

人脑内一含有 ACP 样结构域新基因的发现*

邓艳春 药立波** 刘新平 聂晓燕 王吉村 张晓光 苏成芝

(第四军医大学生物化学与分子生物学教研室, 西安 710032)

摘要 为寻找脑内新基因, 以正常成人全脑 cDNA 为模板, 采用锚定 PCR 方法进行扩增, 将琼脂糖 DNA 电泳鉴定获得的一约 1 200 bp 大小的特异性条带回收, 并克隆入 T easy 载体, 用 310 Genetic Analyzer 进行自动测序, 所得序列进行生物信息学分析: BLAST 相似性分析结果证明所得序列为新序列, 读框分析表明, 该序列中存在一完整编码区, 编码含 357 个氨基酸的蛋白质. ProDom 软件分析发现其含有酰基携带蛋白 (ACP) 样结构域. 随后, 经 3' RACE 法克隆到该基因的全长 cDNA, 其全长为 2 024 bp, 染色体定位在 14q11. 2, 含有 16 个外显子, 15 个内含子, 该基因已登录到 GenBank. 经设计编码区引物, 从 T easy 载体扩增出编码区后再克隆入 pGEX-4T1 表达载体, 经异丙基硫代-D-乳糖苷 (IPTG) 化学诱导表达. 其编码区克隆入 pGEX-4T1 表达载体后, 转入 JM109 宿主菌, 经 IPTG 诱导已得到表达. 点杂交及 RNA 印迹表明, 该基因在正常成人脑内广泛高表达.

关键词 基因, 脑, 乙酰基载体蛋白结构域, 生物信息学

学科分类号 Q73

我们在试图从脑内克隆含 PH 结构域新基因的过程中, 偶然得到一含 ACP 样结构域的新基因(专利号: 00113886. 3, Accession Number: AF159092). 现将其发现过程及在大肠杆菌内的表达报告如下:

PH 结构域是一种新发现的约由 100~120 个氨基酸残基组成的功能性结构域, 广泛分布于单细胞生物、无脊椎、脊椎动物和人的细胞骨架蛋白和信号分子等 100 多种蛋白质分子中, 能与脂类、G 蛋白的 $\beta\gamma$ 亚单位、PKC 等配体相结合, 推测其对蛋白质具有细胞和亚细胞水平的膜定位作用, 并参与调节宿主蛋白的活性, 从而介导信号转导过程中的蛋白质-脂类、蛋白质-蛋白质之间的相互作用^[1]. PH 结构域在生长因子受体结合蛋白 (Grb) 家族中位于蛋白的羧基端, 保守性很高. 该家族中的 Grb2 参与 EGF (epithelia growth factor) 的信号传导^[2], Grb10 参与 IGF (insulin like growth factor)^[3] 的信号传导. 我们知道, 脑内富含脂类和多种生长因子, 其信号传导途径尚不完全清楚, 因此推测, 脑内可能存在 Grb 家族的其他成员, 负责介导其他生长因子的信号传导. 故我们根据 Grb 家族 PH 结构域氨基端的氨基酸同源性, 设计了简并引物作为 PCR 的上游引物; 因 Grb 家族中的 PH 结构域靠近羧基端, 所以, 下游引物采用反转录时加到 cDNA 两端的接头, 作为锚定引物. 为了增加获得目的片段的可能性, 采用了低温 (45 °C) 长时间 (5 min) 退火的 PCR 参数, 将 PCR 产物行琼

脂糖凝胶电泳, 结果却得到比预计大的 (预计 Grb 家族 PH 结构域至 3' 尾约 900 bp) 的一条带. 将其克隆测序, 发现该片段两端都是接头引物, 计算机分析表明其为一个含有完整读框的新基因, 并含有一 ACP 样结构域.

1 材料和方法

1.1 材料

正常成人全脑 cDNA (clontech), PCR 引物 5'-aaywsnctngtngcnatggaytlytengg-3', 5'-ccatcctaa-tacgactcactatagggc-3' (clontech). PCR 引物 5'-cgggatccatggcggagctgcaggaggtg-3', 5'-cggaaatcaac-aagggccatcaacaggagac-3' 及测序引物 5'-agggtttt-cccagtcacgac-3', 5'-acacaggaacagctatgac-3', 5'-tacgctctgtcactttcactgtc-3' (北京赛百盛生物工程公司合成). Taq 酶及限制性内切酶购于华美生物技术公司. 胶回收试剂盒及质粒提取试剂盒购于上海华顺生物技术公司. T easy 载体 (promega). pGEX4T1, 菌种 JM109 本室保存.

1.2 PCR 扩增

1.2.1 片段的扩增: 以成人全脑 cDNA 为模板, 在 50 μ l 反应体积中, 模板 DNA 200 ng, dNTP 各 200 μ mol/L, Taq DNA 聚合酶 3U, 引物分别为

* 国家自然科学基金面上项目和杰出青年基金资助 (39825113).

** 通讯联系人.

Tel: 029-3374513, E-mail: bioyao@fmmu.edu.cn

收稿日期: 1999-12-16, 接受日期: 2000-02-27

20 pmol. 反应参数为 94 °C 30 s, 45 °C 退火 5 min 72 °C 延伸 90 s, 5 个循环. 94 °C 30 s, 52 °C 退火 2 min 72 °C 延伸 90 s, 30 个循环.

1.2.2 编码区的扩增: 以阳性质粒为模板, 在 50 μl 反应体积中, 模板 DNA 100 ng, dNTP 各 200 μmol/L, TaqDNA 聚合酶 3U, 引物分别为 10 pmol. 反应参数为 94 °C 30 s, 68 °C 退火 4 min, 25 个循环.

1.2.3 全长 cDNA 的扩增: 重新设计引物, 用 3' RACE 法, 以成人全脑 cDNA 为模板, 采用 touch down PCR 法克隆出其 3' 端 (初得片段已包括 5' 端). 反应参数为 94 °C 30 s, 72 °C 4 min 5 个循环; 94 °C 30 s, 70 °C 4 min 5 个循环; 94 °C 30 s, 68 °C 退火 4 min, 25 个循环.

1.3 基因克隆及测序

将经琼脂糖凝胶电泳证实的 PCR 扩增产物经胶回收试剂盒回收后, 克隆入 T easy 载体. 用 310

Genetic Analyzer 自动测序. 分别用正、反向通用引物和基因特异性引物从片段的两端和中间测序. 将三次测序结果用 DNAClub 软件进行拼接.

1.4 生物信息学分析

1.4.1 BLAST 相似性分析及染色体定位: 将所得序列去掉载体和引物部分, 用 BLAST 软件经因特网与 GenBank 的所有已知序列进行比较. 染色体定位在 14q11.2 含有 16 个外显子和 15 个内含子.

1.4.2 读框分析: 用 PcGene 软件及 ORF Finder 软件进行基因的读框分析.

1.4.3 功能结构域分析: 用 ProDom 软件进行基因编码区功能结构域的分析.

1.5 表达质粒的构建: 将用 BamH I - EcoR I 酶切 PCR 扩增的编码片段插入到 BamH I - EcoR I 酶切的 pGEX-4T 1 表达载体内, 构成原核表达质粒 pGEX-syld (图 1).

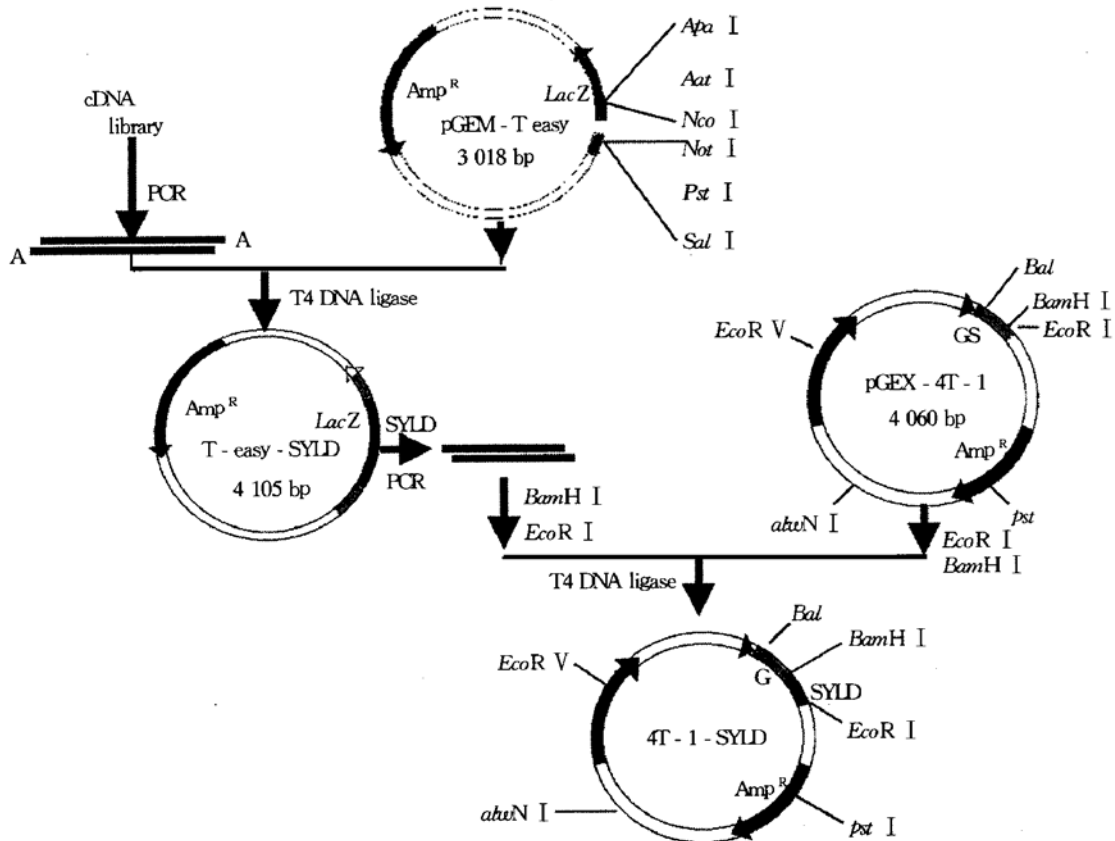


Fig. 1 Construction of GST fusion expression vector

1.6 基因在大肠杆菌中的表达

1.6.1 阳性表达克隆的鉴定: 含重组质粒 pGEX-syld 的 JM109 单菌落接种于含氨苄青霉素 (200 mg) 的 LB 培养基中, 于 37 °C 摇床中培养

2 h, 培养液 A₆₀₀ 达 0.5 ~ 0.6 时, 加入诱导剂 IPTG 至 1.0 mmol/L, 诱导表达 2 h, 收菌, 制备成蛋白质样品后经 12% SDS-聚丙烯酰胺凝胶电泳分析.

1.6.2 表达时间分析: 将表达 SYLD 蛋白的阳性克隆菌液 1:100 转接入 5 ml 含氨苄的 LB 培养基中, 于 37℃ 摇床中过夜培养. 次日按 1:100 比例转接入 10 ml 含氨苄的 LB 培养基中, 于 37℃ 摇床中培养 2 h. 培养液 A_{600} 达 0.5~0.6 时, 加入诱导剂 IPTG 至 1.0 mmol/L, 分别于诱导表达的 1、2、3、4、5 h, 收菌, 制备成蛋白质样品后经 12% SDS-聚丙烯酰胺凝胶电泳分析.

1.6.3 表达形式分析: 取表达的菌体沉淀 0.1 mg 于微量离心管中, 向其中加入 0.5 ml PBS, 震荡混匀后 50 Hz 超声裂菌, 10s/次, 间隔 3 s, 共 2 min.

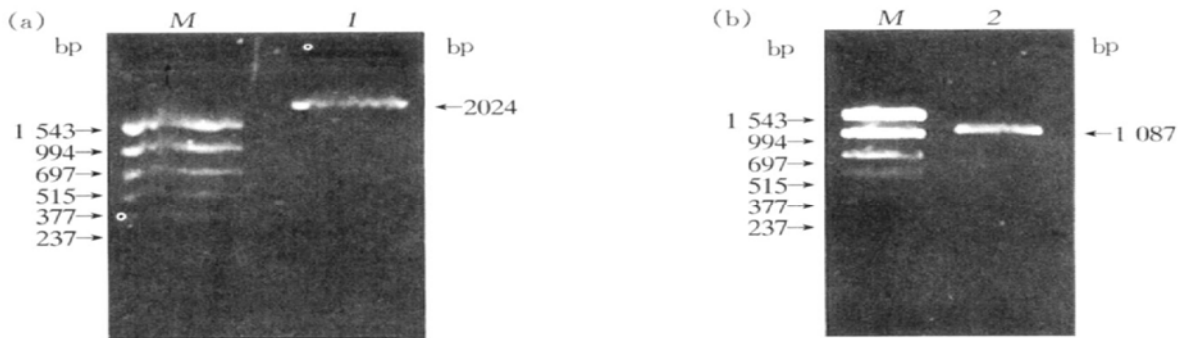


Fig. 2 PCR result of full length cDNA (a) and open reading frame (b)

M: PCR marker; I: full length cDNA; 2: ORF fragment.

2.2 测序结果

测序结果见图 3.

2.3 生物信息学分析结果

2.3.1 BLAST 相似性分析结果: 经 BLAST 软件分析表明无与该序列完全一样的同源基因, 所得序列为新序列.

2.3.2 读框分析结果: 经 PcGene 和 ORF Finder 软件分析, 表明该基因内存在一完整读框. 其编码的氨基酸序列见图 3.

2.3.3 功能结构域分析结果: 用 ProDom 软件分析表明该基因编码的蛋白质分子内含有一酰基携带蛋白 (ACP) 样结构域, 与已知的 ACP 结构域有 53% 的同源性, 和 63% 的同源相似性.

2.4 在大肠杆菌中的表达结果

经 SDS-聚丙烯酰胺凝胶电泳分析表明, 于 66 ku 处见一新生蛋白质条带, 符合 GST 加目的蛋白的分子质量大小 (GST 26 ku, 目的蛋白约 40 ku). 表达时间分析表明 IPTG 诱导培养的 2 h 表达量最高, 表达形式为不溶性的, 存在于包含体中 (图 4a 和 4b).

12 000 r/min 离心 10 min, 取上清和沉淀分别制备蛋白质样品, 进行 12% SDS-聚丙烯酰胺凝胶电泳分析.

2 结 果

2.1 PCR 结果

用 3' 锚定 PCR 法从脑内扩增出两条带, 大小分别约为 1 200 bp 和 700 bp (图 2a). 3' RACE 扩增出 3' 端, 用基因特异引物扩增出该基因的全长 cDNA. 用编码区引物从阳性克隆的质粒内扩增出编码区, 大小约 1 100 bp (图 2b).

3 讨 论

我们试图从脑内克隆含 PH 结构域的 Grb 家族新成员, 因此上游引物是 Grb 家族 PH 结构域的兼并引物, 下游引物是反转录时加到 cDNA 两端的锚定引物, 采用长时间低温退火 PCR 反应参数, 结果却得到特异性条带. 测序结果表明片段两端均为锚定引物. 用非特异引物和非特异条件扩增出特异片段, 表明该基因可能在脑内高峰期表达. 此外该基因编码的蛋白质分子内含有一 ACP 样结构域, 为对其功能研究提供了线索. 在真核细胞生物中, ACP 结构域存在与脂肪酸合成酶的多酶复合体中, 作为一种摆动的手臂来捕获活化的脂肪酸和氨基酸^[4]. 在人体内, ACP 结构域是 N-Myc downstream regulator (NDR) 家族的特有结构域^[5], 因此, 表明该基因可能是人类 NDR 家族的新成员, 推测其可能与原癌基因 N-Myc 的表达调控, 因此与肿瘤的发生及脑的发育有关. 该蛋白质在大肠杆菌中的表达为多克隆抗体的制备等下一步工作打下了基础. 此外, 我们以该基因的 cDNA 做探针, 对

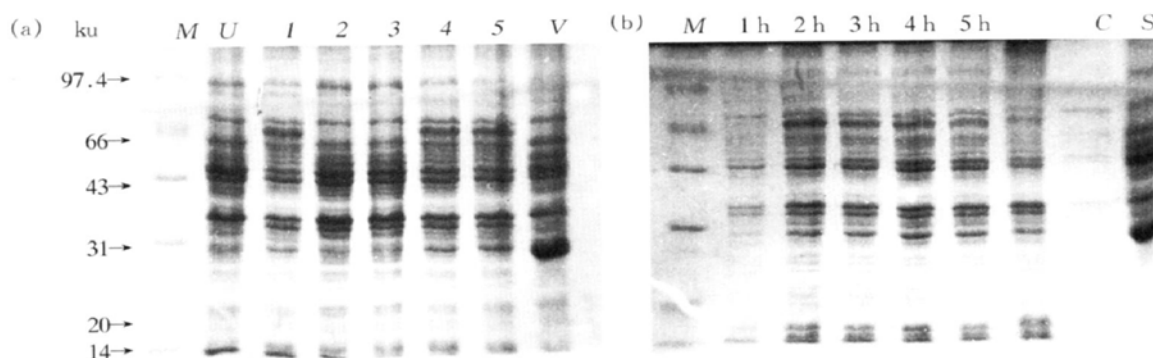


Fig. 4 Result of SDS-PAGE of expression of target gene in *E. coli*

M: protein molecular mass marker; *I*~*5*: the No. of clons; *U*: uninduction; *V*: vector; 1~5 h: time of IPTG induction; *C*: clusion bodies; *S*: supernatant. (a) showed a protein band at molecular mass of 66 ku that equals to the GST (26 ku) plus the aim protein (40 ku) in clon. *I*, *4* and *5*. (b) showed the results of analysis of expression time and style after IPTG induction. The aim protein was expressed in clusion bodies and the higher level of expression can be reached at 2 hour after IPTG induction.

参 考 文 献

- 1 邓艳春, 药立波, 苏成芝. PH 结构域研究进展. 生命科学, 2000, **12** (3): 117~121
- 2 Deng Y C, Yao L B, Su C Z. Life Science, **12** (3): 117~121
- 3 Wong L, Deb T B, Thompson S A, *et al.* A differential requirement for the COOH-terminal region of the epidermal growth factor (EGF) receptor in amphiregulin and EGF mitogenic signaling. J Biol Chem, 1999, **274** (13): 8900~8909
- 4 Dong L Q, Du H, Porter S G, *et al.* Cloning, chromosome localization, expression, and characterization of an Src homology 2 and pleckstrin homology domain-containing insulin receptor binding protein hGrb10γ. J Biol Chem, 1997, **272** (46): 29104~29112
- 5 Tropf S, Reville W P, Bibb M J, *et al.* Heterologously expressed acyl carrier protein domain of rat fatty acid synthase functions in *Escherichia coli* fatty acid synthase and streptomyces coelicolor polyketide synthase systems. Chem Biol, 1998, **5** (3): 135~146
- 6 Altschul S F, Madden T L, Schaffer A A, *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res, 1997, **25**: 3389~3402

Exploring a New Gene Containing ACP Like Domain in Human Brain and Expression It in *E. coli**

DENG Yan-Chun, YAO Li-Bo**, LIU Xir-Ping, NIE Xiao-Yan,
WANG Ji-Cun, ZHANG Xiao-Guang, SU Cheng-Zhi

(Department of Biochemistry and Molecular Biology, The Fourth Military Medical University, Xi'an 710032, China)

Abstract To look for new genes from human brain, get a fragment was obtained using adaptor primer and 3' anchor polymerase chain reaction (PCR) with the human adult whole brain cDNA as template. The fragment was cloned into T easy vector and automatically sequenced with 310 Genetic Analyzer. Later the whole length cDNA of this novel gene was got with the method of 3' rapid amplification of cDNA end (RACE). The whole length of cDNA of this novel gene is 2 024 bp. Chromosome location is at 14q11.2 including 16 exons and 15 introns. After scanning the sequence against GenBank it is proved that the sequence is a new one. ORF analysis showed that there is a complete coding region in it. it can interpretate a protein containing 357 amino acid residues. ProDom analysis result showed that there is an acyl carrier protein (ACP) like domain in it. The gene was banked into GenBank. Then, a pare of primers were designed and were used to amplify the coding region and cloned into pGEX-4T1 expressing vector to express it in *E. coli*. The Dot blotting and Northern blot showed that this novel gene is highly expressed in the normal adult human brain.

Key words gene, brain, acyl carrier protein domain, bioinformatics

* This work was supported by a grant from National Natural Sciences Foundation of China (39825113).

** Corresponding author. Tel: 86-29-3374513, E-mail: bioyao@fmmu.edu.cn

Received: December 16, 1999 Accepted: February 27, 2000