

大鼠脑中血浆硒蛋白 P cDNA 克隆

杨晓光* 田 园 K E Hill¹ R F Burk¹

(中国预防医科院营养与食品卫生研究所, 北京 100050)

提 要 方法: 用大鼠肝中血浆硒蛋白 P cDNA 为探针, 对 λZPA II 大鼠脑 cDNA 文库进行筛选, 对阳性克隆进行了核酸序列测定。结果: 脑中 cDNA 中蛋白编码区与肝中 cDNA 的编码区完全一致, 但在 5'utr 和 3'utr 存在差异。结论: 脑中血浆硒蛋白 P cDNA 的发现说明脑本身可合成血浆硒蛋白 P, 这一结果不支持血浆硒蛋白 P 的硒转运蛋白功能说。

关键词: 硒蛋白 脑 cDNA

中图分类号: R 151.2

硒蛋白是特指蛋白中的硒是以硒半胱氨酸(selenocysteine)形式存在于构成蛋白的多肽链的一级结构中的蛋白, 因此硒蛋白的氨基酸序列分析中应用硒半胱氨酸, 和其蛋白对应的 mRNA 中在蛋白编码区应用 UGA (在 cDNA 中为 TGA)。血浆硒蛋白 P (Selenoprotein P) 是在血浆中发现的分子量 (SDS-PAGE) 为 57kDa 的单链糖蛋白^[1], 在肝中得到其 cDNA 克隆, 其核酸序列分析表明, 在蛋白编码区有 10 个代表硒半胱氨酸的编码 TGA^[2], 但实际蛋白氨基酸序列分析为 7.5 个硒半胱氨酸^[3]。我们对脑中血浆硒蛋白 P 的研究表明, 脑中存在有血浆硒蛋白 P, 其在细胞中的可能位置是细胞膜上^[4]。为确定脑中血浆硒蛋白 P 的来源, 我们进行了如下研究。

1 材 料 和 方 法

1.1 探针制备

将大鼠肝中血浆硒蛋白 P 16C1 cDNA

用 EcoRI 和 XbaI 酶切后, 经琼脂糖电泳分离回收含全部开放阅读框架 (Open Reading Frame ORF) 部分的 2.1kb 片断, 用缺口平移法标记 [α -³²P] dCTP (美国 Du Pont-New England Nuclear Productus)。

1.2 脑 cDNA 文库的筛选

将 λZPA II 大鼠脑 cDNA 文库 (美国 Stratagene 公司) 噬菌体储备液稀释后接种于 E coli BB4 细胞 (美国 Stratagene 公司), 接种浓度为 5×10^4 噬菌体/每平皿。37 培育过夜后将噬菌体 DNA 转移到醋酸纤维膜上。将醋酸纤维膜按顺序浸入下列液体中。(1) 0.1mol/L NaOH, 1.5mol/L NaCl; (2) 0.5mol/L Tris-HCl pH8.0, 1.5mol/L NaCl; (3) 0.2mol/L Tris-HCl pH7.0, $2 \times$ SSC。于 80 真空器内干燥 2 小时。在 42 进行预杂交 2 小时, 预杂交液为 $5 \times$ SSPE, 20% 甲酰胺, 0.1% SDS, $3 \times$ Denhardt's 液。再加入 [³²P] dCTP 标记的 16C1 中 2.1kb 片断于 42 杂

* 男, 1955 年出生, 副研究员, 博士, 课题经费由国家自然科学基金和非教育系统留学回国人员科技活动 D 类资助
¹ 美国 Vanderbilt 大学

交过夜。杂交后在60 °C条件下用2 × SSPE 和 1 × SSPE 各洗膜一次, 每次1小时。压 XAR-5 胶片于- 70 °C 曝光, 对照 X 胶片挑阳性克隆并洗脱于 SM 缓冲液(100mmol/L NaCl, 100mmol/L MgSO₄, 50mmol/L Tris-HCl pH 7. 5)再次筛选, 再到得到完全分离的单个阳性克隆。

1. 3 测序

按 Stratagene 说明, 用带有408 helper phage 的 E. coli XL-1 Blue 细胞进行共培养, 切除阳性 p B luscript phagem id, 用 recircularized 噬菌体感染新的 E. coli XL-1 Blue 细胞。用 Sanger 双氧末端终止法对复制形式成噬菌体 (RF-cDNA) 中插入部分进行双链 DNA 测序。T7 Sequencing kit 购自美国 Pharmacia L KB, 测序策略如图1所示。

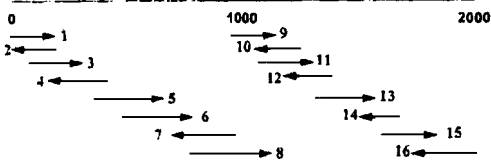


Fig. 1 Sequencing strategy used for selenoprotein P cDNA of rat brain

The arrows represent the segment of cDNA sequenced and the direction of sequencing. T3 primer and T7 primer were used for numbers 1 and 16, respectively. All other sections were sequenced using oligomers specifically synthesized as primers

2 结果与讨论

对 AZPA II 大鼠脑 cDNA 文库筛选, 共得到5个阳性克隆, 每个克隆扩增后用限制性内切酶切下插入片断后进行琼脂糖凝胶电泳, 其中有3个阳性克隆的插入部分小于1kb, 未对其进行测序。对另2个掺入片断为1. 8kb 的 cDNA (11A 1和2A 2) 进行了核酸序列测定, 结果见图2。

5'端非翻译区(5' untranslated region 5'utr)部分完整的11A 1 cDNA 中3'端非翻译区(3' untranslated region 3'utr)部分不完整, 而5'utr部分和ORF前端不完全的2A 2的3'utr部分完整(有polyA 尾)。结合二个 cDNA, 可知脑中全长血浆硒蛋白 P 的 cDNA 序列。脑中血浆硒蛋白 P cDNA 5'utr 有78个碱基比, 肝中血浆硒蛋白 P cDNA 5'utr 多44个碱基。除靠近ORF 端的15碱基与肝中的相同, 其余碱基则不同, 这种差别的意义尚不清楚。脑中 cDNA 中ORF(蛋白编码区)与肝中 cDNA 的ORF 完全一致, 亦是含有10个 TGA 编码。这对肝中血浆硒蛋白 P cDNA 的序列是进一步证实。脑中血浆硒蛋白 P cDNA 3'utr 短于肝中 cDNA 的3'utr, 但仍保留有二个茎突(Stem-loop)结构。而 mRNA 3'utr 中茎突结构是保证硒蛋白在肽链合成时其 mRNA ORF 内的 U GA (cDNA 中为 TGA) 作为硒半胱氨酸的编码而不是终止密码的必需条件^[5]。脑中血浆硒蛋白 P 3'utr 与肝中 cDNA 3'utr 对应部分有二处不同, 一是在1441碱基处, 肝中为 C, 而脑中为 T, 另一处是对应肝 cDNA 1574位后, 脑中 cDNA 多2个 T。因对2A 2和11A 1二个 cDNA 是分别测序, 结果一致。说明这种差别很可能是其 mRNA 本身的差别, 这种差异的意义尚不清楚。脑中发现血浆硒蛋白 P cDNA 克隆及其蛋白编码区与肝中血浆硒蛋白 P 完全一致, 充分证明脑可自身合成血浆硒蛋白 P。因此也有充分理由说明在脑中发现的血浆硒蛋白 P 是脑自身合成而非经血转运。关于血浆硒蛋白 P 的功能有人曾提出是硒的体内转运蛋白, 把肝中硒运送到其它组织^[6], 此脑中发现血浆硒蛋白 P cDNA 克隆的结果不支持此硒蛋白的转运功能说。

2A 2 _____

16C1 GAA GCC TTC TGT AAA AAC GTG TCC TCG GCT ACT GCA AGT AAA ACC ACA GAG CCC

11A 1 _____

2A 2 _____

16C1 TCA GA ⁶⁴⁰ GAG CAT AAC CAC CAC AAG CAC CAT GAC AAA CAT GGG CAT GAG CAT CTT

11A 1 _____

2A 2 _____

16C1 GGG AGC AGT AAG CCT TCA GAG AAT CAG CAA ⁷²⁰ CCA GGG GCA TTA GAT GTT GAG ACA

11A 1 _____

2A 2 _____

16C1 AGT CTT CCT CCT TCA GGC TTG CAC CAC CAC CAC CAC CAT AAG CAC AAG GGC

11A 1 _____

2A 2 _____

16C1 CAG ⁸⁰⁰ CAC AGG CAG GGT CAC TTA GAG AGC TGA GAC ATG GGG GCA AGT GGA GGC TTG

11A 1 _____

2A 2 _____

16C1 CAA CTT TCA CTT GCC CAG AAG AAG CTC ⁸⁸⁰ TGA CGA AGG GGA TGC ATA AAC CAG CTC

11A 1 _____

2A 2 _____

16C1 CTG TGT AAG TTA TCT GAG GAG TCT GGG GCA GCT ACC AGT AGC TGC TGC TGC CAC

11A 1 _____

2A 2 _____

16C1 T ⁹⁶⁰ GC CGA CAC CTC ATA TTT GAG AAG TCA GGA TCT GCA ATC ACT TGA CAG TGT GCC

11A 1 _____

2A 2 _____

16C1 GAA AAC CTC CCA TCC TTG TGT AGC ¹⁰⁴⁰ TGA CAG GGG CTT TTC GCG GAG GAG AAA GTC

11A 1 _____

2A 2 _____

16C1 ATT GAA TCC TGT CAA TGT AGA TCA CCT CCA GCT GCC ¹¹²⁰ TGA CAC AGT CAG CAT GT A

11A 1 _____

2A 2 _____

16C1 AGC CCC ACA GAA GCC AGC CCC AAC TGA AGC TGA AAT AAT AAG ACC AAG AAG

11A 1 _____

2A 2 _____

16C1 TGA AAA TGA AAT TTG AAC TAA ATATTT A ¹²⁰⁰ AAA TAAA GCGTACTCTCCCAACTCCA TC

11A 1 _____

2A 2 _____

16C1 TAAA GACACAA TTTCA TTTCTA GAA TGTTTCCAA TCCA TTTAA TTAA TT A G ¹²⁸⁰IGAA GTAAAA G

11A 1 _____

2A 2 _____

16C1 TA GTTGAAA TTGGA TTTGTGCAAA CA TGGA GAAA TCTACCA CA TTGGCTTCTAAAA TTTA

11A 1 _____

2A 2 _____

16C1 AAA TTTTT A ¹³⁶⁰IGCCA CAAA CCA TTTCA TCCAAA TCA GA TTTGTACCGTGGGGCAACTGAAA

11A 1 _____

2A 2 _____

16C1 A GTGA TTGCGGCCA TTGGTTAA TA TGTC T ¹⁴⁴⁰ICCTTTTTCTTTCTCCA GTGTTCTA GTTACA

11A 1 _____ T _____

2A 2 _____ T _____

16C1 TTGA TGA GAA CA GAA CA TAA ACTA TGACCTA GGGGTTTCTGTTGGA T A ¹⁵²⁰GCTCGTAA TTA

11A 1 _____

2A 2 _____

16C1 A GAA CGGA GAAA GAA CAACAAA GACA TA TTTTCCA GTTTTTTT T ¹⁵⁷⁵CTTTACTTAAACTTT

11A 1 _____ TTTCTTTACTTAAACT

2A 2 _____ TT _____

16C1 CAAAA CAA T ¹⁶⁰⁰AGAAACTTTGTCTTTCTAA TCTTA TACTTTAAACCGA TTAA TCTTTAA CA

11A 1 TTCAAAACAA TA GAAACTTTGTCTTTCTAA TCTTA TACTTTAAACCGA TTAA TCTTTAA

2A 2 _____

16C1 GACTACA TTTTAA TA TCTACTTA TCTTT ¹⁶⁸⁰ITTA TCTCTAA GACTCCTA GTTTGA GTTTCA

11A 1 CA GACTACA TTTTAA TA TCTACTTA TCTTTTTTA TCTCTAA GACTCCTA GTTTGA GTTT

2A 2 _____

16C1 CTACA ¹⁷¹⁶IA TA TCTGTGAA TCTTGTTTTTTTCA TCTAA TGCTGTA TCA G TC ¹⁷⁶⁰ITCTGAGTTGT

11A 1 CACTACA T

2A 2 _____ A TA TCTGTGAA TCTTGTTTTTTTCA TCTAA TGCTGTA TCA GTCITCTGA GTT

16C1 GA GTGACTGTCTTGAAA GA TGTA TGGAA GAAAA GTA TGGTGTTAA TCTGCA TA GTGCTT

2A 2 GTGA GTGACTGTCTTGAAA GA TGTA TGGAA GAAAA GTA TGGTGTTAA TCTGCA TA GTGC

16C1 AA GACA GTA ¹⁸⁴⁰ITTCCA TAA TCAA TGACGGTTTAA TA GA GAAACTGA GTCCTA TGAACTGA

2A 2 TTAA GACA GTA ITTCCA TAA TCAA TGACGGTTTAA TA GA GAAACTGA GTCCTA TGAACT

16C1 ACTCCTTTA TGGCTAA TACAA TTAA GCA A ¹⁹²⁰GAA TGGA GAA TA GAA TTGA TTGGCTACA GTA

2A 2 GAACTCCTTTA TGGCTAA TACAA TTAA GCAA ¹⁹⁸⁴GAA TGGA GAA TA GAA TTGA TTGGCTACA G

16C1 CA GA TTA TCAAAAA TAAA TGCAACTTAAAAA G C ²⁰⁰⁰IGGAAA GTGTGTGTC T ITA TTGTTCAG

2A 2 TACA GA TTA TCAAAAA TAAA TGCAACTTAAAAAAAAAAAAAAAAAAAAA

16C1 CTCA CA TTGAAA GTA GAA GTGCA TCTTTA GA GCCTTAA GAAACTA GGTA A GTTGTTC



```

2080
16C1 TAA TACA CT ΔA GTGCCCTGCTCAAAAACCGCCTCCGA GTGA GGGCGTGTCTTTGGA GGCCG
2160
16C1 CGACGCTGCTCTA GGTCTCGGA TA GTGT T ΔCTGGA GACTTGCAA TTTCTTTGTTCTTTTCC
2240
16C1 TCCTGAA GA TGCTGAA GCTTCTAAA TGAA GCA GAAAAAAAAACTTTGTC A ΔA GCAACTTAG
16C1 AA GTAA GGTTAA GTA TAA TGA ACTACAAA GTA GCAA TCA TAA CA TTTGTACTTTAAAAAC
2320
16C1 TA TCCTA T G ΔACTGGAA GGCCTGTA GCTTCA TTTTTGGTGTGCTTTAAA GA GAAA GTCTA
2400
16C1 GTA TAA GGCTACAAAAA TAA TTTAA TA T A ΔTTAAAA CAAA TA TGGTTT GCCCTGGA GTTA
2480
16C1 TCGGTA TTTTGA TGCTAA TTTCA CTGCCCAA GGA CA GCTGCTTA GTC A ΔA TACTCA GGA
16C1 A TCA GTGACTTCA CCA GAA CCTTCTTCCCA CTGAA TTTGTA AAA TACA GGTGA GGGGCA G
2560
16C1 GTA TA GGA T ΔGAA GGA GGCCTGTCA TTGGA GGA GAA GGAA GGA TGGGCGGGA GA GAA GTT
2640
16C1 TGAA GGAA GA GGA GAA GACTGGAA TGG AA ΔA GA GGAA GA GA CA GGA GGGGA GA GA GA GA
2720
16C1 A GCCA TGGCA GGA GACA TTAA GA TTCTGTTCTGTGTA TTACA GGTTG C ΔA TTAA TA TGT
16C1 TCTTAA GGGTA TGGTA TGGTACTGGGCTTTGTA TGTTTA GGTGGGCAA TTA TA TCTTA TCAA
2800
16C1 TTGGA TCT A Δ

```

Fig. 2 The sequence of cDNA

16C1- the Se-P cDNA of liver, 11A 1 and 2A 2-the Se-P cDNA of brain. From the ATG of the initiation codon (35) to the TAA of the termination codon (1192) of 16C1 is the open reading frame (ORF). The TGA of bold face is the codon of selenocysteine. The line part represents that the sequence is same.

参 考 文 献

- 1 Yang Jian-guo, Morrison-Plummer J, Burk R F. Purification and quantitation of a rat plasma selenoprotein distinct from glutathione peroxidase using monoclonal antibodies. *J Biol Chem*, 1987, 262: 13372
- 2 Hill K E, Lloyd R S, Yang J G, et al. The cDNA for rat selenoprotein P contains 10 TGA codons in the open reading frame. *J Biol Chem*, 1991, 266: 10050
- 3 Read R, Bellow T, Yang J G, et al. Selenium and amino acid composition of selenoprotein P, the major selenoprotein in rat serum. *J Biol Chem*, 1990, 265: 17899
- 4 杨晓光, 田园, Hill KE, 等. 大鼠脑中血浆硒蛋白 P. *营养学报*, 1997, 19: 375
- 5 Berry M J, Banu L, Chen Y, et al. Recognition of UGA as a selenocysteine codon in type I dehydratase requires sequences in the 3' untranslated region. *Nature (Lond)*, 1991, 353: 273
- 6 Motsenbocker M A, Tappel A L. A selenocysteine-containing selenium-transport protein in rat plasma. *Biochim Biophys Acta*, 1982, 719: 147

SELENOPROTEIN P cDNA OF RAT BRAIN

Yang Xiaoguang, Tian Yuan, Hill K E, Burk R F

(*Institute of Nutrition and Food Hygiene, CA PM, Beijing 100050*)

For establishing that selenoprotein P is synthesized in brain, a λ ZA P II rat brain library was screened with selenoprotein P cDNA of liver. Positive clones were obtained and sequenced. The selenoprotein P cDNA of brain (2041 bp with poly A tail) is shorter than the cDNA of liver (3800 bp). Comparing the sequence of selenoprotein P cDNA coming from brain with the sequence of cDNA of liver, the open reading frame is the same, even though there are some differences in 5'utr and 3'utr between the two cDNA. The results indicate the selenoprotein P can be expressed in brain too. This results do not support the proposed function of selenoprotein P as selenium transport protein.

Key words: selenoprotein brain cDNA

收稿日期: 1997-03-17