

# 臺南藥理科技大學專題研究計畫成果報告

## 沙門氏豬霍亂桿菌(*Salmonella choleraesuis*) 脂解酶基因 *ly1* 的核苷酸定序及分析

計畫類別：個別型計畫 整合型計畫

計畫編號：CNFH-92-08

執行期間：92 年 1 月 1 日至 92 年 12 月 31 日

計畫主持人：張淑玉

共同主持人：

計畫參與人員：

執行單位：食品科技系

中華民國 93 年 2 月 20 日

## 摘要

脂解酶 *ly1* 基因，經核苷酸序列分析顯示此基因由 1509 個核苷酸組成，推論的 502 個氨基酸經過比對，得知它是一種 carboxylesterase 基因，而其下游含有兩個完整的 ORF(open reading frame)，分別對應了兩個轉錄方向與 carboxylesterase 基因相反的 LysR transcriptional regulator 基因及 peptidase 基因，在 peptidase 基因之後有一個不完整的 ORF，其對應了 Dipeptide transport protein。

## 前言

沙門氏桿菌(*Salmonella*)屬中的沙門氏豬霍亂桿菌 (*Salmonella choleraesuis*)，經流行病學的研究，是豬特有的病原菌 (serotype-host specificity)(1,2,3,4)，豬感染後的症狀為高弛張熱，但並無腸胃症狀，細菌可轉移至淋巴組織(Peyer's patch 即小腸淋巴集結)及網狀內皮組織(肝、脾、骨髓)導致系統性的疾病(systemic disease)，並引發致命性的菌血症及敗血症(4,5)。雖然在豬飼料中添加抗菌性的添加物，可以控制經由腸胃道的感染，但卻無法完全杜絕沙門氏豬霍亂桿菌的感染與傳播，其罹患率大約是 10%，但其死亡率卻很高，縱然痊癒，豬隻將變成帶原者(carrier)，進而持續性而不定期的由排泄物釋出病原菌(4)，由於發病期的豬隻及帶原者皆不易診斷及偵測，故造成了豬隻感染的潛在危機及經濟上的損失，非僅如此，感染的豬隻亦是人類感染沙門氏豬霍亂桿菌的病源槽(6)。國外研究指出人類感染沙門氏豬霍亂桿菌，常引發菌血症及敗血症，並多方向轉移，引起膿腫性感染，但並無腸胃性症狀，而死亡率為 16~20%(7,8,9)。

本研究從臨床分離到的 *S.*

*choleraesui* 菌株，選殖到可能的致病因子脂解酶基因 *ly1*，作核苷酸的定序與分析，這對將來 *S. choleraesui* 菌株致病因子之探討，致病機制之解明，皆有相當多的助益。

## 材料與方法

### (1) 構築對人體巨噬細胞致死性較原菌株差之沙門氏豬霍亂桿菌變異株

#### (a) 以 Transposon Tn5 構築沙門氏豬霍亂桿菌變異株

採用 Hensel 等人所提之方法稍作改變來構築沙門氏豬霍亂桿菌變異株(10)。將帶有 transposon Tn5 質體之大腸桿菌 *E. coli* S17-1  $\lambda$ pir ( $Tp^r Sm^r recA thi pro hsdR^- M^+$  RP4-2Tc:Mu:Km T7,  $\lambda$ pir) 和對 Tetracycline 具抗性的 *S. choleraesuis* SC-1 進行接合生殖 (conjugation)。

#### (b) 篩選對人體巨噬細胞致死性差之沙門氏豬霍亂桿菌變異株

由於被 *Salmonella* spp. 感染之老鼠或人類巨噬細胞(macrophage)會經由兩個不同的機制引發細胞快速

或遲延地凋凌死亡(apoptosis)而釋出 lactate dehydrogenase (LDH) (11)，因此經由巨噬細胞存活試驗 (macrophage survival assays)(12)及非放射性細胞毒性試驗(The Cyto Tox 96 non-radioactive cytotoxicity assay) (Promega)(11)試驗，加入 LDH 的受質(substrate)，測 490 nm 的吸光值便可知道 LDH 被巨噬細胞釋出之量，而得知巨噬細胞存活的情形。取 *S. choleraesuis* SC-1 的 Tn5 變異株，感染老鼠的 Bone marrow-derived macrophages(BMMs) 或 Peritoneal macrophages (13)，篩選對老鼠巨噬細胞致死性較原菌株(wild type)差之沙門氏豬霍亂桿菌變異株。

## (2) Transposon Tn5 插入位置的分析

抽取各個 Tn5 變異株(*sa1*~*sa2* 突變株)的染色體 DNA，分別以 *Bgl*II, *Eco*RI, *Kpn*I, *Pst*I 和 *Sal*I 作切割 (transposon 不含有這 5 種限制酶的切割位)，經南方點墨法 (Southern hybridization)，以  $\alpha$ -<sup>32</sup>P-labeled kanamycin gene 充當探針，選擇經限制  $\alpha$  切割後而含有 transposon 的 DNA 片段，由這些大小不等的 DNA 片段中，找出較適當的長度者(大於 10 kb，小於 20 kb)，經此實驗決定將來的實驗採用 *Eco*RI 來切割 Tn5 變異株的染色體 DNA。

## (3) 致病基因(virulence gene)之選殖

以 *Eco*RI 分別切割 Tn5 變異株的染色體 DNA，並將其插入同樣經 *Eco*RI

切割過的 pUC19 質體，將這些重組過的質體 DNA 轉型到 *E. coli* XL1B，經含有 Kanamycin 的培養基篩選得到成功的轉型子(transformant)，將此轉型子的質體 DNA(含有 transposon)抽出，並以 P6 和 P7 引子分別讀出 transposon 側邊之核苷酸序列(10)，此即為致病基因的部分核苷酸序列。

完整致病基因的選殖，乃採用菌落雜交法(colony hybridization)得到，取先前已構築好的 *S. choleraesuis* SC-1 基因庫(將 *S. choleraesuis* SC-1 的染色體 DNA 以 *Sau*3AI 切割成 5~10 kb 大小不等的 DNA 片段，並將其插入經 *Bam*HI 切割過的 pBR322 質體)，轉型到 *E. coli* XL1B，以  $\alpha$ -<sup>32</sup>P-labeled 致病基因的部分核苷酸序列充當探針，進行菌落雜交。

## 結果和討論

接合轉型子(transconjugants)經同時含有 Kanamycin 及 Tetracycline 的培養基篩選得到 2000 株，並檢測其對 Ampicillin 的敏感性，確認其為 *S. choleraesuis* SC-1 的 Tn5 變異株。從這些 Tn5 變異株篩選對老鼠巨噬細胞致死性較原菌株(wild type)差之沙門氏豬霍亂桿菌變異株，得到 6 株 Tn5 變異株分別命名為 *sa1*~*sa6* 突變株。用 *Eco*RI 來切割 Tn5 變異株的染色體 DNA，並將其插入同樣經 *Eco*RI 切割過的 pUC19 質體，經含有 Kanamycin 的培養基篩選得到成功的轉型子(transformant)，並以 P6 和 P7 引子分別讀出 transposon 側邊之核苷酸序列。

以  $\alpha$ -<sup>32</sup>P-labeled 致病基因的部分核苷酸序列充當探針，進行菌落雜交，

將雜交到的菌落質體抽出並加以定序，可得到完整的致病基因。將選殖出含有可能致病基因(virulence gene)的DNA片段，進行自動核苷酸序列分析，顯示此DNA片段的Map圖(Fig. 1)，由三個完整的基因及一個部分的基因所構成。由於Transposon Tn5所插入的位置是脂解酶 ly1 基因，因此基因的轉型子可使含有橄欖油的培養基產生透明環，故命名為脂解酶 ly1 基因，而此脂解酶 ly1 基因可能是致病基因。

脂解酶 ly1 基因，經核苷酸序列分析顯示此基因由 1509 個核苷酸組成，推論的 502 個氨基酸經過比對，得知它是一種 carboxylesterase 基因(Fig. 2)，而其下游含有兩個完整的 ORF(open reading fram)，分別對應了兩個轉錄方向與 carboxylesterase 基因相反的 LysR transcriptional regulator(Fig. 3)基因及 peptidase 基因(Fig. 3)，在 peptidase 基因之後有一個不完整的 ORF，其對應了 Dipeptide transport protein(Fig. 4)。

LysR transcriptional regulator 基因，與 carboxylesterase 基因的轉錄方向相反，兩基因間有 14 個核苷酸序列是重疊的，分析 carboxylesterase 基因的上游核苷酸序列，在轉譯啓始點上游 87 個核苷酸處，含有兩個重疊的 TN<sub>11</sub>A 核苷酸序列，符合了 LysR family 的調控特徵，故推論此 carboxylesterase 基因的表現乃受其下游的 LysR transcriptional regulator 基因所調控。其相關的證明實驗及此基因與致病機制之關係，則有待將來進一步的闡明。

## 參考文獻

- Lawson, G. H. K., and C. Dow. 1965.

- The pathogenesis of oral *S. choleraesuis* infection in pigs. J. Comp. Pathol. 75:75-81.
- Baskerville, A., and C. Dow. 1973. Pathology of experimental pneumonia in pigs produced by *Salmonella choleraesuis*. J. Comp. Pathol. 83:207-215.
- Griffith, R. W., and T. T. Kramer. 1981. Sensitivity of smooth *Salmonella choleraesuis* var. Kunzendorff field strains to antibody and complement under various conditions. Am. J. Vet. Res. 45:59-66.
- Wilcock, B. P. and K. J. Schwartz. 1992. Salmonellosis. In diseases of Swine, 7th edn, pp. 570-583. Edited by A. D. Leman and others. Ames, IA: Iowa State University Press.
- Morehouse, L. G. 1972. Salmonellosis in swine and its control. J. Am. Vet. Med. Assoc. 160:593-601.
- Berends, B. R., F. van Knapen, J. M. Snijers, and D. A. Mossel. 1997. Identification and quantification of risk factors regarding *Salmonella* spp. On pork carcasses. Int. J. Food Microbiol. 36:199-206.
- MacCready, R. A., J. P. Reardon, and I. Saphra. 1957. Salmonellosis in Massachusetts: a sixteen-year experience. N. Engl. J. Med. 256:1121-1128.
- Saphra, I., and J. H. Winter. 1957. Clinical manifestations of Salmonellosis in man: an evaluation

- of 7,779 human infections identified at the New York Salmonella Center. N. Engl. J. Med. 256:1128-1134.
9. Schmeiger, H. 1972. Phage P22 mutants with increased or decreased transduction abilities. Mol. Gen. Genet. 119:75-88.
  10. Hensel, M., J. E. Shea, C. Gleeson, M. D. Jones, E. Dalton, and D. W. Holden. 1995. Simultaneous identification of bacterial virulence genes by negative selection. Science 269: 400-403.
  11. Monack, D. M., C. S. Detweiler, and S. Falkow. 2001. *Salmonella* pathogenicity island 2-dependent macrophage death is mediated in part by the host cysteine protease caspase-1. Cell Microbiol. 3(12): 825-837.
  12. Monack, D. M., B. Raupach, A. E. Hromockyj, and S. Falkow. 1996. *Salmonella typhimurium* invasion induces apoptosis in infected macrophages. Proc Natl Acad Sci USA 93: 9833-9838.
  13. Warren, M. K. and S. N. Vogel. 1985. Bone marrow derived macrophages: development and regulation of differentiation markers by colony-stimulating factor interferons. J. Immunol. 134: 982-989.

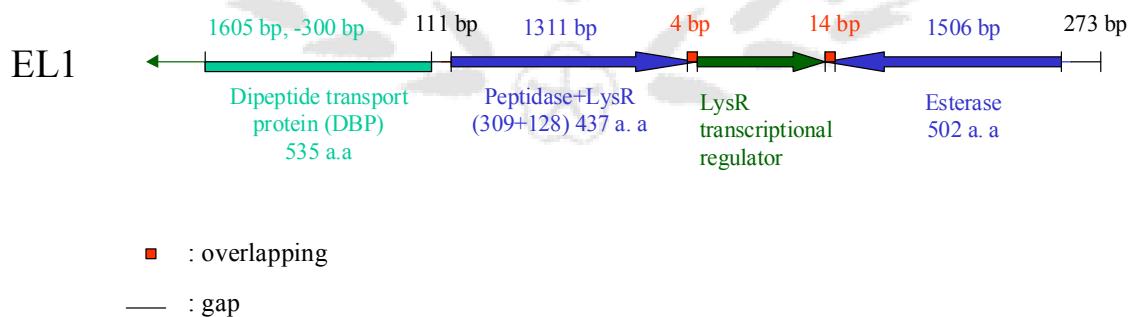


Fig. 1. The map of DNA fragment including virulence gene.

CACCTACGAGGATTTCTGCCCGTCAGCGCGGGGGATCTCCAGTCGAATCTTGGCAA  
 TGAAACGCAGGCAGCGCAGCCACGGCAATGCCAGCCGAATGCCAGTCGGCTCGG  
 CTGTGCGGTGT **ACGACGAGTT** **CGCTTTACGAGGAGGCGGAAGCGCGCAGCAAACAAGC**  
 GTTGCGGTTGCTCTTGAAACCCTACTCTGAAGGGTGTGATGGAAAAAG**AAGGTTGGT**  
**ATG**GAAAAATCCTCCGCCCTGTGGTTGAAACGCCAGGGCGCACTGATTGGTTTACT 60  
**M** E K S S A P V V E T R Q G A L I G F T  
 GAAGGCAGATACCCAGTGTGGTGTGGCATTCCCTATGCCGACCCCTGTTGGCCGTGG 120  
 E G D T H V W C G I P Y A A P P V G P W  
 CGCTGGCGCTCCCCCGCTCCCCCTGCACGCTGGGATGGCGTCCGGCGACGGCGTTC 180  
 R W R S P R P A R W D G V R P A T A F  
 TCCGCCTCCAGCTGGCAGAGCAGCGAAAGCTGTCAGGAGCTGGCGGCGGCGACCCCGC 240  
 S A S S W Q S S E S C Q E L G G G D P G  
 CAGTTCTCTGAAGACTGCCTGTATCTAACGTCTGGTCGCCAGTGGCTCGCGCCGCTCCG 300  
 Q F S E D C L Y L N V W S P V A R A A P  
 CTTCCGGT**ATG**GTCGGCTGCACGGCGAGGATTACCGCTGGCGCTGGCGGGCTGCCT 360  
 L P V **M** V W L H G G G F T L G A G G L P  
 CCGTATAACGGCAGGGCGCTGGCGAAGCGTGGCACGGTGGTGGTACGATCAATTACCGT 420  
 P Y N G R A L A K R G T V V V T I N Y R  
 CTCGGCCACCTCGGCTTTTGCCCACCGCGCTGGAGGGGAGGAAGAGCGCGTGGTG 480  
 L G H L G F F A H P A L E G E E E R V V  
 CATAACTTGCAC TGCTCGATCAGATTCAAGGCCCTGGAATGGGTGCGCGATAACATTGCC 540  
 H N F A L L D Q I Q A L E W V R D N I A  
 GCGTTCGGCGGCGATCCTGAGAACATCACCGTATTGGCGAGTCGGCCGGTGCAGCAGC 600  
 A F G G D P E N I T V F G E S A G A R S  
 GTGCTGTCGCT**ATG**GCTTCCCCGCTTGCAGGAGACTGTTCCATAAGCCATTGTGCAA 660  
 V L S L **M** A S P L A G G L F H K A I V Q  
 AGCGGGTACACGCTGCCGACACCCCGCGCAGCAGGCC**ATG**CATAAAGCGAAGCGATT 720  
 S G Y T L P D T P R E Q A **M** H K G E A I  
 GCCGCCATTTCGGCCTGCACAGTGCTACCGCGGAACAGCTCGCGCATCCCGCTGAG 780  
 A A H F G L H S A T A E Q L R A I P P E  
 GCGTTCTGGCCGCTGACCTCGCCGCTGAATATGCCCTGCGCCATCGTGGGGATTGC 840  
 A F W P L T S P L N I A P A P I V G D C  
 GTTTTGCCCTGAGGCC**ATG**CTCGACGTTTCTCGCGGCCAGCATCCTGTACCGGTG 900  
 V L P E A **M** L D V F F A A R Q H P V P V  
**ATG**ATTGGGTCGAACAGCGACGAAGCCAGCGT**ATG**TCGGTATTGGGGTCGATCTGCC 960  
**M** I G S N S D E A S V **M** S V F G V D L A  
 GGGCAGATCCAGAAGCTCCGCCGTGAGCGCGCTTGGCCTGGGGTTGATAAAGCTGCTT 1020  
 G Q I Q K L R R E R R F G L G L I K L L

TATCCTGGCGTGAAGGGCGATGAGGAAC	TCGGCAGGCAGGTATGCCGCGAC	<b>ATG</b>	GCCTTC	1080
Y P G V K G D E E L G R Q V C R D	M A F			
ACCACC <b>ATG</b> GGATAACGTGGTA <b>ATG</b> CAGGCCAGCAGCGGGCGGCAGCCTGTGCTGGCGA				1140
T T M G Y V V M Q A Q Q R A G G L C W R				
TACTGGTTGATTATGTGGCCGAAGCGGGAGCACGCGACGTACATCAACGGCGCTGGCAC				1200
Y W F D Y V A E A E H A T Y I N G A W H				
GGCAACGAAAAAGTGCCTACGTCTCGATAACCCTGGACAGGTGGAACCTTCGCGGCAG				1260
G N E K V P Y V F D T L G Q V E P S R Q				
TATGTGAATGAACCGCATCTGGCCTTCGCCGCTCAGGTGGCGACTACTGGGTGAGCTC				1320
Y V N E R D L A F A A Q V A D Y W V S F				
GCCCAGGATGCGGGGCACCGCATAGCCTGGCAGGGCCCACCGCCTGGCCCGCCTGCCGG				1380
A R D A G A R D S L A G P T R W P A C R				
AAAGGGCGGGACGTGCTGTTACGTATTGGTGTGAATAAACATGCAGGTTTCGGCTTGAA				1440
K G R D V L L R I G V N K H A G F R L E				
AACCGCTTC <b>ATG</b> CGTGCCGT <b>ATG</b> GACCTCTCAAACGGGT <b>ATG</b> AAACACACCACGTCAGC				1500
N R F M R A R M S L F K R V M K H H V S				
CTCGAC <b>TGA</b> GCAGACAGGCAGAACGCATCCAGCCGTTTCAGGGCCGGGGCGTCGC				1560
L D ***				
GTAATACCAGCCGGTAGCTGCCCGTTTATGCTCGTGTCAAAAGGGCGCATCAGCC				
GTCCGGCGGCACGTCTCCACCAGCGTTCATCCCGATGGCGATGCCAGCCCT				
GAATGGCGCGGTAAATGGCGAGATCCATGGTTCGAAGTGTGATTTGAGCATAGCTG				
GCAGCGGGACCCGGCTGTTGCCAGCCACAGCGTCCAGTCCGTTGCCCCGTGGGT				
GAAGGAAGGTGAGTGCTTCAGCGCAGAGCCGGCCGTAGCGGGCTCATCACCGGGTTAA				
GGCCTCTCGAACAGCAGATGCCGGCGCT				

Fig. 2. The nucleotide sequence of carboxylesterase

GCTTTCATCATTATTCCTGTCTTTAATTGACGGCTAATTACTCTTTGCCATT  
 CATAAAATAACATTAAGTGTGGCGAACACATGAAAATAATACAGAAGGAAACACT  
 ATGGGCTCGGAACCTCCAGACAATTAACCAACGCTTTCCGCTATCTGCCATCACC 60  
**M** G S E L S R Q L T Q R F F R Y L A I T  
 AGCCAGAGCGACCCGAAAGTCAAAACCTGCCCTCCACCCGGGCCAGCACGACATGGCG 120  
 S Q S D P K V K T L P S T P G Q H D M A  
 CGGGAGCTGGCGAAGGAGCTGGAAACGCTGGGGTAGACGATATTGTGATTGATGAGTTC 180  
 R E L A K E L E T L G L D D I V I D E F  
 GCCACCGTTACCGCCGTGAAAAAGGTAAATGTTCCCGCGCCGCGTATTGGTTTATT 240  
 A T V T A V K K G N V P G A P R I G F I  
 ACCCATATCGACACCGTCGACGTCGGTTATCCCCGGATATTCCACAAATATTAACC 300  
 T H I D T V D V G L S P D I H P Q I L T  
 TTTACGGGGATGATCTCTGTCTGAATAAAGAGAAAGATATTGGCTGCGCGTAAAAGAG 360  
 F T G D D L C L N K E K D I W L R V K E  
 CACCCGAAATTCTGGTTATCATGATGAGGAGATTATTCAGCGACGGAACCGCGTA 420  
 H P E I L A Y H D E E I I F S D G T S V  
 TTAGGCGCAGATAATAAAGCGGCCGTACCGTGGTCATGACGGTGTGGAAAACCTCACC 480  
 L G A D N K A A V T V V M T V L E N L T  
 GCTGAGCACAAACCATGGCGATATTGTGGTGGCTTGCCGATGAAGAGATTGGCCTG 540  
 A E H N H G D I V V A F V P D E E I G L  
 TGCGGCGCGAAAGCGCTGGATTAAAGCGCTTCGACGTCGATTTGCCTGGACCATCGAC 600  
 C G A K A L D L K R F D V D F A W T I D  
 TGCTGCGAGCTGGCGAAATTGTTACGAGAACCTTAACCGGGCGGCTGAAATTGCG 660  
 C C E L G E I V Y E N F N A A A A E I R  
 TTTACCGGCACGGCGCACCCGATGTCCGACCGCCAAAGGGGTGCTGGTCATCGCTG 720  
 F T G V T A H P M S A K G V L V N P L L  
 ATGGCAACGGATTTCATCAGCCATTGCGATCGCCGGCAAACCCCGGAATGCACCGAGGGG 780  
 M A T D F I S H F D R R Q T P E C T E G  
 CGCGAAGGTTATATCTGGTTAACGGCATCCAGGCCGGCAAAACGAAGCGGTGCTGAAA 840  
 R E G Y I W F N G I Q A G Q N E A V L K  
 GCCAACATTGCGATTTGACAAAGACGGTTTGCCGCCGCAAGCAGCAAATTGCCGTA 900  
 A N I R D F D K D G F A A R K Q Q I A V  
 GAGACGTTCATCGTGACCGCGCGTCACCTGAATCTTACCCATGCCGCGAAGGAGCTTGT 960  
 E T F I V T A R H L N L T H A A K E L C  
 CTGACGCAGGGCGGTAAAGCCGTAACCGCTGGAAAGCTGGTCGGCTCCCG 1020  
 L T Q G A V S R K I A S L E S W F G F P  
 CTGTTCGAGCGCCATGCCGCGCGCTGCGTCTCTCCTCACAGGGAAAGCGCCCTGCTGCCG 1080  
 L F E R H A R G L R L S S Q G S A L L P  
 GAGCTGCAATCTGCCATTGAGCACCTGCTGAACGTTGCTGAACAGGCCGAACGCCAG 1140  
 E L Q S A F E H L L N V A E Q A R T H Q  
 ACCGTAATTGCGTGTAAAGCGCCTACCTGCGCGATGCGCTGGCTGGTGCCTGCGTGTGTTA 1200  
 T V I R L K A P T C A M R W L V P R L L  
 CAGGTGGAGCGCGAACAGCCGGAACTCCAGATTGCGCTGACACCACCACCGATCACACG 1260  
 Q V E R E Q P E L Q I A L T P P P P I T T

TCAATTCAAAACCGAATCCTGTGACGCCGGATTGTGTTGGGACGCACATGATGACCG	1320
S I S K P N P V T R R L C L G R T ***	
	M M T
CCGGCGATCTGCTGTCGAAGAGGCCTTAACCCCGGTGATGAGCCCGCTACGGCCGGCT	1380
A G D L L F E E A L T P V M S P L R A G	
CTGCGCTGGAAGCACTCACCTCCTCACCCCACGCCGGACAAAACGGACTGGACGCTGT	1440
S A L E A L T F L H P T R D K T D W T L	
GGCTGGCGAAACAGCCGGTCCGCCAGCTATGCTAAAAATCAACACTTCGAAACCA	1500
W L A K Q P G P P A M L K N Q H F E T	
TGGATCTGCCATTACCGCCGCCATTCAAGGGCTGGGCATGCCATCGCGATGAAACGC	1560
M D L A I T A A I Q G L G I A I A D E T	
TGGTGGAGGAGGACGTGCGCGCCGGACGGCTGATGCGCCCTTTGACACGAGCATAAAA	1620
L V E E D V R A G R L M R P F D T S I K	
CGGGGCGAGCTACCGGCTGGTATTACCGCACGCCCGCCCTGAAAACGGGCTGGATG	1680
T G A S Y R L V L R D A P G P E N G L D	
CGTTCGCGCCTGCTGCTCAGTCGAGGCTGACGTGGTGTTCATCACCGTTGAAGAG	1740
A F R A C L L S R G ***	
GCTCATACGGCACGCATGAAGCGGTTTCAAGCCGAAAACCTGCATGTTATTACACC	
AATACGTAACAGCACGTCCGCCCTTCCGGCAGGCCAGCGCTGGCCCTGCCAG	

Fig. 3. The nucleotide sequence of dipeptidase

TTCCGAGCCCATACTGTTCTTCTGTATTATTTCATGTGTTGCCATCACTTA  
 ATGTTATTTATGAAATGGAAAAGAAGTAATTAGCCGTCGAATTAAAAGACAGGAATAA  
**ATGAT**AAAAGCACATTACA**ATG**AACGCTGGCGCTGCACTGACGGTCAGTTCC 60  
**M M K S T F T M I T L A L A A L T V S S**  
 ACCGTCGCGGCAAAACGCTGGTATTGCTCCGAAGGATGCCGGAAAATTCAATCCT 120  
**T V A A K T L V Y C S E G S P E N F N P**  
 CAGCTCTATACTCGGGGACCAGCGTGGACGCCAGCGCGTACCGTTATAACCGTCTG 180  
**Q L Y T S G T S V D A S A V P V Y N R L**  
 GTCGATTCAAACCGGGACTACCGAACCTGGTACCGAGCCTGGCGAAAGCTGGAGGTA 240  
**V D F K P G T T E L V P S L A E S W E V**  
 AGCGAGGATGGCAAGGTCTACACCTCACCGCAAAGGGTGAAATTCCACAGTAAT 300  
**S E D G K V Y T F H L R K G V K F H S N**  
 AAGCTGTTCACGCCGACGCCGACTTCAACGCCGACGTGATTCTCGTTTGC 360  
**K L F T P T R D F N A D D V I F S F L R**  
 CAGAAGGATGTGAATCATCCTTACCATACGTCTCCAACGGCAGTTATTCAAATTGAA 420  
**Q K D V N H P Y H N V S N G S Y S N F E**  
 AGTCTGGAGTTCGGCAGCCTGATTACCGCATTGATAAAAGTTGACGATCGCACCGTGC 480  
**S L E F G S L I T A I D K V D D R T V R**  
 TTCACCCCTGGCGCATCCGGAAGCGCCGTTGCTGACCTGGCGTGGTACTTGC 540  
**F T L A H P E A P F V A D L A W Y F A S**  
 ATTCTGTCGGAGTACGCCGATGCC**ATG**CTGAAAGCGGGACGCCAGAAAAGGTCGAT 600  
**I L S A E Y A D A M L K A G T P E K V D**  
**ATG**CAGCCGATTGGCACCGGCCGTTAACGCTGTCGAATATCAGAAGGATTCCGCATCC 660  
**M Q P I G T G R L S C R N I R R I P A S**  
 TCTTTACCGCTTCCCTGACTACTGGCAGGGAAAATCGAACGCTGGATCGTCTGGTGTCA 720  
**S L P L S L T T G R E N R S W I V W C S**  
 CCATCACGCCGGACGCCCTCGGTACGTTGCCAAAGTTGAGAAGAATGAGTGTCAAGGTG 780  
**P S R R T P R Y V L P K F E K N E C Q V**  
**ATG**CCGTTCCCGAACCCGGGACCTGCCGCT**ATG**AAGGCTAACAAAGACATCAACCTG 840  
**M P F P N P A D L P R M K A N K D I N L**  
**ATG**AGCAAGGCCGGCTGAATACCGGTTCCCTGGCGTTAACGCAAAAGCCGCCGCTG 900  
**M S K A G L N T G F L A F N T Q K P P L**  
 GATAACGTACAAGTGCGCCAGGCGCTGGCG**ATG**GCGATTAACAAACCGGCCATCATTGAG 960  
**D N V Q V R Q A L A M A I N K P A I I E**  
 GCGGTTTCCACGGCACCGGCACGGCGAAAAACCTGCTGCCGCTGGCGTCTGGAGT 1020  
**A V F H G T G T A A K N L L P P G V W S**  
 GCCGACAGTGAGCTGAAAGACTACGATTACGATCCAGAAAAAGCGAAAGCGCTGTTAAAG 1080  
**A D S E L K D Y D Y D P E K A K A L L K**

GAGGCGGGGTTGCCAACGGCGTAAGCATCGATCTGTGGGCC	<b>ATG</b> CCC GTGCAACGGCCG	1140
E A G F A N G V S I D L W A	<b>M</b> P V Q R P	
TATAACCGAACGCGAAGCGT	<b>ATG</b> CGGGAG	1200
Y N P N A K R	<b>M</b> A E	
<b>M</b> I Q A D W A K V G		
GTGCAGACCAAAATCGTGACCTACGAATGGGCGAATACCTGAAGCGCGTGAAGGGCGGG		1260
V Q T K I V T Y E W G E	Y L K R V K G G	
GAGCATCAGGCTGGCCT	<b>ATG</b> GGCTGGACAACGGCGACGGGGAT	1305
E H Q A A L	<b>M</b> G W T T A T G D	

Fig. 4. The partial nucleotide sequence of dipeptide transport protein

