

利用蛋白质限制性水解法解析 大肠杆菌 T 蛋白的独立结构域*

陈 静 陈枢青 **

(浙江大学药学院生物制药研究室, 杭州 310031)

摘要 为了解析分支酸变位酶和预苯酸脱氢酶在大肠杆菌 T 蛋白的定位, 根据 T 蛋白限制性水解结果, 分段克隆分支酸变位酶和预苯酸脱氢酶。T 蛋白限制性水解结果显示, 第 93 位氨基酸是大片段的 N 端, 分段克隆的 1~93 片段测定得到分支酸变位酶活性, 96~373 片段得到了预苯酸脱氢酶活性。研究表明, 大肠杆菌 T 蛋白由两个独立结构域组成, N 端 93 个氨基酸组成了分支酸变位酶, C 端 277 个氨基酸组成了预苯酸脱氢酶。

关键词 T 蛋白, 结构域, 限制性水解

学科分类号 Q812

在细菌、真菌及植物中, 分支酸 (chorismate) 是一种位于关键分叉点上的中间代谢产物, 是所有芳香族氨基酸合成的共同前体。它可在双功能酶分支酸变位酶 (chorismate mutase, CM) 和预苯酸脱水酶 (prephenate dehydratase, PDT) 的催化下合成苯丙氨酸, 在另一个双功能酶分支酸变位酶和预苯酸脱氢酶 (prephenate dehydrogenase, PDH) 的催化下合成酪氨酸。前者被称为 P 蛋白^[1], 后者为 T 蛋白^[2]。由于人体不能合成芳香族氨基酸, 没有该类酶的存在, 所以其抑制剂无疑将成为高效、无毒、广谱的抗菌药物。因此, 近年来 P 蛋白和 T 蛋白作为研发新药的靶位点在国际上受到了重视^[3]。

大肠杆菌 P 蛋白和 T 蛋白都是双功能酶, 分子质量大小类似, 空间结构可能也有相似之处。1990 年, 美国康奈尔大学的 Ganem 研究小组^[4]发现大肠杆菌 P 蛋白的 N 端 113 个氨基酸是一个独立的结构域, 其能表达全部 CM 的活性。同年该研究小组^[5]解析了 CM 的空间结构, 它由三段 α 融合螺旋组成, 其中一段 α 融合螺旋上有亮氨酸拉链结构, 正是通过这段亮氨酸拉链结构使 P 蛋白组成了二聚体。随后几年, 该研究小组^[6]又克隆了 PDT 和 P 蛋白的调节结构域 (R domain), 发现其都有独立的功能结构域, 可以分别表达并显示完整的活性。这样, 拥有 386 个氨基酸的 P 蛋白被发现含有三个独立的功能结构域, 其 N 端 113 个氨基酸为 CM, C 端 100 个氨基酸是调节结构域, 中间 173 个氨基酸是 PDT。P 蛋白调节结构域的天然作用是结合终产物苯丙氨酸并反馈抑制 PDT 的活性, 他们发现

克隆的独立 PDT 结构域不受苯丙氨酸的反馈抑制, 也发现克隆的独立调节结构域仍然有结合苯丙氨酸的活性, 尤其有意义的是将独立的调节结构域和 CM 直接克隆成一个融合蛋白, 发现苯丙氨酸能够反馈抑制与其直接结合的 CM 活性, 说明苯丙氨酸与调节结构域的结合, 触发了调节结构域的空间结构发生改变, 这种改变可以影响与其相连的结构域的活性。

大肠杆菌 T 蛋白是拥有分支酸变位酶和预苯酸脱氢酶的双功能酶, 是酪氨酸合成代谢步骤中的关键酶, 其分子质量为 42 ku, 含 373 个氨基酸, 并能够被终产物酪氨酸反馈抑制^[2]。大肠杆菌 T 蛋白 CM 的催化活性与 P 蛋白 CM 的活性完全一样, 但其相对应的氨基酸序列却只有 25% 的同源性, 用计算机对该段序列进行空间结构预测, 发现预测的空间结构与 P 蛋白 CM 的实际空间结构极为相似。Turnbull 等^[7]曾报道 T 蛋白的 CM 和 PDH 有各自的活性中心, 但其 CM 和 PDH 活性又相互依赖, 这都提示 T 蛋白的结构与 P 蛋白虽有类似之处, 区别也很大。

本研究拟利用蛋白质限制性水解方法, 预测 T 蛋白中结构域之间的连接部位, 再分段克隆 T 蛋白片段, 从而确定 CM 和 PDH 在 T 蛋白中的序列定位, 为研究 T 蛋白作用机理, 和进一步研发高效无毒的抗菌药物奠定基础。

* 浙江省自然科学基金资助项目 (302110)。

** 通讯联系人。

Tel/Fax: 0571-87217406

E-mail: chenshuqing@zju.edu.cn, 9741426@163.com

收稿日期: 2003-12-10, 接受日期: 2004-01-25

1 材料与方法

1.1 材料

T蛋白由本实验室制备^[8]; 含有TyrA序列的质粒pKB45由美国康奈尔大学Wilson教授赠送; 表达质粒pET26b+是Novagen公司产品; 转化态细胞BL21 Gold (DE3)购于Stratagene公司; 1 kb DNA Ladder、连接酶和限制性内切酶购于New England Biolabs; PCR引物由康奈尔大学生物技术中心合成, 超纯琼脂糖和SDS-聚丙烯酰胺凝胶电泳(SDS-PAGE)蛋白质分子质量标准Benchmarker是GibcoBRL的产品; 质粒抽提试剂盒、PCR产物纯化试剂盒和凝胶中DNA纯化试剂盒是QIAGEN公司产品; LB培养基是United States Biological产品; GC-rich PCR试剂盒是Roche公司产品; 木瓜蛋白酶、分支酸和预苯酸等其他试剂均为Sigma公司产品。

1.2 仪器

PCR仪是Thermolyne公司产品, 分光光度计UV-1201是Shimadzu公司产品, 冷冻离心机J2-21是Beckman公司产品, DNA测序仪ABI 3700和蛋白质N端序列分析仪Precise 492是PE公司产品。

1.3 方法

1.3.1 T蛋白的限制性水解: 利用不同摩尔数比例的木瓜蛋白酶和不同的水解时间对纯化了的T蛋白进行限制性水解。将20 μg T蛋白溶解于100 μl水解缓冲液(0.1 mol/L NH₄OAc, 0.004 mol/L EDTA, 0.01 mol/L半胱氨酸pH 6.8), 加2 μl浓度为0.01 g/L(木瓜蛋白酶:T蛋白=1:1 000)的木瓜蛋白酶, 然后37℃孵育。每隔0 min, 15 min, 30 min, 45 min, 60 min, 90 min, 120 min取出10 μl水解样本和5 μl SDS加样缓冲液, 经沸水浴3 min后, SDS-PAGE电泳鉴定。由于酪氨酸是T蛋白的特异性抑制剂, 在蛋白质上一定有其特异性结合位点。为了研究酪氨酸对T蛋白限制性水解造成的影响, 另外设置了一组添加300 μmol/L酪氨酸的对照组。

1.3.2 蛋白质N端序列分析: 首先将上述SDS-PAGE电泳胶进行蛋白质印迹转移到Immobilien-P薄膜上。电转移的条件是50 V, 1 h。薄膜经0.1%考马斯亮蓝染色10 min, 再在90%甲醇和7%醋酸中脱色至背景无色。切下需要进行蛋白质N端序列分析的条带, 即刻寄送样本至美国康奈尔大学生物资源中心进行蛋白质N端序列分析。

1.3.3 质粒构建和转化: 根据Tyr A基因序列和T蛋白限制性水解结果, 设计PCR特异性扩增引物对T蛋白片段的基因进行扩增。在设计引物时引入Xho I酶切位点CTCGAG和Nde I的酶切位点CATATG, 并同时在终止密码前引入了6个组氨酸的密码以利产物的分离纯化。PCR扩增产物经纯化、酶切和表达质粒pET26b+连接。pET26b+含有T7启动子和lac操纵子, 可以经IPTG诱导大量扩增目的基因。pET26b+还含有抗卡那霉素基因, 可用于质粒的筛选。将连接后的质粒转化进入宿主BL21 Gold (DE3)转化态细胞, 该细胞既可用于质粒制备也可以用于诱导表达。挑取转化克隆进行单克隆培养, 并抽提质粒进行酶切插入序列鉴定和插入序列的DNA序列测定, 以确认在PCR扩增时没有引入突变。

1.3.4 诱导表达: 接种含上述质粒的菌株于20 ml含60 mg/L卡那霉素的LB培养基, 37℃震摇培养过夜。菌液再次接种到不含卡那霉素的LB培养基, 接种量为1%。37℃震摇培养约2 h至菌液 $A_{600} = 0.6 \sim 0.8$ 。加IPTG使终浓度为1 mmol/L, 再继续30℃震摇培养2 h, 8 000 r/min离心30 min收集菌体, -20℃储存备用。

1.3.5 表达产物的分离纯化: 相当于250 ml培养液的菌体加25 ml结合缓冲液(5 mmol/L咪唑, 0.5 mol/L NaCl, 20 mmol/L Tris-HCl pH 7.9), 充分混悬, 利用细胞破碎仪裂解细胞, 压力 1.379×10^7 Pa, 9 000 r/min离心30 min, 弃去沉淀。上清加0.15倍体积的30%硫酸链霉素沉淀核酸, 冰浴30 min, 9 000 r/min再离心30 min, 上清液即可用于his-tag亲和层析。亲和层析柱的填料是5 ml NTA-Ni-树脂, 首先用15 ml蒸馏水洗去树脂保存液, 再用25 ml荷载缓冲液(50 mmol/L NiSO₄)使树脂荷载金属离子镍。用15 ml结合缓冲液平衡柱子并洗去未结合的镍离子, 即可加样。对于5 ml的填料, 加样量根据带his-tag的蛋白质计算可达40 mg, 加样量越接近饱和, 纯化蛋白质的纯度就越高。加样后用50 ml结合缓冲液冲洗未结合杂蛋白, 并使带his-tag的T蛋白与镍离子结合。再用30 ml洗涤缓冲液(60 mmol/L咪唑, 0.5 mol/L NaCl, 20 mmol/L Tris-HCl pH 7.9)充分洗去杂蛋白。然后用30 ml洗脱缓冲液(1 mol/L咪唑, 0.5 mol/L NaCl, 20 mmol/L Tris-HCl pH 7.9)洗下带his-tag的T蛋白, 收集该组分即为纯品T蛋白。树脂还可以继续用15 ml剥离缓冲

液 (100 mmol/L EDTA, 0.5 mol/L NaCl, 20 mmol/L Tris-HCl pH 7.9) 彻底洗去镍离子, 树脂可以较长期储存于该溶液中。收集到的纯品 T 蛋白, 立即加入 0.5 mol/L DTT 和 0.5 mol/L EDTA 至终浓度为 1 mmol/L, 于透析液 (0.1 mol/L MES, 0.051 mol/L N-乙基吗啉, 0.01 mol/L 二乙醇胺, 1 mmol/L EDTA, 1 mmol/L DTT, 10% 甘油) 中透析过夜, 分装后 -70℃ 保存。

1.3.6 产物鉴定: 分支酸变位酶和预苯酸脱氢酶的活性测定按照 Davidson 等^[9]的方法进行。活性单位定义为在测定条件下每分钟产生 1 μmol 产物所需的酶量。比活性为每毫克蛋白质所拥有的酶活性单位数。表达产物的分子质量测定和纯度鉴定利用 SDS-聚丙烯酰胺凝胶电泳, 蛋白质浓度测定按照 Bradford^[10] 法进行, BSA 作为标准。酪氨酸调节作用按 Dixon 作图法测定。

2 结 果

2.1 T 蛋白的限制性水解

按 1 分子木瓜蛋白酶对 1 000 分子 T 蛋白的比率进行 37℃ 孵育, 在不同的水解时间停止孵育, 样本的 SDS-PAGE 电泳结果见图 1。随着水解时间的延长, T 蛋白被渐渐水解, 有一条分子质量大约为 25 ku 到 30 ku 的谱带较为持久地存在, 其可能为一个独立的结构域。300 μmol/L 酪氨酸的存在对限制性水解的结果没有影响。

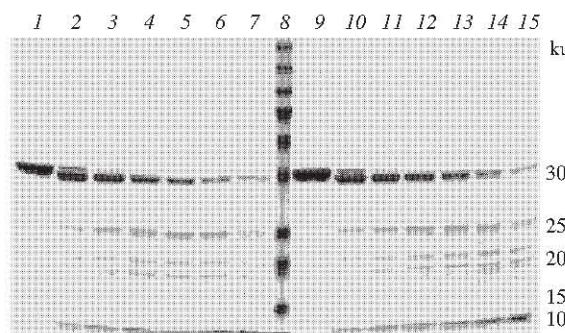


Fig. 1 SDS-PAGE analysis of samples by limited proteolytic digestion

1 ~ 7: Added 300 μmol/L tyrosine group; 9 ~ 15: Not added tyrosine group; 8: Protein marker; 1 and 9: Samples digested for 0 min; 2 and 10: samples digested for 15 min; 3 and 11: samples digested for 30 min; 4 and 12: samples digested for 45 min; 5 and 13: samples digested for 60 min; 6 and 14: samples digested for 90 min; 7 and 15: samples digested for 120 min.

2.2 N 端序列分析

把上述分子质量大约为 30 ku 的条带切割下来, 经 PE 公司的蛋白质 N 端序列分析仪 Procise 492 测得序列如下: TLCPSLRPVVIV。将该段序列与 T 蛋白的全序列进行比较, 发现其为 T 蛋白中的第 93 ~ 104 位氨基酸。也就是说, 该水解片段的 N 端位于第 93 位, 可以推测 T 蛋白至少有 2 个独立结构域, 93 位是其分界点 (图 2)。

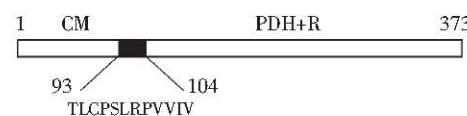


Fig. 2 Distribution of CM, PDH and R domain in T-protein

大肠杆菌 T 蛋白含有 373 个氨基酸, 限制性酶切结果发现有 30 ku 的片段较耐酶切, 该片段的 N 端氨基酸序列测定结果说明, 其 N 端定位于 T 蛋白的 93 位氨基酸。根据同源性比较推测其 N 端片段是 CM 结构域、C 端片段是 PDH 和 R 结构域。

2.3 独立结构域的克隆表达

根据上述 N 端分析的结果, 设计表达如下片段: T1 ~ 93, T1 ~ 108, T93 ~ 373, T96 ~ 373, T101 ~ 373。克隆株经诱导表达, 发现有天然 N 端的 T1 ~ 93 和 T1 ~ 108 得到了大量表达, 表达量达到 150 ~ 200 mg/L。没有天然 N 端的 T93 ~ 373、T96 ~ 373 和 T101 ~ 373 的表达量较少, 为 25 ~ 50 mg/L。由于每一个片段在 C 端都设计了由 6 个组氨酸组成的 His-tag, 故利用 His-tag 亲和层析比较容易地纯化, 得到较大量的表达产物 (图 3)。

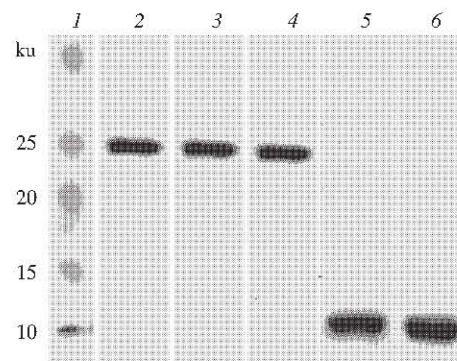


Fig. 3 Purified expression of different T protein segment

1: Protein marker; 2: T-protein 93 ~ 373 segment; 3: T-protein 96 ~ 373 segment; 4: T-protein 101 ~ 373 segment; 5: T-protein 1 ~ 93 segment; 6: T-protein 1 ~ 108 segment.

表 1 是用于克隆 T1 ~ 93, T1 ~ 108, T93 ~ 373, T96 ~ 373, T101 ~ 373 片段的引物, 表 2 是构建克

隆 T 蛋白片段的质粒时所使用的引物对.

Table 1 Primers applied in cloning T1 ~ 93, T1 ~ 108, T93 ~ 373, T96 ~ 373, T101 ~ 373 segment

Primers	Sequence (5'→3')
T01	GGT AGA CTC GAG TCA GTG GTG GTG GTG CTG GCG ATT GTC ATT CGC CTG ACG C
T02	GCT TAA GAG GTT TCA TAT GGT TGC TGA ATT G
PDH93	CGA CAA AGG ACA TAT GCA ACT TTG TCC GTC ACT GCG
PDH96	GGA TTT AAA ACA CAT ATG CCG TCA CTG CGT CCC GTG
PDH101	CCG TCA CTG CAT ATG GTG GTT ATC GTC GGC G
CM1 ~ 94	GCC ACC GCC GAC CTC GAG TCA GTG GTG GTG GTG AAG TGT TTT AAA TCC TTT GTC
CM1 ~ 108	CGA GAG GGT CAG CTC GAG TCA GTG GTG GTG GTG ACC GCC ACC GAC GAT

Table 2 Primers pairs applied in constructing plasmid to express T-protein segment

Plasmid	Primers pairs	T-protein segment
pSQC1	T02, T01	T-protein
pSQC2	PDH93, T01	T93 ~ 373
pSQC3	PDH96, T01	T96 ~ 373
pSQC4	PDH101, T01	T101 ~ 373
pSQC9	T02, CM1 ~ 94	T1 ~ 94
pSQC10	T02, CM1 ~ 108	T1 ~ 108

2.4 活性测定结果

对纯化了的表达产物进行分支酸变位酶和预苯酸脱氢酶活性测定(表 3). 发现 T1 ~ 93 和 T1 ~ 108 拥有分支酸变位酶活性, 却没有预苯酸脱氢

成了预苯酸脱氢酶独立结构域, 96 ~ 101 位氨基酸对预苯酸脱氢酶具有非常重要的作用.

4 讨 论

大肠杆菌 P 蛋白和 T 蛋白在结构与功能上有许多类似之处, 都由 2 个酶组成, 都受终产物所反馈抑制, 分子质量大小类似. 同源性比较发现, P 蛋白和 T 蛋白的 N 端有类似之处, 同源性虽然只有 25%, 可是由 DNAstar 预测的 T 蛋白 N 端 100 个氨基酸序列空间结构与 P 蛋白已知的空间结构极为相似, 都是由 3 段 α 融合组成的类似于一个“4”字形结构, 推测其可能就是分支酸变位酶的独立结构域. 但是, 要准确定位独立结构域的分界点, 还必须借助蛋白质限制性水解法提供指导性依据.

所谓蛋白质限制性水解是利用有限量的蛋白酶去水解相对较多量的蛋白质, 从而使蛋白酶选择被水解蛋白质的薄弱部分进行剪切, 被水解蛋白质的致密部分就可以完整保留. 利用蛋白质限制性水解法对蛋白质的空间结构进行预测是有一些帮助的, 尤其是对独立结构域在蛋白质氨基酸序列上的定位和分布, 指导意义较为明显.

本研究发现 T 蛋白的限制性水解结果中有一条谱带较为持久地存在, 又具有较大的分子质量. 因此, 首先选择了这个组分进行 N 端氨基酸序列分析, 结果显示其序列与 T 蛋白中 93 ~ 104 位的序列相同, N 端在 93 位. 提示该组分可能是 T 蛋白 93 位氨基酸至 C 端的片段, 是一个完整的独立结构域. 根据大肠杆菌 P 蛋白的结构域分布, 推测 T 蛋白的 N 端 93 个氨基酸组成了分支酸变位酶独立结构域, C 端 280 个氨基酸为预苯酸脱氢酶以及反馈抑制剂酪氨酸的结合位点.

Table 3 CM and PDH's specific activity of different purified segment of T protein

Plasmid	Segment of T protein	Molecular mass	Activity(CM)/(U · mg ⁻¹)	Activity(PDH)/(U · mg ⁻¹)
pSQC1	T1 ~ 373 (T-protein)	42 ku	130	98
pSQC2	T93 ~ 373		0	25.2
pSQC3	T96 ~ 373		0	55.0
pSQC4	T101 ~ 373		0	0
pSQC9	T1 ~ 94		11.4	0
pSQC10	T1 ~ 108		8.1	0

酶的活性. T93 ~ 373 和 T96 ~ 373 拥有预苯酸脱氢酶的活性, 却没有分支酸变位酶的活性. T101 ~ 373 既无分支酸变位酶活性也没有预苯酸脱氢酶的活性. 这些结果说明 T 蛋白的 1 ~ 93 位氨基酸形成了分支酸变位酶独立结构域, 96 ~ 373 位氨基酸组

根据上述推测选择性地克隆了 T 蛋白 T1 ~ 94 片段和 T93 ~ 373 片段, 为了保险起见, 同时克隆了 T1 ~ 108、T96 ~ 373 和 T101 ~ 373 片段, 以保证克隆到有活性的独立酶片段。克隆片段经分离纯化, 测得其比活性见表 3。显然, 蛋白质限制性水解法提供的依据是准确的, 克隆片段 T1 ~ 94 的确组成了分支酸变位酶, 片段 T96 ~ 373 表达预苯酸脱氢酶。如果把分界点位移至 101 位, 预苯酸脱氢酶的活性完全消失。说明蛋白质限制性水解结合片段 N 端氨基酸测序, 用于指导蛋白质中结构域的定位是一种事半功倍的方法。

值得注意的一个问题是表 3 中克隆的 T 蛋白片段的比活性都不如天然的高, 分支酸变位酶的比活性只有大约 5%, 预苯酸脱氢酶的比活性也只有大约 20%。多次重复诱导表达和分离纯化过程, 结果基本一致, 排除了变性造成的比活性降低的可能性。因而推测在天然 T 蛋白中, 分支酸变位酶和预苯酸脱氢酶有协同作用, 分支酸变位酶可以直接将催化产物预苯酸交给预苯酸脱氢酶, 反之, 预苯酸脱氢酶也在帮助分支酸变位酶释放反应产物。因此, 单独存在的分支酸变位酶和预苯酸脱氢酶的比活性就远远不如天然存在的 T 蛋白。Turnbull 等^[7]的结果可作为这个观点佐证, 更直接的证据尚需进一步的研究工作。

另外, 酪氨酸的存在没有影响 T 蛋白限制性水解的结果, 说明酪氨酸与 T 蛋白中调节结构域的结合, 虽然导致催化结构域发生结构变化, 从而

抑制 PDH 的活性, 但其对整个 T 蛋白空间结构的影响还是比较小的, 没有结构域之间交联的发生和转移。

参 考 文 献

- Schmit J C, Zalkin H. Chorismate mutase-prephenate dehydratase. Partial purification and properties of the enzyme from *Salmonella typhimurium*. *Biochemistry*, 1969, **8** (1): 174 ~ 181
- Koch G L, Shaw D C, Gibson F. Studies on the subunit structure of chorismate mutase-prephenate dehydrogenase from *Aerobacter aerogenes*. *Biochim Biophys Acta*, 1970, **212** (3): 387 ~ 395
- Husain A, Chen S, Wilson D B, et al. A selective inhibitor of *Escherichia coli* prephenate dehydratase. *Bioorg Med Chem Lett*, 2001, **11** (18): 2485 ~ 2488
- Stewart J, Wilson D B, Ganem B. A genetically engineered monofunctional chorismate mutase. *J Am Chem Soc*, 1990, **112**: 4582 ~ 4584
- Lee A Y, Stewart J D, Clardy J, et al. New insight into the catalytic mechanism of chorismate mutases from structural studies. *Chem Biol*, 1995, **2** (4): 195 ~ 203
- Zhang S, Pohnert G, Kongsaeree P, et al. Chorismate mutase-prephenate dehydratase from *Escherichia coli*. Study of catalytic and regulatory domains using genetically engineered proteins. *J Biol Chem*, 1998, **273** (11): 6248 ~ 6253
- Turnbull J, Morrison J F. Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. 2. Evidence for two different active sites. *Biochemistry*, 1990, **29** (44): 10255 ~ 10261
- Yu X X, Chen S Q. High efficient expression and purification of the *Escherichia coli* T-protein. *Chin Pharm J*, 2003, **38** (9): 707 ~ 710
- Davidson B E, Hudson G S. Chorismate mutase-prephenate dehydrogenase from *Escherichia coli*. *Methods Enzymol*, 1987, **142**: 440 ~ 450
- Bradford M M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, 1976, **72**: 248 ~ 254

Independent Domain Localization of *Escherichia coli* T-protein by Limited Proteolytic Digestion*

CHEN Jing, CHEN Shu-Qing **

(Department of Biopharmaceutics, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310031, China)

Abstract In order to elucidate the location of chorismate mutase and prephenate dehydrogenase on *E. coli* T-protein, protein limited digestion and fragmentation cloning were employed. Fragment 1 ~ 93 and fragment 96 ~ 373 were cloned and expressed according to the results of limited digestion of T-protein respectively. Two fragments were found to have chorismate mutase and prephenate dehydrogenase activities respectively. In a conclusion, *E. coli* T-protein do have independent domains, N-terminal 93 amino acids belongs to chorismate mutase and C-terminal 277 amino acids belongs to prephenate dehydrogenase.

Key words T-protein, protein limited digestion, location of activity domain

* This work was supported by a grant from The National Natural Sciences Foundation of Zhejiang (302110).

** Corresponding author. Tel/Fax: 86-571-87217406, E-mail: chenshuqing@zju.edu.cn, 9741426@163.com

Received: December 10, 2003 Accepted: January 25, 2004