

综述

人组织型纤溶酶原激活剂 (tPA) 基因

钱民章

宋后燕

(遵义医学院生化教研室, 贵州遵义) (上海医科大学生化教研室, 上海)

纤溶酶原激活剂 (PA) 是以丝氨酸为活性中心的蛋白酶。它能将纤溶酶原转化为纤溶酶, 在纤溶系统中有重要作用^[1]。同时, PA 与细胞转移、血液凝固、组织重建、激肽产生及原发性肿瘤的发生等生理和病理生理现象有密切的关系^[2-4]。

人纤溶酶原激活剂有尿激酶型 (uPA) 及组织型 (tPA) 两种主要类型。这两种蛋白质免疫性不同, 分子量有区别^[5,6], 由不同的基因编码^[7]。

tPA 主要是由血管内皮细胞合成的单链多肽, 分子量约为 68000, 在纤溶酶或胰蛋白酶作用下, 水解 Arg₂₇₅-Ilu₂₇₆ 肽键, 形成由二硫键连接的双链结构^[8]。氨基末端在重链, 羧基末端在轻链。重链有两个三角区结构^[9], 并有与纤维结合素指形结构有同源性的区域^[10] 和生长因子样区域等^[11]。轻链含有由 His(322)、Asp(371)、Ser(478) 组成的酶活性中心^[12]。

tPA 对纤维蛋白原亲和力很低, 对纤维蛋白亲和力极高, tPA-纤维蛋白复合物能高效且特异地激活血凝块中的纤溶酶原, 形成纤溶酶, 后者溶解血栓中的纤维蛋白, 但几乎不激活循环血液中的纤溶系统, 因此, tPA 是一种较为理想的血栓溶剂^[13]。

tPA 在组织和体液中含量极微, 从天然组织提取或人工合成药用 tPA 均不切实际, 唯有基因工程是合理途径。1983年, 美国 Genentech 公司构建了 cDNA 克隆,

并在原核和真核系统中均成功地表达了 tPA, 其产品已试用于临床^[9]。1985年, 英国 Browne 等分离了 tPA 基因, 在哺乳类细胞中表达了 tPA^[14]。本文简要综述 tPA 基因的分离、结构与功能的关系及其染色体定位。

tPA 基因的分离

Lindenmaier 等从人胎盘组织提取 DNA, 与 cosmid 构建成人染色体基因库^[15]。1984年, Ny 等人用编码人 tPA B 链的 cDNA 作探针^[16], 从这一基因库中分离到人 tPA 基因^[17]。他们先将基因库分为 12 个组分, 用限制性内切酶水解, 经 Southern 固相分子杂交优选^[18], 得到一个强杂交组分。将该组分转化细胞后, 用菌落原位杂交, 从 150000 个转化子中分离到 7 个阳性克隆。这些克隆的 DNA 限制性酶谱都十分相似。最后用 DNA 顺序分析鉴定了他们分离到的 tPA 基因——PA 01, 该基因总长度约 20kb, 但缺少 mRNA 5'-末端 58 个核苷酸^[17]。

Ny 等人还合成了特异性的 5'-末端和 3'-末端寡聚核苷酸探针, 用以测定 tPA 基因 5'→3' 的方向。其 5'-特异性探针与 tPA mRNA 91—108 核苷酸互补, 与 PA-01 BamH I/Cla I 的 6.9kb 片段杂交。3'-特异性探针与 tPA mRNA 的 2168-2185 核

本文于 1986 年 12 月 26 日收到。

苷酸互补,与 PA-01 Bgl II /EcoR I 的 440 bp 杂交^[17]。

1985年, Browne 等人也用 tPA cDNA 探针,筛选了人染色体基因库,分离到 9 个含有人 tPA 基因的阳性克隆,其中一个所携带完整的 tPA 基因,全长约 40kb^[14]。他们采用了两个寡聚核苷酸探针分离完整的 tPA 基因。第一个的顺序为 5'-d(CCT-GGTTACGGTCGCATGT)-3', 此段顺序跨越编码 tPA B 链的部位和 mRNA 3'-非翻译区之间^[9], 第二个的顺序为 5'-d(CC-CTCTCTTCATTGCATCC)-3', 相应于 tPA 疏水信号肽的 N-末端部位^[9]。首先用探针 (1) 筛选出 tPA cDNA 克隆, 从中分离 cDNA 作为探针, 再与探针 (2) 一起筛选染色体基因库, 得到与两个探针均杂交的阳性克隆。用限制性内切酶谱证明该克隆含有完整的 tPA 基因^[14]。

Browne 的 cosmid 载体 (pTCF) 含有氨基糖苷 3'-磷酸转移酶选择标志^[19], 他们将筛选到的含完整 tPA 基因的 pTCF 重组子转移到鼠 L 细胞中, 有一半的转化细胞产生了人 tPA, 这种 tPA 的纤溶活性能被特异性的抗 tPA 抗体抑制, 并用凝胶电泳证明了它与人 Bowes 细胞株分泌的 tPA 分子量大小一致, 从而进一步证明了他们分离到的阳性克隆携带了完整的 tPA 基因, 包括编码基因和表达所需的其他结构元件 (图 1)^[14]。

1986年, Sandra 用 Alu I/Hae III 水解人胎肝染色体 DNA 制成基因库, 并用 tPA cDNA 中段顺序为探针, 从基因库中分离到 7 个阳性噬斑; 用 tPA cDNA 5'-末端顺序为探针分离到 12 个阳性噬斑, 经限制性内切酶图谱等分析证明他们分离到完整的 tPA 基因^[20]。

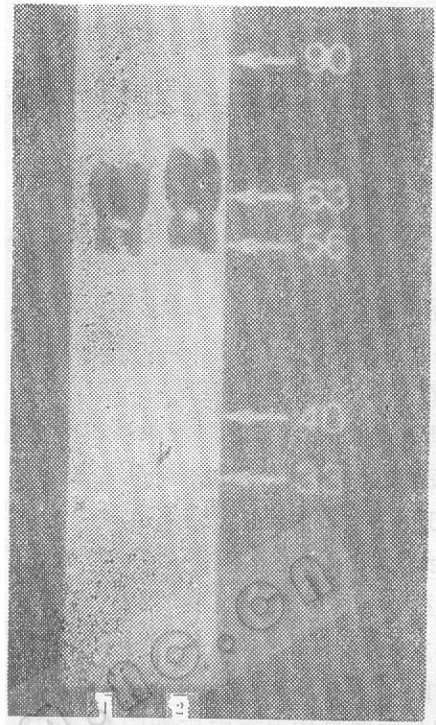


图 1 tPA 的 SDS 聚丙烯酰胺凝胶电泳图谱^[14]
1. 20 μ l Bowes 细胞分泌的 tPA
2. 20 μ l 鼠 L 细胞中重组子表达的 tPA

tPA 基因的结构特征与功能

1983年, Pennia 等成功地构建了编码 tPA 的完整 cDNA 克隆, 并在大肠杆菌中进行了表达^[9]。该克隆的 tPA ds-cDNA 的限制性内切酶位点, 也同样出现在 tPA 基因中^[17]。

Ny 等的 tPA 基因, 用限制性内切酶水解, 琼脂糖凝胶电泳, 与 tPA 探针作 Southern 固相分子杂交, 然后将阳性片段克隆到 M13, mP8 或 mP9 中, 进行顺序分析, 显示了有以下结构特征:

tPA 基因至少由 13 个内含子分隔为 14 个外显子, 每一个内含子 5' 端以 G-T 开始, 3'-端以 A-G 结束 (图 2)。

tPA 基因的 14 个外显子, 分别编码 tPA 各个功能区域^[21, 22]。

Exon I, 位于 tPA 基因的 5' 端, 编码 mRNA 非翻译区^[17]。

Exon II, 编码信号肽及其下游区的 1—3 个氨基酸。信号肽与 tPA 分泌有关, 其确切大小尚不清楚^[17, 23, 24]。

Exon III, 编码前导肽样结构^[17]。

据 Genentech 报道, 在 tPA 分子中—35—1 氨基酸属于信号肽和前导肽。其中前 20—23 个氨基酸组成疏水的信号肽后 12—15 个组成亲水的前导肽。在 tPA 基因中, 信号肽及前导肽分别由不同的外显子编码^[25]。

Exon IV, 编码重链“指形区”^[17]。tPA 的重链“指形区”, 是 tPA 与纤维蛋白亲和结合的位点, 并涉及到粘附结合功能^[26]。鸡粘连蛋白的基因至少含有 48 个外显子, 其中一个已测定了顺序^[27], 这个外显子与人 tPA 基因中第 4 个外显子大小类似, 并有同源性。tPA 指形区与纤维结合素的指形区也具有同源性。

Exon V, 编码生长因子样区域。人高分子量 uPA、牛凝血因子区、牛凝血酶原、人鼠转化生长因子等分子中也都存在生长因子样结构, 并都具有同源性。这种编码结构是否与这些蛋白质的一种共同的功能相关联, 现在还不能肯定^[10, 17]。

Exon VI—IX, 编码重链二个三角区部位。tPA 重链含有二个三角区, 两者之间仅有二硫键连接。这种结构也存在于凝血酶原、纤溶酶原和尿激酶^[11], 决定了这些蛋白质均能与纤维蛋白结合^[28]。这两个三角区有十分相似的内含子—外显子图谱, 编码顺序很象来源于同一祖先。内含子 F 和 H 均通过隔裂一个精氨酸密码把三角区的编码部分分为两个外显子。在 tPA 的指形区和三角区都存在纤维蛋白结合位点。显然这两个部位对于 tPA 与纤维蛋白的结合是十分重要的^[17]。

Exon X—XIII, 编码轻链部分。含有活性中心的 tPA 轻链由 4 个内含子分隔的 5 个外显子所编码^[17]。糜蛋白酶、胰蛋白酶和弹性蛋白酶都是以丝氨酸为活性中心的蛋白酶, 它们的内含子部位已明确^[29]。比较这些蛋白质和 tPA 内含子酶谱, 可看到它们都具有与 tPA 基因的 J, L, M 相应的内含子。这说明以丝氨酸为活性中心的蛋白酶家系不仅有同源氨基酸保守区域, 而且还有一个保守的内含子—外显子结构, 这似乎表明了编码这些蛋白酶的基因是从一个共同的祖先衍生的基因家系^[17, 29]。

1986 年, Sandra 对 tPA 基因 36594bp 进行了顺序测定。包括转录起始部位到 poly (A) 尾。他们也得到了 13 个内含子将 tPA 基因分为 14 个编码部位的同样结论, 并测定了外显子大小范围在 43—914bp, 3'-非翻译区含有最大的外显子, 其余外显子平均长度是 150bp^[20], 与 Naora 和 Blake 报道的高等真核细胞中外显子大小相接近^[30, 31]。内含子大小范围在 111—14257bp^[20]。tPA 基因中内外含子的连结与 Mount 1982 年提出的原则颇为一致^[32]。

tPA 基因中, 最多的重复顺序是 Alu I, 共有 28 个分布在各个内含子及 5' 邻近区, 占总核苷酸顺序的 22%^[17, 20]。tPA 基因中还含有 Kpu I 重复顺序, 位于转录起始部位上游 (—1418——1848), 它与几种人和其他灵长类的 Kpu I 顺序有 70% 的同源性。Rpu I 重复顺序的长短对 tPA 基因的多态性结构影响是值得注意的。此外, 在 tPA 基因中还存在几种较短的重复顺序^[20]。

Sandra 等人用 S₁ 核酸酶和核酸外切酶测定了 tPA 基因的转录起始部位, “TA-TA” 及 “CAAT” 盒的位置^[20]。用人 tPA 基因外显子与 tPA cDNA 顺序比较,

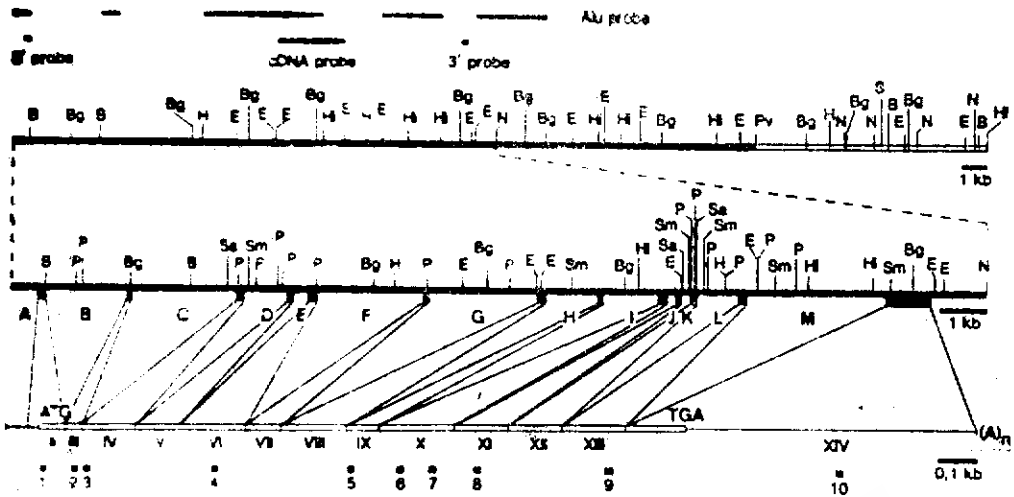


图 2 人tPA基因结构^[17]

1. 图上部第(1)条线表示pCOSpA01限制性内切酶图谱。细线为载体部分，粗线为插入的人DNA顺序
2. 图中部第(2)条线表示tPA基因，外显子Ⅰ—XⅣ及内含子A—L的位置，经Southern固相分子杂交及顺序分析证明
3. 图下部第(3)条线描述mRNA结构
5'探针、3'探针、cDNA探针、Alu I 探针在cosmidDNA上的定位以及用于 Southern 固相杂交及顺序分析的低聚核苷酸1—10的位置等均在图中表示

限制性内切酶位点的表示如下：

B, BamH I, Bg, Bgl I, E, EcoR I, Hi, Hind I,
H, Hpa I, N, Nru I, P, Pst I, Pv, Pvu I,
Sa, Sac I, S, Sal I, Sm, Sma I,

Sal, Sma I 及 Pst I 的限制性内切酶位点在第一条线上未表示。

除了几个摇摆密码改变，其他都是一致的。据推测，这些摇摆碱基是正常等位变异所导致^[9,17,20]。

tPA 基因的染色体定位

tPA 基因以单拷贝存在于染色体中^[20]。1985年，Rajput 等人用人-鼠体细胞融合，进行基因转移，研究tPA与uPA基因的染色体定位^[7]。

他们用限制性内切酶 HindⅢ 分别消化人和鼠染色体DNA，将水解片段与tPA cDNA探针杂交，结果为tPA探针与人DNA/HindⅢ 的强杂交带为 7kb，与鼠 DNA/

HindⅢ 的强杂交带为 23kb。

然后他们随机选择了32个不同的融合细胞，其细胞DNA/HindⅢ 与tPA探针杂交的结果为：32个细胞中都有23kb条带，18个细胞中尚有 7kb 强杂交带。其余24个融合细胞中 7kb 强杂交带消失。进一步作染色体组型核型鉴定发现：含有 7kb 杂交条带的18个融合细胞中都有第 8 对染色体，而24个不含 7kb 条带的融合细胞中均不存在第 8 对染色体。因此证明人 tPA 基因定位于第 8 对染色体^[7]。用同样的方法，他们又确定了人 uPA 基因定位于第 10 对染色体^[7]。这个结论与过去提出的 PA 基因定位在第 6 对染色体是不相同的^[88]。

参 考 文 献

- [1] Christman, J. R. et al., In: *Proteinases in Mammalian Cells and Tissues*, ed. Barret, A.J. (Elsevier, Amsterdam), p. 91, 1977.
- [2] Mullins, D. E. et al., *Biochim. Biophys. Acta*, 685:17.7, 1982.
- [3] Reich, E. et al., *Cancer. Res.*, 43:5467, 1983.
- [4] Ossowski, L. et al., *Cell*, 35:611, 1983.
- [5] Steffens, G.J. et al., *Physiol. chem.*, 363:1043, 1982.
- [6] Pohl, G. et al., *Biochemistry*, 23:3701, 1984.
- [7] Rajout, B. et al., *Science*, 230(4726):672—674, 1985.
- [8] Wallen, P. et al., *Eur. J. Biochem.*, 132: 681—686, 1983.
- [9] Pennica, D. et al., *Nature (London)*, 301:214—221, 1983.
- [10] Banyal, L. et al., *FEBS Lett*, 163:37—41, 1983.
- [11] Günzler, W.A. et al., *Hoppe-Seyler's Z. Physiol. Chem.*, 363:1155—1165, 1982.
- [12] Strasburger, W. et al., *FEBS Lett*, 157:219—223, 1983.
- [13] Welmar, W. et al., *Lancet* ii 1018—1020, 1981.
- [14] Browne, M. J. et al., *Gene*, 33:279—284, 1985.
- [15] Lindenmaier, W. et al., *Nucleic Acid Res.*, 10:1243—1256, 1982.
- [16] Ny, T. et al., *Proc. Natl. Acad. Sci. USA*, 80:349—352, 1983.
- [17] Ny, T. et al., *Proc. Natl. Acad. Sci. USA*, 81:5355—5359, 1984.
- [18] Lund, B. et al., *Proc. Acad. Sci. USA*, 81:2435—2439, 1984.
- [19] Grosveld, F.G. et al., *Nucleic Acid Res.*, 10:6715—6732, 1982.
- [20] Sandra, J. et al., *J. of Bio. chem.*, 261(15):6972—6985, 1986.
- [21] Gilbert, W., *Nature (London)*, 271:501, 1978.
- [22] Blake, C.C. F., *Nature (London)*, 273:267, 1978.
- [23] Houck, C.M. et al., *J. Mol. Biol.*, 132:289—306, 1979.
- [24] Austen, B. M., *FEBS. Lett*, 103:308—312, 1979.
- [25] Jounvall, H. et al., *FEBS Lett*, 156:47—50, 1983.
- [26] Petersen, T.E. et al., *Proc. Natl. Acad. Sci. USA*, 80:137—141, 1983.
- [27] Hirano, H. et al., *Proc Natl. Acad. Sci. USA*, 80:46—50, 1983.
- [28] Sottrup-Jensen, L. et al., In: *Progress in chemical Fibrinolysis and Thrombolysis*, eds. Davidson, J.F. Rowan, R.M. Samama, M.M. & Desnoyers, P.C. (Raven, New York) Vol. 3, pp. 191—209, 1978.
- [29] Graik, C.S., *Science*, 220:1125—1129, 1983.
- [30] Naoro, H. et al., *Proc. Natl. Acad. Sci. USA*, 79:6196—6200, 1982.
- [31] Blake, C., *Nature*, 306:535—537, 1983.
- [32] Mount, S.M., *Nucleic Acid Res.*, 10:459—472, 1982.
- [33] Kacherlapati, R. et al., *Cell*, 15:1331, 1978.