PaP3,其全基因组测序已完成。噬菌体 PaP3 全基因组由 45503 bp 组成,经 ORF finder 搜索表明 PaP3 基因组中具有 256 个(大于 100 bp)ORFs,基因预测分析提示有 71 个编码基因(Coding sequences, CDS)^[1,2]。在基因组测序完成后,我们展开了对该噬菌体基因功能的研究,以便更全面、深入地了解该噬菌体。本文介绍了对 PaP3 噬菌体 57 kD 蛋白编码基因的确定和功能的初步探讨。

1 材料和方法

1.1 铜绿假单胞菌噬菌体 PaP3

铜绿假单胞菌噬菌体 PaP3 为本实验室自行分离鉴定株。我们从临床标本中分离到其宿主菌铜绿假单胞菌,然后在医院下水道中分离到该噬菌体,详情参见文献[3]。其全基因组测序和初步注释已完

成^[1,2], GenBank 登录号为 :NC_004466, AY078382。

1.2 噬菌体 PaP3 颗粒的纯化

参照文献[4]的噬菌体纯化方法,配备 CsCl 3 层梯度液,密度分别为 1.45、1.50、1.70g/mL。将培养的 PaP3 和宿主菌混合悬液经 DNase I 及 RNase A 消化后,NaCl 沉淀细菌。离心取上清液,用 PEG 8000 冰浴沉淀噬菌体颗粒。将沉淀的噬菌体颗粒用 2mL 的 TM 液溶解稀释后,加在 3 层 CsCl 梯度液上,22000r/min 高速离心 2h,收集其中相应的灰色区带即为铜绿假单胞菌噬菌体 PaP3 颗粒纯化物。

1.3 PaP3 噬菌体 SDS-PAGE 分析

参照《分子克隆实验指南》第二版的方法,配备 12%分离胶和 5%积层胶,将噬菌体 PaP3 纯化颗粒和中分子量标准蛋白质,加入等量的蛋白加样缓冲液混合,煮沸变性处理后上样,恒压 80V 电泳约 4h 待溴酚蓝指示剂到达凝胶下缘时,停止电泳。切下含中分子量标准蛋白质的凝胶,用考马斯亮蓝 R-250 染色观察。其余各泳道的蛋白凝胶切去积层胶部分进行电转印。

1.4 噬菌体 PaP3 蛋白转印 PVDF 膜

参照文献[5]的方法进行。将余下的 PaP3 蛋白凝胶切去积层胶部分,与 PVDF 膜一起,在电转印缓冲液中平衡 30min。然后在夹板中制备三明治夹层,由正极到负极依次为:3 层浸湿的 Whatman 滤

基金项目 国家自然科学基金(30070037)

* 通讯作者。Tel 86-23-68752240, E-mail hufuquan@yahoo.com

作者简介 朱军民(1975-),男,山东省淄博市人,助教,硕士,主要从事噬菌体基因组学和生物信息学方面的研究。E-mail :zhujunmincn@yahoo.com.cn

收稿日期 2003-04-10,修回日期 2003-07-15

纸、PVDF 膜、稍小于 PVDF 膜的凝胶块、稍小于凝胶块的 3 层 Whatman 滤纸。整个操作在电转印缓冲液中进行 ,注意各层间不能有气泡。将夹板置于电转印槽中 ,置 4 ℃ 冰箱中恒压 30 V 电转印过夜 ,将 PVDF 膜取出用考马斯亮蓝染色 ,脱色 ,直至出现明显的蛋白条带而背景色脱净为止。

1.5 噬菌体 PaP3 蛋白 N-端测序

将转至 PVDF 膜上 57 kD 左右的蛋白样品委托北京大学生命科学院进行蛋白 N-端测序 ,根据 Ed-man 降解法 ,完成 N 端的 5 个氨基酸的测序。

2 结果

2.1 噬菌体 PaP3 蛋白的 SDS-PAGE 分析

将用 PEG 8000 沉淀和 CsCl 梯度密度离心法提取纯化的噬菌体 PaP3 颗粒 ,经 SDS-PAGE 分析 ,结果如图 1 所示。可见在中分子量标准蛋白范围内 ,凝胶上共可以出现 9 条蛋白条带。分子量依次可能为 114 kD、90 kD、57 kD、53 kD、35 kD、33 kD、29 kD、27 kD、25 kD(其中的 57 kD 和 53 kD 相距非常接近 ,而 53 kD 蛋白量较少 ,有些模糊)。结果表明噬菌体 PaP3 中至少有 9 个 ORFs 可表达出蛋白质。

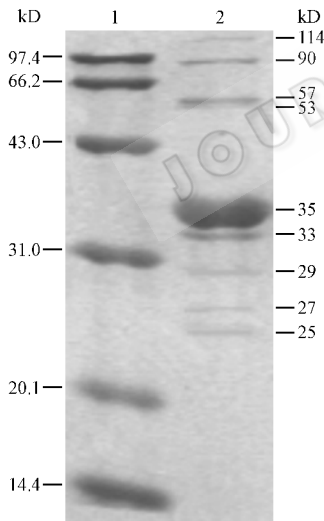


图 1 PaP3 噬菌体蛋白 SDS-PAGE 图谱

Fig.1 SDS-PAGE of bacteriophage PaP3

1. Marker ; 2. Protein of bacteriophage PaP3.

2.2 噬菌体 PaP3 蛋白 N-端测序

在先前的研究中 ,我们只能从噬菌体 PaP3 颗粒中分离出 4 条蛋白带。其中 35 kD 蛋白带的含量最高 ,当时即已完成该蛋白的氨基酸测序鉴定。在本研究中 ,我们通过对 SDS-PAGE 条件的优化 ,分离到 9 条蛋白带。转印 PVDF 膜后 ,由于电转印的效率在 PVDF 膜上只出现 3 条明显的蛋白带 :90 kD、57 kD、

35 kD。达到氨基酸可测序量的只有 57 kD 和 35 kD 蛋白带 ,故将 57 kD 蛋白条带送北京大学生命科学院进行蛋白 N-端测序。

本次蛋白质 N-端测序前 5 个氨基酸结果依次为 :A(丙氨酸) I(亮氨酸) E(谷氨酸) R(精氨酸) Q(谷氨酰胺) ,加上第一位的 M(蛋氨酸) ,故其 N-端排列顺序为 :MALERQ。

2.3 噬菌体 PaP3 57kD 蛋白编码基因的确认

我们确认 57kD 蛋白编码基因的依据有两个 :一是 N-末端氨基酸的序列 ,二是分子量的大小。首先 ,从 256 个 ORFs 中 ,将编码接近 57kD 蛋白产物的 ORF 全部列出来 ,并同时将它们 N 末端氨基酸序列列出来 ,与 57kD 蛋白的实际测序结果进行比对 ,以确认哪一个 ORF 是 57kD 蛋白的编码基因。

将噬菌体 PaP3 编码近 57 kD 蛋白的 ORFs 基因产物的 N-末端氨基酸序列进行比较(表 1)。所测序列 MALERQ 具有唯一性 ,对应的是 ORF₃₄₇₉₃ 的 N-端的前 6 个氨基酸。对 ORF₃₄₇₉₃ 进行序列分析 ,其理论计算分子量为 57.4kD ,与我们用中分子量标准蛋白所测的 57kD 基本吻合。因此 ,根据 ORF₃₄₇₉₃ 的 N-端氨基酸序列 ,计算理论分子量大小 ,可以肯定 57kD 的蛋白条带所对应的 ORF 即为 ORF₃₄₇₉₃。且在整个噬菌体基因序列中 ,ALERQ 顺序也具有唯一性 ,只出现在 ORF₃₄₇₉₃ 的 N-末端。

表 1 噬菌体 PaP3 所有近 57kD 推定蛋白的 N-端氨基酸序列比较

Table 1 N-terminal amino acid sequence comparsion of all 57 kD-proteins coded by PaP3 genome

ORF	N-terminal sequence	Theory MW/kD
ORF ₂₁₆₉₇	MKMENY	37.1
ORF ₃₈₈₄₆	MDLENQ	37.6
ORF ₅₉₂₈	MARLKY	38.1
ORF ₉₂₂₅	MSALYS	44.5
ORF ₃₀₈₂₂	MASMAY	53.7
Terminase gene	MDTQER	54.6
ORF ₁₀₄₃₄	MCGLVG	56.8
ORF ₃₄₇₉₃	MALERQ	57.4
DNA polymerase2	VDWRKS	61.2
Primase/helicase gene	MKKIIG	64.2
ORF ₄₁₂₀₈	MAKRRK	80.9

2.4 PaP3 噬菌体 ORF₃₄₇₉₃ 基因序列分析

ORF₃₄₇₉₃ 位于噬菌体 PaP3 基因组的负链上 ,从
© 中国科学院微生物研究所期刊联合编辑部 http://journals.im.ac.cn

34793 至 32323 号碱基,全长 1542 bp, A = 24.71%, G = 24.19%, T = 26.13%, C = 24.97%, A + T = 50.84%, G + C = 49.16%。该基因编码产物为 514 个氨基酸,分子量为 57.4kD,等电点(PI)为 5.879。

3 讨论

噬菌体分别由 Frederick Twort 于 1915 年和 Felix d'Herelle 于 1917 年发现。它是一种比细菌更小、更简单、在进化阶梯上更为原始的非细胞形态的生命现象。最简单的噬菌体只有 10 多个基因,最复杂的噬菌体有数百个基因。噬菌体生活在一个极为特殊的环境之中,首先,它们生活在细菌体内,而细菌又生活在人体内,这就构成了双重生境嵌套的特殊环境。因此,噬菌体是研究噬菌体、细菌及人体三者相互作用关系的一种极好的模式生物。

本课题组于 1999 年开始了铜绿假单胞菌噬菌体的研究,自行分离到一株铜绿假单胞菌噬菌体,命名为 PaP3。已经弄清噬菌体 PaP3 全基因组由 45503bp 组成,经 ORF finder 搜索表明 PaP3 基因组中具有 256 个大于 100 bp 的 ORFs,基因预测分析提示有 71 个 ORFs 可能是编码基因。经 BLAST ORF 序列同源性比对,在蛋白质氨基酸水平上,发现有 6 个 ORFs 编码产物有同源性序列存在,而其余均是功能尚不清楚的新基因。应该注意的是,上述基因的预测,都是现行常用生物信息学分析软件的分析结果。但这些计算机分析结果最终还是需要实验证据来证实,才能够最终敲定预测结果的真实性。

先前的研究我们只能从噬菌体 PaP3 颗粒中分离出 4 条蛋白带,对于一个 4 万多碱基的噬菌体基因组,其结构蛋白不应该仅有 4 个。本研究采用了网上介绍的改良方法,对整个实验条件进行了优化。目前,已经知道的与 PaP3 基因组大小相差不多的几个噬菌体,其噬菌体颗粒在 SDS-PAGE 凝胶上,通常可出现几个到十几个不同的条带^[6,7],而对于与其相似的铜假单胞菌噬菌体如约 36 kb 的铜假单胞菌噬菌体 phi CTX 在 SDS-PAGE 凝胶上共出现 10 条带^[8],56 kb 的噬菌体 D3 在 SDS-PAGE 凝胶电泳后显示有 6 条蛋白条带^[9]。而我们的噬菌体 PaP3 共有 9 条带,与它们都是相类似的,可能已经接近或达到了噬菌体 PaP3 在 SDS-PAGE 凝胶上的结构蛋白数。

分离纯化了蛋白条带,我们就有条件确定该蛋白的编码基因和编码产物的氨基酸序列。在本研究中,采用 N-末端氨基酸测序结合分子量大小这两个

指标来确认 57kD 蛋白的编码基因。测序结果表明 57kD 蛋白的 N-末端氨基酸序列为“ALERQ”。我们检查了噬菌体 PaP3 的全部 256 个 ORFs,发现只有处在负链上的 ORF_{34793~33252}(根据起始位点命名为 ORF₃₄₇₉₃)的编码基因其 N 端的 5 个氨基酸序列为“ALERQ”。随后,我们对 ORF₃₄₇₉₃进行了序列分析,根据计算,其分子量为 57.4 kD。由表 1 可见,与其分子量最为接近的有两个 ORFs,分别是 ORF₁₀₄₃₄和 ORF₃₄₇₉₃,但 ORF₁₀₄₃₄的序列为“MCGLVG”,与我们所测的结果完全不同,而 ORF₃₄₇₉₃的序列为“MALERQ”,与我们所测的结果完全一致。因此可以确认 ORF₃₄₇₉₃就是 57 kD 蛋白的编码基因。

噬菌体基因组所编码的蛋白中,结构性蛋白通常表达较稳定,且表达拷贝数较高。而调节性蛋白的转录和翻译水平相对很低,甚至只在噬菌体生活周期的某些阶段一过性地表达。调节性蛋白即使表达也仅存在于宿主菌的胞浆之中,而不是噬菌体颗粒内^[10,11]。由于我们提取的是噬菌体的颗粒,而不是细菌胞浆内含物来完成本研究。故本研究中分离纯化噬菌体颗粒后检测到的上述 9 条蛋白带,应该都是噬菌体的结构性蛋白。因此,本研究的 57kD 蛋白也应该是噬菌体 PaP3 的结构蛋白之一,很可能是噬菌体衣壳蛋白中的一种壳微粒。至于该蛋白是位于头部、尾部或其它位置,有待进一步研究。

参 考 文 献

- [1] 张克斌,金晓琳,朱军民,等.铜绿假单胞菌噬菌体 PaP3 基因组测序.第三军医大学学报,2002,24(4):378-380.
- [2] 张克斌,金晓琳,朱军民,等.铜绿假单胞菌噬菌体 PaP3 基因组的初步注释.第三军医大学学报,2002,24(4):381-384.
- [3] 张克斌,陈志强,金晓琳,等.铜绿假单胞菌噬菌体分离鉴定及耐噬菌体突变频率测定.微生物学通报,2002,29(1):40-45.
- [4] 萨姆布鲁克 J,弗里奇 E F,曼尼阿蒂斯 T.分子克隆实验指南第 2 版.金冬雁,黎孟枫,译.北京:科学出版社,1996,136-137.
- [5] http://www.expasy.ch/ch2d/protocols/protocols_fm8.html
- [6] Tremblay D M, Monieau S. Complete genomic sequence of the lytic bacteriophage DT1 of *Streptococcus thermophilus*. *Virology*, 1999, 255(1):63-76.
- [7] Volkmar B J R, Stefan H, Horst N, et al. Taxonomic differentiation of bacteriophages of *Lactococcus lactis* by electron microscopy, DNA-DNA hybridization, and protein profiles. *J Gen Microb*, 1989, 135: 2551-2560.
- [8] Elsbagg H, Xiong G, Lutz F. Nucleotide sequence of attP and cos sites of phage CTX and expression of cytotoxin in *Pseudomonas aeruginosa*. © 中国科学院微生物研究所期刊联合编辑部 <http://journals.im.ac.cn>

inosa PA158. *Mol Gen Genet* ,1993 **237**(3) :421 – 428.

[9] Gilakjan Z A ,Kropinski A M. Cloning and analysis of the capsid morphogenesis genes of *Pseudomonas aeruginosa* bacteriophage D3 : another example of protein chain mail ? *J Bacteriol* ,1999 **181**(23) : 7221 – 7227.

[10] 司稚东 ,何晓青 ,汪家驯 ,等. 噬菌体学. 北京 :科学出版社 , 1996 63 – 88.

[11] Calender R. The Bacteriophages. New York :Plenum Press ,1988 ,117 – 167.

Identification of 57 kD-protein Gene of Bacteriophage PaP3

ZHU Jun-Min JIN Xiao-Lin RAO Xian-Cai SHEN Xiao-Dong HU Fu-Quan*
(Department of Microbiology ,Third Military Medical University ,Chongqing 400038 ,China)

Abstract : The coding gene of 57 kD protein of bacteriophage PaP3 was identified and primarily ascertained its function. bacteriophage PaP3 particles were isolated and purified by using. PEG-sedimentation combined with CsCl gradient centrifugation then the total proteins were isolated on SDS-PAGE to figure out how many proteins making up of the virion particle. The protein bands on SDS-PAGE were transfered onto PVDF membrane by electronic transfer method. Afterward , the 6 amino acid sequence of the N-terminal of 57 kD protein was determined by using Edman degradation method , and then clarified coding gene of 57 kD protein from 256 ORFs of genome PaP3 based on the sequence result. Results showed that constituent proteins of the phage were at least consisted of 9 proteins. The 6 amino acid sequence of the N-terminal of 57 kD protein is “ MALERQ ” , and based on this sequence , it may be confirmed that 57 kD protein is encoded by ORF₃₄₇₉₃ , which is composed of 1542 base pairs , with G + C content of 49.16% , and coding 514 amino acids , PI = 5.879. The theory MW of ORF₃₄₇₉₃ coding protein is 57.4 kD ,extremely closed to actually detection value. The 514 amino acid sequence of 57 kD protein has been clarified. The experiment showed this protein is one of capsomeres making up of the phage capsid proteins.

Key words : Bacteriophage PaP3 , Structural protein , Coding gene

Foundation item :Chinese National Natural Science Found(30070037)
* Corresponding author. Tel 86-23-68752240 ; E-mail :hufuquan@yahoo.com
Received date :04-10-2003

欢迎订阅《微生物学报》

《微生物学报》(双月刊)创刊于 1953 年 ,是我国微生物学领域唯一的综合性学报级期刊和国家自然科学核心期刊。主要报道普通微生物学 ,工业、农业、医学和兽医微生物学 ,免疫学以及与微生物学有关的生物工程等方面的研究成果和科研进展。

2004 年本刊全新改版 ,更换了彩色封面 ,由原来的小 16 开本改为标准大 16 开本(210 × 297) ,双月刊 ,每册 128 页 ,发表周期缩短 ,内容更加丰富详实。欢迎广大读者到邮局或直接与本刊编辑部联系购买 ,每册定价 28 元 ,全年 168 元 ,我们将按期免费邮寄。

另 ,本刊编辑部现存有少量过期期刊 ,如有需要者可直接与编辑部联系 ,每册定价 20 元 ,款到即免费寄上(请在汇款单上注明所购刊物的年代、卷、期和数量)。

邮购地址 :100080 北京海淀中关村 中国科学院微生物研究所内
《微生物学报》编辑部
Tel (010) 62630422 ; Fax (010) 62554303 ; E-mail :actamicro@sun.im.ac.cn