Molecular Cloning of the cDNA for Porcine Parotid Hormone

Qian Zhang

Follow this and additional works at: https://scholarsrepository.llu.edu/etd

Part of the Genetics and Genomics Commons, and the Physiology Commons

Recommended Citation
Zhang, Qian, "Molecular Cloning of the cDNA for Porcine Parotid Hormone" (2000). Loma Linda University Electronic Theses, Dissertations & Projects. 822.
https://scholarsrepository.llu.edu/etd/822

This Dissertation is brought to you for free and open access by TheScholarsRepository@LLU: Digital Archive of Research, Scholarship & Creative Works. It has been accepted for inclusion in Loma Linda University Electronic Theses, Dissertations & Projects by an authorized administrator of TheScholarsRepository@LLU: Digital Archive of Research, Scholarship & Creative Works. For more information, please contact scholarsrepository@llu.edu.
Molecular Cloning of the cDNA for Porcine Parotid Hormone

by

Qian Zhang

A Dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Physiology

June, 2000
Each person whose signature appears below certifies that this dissertation in their opinion is adequate, in scope and quality, as a dissertation for the degree Doctor of Philosophy.

John Leonora, Professor of Physiology

Raymond G. Hall, Associate Professor of Physiology

Kerby C. Oberg, Assistant Professor of Human Anatomy and Pathology

Aladar A. Szalay, Professor of Biochemistry

Jean-Marc Tieche, Assistant Research Professor of Physiology
ACKNOWLEDGEMENTS

I am grateful to Loma Linda University for providing the facilities and financial supports. I would like to express my appreciation to the individuals who helped me complete this study. I wish to thank my advisors, Dr. John Leonora and Dr. Aladar A. Szalay, for their guidance and support. I am also grateful to my committee members, Dr. Kerby C. Oberg, Dr. Jean-Marc Tieche, and Dr. Raymond G. Hall, for their advice and comments. Special acknowledgment to Dr. Paul McMillan of Anatomy Department for his help in immunohistochemistry studies, to Dr. Eru Kyeyune-Nyombi of Biochemistry Department for his help in the construction of cDNA library, and to Dr. George Bagi of Biochemistry Department for his help in Western blotting studies.
TABLE OF CONTENTS

Approval Page................................................................. ii
Acknowledgements............................................................. iii
Table of Contents............................................................ iv
List of Figures................................................................. vii
List of Tables................................................................. ix
Abbreviations................................................................ x
Abstract................................................................................ xi

CHAPTER I: BACKGROUND...................................................... 1
A. Introduction........................................................................ 1
B. Diversity of Salivary Glands........................................... 2
C. Dental Caries and Salivary Glands.................................... 5
D. Parotid Hormone and Proline-rich Proteins..................... 8
E. Objective and Specific Aims.............................................. 14

CHAPTER II: MATERIALS AND METHODS............................. 15
A. Materials.......................................................................... 15
   1. Reagents...................................................................... 15
   2. Enzymes...................................................................... 16
   3. Plasmid Vectors.......................................................... 19
   4. *E.coli* Strains............................................................ 20
   5. Source of Animal Tissue and Mammalian Cell Lines......... 22
   6. Primers for PCR and DNA Sequencing.......................... 22
   7. Kits and Systems........................................................ 23
   8. Supplies........................................................................ 23
   9. Solutions....................................................................... 24
  10. Gels.............................................................................. 27
  11. Computer Software and Databases............................... 28
B. Methods........................................................................... 29
   1. Molecular Cloning of Porcine Parotid Hormone.............. 29
      a. Tissue Collection.................................................... 29
      b. TA Cloning of RT-PCR Products............................ 29
         1). Isolation of Total RNA........................................ 30
         2). Designing Primers............................................. 31
         3). RT-PCR............................................................ 34
         4). Cloning and Screening of Target Genes............... 34
      c. Construction of cDNA Library................................. 37
         1). Isolation of Messenger RNA.............................. 37
         2). First Strand Synthesis...................................... 40
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>3)</td>
<td>Second Strand Synthesis</td>
</tr>
<tr>
<td>4)</td>
<td>Addition of Sal I Adapters</td>
</tr>
<tr>
<td>5)</td>
<td>Digestion of cDNA with Not I</td>
</tr>
<tr>
<td>6)</td>
<td>Size Fractionation of cDNA</td>
</tr>
<tr>
<td>7)</td>
<td>Ligation of cDNA to the Vector DNA</td>
</tr>
<tr>
<td>8)</td>
<td>Introduction of cDNA constructs into E.coli</td>
</tr>
<tr>
<td>D</td>
<td>Analysis of the cDNA Library by Colony Hybridization</td>
</tr>
<tr>
<td>1)</td>
<td>Making Replica Membranes</td>
</tr>
<tr>
<td>2)</td>
<td>Labeling Probes</td>
</tr>
<tr>
<td>3)</td>
<td>Prehybridization Conditions</td>
</tr>
<tr>
<td>4)</td>
<td>Hybridization Conditions</td>
</tr>
<tr>
<td>5)</td>
<td>Posthybridization Washing</td>
</tr>
<tr>
<td>6)</td>
<td>Purifying Positive Clones</td>
</tr>
<tr>
<td>E</td>
<td>Construction of Progressive Unidirectional Deletions of Inserted DNA</td>
</tr>
<tr>
<td>1)</td>
<td>Preparation of the Starting DNA</td>
</tr>
<tr>
<td>2)</td>
<td>Restriction Digestion of Plasmid DNA</td>
</tr>
<tr>
<td>3)</td>
<td>Exonuclease III Deletion, Ligation and Transformation</td>
</tr>
<tr>
<td>4)</td>
<td>Screening of the Truncated Insert Containing Subclones</td>
</tr>
<tr>
<td>5)</td>
<td>DNA Sequencing and Data Compiling</td>
</tr>
<tr>
<td>F</td>
<td>Analysis of Positive cDNA Clones</td>
</tr>
<tr>
<td>1)</td>
<td>Partial Sequence Analysis</td>
</tr>
<tr>
<td>2)</td>
<td>Restriction Analysis of Positive cDNA Clones</td>
</tr>
<tr>
<td>3)</td>
<td>PCR Screening with Specific Primers</td>
</tr>
<tr>
<td>2</td>
<td>Tissue Distribution of PH Gene Product(s)</td>
</tr>
<tr>
<td>a</td>
<td>Northern Analysis</td>
</tr>
<tr>
<td>1)</td>
<td>Isolation of Total RNA and mRNA</td>
</tr>
<tr>
<td>2)</td>
<td>Preparation of Northern Blots</td>
</tr>
<tr>
<td>3)</td>
<td>Hybridization Analysis</td>
</tr>
<tr>
<td>b</td>
<td>In Situ Hybridization and Immunohistochemistry</td>
</tr>
<tr>
<td>1)</td>
<td>Tissue Preparation</td>
</tr>
<tr>
<td>2)</td>
<td>Preparation and Synthesis of the Riboprobes</td>
</tr>
<tr>
<td>3)</td>
<td>Hybridization Steps</td>
</tr>
<tr>
<td>4)</td>
<td>Autoradiography and Visualization</td>
</tr>
<tr>
<td>5)</td>
<td>Hematoxylin and Eosin Staining of Tissue Sections</td>
</tr>
<tr>
<td>6)</td>
<td>Immunohistochemistry for Detection of Parotid Hormone in Tissue Sections</td>
</tr>
<tr>
<td>c</td>
<td>Southern Analysis of Genomic DNAs</td>
</tr>
<tr>
<td>1)</td>
<td>Isolation of Genomic DNA</td>
</tr>
<tr>
<td>2)</td>
<td>Preparation of Southern Blots</td>
</tr>
<tr>
<td>3)</td>
<td>Hybridization Analysis</td>
</tr>
<tr>
<td>3</td>
<td>Expression of Cloned PH cDNAs</td>
</tr>
<tr>
<td>a</td>
<td>Expression in Cell-free System</td>
</tr>
<tr>
<td>1)</td>
<td>Construction of Expression Plasmids</td>
</tr>
<tr>
<td>2)</td>
<td>Translation Procedure</td>
</tr>
<tr>
<td>3)</td>
<td>Detection of Translated Proteins</td>
</tr>
</tbody>
</table>
b. Expression of PH containing cDNAs in *E.coli*. 
1) Construction of Expression Plasmids.................................84
2) The Choice of Expression Hosts...........................................85
3) Promoter Activation with IPTG............................................86
4) Detection of Expressed Proteins.........................................87
c. Transient Expression in COS-7 Cells..................................87
1) Construction of Expression Plasmid pCEP4-TP<sub>23</sub>............87
2) Transfection of COS-7 Cells with FuGENE 6...........................88
3) Detection of Expressed Proteins.........................................91

CHAPTER III: RESULTS AND DISCUSSIONS................................93

A. Cloning of Porcine Parotid Hormone cDNA...........................93
1) TA Cloning of RT-PCR Products.........................................93
2) Construction and Screening of Porcine cDNA Library...............102
3) Sequence Comparison of Three cDNAs..................................114
4) Analysis of Positive cDNA Clones......................................153
B. Tissue Distribution of PH Gene Products..............................159
1) Northern Analysis............................................................159
2) *In Situ* Hybridization and Immunohistochemical Staining.........163
3) Southern Analysis of Genomic DNAs....................................170
C. Expression of Cloned PH cDNAs..........................................175
1) Translation of Cloned PH cDNAs in Cell-free System...............175
2) Expression of Cloned PH cDNAs in *E.coli*.............................178
3) Expression of Cloned PH cDNAs in Mammalian Cells................183

CHAPTER IV. CONCLUSIONS AND FUTURE STUDIES.........................187

A. Conclusions...........................................................................187
B. Future Studies......................................................................189

CHAPTER V. REFERENCES.......................................................192
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Design of oligonucleotide primers based on the available partial amino acid</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>sequences of four peptides from the porcine parotid hormone.</td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Outline of the SuperScript Plasmid System procedure.</td>
<td>38</td>
</tr>
<tr>
<td>3.</td>
<td>Colony hybridization: identification of a specific clone from a cDNA library</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>by membrane hybridization to a radiolabeled probe.</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Construction of nested deletions for DNA sequencing.</td>
<td>52</td>
</tr>
<tr>
<td>5.</td>
<td>Plasmid DNA pSPORT 1.</td>
<td>55</td>
</tr>
<tr>
<td>6.</td>
<td>Design of oligonucleotide primer PH5P for screening of positive cDNA clones</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>using PCR.</td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Construction of plasmid DNAs for <em>in vitro</em> translation and for expression in</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td><em>E. coli</em> cells.</td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Construction of plasmid DNA pCEP4-TP23 for expression in COS-7 cells.</td>
<td>89</td>
</tr>
<tr>
<td>9.</td>
<td>Separation of RT-PCR products by gel electrophoresis.</td>
<td>94</td>
</tr>
<tr>
<td>10.</td>
<td>Comparison of nucleotide sequences of five PCR clones.</td>
<td>97</td>
</tr>
<tr>
<td>11.</td>
<td>Nucleotide sequence and amino acid sequence derived from the insert of PCR</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>clone S7-3.</td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Autoradiogram of cDNA library hybridized with radiolabeled probes.</td>
<td>104</td>
</tr>
<tr>
<td>13.</td>
<td>Progressive unidirectional deletions of a cDNA insert.</td>
<td>110</td>
</tr>
<tr>
<td>14.</td>
<td>A sample of DNA sequencing data.</td>
<td>112</td>
</tr>
<tr>
<td>15.</td>
<td>Nucleotide sequence and deduced amino acid sequence from D2A cDNA insert.</td>
<td>115</td>
</tr>
<tr>
<td>16.</td>
<td>Nucleotide sequence and deduced amino acid sequence from TP23 cDNA insert.</td>
<td>118</td>
</tr>
<tr>
<td>17.</td>
<td>Comparison of nucleotide sequences of D2A and TP23 cDNAs.</td>
<td>121</td>
</tr>
<tr>
<td>Page</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td>Comparison of amino acid sequences derived from D₂A and TP₂ cDNAs...</td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td>Nucleotide sequence and deduced amino acid sequence from SL₄₄-₁ cDNA insert...</td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td>Comparison of nucleotide sequences of TP₂₃ and SL₄₄-₁ cDNAs...</td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td>Comparison of amino acid sequences derived from TP₂₃ and SL₄₄-₁ cDNAs...</td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td>Four domains of amino acid sequences derived from the cDNA inserts of D₂A, TP₂₃ and SL₄₄-₁...</td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td>Schematic structure of proline rich proteins derived from porcine cDNA clone D₂A, TP₂₃ and SL₄₄-₁...</td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td>Comparison of consensus DNA and amino acid sequences of TP₂₃ and D₂A cDNA (from BstN I repeats) to other PRPs...</td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td>Comparison of signal peptide, repeat, and C-terminal sequences of PRPs...</td>
<td></td>
</tr>
<tr>
<td>26.</td>
<td>Restriction analysis of plasmids carrying cDNA inserts hybridized to oligonucleotide probe PHiA...</td>
<td></td>
</tr>
<tr>
<td>27.</td>
<td>Screening of positive cDNAs with primer PH5P using PCR...</td>
<td></td>
</tr>
<tr>
<td>28.</td>
<td>Northern blot analysis of polyA (+) RNAs from different porcine tissues...</td>
<td></td>
</tr>
<tr>
<td>29.</td>
<td>Localization of PH gene transcripts by in situ hybridization and the PH protein by immunostaining in porcine parotid gland tissue...</td>
<td></td>
</tr>
<tr>
<td>30.</td>
<td>Localization of PH gene transcripts by in situ hybridization and the PH protein by immunostaining in porcine sublingual gland tissue...</td>
<td></td>
</tr>
<tr>
<td>31.</td>
<td>Localization of PH gene transcripts by in situ hybridization in porcine submandibular gland tissue...</td>
<td></td>
</tr>
<tr>
<td>32.</td>
<td>Southern blot analysis of pig, human and rat genomic DNAs...</td>
<td></td>
</tr>
<tr>
<td>33.</td>
<td>Western blot analysis of translation products encoded by the 1650 bp insert of plasmid pET 28a(+) - D₂A in a coupled transcription and translation system...</td>
<td></td>
</tr>
<tr>
<td>34.</td>
<td>Separation of protein extracts from E.coli strain BL21(DE3) transformed with plasmid pET 28a(+) - D₂A and pET 28a(+) - Trd₈ by gel electrophoresis...</td>
<td></td>
</tr>
<tr>
<td>35.</td>
<td>Western blot analysis of proteins expressed in COS-7 cells transfected with plasmid pCEP4-TP₂₃...</td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. \textit{E.coli} Strains Used as Cloning and Expression Host in This Study</td>
<td>21</td>
</tr>
<tr>
<td>2. Utilization of Oligonucleotide Primers in RT-PCR Cloning of Porcine Parotid Hormone cDNA</td>
<td>35</td>
</tr>
<tr>
<td>3. PCR Programs for RT-PCR Cloning of Porcine Parotid Hormone cDNA</td>
<td>36</td>
</tr>
<tr>
<td>4. List of Nine Clones Obtained from Cloning of RT-PCR Products</td>
<td>96</td>
</tr>
<tr>
<td>5. Partial Characterization of 20 Positive Clones from Porcine cDNA Library</td>
<td>108</td>
</tr>
<tr>
<td>6. Comparison of Amino Acid Compositions</td>
<td>147</td>
</tr>
</tbody>
</table>
ABSTRACT OF THE DISSERTATION

Molecular Cloning of the cDNA for Porcine Parotid Hormone

by

Qian Zhang

Doctor of Philosophy, Graduate Program in Physiology
Loma Linda University, June, 2000
Dr. John Leonora, Chairperson

It has been found that parotid glands secrete a factor into blood that stimulates dentinal fluid movement, which in turn, prevents the development of dental caries. This factor, designated as parotid hormone, has been isolated from porcine parotid glands and its partial amino acid sequence has been determined. This study was developed to isolate and identify the cDNA sequence for porcine parotid hormone.

A porcine parotid cDNA library was constructed and hybridized with oligo probes designed based on the partial amino acid sequences of the isolated porcine parotid hormone. Over 100 positive clones were isolated and the complete nucleotide sequences were determined for three cDNA clones. The derived amino acid sequences indicate that these three proteins, ranging from 48.5 to 62.3 kDa, are unusually high in proline (51.8-60.7%). The structural analysis demonstrates that they contain four general domains characteristic to proline-rich proteins: a signal peptide region, a transition region, a proline rich repeat region, and a carboxyl-terminal region. The last 30 amino acids encoded by two of the cDNAs have 100% identity with the isolated parotid hormone sequence, suggesting that the isolated parotid hormone was probably derived from the end of these cDNAs.
Northern analysis of porcine parotid tissue revealed five major mRNA size classes, ranging from 890 bases to 3550 bases, hybridized to the cDNA probe generated from one of the porcine cDNA clones, suggesting the existence of homologous sequences in these mRNAs, or that these transcripts were derived from closely related genes—a multigene family. *In situ* hybridization and immunohistochemical staining of porcine tissues indicated that the acinar cells of the parotid glands are the primary location for both parotid hormone related mRNAs and the translation products.

The expression of cloned parotid hormone containing cDNAs produced a single protein in the cell-free system and multiple products in transiently transfected COS-7 cells, as recognized by the antibodies made against the isolated parotid hormone, suggesting that post-translational and/or post-secretory modifications may have taken place.
A. Introduction

Dental caries is a multifactorial disease, and the body’s defense system is a complex one with numerous interacting factors. It is generally accepted that salivary exocrine secretion influences host resistance to caries attack, mainly by its rate of flow and by its content of fluoride. Increasing evidence, however, suggests that hormonal factor(s) from parotid glands may play an even more important role in determining the resistance or susceptibility to dental caries.

Parotid glands secrete a factor that is released into blood and stimulates dentinal fluid movement (DFM), which in turn, prevents inward diffusion of oral cytotoxic bacterial products and the cariogenic process. This factor, designated as parotid hormone (PH), has been isolated from porcine parotid glands and its partial amino acid sequence has been determined. Despite many supporting studies, the role of parotid hormone in protecting against dental caries cannot be directly proven because the parotid endocrine function cannot be easily separated from its exocrine functions. New experimental approaches are needed that allow direct and independent assessment of the relationship between parotid hormone, dentinal fluid movement, and dental caries.

This study aimed to isolate and identify the cDNA sequence of porcine parotid hormone based on its unique partial amino acid sequence. It was expected that following the cloning of the porcine gene for parotid hormone, more strategies could be developed and implemented to elucidate systemic mechanisms essential for maintaining dental health, particularly those related to the endocrine function of parotid glands.
The following sections in this chapter provide a review on related background information, starting with the diversity of salivary glands and its relationship with dental caries, focusing mainly on the evidence supporting research in this direction, and the objective and specific aims for this study.

B. Diversity of Salivary Glands

The salivary glands of mammals are a generally similar but specifically diverse group of "exocrine" organs. The major salivary glands, i.e., the parotid, submandibular, and sublingual glands, are large paired organs that lie outside the oral cavity proper, and which communicate with the oral mucosa through major ducts. Minor salivary glands are found in the submucosal of most regions of the oral cavity (See Hand, 1985; Klein, 1987, for general description of salivary glands). There are remarkable differences, with respect to both form and function, between the salivary glands of different species, as well as between different salivary glands of the same species.

The morphological diversity of salivary glands is well illustrated in several monographs (Tandler, 1978; Young & van Lennep, 1978; Pinkstaff, 1979). Examination of sections of salivary gland secretory endpieces stained with hematoxylin and eosin allows one to describe two types of secretory cells, serous and mucous. Parotid glands are branched acinous glands and their secretory portions are composed almost exclusively of serous cells. Submandibular and sublingual glands, on the other hands, are generally thought to be branched tubuloacinar glands, having mucous tubuloacini with associated serous demilunes (Young & van Lennep, 1978). The serous cell has been described as a cell that is primarily a protein secreting cell and contains little, if any,
mucin. When histochemical staining methods for glycoconjugates are used, however, it is found many serous cells contain appreciable concentrations of neural and/or acidic glycoconjugates, thus termed seromucous cells by Munger (1964).

Therefore, salivary gland secretory endpieces are said to consist of some combination of serous, mucous, or seromucous secretory cells. These terms, derived from histological and histochemical staining, serve to superficially categorize the endpiece cells in terms of their secretory products. However, examination of the ultrastructure of acinar cells in the parotid, submandibular, and sublingual salivary glands of hundreds of species of mammals suggests that acinar cells are apparently extremely diversified (Philips et al., 1993). The diversity is so great in the structure of acinar cells and their secretory products that the traditional histological and histochemical groupings are both overly simplistic and misleading. According to Philips et al. (1993), salivary gland epithelium consists of the most diversified group of cells in mammals. When salivary glands are compared across species, they obviously are more diversified than any other organ systems.

The content of mucous and serous cells varies among salivary glands. Within each of these major classes, further diversity in the expression of different secretory proteins is evident. In a given species, a particular secretory protein may be expressed in one specific gland, while other species may lack this protein or express it in a different gland (Sreebny et al., 1967; Tan & Teng, 1979; Dickinson et al., 1989). Most of the diversity in salivary proteins probably results directly from some type of posttranslational modification of gene products (Bennick, 1987; Mamula et al., 1988). As an example, in human beings the diversified proline-rich proteins (PRPs) are derived from a small
C. Dental Caries and Salivary Glands

As described in a classic diagram by Keyes and Jordan (1963), the triad of factors indispensable for caries activity consists of the microbiota (dental plaque), the diet (carbohydrates), and the host with (susceptible) teeth. Caries development requires the simultaneous presence of all three factors for a significant time.

Studies over the past 60 years have provided scientific support for the essential roles of microorganisms and diet in the development of dental caries (McLure and Hewitt, 1946; Orland et al., 1954; Kite et al., 1950). It is believed that dental caries results from demineralization brought about by frequent pH drops in plaque. The low pH generated by a high carbohydrate diet gives a selective advantage to the more aciduric microorganisms (Harper and Loesche, 1984; Bowden and Hamilton, 1987; Bradshaw et al., 1989), among which mutans streptococci (MS) have been identified as the major contributor to dental caries in humans (Loesche, 1986; Bowden and Edwardsson, 1994). However, not all individuals with poor oral hygiene and frequent sugar consumption develop caries, a finding that gave rise to the concept of more or less caries-resistant teeth.

In exploring the factors determining the resistance to caries, studies have led to inconclusive results. For more than two decades, saliva has been acknowledged as indispensable to the integrity of the teeth and mouth tissues (Mandel and Wotman, 1976; Mandel, 1984; Baum et al., 1985; Mandel, 1989). Chronic hyposalivation in adults leads to an increased risk of dental caries (Mandel, 1974; Baum et al., 1985; Fox et al., 1985; Vissink et al., 1988). Thus, most studies have focused on the exocrine secretion of salivary glands, in searching for the components in saliva that may have a potential role
in caries prevention. It is now generally accepted by dental professionals that salivary secretion influences resistance to individual caries attack, mainly by its rate of flow and its content of fluoride (reviewed by Lagerdof and Oliveby, 1994). Other caries-protective factors in saliva include its buffering capacity, its content of calcium, inorganic phosphate, and antimicrobial agents such as secretory immunoglobulins (sIgA), peroxidases, lactoferrin, lysozyme, and agglutinins (Xu and Oppenheim, 1993; Edgar et al., 1994; Lagerdof and Oliveby, 1994). For the most part, however, these factors have been inferred to be important in caries prevention and direct experimental proof for such a role in vivo is still lacking.

A new insight into the fundamental basis of resistance to caries has been provided by research in dentinal ultrastructure and internal dynamics. The odontoblastic cells at the pulp-dentin junction have cytoplasmic extensions stretching from the predentin area to the dentin-enamel junction through narrow canaliculi (Katchburian and Holt, 1968; Weinstock and Leblond, 1974; Hansson et al., 1975). Accordingly, dentin can be regarded as a barrier to bidirectional diffusive and convective transport between the mouth and the underlying pulp (Pashley, 1992 & 1994). Studies in various species indicate the existence of dentinal fluid movement that can occur outward from the pulp or inward, depending on experimental conditions (Steinman, 1968; Pashley et al, 1981; Potts et al, 1985; Maita et al, 1991; Vongsavan and Matthews, 1991 & 1992; Ciucchi et al., 1995). In exposed dentin, the inward diffusive passage of bacterial cytotoxic products is opposed by a continuous outward flow of dentinal fluid, which may represent a physiological means by which integrity of the avascular dental tissues is maintained despite the potential for ingressive diffusion of exogenous substances (Vongsavan and
Matthews, 1992; Pashley and Matthews, 1993; Matthews and Vongsavan, 1994; Gerzina and Hume, 1995). The resistance of teeth to decay, therefore, may involve maintenance of adequate endogenous centrifugal DFM.

The significance of DFM in relation to dental health has been suggested by caries studies in rats that showed an inverse relationship between the cariogenicity of certain dietary factors and pharmacological agents and the ability of the latter to stimulate DFM under acute administration (Steinman, 1968; Steinman and Leonora, 1975a, 1975b). It was observed that the cariogenicity of a high sucrose diet could be significantly inhibited (85-100%) with the addition of dietary supplements, such as carbamyl-DL-aspartic acid and carbamyl phosphate that stimulate dentinal fluid movement in intact rats. The supplements were ineffective in sialoadenectomized animals (Steinman and Leonora, 1975b & 1980), which suggested that these supplements exercise their cariostatic effect through mechanism(s) involving salivary glands. In desalivated rats, the highest incidence of decay occurred in intact teeth, and the least in endodontically-treated (pulpless teeth), suggesting that the pulp has an essential role in determining susceptibility to decay in response to desalivation. The increase in dental caries associated with sialoadenectomy is the result of a malfunctioning pulp, due to the elimination of endocrine factors secreted by salivary glands; while the absence of saliva associated with sialoadenectomy only has a minor role in the development of dental caries (Steinman et al., 1980).
D. Parotid Hormone and Proline-rich Proteins

The direct evidence for the existence of an endocrine system that regulates the intradentinal fluid movement was first reported in 1968 by Leonora and Steinman that the infusion of crude extracts of parotid glands or hypothalamic tissue stimulated intradentinal dye penetration in rats. The concept of a hypothalamus-parotid gland endocrine axis was proposed: the parotid glands secrete a parotid hormone, presumably under the control of a hypothalamic-releasing factor, that stimulates dentinal fluid movement (Leonora and Steinman, 1968). Supporting this hypothesis, the factor that stimulates DFM was isolated from porcine parotid glands (Tieche et al, 1980), and similarly active factors have been isolated from bovine (Yamamoto et al, 1986a), and rat (Yamamoto et al, 1986b) parotid gland extracts. This purified porcine parotid hormone was further used to make a polyclonal antibody that can be used to measure the circulating concentration of this hormone, as immunoreactive PH (iPH). In addition, a PH-releasing factor was partially purified from porcine thalamus-hypothalamus tissue, and this releasing factor stimulated DFM in rats or triggered the release of iPH in pigs only when infused into intact animals. Parotidectomy completely blunted this response, indicating that parotid gland is the primary source of the parotid hormone (Leonora et al., 1987; Tieche and Leonora, 1989).

Studies with carbamyl-DL-aspartic acid (CAA), a compound capable of stimulating dentinal fluid movement, have provided further evidence for the existence of a hypothalamus-parotid gland endocrine axis. It was demonstrated that infusion of plasma from CAA treated rats was effective in stimulating DFM in both intact and parotidectomized rats, whereas the plasma obtained from parotidectomized, CAA-treated
rats was effective only in intact rats but not in parotidectomized ones. Moreover, infusion of plasma from rats treated with CAA after an electrolytic lesion of the hypothalamus was ineffective, indicating that the DFM-stimulating activity of CAA involves the sequential release of a hypothalamic factor and a parotid factor. These two factors may represent parotid hormone from parotid glands, and the parotid hormone releasing factor from hypothalamus (Leonora et al., 1993b).

Like many other hormones, the release of parotid hormone is modified by certain factors. Physiologically, feeding is the only known stimulus for iPH secretion, and the quality of the food dictates the magnitude of the response (Leonora et al., 1987; Leonora et al., 1993a; Tieche and Leonora, 1995). Neither smell of food, nor the sight of food, nor the chewing of nonnutritive substance, crushed ice, is effective in initiating the secretion of parotid hormone (Leonora et al., 1987), indicating that different secretion mechanisms are involved in the exocrine salivary and endocrine functions of the parotid glands. A possible pathway leading to feeding-related secretion of parotid hormone has been postulated by Leonora et al (1987 & 1993b), which includes stimulation of various sensory receptors in the oropharyngeal cavity, afferent to the central nervous system where integration and cognitive processes take place, leading to efferent signals to the hypothalamus that trigger the secretion of a parotid hormone-releasing factor and ultimately the release of parotid hormone which, upon reaching the pulpal capillary-odontoblast space, would activate the passage of fluid from the circulation into the dentinal tubules. Parotid hormone could induce an increase in intracellular cAMP through an expected hormone-tissue receptor type of activation, a common second
message pathway for protein/peptide hormones. But systemically, the function of this axis and the mechanism through which PH acts remain to be elucidated.

It is very intriguing to point out the relationship between PH secretion and high sucrose diet. Fermentable carbohydrates, especially sucrose, have been indicated as the most important cause of dental caries by a variety of animal experiments and clinical observations (Burt and Ismail, 1987; Bradshaw et al., 1989; Scheie et al., 1984; Sgan-Cohen et al., 1988). It was suggested that high sucrose diet favors the growth of mutans streptococci by playing a crucial role in the formation of extracellular glucans, which make the mutans streptococci highly capable of attaching to the tooth surface and thus making it “more pathogenic” (Newbrun, 1967). Studies by Leonora and Tieche, however, have led to a different view that sucrose may also exert a caries inducing effect through systemic mechanism(s) involving parotid glands. It was found that high sucrose diet inhibits the basal secretion of parotid hormone in both rats (Leonora et al., 1992) and pigs (Tieche et al., 1994). Acute feeding of a high sucrose diet triggered a lower iPH response in pigs when compared to feeding of standard chow (Tieche and Leonora, 1995). The inhibition of DFM after ingesting a high-sucrose diet could be the result of depressed parotid hormone secretion, which suppresses the dentinal fluid movement from the interior of the tooth through the enamel, allowing the cariogenic process to proceed and caries to ensue.

The partial amino acid sequence of porcine parotid hormone indicates that it is a small glycoprotein with an estimated molecular weight of 8,100 dalton, rich in proline and glycine (Tieche et al., 1980). Parotid glands are known to secrete a number of proline rich proteins. Proline-rich proteins (PRPs) are a unique family of molecules,
which collectively comprise more than 70% of the secreted proteins in human saliva (Bennick, 1982). More than 20 PRPs have been described in humans, and they can be classified as acidic or basic on the basis of their charge, and some can be glycosylated or phosphorylated (Kauffman et al., 1991). These unusual proteins are synthesized constitutively in humans, but families of similar proteins are dramatically increased or induced in parotid and submandibular glands of rats, mice, and hamsters by isoproterenol treatment or by dietary tannins (Johnson, 1984; Ann et al., 1987; Mehansho et al., 1983; Mehansho et al., 1987a). For human PRPs, a small number of genes gives rise to a large number of secreted proteins, resulting from mRNA splicing, post-translational cleavages and modifications, as well as postsecretory cleavages (Maeda, 1985; Maeda et al, 1985; Kauffman et al., 1991). Sequences of human, rat and mouse PRPs have a high degree of similarity, and individual proteins from various species exhibit a large number of internal repeats, which probably arose by duplication of an ancestral gene (Ziener et al, 1984; Clements et al, 1985). The high conservation of the sequences and structures of PRP genes and PRPs argues for specific biological functions for these unusual proteins.

Several functions, such as calcium binding, hydroxyapatite binding, and formation of the dental acquired pellicle, have been attributed to the human acidic proline-rich proteins (Hay, 1973; Bennick, 1976, 1977, 1982). Acidic PRPs may play an important role in enamel homeostasis and inhibition of crystal growth of calcium phosphate on the tooth surface (Hay et al., 1979; Moreno et al., 1979). The binding sites for calcium and hydroxyapatite are located in the N-terminal proline-poor part of the protein, containing mainly negatively charged amino acids (Bennick, 1982; Moreno et al., 1982). In addition, hydroxyapatite-bound acidic proline-rich proteins mediate the
attachment of several important bacteria to the tooth surface (Gibbons and Hay, 1988), suggesting a possible role in the formation of dental plaque.

Proline rich proteins may also play a role in chelating dietary compounds thought to be harmful. Tannins are complex phenol compounds found in foods derived from plants such as barley, legumes, fruits and berries. Consumed in large amounts, tannins can cause carcinomas, liver disease, growth retardation, and other pathological problems (Mehansho et al., 1987b). Studies have shown that diets rich in tannins are capable of producing dramatic increases in parotid gland weight and a 12-fold increase in the amount of proline-rich proteins secreted from the parotid glands (Mehansho et al., 1983; Mehansho et al., 1985). The induction of PRP synthesis by dietary tannins has led to the proposal that PRPs may act as a first line of defense against the detrimental effects of the tannins (Mehansho et al., 1983; Warner and Azen, 1988).

The biological functions of basic proline-rich proteins, however, are less defined. Johnson et al. (1991) found a lower concentration of basic PRPs in caries-active subjects. There are reports of a decrease in basic proline rich proteins in conditions of hypothyroidism (Carbone et al., 1966; Johnson and Cortez, 1985), glucocorticoid treatment (Johnson et al. 1987), zinc deficiency (Johnson and Alvares, 1984), alloxan-induced diabetes (Anderson and Johnson, 1981), and Harvard caries-susceptible rats (Sweeney et al., 1962). Each of these circumstances has been associated with greater caries susceptibility (Carbone et al., 1966; Haldi et al., 1962; Muhler et al., 1956; Liu and Lin, 1969; Fang et al., 1980; Sweeney et al., 1962).

A comparison with other reported PRPs indicates that porcine parotid hormone is unique in both amino acid composition and sequence, and it relates more closely to basic
PRPs in amino acid composition. The dentinal fluid transport (DFT) stimulating factors from bovine (Yamamoto et al, 1986a) and rat (Yamamoto et al, 1986b) are quite different from the isolated porcine PH in terms of size, amino acid composition and sequence, suggesting that there might be more than one factor in parotid glands which possess the DFT stimulating activity.

Despite all the supporting evidence presented above, the role of parotid hormone in protecting against dental caries cannot be directly proven, because the effects of other critical factors such as salivary function and microflora cannot be excluded. In addition, for technical reasons, simultaneous measurement of iPH and DFM in the same animal remains to be developed. New experimental approaches that allow direct and independent assessment of the relationship between parotid hormone, dentinal fluid movement, and dental caries are needed.

The isolation of porcine parotid hormone and determination of its partial amino acid sequence have provided the basis for further characterization of parotid hormone and investigation of its function(s). One way of characterizing PH is to identify the cDNA sequence of porcine parotid hormone by screening a porcine parotid cDNA library with specific probe(s) corresponding to the isolated porcine PH. Once the cDNA of the porcine parotid hormone has been identified, it can be used to screen human, rat or mouse cDNA libraries for the counterpart gene. An expression system can be generated to express cloned PH cDNA to produce high quantities of this protein. The overexpressed proteins can then be used in *in vivo* and *in vitro* studies, and to generate monoclonal antibodies for further physiological studies. The isolation of the rat and/or mouse cDNA would permit design of transgenic animals to provide direct proof of the role of parotid
hormone in protection of dental caries. For example, null alleles can be created to specifically eliminate expression of the PH gene product in mice (Capecchi, 1989) to explore whether absence of PH has any influence on subsequent disease. Ultimately, subtle mutations could be introduced to ascertain in vivo which domains of PH are required for activity. Given the wide use of rats as a model to study dental caries, creation of PH defective rats offers great promise for the future. Furthermore, identification of human PH gene and understanding of its function(s) will provide opportunity for developing potential diagnostic and/or predictive caries tests, and new treatment for caries in humans.

Obviously, cloning the gene for porcine parotid hormone will provide means to facilitate further studies towards better understanding the natural systemic defensive mechanism against dental caries.

E. Objective and Specific Aims

The objective of current study was to isolate and identify the cDNA sequence for porcine parotid hormone that has been shown to stimulate dentinal fluid movement. The specific aims included:

1. Construction of a porcine parotid cDNA library and screening the library to identify the clone(s) containing the sequence for parotid hormone.

2. Characterization of the PH gene product(s) in terms of tissue distribution.

3. Expression of the cloned PH cDNA in E. coli, in mammalian cells, and in vitro using coupled transcription/translation system and verification of the expressed products by specific antibody assays.
A. Materials

1. Reagents

Dulbecco's Modified Eagle's Medium (DMEM), 1x Trypsin-EDTA solution (0.05% Trypsin, 0.53 mM EDTA-4Na in HBSS, without Ca, Mg, NaHCO₃) were purchased from Cellgro, Mediatech Inc. (Herndon, VA). Fetal bovine serum (FBS) was purchased from GibcoBRL, Life Technologies, Inc. (Gaithersburg, MD). Bacto-tryptone, Bacto-yeast extract, and Bacto-agar were purchased from Difco, Inc. (Livonia, MI).

Bovine serum albumin (BSA), Kanamycin, Ampicillin, aprotinin, Phenylmethylsulfonyl fluoride (PMSF), Triton X-100, Ficoll, dextran sulfate, polyvinylpyrrolidone, β-mercaptoethanol (β-ME), sodium dodecyl sulfate (SDS), and diethyl pyrocarbonate (DEPC) were purchased from Sigma Chemical Company (St. Louis, MO).

Acrylamide, bis-acrylamide, ammonium persulfate, phenol, choloroform, methanol, permont and paraformaldehyde (PFA) were purchased from Fisher Scientific (Tustin, CA). N,N,N',N'-tetramethylethylenediamine (TEMED), Coomassie Brilliant Blue, and Bradford Protein Determination Reagent were purchased from Bio-Rad Laboratories (Richmond, CA).

Radiolabeled [α-³²P] dCTP (3000 Ci/mmol) and [γ-³²P] ATP (7000 Ci/mmol) were purchased from Amersham Life Science (Piscataway, NJ). [α-³⁵S] UTP (800 Ci/mmol) was purchased from NEN Life Science Products, Inc. (Boston, MA).

Herring sperm DNA, Salmon sperm DNA, rNTPs, RNasin, yeast tRNA, dithiothreitol (DTT), Formamide, RNA molecular weight markers and FuGENE 6
Transfection Reagent were purchased from Boehringer Mannheim Corp. (Indianapolis, IN). DNA molecular weight standards were purchased from New England BioLabs, Inc. (Beverly, MA).

Paraffin wax (Paraplast tissue embedding medium) was purchased from Oxford Labware (St. Louis, MO). Hemo-De solvent and clearing agent (Histo-clear) was purchased from Fisher Scientific (Tustin, CA). Mineral spirit was purchased from Bortz Oil Company (Santa Fe Springs, CA). Ethanol was purchased from AAPER Alcohol & Chemical Co. (Shelbyville, KY). Isopropanol was purchased from EM Science (Gibbstown, NJ).

RNA Zol B was purchased from TEL-TEST, Inc. (Friendswood, TX). QuickHyb Hybridization Solution was purchased from Stratagen (La Jolla, CA).

Two polyclonal antibodies were obtained from Dr. Tieche, Loma Linda University. One antibody was made against the isolated porcine parotid hormone, designated as AbPH, and the other against its synthetic peptide (the first 11 amino acids of isolated PH), named AbPH11.

All other reagents were obtained from Sigma Chemical Company (St. Louis, MO), VWR (Brisbane, CA), J.T. Baker Chemical Co. (Phillipsburg, NJ), and Fisher Scientific (Tustin, CA).

2. Enzymes

All enzymes were purchased from Promega (Madison, WI), New England Biolabs, Inc. (Beverly, MA), Boehringer Mannheim Corp. (Indianapolis, IN), and Beckman Instruments, Inc. (Palo Alto, CA). Only the specific applications involved in
this study are described here. Some of the enzymes provided with commercial kits are not included.

a. Restriction Endonucleases: These restriction enzymes (Type II restriction endonucleases) cleave double-stranded DNA at specific sites known as recognition sequences that are usually four to eight nucleotides in length and display twofold symmetry. Some cleave both strands exactly at the axis of symmetry, generating fragments of DNA that carry blunt ends; others cleave each strand at similar locations on opposite sides of the axis of symmetry, creating fragments of DNA that carry protruding single-stranded termini. DNA restriction fragments with compatible cohesive termini or blunt ends can be joined by DNA ligases. The availability of restriction enzymes and DNA ligases has facilitated the development of recombinant DNA technology, the dominant approach for studying many basic biological processes.

b. Reverse Transcriptase: Avian myeloblastosis virus (AMV) reverse transcriptase was used to transcribe mRNA into complementary DNA in reverse transcription reactions (RT).

c. DNA Polymerase: The large fragment of E. coli DNA polymerase I, also known as Klenow fragment, was used to facilitate cloning of blunt ended DNA fragments by filling in the recessed 3' termini created by digestion of DNA with restriction enzymes or resection with Exonuclease III and S1 nuclease. It was also used for random labeling of cDNA probes. The *Taq* DNA polymerase, a thermostable DNA-dependent DNA polymerase, was
used to amplify specific sequences of DNA *in vitro* by polymerase chain reaction (PCR) assays.

d. T4 DNA Ligase: Bacteriophage T4 DNA ligase was used to join DNA molecules with compatible cohesive termini or blunt ends, occurring between 5’-phosphate and 3’-hydroxyl termini.

e. T4 Polynucleotide Kinase: Bacteriophage T4 polynucleotide kinase catalyzes the transfer of the γ-phosphate of ATP to a 5’ terminus of DNA or RNA. It was used to endlabel oligonucleotides for use as hybridization probes.

f. Alkaline Phosphatase: Calf intestinal alkaline phosphatase (CIP) was used to catalyze the removal of 5’-phosphate residues from DNA fragments to prevent self-ligation.

g. T7 RNA Polymerase: Bacteriophage T7 synthesizes DNA-dependent RNA polymerase that recognizes and initiates synthesis of RNA on double-stranded DNA templates that carry the T7 promoter. T7 RNA polymerase was used to synthesize single-stranded RNA for use as hybridization probes.

h. Exonuclease III: It catalyzes the stepwise removal of 5’ mononucleotides from the 3’-hydroxyl termini of double-stranded DNA. It was provided in the Erase-a-Base System (Promega) for generation of nested sets of deletions in the terminal sequences of double-stranded linear DNA fragments.

i. Nuclease S1: It was provided in the Erase-a-Base System for removal of single-stranded tails from DNA fragments to produce blunt ends.

j. RNase A: Ribonuclease A is an endoribonuclease that specifically attacks single-stranded RNA 3’ to pyrimidine residues and cleaves the phosphate
linkage to the adjacent nucleotide. It was used to remove unhybridized regions of RNA from DNA:RNA hybrids, and to degrade RNA in plasmid DNA preparations. DNase-free RNase A was made by dissolving pancreatic RNase A at a concentration of 10 mg/ml in 10 mM Tris-HCl, pH 7.5, 15 mM NaCl, heating at 100°C for 15 minutes and cooling slowly to room temperature. The solution was dispensed into aliquots and stored at –20°C.

k. DNase I: Deoxyribonuclease I is an endonuclease that hydrolyzes double-stranded or single-stranded DNA preferentially at sites adjacent to pyrimidine nucleotides. It was used to remove DNA during RNA probe preparation.

l. Proteinase K: It was used to degrade proteins in the preparation of nucleic acids.

3. Plasmid Vectors

a. pCR 2.1 (Invitrogen, size 3.9 kb): It carries a lac promoter for bacterial expression of the lacZα fragment for α-complementation (blue-white screening), and both Kanamycin (Kan) and Ampicillin (Amp) resistance genes for selection and maintenance in E. coli. It also contains T7 promoter and priming site for in vivo or in vitro transcription of sense RNA, and sequencing of insert. It was provided in Original TA Cloning Kit as a linearized form to facilitate direct cloning of PCR products.

b. pSPORT 1 (GibcoBRL, size 4.1 kb): It was provided in the SuperScript Plasmid System as the cloning vector for construction of cDNA libraries. It contains unique Sal I and Not I sites needed to clone the cDNA directionally. It has been engineered to include several useful features such as construction
of nested deletions, expression of cloned genes by inducing the \textit{lac} promoter with isopropylthio-beta-galactoside (IPTG), and \textit{in vitro} transcription from phage promoter SP6 and T7. It carries the gene for resistance to Ampicillin, useful for selection of transformants in \textit{E. coli}.

c. pET-28a(+) (Novagen, size 5.4 kb): The target genes are cloned in the vector under control of the bacteriophage T7 promoter, and expression is induced by providing a source of T7 RNA polymerase in the host cells, which was accomplished by inserting a T7 polymerase gene copy into the \textit{E. coli} genome. This vector carries the gene for Kanamycin resistance, which allows the selection of transformants in \textit{E. coli}. In addition, an N-terminal HisTag/thrombin/T7Tag configuration plus an optional C-terminal HisTag sequence are provided for easy detection and purification of expressed proteins.

d. pCEP 4 (Invitrogen, size 10.4 kb): It carries a cytomegalovirus (CMV) promoter and a simian virus 40 (SV40) Poly A signal for highly efficient expression of proteins encoded by the insert DNA in mammalian cells. This vector contains two antibiotic resistance genes, encoding hygromycin phosphotransferase (Hyg$^\circ$) (driven by Thymidine Kinase promoter) and $\beta$-lactamase (Amp$^\circ$) as selection markers for either mammalian or \textit{E. coli} cell culture.

4. \textit{E. coli} Strains

The \textit{E. coli} host strains used in this study are presented in Table 1.
Table 1. *E. coli* Strains Used As Cloning and Expression Host in This Study

<table>
<thead>
<tr>
<th>Strain (Supplier)</th>
<th>Genotype</th>
<th>Application</th>
<th>Antibiotic Resistance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH5α (GibcoBRL)</td>
<td>$F^\phi80\ lacZ\Delta M15\ (\Delta(lacZYA-argF))$ $U169\ deoR\ recA1\ endA1$ $hsdR17(r_{K^+}, m_{K^+})\ phoA\ supE44\chi$ $thi-1\ gyrA96\ relA1$</td>
<td>A recombination-deficient suppressing strain commonly used for plating and growth of plasmids</td>
<td>none</td>
</tr>
<tr>
<td>INVoF’ (Invitrogen)</td>
<td>$F^\phi80\ lacZ\Delta M15\ (\Delta(lacZYA-argF))$ $U169$</td>
<td>Ideal for propagating plasmids, allowing stable replication of high copy number plasmids</td>
<td>none</td>
</tr>
<tr>
<td>DH10B (GibcoBRL)</td>
<td>$F^\phi80\ lacZ\Delta M15\ (\Delta(lacZYA-argF))$ $U169$</td>
<td>Ideal for use in generating more representative cDNA and genomic libraries</td>
<td>none</td>
</tr>
<tr>
<td>NovaBlue (Novagen)</td>
<td>endA1 $hsdR17(r_{K^+}, m_{K^+})\ supE44$ $thi-1\ recA1\ gyrA96\ relA1$ $lac[F^{proA+B+}\ lac^{prZ\Delta M15\ ::\ Tn10}]$</td>
<td>Suitable for expression from <em>E. coli</em> promoters such as <em>lac</em>, <em>tac</em>, and <em>trp</em></td>
<td>Tetracycline (12.5 µg/ml)</td>
</tr>
<tr>
<td>NovaBlue (DE3) (Novagen)</td>
<td>endA1 $hsdR17(r_{K^+}, m_{K^+})\ supE44$ $thi-1\ recA1\ gyrA96\ relA1$ $lac[F^{proA+B+}\ lac^{prZ\Delta M15\ ::\ Tn10}]$ (DE3)</td>
<td>Carries the gene for T7 RNA polymerase under lacUV5 control, suitable for expression from T7 promoter</td>
<td>Tetracycline (12.5 µg/ml)</td>
</tr>
<tr>
<td>BL21 (Novagen)</td>
<td>$F^\phiompThsdS_{8}(r_{8}, m_{8})\ gal\ dcm$</td>
<td>Suitable for expression from <em>E. coli</em> promoters such as <em>lac</em>, <em>tac</em>, and <em>trp</em></td>
<td>none</td>
</tr>
<tr>
<td>BL21(DE3) (Novagen)</td>
<td>$F^\phiompThsdS_{8}(r_{8}, m_{8})\ gal\ dcm$ (DE3)</td>
<td>Suitable for expression from T7 promoter</td>
<td>none</td>
</tr>
</tbody>
</table>

* The appropriate drug to select for the target plasmid must also be added.
5. **Source of Animal tissue and Mammalian Cell Lines**

   a. **Pigs:** The tissue samples were obtained from Land Race York Hamshire castrated pigs from Dr. Tieche’s laboratory at Loma Linda University.

   b. **Rats:** The tissue samples were obtained from Sprague-Dawley male rats from Dr. Leonora at Loma Linda University.

   c. **COS-7 cells:** The COS-7 cells are derived from kidney cells of the African green monkey, immortalized by transformation with SV40. The COS-7 cells were purchased from American Type Culture Collection (ATCC, CRL1651) (Manassas, VA).

6. **Primers for PCR and DNA Sequencing**

   All primers, except those indicated, were synthesized or provided by the Core Facility at the Center for Gene Therapy, Loma Linda University.

   a. **Oligo A:** 5’ GCNCCNCCNNGGNGCNMGNCC 3’

   b. **Oligo B:** 5’ CCYTGNGGYTCNGGNCCNGC 3’

   c. **Oligo C:** 5’ CCNGGYTGNGGNCKNGGYTG 3’

   d. **Oligo D:** 5’ GGTYGNCCNGCNGGYTGNGG 3’

   e. **PHiA:** 5’ GCICCICICIGGICIMGICC 3’ (Integrated DNA Technologies, Inc. Coralville, IA)

   f. **PHiD:** 5’ GGICCIGGIGIGGICKYTG 3’

   g. **PH5P:** 5’ GTCCCAGCCCACCAAGACC 3’ (Sigma-Genosys, The Woodlands, TX)

   h. **T7 Promoter Primer:** 5’ TAATACGACTCCTATAGGG 3’

   i. **M13 Sequencing Primer:** 5’ CCCAGTCACGACGTTCCTAAAACG 3’
7. Kits and Systems

Major commercial kits and systems used in this study are listed below. Usages are described or outlined in the method sections. For further details, see the manufacturer’s instructions.

a. Original TA Cloning Kit (Invitrogen, Carlsbad, CA)
b. SuperScript Plasmid System for cDNA Synthesis and Plasmid Cloning (GibcoBRL Life Technologies, Inc.)
c. Erase-a-Base System (Promega, Madison, WI)
d. TNT Quick Coupled Transcription/Translation Systems (Promega)
e. Transcend Non-Radioactive Translation Detection Systems (Promega)
f. Opti-4CN Detection Kit, Goat-anti-Rabbit (BioRad)
g. QIAfilter Plasmid Midi Kit (Qiagen, Valencia, CA)
h. QIAquick Gel Extraction Kit (Qiagen)
i. Oligotex mRNA Kit (Qiagen)
j. Megaprime DNA Labeling System (Amersham)
k. DAKO LSAB+ Kit Peroxidase: Labeled Streptavidin Biotin Method for specific staining of primary mouse, rabbit or goat antibody (DAKO Corporation, Carpinteria, CA)
l. AEC Supersensitive Detection Kit (BioGenex, San Ramon, CA)

8. Supplies

Duralon-UV and Duralose-UV membranes were from Stratagen (La Jolla, CA). NitroBind nitrocellulose transfer membrane (0.45 micron) was purchased from MSI (Westboro, MA). MicroSpin G-10 columns were from Pharmacia Biotech (Piscataway,
NJ). Micro Slides (precleaned, superfrost plus, 25x75x1 mm) and Micro cover glasses were purchased from VWR (West Chester, PA). Kodak BioMax X-ray films were purchased from Fisher Scientific (Tustin, CA).

9. Solutions

The recipes for solutions containing more than one substance are listed as following. More recipes can be found in the method sections.

a. Luria-Bertani (LB) Medium (per Liter): 10 g of NaCl, 10 g of tryptone, 5 g of yeast extract, pH adjusted to 7.0; autoclaved for 20 minutes at 15 lb/sq.in.

b. LB Agar (per Liter): 10 g of NaCl, 10 g of tryptone, 5 g of yeast extract, 20 g of agar, pH adjusted to 7.0; autoclaved, and poured into petri dishes (~25 ml/100-mm plate).

c. SOC Medium (per Liter): 10 g of tryptone, 5 g of yeast extract, 5 g of NaCl, 10 mM MgSO$_4$, 10 mM MgCl$_2$, pH adjusted to 7.0; autoclaved, and filter-sterilized glucose added to a final concentration of 20 mM.

d. TE buffer: 10 mM Tris-HCl pH 8.0, 1 mM EDTA.

e. TEN buffer: 10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 25 mM NaCl.

f. TE-saturated phenol:chloroform:isoamyl alcohol (25:24:1): Mix equal parts of TE buffer and phenol and allow the phases to separate. Then mix 1 part of the lower, phenol phase with 1 part of chloroform:isoamyl alcohol (24:1).


h. Phosphate-buffered saline (PBS) (per Liter): 8 g of NaCl, 0.2 g of KCl, 1.44 g of Na$_2$HPO$_4$, and 0.24 g of KH$_2$PO$_4$, pH adjusted to 7.4.
i. 1x PBST: 1x Phosphate-buffered saline, 0.1% Tween 20.

j. Tris-buffered saline (TBS): 20 mM Tris, 500 mM NaCl, pH 7.5.

k. 1x TBST: 1x Tris-buffered saline, 0.1% Tween 20.

l. Extraction buffer (for genomic DNA extraction): 10 mM Tris-HCl pH 8.0, 0.1 M EDTA pH 8.0, 20 µg/ml pancreatic RNAase, 0.5% SDS.

m. 20x SSC (per liter): 175.3 g of NaCl, 88.2 g of sodium citrate, pH adjusted to 7.0.

n. 50x Denhardt’s Reagent (per 500 ml): 5 g of Ficoll, 5 g of polyvinylpyrrolidone, 5 g of bovine serum albumin (BSA); filtered through a disposable filter.

o. Hybridization mix (for in situ hybridization): 50% Formamide, 0.3 M NaCl, 20 mM Tris-HCl pH 8.0, 5 mM EDTA, 10% dextran sulfate, 0.02% Ficoll, 0.02% BSA, 0.02% polyvinylpyrrolidone, 0.5 mg/ml yeast RNA.

p. 2x SDS gel-loading buffer: 100 mM Tris-HCl pH 6.8, 200 mM dithiothreitol (DTT), 4% SDS, 0.2% bromophenol blue, 20% glycerol.

q. Tris-glycine electrophoresis buffer (pH 8.3): 25 mM Tris base, 250 mM glycine, 0.1% SDS.

r. Transfer buffer (pH 8.3) (for Western blotting): 39 mM glycine, 48 mM Tris base, 0.037% SDS, 20% methanol.

s. Blocking buffer (for Western blotting): 1% BSA in 1x PBS.

t. Antibody dilution buffer (for Western blotting): 1% BSA in 1x PBST.

u. 0.5x TBE: 0.045 M Tris, 0.045 M boric acid, 1 mM EDTA (pH 8.0).
v. 10x MOPS: 0.2 M 3-[N-morpholino]propanesulfonic acid (MOPS), 0.05 M NaOAc trihydrate, and 5 mM EDTA, pH adjusted to 7.0.

w. Sample loading dyes (for Northern blotting): 0.72 ml of deionized formamide, 0.32 ml of 5x MOPS, 0.26 ml of 37% formaldehyde, 0.1 ml of 80% glycerol, 4 mg of bromophenol blue, 4 mg of xylene cyanole, and 0.1 ml of DEPC-treated sterile water; mixed thoroughly and stored at −20°C.

x. S1 nuclease mix (for Erase-a-Base system) (for 25 time points): 172 µl of deionized water, 27 µl of S1 7.4x buffer, 60 units of S1 Nuclease.

y. Klenow mix (for Erase-a-Base system): 30 µl of Klenow buffer, 3-5 units of Klenow DNA polymerase.

z. Ligase mix (for Erase-a-Base system) (for 25 ligations): 790 µl of deionized water, 100 µl of Ligase 10x buffer, 100 µl of 50% polyethylene glycol (PEG), 10 µl of 100 mM DTT, 5 units of T4 DNA Ligase.

aa. Coomassie Blue staining solution (per 100 ml): 0.25g of Coomassie Brilliant Blue R250, 90 ml of methanol:H₂O (1:1 v/v), 10 ml of glacial acetic acid.

bb. Destaining solution (per 200 ml): 90 ml of methanol, 90 ml of deionized water, 20 ml of glacial acetic acid.

cc. Single-detergent lysis buffer: 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.02% sodium azide, 100 µg/ml phenylmethylsulfonyl fluoride (PMSF), 1 µg/ml aprotinin, 1% Triton X-100. PMSF should be added to the lysis buffer just before use.

dd. 4% Paraformaldehyde (PFA) (per 500 ml): in 300 ml of hot deionized water (60-70°C), add 1 ml of 2 N NaOH, 20 g of paraformaldehyde powder, 2 ml of
1 N HCl, 50 ml of 10x PBS; adjust pH to 7.4, then add deionized water to 500 ml, filter through a disposable filter on ice, and store at 4°C.

ee. Solution I (for plasmid miniprep): 25 mM Tris-HCl pH 8.0, 10 mM EDTA, 50 mM glucose.

ff. Solution II (for plasmid miniprep): 0.1 N NaOH, 1% SDS.

gg. Solution III (for plasmid miniprep) (per 100 ml): 60 ml of 5 M potassium acetate, 11.5 ml of glacial acetic acid, 28.5 ml of H2O.

hh. Cracking 2x buffer (per 50 ml): 2 ml of 5 M NaOH, 2.5 ml of 10% SDS, 10 g of sucrose.

ii. Diethylpyrocarbonate (DEPC)-treated water: 0.1% diethyl pyrocarbonate in water, mixed thoroughly, stood for at least 12 hours before autoclaved for 15 minutes. DEPC is a strong inhibitor of RNAses.

10. Gels

a. 0.7% agarose gel (50 ml): 0.35 g of agarose dissolved in 50 ml of 0.5x TBE.
   Running buffer: 0.5x TBE.

b. 1% agarose gel (50 ml): 0.5 g of agarose dissolved in 50 ml of 0.5x TBE.
   Running buffer: 0.5x TBE.

c. 2% agarose gel (50 ml): 1 g of agarose dissolved in 50 ml of 0.5x TBE.
   Running buffer: 0.5x TBE.

d. 1.5% agarose formaldehyde gel (50 ml): 0.75 g of agarose dissolve in 35 ml of DEPC-treated H2O, mix with 5 ml of 10x MOPS, 8.9 ml of 37% of formaldehyde. Running buffer: 1x MOPS.
e. 6% polyacrylamide gel (10 ml): mix 2 ml of 30% acrylamide mix
(acrylamide:bisacrylamide 29:1), 5.93 ml of water, 2 ml of 5x TBE, 0.07 ml
of 10% ammonium persulfate, and 8 µl of TEMED. Running buffer: 1x TBE.

f. 8% polyacrylamide gel (10 ml): mix 2.66 ml of 30% acrylamide mix, 5.27 ml
of water, 2 ml of 5x TBE, 0.07 ml of 10% ammonium persulfate, and 6 µl of
TEMED. Running buffer: 1x TBE.

g. 12% polyacrylamide resolving gel (10 ml): mix 4 ml of 30% acrylamide mix,
3.3 ml of water, 2.5 ml of 1.5 M Tris (pH 8.8), 0.1 ml of 10% SDS, 0.1 ml of
10% ammonium persulfate, and 4 µl of TEMED. Running buffer: Tris-
glycine electrophoresis buffer.

h. 5% polyacrylamide stacking gel (5 ml): mix 0.83 ml of 30% acrylamide mix,
3.4 ml of water, 0.63 ml of 1 M Tris (pH 6.8), 0.05 ml of 10% SDS, 0.05 ml
of 10% ammonium persulfate, and 5 µl of TEMED. Running buffer: Tris-
glycine electrophoresis buffer.

11. Computer Software and Databases

a. PC/Gene: the nucleic acid and protein sequence analysis software system
(A.Bairoch, University of Geneva, Switzerland; IntelliGenitis Inc., release
6.85, 1995).

b. Omega 2.0 (Oxford Molecular Ltd., Genetics Computer Group, An Oxford
Molecular Company).

c. BLAST program under www.ncbi.nlm.nih.gov
B. Methods

1. **Molecular Cloning of Porcine Parotid Hormone**

   a. **Tissue Collection**

   Porcine parotid glands were obtained from pigs under anesthesia. Other tissues, including submandibular glands, sublingual glands, pancreas, liver, small intestine, kidney, trachea, lung, skeletal muscle, hypothalamus and pituitary gland, were collected from pigs immediately after euthanized. The tissue samples were excised and frozen in liquid nitrogen immediately, and later transferred to -70°C for storage. Some of the tissues were directly fixed in 4% paraformaldehyde in PBS, for analysis by *in situ* hybridization.

   Parotid, submandibular, and sublingual glands were removed from rats under anesthesia. Human parotid gland tissue samples were obtained from the surgery department at Jerry L. Pettis Memorial Veterans Medical Center, Loma Linda, CA. In either case, the tissue samples were frozen in liquid nitrogen immediately, followed with storage at −70°C until processed.

   b. **TA Cloning of RT-PCR Products**

   The Original TA Cloning Kit from Invitrogen was used to clone RT-PCR products. It is designed for cloning PCR products directly from a PCR reaction without the need for purification or restriction digestion. The cloning vector pCR 2.1 contains the *lacZα* complementation fragment for easy blue-white color screening.
The reverse transcriptase (RT) was used to transcribe mRNA to complementary DNA, and then the first strand cDNA was used as a template in polymerase chain reaction (PCR). RT-PCR is commonly used for generation of probes specific for genes or cDNAs that have not been cloned.

1). Isolation of Total RNA

The total RNA was isolated with RNAzol B solution. This method is based on the unique property of RNAzol that promotes formation of complexes of RNA with guanidinium and water molecules, and abolishes hydrophilic interactions of DNA and proteins. Therefore, DNA and proteins are efficiently removed from the aqueous phase while RNA remains in this phase during the sample extraction with RNAzol.

The tissue samples were homogenized in RNAzol B solution (2ml per 100 mg tissue) with a few strokes in a glass-Teflon homogenizer on ice. About 0.2 ml of chloroform was added per 2 ml of homogenate, and the samples were shaken vigorously for 15 seconds and placed on ice for 5 minutes. After the suspension was centrifuged at 12,000g at 4°C for 15 minutes, the upper aqueous phase was transferred to fresh tubes, and an equal volume of isopropanol was added. The samples were stored for 15 minutes at 4°C, then centrifuged for 15 minutes at 12,000g (4°C). The supernatant was removed and the RNA pellet was washed with 75% ethanol by vortexing and subsequent centrifugation for 8 minutes at 7,500g (4°C). The pellet was dried under vacuum for 10 minutes, and dissolved in DEPC-treated water. The concentration of RNA was determined by OD$_{260}$. The total RNA samples were stored at $-70^\circ$C until used.
2). Designing Primers

The four fragments of the porcine parotid hormone for which the amino acid sequence has been determined are shown in Figure 1. Since the molecular weight of PH estimated by SDS-PAGE was about 8.1 kDa, the number of amino acids was assumed to be approximately 90. A total of 30 amino acids were determined for the native PH coupled to the activated porous glass beads (PH-APG). This sequence was thought to be partial, therefore was designated as N-terminal fragment, or fragment A. The sequences obtained from the digests of the native PH were named fragment B (pronase digest), fragment C (pepsin digest) and fragment D (pronase digest of PH-APG), all of which were speculated to be downstream to the fragment A, although the alignment of these fragments was not clearly determined.

Four degenerate primers were designed and synthesized according to those amino acids sequences. Because the degeneracy of the genetic code, all related oligonucleotides were synthesized in a comprehensive pool to represent all possible ways to code for a given short sequence of amino acids. The sequence of each primer was selected based on the criteria that 1) each oligonucleotide is 20 nucleotides in length, 2) the selected amino acid sequence is relatively unique for that fragment and unique in comparing to other known PRPs, 3) the degeneracy of those selected amino acids is relatively low. Oligo A (5’ primer) and Oligo B, C, D (3’ primers) were designed corresponding to a stretch of amino acid sequence for each fragment. To decrease the degeneracy, another two primers corresponding to part of the amino acid sequences in fragment A and fragment D were synthesized, designated as PHiA and PHiD, with deoxyinosine incorporated into the position where the degeneracy is higher than 2 (Figure 1).
Figure 1. Design of oligonucleotide primers based on the available partial amino acid sequences of four peptides from the porcine parotid hormone. The amino acid sequences from which the primers were designed are underlined. Nucleotide codes: M=A or C; Y=C or T; R=A or G; K=G or T; N=A or T or C or G; I=deoxyinosine
<table>
<thead>
<tr>
<th>Fragment A</th>
<th>NH$_2$-APPGARPPPGPPPPGPSPRPPPPGPPQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA sequence</td>
<td>Ala Pro Pro Gly Ala Arg Pro</td>
</tr>
<tr>
<td>mRNA coding</td>
<td>GCN CCN CCN GGN GCN MGN CCN</td>
</tr>
<tr>
<td>DNA coding</td>
<td>CGN GGN GGN CCN GGN KCN GGN</td>
</tr>
<tr>
<td>Oligo A</td>
<td>5’ GCNCCNCNGGNCMNCCGCCG 3’</td>
</tr>
<tr>
<td>PHiA</td>
<td>5’ GCICICICICICICICICICCCG 3’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fragment B</th>
<th>APPGAGPEPQGPPPSG(E/P)RPQE(P/R)PQGP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA sequence</td>
<td>Ala Gly Pro Glu Pro Gln Gly</td>
</tr>
<tr>
<td>mRNA coding</td>
<td>GCN GGN CCN GAR CCN CAR GGN</td>
</tr>
<tr>
<td>DNA coding</td>
<td>CGN CCN GGN CTY GGN GTY CCN</td>
</tr>
<tr>
<td>Oligo B</td>
<td>5’ CCYTNGGTYTNGGNCNCCGC 3’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fragment C</th>
<th>APPGARG(E/P)(P/D)GPPPPPGPPSQPRQPQGPPQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA sequence</td>
<td>Gln Pro Arg Pro Gln Pro Gly</td>
</tr>
<tr>
<td>mRNA coding</td>
<td>CAR CCN MGN CCN CAR CCN GGN</td>
</tr>
<tr>
<td>DNA sequence</td>
<td>GTY GGN KCN GGN GTY GGN CCN</td>
</tr>
<tr>
<td>Oligo C</td>
<td>5’ CCNNGYTNGGNCNKNGGTYTNG 3’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fragment D</th>
<th>APQPAGQPPPQGPPSQRPPPQGPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA sequence</td>
<td>Pro Gln Pro Ala Gly Gln Pro</td>
</tr>
<tr>
<td>mRNA coding</td>
<td>CCN CAR CCN GGN GGN CAR CCN</td>
</tr>
<tr>
<td>DNA coding</td>
<td>GGN GTY GGN CCN GTY GGN</td>
</tr>
<tr>
<td>Oligo D</td>
<td>5’ GGYTGCCNGGNCNGGTYTNGG 3’</td>
</tr>
</tbody>
</table>

| AA sequence        | Gln Arg Pro Pro Pro Gly Pro    |
| mRNA coding        | CAR MGN CCN CCN GGN CCN       |
| DNA coding         | GTY KCN GGN GGN CCN GGN CCN   |
| PHiD               | 5’ GGICICICICICICICICICICICKYTG 3’ |
3). RT-PCR

The Oligo B, C, D and PHiD were used in the separate reactions as a 3' primer to generate the first strand of cDNA from parotid mRNA using AMV reverse transcriptase. The reverse transcription was set up as follows: 1 µl of total RNA (1 µg/µl) from pig parotid glands was mixed with 8 µl of DEPC-treated water and 1 µl of reverse primer (Oligo B, C, D or PHiD, at 20 pmole/µl). The tubes were first heated in a 70°C water bath for 5 minutes, then cooled down slowly to 30°C, and put on ice for 5 minutes. Six µl of DEPC-treated water, 2 µl of 10x PCR buffer, 0.8 µl of 25 mM MgCl₂, and 0.2 µl of 100 mM dNTPs were added, and the mixture was incubated at 42°C for 2 minutes. One µl of AMV reverse transcriptase (40 units/µl) was added and mixed by gentle tapping, followed by incubation at 42°C for 15 minutes.

The mRNA:cDNA hybrids were then subjected to PCR in a DNA thermal cycler using Oligo A or PHiA as 5' primer in conjunction with each of the other four 3' primers as presented in the Table 2. Each PCR was set up containing: 40.5 µl sterile deionized water, 1 µl of RT product, 5 µl of 10x PCR buffer, 0.5 µl of 100 mM dNTPs, 1 µl of 5' forward primer (Oligo A or PHiA, 20 pmole/µl), 1 µl of 3' reverse primer (Oligo B, C, D, or PHiD, 20 pmole/µl), and 1 µl of Taq DNA polymerase (5 units/µl). All PCR amplifications were carried out using either the “touch down” procedure or the “stringent” procedure as listed in Table 3.

4). Cloning and Screening of Target Genes

The PCR products were ligated with the pCR 2.1 vectors and then transformed into the INVaF’ competent cells according to manufacturer’s instructions. White colonies and some of the white with blue center or light blue colonies were picked for
Table 2. Utilization of Oligonucleotide Primers in RT-PCR Cloning of Porcine Parotid Hormone cDNA

<table>
<thead>
<tr>
<th>Reaction</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT primer</td>
<td>Oligo B</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oligo C</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oligo D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHiD</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR 5’primer</td>
<td>Oligo A</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>PHiA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>PCR 3’primer</td>
<td>Oligo B</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oligo C</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oligo D</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHiD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. PCR Programs for RT-PCR Cloning of Porcine Parotid Hormone cDNA

<table>
<thead>
<tr>
<th>Program</th>
<th>Step</th>
<th>Temperature</th>
<th>Length</th>
<th>Cycle number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>94 °C</td>
<td>2 min</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>94 °C</td>
<td>5 sec</td>
<td></td>
</tr>
<tr>
<td>Touch Down</td>
<td>2</td>
<td>66 °C, -1 °C each cycle down</td>
<td>30 sec</td>
<td>20</td>
</tr>
<tr>
<td>Program</td>
<td></td>
<td>72 °C</td>
<td>1 min</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>46 °C</td>
<td>5 sec</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72 °C</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>72 °C</td>
<td>1 min</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>94 °C</td>
<td>2 min</td>
<td>1</td>
</tr>
<tr>
<td>Stringent</td>
<td>2</td>
<td>94 °C</td>
<td>1 min</td>
<td>30</td>
</tr>
<tr>
<td>Program</td>
<td></td>
<td>67 °C</td>
<td>1 min</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>72 °C</td>
<td>1 min</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72 °C</td>
<td>7 min</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. Outline of the SuperScript Plasmid System procedure
screening. The cDNA inserts were released with *EcoR I* restriction, and the size of inserts was determined after separation of the digest in 8% polyacrylamide gels. The cDNA inserts of representative clones were sequenced with an automated DNA sequencer. The sequences of both strands for each cloned DNA were translated in three reading frames and the deduced amino acid sequences were compared for homology with the partial amino acid sequence of the isolated porcine parotid hormone.

c. Construction of cDNA Library

The porcine parotid cDNA library was constructed using SuperScript Plasmid System from GibcoBRL according to manufacturer’s instructions with few modifications. A schematic outline of procedures is presented in Figure 2.

1). Isolation of Messenger RNA

The poly A⁺ messenger RNA was purified from total RNA with Oligotex Spin Column Protocol from QIAGEN. The Oligotex procedure for isolation and purification of poly A⁺ mRNA takes advantage of the fact that most eukaryotic mRNAs carry a homopolymer of 20-250 adenosine nucleotides, known as the poly A tail. In contrast, RNA such as ribosomal RNA (rRNA) and transfer RNA (tRNA) are not polyadenylated. Therefore, separation of poly A⁺ RNA from rRNA and tRNA can be achieved by hybridizing the polyadenylated tails of mRNA molecules to oligo dT primers, which are coupled to latex particles (Oligotex). The poly A⁺ mRNA can subsequently be released by lowering the ionic strength and destabilizing the dT:A hybrids.

Total RNA in DEPC-treated water was mixed with 2x Binding Buffer and Oligotex Suspension according to manufacturer’s instruction. The mixture was first
mRNA

First strand synthesis

Second strand synthesis

Sal I adapter addition

Not I digestion
Size fractionation

Ligation to Plasmid pSPORT1, Not I-Sal I-Cut

cDNA ready for transformation
incubated for 3 minutes at 65°C to disrupt secondary structure in the RNA, and then for 10 minutes at room temperature to allow hybridization between the oligo dT<sub>30</sub> and the poly A tail of the mRNA. The tube was then centrifuged for 2 minutes at 14,000g and the supernatant was discarded. The Oligotex resin was resuspended in Wash Buffer and transferred onto a spin column, and the spin column was centrifuged at 14,000g for 30 seconds. The washing step was repeated once, and the flow-through was discarded. Preheated (70°C) Elution Buffer, 20 µl to 40 µl, was added onto the spin column, and the mRNA fraction was eluted from the resin by centrifugation at 14,000g for 1 minute. The elution step was repeated once to ensure maximum yield. The concentration of mRNA was measured by OD<sub>260</sub>, and the purity was estimated by A<sub>260/A<sub>280</sub> ratio. The mRNA was stored in elution buffer at -70°C.

2). First Strand Synthesis

The SuperScript Plasmid System features a Not I primer-adapter designed to prime first strand cDNA synthesis at the poly A tail of mRNAs. This primer-adapter contains 15 dT residues and 4 restriction endonuclease sites chosen for their relative rarity in mammalian genomes, including restriction site for Not I, a 8-base recognition sequence occurring approximately once in 10<sup>6</sup> bp.

SuperScript RT II is the reverse transcriptase supplied in the system for first stand cDNA synthesis reaction. Like other reverse transcriptases, this enzyme can polymerize deoxynucleoside triphosphates into a complementary DNA strand using an RNA molecule as template. Unlike others, however, this enzyme does not contain any RNase H activity, which is detrimental to the first strand cDNA synthesis, but exhibits significantly improved polymerase activity.
About 3 µg of mRNA was diluted with DEPC-treated water into a volume of 7 µl, and 2 µl of Not I primer-adapter was added. The mixture was heated to 70°C for 10 minutes, and quick-chilled on ice. The reaction mixture was centrifuged to the bottom of the tube, and the following components were added: 4 µl of 5x first strand buffer, 2 µl of 0.1 M DTT, 1 µl of 10 mM dNTP mix, and 1 µl of [α-32P] dCTP (10 µCi/µl). The reaction mixture was vortexed gently and placed at 37°C for 2 minutes. In the next step, 3 µl of SuperScript RT II was added into the tube, mixed, and incubated at 37°C for 1 hour.

The tube was then placed on ice to terminate the reaction. An aliquot of 2 µl was removed from the reaction mixture to calculate first strand yield. The remaining 18 µl of the first strand reaction was used to continue immediately the second strand reaction.

3). Second Strand Synthesis

To the first strand reaction containing tube on ice, the following reagents were added in the order shown: 93 µl of DEPC-treated water, 30 µl of 5x second strand buffer, 3 µl of 10 mM dNTP mix, 1 µl of E. coli DNA ligase (10 units/µl), 4 µl of E. coli DNA polymerase I (10 units/µl), and 1 µl of E. coli RNase H (2 units/µl). The tube was vortexed gently and incubated for 2 hours at 16°C. After incubation, 2 µl of T4 DNA polymerase (10 units) was added, and the mixture was placed at 16°C for 5 minutes.

The reaction was terminated on ice by the addition of 10 µl of 0.5 M EDTA. The mixture was treated with 150 µl of phenol:chloroform:isoamyl alcohol (25:24:1), vortexed thoroughly, and centrifuged at room temperature for 5 minutes at 14,000g to separate the phases. The upper aqueous layer of approximately 140 µl was removed carefully and transferred to a fresh 1.5-ml microcentrifuge tube. To precipitate DNA
molecules, 70 µl of 7.5 M NH₄OAc and 0.5 ml of ethanol (-20°C) were added into the mixture, vortexed thoroughly, and centrifuged immediately at room temperature for 20 minutes at 14,000g. The pellet was washed with 0.5 ml of 70% ethanol (-20°C), centrifuged for 2 minutes and the supernatant was removed. The pelleted cDNA was then dried at 37°C for 10 minutes to evaporate residual ethanol.

4). Addition of Sal I adapters

The 4-base, 5' extension of Sal I adapters was ligated to the blunt-ended cDNA products to introduce asymmetry and maximize the ligation efficiency of cDNA to the linearized vector DNA. The reaction was set up on ice by adding the following reagents in the order shown: 25 µl of DEPC-treated water, 10 µl of 5x T4 DNA ligase buffer, 10 µl of Sal I adapters, and 5 µl of T4 DNA ligase (1 unit/µl). The reagents were mixed gently, and incubated at 16°C for a minimum of 16 hours.

After incubation, 50 µl of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the reaction and mixed thoroughly, followed by centrifugation at room temperature for 5 minutes at 14,000g. The upper aqueous layer, about 45 µl, was carefully removed and transferred into a fresh 1.5-ml microcentrifuge tube. In the next step, 25 µl of 7.5 M NH₄OAc and 150 µl of absolute ethanol (-20°C) were added, vortexed thoroughly, and centrifuged at room temperature for 20 minutes at 14,000g. The pelleted DNA was washed with 0.5 ml of 70% ethanol (-20°C), centrifuged for 2 minutes to remove the supernatant, and dried at 37°C for 10 minutes.

5). Digestion of cDNA with Not I

Restriction with Not I before ligation of cDNA to the vector exposes the asymmetric element needed to obtain directional clones. The following reagents were
added on ice to the cDNA pellet from last step: 41 µl of DEPC-treated water, 5 µl of REact 3 buffer, and 4 µl of Not I (15 units/µl). The reaction was mixed gently and incubated for 2 hours at 37°C.

The phenol extraction and ethanol precipitation were carried out as in the last step. The cDNA pellet was dried at 37°C for 10 minutes, resuspended and subjected to column chromatography.

6). Size Fractionation of cDNA

Column chromatography was used to separate cDNA from residual adapters, Not I fragment released by restriction digestion, and other low molecular weight DNAs.

The cDNA pellet was resuspended on ice in 100 µl TEN buffer (10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, and 25 mM NaCl), while the 1-ml prepacked Sephacryl S-500 HR chromatography column was equilibrated with TEN buffer by loading 0.8 ml of TEN buffer each time and letting it drain completely for 4 times. Sterile microcentrifuge tubes were labeled from 1 to 20 and used to collect the effluent. The entire cDNA sample was loaded to the center of the top frit and the effluent was collected into tube 1. Then, 100 µl of TEN buffer was added to the column, and the effluent was collected into tube 2. In the next step, another 100 µl of TEN buffer was added, followed with more as needed. The single-drop fractions were collected into individual tubes until a total of 18 drops was collected into tube 3 through 20.

The volume in each tube was measured, and the Cerenkow count for each fraction was obtained. The amount of cDNA in each fraction was calculated and the earliest fraction containing >10 ng of cDNA at >0.7 ng/µl was used in the vector ligation reaction.
7). Ligation of cDNA to the Vector DNA

The cloning vector included in the SuperScript Plasmid System is plasmid pSPORT 1, which contains unique Sal I and Not I sites needed to clone the cDNA directionally. Plasmid pSPORT 1 is supplied predigested with Sal I and Not I and has been carefully prepared to provide a low background of nonrecombinant colonies.

The following components were mixed in a sterile 1.5-ml microcentrifuge tube at room temperature: 4 µl of 5x T4 DNA ligase buffer, 1 µl of pSPORT 1, Not I-Sal I-Cut, 10 ng of cDNA, and sufficient DEPC-treated water to bring the volume to 19 µl, and 1 µl of T4 DNA ligase. The reaction was incubated for 3 hours at room temperature.

8). Introduction of cDNA Constructs into E. coli

The ligated cDNA was introduced into ElectroMax DH10B Cells by electroporation as follows.

After completion of the ligation reaction, 5 µl of yeast tRNA, 12.5 µl of 7.5 M NH₄OAc and 70 µl of ethanol (-20°C) were added to precipitate the DNA. The DNA was pelleted immediately at room temperature for 20 minutes at 14,000g. The pellet was washed with 0.5 ml of 70% ethanol (-20°C) and centrifuged for 2 minutes to remove the supernatant. The cDNA pellet was dried at 37°C for 10 minutes to evaporate residual ethanol, and dissolved in 3 µl of sterile, distilled water.

An aliquot of 1 µl of the ligated cDNA was added to 40 µl of ElectroMax DH10 Cells. The electroporation was performed with BioRad Gene Pulser at 2.5 kV in a 0.1-cm gap chamber at settings of 100 Ohms and 25 µF. One ml of S.O.C. medium was added to the electroporated cells, which were then incubated at 37°C for 1 hour with vigorous aeration.
The equivalent of 1, 0.1, and 0.01 µl of cells were prepared by serial dilution into LB medium, and the culture was plated on LB agar plates supplemented with 100 µg/ml ampicillin. The plates were incubated overnight at 37°C, and the remaining transformed cells were stored at 4°C.

After 16-18 hours of incubation, the number of colonies was counted for each plate. The remaining cells were then plated onto twenty 150-mm plates with estimated density about 50,000 colonies per plate, and incubated at 37°C overnight.

d. Analysis of the cDNA Library by Colony Hybridization

Membrane hybridization with specific probes was the method used to screen the cDNA library. A schematic outline of the procedures is presented in Figure 3.

1). Making Replica Membranes

The master plates of cDNA library (twenty 150-mm plates) were removed from 37°C and stored at 4°C for 1-2 hours. Duralon-UV membranes (137-mm circular) were used to lift the colonies from the master plates. The dry membrane was aligned with the plate, and a small hot needle was used to mark the membrane filter to aid in realignment after hybridization. The filter was carefully removed from the plate and placed colony side up on a fresh LB/ampicillin plate. Both master plate and replica filter were incubated at 37°C for at least 4 hours. Another set of filters was made as above after the regeneration of colonies on the master plates. The master plates were sealed with Parafilm, and stored at 4°C.

The replica membranes were placed colony side up on the wet blotting paper saturated with 0.5 M NaOH for 5 minutes to denature the DNA, and blotted on dry
Figure 3. Colony hybridization: identification of a specific clone from a cDNA library by membrane hybridization to a radiolabeled probe.
Plate and grow cDNA library at appropriate density

Place filter on plate to pick up cells from each colony

Incubate the filter in alkaline solution to lyse and denature released plasmid DNA

UV crosslink the released DNA to filter

Block the filter to decrease nonspecific binding; hybridize with labeled probe

Wash the filter at appropriate stringency

Perform autoradiography

Signal appears over plasmid DNA that is complementary to probe

Align the film with the master plate to identify positive clones

Clones from the identified region are replated at low density and the pure isolates can be obtained by repeating the hybridization as shown in the figure.
blotting paper for 5 minutes. The membranes were then placed on the blotting paper saturated with a solution of 1 M Tris-HCl (pH 7.5) for 2 minutes, and blotted dry for 2 minutes. Finally, the membranes were placed on the paper saturated with a solution of 0.5 M Tris-HCl (pH 7.5) and 1.5 M NaCl for 10 minutes, dried for 2 minutes, and followed by crosslinking the DNA to the membrane with the autocrosslink setting on the Stratalinker UV crosslinker.

2). Labeling Probes

a). Endlabeling of Oligonucleotide Probes

T4 Polynucleotide Kinase (PNK) was used to catalyze the transfer of the γ-phosphate from ATP to the 5’-terminus of polynucleotides to end-label oligonucleotide probe PHiA. About 100 ng of PHiA was mixed with sterile, deionized water to bring the volume to 7 µl, and 1 µl of 10x kinase buffer, 1 µl of [γ-32P] ATP (167 µCi/µl), and 1 µl of T4 PNK (1 unit/µl) were added to the tube. The reaction was carried out at 37°C for 30 minutes, and stopped by heating to 65°C for 15 minutes to inactivate the kinase. The unincorporated counts were removed using a MicroSpin G-10 column.

b). Random Labeling of Double Stranded cDNA Probe

The random labeling of cDNA probe was carried out using Amersham “Megaprime DNA Labeling System”. About 25 ng of cDNA template, the EcoR I fragment of PCR clone S7,3 (300 bp) or cDNA clone D2A (1500 bp), was mixed with 5 µl of Random Primers and sterile deionized water to a volume of 26 µl. The mixture was denatured at 100°C for 5 minutes, and then brought to room temperature. The following reagents were added into the mixture: 5 µl of Reaction Buffer, 12 µl of dATP, dTTP, and dGTP (4 µl each), 5 µl of [α-32P] dCTP (10 µCi/µl) and 2 µl of Klenow Fragment. The
reaction was mixed gently, and incubated at 37°C for 30 minutes. The reaction was stopped with the addition of 5 µl Stop Buffer (0.2 M EDTA). A MicroSpin G-10 column was used to remove unincorporated dNTPs.

3). Prehybridization Conditions

For hybridization with oligonucleotide probe PHiA, the membranes were prehybridized in a prehybridization solution of 6x SSC buffer, 5x Denhardt’s reagent, 1% (w/v) SDS, 20 mM NaH₂PO₄, and 500 µg/ml sheared, denatured herring sperm DNA (by boiling for 5 minutes before adding to warmed prehybridization solution), at 0.25 ml/cm² of membrane for 1-2 hours at 42°C in the hybridization bottles.

When the membranes were hybridized with cDNA probes, the prehybridization was carried out at 68°C in a solution of 6x SSC, 5x Denhardt’s reagent, 1% SDS, and 100 µg/ml of sheared, denatured herring sperm DNA.

4). Hybridization Conditions

For oligonucleotide probes, the hybridization was carried out at 42°C overnight with a minimum volume of the following hybridization solution: 6x SSC buffer, 20 mM NaH₂PO₄, 1% SDS, 500 µg/ml of sheared, denatured herring sperm DNA, and ³²P labeled PHiA at 1x10⁶ cpm/ml of hybridization solution.

When a double stranded cDNA probe was used, about 1-2 x 10⁶ cpm/ml of probe were boiled with the salmon sperm DNA for 5 minutes before adding to prewarmed hybridization solution: 6x SSC, 5x Denhardt’s reagent, 1% SDS, and 100 µg/ml of sheared, denatured herring sperm DNA. The hybridization was carried out at 68°C overnight.
5) **Posthybridization Washing**

For oligonucleotide probes, the membranes were washed once at room temperature for 10 minutes with a 6x SSC buffer and 0.1% SDS wash solution to remove any unbound probe and the hybridization solution. Then the membranes were washed twice for 15 minutes each at 50°C, twice at 55°C, once at 60°C if needed. When the washes were complete, the membranes were briefly blotted on blotting papers and sealed damp in plastic wrap, before exposing to X-ray film in an X-ray cassette with an intensifying screen at -80°C overnight.

Post-hybridization wash for cDNA probes was started with a wash for 10 minutes at room temperature in a solution of 1x SSC and 0.1% SDS, then three times for 15 minutes each in a 0.1x SSC buffer and 0.1% SDS wash solution at 55-65°C, before exposing to X-ray film overnight at -80°C.

The damp membranes can be washed for 15 minutes at 65°C in deionized water to strip off the probe for reusing with other probes.

6) **Purifying Positive Clones**

Because of the high density of colonies on the master plate, it is usually impossible to pick a purified colony directly. To isolate and purify positive clones, the following steps were carried out. The autoradiogram was aligned with the corresponding master plate, and a 1-5 mm region surrounding the hybridization signal was scraped with a sterile toothpick and inoculated into 100 µl of LB broth containing 100 µg/ml ampicillin. The bacterial culture was diluted and plated onto 150-mm LB agar plates, supplemented with 100 µg/ml ampicillin, to give 100 to 1000 colonies per plate. The
screening steps, from colony lifting to hybridization with specific probe, were repeated to obtain individually purified, positive clones.

e. Construction of Progressive Unidirectional Deletions of Inserted DNA

Most of the cDNA inserts of positive clones were too long to be sequenced conveniently from a single primer-binding site on the vector. There are two methods available to obtain the sequences of the target DNA kilobases in length. First, the sequence of target DNA can be obtained in a step-wise fashion by using the nucleotide sequences acquired from one set of reactions to design a new oligonucleotide that is then used to prime the subsequent set of reactions. Alternatively, the entire sequence of the target DNA can be obtained by sequencing a nested set of unidirectional deletions, which progressively move the universal priming site closer to the remote region of the target DNA. The first method was a failure in this study due to nature of the sequence from the positive clone(s), therefore, the alternative method was utilized and described here.

The Erase-a-Base System is designed for rapid construction of plasmid subclones containing progressive unidirectional deletions of any inserted DNA. This system is developed based on the fact that exonuclease III (Exo III) can specifically digest insert DNA from a 5’ protruding or blunt end restriction site, while the adjacent sequencing primer binding site is protected from digestion by a 4-base 3’ overhang restriction site or by an α-phosphorothioate filled end. The schematic diagram of the Erase-a-Base System is presented in Figure 4.
Figure 4. Construction of nested deletions for DNA sequencing. Schematic diagram of the Erase-a-Base System (adopted from Promega Erase-a-Base System Manual).
plasmid

- priming site

**Prepare DNA by restriction digestion with appropriate enzymes**

- 5' overhang or blunt end

- 3' overhang or filled in with α-phosphorothioate dNTPs

- Exo III
  - remove timed aliquots

- S1 Nuclease

- Klenow, dNTPs

- T4 DNA Ligase

- transform, plate, select colonies/plaques, sequence
1). **Preparation of the Starting DNA**

The starting DNA was prepared with QIAfilter Plasmid Midiprep Kit following the manufacturer’s instructions. Each preparation with 25 to 50 ml of cell culture was expected to yield 75 to 100 µg of DNA for high-copy plasmids, and 10 to 50 µg for low-copy plasmids.

2). **Restriction Digestion of Plasmid DNA**

The target DNA was cloned in the multiple cloning site of pSPORT 1, clustered with several restriction sites that yield 3' and 5' extensions (Figure 5). To make deletions from T7 promoter side, *Pst I* was chosen to produce a 3' overhang end downstream of T7 primer binding site, and *Sal I* to generate a 5' protruding end close to the insert DNA. Since these two enzymes require different restriction buffers, digestion with *Pst I* was carried out first.

About 10 µg of plasmid DNA was used to set up a 100-µl reaction, with 40 units of *Pst I*, incubated at 37°C for 4 hours. An aliquot of the reaction (about 0.3 µg) was checked for digestion completion by electrophoresis in a 1% agarose gel. When the digest was complete, the proteins were extracted once with 1 volume of TE-saturated phenol:chloroform:isoamyl alcohol (25:24:1), and once with 1 volume of chloroform:isoamyl alcohol (24:1). The upper, aqueous phase was then transferred into a fresh tube, and 0.1 volume of 3 M sodium acetate and 2 volumes of ethanol were added and mixed well, followed by centrifugation at 12,000g for 5 minutes. The pellet was washed with 1ml of 70% ethanol, and dried under vacuum. The digestion with *Sal I* was set up essentially the same as above. After digestion, the plasmid DNA was purified by extraction with phenol and precipitated with ethanol.
Figure 5. Plasmid DNA pSPORT 1. Multiple cloning site and primer binding regions on the vector (A) and list of restriction endonucleases for generating nested deletions with exonuclease III (B).
(A)

M13 Forward Sequencing Primer

5' -CCCAGTCAGGATGGTAAAAACG-3'   SP6 promoter
5' -CCCAGTCAGGATGGTAAAAACGACGTTGAGAAATTAGTGAACACTATAGA

Sph I  Mlu I  SnaB I  Hind III  BamHI  Xba I  Not I
AGAGCTATGACGTCATGCACGCGTACGTAAGCTGGATCCTCTAGAGCGGCCGC

Sal I  Sma I  EcoRI  Rsr II  Kpn I  Pst I
GTCGACCCGGAATTCCCGGCCGTCGACC

cDNA Insert

T7 promoter

CGTACGCTTTCCCTAGTGAGCTGTATTA-3'
3' -GGGATATCACCTAGCATAAT-5'

← T7 Promoter Primer

(B)

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Restriction ends</th>
<th>Resistance to Exo III</th>
</tr>
</thead>
<tbody>
<tr>
<td>M13 primer side:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sph I</td>
<td>3' overhangs</td>
<td>yes</td>
</tr>
<tr>
<td>Mlu I</td>
<td>3' overhangs</td>
<td>yes</td>
</tr>
<tr>
<td>SnaB I</td>
<td>blunt ends</td>
<td>no</td>
</tr>
<tr>
<td>Hind III</td>
<td>5' overhangs</td>
<td>no</td>
</tr>
<tr>
<td>BamHI</td>
<td>5' overhangs</td>
<td>no</td>
</tr>
<tr>
<td>Xba I</td>
<td>5' overhangs</td>
<td>no</td>
</tr>
<tr>
<td>Not I</td>
<td>5' overhangs</td>
<td>no</td>
</tr>
</tbody>
</table>

| T7 primer side: | | |
| Sal I | 5' overhangs | no |
| Sma I | blunt ends | no |
| EcoRI | 5' overhangs | no |
| Rsr II | 5' overhangs | no |
| Kpn I | 3' overhangs | yes |
| Pst I | 3' overhangs | yes |
3). **Exonuclease III Deletion, Ligation and Transformation**

About 5 µg of doubly digested plasmid DNA was used for Exonuclease III deletion, where 25 time points were taken. Six µl of the 10X Exo III buffer and enough water were added into DNA to make a volume of 60 µl. Twenty-five microcentrifuge tubes were set up on ice, and 7.5 µl of S1 nuclease mix was added into each of 24 tubes. The DNA mixture was warmed up to 37°C in a water bath, and a 2.5 µl sample was transferred to the S1 tube (containing no S1 nuclease mix) on ice as a zero time control. About 500 units of Exo III was then added into the DNA mixture, mixed rapidly, and 2.5 µl samples were removed at 30-second intervals into the S1 tubes on ice. After all the samples were taken, the tubes were moved to room temperature for 30 minutes to allow S1 nuclease to remove the remaining single-stranded tails. The reactions were stopped with addition of 1 µl of S1 Nuclease Stop Buffer to each tube and heating at 70°C for 10 minutes to inactivate the S1 nuclease. The extent of deletion was examined by electrophoresis of 2 µl of samples from each time point in a 1% agarose gel. The resection with *Exo III* usually proceeds at about 450 bases/minute at 37°C for most constructs.

The remaining samples were precipitated after addition of 0.3 volume of 7.5 M ammonium acetate and 2 volumes of ethanol for 15 minutes at –20°C. The samples were centrifuged at 12,000g for 5 minutes and the pellets were washed with 0.5 ml of 70% ethanol. The DNA pellets were dried under vacuum, and resuspended in 9 µl of TE buffer. The samples were transferred to 37°C and 1 µl of Klenow mix was added to each sample. After 3 minutes of incubation, 1 µl of the dNTP mix was added, and the incubation was continued for additional 5 minutes. The Klenow was then inactivated at
65°C for 10 minutes. Incubation with Klenow DNA Polymerase serves to flush both ends of the deletion-containing DNA.

The samples were placed at room temperature, and 40 µl of ligase mix was added to each sample. The ligation was carried out for 1-2 hours at room temperature. After ligation was complete, 10 µl of ligation products from each time point was mixed with 20 µl of DH5α competent cells, and incubated on ice for 30 minutes. The competent cells were heat shocked at 42°C in a water bath for 1 minute, and then placed on ice for 2 minutes, before adding to 200 µl of SOC medium and incubating at 37°C with agitation for 1 hour. The cell mixture was plated onto LB agar plates containing 100 µg/ml ampicillin, and incubated overnight at 37°C.

4) Screening of the Truncated Insert Containing Subclones

a). Rapid Screening

About 10 colonies from each time point were randomly selected for using with the cracking procedure, developed by Barnes (1977), which allows rapid estimation of plasmid size. The individual colonies were picked with sterile pipette tips, smeared against the bottom of a microcentrifuge tube, and then streaked onto a fresh LB/ampicillin plate. The plate was incubated at 37°C for several hours to allow the regeneration of the colonies for later analysis.

A volume of 50 µl of 10 mM EDTA, pH 8.0, was added to each tube, and the cells were resuspended with vortexing. Freshly made cracking 2x buffer of 50 µl was added to each tube, vortexed, and incubated at 70°C for 5 minutes. The tubes were cooled down to room temperature, before 1.5 µl of 4 M KCl and 0.5 µl of 0.4% bromophenol blue were added, followed by incubation on ice for 5 minutes, and
centrifugation at 12,000g for 3 minutes at 4°C. About 25-50 µl of each supernatant was checked in a 1% agarose gel, and the plasmid sizes were estimated using supercoiled DNA as size markers. Clones containing deletions of appropriate sizes were grown in overnight cultures and plasmid minipreps were done to obtain DNA sufficient for sequencing.

b). Preparation of Plasmid DNA

About 1.5 ml of the overnight culture was transferred into a microcentrifuge tube and centrifuged at 12,000g for 1 minute. The bacterial pellet was resuspended into 100 µl of ice-cold miniprep lysis buffer (solution I), and incubated at room temperature for 5 minutes. Freshly made solution II of 200 µl was then added and mixed by inversion, followed by incubation on ice for 5 minutes. About 150 µl of ice-cold potassium acetate solution (solution III), pH 4.8, was added, mixed by inversion, and incubated on ice for 5 minutes. The tube was centrifuged at 12,000g for 5 minutes and the supernatant was transferred to a fresh tube. DNase-free RNase A was added to a final concentration of 20 µg/ml, and the incubation was carried out at 37°C for 20 minutes. The plasmid DNA was treated once with 1 volume of TE-saturated phenol:chloroform:isoamyl alcohol (25:24:1), and once with 1 volume of chloroform:isoamyl alcohol (24:1). The upper, aqueous phase was then mixed with 2.5 volumes of ethanol, centrifuged at 12,000g for 15 minutes. The pellet was rinsed with 70% ethanol and dried under vacuum. The plasmid DNA was dissolved in sterile deionized water for further analysis.

c). Restriction Analysis

Each plasmid DNA was digested with both EcoRV, its cleavage site was located on the vector outside the T7 promoter region, and XbaI to release the insert. The
restriction products were analyzed in 1% agarose gels, and the representative clones containing appropriate truncations were sequenced to yield overlapping sequences.

5). DNA Sequencing and Data Compiling

The sequencing was performed using universal T7 promoter primer by an automated DNA sequencer. Three cDNA clones, D2A, TP23, and SL44.1, were taken through the process of progressive deletions and the sequences of their subclones were obtained, analyzed and compiled using the PC/Gene program, with manual inspection and manipulation.

f. Analysis of Positive cDNA Clones

About 100 positive clones were isolated and purified. It was impractical to obtain complete DNA sequence for each of the positive clones, ranging from 400 bps to 2.4 kb in length. The following strategies were employed to aid in analysis.

1). Partial Sequence Analysis

About 20 of the positive clones with various insert sizes were sequenced in both directions using T7 and M13 universal primers. The partial sequences were analyzed to get a general picture from positive clones.

2). Restriction Analysis of Positive cDNA Clones

The complete cDNA sequences of TP23, D2A, and SL44.1 were analyzed with restriction mapping using Omega 2.0 program. Several Restriction enzymes were selected based on the criteria that all the cutting sites should locate within the inserts only, and the frequency of cutting should be greater than 2 but less than 9, so that the
restriction pattern would be unique to the insert and the number of restriction fragments manageable for further analysis.

About 0.5 to 1 µg of plasmid DNA from each positive clone was used to set up restriction reactions in appropriate buffers with selected restriction enzymes. The restriction products were analyzed by conventional gel electrophoresis, either in polyacrylamide gels, or in agarose gels, depending on the size of restriction products. The restriction products of TP23, D2A, and SL44-1 were used as comparisons.

3). PCR Screening with Specific Primers

A 19-base primer, designated as PH5P, was designed based on the 3' end sequences of TP23, which was also shared by SL44-1 (Figure 6). This primer was used as 5' forward primer in the PCR, coupled with M13 primer on the vector as 3' reverse primer, to screen other positive clones.

The PCR was set up by mixing the following reagents: 15 µl of sterile deionized water, 1 µl of plasmid DNA (10 ng/µl), 2.5 µl of 10x PCR buffer, 0.25 µl of dNTPs (15 mM each), 2.5 µl of 60 mM (NH4)2SO4, 2.5 µl of 50% glycerol, 0.5 µl of PH5P (20 pmol/µl), 0.5 µl of M13 primer (20 pmol/µl), and 0.25 µl of Taq DNA Polymerase (5 units/µl). The PCR cycles were carried out as follows: 94°C, 2 minutes, one cycle; 94°C, 15 seconds, then 65°C, 15 seconds, 30 cycles; and 72°C, 5 minutes, one cycle. The PCR products were analyzed in 2% agarose gels, and the PCR products of TP23, SL44-1, and D2A were used as comparisons.
Figure 6. Design of oligonucleotide primer PH5P for screening of positive cDNA clones using PCR. (A) PH5P was designed based on the 3' end sequence of the cDNA insert from clone TP\textsubscript{23}. The sequence of PH5P is underlined, and the nucleotide position numbers are given in the right margin. The amino acid position numbers are given in superscripts, and the length of poly A tail is in subscript. The putative stop codon is in bold type. (B) The C-terminal amino acid sequences derived from cDNA inserts of TP\textsubscript{23} and SL\textsubscript{44.1} are compared to the N-terminal sequence of the isolated porcine parotid hormone.
(A)

TP23:  

\[
\begin{align*}
\text{CCGCCACCTGTGCGACACCCCGCCGACCCGCACACGACGCCACCTG} & \quad 1620 \\
\text{PAPPGAP520RPPPAGPPPPPPG} & \quad 1620 \\
\text{GACCAGCCCCAAGCTGTGCCAGGCTCTCCGCCGGACCTCCCCACCTG} & \quad 1680 \\
\text{GPPAPPG540AARPPPPGPPPPGGPP550P} & \quad 1680 \\
\text{GCCCAAGACCTTACACCAGGCCCACGGACGTGAACCGGAATTTCAAGGACGAAA} & \quad 1740 \\
\text{SPPPSP560P} & \quad 1740 \\
\text{TGAGCGGAACCAACCGCTCCTGGAGGGCCTGACATTTGGAATTGAGCTT} & \quad 1800 \\
\text{CGATTACAAATAATCATCTCTATCC(A)15} & \quad 1843 \\
\end{align*}
\]

PH5P: 5’ GTCCCAAGGGGCAAGGACC 3’

(B)

TP23 3’ end:  ...PPPPGPAPPGARPPPPGPPPPGSPPRPPPPGPPPQ end
SL44 3’ end:  ...PPPPGPAPPGARPPPPGPPPPGSPPRPPPPGPPPQ end
Isolated PH:  NH2–APPGARPPPPGPPPPGSPRPPPPGPPPQ...
2. Tissue Distribution of PH Gene Product(s)

a. Northern Analysis

1). Isolation of Total RNA and mRNA

The methods for isolation of total RNA and mRNA were as described in previous sections (p30 and p37).

2). Preparation of Northern Blots

About 20 μg of total RNA from each tissue was loaded onto 12-well, 11x14 cm size, 1.5% agarose formaldehyde gel, running in 1x MOPS buffer at 80-100 Volts for about 4 hours. After electrophoresis, the gel was soaked in a 0.05 N NaOH and 0.15 M NaCl solution for 30 minutes, then neutralized in a 0.1 M Tris-HCl (pH 7.5) and 0.15 M NaCl solution for 30 minutes. The Duralose-UV membrane was prewetted first in deionized water, then in transfer buffer (20x SSC). The RNA was transferred by traditional capillary method with 20x SSC for overnight. After transfer, the location of the wells was marked on the membrane, and the transferred RNA was crosslinked to the membrane using the Stratalinker UV crosslinker. The membrane was then stained with 0.02% Methylene Blue for 1 minute, and destained in DEPC-treated water until the ribosome RNA bands were distinctive. The membrane was photographed for later reference.

About 0.2-0.5 μg of messenger RNA from each tissue was separated by electrophoresis in a 1.5% agarose formaldehyde gel, 7.2x10cm in size. The lane with RNA molecular weight marker was cut off and stained with 5μg/ml ethidium bromide in 1x MOPS buffer for 5 minutes with gentle shaking, then destained in deionized water.
The gel strip was photographed under UV light with a ruler for later reference. The membrane transfer and UV crosslink were carried out as described previously.

3). Hybridization Analysis

QuickHyb Hybridization Solution was used in this Northern Analysis. This solution reduces the time required for hybridization from the customary 12-24 hours down to 1-2 hours, resulting in sensitive detection of target sequences.

The double-stranded cDNA from S7,3 or D2A (EcoR I fragment) were randomly labeled with $^{32}$P as previously described (p48). The membrane was prehybridized in QuickHyb solution (minimum 33 µl/cm² of blot) at 68°C for 10-20 minutes. About 2.5 x 10⁶ total counts of labeled probe per 2 ml of hybridization solution was boiled with 100 µl of 10 mg/ml sonicated salmon sperm DNA for 2 minutes, and then added into the prehybridization solution. The hybridization was carried out at 68°C for 1 hour.

After hybridization, the membrane was washed with 2x SSC and 0.1% SDS (~2.5 ml/cm² of blot) twice for 15 minutes each at room temperature. The final wash was in 0.1x SSC and 0.1% SDS at 60-65°C for 30-60 minutes. The membrane was briefly blotted on blotting paper before wrapped in plastic wrap and exposed to X-ray film with an intensifying screen at −80°C. The exposure time was from 1 hour to 48 hours, depending on the intensity of the signal.

b. In Situ Hybridization and Immunohistochemistry

The following steps were adopted from the method developed by Albrecht et al (1996) with few modifications.
1. Tissue Preparation

a). Harvesting Tissues

Various tissues were harvested from adult pigs under anesthesia or immediately postmortem, and frozen in liquid nitrogen and later transferred to -70°C. Some of the tissues were directly fixed in 4% paraformaldehyde in PBS.

b). Fixation

Frozen or fresh tissues were placed in 4% PFA, and large tissue was cut into smaller pieces as appropriate. The dissections were performed in sterile, ice-cold PBS, and the tissue pieces were rinsed in fresh PBS after dissection was complete. The tissue pieces were placed into 4% PFA, and the volume of the fixative was about 30 times or more than that of the tissue. The fixations were performed at 4°C with gentle rocking overnight.

c). Dehydration

The tissue samples were dehydrated in the following solutions, 30-60 minutes for each step, depending on the size of the tissue to be dehydrated. The dehydration was performed at 4°C on a rocking platform moving at a gentle rate (20 times per minute).

The composition of the dehydration solution was:

1. 0.9 % NaCl
2. 50% ethanol in 0.9% NaCl
3. 70% ethanol in H₂O
4. 90% ethanol in H₂O
5. 100% ethanol
d). Embedding

Prior to embedding, the tissue samples were processed by placing them into the following solutions for 60 minutes each step:

1. 100% ethanol, at room temperature, with gentle rocking.
2. Ethanol/xylene 1:1, at room temperature, with gentle rocking.
3. Xylene, at room temperature, with gentle rocking.
4. Xylene/paraffin wax (melted at 58°C) 1:1, at 58°C
5. Paraffin wax, 3 times at 58°C

The tissue samples were embedded in sterilized metal mold. Care was taken to make sure there was paraffin entirely surrounding the tissue and no air bubble was trapped inside.

e). Sectioning

The tissue-containing block was trimmed to yield crisp face edges angled at 90° to help form nice ribbons. The tissue was cut into 5 µm sections, and a ribbon of 7 to 10 sections was taken at a given level, and gently laid out on the water bath (DEPC-treated water between 45 to 50°C). The sections were separated with clean angled forceps and placed on micro slides (25x75mm). The slides were labeled and dried on a slide warmer at 38°C for 2 hours, then stored in a slide box in a cool dry place.

2). Preparation and Synthesis of the Riboprobes

Single-stranded RNA probes can be generated using plasmid vector containing a polylinker bordered by promoters of T3, T7 or SP6 bacteriophage RNA polymerases. The plasmid DNA used in this study was pCR 2.1 vector, containing the cloned PCR product linked to the T7 promoter. Two plasmids, S7-3 and T4-17, containing similar
cDNA sequences (approximately 300 bps in length) but inserted in opposite orientations, were chosen to generate sense and antisense RNA probes respectively.

The plasmid was linearized first by digesting with BamHI, which cut at a single site downstream of the insert, so that the synthesized RNA probe would be terminated as a result of polymerase run-off. Approximately 10 µg of DNA was linearized. The digestion mixture was then treated with phenol/chloroform, and the DNA template was precipitated with ethanol. After the pellet was washed twice with 70% ethanol, the template was dried and dissolved in DEPC-treated water, and the concentration of template DNA was estimated by OD$_{260}$.

The transcription reaction was carried out as follows: 11 µl of linearized DNA template (total of 1 µg/reaction) diluted in DEPC-treated water, 3 µl of 10x transcription buffer, 3 µl of rCTP, rGTP, rATP (10 mM each), 1 µl of RNasin, 0.8 µl of 1 M DTT, 10 µl of [α-$^{35}$S] UTP (10 µCi/µl), and 1 µl of T7 RNA polymerase (10 units/µl), were mixed well, and incubated at 37°C for one hour. One µl of T7 RNA polymerase was added and the reaction mixture was incubated for an additional hour.

The template DNA was removed by incubation of the mixture with 1 µl of RNase free DNase I (2 units) for 15 minutes at 37°C.

The RNA probe was precipitated by adding the following to the transcription reaction: 200 µl of DEPC-treated water, 5-10 µl of yeast tRNA (10 mg/ml), 250 µl of 4 M ammonium acetate, and 1 ml of ice cold ethanol. The mixture was vortexed well, and centrifuged at 14,000g for 20 minutes. The pellet was rinsed with 70% ethanol twice, and dissolved in 50 µl of warm (50°C) hybridization solution. The incorporation of $^{35}$S UTP was measured in a scintillation counter by counting an aliquot of dissolved pellet.
3). **Hybridization Steps**

Two sets of slides were prepared for hybridization, one set with the antisense probe, the other with the sense probe as negative controls.

a). **Preparation of Sections**

All solutions used for dewaxing sections should be made RNase free. Most working solutions can be used 5 times, while PFA can be reused 2 times, and Proteinase K and acetylation solutions need to be made fresh each time.

The slides were placed in a slide rack, and dipped into the following solutions in the order listed:

1. Histo-Clear to remove the wax, 10 minutes, twice
2. Ethanol series, 0.9% NaCl as dilutant, from 100% ethanol, progressing to 95, 85, 70, 50, and 30%, 20 seconds each step with agitation
3. 0.9% NaCl, 5 minutes
4. 1x PBS, 5 minutes
5. 4% PFA, 20 minutes (this solution can be re-used for step 9)
6. 1x PBS, 5 minutes
7. Proteinase K (20 µg/ml proteinase K in 50 mM Tris-HCl pH 7.6, 5 mM EDTA pH 8.0), 5 minutes
8. 1x PBS, 5 minutes
9. 4% PFA, 20 minutes
10. Acetylation: 600 µl of acetic anhydride was added into 0.1 M triethanolamine-HCl (TEA, pH 8.0), 5 minutes; another 600 µl more
of acetic anhydride was added, 5 minutes. This step was performed in a ventilated hood, and the slide rack was suspended above a rapidly rotating magnetic stir bar in the solution.

(11). 1x PBS, 5 minutes
(12). 0.9% NaCl, 5 minutes
(13). 30% and 50% ethanol, 1 minute each with agitation
(14). 70% ethanol, 5 minute (longer to remove salts)
(15). 80, 95, 100, and 100% ethanol, 20 seconds each with agitation

The slides were air-dried in a dust-free, RNase-free place, and then stored in a sealed container at room temperature for up to a few days.

b). Prehybridization

The slides were incubated first with a prehybridization solution in order to reduce the background. The hybridization mix (about 100 µl per slide, see p for recipe) was heated to 42°C in a heating block, and 20 µl of 12 mM alpha-S-thio ATP, 10 µl of 1 M DTT per ml were added as prehybridization mix. The prehybridization mix was applied onto the dry slide, and covered with a glass cover slip. Care was taken to make sure that prehybridization mix completely covered every section, and no air bubble was trapped. The slides were incubated horizontally for 2 hours at 55-65°C in a humidified chamber, which was achieved by placing several layers of Whatman filter paper soaked with 50% Formamide and 2x SSC underneath the plastic slide rack.

c). Hybridization

The hybridization mix (about 70 µl per slide) was heated to 42°C, and 20 µl of 12 mM alpha-S-thio ATP, 10 µl of 1 M DTT per ml were added. The antisense and sense
probes were heated to 70°C, and the calculated amount of probes (about 2 x 10^6 cpm per slide) was added to hybridization solution. The prehybridization cover slip was removed from each slide, and 70 µl of hybridization/probe mix was applied, followed by covering with a fresh cover slip. The slides were incubated horizontally in the chamber overnight at desired temperature (50°C or lower for low stringency, 60°C for high stringency).

d). Posthybridization Washes

The cover slips were carefully removed, and the slides were placed in a slide rack, then into a solution of 50% formamide, 2x SSC, and 40 mM β-mercaptoethanol (stringent wash) for 30 minutes at desired temperature (between 55 and 65°C, higher than hybridization temperature). The slides were then washed in 0.5 M NaCl, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, and 40 mM β-mercaptoethanol for 30 minutes at 37°C. The nonspecifically bound probe was hydrolyzed with 20 µg/ml RNase A, in 0.5 M NaCl, 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA, for 30 minutes at 37°C. The slides were washed again in the above solution (without RNase A) for 15 minutes at room temperature. The stringent wash was repeated once, followed by washing with 2x SSC for 15 minutes, and with 0.1x SSC for 15 minutes at room temperature. The slides were dehydrated through an ethanol/NH₄ acetate series: 30, 50, 70, 95% ethanol, 30 seconds each, then twice in 100% ethanol. The slides were finally air-dried.

4). Autoradiography and Visualization

The length of time that slides should be exposed to emulsion was determined by first exposing the slides to a high performance autoradiographic film (Kodak BioMax) with a Cronex intensifying screen at room temperature overnight.
a). Application of Emulsion

The stock solution of emulsion was made ahead of time by mixing 118 ml of Kodak NTB-2 emulsion with 200 ml of autoclaved water after both heated to 42°C for 35 minutes. The emulsion was aliquoted into glass scintillation vials (about 20 ml each), wrapped in two layers of aluminum foil, then placed inside a light-tight box. The whole process was carried out in a dark room with a dark red (number 2) safelight. The box was stored at 4°C and kept away from all forms of beta radiation.

Also in the dark room, an emulsion aliquot was warmed up in a 42°C water bath for 12 minutes, and then poured into a small plastic slide mailer. The slide mailer was placed in a Koplin jar in the 42°C water bath. The slide was inserted into the emulsion gently for 3 to 4 seconds, then placed horizontally onto a slide rack after the back of the slide was wiped clean with Kimwipe paper. The slide rack was placed into a drying dark box for 8 hours to overnight at room temperature.

Once dry, the slides were placed into a sealed black box, wrapped in aluminum foil, and stored at 4°C. The emulsion exposure time was about 3 to 4 times the length of the Cronex exposure.

b). Developing

The slides were brought to room temperature first. The following steps were performed in the dark room: developed in Kodak D-19 for 2 minutes, washed in water for 30 seconds, placed in fixer for 5 minutes, followed by rinsing in water for 10 minutes. After developing, the slides were counter-stained with a solution of 2 µg/ml Hoechst dye (33258) in water for 2 minutes, followed by a 2-minute rinse in water. The slides were
air-dried in a dark place. Once dry, the slides were overlaid with 50 µl of a solution of 5 g Canada balsam in 10 ml methyl salicylate, and then covered with a glass cover slip.

c). Viewing and Photography

The images of in situ hybridization specimens were captured with a videocamera linked to a computer. The image revealed by blue Hoechst (a nuclear dye) fluorescence was superimposed with a darkfield image. The silver grains, which appeared as white on the image, indicate the expression of the gene of interest. It can be pseudocolored to yellow to enhance the contrast using the Adobe Photoshop program.

5). Hematoxylin and Eosin Staining of Tissue Sections

Hematoxylin and Eosin (H&E) staining was used to show the microscopic structure of tissue sections. The slides were placed in a slide rack, taken through successive changes in the following solutions for deparaffinization, hydration, H&E staining, and dehydration, as listed below:

(1). Xylene, 2 minutes each, with 3 changes
(2). 100% ethanol, 1 minute each, with 3 changes
(3). 95% ethanol, 1 minute each, with 3 changes
(4). 80% ethanol, 1 minute
(5). Water, 3 minutes
(6). Hematoxylin solution, 45 seconds
(7). Running tap water, about 10 minutes
(8). 80% ethanol, 1 minute
(9). 95% ethanol, 1 minute
(10). Eosin solution, 2 minutes
(11). 95% ethanol, 1 minute each, with 3 changes
(12). 100% ethanol, 1 minute each, with 3 changes
(13). Xylene, 2 minutes each, with 3 changes

Finally, the slides were mounted with Permount with cover slips.

6). Immunohistochemistry for Detection of Parotid Hormone in Tissue Sections

Two sets of slides were used in this study, one set with the primary antibodies against porcine parotid hormone, and the other with normal rabbit serum as controls. Either the antibodies produced in rabbits against the isolated porcine PH (AbPH) or the antibodies made against the synthetic peptide (AbPH_{11}) were used as primary antibody.

The slides were first taken through three 2-minute changes of Mineral Spirits and 2-minute changes of pure Isopropyl alcohol, followed by rinsing in tap water for several minutes. Rubber cement was applied to the slides to circle the sections, while the sections were covered with TBST. After the cement circle dried, 2 drops of 2 N HCl were applied to the tissue sections and incubated for 2 minutes at room temperature. The slides were then washed with TBST for 4 times, and 2 drops of Peroxidase Blocking Reagent (from DAKO) were applied to the sections to quench endogenous peroxidase activity. The slides were placed horizontally on several layers of paper wetted with water in a petri dish with cover on for about 15 minutes.

The slides were washed with TBST for 4 times, and 2-3 drops of 5% nonfat milk (in TBS) blocking solution were applied. After incubation at room temperature for 10 minutes, the blocking solution was removed, and 2 drops of either normal rabbit serum or the primary antibody (1:50 dilution) were applied to the tissues sections. The slides were then incubated in the petri dish on a 35°C heating plate for 1 hour.
After incubation with the primary antibody was complete, the slides were washed with TBST thoroughly (about 8 times) over a 15-minute span. 2 drops of Link (biotinylated anti-mouse, anti-rabbit, anti-goat immunoglobins from DAKO) were applied and the slides were incubated at 35°C for additional 30 minutes.

The slides were washed with TBST thoroughly before 2 drops of Peroxidase Conjugated Streptavidin were applied and the incubation was carried out at 35°C for 30 minutes. Again, the slides were washed thoroughly at room temperature.

For color development, AEC (AEC chromogen plus H₂O₂ substrate, from BioGenex Supersensitive Detection Kit) was applied to the tissue sections, and the slides were washed with TBST after the color developed to desired intensity. Hematoxylin was used to counterstain the sections, and excess dye was rinsed away under running tap water. The slides were mounted with Permount.

c. Southern Analysis of Genomic DNAs

1). Isolation of Genomic DNA

This method was adopted from Sambrook’s Molecular Cloning, A Laboratory Manual: Isolation of DNA from Mammalian Cells (Protocol I). Frozen tissue sample, parotid gland tissue from pig, human, or rat, was removed from −70°C, and ground into fine powder in liquid nitrogen. After the liquid nitrogen evaporated, the powdered tissue was added little by little into approximately 10 volumes of extraction buffer in a 50-ml centrifuge tube with shaking. The tube was incubated for 1 hour at 37°C, and then proteinase K was added to a final concentration of 100 µg/ml. The tube was placed in a 50°C water bath for 3 hours with periodical swirling of the viscous solution. The solution
was then cooled down to room temperature, and an equal volume of phenol (equilibrated with 0.5 M Tris-HCl pH 8.0) was added. The tube was turned slowly to allow two phases to form an emulsion. The separation of the two phases was achieved by centrifugation at 5,000g for 15 minutes at room temperature.

The viscous upper, aqueous phase was transferred with a wide-bore pipette to a clean centrifuge tube, and the extraction with phenol was repeated twice. After the third extraction, the aqueous phase was transferred to a fresh centrifuge tube and 0.2 volume of 10 M ammonium acetate and 2 volumes of ethanol (room temperature) were added. When the solution was thoroughly mixed, the DNA precipitate was immediately formed, which was removed from the ethanolic solution with a U-shaped Pasteur pipette. The DNA was dissolved in TE buffer, and the concentration of the DNA was determined by OD$_{260}$. The DNA samples were stored at 4°C.

2). Preparation of Southern Blots

Restriction endonuclease BamH I, EcoR I, Hind III, Pst I, Sac I, and Xba I were used to digest pig genomic DNA separately, 10 µg each. EcoR I was used to digest human and rat genomic DNA. After restriction, the DNA samples were precipitated in ethanol and loaded onto a 0.7 % agarose gel, and the electrophoresis was performed using 0.5x TBE running buffer at 1 V/cm for about 18 hours. The gel was then stained with ethidium bromide and photographed under UV light with a ruler for later reference.

The following steps were performed to allow more complete transfer of DNA to the membrane. For depurination of DNA, the gel was soaked in 0.25 N HCl for 10 minutes, and rinsed in deionized water. The gel was then denatured in a solution of 1.5 M NaCl and 0.5 M NaOH for 20 minutes, followed by neutralizing in a solution of 1.5 M
NaCl and 0.5 M Tris-HCl (pH 8.0) for 20 minutes. The gel was soaked briefly in transfer buffer (20x SSC) before membrane transfer.

The Duralose-UV membrane was prewetted with deionized water, and then with transfer buffer. DNA samples were transferred to the membrane by traditional capillary blotting with 20x SSC. After the transfer was completed, the position of the gel wells was marked on the membrane, and the transferred DNA fragments were crosslinked to the membrane using the Stratalinker UV crosslinker.

3). Hybridization Analysis

The prehybridization and hybridization procedures were essentially identical to that of Northern analysis. Post-hybridization wash was carried out first in 2x SSC and 0.1% SDS twice for 15 minutes each at room temperature. The membrane was then exposed to X-ray film if a low-stringency wash was desired. For high-stringency wash, the membrane was washed once for 60 minutes at 60°C, or even at 68°C, with 0.1x SSC and 0.1% SDS. The exposure time varied from overnight to 48 hours, dependent on the intensity of the signals.

3. Expression of Cloned PH cDNA(s)

a. Expression in Cell-free System

The TNT Quick Coupled Transcription / Translation System from Promega was used to transcribe and translate cDNAs cloned downstream from the T7 RNA polymerase promoters. Compared with standard rabbit reticulocyte systems, this system simplifies the process by combining the RNA polymerase, nucleotides, salts and RNasin Ribonuclease Inhibitor with the reticulocyte lysate solution to form a single TNT Quick
Master Mix, thus greatly reduces the time required to obtain *in vitro* translation results. This system provides an easy and fast way for verification and characterization of cloned gene products.

1). Construction of Expression Plasmids

Plasmid DNAs from cDNA library can be used directly in the TNT Quick Coupled Transcription/Translation System, because the DNA sequences were cloned in the pSPORT 1 vectors downstream from the T7 RNA polymerase promoters. However, the translation efficiency is affected by the secondary structure in the 5'-untranslated region (UTR), and optimal translation occurs if the AUG initiation codon is in a “Kozak consensus” context (A/GCCAUGG) (Kozak, 1986) in the absence of inhibiting secondary structure. The AUG initiation codon of the gene(s) of interest was in a context of “CAAGATGC”, differing considerably from Kozak consensus. Therefore, the target cDNA was subcloned into the vector pET-28a(+), which features an N-terminal HisTag sequence downstream from a T7 promoter, and the initiation codon of HisTag is in a context of “ACCATGG”, matching the Kozak consensus sequence (Figure 7A).

The cDNA insert of D2A was cut out by doubly digestion with Sal I and Not I, and separated from the vector by gel electrophoresis. The insert fragment, about 1.7 kb in size, was purified with QIAGEN gel extraction kit. Vector pET-28a(+) was also digested with Sal I and Not I, dephosphorylated with alkaline phosphatase, and purified by gel extraction. The ligation of D2A fragment with linearized pET-28a(+) was carried out at room temperature for 2 hours, using the ligase mix from Promega’s “Erase-a-Base System”. The sequence of the fusion was obtained by sequencing with T7 promoter primer, and the fusion was confirmed to be in-frame with the HisTag (Figure 7B).
Figure 7. Construction of plasmid DNAs for *in vitro* translation and for expression in *E. coli* cells.

(A) Cloning and expression region of plasmid vector pET-28a(+)
(B) Construction of plasmid DNA pET 28a(+)-D2A
(C) Construction of plasmid DNA pET 28a(+)-Trd8
(A) Plasmid vector pET-28a(+) cloning/expression region

\[ \text{Bgl II} \quad \text{T7 promoter} \rightarrow \quad \text{lac operator} \quad \text{Xba I} \quad \text{rbs} \]

\[
\begin{align*}
\text{AGATCTCGATCCCGAAATTAAATACGACTCACTATAGGGGAATTTGTGACGGGATAACAAATTTCCCTCTAGAAATAATTTTTAATCTTTAAGAAGGAGA}
\end{align*}
\]

\[
\begin{align*}
\text{Nco I} & \quad \text{HisTag} & \quad \text{Nde I} & \quad \text{Nhel} & \quad \text{T7Tag} \\
\text{TATACCATGGGCAGCAGCAGCATCATCATTATCATCACACGACGCGGCTGAGTGCAGCAGGAGCGGCTGCTAACAAAGCCCGAAAGGA} & \quad \text{MetGlySerSerHisHisHisHisHisGlyLeuValProArgGlySerHisMetAlaSerMetThrGlyGlyGlnGlnMetGly} \\
& \quad \text{thrombin}
\end{align*}
\]

\[
\begin{align*}
\text{BamHI} & \quad \text{EcoRI} & \quad \text{SacI} & \quad \text{SalI} & \quad \text{HindIII} & \quad \text{NotI} & \quad \text{XhoI} & \quad \text{HisTag} \\
\text{CGCGGATCCGAATTCCGTAGCTCCGTACAGCTCGGTGAGCGGTCCAGACCCCAGCACGACGCGCAGATCGGAACAC} & \quad \text{ArgGlySerGluGluLeuArgArgGlnAlaCysGlyArgThrArg...} \\
& \quad \text{thrombin}
\end{align*}
\]

\[
\begin{align*}
\text{T7 terminator} \\
\text{AGCTGAGTGGCTGGTGGCAACTACGTCCGTACACGTTGGAATACCCCTTGGGGCTCTACATTACCGGCTTTTGGGTTTTTTT}
\end{align*}
\]
(B) Construction of plasmid DNA pET 28a(+)–D₂A

The cDNA insert of D₂A was released from pSPORT 1 with Sal I and Not I digestions, which was then ligated to pET-28a(+) vector, precut with Sal I and Not I. The result was an in frame fusion with HisTag, as confirmed by sequencing data.

Characters in bold are insert sequences from D₂A. The initiation codon for the insert is marked with ‘*’.

<table>
<thead>
<tr>
<th>Neo I</th>
<th>HisTag</th>
<th>Nde I</th>
<th>Nhel</th>
<th>T7Tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>TATACCATGGCAGCAGCACTCATCATCATCATCAGCAGCAGCGCTGGTGCCGCGCAGCCAGCATATGGCTAGCATGACTGTTGGACAGCAGCAATGGGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetGlySerSerHisHisHisHisHisSerGlyLeuValProArgGlySerHisMetAlaSerMetThrGlyGlyGlnGlnMetGly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BamH I</th>
<th>EcoR I</th>
<th>Sac I</th>
<th>Sal I</th>
<th>***</th>
<th>Not I</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGCGGATCCGAAATTGAGCTCCGTGAACCCACGCGTCCGAGAACAACCGCAGCACTTCTCCAAGATGCTGCGGATC---PolyA---GGCCGGCGCGCACT---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ArgGlySerGluPheGluLeuArgArgProThrArgProGluAsnAsnArgSerThrSerSerLysMetLeuProIle---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(C) Construction of plasmid DNA pET 28a(+)-Trd₈

The insert of Trd₈ was released from pSPORT 1 with Sma I and BamH I, which was then ligated to pET-28a(+), precut with Nde I (blunt ended with Klenow) and BamH I. The resulted fusion was in frame with HisTag, as confirmed by sequencing.

Characters in bold are insert sequences from Trd₈. The initiation codon for the insert is marked with ‘*’.

<table>
<thead>
<tr>
<th>Nco I</th>
<th>HisTag</th>
<th>Nde I SmaI</th>
</tr>
</thead>
<tbody>
<tr>
<td>TATACCATGGCGAGCACATCATCATCATCAGCAGCGCTGGTGCGCGCGCGCGCCATAGGGTCGACCCACGCGTCCGGAACAACCGCAGCACTTCC</td>
<td>MetGlySerSerHisHisHisHisHisHisSerSerGlyLeuValProArgGlySerHisArgValAspProArgValArgAsnAsnArgSerThrSer</td>
<td></td>
</tr>
<tr>
<td>TCCAGATGCTGCCATCTC---PolyA---GGCCGGCCGTCTAGACGATCGAA---</td>
<td>SerLysMetLeuProIleLeu---</td>
<td></td>
</tr>
</tbody>
</table>

BamH I
2). **Translation Procedure**

The translation procedure using Transcend Colorimetric Detection System is presented here. Using this system, biotinylated lysine was added to the translation reaction as a precharged ε-labeled, biotinylated lysine-tRNA complex (Transcend tRNA). The biotinylated lysine residues were incorporated into nascent proteins during translation.

The standard reaction was set up by mixing the following reagents: 40 µl of TNT Quick Master Mix, 1 µl of 1 mM Methionine, 2 µl of DNA template, pET-28a(+)-D2A (0.5 µg/µl), and 1 µl of Transcend Biotin-Lysyl-tRNA. The reaction was incubated at 30°C for 60 minutes.

Since different gene constructs may require different concentrations of Mg²⁺ and K⁺ for optimal expression, additional magnesium acetate and potassium chloride were added to the reaction to find the optimal Mg²⁺ and K⁺ levels. Magnesium acetate was added in 0.1 mM increments up to 0.5 mM, and potassium chloride in 10 mM increments up to 50 mM.

3). **Detection of Translated Proteins**

Once the translation reaction was complete, a 2-µl aliquot was removed and added into 20 µl of 1x SDS gel loading buffer, and heated at 100°C for 2 minutes to denature the proteins. About 5-10 µl of the denatured sample was then loaded onto a 12% SDS polyacrylamide gel with 5% stacking gel. The electrophoresis was carried out at constant current of 15 mA in the stacking gel and at 30 mA during separation.

After electrophoresis, the proteins were transferred onto nitrocellulose membrane by electro blotting using BioRad semi-dry blotting system. The biotinylated proteins
were visualized by binding Streptavidin-Alkaline Phosphatase (Streptavidin-AP), followed by colorimetric detection.

The details of the detection process include the following steps. First, the membrane was rinsed briefly with TBS after blotting. After rinsing, 15 ml of TBS containing Tween 20 (0.5%) was added to submerge the membrane, protein side up, and incubated at room temperature for 60 minutes with gentle agitation. After removal of the blocking solution, 6 µl of Streptavidin-AP in 15 ml of fresh TBS + Tween 20 (0.5%) was added. The Streptavidin-AP binding was carried out at room temperature for 45 minutes with gentle rocking. The membrane was washed 2 times for 1 minute each with 30 ml of TBS containing Tween 20 (0.5%), and then 2 more times with 30 ml of TBS.

The color reaction was started with incubating the membrane in 5 ml of Western Blue Stabilized Substrate solution. After the color of the bands developed to the desired intensity, the reaction was stopped by washing the membrane in deionized water for several minutes with a couple of changes. The membrane was air-dried and photographed.

The detection of translated products was also performed with Western Blotting using the polyclonal antibody AbPH, as described later in “transient expression in COS-7 cells”.

**b. Expression of PH Containing cDNAs in E. Coli**

1). **Construction of Expression Plasmids**

Clones from the cDNA library can be used directly for expression of cloned cDNAs in *E.coli* by inducing the *lac* promoter. The vector pSPORT 1 contains a single
copy of the \textit{lac} repressor gene (\textit{lacI}), so that the transcription of cDNA from the \textit{lac} promoter is effectively repressed unless the promoter is induced by the addition of 1 mM isopropylthio-\textit{\beta}-galactoside (IPTG).

Target genes cloned in pET-28a(+), on the other hand, are under control of bacteriophage T7 transcription signals, and expression is induced by providing a T7 RNA polymerase in the host cell. High expression strains contain a chromosomally integrated copy of the T7 RNA polymerase gene under \textit{lacUV5} control, therefore the activation of the \textit{lacUV5} promoter can be induced by the addition of IPTG.

In addition to pET-28a(+)\textit{-D2A}, another positive cDNA Trd\textsubscript{8} (in vector pSPORT 1) was subcloned into pET-28a(+) and used to transform \textit{E. coli}. The construction of pET-28a(+)\textit{-Trd8} was similar to that of pET-28(+)\textit{-D2A}, except that the cDNA insert of Trd\textsubscript{8} was released from pSPORT 1 by restriction with \textit{Sma I} (generating fragments with blunt ends) and \textit{BamHI}, and the vector DNA pET-28a(+) was cut with \textit{Nde I} and \textit{BamHI}. The digestion with \textit{Nde I} was followed by incubation with T4 DNA polymerase to produce blunt-end products, before subjecting the DNA fragment to restriction with \textit{BamHI}. The ligated product was an in-frame fusion with HisTag as confirmed by sequencing (Figure 7C).

2). The Choice of Expression Hosts

Different expression constructs may require different expression hosts. For example, the \textit{E. coli} strains DH\textsubscript{5x}, NovaBlue and BL21 do not contain the gene for T7 RNA polymerase and therefore cannot be used for T7 promoter driven gene expression studies. These strains are not suitable for expression of the genes cloned in pET-28a(+), but may be suited for gene expression from \textit{E. coli} promoters such as \textit{lac}, \textit{tac} and \textit{trp}, as
in vector pSPORT 1. Host strains BL21(DE3) and NovaBlue(DE3), on the other hand, carry the gene for T7 RNA polymerase under lacUV5 control, thus are suited to expression from T7 promoters.

Plasmid TP23 and SL44-1 were therefore introduced into NovaBlue and BL21 cells, and the expression was induced by IPTG. Plasmid pET-28a(+) D2A and pET-28a(+) Trd8 were transferred into the NovaBlue(DE3) and BL21(DE3) host strains, and the expression was also induced by IPTG. Ampicillin was used to select for transformants containing plasmid pSPORT 1 at a concentration of 100 µg/ml, and Kanamycin (15 µg/ml) for transformants containing plasmid pET-28a(+).

3). Promoter Activation with IPTG

The general procedure for IPTG induction was similar for all constructs tested, including TP23, SL44-1, pET28a(+) D2A, and pET 28a(+) Trd8. First, a single colony from each construct was grown in 2 ml of LB medium containing appropriate antibiotic overnight at 37°C. LB medium of 5 ml containing the appropriate antibiotic was inoculated with 50 µl of the overnight culture. The culture was then incubated for 2 hours at 37°C with aeration. One ml of uninduced culture was removed afterward and transferred to a microcentrifuge tube. IPTG was added to the culture to a final concentration of 1 mM and the incubation was continued at 37°C with aeration. An aliquot of 1 ml of induced culture was transferred to a microcentrifuge tube at different time points during the incubation (i.e., 1, 2, 3, and 4 hours), and centrifuged at 12,000g for 1 minute at room temperature. The pellet was resuspended in 100 µl of 1x SDS gel-loading buffer, heated to 100°C for 3 minutes, and then centrifuged at 12,000g for 1
minute at room temperature. The suspension of cells was then ready for SDS-polyacrylamide gel analysis.

4). Detection of Expressed Proteins

About 15 µl of each suspension was loaded onto a 6% SDS-polyacrylamide gel, and the electrophoresis was carried out at a constant voltage of 15 V/cm until the bromophenol blue tracking dye reached the bottom of the gel.

The gel was then stained with at least 5 volumes of Coomassie Brilliant Blue staining solution for a minimum of 4 hours at room temperature on a slowly rotating platform. The gel was then soaked in the destaining solution on the rocking platform for 4 to 8 hours, with 3 to 4 changes. The photograph of the stained gel was taken afterward. It was expected that the induced protein should appear as a novel band compared to the negative control. Both uninduced cells and cells transformed with vector alone can serve as the negative control here.

c. Transient Expression in COS-7 Cells

1). Construction of Expression Plasmid pCEP4-TP23

The cDNA insert of TP23 was subcloned into vector pCEP4, an expression vector containing CMV promoter, for expression in mammalian cells. TP23 was chosen because it contains a 3’ end sequence identical to the first 30 amino acids of the isolated porcine parotid hormone.

The insert of TP23 was released by sequential digestion with Kpn I and BamH I, and the fragment was purified from the gel using the QIAquick Gel Extraction Kit. Plasmid DNA pCEP4 was also digested with Kpn I and BamH I, dephosphorylated with
alkaline phosphatase, and purified from the gel. The TP_{23} fragment was then ligated to the linearized pCEP4 vector with ligase mix from Promega’s “Erase-a-Base System”. The ligation product was introduced into DH_{5α} competent cells, and the recombinant plasmid DNA was identified by plasmid miniprep, followed by restriction analysis with \textit{Kpn I} and \textit{BamH I} (Figure 8).

2). Transfection of COS-7 Cells with FuGENE 6

FuGENE 6 Transfection Reagent is a multi-component lipid-based transfection reagent that complexes with and transports DNA into cells during transfection. The expression plasmid pCEP4-TP_{23} was prepared with QIAfilter Midiprep Kit, and dissolved in sterile deionized water at a concentration of 1 µg/µl.

COS-7 cells were plated one day before the transfection experiment. The plating density was about 1-3 \times 10^5 cells in 2 ml media in a 35 mm culture dish so that the culture was approximately 50-80% confluent after overnight incubation. The cells were grown in DMEM medium supplemented with 10% FBS under a 5% CO_{2} atmosphere at 37°C.

About 3 µl of FuGENE 6 Reagent were used for 1 µg of DNA. For transfection of cells on a 35 mm plate, 97 µl of serum-free medium was pipetted into a small sterile tube first as diluent, and then 3 µl of FuGENE 6 Reagent was added directly into this medium, mixed by gently tapping. About 1µg DNA was added, and mixed gently. The complex mixture was incubated at room temperature for a minimum of 15 minutes before adding it dropwise to the cells. The plates were swirled to ensure even dispersal, followed by incubation at 37°C. Several plates of cells were transfected each time so that the expression levels could be determined at various time points after transfection.
Figure 8. Construction of plasmid DNA pCEP4-TP\textsubscript{23} for expression in COS-7 cells. The cDNA insert of clone TP\textsubscript{23} was released with $Kpn$ I and $BamH$ I from vector pSPORT 1, and ligated to vector DNA pCEP4, which was precut with $Kpn$ I and $BamH$ I.
pSPORT-TP23: multiple cloning site and expression region

\[ \text{Hind III BamH I Xba I Not I TP23 Insert Sal I Sma I EcoRI Rsr II Kpn I Pst I} \]

--- [T7 Promoter]---

pCEP4: multiple cloning site and expression region

\[ \text{BamH I Sfi I Xho I Not I Nhe I Hind III Nhe I Pvu I Kpn I} \]

--- [SV40 Poly A Signal]---

--- [CMV Promoter]---

pCEP4-TP23: multiple cloning site and expression region

\[ \text{BamH I Xba I Not I TP23 Insert Sal I Sma I EcoRI Rsr II Kpn I} \]

--- [SV40 Poly A Signal]---

--- [CMV Promoter]---
3). Detection of Expressed Proteins

The transfected cells were harvested after transfection by removing the plates from incubator and placing them on crushed ice. The culture medium was transferred into several microcentrifuge tubes and stored at \(-70^\circ\text{C}\) before analysis. The attached cells were washed twice with ice cold PBS, and the last traces of PBS were drained from the plate. About 0.25 ml of ice-cold lysis buffer (single-detergent lysis buffer) was applied to the cell monolayer in 35 mm plate. The plate was incubated on ice for 20 minutes before the cell layer was scraped off with a rubber policeman. The cell debris in the lysis buffer was transferred to a chilled microfuge tube, and the lysate was centrifuged at 12,000g for 2 minutes at \(4^\circ\text{C}\). The supernatant was transferred to a fresh microfuge tube, and stored at \(-70^\circ\text{C}\) before analysis.

About 10 µg of total proteins taken at each time point (quantified with BioRad Bradford Protein Determination reaction) was mixed with an equal volume of 2x SDS gel loading buffer. The protein samples were boiled at 100\(^\circ\text{C}\) for 3 minutes before separation by SDS-PAGE.

The SDS-PAGE was performed in 12% separating gel using a 5% stacking gel, or in 8% separating gel, or in 4-12% gradient gel NuPAGE from Novagen. After electrophoresis, the proteins were transferred to nitrocellulose membranes via BioRad semi-dry blotting system for immunological detection of immobilized proteins.

The Opti-4CN Detection Kit (Goat-anti-Rabbit) from BioRad was used in Western Blotting according to manufacturer’s instructions. After the transfer was complete, the membrane was rinsed in 1x PBST, twice for 5 minutes each. The membrane was incubated with the blocking solution for at least one hour (or overnight at
4°C). The membrane was washed in PBST twice before subjecting it to primary antibody AbPH. The primary antibody was diluted 5000 fold with an antibody dilution buffer and the incubation was carried out for 1 hour with gentle shaking on a rocking platform. An additional two washes were applied before incubation with the secondary antibody, the horseradish peroxidase conjugated goat-anti-rabbit antibody (HRP-GAR), which was diluted 3000 fold. After 1 hour of incubation with HRP-GAR, the membrane was washed twice with PBST, and the Opti-4CN substrate was applied for color development. The membrane was photographed after the color developed to the desired intensity.
CHAPTER III
RESULTS AND DISCUSSION

A. Cloning of Porcine Parotid Hormone cDNA

1. TA Cloning of RT-PCR Products

The RT-PCR amplifications were set up as shown in Table 2, and one set of reaction was carried out with the PCR “touch-down” procedure (designated as Tn), and another set was done using the PCR stringent procedure (designated as Sn) (Table 3). The RT-PCR products were separated by gel electrophoresis, and the results are shown in Figure 9.

In a follow up experiment, the PCR products from the reaction S7, S12, T2, T3 and T4 were ligated with the pCR 2.1 vectors using the Original TA Cloning Kit. These five reactions were chosen because they generated PCR products of reasonable size (above 100 bps). White colonies and some of the white with blue center or light blue colonies were picked from each transformation experiment. The insert size varied from 60 to 350 bps. Plasmid DNA isolated from clones carrying different size of inserts from each transformation was sequenced in both directions and the obtained sequences were analyzed in all three reading frames. The deduced amino acid sequences were compared with the partial amino acid sequence of the isolated porcine parotid hormone for homology.

The obtained DNA sequences of nine clones, one from S7, and two from each of the other four transformations, are presented in Table 4. Even though they were obtained from RT-PCR reactions with different primers, it is noticeable that five of the nine clones may originate from the same gene transcript or from transcripts of closely related genes.
Figure 9. Separation of RT-PCR products by gel electrophoresis. RT-PCR was set up as presented in Table 2 and 3. (A) shows the products obtained from the “stringent” procedure and (B) presents the products from the “touch down” procedure. M, DNA molecular weight marker.
Table 4. List of Nine Clones Obtained from Cloning of RT-PCR Products

<table>
<thead>
<tr>
<th>Clone</th>
<th>RT primer</th>
<th>5' primer</th>
<th>3' primer</th>
<th>Insert size*</th>
<th>Nucleotide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>S7-3</td>
<td>Oligo D</td>
<td>Oligo A</td>
<td>Oligo D</td>
<td>348 bps</td>
<td>see Figure 11</td>
</tr>
<tr>
<td>S12-2</td>
<td>Oligo B</td>
<td>PHiA</td>
<td>Oligo B</td>
<td>154 bps</td>
<td>similar to 18S rRNA</td>
</tr>
<tr>
<td>S12-9</td>
<td>Oligo B</td>
<td>PHiA</td>
<td>Oligo B</td>
<td>130 bps</td>
<td>similar to 18S rRNA</td>
</tr>
<tr>
<td>T2-8</td>
<td>PHiD</td>
<td>Oligo A</td>
<td>Oligo C</td>
<td>~320 bps</td>
<td>see Figure 10</td>
</tr>
<tr>
<td>T2-16</td>
<td>PHiD</td>
<td>Oligo A</td>
<td>Oligo C</td>
<td>~140 bps</td>
<td>no open reading frame</td>
</tr>
<tr>
<td>T3-12</td>
<td>PHiD</td>
<td>Oligo A</td>
<td>Oligo D</td>
<td>~340 bps</td>
<td>see Figure 10</td>
</tr>
<tr>
<td>T3-15</td>
<td>PHiD</td>
<td>Oligo A</td>
<td>Oligo D</td>
<td>306 bps</td>
<td>see Figure 10</td>
</tr>
<tr>
<td>T4-1</td>
<td>PHiD</td>
<td>Oligo A</td>
<td>PhiD</td>
<td>106 bps</td>
<td>GC rich</td>
</tr>
<tr>
<td>T4-17</td>
<td>PHiD</td>
<td>Oligo A</td>
<td>PhiD</td>
<td>333 bps</td>
<td>see Figure 10</td>
</tr>
</tbody>
</table>

* All sequences were obtained using universal T7 promoter primer as sequencing primer. The sequences of T2-8, T2-16, and T3-12 were not full cDNA sequences. The insert sizes were estimated by gel electrophoresis after releasing the insert.
Figure 10. **Comparison of nucleotide sequences of five PCR clones.** All sequences were obtained using T7 promoter primer as sequencing primer. The sequences of T3-12 and T2-8 inserts were not complete on one end. The sequence of clone S7-3 was inverted and complemented before the alignment was performed using PC gene program.

'\*' shows a position in the alignment that is perfectly conserved. '. ' indicates a position that is well conserved. 'N' indicates an ambiguous nucleic acid is read.
Figure 11. Nucleotide sequence and amino acid sequence derived from the insert of PCR clone S\textsubscript{7,3}. The nucleotide sequence is translated into amino acid sequences in all three frames. The reading frame one is an open reading frame, with first 14 amino acid sequence (underlined) identical to the N-terminal of isolated porcine parotid hormone.
GCACCGCCGGGTGAAGACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC ACC

GCAGCC
Ala
Gln
Ser
This transcript could code for a proline rich protein that shares homology with the parotid hormone. The sequence of S\textsubscript{7,3} seems to be identical to the first 14 amino acid sequence of the isolated parotid hormone (Figure 11), which, in addition to the oligonucleotide probe PHiA, was chosen to generate cDNA probes for screening of the cDNA library.

The PCR provides a powerful approach for cloning a cDNA deduced from the partial amino acid sequence of a purified protein. When degenerate oligonucleotide mixtures are used as primers, the PCR procedure effectively selects the correct oligonucleotides for priming DNA synthesis, because only DNA synthesized from the correct cDNA template will hybridize to oligonucleotides present in both degenerate primer pools, and will allow exponential amplification to take place. However, this advantage may lead to false amplification, if one of the degenerate oligonucleotide pools was designed based on an incorrectly determined peptide sequence.

When RT-PCR is employed to clone a gene, the requirement for sequence accuracy is even more stringent, particularly that of the 3’ primer, because the cDNA template for PCR is provided by reverse transcription with the priming of 3’ primer. If the 3’ primer fails to recognize the correct mRNA sequence for transcription, no correct template will be available for PCR amplification. The success of cloning of an unknown gene by RT-PCR depends on the “specificity” of primers, or the accuracy of amino acid sequences.

In this study, the PHiA sequence seemed to be the only “accurate” primer, Oligo dT, a universal reverse primer was also used in RT-PCR to overcome the potential problems associated with other 3’ primers. Numerous products were generated from this
combination. Several major products were seen ranging from 75 bps to 500 bps, together with many minor products over 1 kb (data not shown). These PCR products were cloned and more than 10 clones were sequenced using standard DNA sequencing procedure. The sequencing was made difficult due to early termination or abrupt loss of the signals. It was later discovered that the primer PHiA bound to multiple sites on a GC-rich template. This could explain why PHiA gave rise to multiple major PCR products, and the sequencing process was difficult.

It was expected that the “touch down” procedure produced more products than the stringent procedure. Lower stringency with the “touch down” procedure allows more bindings with mismatch. Multiple products suggest that the primers were not specific and/or no target template was available for the primers. It was also noticed from the sequence data that some of the clones did not even contain the correct priming site. These findings are apparently the results of “random” binding under low stringent conditions. It is therefore not surprising to see that some of the products were not related to the target gene. Nevertheless, even with the not very precise 3’ primers, five clones obtained from 4 different RT-PCR reactions contained similar if not identical gene transcript(s) for the target gene, suggesting the robustness of RT-PCR technology.

2. Construction and Screening of Porcine cDNA Library

The Poly A⁺ mRNAs from the parotid glands of two pigs, one physiologically stimulated by feeding, the other treated daily with 0.6 mg/kg isoproterenol for 4 days, were prepared using Oligotex mRNA Kit according to manufacturer’s instructions. About 3μg of mRNA was used to construct a porcine parotid cDNA library with GibcoBRL SuperScript Plasmid System, and the ligated cDNA was introduced into
DH10B cells by electroporation. Approximately $10^6$ individual colonies were obtained in total. The analysis of twenty randomly picked individual clones indicated that 15% of the clones was apparent “empty”. The average insert size determined was 1150 bps, ranging from 400 bps to 2500 bps.

Since the production of porcine parotid hormone was shown to be stimulated by feeding as well as by treatment with isoproterenol (observation by Dr. Tieche), including mRNAs from both sources may have increased the abundance of target transcript(s). However, isoproterenol was found to stimulate overall proline-rich protein productions in other animals (Mehansho et al., 1985; Ann et al., 1987). If this was also the case with pigs, there might be a potential problem to screen for the proline rich PH cDNAs among the stimulated other proline rich cDNAs.

Oligonucleotide primer PHiA was selected as the primary probe to screen the cDNA library. The underlying reasons include: 1) PHiA was designed corresponding to the first 7 amino acid sequence (APPGARP) of the isolated porcine PH, which was confirmed by repeated AA sequencing; 2) the Genbank database search indicated this sequence was unique in both amino acid sequence and nucleotide sequence, when compared to all reported proteins; 3) PHiA had a codon degeneracy of 2, which was much lower than Oligo A (over 6000), so that adjusting the concentration of probe to allow for codon usage was easier.

About 100 positive clones were purified and confirmed by several rounds of screening with PHiA. The insert sizes varied from 400 bps to 2.4 kb. Positive clones identified by PHiA also showed cross hybridization with the cDNA probe from PCR clone S7-3. As shown in Figure 12, the number of hybridization signals obtained with...
Figure 12. Autoradiogram of cDNA library hybridized with radiolabeled probes. (A) Hybridization was carried out with $^{32}$P labeled PHiA overnight at 42°C. The membrane was washed once in 6x SSC/0.1% SDS for 10 minutes at room temperature, twice for 15 minutes each at 50°C, twice for 15 minutes each at 55°C, and once for 30 minutes at 60°C. (B) The membrane was probed with the 300 bp insert of PCR clone S$_{7,3}$ overnight at 68°C. The membrane was first washed in 1x SSC/0.1% SDS for 10 minutes at room temperature; then three times for 15 minutes each in 0.1x SSC/0.1% SDS at 55°C to 60°C. The last wash was 30 minutes at 65°C.
PHiA was similar to that of S7-3, but the intensity of signals became stronger after hybridization with the cDNA probe. The length of the cDNA insert from S7-3 was about 300 bps, much longer than oligonucleotide PHiA, a 20-mer probe, therefore, it is not surprising to see stronger signals generated from hybridization with the cDNA probe. The fact that similar hybridization density and cross reactivity existed between PHiA and the cDNA insert of S7-3 suggests that both probes most likely recognized the same transcript(s). The strength of the signals indicated further that the cDNA insert of S7-3 contains a sequence that hybridized to mRNA species of fairly high abundance.

There are several advantages associated with using cDNA probes instead of oligonucleotides for the identification of genes with low levels of transcriptional activity. The transcripts hybridized to the probe would be more specific, and the intensity of the signal would be stronger. The oligonucleotide probe PHiA in this study, however, worked almost equally well, possibly due to the abundance of the target mRNA(s), and the multiple hybridizing sites for PHiA on a single transcript as revealed subsequently by sequencing some of the positive cDNAs.

Since the PH protein contains over 65% proline and glycine, coded by CCN and GGN respectively, their mRNA might form intramolecular and/or intermolecular secondary structures that might interfere with second strand synthesis. However, the abundance of mRNA transcripts detected by PHiA or S7-3 suggests that the RNA:DNA hybrid produced was a reasonable template for second strand synthesis for DNA polymerase I.

Twenty positive clones were randomly selected for initial analysis. The insert sizes were from 400 bp to 2.3 kb, as estimated by gel electrophoresis. The cloned cDNA
fragments were initially sequenced from both directions using universal primers available for cloning vectors and the partial sequences were analyzed to locate the site complementary to PHiA. As shown in Table 5, all but two cDNAs contained the sequence segment(s) identical to PHiA. Thirteen of the clones were found to be full-length cDNAs which extend from 17–27 nucleotides 5' of the translation initiation codon ATG to the poly(A) tail on the 3' end, ranging from 15 to 142 bps, and their nucleotide sequences from 5' were identical. Four of the clones did not contain the initiation ATG, but shared the same poly (A) pattern and/or 3' nucleotide sequences as those full-length cDNAs, suggesting that they might be missing the 5' end of mRNA due to degradation. There were two clones which showed no recognition site for PHiA hybridization, and contained neither initiation codon ATG nor poly A tail. Their partial sequences indicated a high GC content, so they may cross hybridize with PHiA.

Three of the full-length cDNA clones, TP23, SL44.1, and D2A, were further studied by complete sequence analysis. After the first set of sequences were obtained using the universal T7 primer, an oligonucleotide primer complementary to the most distal tract of reliable 5' sequence of D2A was synthesized and used for continued sequencing. The results indicated that this primer may have bound to more than one site and generated therefore a mixture of signals. A close analysis of the nucleotide sequences, specially those from the “degraded” cDNAs, revealed that these cDNAs contain a long segment of internal tandem repeats, which makes it impossible to obtain a complete sequence by directed sequencing with progressively designed oligonucleotides. Therefore, the Erase-a-Base System from Promega was utilized to make progressive unidirectional deletions of the inserted cDNAs.

107
Table 5. Partial Characterization of 20 Positive Clones Isolated from Porcine cDNA Library

<table>
<thead>
<tr>
<th>Clone</th>
<th>Insert Size&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Initiation codon&lt;sup&gt;b&lt;/sup&gt;</th>
<th>poly A tail&lt;sup&gt;c&lt;/sup&gt;</th>
<th>PHiA site&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>A&lt;sub&gt;1&lt;/sub&gt;A</td>
<td>0.4 kb</td>
<td>No</td>
<td>None</td>
<td>No</td>
</tr>
<tr>
<td>A&lt;sub&gt;5&lt;/sub&gt;A</td>
<td>1.4 kb</td>
<td>No</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>A&lt;sub&gt;9&lt;/sub&gt;A</td>
<td>2.0 kb</td>
<td>No</td>
<td>Medium</td>
<td>Yes</td>
</tr>
<tr>
<td>D&lt;sub&gt;2&lt;/sub&gt;A</td>
<td>1.7 kb</td>
<td>Yes</td>
<td>Long</td>
<td>Yes</td>
</tr>
<tr>
<td>C&lt;sub&gt;56&lt;/sub&gt;</td>
<td>1.0 kb</td>
<td>No</td>
<td>None</td>
<td>No</td>
</tr>
<tr>
<td>C&lt;sub&gt;81&lt;/sub&gt;</td>
<td>1.6 kb</td>
<td>Yes</td>
<td>None</td>
<td>Yes</td>
</tr>
<tr>
<td>C&lt;sub&gt;633&lt;/sub&gt;</td>
<td>2.2 kb</td>
<td>Yes</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>C&lt;sub&gt;646&lt;/sub&gt;</td>
<td>2.2 kb</td>
<td>Yes</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>SL&lt;sub&gt;44-1&lt;/sub&gt;</td>
<td>2.3 kb</td>
<td>Yes</td>
<td>Medium</td>
<td>Yes</td>
</tr>
<tr>
<td>TP&lt;sub&gt;1&lt;/sub&gt;</td>
<td>1.0 kb</td>
<td>No</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>TP&lt;sub&gt;6&lt;/sub&gt;</td>
<td>1.6 kb</td>
<td>No</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>TP&lt;sub&gt;12&lt;/sub&gt;</td>
<td>1.5 kb</td>
<td>Yes</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>TP&lt;sub&gt;16&lt;/sub&gt;</td>
<td>2.0 kb</td>
<td>Yes</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>TP&lt;sub&gt;20&lt;/sub&gt;</td>
<td>2.0 kb</td>
<td>Yes</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>TP&lt;sub&gt;23&lt;/sub&gt;</td>
<td>1.9 kb</td>
<td>Yes</td>
<td>Short</td>
<td>Yes</td>
</tr>
<tr>
<td>Trd&lt;sub&gt;5&lt;/sub&gt;</td>
<td>2.1 kb</td>
<td>Yes</td>
<td>Long</td>
<td>Yes</td>
</tr>
<tr>
<td>Trd&lt;sub&gt;7&lt;/sub&gt;</td>
<td>2.1 kb</td>
<td>Yes</td>
<td>Long</td>
<td>Yes</td>
</tr>
<tr>
<td>Trd&lt;sub&gt;8&lt;/sub&gt;</td>
<td>2.0 kb</td>
<td>Yes</td>
<td>Medium</td>
<td>Yes</td>
</tr>
<tr>
<td>Trd&lt;sub&gt;11&lt;/sub&gt;</td>
<td>2.0 kb</td>
<td>Yes</td>
<td>Medium</td>
<td>Yes</td>
</tr>
<tr>
<td>Trd&lt;sub&gt;12&lt;/sub&gt;</td>
<td>2.0 kb</td>
<td>Yes</td>
<td>Medium</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<sup>a</sup> The insert size was estimated by gel electrophoresis after releasing the insert.

<sup>b</sup> The 5' sequence of the insert contains "ATG" as putative initiation codon.

<sup>c</sup> The length of poly A tail: short=15-38 bases, medium=60-70 bases, long=130-140 bases.

<sup>d</sup> The partial sequence of the insert contains sequence segment(s) identical to the sequence of PHiA
For each cDNA clone, overlapping subclones (each overlap was at least 100 nucleotides long) were generated and introduced into DH<sub>5</sub>a competent cells. Figure 13 shows the subclones which were selected through rapid screening and restriction analysis. The DNA sequence of the truncated insert from these subclones was obtained using universal T7 primers and the subclone sequences were compiled to give complete sequence of the original clone.

Several complications were encountered during the sequencing effort, which were caused by the high GC content of the cloned cDNAs. The secondary structure formation produced shorter readable sequences, and the long stretches of tandem repeats may have made this problem worse. Even with a sequencing reaction suitable for GC-rich template, which includes 5% DMSO in the reaction mix, elevating the denaturing temperature, and addition of more Taq polymerase, it was still difficult to sequence these cDNA fragments. Figure 14 shows an example of the sequencing data in which a long stretch of GC rich region led to early termination of the sequencing signals. Therefore, more subclones spanning those problematic segments had to be made and sequenced to obtain complete sequences.

The long stretches of tandem repeat also caused difficulties in aligning sequences from progressive deletions. The insert size of each subclone sequenced was carefully estimated by gel electrophoresis, and the sequence data were complied with extra attention. It should be noted, however, that the accuracy of the complete sequences is not definite without the complete sequence data from the complementary strand of the cDNAs.
Figure 13. Progressive unidirectional deletions of a cDNA insert. Twenty one deletion fractions were obtained from cDNA clone TP23 and the extent of deletion was examined by gel electrophoresis (A). Several clones from each deletion fraction were subjected to agarose gel electrophoresis to estimate plasmid size by the rapid screening procedure (B). The representative subclone from each deletion fraction was digested with EcoRV and XbaI, and the size of released insert was estimated by gel electrophoresis (C). M, DNA molecular weight marker.
Figure 14. A sample of DNA sequencing data. The GC rich region of the template produced shorter readable sequence.
3. **Sequence Comparison of Three cDNAs**

The complete nucleotide sequences were determined for three cDNA clones, D₂A, TP₂₃, and SL₄₄. The nucleotide sequences of D₂A and TP₂₃ reveal open reading frames which encode polypeptides in length of 511 and 566 amino acids, respectively. The cDNA insert of D₂A contains the complete coding sequence for a putative 48.5-kDa proline-rich protein. In addition, it contains 27 nucleotides of 5'-noncoding sequence and 33 nucleotides of 3'-noncoding sequence, plus a poly A tail of 132 nucleotides. There is a poly A addition site consensus sequence, AATAAAA, 14 nucleotides upstream from the beginning of the poly A tail. The cDNA insert of TP₂₃ contains the sequence coding for a proline-rich protein of 53.2 kDa. The cDNA sequence also includes 17 nucleotides 5' to the putative initiation codon ATG, 109 nucleotides of 3'-noncoding sequence, and 16 bases of a poly A tail. A consensus sequence for poly A addition occurs 12 nucleotides upstream from the beginning of the poly A tail. The complete nucleotide sequences together with the deduced amino acid sequences for these two cDNAs are presented in Figure 15 and 16. Sequence comparison of D₂A cDNA with that of TP₂₃ cDNA reveals a complete identity with the exception that D₂A has a 243 bp “deletion” in the 3' region of the sequence, which is translated to 78-amino acid proline-rich C-terminal for TP₂₃. This “deletion” also removed the stop codon so that a part of 3'-untranslated region for TP₂₃ is translated into a 23-amino acid, non proline-rich C-terminal domain for D₂A (Figure 17 and 18). This deletion could be introduced during the cDNA synthesis, or the result of alternative splicing of mRNAs.

The deduced amino acid sequence from the cDNA insert of D₂A or TP₂₃ also shows secondary modification sites such as two potential N-glycosylation sites at Asn...
Figure 15. Nucleotide sequence and deduced amino acid sequence from D₂A cDNA insert. The start and stop codons are in bold type, and the consensus polyA addition site is underlined. The length of poly A tail is given as subscript. The nucleotide position numbers are given in the right margin, and amino acid position numbers are given as superscripts.
CCACCGCCACCCGGACCAGCCCCACCTGGTGCCAGACCCCCGCCCGGACCGCCCCCACC

GGACCGCCGCCACCCGGACCAGCCCCACCTGGTGCCAGACCCCCGCCCGGACCGCCCCCACC

CCCGGACCGCCGCCACCCGGACCAGCCCCACCTGGTGCCAGACCCCCGCCTGGACCGCCCCCACC

CCACCGGACCAGCCCCACCCGGACCAGCCCCACCTGGTGCCAGACCCCCGCCCGGACCGCCCCCACC

CCCCCACCCGCTGACGAACCCCAGCAGGGACCAGCCCCATCCGGTGACAAACCAAAGAAG

AAACCACCCCCCACCACCGCTGACGAACCCCAGCAGGGACCAGCCCCATCCGGTGACAAACCAAAGAAG

CAGAAAATGAGCGCGAAAACACCCGTCCTTCGGAGAGCCGTGACATTGGAATGTGACGGT

TGAGCTTCGATTACCAATAAAAATCATCTGCATCCAG(A)
Figure 16. Nucleotide sequence and deduced amino acid sequence from TP$_{23}$ cDNA insert. The start and stop codons are in bold type, and the consensus polyA addition site is underlined. The length of poly A tail is given as subscript. The nucleotide position numbers are given in the right margin, and amino acid position numbers are given as superscripts.
CCGGACCAGCCCACCTGTGGTGCAGACCCCGCCCCGCCCGAGCCGCGCCACCCCGAGCCGCGC
PGAPAPP3GAPAPRPPPPPP3GAP390PGAAP3GAPRPPPPPP3GAP410PGAP
CACCCGGACCAGCCCACCTGTGGTGAGACCCCGCCCCGCCGGACCGCCCGCCTGGAGA
PPPPGAPP3GAP400PAPRPPPPPPGAP420PAPRPPPPPPGAP430PAP
CACCGCCACCCGGACCGCCCCACCTGTGGTGACGACCAGCCCGCCACCGACCCCGGC
PPPPGAP440APAPAPRPPPPPP450GAPPPPP
CTGACGACGCCACCCAGGACGACGCCCACCTCGGTGGACAACAAACAAAGAAGAA
ADPQ460GPAPSGDGP70KPKPP
CACCCGCTGGCCACCCCCGCCAGACCCCGCCACCGGACCCCGCCACCGACCCCGGAC
PPAGPAP480PPPPPPPPPPPGAP490PGAP500PAPRPPPPPGAP510PGAP
CCCCACCTGTGGTGACGACCCGGACCCGACCTCGACCACCGGACCCGGACCCCGGAC
APPGAPR500PPRPAGPPPPPGAP510PGAP520PPPPPPGAP530PGAP
CGGCCCCACCTGTGGTGACGACCCCGCCGACGCCAGCCGCCAGCCGCCACGACCAC
PAPPGAP540ARPPPPPPPPGAP550PAPRPPPPPGAP560PGAP
GACCAGCCACCCACCTGGTGAGACCTCGGCCGGAGCTCCCCACCCACCGGCCTGTC
GPPAPPPG540APRPPPPPPGAP550PAPRPPPPPGAP560PGAP
GCCCACAGCGAGCCACCCACCGAGCCGCCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
TGGACACCGGAATTCAGGACAGAAA1740
SPPPPGPPP560GAPPPPP
TGACGCCGAAAAACACCCGTCTTGCGAGAGCCCCGTGACATTGGAATGTCGACA
GCTTTGAGCTT1800
CGATTACCAATAAAAAATCATCTGCATCC(A)161843
Figure 17. Comparison of nucleotide sequences of D$_2$A and TP$_{23}$ cDNAs. Identities of the two sequences are shown by the vertical bars. The putative start and stop codons are in bold type, and the segment shared also by PCR clone S$_{7,3}$ and T$_{4,17}$ is underlined. The nucleotide position numbers are given in the right margin.
The two sequences to be aligned are:

TP23 Total number of bases: 1828.
D2A Total number of bases: 1597.

Open gap cost: 10
Unit gap cost: 10
<table>
<thead>
<tr>
<th></th>
<th>TP23</th>
<th>D2A</th>
</tr>
</thead>
<tbody>
<tr>
<td>590</td>
<td>ACCCTCCCAGACCACCCCCACC T GC T AACGAATCCCAGCCGGGACCCAGA</td>
<td>-590</td>
</tr>
<tr>
<td>600</td>
<td>CCTCCGCCGGGACCACCATCACCACCCGCTAACGATAGCCAGGAGGGATC</td>
<td>-600</td>
</tr>
<tr>
<td>640</td>
<td>ACCCCCATCCGCTGACGGACCCCAGCAAGGACCAGCCCCATCCGG T GACA</td>
<td>-640</td>
</tr>
<tr>
<td>650</td>
<td>ACCCCCATCCCGCTGACGGACCCCAGCAAGGACCAGCCCCATCCGGTGACA</td>
<td>-650</td>
</tr>
<tr>
<td>690</td>
<td>ACCCGGACCAGCCCCACCTGGCGCCAGACCCCCGCCGGGAC</td>
<td>-690</td>
</tr>
<tr>
<td>700</td>
<td>ACCCGGACCAGCCCCACCTGGCGCCAGACCCCCGCCGGGAC</td>
<td>-700</td>
</tr>
<tr>
<td>740</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-740</td>
</tr>
<tr>
<td>750</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-750</td>
</tr>
<tr>
<td>790</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-790</td>
</tr>
<tr>
<td>800</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-800</td>
</tr>
<tr>
<td>840</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-840</td>
</tr>
<tr>
<td>850</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-850</td>
</tr>
<tr>
<td>890</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-890</td>
</tr>
<tr>
<td>900</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-900</td>
</tr>
<tr>
<td>940</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-940</td>
</tr>
<tr>
<td>950</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-950</td>
</tr>
<tr>
<td>990</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-990</td>
</tr>
<tr>
<td>1000</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1000</td>
</tr>
<tr>
<td>1040</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1040</td>
</tr>
<tr>
<td>1090</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1090</td>
</tr>
<tr>
<td>1100</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1100</td>
</tr>
<tr>
<td>1140</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1140</td>
</tr>
<tr>
<td>1150</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1150</td>
</tr>
<tr>
<td>1190</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1190</td>
</tr>
<tr>
<td>1200</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1200</td>
</tr>
<tr>
<td>1240</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1240</td>
</tr>
<tr>
<td>1250</td>
<td>CCCCCACCCGGACCACCGCC T CC T GGACCAGCCCCACC T GGCGCCAGACC</td>
<td>-1250</td>
</tr>
</tbody>
</table>
Figure 18. Comparison of amino acid sequences derived from D2A and TP23 cDNAs. Identities of the two sequences are shown by the vertical bars. The amino acid position numbers are given in the right margin.
The two sequences to be aligned are:

D2A Total number of residues: 511.
TP23 Total number of residues: 566.

Open gap cost: 7
Unit gap cost: 1
D2A  - -----------------IQGQKMSAKTPVLRAVTLECDG -511
TP23  - PPGPSPRPFPGGPPPQ -566

Identity: 488 (95.31%)
Number of gaps inserted in D2A: 1
Number of gaps inserted in TP23: 0
184 and 202, one N-myristoylation site at Gly 207, and one potential casein kinase II phosphorylation site at Ser 52. It is important to know that D2A was identified with PHiA screening, while TP23 was isolated from S7.3 screening, but both clones contain the cDNA sequence of S7.3 as part of their sequences. In other words, screening with either an oligonucleotide probe or with a cDNA probe has identified clones with highly conserved cDNA sequence.

The cDNA insert of SL44:1 contains the complete coding sequence for a putative 62.3-kDa proline-rich protein (676 amino acids) with 23 nucleotides of 5'-noncoding sequence and 114 nucleotides of 3'-noncoding sequence, plus a poly A tail of 64 nucleotides. The consensus sequence for poly A addition site, AATAAAAA, occurs 17 nucleotides from the poly A tail (Figure 19). Compared to TP23, the nucleotide sequence alignment indicates an 89.6% overall identity between these two cDNAs (Figure 20), and when the derived amino acid sequences are compared, the identity is about 88.3% (Figure 21). SL44:1 shares the same 3' end sequences up to 64 amino acids with TP23.

Unlike TP23 and D2A, which may be derived from the same gene transcript, SL44:1 seems to originate from a different transcript. The difference between SL44:1 and TP23 are distinct, and is not likely introduced by insertion, deletion, or alternative splicing. They all contain the same 21-amino acid repeat, however, the codon usage is somehow different, as revealed by alignments in Figure 20 and 21.

These three porcine cDNAs encode polypeptides consisting of a series of domains that are characteristic of proline-rich proteins: signal sequence, transition region (referring to the N-terminal segment between the signal sequence and the proline-rich region), highly repeated proline-rich domain, and C-terminal domain (Clements et al,
Figure 19. Nucleotide sequence and deduced amino acid sequence from SL_{44.1} cDNA insert. The start and stop codons are in bold type, and the consensus polyA addition site is underlined. The length of poly A tail is given as subscript. The nucleotide position numbers are given in the right margin, and amino acid position numbers are given as superscripts.
ACAACCGCAGCACTTCTCCAAGATGCTGCGATCTCTACTGCTGCTGGCTCTACGTTGCTGGCC

M1 L P I L L S V A L 10 L A

TGAGCTCAGCTGGAGCCGAGGGCTTTTTTGATATTAGAGGATGCAAATCAAACTCCCGCTGAGA

L S S A R S P E20 F D L E D A N S N S30 A E

AGTTTCTGAGACACCTCCTGGAAGCGACCACCGAGACCCAGCTCTCCTGAGGAAAGACG

Q G E G H Q K R40 P R P G D G P E Q70 G E

CCCCCACCTGTTGCGACAGCCCCGGCGGGACCGGACCGCCAGCCACCGAACGACCCGACCACCCG

A P P G A R P E80 P G P P P G P P P P E90 G P

CAGCCCCACCTGTTGCGACAGCCCCGGCGGGACCGGACCGCCAGCCACCGAACGACCCGACCACCC


GACCAACGCTACCTCCAGCCAGCAGCCCCGCGCCGACGGCAGCCCCGCGCCGACGGCAGGACGACG

G P P A P D A R120 P P G P P P P P P P P E130 G P

CGGGACCAGCTCACCAGTTGCAGACAGCCCCGCCCCGGAGCGCCGCCCCGGACCCCGAGGACCGACC

G P G P A P A R140 P R P P G P P P P G150 P

CACCCGGACAGCCCGAACCTTCTGTTGCAGACAGCCCCGCCCCGGAGCGCCGCCCCGGACCCCGAC

G P P G A R P160 A R P P P P P P P P P

CGCCACGGGACCAGCCCCACCTTCTGTTGCAGACAGCCCCGCCCCGGAGCGCCGCCCCGGACCCCGAC

G P P G A R P180 G A R P P P P P P P P

CACCACCGGACCAGCCGACCAGCTCCAGCCAGCCAGCCCCGCCCCGGAGCGCCGCCCCGGACCCCGAC

G P P P G A R P200 P G A R P P P P P P P

GACCACCGGACCAGCCGACCAGCTCCAGCCAGCCAGCCCCGCCCCGGAGCGCCGCCCCGGACCCCGAC

G P P P G A R P220 P G A R P P P P P P

CCGGACCGCCCCACCCGGACGAGCTCCACCCAGGTTGCCAGACCCCCGCCCCGGAGCGCCGCCCCAC


CACTCGGACCGACCCGGACCGGAGCAAGCCGCCCCACTCTGTTGCGACAGACCCCCGCCCCGAGGACG


CCACCCGGACCGGCCACCCGGACGGAGACGGCTCCACCCAGGTTGCCAGACCCCCGCCCAGGACG


CCACCCACCGGACCGGCCACCCGGACGGAGACGGCTCCACCCAGGTTGCCAGACCCCCGCCCAGGACG

G P P P G P A300 G P A R P P P P P P P

GACCGCCCCACCGGACCAGCCGACCAGCCGACGGAGACGGCTCCACCCAGGTTGCCAGACCCCCGCC


CCGGACCGCCCCACCGGACGCGGACCGCCACCGGACGGAGACGGCTCCACCCAGGTTGCCAGAC


130
Figure 20. Comparison of nucleotide sequences of TP_{23} and SL_{44-1} cDNAs. Identities of the two sequences are shown by the vertical bars. The putative start and stop codons are in bold type, and the nucleotide position numbers are given in the right margin.
The two sequences to be aligned are:

TP23 Total number of bases: 1828.
SL44-1 Total number of bases: 2167.

Open gap cost: 10
Unit gap cost: 10

TP23
```
GCGACATCTCCCTCCGAGAGTGCTTGCGATTCCTCTCTCAGTGGC CTG
```

SL44-1
```
ACAAACCCAGACTCTCCCTCCGAGAGTGCTTGCGATTCCTCTCTCAGTGGC CTG
```

TP23
```
TGTGGCTGGCCCAGAGCATCTTGAGAACAGAGCTCGACGGACGGGCCGGACG
```

SL44-1
```
TGTGGCTGGCCCAGAGCATCTTGAGAACAGAGCTCGACGGACGGGCCGGACG
```

TP23
```
AATTAATCCTCGCTGAGATTTCTTGAGACCATCCTCGAGGGCGGACG
```

SL44-1
```
AATTAATCCTCGCTGAGATTTCTTGAGACCATCCTCGAGGGCGGACG
```

TP23
```
CACCCAGGCCCCCTCCGCCAGGCCGCCTCTCTGAGGGAAAGCTCAGTGGC
```

SL44-1
```
CACCCAGGCCCCCTCCGCCAGGCCGCCTCTCTGAGGGAAAGCTCAGTGGC
```

TP23
```
AGACCCAGCAGACCTCGGTACGGACGGACGAGGCGGACGAGGACCGGAC
```

SL44-1
```
AGACCCAGCAGACCTCGGTACGGACGGACGAGGCGGACGAGGACCGGAC
```

TP23
```
TGCCAGACCCCCGCCGGACCGCCCCACCTCTGAGGGAAAGCTCAGTGGC
```

SL44-1
```
TGCCAGACCCCCGCCGGACCGCCCCACCTCTGAGGGAAAGCTCAGTGGC
```

TP23
```
CCGCCACCTCGGTGGCTGAGGACGACCCGGACGAGGCGGACGAGGACCGGAC
```

SL44-1
```
CCGCCACCTCGGTGGCTGAGGACGACCCGGACGAGGCGGACGAGGACCGGAC
```

TP23
```
CCCGCCACCCGGACCAGCCCCACCTCTGAGGGAAAGCTCAGTGGC
```

SL44-1
```
CCCGCCACCCGGACCAGCCCCACCTCTGAGGGAAAGCTCAGTGGC
```

TP23
```
CGCCCACTCCGGGTGGCTGAGGACGACCCGGACGAGGCGGACGAGGACCGGAC
```

SL44-1
```
CGCCCACTCCGGGTGGCTGAGGACGACCCGGACGAGGCGGACGAGGACCGGAC
```

TP23
```
CCCGCCACCCGGACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGAC
```

SL44-1
```
CCCGCCACCCGGACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGAC
```

TP23
```
CCCCACCCGGACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGAC
```

SL44-1
```
CCCCACCCGGACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```

TP23
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```

TP23
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```

TP23
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```

TP23
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```

TP23
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```

TP23
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```

TP23
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```

TP23
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
ACCAGCCCCACCTCGGTGGCTGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

TP23
```
GGCGGACTTCACGTACGAGGACGACCCGGACGAGGACCGGACGAGGACCGGAC
```

SL44-1
```
GGTGCCAGACCCCCGCCGGGACCTCCACCACCGGCCACCGCCACCGCCACCGCC
```
TP\textsubscript{23} - CCACCACCCCCCTGTCGCCAGCCACCAAGACCTCACCAGGCCGCCACCC -1711
SL\textsubscript{44-1} - CCACCACCCCCAGCTCCAGCCACCAAGACCTCACCAGGCCGCCACCC -2047
TP\textsubscript{23} - CCAGTGAACCGGAAATGAGGACAGAAAAATGAGCGGAAAACACCCGTCC -1761
SL\textsubscript{44-1} - CCAGTGAACCGGAAATGAGGACAGAAAAATGAGCGGAAAACACCCGTCC -2097
TP\textsubscript{23} - TTCGGAGAGCCGTGACAATTGGAAATGACGCCGTGAGCCCATGATTACCAAT -1811
SL\textsubscript{44-1} - TTCGGAGAGCCGTGACAATTGGAAATGACGCCGTGAGCCCATGATTACCAAT -2147
TP\textsubscript{23} - AAAATCATCTGCATCCA -1828
SL\textsubscript{44-1} - AAAATCATCTGCATCCAGTT -2167

Identity: 1638 (89.61%)
Number of gaps inserted in TP\textsubscript{23}: 26
Number of gaps inserted in SL\textsubscript{44-1}: 2
Figure 21. *Comparison of amino acid sequences derived from TP\textsubscript{23} and SL\textsubscript{44-1} cDNAs.* Identities of the two sequences are shown by the vertical bars. The amino acid position numbers are given in the right margin.
The two sequences to be aligned are:

TP₂₃ Total number of residues: 566.
SL₄₄ Total number of residues: 676.

Open gap cost: 7
Unit gap cost: 1

**TP₂₃**
MLPILLSVALLALSSARSPFFDLEDANSNASEKFLRPYPPPPRPPPPPE -50

**SL₄₄**
MLPILLSVALLALSSARSPFFDLEDANSNASEKFLRPYPPPPRPPPPPE -50

TP₂₃
ESQEGHQKRPRPPGDGPQGPAPPAGRPYPPPPGPPPPGPAAPPAGRP -100

SL₄₄
ESQEGHQKRPRPPGDGPQGPAPPAGRPYPPPPGPPPPGPAAPPAGRP -100

TP₂₃
PPGPYPGPPPPGPAAPPAGRPYPPPPGPPPPGPAAPPAGRPPPPGPP -150

SL₄₄
PPGPYPGPPPPGPAAPPAGRPYPPPPGPPPPGPAAPPAGRPPPPGPP -150

TP₂₃
GGLQQGPAPSHVGPKKPPPPPGA---------GHPRRPPPANESQPGPR -191

SL₄₄
------------------------PPPPGPAPPAGRPPPPGPAPPAGRPP -209

TP₂₃
PPGPYPGPPPPGPAAPPAGRPPPPGPPPPGPAAPPAGRPPPPGPP -291

SL₄₄
PPGPYPGPPPPGPAAPPAGRPPPPGPPPPGPAAPPAGRPPPPGPP -291

TP₂₃
GPAHGARPPGPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGR -341

SL₄₄
GPAHGARPPGPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGR -341

TP₂₃
PPGPYPGPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPG -391

SL₄₄
PPGPYPGPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPG -391

TP₂₃
GPPGPAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGR -441

SL₄₄
GPPGPAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGR -441

TP₂₃
PPGARPPPPPAPPAPDEPQQGPAPSPGKPPKKPPPPGA--------- -485

SL₄₄
PPGARPPPPPAPPAPDEPQQGPAPSPGKPPKKPPPPGA--------- -485

TP₂₃
------------------------PPPPGPAPPAGRPPPPGPAAPPAGRPPPPG -509

SL₄₄
------------------------PPPPGPAPPAGRPPPPGPAAPPAGRPPPPG -509

TP₂₃
- --------------------------------- -485

SL₄₄
- --------------------------------- -485

TP₂₃
PPGPYPGPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPG -559

SL₄₄
PPGPYPGPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPGPAAPPAGRPPPPG -559
Identity: 500 (88.34%)
Number of gaps inserted in TP_{23}: 2
Number of gaps inserted in SL_{44-1}: 3
Figure 22. Four domains of amino acid sequences derived from the cDNA inserts of D2A (A), TP23 (B), and SL44.1(C). The derived amino acid sequence is arranged to align homologous regions. Variations in the repeat regions are indicated in bold type, and two 34-amino acid repeats are underlined.
(A)

Signal Sequence: MLPILLSVALLALSSA$^{16}$

Transition Region: RSPFFFDLEDANSSEKFLRPPGGPRPPPPEE$^{51}$ SQGEHQQKRPPPGDPFQGPAPPGAR$^{78}$

Repeat Region: PPPGPPPPGGPPGPAPPGAR$^{99}$ PPPGPPPPGGPPGPAPPGAR$^{120}$ PPPGPPPPGGPPGPAPPGAR$^{141}$

PPPPPPAGGLQQGPAPSHV$^{162}$
GPKIKPPPGAGHPRPPP$^{182}$
ANESQGPRPPPGPPSPE$^{200}$
ANDSQEIGSPPS$^{211}$
ADQPGQGPAEPSDGKPKKPPPPAGPPPPPPPPFG$^{245}$
PPPPGPAPPGAR$^{257}$

Repeat Region: PPPGPPPPGGPPGPAPPGAR$^{278}$
PPPGPPPPGGPPGPAPPHGAR$^{299}$
PPPGPPPPGGPPGPAPPGAR$^{320}$
PPPGPPPPGGPPGPAPPGAR$^{341}$
PPPGPPPPGGPPGPAPPGAR$^{362}$
PPPGPPPPGGPPGPAPPGAR$^{383}$
PPPGPPPPGGPPGPAPPGAR$^{404}$
PPPGPPPPGGPPGPAPPGAR$^{425}$
PLLPPPGPPGPAPPGAR$^{446}$

PPP PPP$^{454}$
ADEPQGGPAPSGDKPKKPPPPAGPPPPPPPPFG$^{488}$

C-terminal: IQGQKMSAKTPVLRRAVTLG$^{511}$
(B)

Signal Sequence: MLPILLSVALLALSSA

Transition Region: RSPFFDLEDANSNSAEKFLRPPGGPRPPPEE SQGEHQKRPRPPGDGPEQGPAPPGAR

Repeat Region: PPPGPPPPGPAGPPAPPPGAR

Repeat Region: PPPGPPPPGPAGPPAPPPGAR

Repeat Region: PPPGPPPPGPAGPPAPPPGAR

Repeat Region: PPPGPPPPGPAGPPAPPPGAR

C-terminal: PPPGPPPPGPSPRPPPPGPPGPO
**Signal Sequence:**
MLPILLSVALLALSSA\(^{16}\)

**Transition Region:**
RSPFFDLEDANSNSAEKFLRPPGGPPPRPPPEE\(^{51}\)
SQGEHGQKRPRPPDGPEQGPAPPGAR\(^{78}\)

**Repeat Region:**

\[
\begin{align*}
&\text{PPP}GPPPPGPPPPGPAPPGAR^{99} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{120} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{141} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{162} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{183} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{204} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{225} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{246} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{267} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{288} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{309} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{330} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{351} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{372} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{393} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{414} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{422} \\
&\text{ADEPQQGPAPSGDKKKP}^{442} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{463} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{484} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{505} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{526} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{547} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{568} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{589} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{610} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{631} \\
&\text{PPP}GPPPPGPPPPGPAPPGAR^{652} \\
\end{align*}
\]

**C-terminal:**

\[
\text{PPP}GPPPPPPGPPSPPRPPPGPPQ^{676}
\]
Figure 23. Schematic structure of proline rich proteins derived from porcine cDNA clone D2A, TP23 and SL44.1. The amino acid sequences for the 21-amino acid repeat and the C-terminal region of TP23 and SL44.1 are provided at the bottom of the diagram. The sequence obtained from the isolated porcine parotid hormone is also given for comparison. S, signal sequence; T, transition region; R, 21-amino acid repeat region; P', proline rich region with variations; C, C-terminal region; PH, parotid hormone.
STRP'RC

D2A

STRP'RC

TP23

STRP'RC

SL44-1

PPP PPPPGAPPPGAR

PPP PPPPPGSPPRPPPQ

APP GARPPP PPPPGSPPRPPPQ
The deduced amino acid sequences of D2A, TP23 and SL44-I are presented in Figure 22 and 23 to show these four domains, and some variations are observed in the repeat regions. The most striking feature is that these cDNAs contain 21 amino acid repeats tandemly arranged, from 12 repeats in D2A, 14 repeats in TP23, to 26 repeats in SL44-I.

The amino acid compositions of D2A, TP23, and SL44-I are summarized in Table 6, which are very similar to the amino acid composition of the isolated porcine parotid hormone ("partial sequence"), except the contents of glutamine/glutamate are lower. They are more different from human acidic proline rich proteins than from human basic PRPs.

It has been found that in human basic proline rich proteins, sites for the restriction enzyme BstNI occur in the repeated regions at approximately every 63 bases. They were called BstNI repeats by Azen et al. (1984). The typical BstNI repeats encoding human basic PRP are also present in the repeat region of D2A, TP23 or SL44-I, supporting the notion that these three cDNAs may belong to salivary basic PRP multigene families. BstNI cuts the DNA sequence C-C-A/T-G-G, which is within one of the coding sequences for the dipeptide Pro-Gly, C-C-N-G-G-N. Figure 24 shows the DNA and the decoded amino acid sequences of D2A and TP23 compared to the consensus nucleotide and amino acid sequences from several basic PRPs (Azen et al., 1984; Lyons et al., 1988). The BstNI site appears more frequently in TP23 or D2A. In one 21-amino acid repeat, the dipeptide Pro-Gly occurs four times.

Genbank database searches failed to identify significant similarities between porcine cDNA D2A, TP23 or SL44-I and any known protein or nucleotide sequence, except
Table 6. Comparison of Amino Acid Compositions

The amino acid compositions deduced from the cDNA sequences of D2A, TP23 and SL44-1 are compared with the amino acid composition of the isolated porcine parotid hormone (partial sequence), and human acidic (APRPs) and basic proline-rich proteins (BPRPs) (Bennick, 1982; Clements et al., 1985). "-" indicates either "not detected" or "not available".

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>D_{2A}</th>
<th>TP_{23}</th>
<th>SL_{44-1}</th>
<th>PH (partial)</th>
<th>BPRPs</th>
<th>APRPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
<td>51.8</td>
<td>55.8</td>
<td>60.7</td>
<td>54.7</td>
<td>42</td>
<td>29</td>
</tr>
<tr>
<td>Gly</td>
<td>15.6</td>
<td>16.0</td>
<td>17.3</td>
<td>15.4</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>Ala</td>
<td>9.0</td>
<td>8.8</td>
<td>9.1</td>
<td>6.8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Arg/Lys</td>
<td>7.4</td>
<td>6.7</td>
<td>5.6</td>
<td>5.1</td>
<td>7-10</td>
<td>-</td>
</tr>
<tr>
<td>Gln/Glu</td>
<td>4.4</td>
<td>3.5</td>
<td>1.8</td>
<td>12.9</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Ser</td>
<td>3.1</td>
<td>2.8</td>
<td>1.4</td>
<td>3.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Leu</td>
<td>2.3</td>
<td>1.7</td>
<td>1.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Asp/Asn</td>
<td>2.4</td>
<td>2.1</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>9-11</td>
</tr>
<tr>
<td>His</td>
<td>0.7</td>
<td>0.7</td>
<td>0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Val</td>
<td>0.7</td>
<td>0.3</td>
<td>0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Phe</td>
<td>0.5</td>
<td>0.5</td>
<td>0.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Met</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Thr</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ile</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cys</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 24. Comparison of consensus DNA and amino acid sequences of TP23 and D2A cDNAs (from BstN I repeats) to consensus DNA and amino acid sequences of the repeat region from basic PRPs described by others (Azen et al., 1984; Lyons et al., 1988). The nucleic acid codes are: R=G A, Y=T C, K=G T, M=A C, S=G C, W=A T, B=G T C, and V=G C A. The potential BstN I restriction sites (CCA/TGG) are underlined.
TP$_{23}$ and D$_2$A:

```
CCCCCGCCBGGWCCKCCMCCVCCMGGACCRCGCCACCCGGACCMGCCCCACCTGGYGCAGA
PPPGPPPPPPPGPPPPPPPAPPPGAR
```

PRB1 and PRB2:

```
CCTCCTCCAGAAAGCCACAAGGACCAACCACCAAGGAGGCAACAAGCCCAAGGTCCCCCA
PPPGKPPQGPQGGNGKPQGP
```

PRB3 and PRB4:

```
CCTCSTCCRGAAAGCCAGAAGGACCAACCACCAAGGAGGAAACCAGTCCCAAGGTCCCCCA
PPPGKPPQGPQGGNGQSQGP
```
Figure 25. Comparison of signal peptide, repeat, and C-terminal sequences of human, rat, mouse proline-rich proteins to those derived from the cDNA inserts of TP23 or SL44-1 or D2A (denoted as porcine PRPs). The sequences for human acidic and basic proline rich proteins are consensus sequences. The sequences for rat or mouse were obtained from individual cDNAs. The sequences are aligned to give maximum homologies. The conserved amino acids across species are in bold type. The similarities between porcine PRPs and human basic PRPs are also underlined. S, signal peptide; R, proline rich repeat; C, carboxyl-terminal region.
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S</strong>human PRPs:</td>
<td>MLILLSVALLALSSA</td>
</tr>
<tr>
<td><strong>S</strong>porcine PRPs:</td>
<td>MLPILLSVALLALSSA</td>
</tr>
<tr>
<td><strong>R</strong>human acidic PRPs:</td>
<td>PPGKPQGPPQGGHPQGGPP</td>
</tr>
<tr>
<td><strong>R</strong>human basic PRPs:</td>
<td>PPGKPQGPPQGGKPPQGPP</td>
</tr>
<tr>
<td><strong>R</strong>porcine PRPs:</td>
<td>PPGPPPPPPPAPPGAR</td>
</tr>
<tr>
<td><strong>R</strong>rat pRP25:</td>
<td>PPGQKPQGPPQGGPQ-</td>
</tr>
<tr>
<td><strong>R</strong>mouse pMP1:</td>
<td>PPGQKPQGPPQGGPQ-</td>
</tr>
<tr>
<td><strong>C</strong>TP23 or SL44-1:</td>
<td>PAPPGARPQPPGPPQPGPQPQPGPQQ</td>
</tr>
<tr>
<td><strong>C</strong>human basic PRPs:</td>
<td>GNPQQPQAAGQPQGPPQGQPPHPSRPPQ</td>
</tr>
<tr>
<td><strong>C</strong>human acidic PRPs:</td>
<td>GNPQQPQPQGQPPQGQPPQGQP</td>
</tr>
<tr>
<td><strong>C</strong>mouse pUMP4:</td>
<td>IQGQKMSAKTPVLRRAVTLECDG</td>
</tr>
<tr>
<td><strong>C</strong>D2A:</td>
<td>YLFSLFA</td>
</tr>
<tr>
<td><strong>C</strong>mouse pUMP12s:</td>
<td>PAQDATHEQPSYLWFSS</td>
</tr>
<tr>
<td><strong>C</strong>rat pRP33:</td>
<td>PAQDATHEQPSYLWFSS</td>
</tr>
</tbody>
</table>
the 5'-flanking sequence and the putative signal peptide region, which have strong homologies with other PRP mRNAs (>90% identity). Compared to human PRPs (Figure 25), the amino acid sequence of putative signal peptide of these porcine cDNAs is very similar to the consensus sequence of signal peptide for human PRPs (91% identity), but the homology in the repeat region is lower than 50%, while the similarities between rat and mouse or human sequence are higher than 70%. There appears to be little, if any conservation or homology in the carboxyl-terminal regions of the various derived peptide sequences. There is a cluster of aromatic amino acids and leucine in the C-terminal region of D2A, if this is not due to the deletion error happened during cDNA synthesis. In TP23 and SL44-1, remnants of the repeat regions seem to be present in the C-terminal region (Figure 22 and 25).

The most important finding seems to lie with the sequence of last 30 amino acids of TP23 and SL44-1, which is identical to the first 30 amino acid sequences of isolated parotid hormone (Figure 6). This finding is significant in several ways. First, the isolated porcine parotid hormone was the most active peptide purified from parotid extracts, and it may represent the mature protein processed from a larger precursor, in other words, from a larger mRNA transcript. Secondly, it has been found from other studies that all full-size cDNA clones that contain the BstN I-type repeats would code for proteins of larger sizes than all of the basic salivary proteins that have been sequenced, and their amino acid sequences are almost identical to the decoded sequences of repeated elements from the middle or end of the cDNA clones (reviewed by Maeda, 1985). It is therefore reasonable to assume that parotid hormone, similar to other basic proline rich proteins, was derived from the 3' end of the cDNA clones, such as TP23 or SL44-1.
Furthermore, as of this date, the “first” 30 amino acids from the isolated parotid hormone is the only peptide of the hormone that has been confirmed by AA sequencing. Taken all together, the active form of porcine parotid hormone was very likely derived from the 3’ end of these cDNA clones, e.g., TP23 or SL44-1, and was actually 30 amino acids in size. Alternatively, it was derived from a different transcript closely related to TP23 and SL44-1 cDNAs.

The discrepancy between the estimated molecular weight reported for porcine parotid hormone (8100 Da by SDS-PAGE) and the calculated molecular weight for these 30 amino acids (2858 Da) needs further investigation. One possible explanation is that proline rich proteins typically exhibit anomalous migration on SDS-PAGE so that the apparent molecular weight values of proline rich protein peptides are greater than the molecular weights predicted from the cDNA sequences. This will be discussed in more detail in the later sections.

4. Analysis of Positive cDNA Clones

EcoO109 I and AccB7 I were used to digest the plasmid DNAs from all the positive cDNA clones, and D2A, TP23, and SL44-1 were included as comparison. The restriction products were separated by gel electrophoresis, and a sample of results is presented in Figure 26.

From the restrictions with EcoO109 I, at least 16 different digestion patterns were observed. When clones with similar EcoO109 I restriction patterns were digested with AccB7 I, the restriction results varied, indicating an even greater diversity. Summarizing the results from both restrictions have generated a total of at least 45 different patterns.
Figure 26. Restriction analysis of plasmid DNAs carrying inserts hybridized to oligonucleotide probe PHiA. (A) Several positive cDNA clones were digested with EcoO109 I and the digested DNA was subjected to polyacrylamide gel electrophoresis. (B) Plasmid clones were digested with AccB7 I and Xba I and the digested DNA was subjected to agarose gel electrophoresis. The restriction products of cDNA clone D2A, TP23 and SL44-1 were included as comparisons. M, DNA molecular weight marker.
Figure 27. Screening of positive cDNAs with primer PH5P using PCR. The PCR was performed using PH5P as forward primer and an M13 specific primer as reverse primer, and the PCR products were separated by agarose gel electrophoresis. D$_{2A}$, SL$_{44\_1}$ and TP$_{23}$ were used as negative or positive controls. M, DNA molecular weight marker.
Since restriction analysis suggested such a great variety, sequencing all representative clones from each group may become impractical. Previously parotid hormone had been estimated as 90 AA, although it is clear that proline rich proteins exhibit altered gel migration, it is possible that the protein encoded by the cDNA of TP\textsubscript{23} or SL\textsubscript{44-1} is “truncated” version of the complete protein. Thus to minimize the chance of missing a clone with the “complete” sequence for parotid hormone, a specific primer, PH5P, was designed based on the 3’ end sequences of TP\textsubscript{23} or SL\textsubscript{44-1} (Figure 6).

All positive clones were analyzed with PCR, using specific primer PH5P and universal primer M13, to identify clones containing 3’ end sequence similar to TP\textsubscript{23} or SL\textsubscript{44-1}, yet extending beyond the stop codon. A sample of PCR results is presented in Figure 27. Out of about 100 positive clones, four clones showed no major PCR products, including clone D\textsubscript{2}A, which had a different 3’ end sequence, or a deletion at the 3’ end. All others showed only one major product ranging from 285 bps to 350 bps. The product of TP\textsubscript{23} was about 285 bps, and that of SL\textsubscript{44-1} was 333 bps. Four clones with product apparently larger than that of SL\textsubscript{44-1} were sequenced with PH5P as sequencing primer. The results indicated that they all shared the same 3’ end sequences as that of TP\textsubscript{23} or SL\textsubscript{44-1}, except that their poly A tails were longer. Once this was discovered, no further sequencing was performed on the 5’ end of these clones.

Both Eco\textsubscript{O}I09 I and AccB7 I are sensitive to overlapping dcm methylation, which may have accounted for the differences in the cleavage seen here, because dcm modification would generate partial digestion patterns. The difference in the length of poly A tail may have also contributed partially to the diversity, but the overall patterns suggest it originated primarily from the difference in DNA sequence. It may also be due
to a single base difference within the restriction recognition sites. The complete sequence data on these three clones, plus all available sequence information from other positive clones, suggest that all these positive clones contain similar 5' and/or 3' sequences, and the variants arise mainly from length polymorphism in the repeated proline rich regions, and/or possible some other minor difference in the sequences.

B. Tissue Distribution of PH Gene Product(s)

1. Northern Analysis

Northern blots of total and mRNA from pig tissues were probed with $^{32}$P labeled 1.5 kb cDNA insert from D$_2$A. D$_2$A was chosen because it was not shown at that time that the parotid hormone was derived from the 3' end of the cDNA clones such as TP$_{23}$ or SL$_{44.1}$. The picture obtained with D$_2$A probe may represent a broad view for porcine proline rich proteins, not necessarily specific for parotid hormone. On the other hand, if D$_2$A was truncated form of TP$_{23}$, the results with probe from the 3' end of TP$_{23}$ or SL$_{44.1}$ will be no different.

Northern blots demonstrated the presence of at least five RNA species in porcine parotid gland, with estimated size at 890, 1230, 1995, 2820, and 3550 bases (Figure 28). The hybridization signals were more diffused at two smaller size classes. Messenger RNAs in the above size range could be translated to proteins between 32.6 kDa and 130 kDa, which are larger than most of the known human parotid PRPs. In addition, two weak bands were seen at estimated sizes of 7100 and 7900 bases. These two bands remained after 1 hour of wash at 68°C in 0.1x SSC and 0.1% SDS. These two species may represent unprocessed pre-mRNA. There was no hybridization with mRNAs from
Figure 28. Northern blot analysis of polyA (+) RNAs from different porcine tissues. PolyA (+) RNAs were obtained from porcine hypothalamus (Lane 1), parotid gland (Lane 2), submandibular gland (Lane 3), sublingual gland (Lane 4), lung (Lane 5), liver (Lane 6), pancreas (Lane 7) and skeletal muscles (Lane 8). The blot was probed with the 1.5 kb cDNA insert from cDNA clone D2A. The blot was washed 2 times for 15 minutes each in 2x SSC and 0.1% SDS at room temperature, followed by a wash in 0.1x SSC and 0.1% SDS for 60 minutes at 68°C. The sizes of major RNA bands (estimated from a RNA ladder) are shown on the left.
other porcine tissues, including submandibular, sublingual, liver, pancreas, hypothalamus, lungs, and skeletal muscle.

From the number of hybridizing bands in Figure 28, it is apparent that the cDNA insert of D2A contains sequences shared by several mRNAs. Hybridization of several mRNA species to the insert of D2A emphasizes the conserved structure in the mRNAs. This could be the result of the synchronous expression of a family of very closely related genes or from the synthesis of multiple mRNAs followed by different splicing patterns from the same gene(s) (Ziemer et al., 1984; Azen et al., 1984; Maeda, 1985).

Data taken together from those positive cDNA clones and from Northern analysis suggest that the mRNA species of small sizes, e.g., smaller than 2.5 kb, were favorably synthesized during the cDNA library construction. The large mRNA species (greater than 2.5 kb) may have not been included due to the possibility of incomplete cDNA synthesis. This may explain why all positive clones identified so far had an insert smaller than 2.5 kb. In addition, clones with stronger hybridizing signals were picked during the screening of cDNA library. Therefore, most of the positive clones isolated had an insert size above 2 kb. This can be explained as the cDNA clones with larger insert are likely to have more tandem repeats, hence increased sites for PHiA binding and stronger hybridization signals, and the smaller mRNA species (800 to 1200 bps) may have been left out because of this screening bias. The sequence information from a couple of smaller clones, however, suggests that they were probably the degradation products of larger mRNAs.
2. *In Situ* Hybridization and Immunohistochemical Staining

A riboprobe made from T4-17 was used to determine tissue localization of gene expression by *in situ* hybridization. In addition, immunohistochemical staining was performed to localize parotid hormone, in which the primary antibody produced against the synthetic peptide of PH (AbPH₁₁) was used. T₄₋₁₇ was also obtained from cloning of PCR products, but its insert was in the opposite orientation as that of S7-3. The sequence was similar to that of S₇₋₃, and contained a segment that is conserved between D₂A and TP₂₃ (Figure 17). The transcription of T₄₋₁₇ with T7 RNA polymerase would give an anti-sense probe to the target gene(s). Figure 29 to 31 show the results of both experiments.

*In situ* hybridization of parotid gland slices with T₄₋₁₇ cRNA probe showed a high density of silver grains over acinar cell cytoplasm, with occasional grains over nuclei. Grains overlying ducts or connective tissue were not significantly different from background levels. There were also silver grains seen in some of the serous cells in the sublingual glands, while the submandibular glands were negative. No significant silver grain was seen in other porcine tissues such as liver, pancreas, small intestine, trachea, lungs, kidney, hypothalamus, pituitary, and skeletal muscles.

Immunohistochemical staining of parotid gland slices with the AbPH₁₁ was most intense over acinar tissue. There was minimal staining seen in the ductal and other non-acinar areas. Some of the serous cells from the sublingual glands were also stained, correlating with the finding from the *in situ* hybridization. No significant staining was seen on other porcine tissues.

There was no evidence for the presence of mRNA encoded by D₂A cDNA in sublingual tissues as shown by Northern analysis even when the exposure time was
Figure 29. Localization of PH gene transcripts by *in situ* hybridization and the PH proteins by immunohistostaining in porcine parotid gland tissue. (A) Hematoxylin & Eosin staining provides histological reference; (B) *In situ* hybridization of parotid hormone mRNA with an antisense RNA probe generated from the 300 bp insert of PCR clone T₄₁₇. Autoradiographic grains are colored yellow to enhance the contrast; (C) Immunocytochemical localization of parotid hormone with anti-PH serum (AbPH₁₁). The amplification levels are indicated at the right.
Figure 30. Localization of PH gene transcripts by in situ hybridization and the PH proteins by immunohistostaining in porcine sublingual gland tissue. (A) Hematoxylin & Eosin staining provides histological reference; (B) In situ hybridization of parotid hormone mRNA with an antisense RNA probe generated from the 300 bp insert of PCR clone T$_{4-17}$. Autoradiographic grains are colored yellow to enhance the contrast; (C) Immunocytochemical localization of parotid hormone with anti-PH serum (AbPH$_{11}$). The amplification levels are indicated at the right.
Figure 31. Localization of PH gene transcripts by \textit{in situ} hybridization in porcine submandibular gland tissue. (A) Hematoxylin & Eosin staining provides histological reference; (B) \textit{In situ} hybridization of parotid hormone mRNA with an antisense RNA probe generated from the 300 bp insert of PCR clone T$_4$-17. Autoradiographic grains are colored yellow to enhance the contrast. The amplification level is indicated at the right.
increased to 48 hours. In contrast, small amounts of silver grain, as well as immunohistochemical staining were found in some of the demilunar cells in the sublingual gland. It is likely that the Northern analysis was not sensitive enough to detect the mRNAs in whole-gland extracts.

In summary, it is demonstrated that there was an abundant synthesis of parotid hormone related proteins in porcine parotid gland, restricted expression in sublingual glands, and none in submandibular glands. This pattern is similar to the distribution of basic PRPs in humans and rabbits. In both species, BPRPs have been found to present in parotid gland and in a few sublingual demilune cells but not in the submandibular gland (Robinson et al., 1989; Ferrerra et al., 1992).

*In situ* hybridization using the porcine probe was also performed on tissue sections from human and rat parotid glands, and on rat submandibular and sublingual samples. Although *in situ* hybridization is very sensitive and specific, it was hoped that at low stringency, some interspecies cross binding might be observed. Unfortunately, no such cross binding was observed. Negative results on tissue sections from other species suggest that the counterpart gene transcripts are not completely complementary to the porcine sequence.

3. **Southern Analysis of Genomic DNAs**

The cDNA insert from PCR clone S7.3 was used as a probe to examine the distribution of PH encoding sequences in pig genomic DNA. Two major hybridization bands were observed in genomic DNA restricted with *BamH I*, *Hind III*, *Sac I*, and *Xba I*, and five major bands were detected after restriction with *EcoR I* and *Pst I*. However, as shown in Figure 32, the intensity of hybridization bands was different. In the *BamH I*
Figure 32. Southern blot analysis of pig, human and rat genomic DNAs. Genomic DNAs were prepared from parotid tissues of pig, human and rat and digested with restriction enzyme indicated above each lane. The membrane was probed with the 300 bp cDNA insert from PCR clone S7-3. The blot was washed 2 times for 15 minutes each in 2x SCC and 0.1% SDS at room temperature, and once for 60 minutes in 0.1x SSC and 0.1% SDS at 55°C. The sizes of marker DNAs (in kb) are shown on the left.
<table>
<thead>
<tr>
<th>kb.</th>
<th>M</th>
<th>Pig</th>
<th>Pig</th>
<th>Pig</th>
<th>Pig</th>
<th>Pig</th>
<th>Pig</th>
<th>Human</th>
<th>Rat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BamH I</td>
<td>EcoR I</td>
<td>Hind III</td>
<td>Pst I</td>
<td>Sac I</td>
<td>Xba I</td>
<td>EcoR I</td>
<td>EcoR I</td>
</tr>
</tbody>
</table>

10 8 6 5 3
digest, two major bands were seen at size of 10.72 kb and 9.55 kb, and a weak band at 8.32 kb disappeared after the membrane was washed at 60°C for 1 hour. The \textit{Hind III} digest yielded two major bands at 10.72 kb and 9.33 kb, and \textit{Sac I} digest had two major bands at 8.51 kb and 7.85 kb, and a minor one at 10.96 kb, which disappeared after more stringent wash. In the \textit{Xba I} digest, two major bands were seen at 10 kb and 9.33 kb, and a weak band at 3.63 kb. Both \textit{EcoR I} and \textit{Pst I} gave 5 major bands with similar sizes, namely, 3.55, 2.69, 2.40, 2.04, and 1.86 kb. \textit{EcoR I} digestion also yielded two minor bands at 7.24 and 1.45 kb, and \textit{Pst I} digestion yielded two minor bands at 5.75 and 1.41 kb.

When the hybridization was repeated with the cDNA insert from D$_2$A, the hybridization pattern was about the same as that of S$_{7,3}$ probe, except that the intensity of each major band within each digest appeared now to be equal (data not shown).

There was no hybridization of the cDNA probe from D$_2$A with human and rat genomic DNAs, even when the posthybridization washing stringency was lowered to 2x SSC and 0.1% SDS twice for 15 minutes each at room temperature. Again, interspecies variation is great enough, but negative results do not rule out the presence of a homologous gene.

Human proline rich proteins are divided into two subfamilies according to Maeda (1985). One group contains \textit{Hae III} repeats coding for acidic PRPs and the other \textit{BstN I} repeats coding for basic and glycosylated PRPs. When the Southern blots containing \textit{EcoR I} digested human DNA was hybridized with a probe made from each of the two types of repeats, both probes showed hybridization to six identical \textit{EcoR I} fragments under nonstringent conditions (in 3x SSC with 0.5% SDS twice for 1 hr each at 68°C).
After washing under stringent conditions, however, two bands remained visible with the
probe having \textit{Hae III} repeats at positions of 8.4 and 4.6 kb, and four bands hybridized to
the probe having \textit{BstN I} repeats at positions of 6.2, 3.9, 3.6, and 3.3 kb. In addition to
\textit{EcoR I} digests, Southern blot hybridization to human DNA digested with several other
enzymes using a probe containing \textit{BstN I} or \textit{Hae III} repeats showed also six hybridizing
bands. The data suggest that six loci control the synthesis of human salivary proline rich
proteins, including two proline rich protein loci forming the \textit{Hae III}-type subfamily and
four proline rich protein loci forming the \textit{BstN I}-type subfamily (Maeda, 1985).

Since \textit{D}_2\textit{A} contains \textit{BstN I}-type repeats, it is justified to compare its hybridization
pattern with that of human basic PRP genes. Digestion with \textit{EcoR I} or \textit{Pst I} generated
more hybridizing fragments in porcine genomic DNA than that with other restriction
enzymes, suggesting there were cutting sites within the genes for these two enzymes.
Since all other enzyme digests have resulted in two major hybridizing bands, it is
speculated that the basic proline rich protein genes recognized by \textit{D}_2\textit{A} cDNA probe are
localized to two loci in porcine, instead of four as in human.

The fact that no hybridization was observed with the porcine cDNA probe to
human genomic DNA suggests that the porcine basic proline rich proteins are
considerably different from human basic PRPs. This data calls into question that the
nucleotide sequences of proline rich proteins are well conserved across species, as
demonstrated by cross-hybridization among the rat cDNAs (Clements et al., 1985), by
hybridization of the rat \textit{pRP33} to human and mouse (Clements et al, 1985), and by
hybridization of the human \textit{PRP1} to rabbit cDNAs (Ferreira et al., 1992).
C. Expression of Cloned PH cDNA(s)

1. Translation of Cloned PH cDNAs in Cell-free Systems

Plasmid DNA pET-28a(+)\text{-}\text{D}_2\text{A} was used in the TNT Quick Coupled Transcription/Translation System to obtain cell free protein synthesis. Separation of the translation mixture revealed one major product with molecular weight of 90 kDa, which is much higher than 54.4 kDa deduced from the DNA sequence, as shown by the Transcend Colorimetric Detection System (Figure 33A). The Western blotting with antibodies made against the isolated parotid hormone revealed similar result, one major product with apparent molecular weight of 90 kDa (Figure 33B).

Figure 33A shows that the concentration of Mg$^{2+}$ and K$^+$ significantly affected the translation efficiency of pET-28a(+) DNA. In the absence of magnesium acetate and potassium chloride, only limited translation products were obtained. With magnesium at 0.2 mM, and potassium at 20 mM, however, the translation efficiency of pET-28a(+)\text{-}\text{D}_2\text{A} was greatly improved to the level comparable to that of luciferase control reaction provided in the system.

When plasmid TP$_{23}$ or D$_2$A was used directly in the TNT transcription and translation system, the translation products were barely visible (data not shown). When the cDNA insert of D$_2$A was ligated downstream to the HisTag encoding sequence in the plasmid pET 28a(+), an efficient translation was obtained. This finding supports what has been described previously, that optimal translation occurs if the AUG initiation codon is in a "Kozak consensus" context (A/GCCAUGG) in the absence of inhibiting secondary structure.
Figure 33. Western blot analysis of translation products encoded by the 1650 bp insert of plasmid pET 28a(+)--D_2A in a coupled transcription and translation system. The cDNA insert of plasmid pET 28a(+)--D_2A was transcribed and translated in vitro using the TNT Quick Coupled Transcription & Translation System in the presence of biotinylated lysine. (A) Western blot analysis of translated proteins using the Transcend Detection System, in which the biotinylated proteins were visualized by binding streptavidin-alkaline phosphatase, followed by colorimetric detection. Lane 1 represents the translation product (61 kDa) obtained from the luciferase control vector provided by the TNT system. Lane 2 through 7 represents translation products of pET 28a(+)--D_2A encoded transcripts under increasing concentrations of Mg\(^{2+}\) and K\(^+\) as indicated above each lane. The sizes of protein molecular weight markers are given at the left in kDa. (B) Western blotting of translated proteins using the polyclonal antibodies made against porcine parotid hormone (AbPH). Lane 1 represents the translation products obtained from the luciferase control provided by the TNT system. Lane 2 represents the translation products of pET 28a(+)--D_2A encoded transcripts in the presence of 0.2 mM of Mg\(^{2+}\) and 20 mM of K\(^+\). The sizes of protein molecular weight markers are indicated at the left in kDa.
The molecular weight of the translation products was much larger than predicted, but similar to products expressed in COS-7 cells (for details see the following sections). Post-translational modifications cannot be the reason for the size difference since the TNT transcription and translation system lacks such functions. It has been found in several other studies that proline rich proteins tend to migrate at a larger size (Muenzer et al., 1979; Ziemer et al., 1982). In these studies, seven proline rich proteins showed molecular weight between 15 and 25 kDa by sedimentation equilibrium, which by gel filtration gave molecular weights of above 70 kDa, and by SDS PAGE, molecular weight between 22 and 36 kDa. According to the authors, the unexpected behavior in gel filtration and PAGE analysis apparently reflects the high axial ratio (>25) of these proteins, which exist as highly elongated rods. The structural analysis of D2A and TP23 suggests that these PRPs have an elongated bending and twisting, leading to similarly aberrant behavior on gel.

2. Expression of Cloned PH cDNAs in E.coli

Several cDNA constructs were transformed into suitable host strains as described in the method sections. The individual strains containing the expression constructs were then induced by the addition of IPTG. The total protein was separated by gel electrophoresis, and stained with Coomassie blue. Figure 34 shows the protein patterns obtained from E.coli strain BL21(DE3) cells transformed with plasmids pET-28a(+)-D2A and pET-28a(+)-Trd8, before and after induction with 1 mM IPTG. There was no significant difference in the protein expression pattern with IPTG induction. Similar results were observed in NovaBlue(DE3) cells transformed with plasmid DNA pET-
Figure 34. Separation of protein extracts from *E. coli* strain BL21(DE3) transformed with plasmid pET 28a(+)\-D2A and pET 28a(+)\-Trd8 by gel electrophoresis. The pET 28a(+)\-D2A transformed cells were collected before the addition of IPTG (Lane D0), and 1, 2, 3, and 4 hours after IPTG induction (Lanes D1-4). The pET 28a(+)\-Trd8 transformed cells were collected before the addition of IPTG (Lane T0), and 1, 2, 3, and 4 hours after IPTG induction (Lanes T1-4). The cell lysates were separated by SDS PAGE (12% separating gel) and the gel was stained with Coomassie Blue. The sizes of protein molecular weight markers are indicated at the left in kDa.
28a(+)-D$_2$A or pET-28a(+)-Trd$_8$, and BL21 or NovaBlue cells transformed with plasmid TP$_{23}$ or SL$_{44.1}$ (data not shown).

In spite of the fact that the pET expression system has been used to produce substantial amounts of protein from a wide variety of genes of both prokaryotic and eukaryotic sources, in this study, disappointingly small amounts of the protein were made. The explanation may be that this particular protein itself may interfere with gene expression or maybe toxic to the cell. It has been reported for some proteins by pulse labeled studies, in which a gradual or rapid decrease in the rate of protein synthesis was observed as foreign protein accumulated, or sometimes all protein synthesis stopped before any target protein could be detected.

The other reason may be the instability of certain proteins in the host cell. The foreign protein stability may be influenced by the amino acid immediately following the N-terminal methionine (penultimate amino acid). The amino acid at this position determines the removal of N-terminal fMet (Hirel et al., 1989; Lathrop et al., 1992). The catalytic efficiency of methionyl aminopeptidases (MAP) decreases as the size of the penultimate amino acid side chain increases. In these studies, little or no processing was observed when the following amino acids occupied the penultimate position: His, Gln, Glu, Phe, Met, Lys, Tyr, Trp, and Arg. Processing ranged from 16%-97% when the remaining amino acids occupied this position. Tobias et al (1991) have determined the relationship between a protein’s amino terminal amino acid and its stability in bacteria, i.e., the N-end rule. They reported protein half-lives of only 2 minutes when Arg, Lys, Phe, Leu, Trp, or Tyr was present at the amino terminus of the test protein. In contrast, all other amino acids conferred half-lives of > 10 hours when present at the amino
terminus in the same protein. Taken together, these studies suggest that Leu in the penultimate position would be a poor choice, because it would likely be exposed by fMet processing and then be targeted for rapid degradation. In this study, the PH protein has Leu in the penultimate position, and it may explain why the TP23 or SL44-1 expression was not found in significant amounts in *E. coli*. However, the above reasoning does not apply to the other two constructs involving pET-28a(+), in which Gly was in the penultimate position in the HisTag fused proteins.

Besides toxicity and degradation, the expression level of PH proteins may be influenced by other factors, which include secondary structure of the mRNA and excessive rare codon usage. Secondary structure in the mRNA can interfere with the AUG translation initiation codon and/or the ribosome-binding site (Tessier et al., 1984; Looman et al., 1986; Lee et al., 1987). Searching the coding strand of the PH cDNA fragment failed to find stretches of sequence complementary to the mRNA generated from transcription of pET-28a(+) DNA (5'…AAGAAGGAGAUACCAUGG…3'). This suggests that secondary structure was not a plausible cause. The fact that pET-28a(+)-D2A DNA was successfully translated in the cell free system also supported the assumption that secondary structure was not the reason.

Rare codon usage in the target gene has also been implicated as a cause for low-level expression (Zhang et al., 1991; Sorensen et al., 1989). In Sorensen’s study, the difference in translation rate between infrequent codons and common codons was in the order of sixfold. The PH proteins in this study contain more than 5% arginine, all encoded by AGA, the least used arginine codon in *E. coli*. This may have contributed to low levels of expression in *E.coli*. 

182
3. **Expression of Cloned PH cDNAs in Mammalian Cells**

The COS-7 cells were transfected with pCEP4-TP$_{23}$ DNA using FuGene 6 Reagent. The transfected cells were harvested 24, 48, and 72 hours after transfection. Both the culture medium and the cell lysate were collected and analyzed by SDS-PAGE and Western Blotting. The antibody made against the isolated porcine parotid hormone (AbPH) was used as the primary antibody in the Western Blotting. COS-7 cells transfected with vector pCEP4 DNA (mock transfection) were used as negative control.

Figure 35 shows the gel separation of expressed proteins recognized by AbPH. In pCEP4-TP$_{23}$ transfected cells, a major band was seen with an estimated molecular weight of 110 kDa, which is much larger than the predicted 53.2 kDa. In addition, a strong band was seen at the position of 31 kDa, and a weak band at the position of 76 kDa. In the culture medium, a major band was seen at the position of 107 kDa, and two more bands at the position of 33 and 30 kDa. A number of weak bands at smaller size were also observed. No significant cross-reaction was seen in mock-transfected cells and culture medium.

It is likely that the band at 110 kDa in plasmid pCEP4-TP$_{23}$ transfected cell extracts was the correct protein. Again, the apparent molecular weight was much higher than predicted. Since the protein encoded by TP$_{23}$ DNA contains a putative signal peptide, the specific band seen in culture medium at 107 kDa was apparently the secreted form and may have been processed by removing the signal peptide.

The smaller protein bands, which were detected by the PH specific antibodies, may be resulted from protein processing or endopeptidase activities. Apparently, post-translational modifications and/or post-secretory cleavage may generate smaller proteins.
Alternative splicing of mRNAs may also play a role. These questions should be addressed in future studies.
Figure 35. Western blot analysis of proteins expressed in COS-7 cells transfected with plasmid pCEP4-TP23. The expressed proteins were resolved by a 4-12% gradient gel and detected by the polyclonal antibodies made against porcine parotid hormone (AbPH). Lane 1 represents the cell extracts obtained 72 hours after the transfection with pCEP4 vector (with no insert). Lane 2 to 4 represents the cell extracts collected 24, 48 and 72 hours after the transfection with plasmid pCEP4-TP23 (1.7 kb insert), respectively. Lane 5 represents the culture medium collected 72 hours after the transfection with pCEP4 vector, and Lane 6 to 8 represents the culture medium collected 24, 48, and 72 hours after the transfection with pCEP4-TP23. The sizes of protein molecular weight markers are indicated at the right in kDa.
A. Conclusions

To isolate and identify the cDNA sequence of porcine parotid hormone, a porcine parotid cDNA library was constructed and hybridized with oligo probes designed based on the partial amino acid sequences of the isolated porcine parotid hormone. Over 100 positive clones were isolated, and the complete nucleotide sequences were determined for 3 cDNA clones, D_{2A}, T{P}_{23} and S{L}_{44.1}. The sequence analysis indicated that these three cDNA clones contain sequences encoding proteins of 48.5 kDa, 53.2 kDa, and 62.3 kDa, respectively. These proteins are unusually high in proline (51.8-60.7%) and glycine (15.6-17.3%), and contain considerably high amounts of arginine and lysine (5.6-7.4%).

The cDNA sequences exhibit four general domains characteristic for proline-rich proteins: a signal peptide region, a transition region, a repeat region, and a carboxyl-terminal region. The most striking feature of these proteins is that the repeat region contains 12-26 tandemly arranged repeats of 21 amino acids with the prototype sequence of PPPGPPPGPPPGPAPPGAR. A Genbank data search revealed that the 5' untranslated sequences and regions encoding the signal peptides of these three proteins are homologous (90% identity) to other known salivary proline rich proteins. The basic amino acid content and typical BstN I sites in the repeats suggest that these cDNAs are part of basic, proline rich protein multigene family. In addition, the available partial amino acid sequences from other positive cDNA clones indicated that they contain very similar if not identical sequences to these three clones. The apparent variances detected by restriction analysis are probably due to the length polymorphism in the repeat regions.
The last 30 amino acids encoded by TP_{23} or SL_{44-1} have 100% identity with the isolated parotid hormone sequence, suggesting that the isolated PH was probably derived from the end of these cDNAs.

Northern analysis of porcine parotid tissue with porcine cDNA probe revealed five major mRNA size classes, ranging from 890 bases to 3550 bases. Hybridization of several mRNA species to the same cDNA probe suggests the existence of homologous sequences in the mRNAs. *In situ* hybridization and immunohistochemical staining of porcine tissues indicated that the acinar cells of the parotid glands are the primary location for both parotid hormone related mRNAs and the translation products. The hybridization signals and immunohistochemical staining were also seen in a few serous demilune cells in the sublingual glands, although the Northern analysis of the same tissue showed no signals. Submandibular glands, and tissues from liver, pancreas, small intestine, hypothalamus, pituitary, trachea, lungs, kidney and skeletal muscles did not show hybridization to PH probes or immunoreaction to PH antisera. Southern analysis of genomic DNA from human, rat and porcine demonstrated no cross hybridization with human and rat genes, suggesting that these porcine basic proline rich proteins are considerably divergent from human and rat BPRPs.

Using the TNT transcription and translation system, the cloned porcine cDNA (D_{2}A) produced a protein of approximately 90 kDa, which was recognized by the primary antibodies made against porcine parotid hormone. The expression of TP_{23} in COS-7 cells resulted a major product at 110 kDa. Interestingly, this protein was secreted into the culture medium. There were multiple smaller proteins detected in the cell extracts and in the culture medium. These findings suggest that post-translational modifications and/or
The last 30 amino acids encoded by TP_{23} or SL_{44-1} have 100% identity with the isolated parotid hormone sequence, suggesting that the isolated PH was probably derived from the end of these cDNAs.

Northern analysis of porcine parotid tissue with porcine cDNA probe revealed five major mRNA size classes, ranging from 890 bases to 3550 bases. Hybridization of several mRNA species to the same cDNA probe suggests the existence of homologous sequences in the mRNAs. In situ hybridization and immunohistochemical staining of porcine tissues indicated that the acinar cells of the parotid glands are the primary location for both parotid hormone related mRNAs and the translation products. The hybridization signals and immunohistochemical staining were also seen in a few serous demilune cells in the sublingual glands, although the Northern analysis of the same tissue showed no signals. Submandibular glands, and tissues from liver, pancreas, small intestine, hypothalamus, pituitary, trachea, lungs, kidney and skeletal muscles did not show hybridization to PH probes or immunoreaction to PH antisera. Southern analysis of genomic DNA from human, rat and porcine demonstrated no cross hybridization with human and rat genes, suggesting that these porcine basic proline rich proteins are considerably divergent from human and rat BPRPs.

Using the TNT transcription and translation system, the cloned porcine cDNA (D_{2A}) produced a protein of approximately 90 kDa, which was recognized by the primary antibodies made against porcine parotid hormone. The expression of TP_{23} in COS-7 cells resulted a major product at 110 kDa. Interestingly, this protein was secreted into the culture medium. There were multiple smaller proteins detected in the cell extracts and in the culture medium. These findings suggest that post-translational modifications and/or
The last 30 amino acids encoded by TP_{23} or SL_{44,1} have 100% identity with the isolated parotid hormone sequence, suggesting that the isolated PH was probably derived from the end of these cDNAs.

Northern analysis of porcine parotid tissue with porcine cDNA probe revealed five major mRNA size classes, ranging from 890 bases to 3550 bases. Hybridization of several mRNA species to the same cDNA probe suggests the existence of homologous sequences in the mRNAs. *In situ* hybridization and immunohistochemical staining of porcine tissues indicated that the acinar cells of the parotid glands are the primary location for both parotid hormone related mRNAs and the translation products. The hybridization signals and immunohistochemical staining were also seen in a few serous demilune cells in the sublingual glands, although the Northern analysis of the same tissue showed no signals. Submandibular glands, and tissues from liver, pancreas, small intestine, hypothalamus, pituitary, trachea, lungs, kidney and skeletal muscles did not show hybridization to PH probes or immunoreaction to PH antisera. Southern analysis of genomic DNA from human, rat and porcine demonstrated no cross hybridization with human and rat genes, suggesting that these porcine basic proline rich proteins are considerably divergent from human and rat BPRPs.

Using the TNT transcription and translation system, the cloned porcine cDNA (D_{2A}) produced a protein of approximately 90 kDa, which was recognized by the primary antibodies made against porcine parotid hormone. The expression of TP_{23} in COS-7 cells resulted a major product at 110 kDa. Interestingly, this protein was secreted into the culture medium. There were multiple smaller proteins detected in the cell extracts and in the culture medium. These findings suggest that post-translational modifications and/or
post-secretory processing might be responsible for generating the smaller proteins. Both the in vitro synthesized as well as the in vivo expressed proteins exhibited aberrant migrations on SDS PAGE, possibly due to their high proline content.

In summary, several cDNA clones were isolated from the porcine parotid cDNA library which encode 30 amino acids identical to the isolated porcine parotid hormone. However, there were several mRNA species sharing homology with each other, suggesting that they were derived from closely related genes – a multigene family. In situ hybridization and immunohistochemistry indicated that the acinar cells in parotid glands are the primary location for both parotid hormone related mRNA synthesis and that of the proteins. In both coupled transcription/translation system and COS-7 cells, cloned PH containing cDNAs yielded proteins which were recognized by anti-PH antibodies. These proteins demonstrated slower migration on SDS PAGE.

B. Future Studies

The cloning of cDNAs for porcine parotid hormone has raised many intriguing questions for further studies. The most important question is whether the 30 amino acid segment encoded by the end of these cDNA clones has parotid hormone activities. To answer this question, the 3’ end of the cDNA clones which encode for the first 30 amino acid sequence of isolated parotid hormone may be engineered into an expression vector, followed by expression in mammalian cells or in cell free system. The newly synthesized proteins can be studied in animal models to examine their dentinal fluid stimulating activity using established procedures.
It would also be interesting to find out the mechanism which leads to the processing of these proteins. The cloned cDNAs can be expressed in the TNT system, and the co-translational processing and post-translational modifications can be investigated using canine pancreatic microsomes. The cDNA clones can also be expressed in different mammalian cell lines, including parotid acinar cell line, to see whether different cell types trigger different expression and processing patterns. The specific tag(s) can be linked to the N- and/or C-terminal of the cDNA sequence, and antibodies specific to the tag(s) can be used to aid the analysis. The expression can also be carried out in the presence of protease inhibitor(s) to find out if the expressed protein is self-cleaved, or cleaved by cellular proteases. Mutations can also be introduced into the putative cleavage sites to determine whether these sites are critical for the processing of the protein. Electron microscopic immunocytochemistry can be performed to examine the subcellular localization of parotid hormone that would provide valuable information regarding the secretory processing of the hormone.

At genomic DNA level, the chromosomal walking can be used to obtain gene sequences 5' upstream to study the gene structure and its regulation. The complete PH gene sequences can be obtained by screening porcine genomic library with the cDNA as probes and the sequences can be analyzed and compared to all other known genes for proline rich proteins.

If the functional study gives promising results, the counter part gene(s) in human, rat and mouse can be also explored. Since these porcine cDNAs contain considerable differences compared to any other genes in human and rat, the most closely related
protein(s) would then be tested in animal models to investigate their dentinal fluid stimulating activity.
CHAPTER V

REFERENCES


