INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
A study of the guinea pig relaxin gene(s)

Lee, Yee Andrew, Ph.D.

University of Hawaii, 1991
A STUDY OF THE GUINEA PIG RELAXIN GENE(S)

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF
THE UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN BIOMEDICAL SCIENCES
(ANATOMY AND REPRODUCTIVE BIOLOGY)

AUGUST 1991

BY
YEE LEE

DISSE ss sion COMMITTEE:
Gillian D. Bryant-Greenwood Chairperson
Frederick C. Greenwood
Vincent J. Defeo
Ronald Hammer
Morton Mandel
Dorsey Stuart
THIS DISSERTATION IS DEDICATED TO

MY FATHER AND MY DAUGHTER
ACKNOWLEDGEMENTS

Special acknowledgments are extended to the following who help me to make this dissertation study a reality: -

- To Mary Ellen-Grace, Theresa Li-Lum, Yao De Zhi, Dr. Wang Yi-Fei and Dr. Ruth Kleinfeld for sponsoring, encouraging and helping me to come to United States to pursue the Doctor Degree of Philosophy in Biomedical Science.

- To my advisor Dr. Gillian D. Bryant-Greenwood for her excellent guidance, support and criticism through this critical period of my career. My thanks especially for suggesting a challenging and rewarding project giving such a wonderful opportunity to practice many molecular techniques. It would not have been possible to complete this dissertation study without her understanding and helping.

- To Dr. Frederick C Greenwood for his continual guidance, encouragement and understanding, and particularly for his strong financial support and help in improving my second language. It would also not have been possible to reach goal without his aid.

- To Dr. Morton Mandel and Dr Dorsey Stuart for sharing their experience in molecular biology and for their help with experimental materials and protocols.
- To Dr. Vincent J. Defeo and Dr. Ronald Hammer for their encouragements and for their guidance as the members of my dissertation committee.
- To Dr. Lily Tashima and Dr. Deborah Hansell for their helpful discussions and gentle criticism.
- To my colleagues in the Molecular Endocrinology group for their day-to-day support and cooperation. I wish to thank Mrs. Sandy Y. Yamamoto for her assistance.
- To Dr. R.K. Craig of Imperial College, London for the lactating guinea pig mammary gland cDNA library, Dr. Simon Kwok of the University of Kansas Medical Center for the pig preprorelaxin cDNA probe and Dr. M.S. Soloff of the Medical College of Ohio, Toledo, OH for the rat preprorelaxin cDNA probe.
- To Dr. Neil S. Reimer and his colleagues in Biotechnology Molecular Biology Instrumentation and Training Facility for providing oligonucleotide primers.
- To Mrs. Judy Horio in the Department of Anatomy and Reproductive Biology and Mrs. Andrea Matsuda in the Pacific Biomedical Research Center for their generous support.
- To My aunt Yu Hua Chang, my brother, sister and their family for their warm encouragement and financial support to help me come to the United States.
I wish to record the followings:-

- My special thanks to my wife Chien Lee and my son Matthew Zuo Lee for their understanding, encouragement and support throughout these five hard but productive years.

- My special appreciation to my father and mother for their continuous encouragement and support throughout my life. Especially for my father who gave me his precious advice and warm encouragement till the last minute of his life.

- My special memory of and acknowledgement to my daughter Yin Lee for her sacrifice to my scientific career.

Finally, I would like to dedicate this dissertation to my father who died of cancer in 1990, and to my daughter Yin who died of the inability of medicine in 1984.
ABSTRACT

This dissertation study was designed to elucidate the nucleotide sequence of the guinea pig relaxin gene and hence to derive the amino acid sequence of guinea pig preprorelaxin, to show how many relaxin gene(s) are present in the guinea pig genome, and to study the transcription of a relaxin gene in the guinea pig endometrium during the reproductive cycle.

A lactating guinea pig mammary gland cDNA library was screened with a radioactive full-length rat preprorelaxin cDNA probe. Seven positive clones were identified. The characterization of the inserts with the PCR (Polymerase Chain Reaction) technique and Southern hybridization suggested that they are truncated molecules. In order to obtain general information of the guinea pig relaxin gene, the clone containing the longest insert (approximately 600 bps) was submitted to sequence analysis. Computer-assisted sequencing data analysis show that this insert was similar to the mRNA sequence of the pig preprorelaxin. The conclusion from this study was that the guinea pig relaxin gene sequence is similar to that of the pig relaxin gene.

Based on the preliminary information from these studies, several sets of oligonucleotide primers were selected from different regions of the mRNA sequence of pig preprorelaxin. These synthetic primers were used to screen a cDNA "pool" prepared from the guinea pig endometrium in late pregnancy by a PCR technique. One set of these PCR primers successfully
amplified a relaxin related cDNA fragment (286 bps). Sequence analysis of this PCR product confirmed that it encodes an intact B chain of the guinea pig preprorelaxin plus part of the signal peptide and C peptide of this molecule. The rest part of the guinea pig endometrial relaxin gene was elucidated by a RACE-PCR and subcloning strategy. This is the first report of any part of the guinea pig relaxin gene sequence. Availability of this sequence allowed a study of the physiology of guinea pig relaxin.

The second question of this dissertation was studied by a Southern analysis of guinea pig genomic DNA digests with a guinea pig relaxin specific cDNA probe. It was shown that there are two different relaxin genes in the guinea pig genome, a situation similar to the human genome, but different to the genomes of other mammals studied thus far. One relaxin gene has been shown to be expressed in the guinea pig endometrium in this dissertation using PCR and direct sequencing. Whether, or where, the second relaxin gene is expressed needs further investigation.

The availability of the guinea pig relaxin gene sequence allowed a more convincing study of the expression of this gene in the guinea pig endometrium. Guinea pig endometrial mRNAs from different stages of the reproductive cycle were hybridized with the radioactive guinea pig relaxin cDNA probe. Northern analyses obtained from this study showed that the transcription pattern of the relaxin gene in this organ is
maximal in the late pregnancy, minimal during mid pregnancy, appears in the estrous cycle and disappears rapidly after parturition. This is the first study of the transcription of the relaxin gene in the guinea pig endometrium during the reproductive cycle with a species-specific cDNA probe. It provides proof that relaxin is indeed synthesized in the guinea pig endometrium, not sequestered from other sources. The loss of transcription activity of the relaxin gene, in the endometrium, during lactation indicates that there are other sites of relaxin synthesis as sources of the plasma relaxin found in lactation.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGMENTS</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xvi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xvii</td>
</tr>
<tr>
<td>CHAPTER ONE. INTRODUCTION AND BACKGROUND</td>
<td></td>
</tr>
<tr>
<td>I. THE EUKARYOTIC GENOME AND THE PROTEIN-CODING</td>
<td></td>
</tr>
<tr>
<td>GENE: STRUCTURE AND EXPRESSION</td>
<td></td>
</tr>
<tr>
<td>1. THE EUKAROTIC AND THE PROTEIN-CODING GENE.........</td>
<td>2</td>
</tr>
<tr>
<td>2. THE EXPRESSION OF A PROTEIN-CODING GENE............</td>
<td>9</td>
</tr>
<tr>
<td>II. THE INSULIN FAMILY OF HORMONES.....................</td>
<td>17</td>
</tr>
<tr>
<td>III. THE RELAXIN GENES AND RELAXINS IN DIFFERENT</td>
<td></td>
</tr>
<tr>
<td>SPECIES</td>
<td></td>
</tr>
<tr>
<td>1. THE RELAXIN GENES</td>
<td>24</td>
</tr>
<tr>
<td>2. RELAXINS</td>
<td>31</td>
</tr>
<tr>
<td>IV. RELAXIN AND THE GUINEA PIG ENDOMETRIUM............</td>
<td>35</td>
</tr>
<tr>
<td>V. RELAXIN AND THE MAMMARY GLAND</td>
<td></td>
</tr>
<tr>
<td>1. THE ACTION OF RELAXIN ON THE MAMMARY GLAND.........</td>
<td>38</td>
</tr>
<tr>
<td>1.1. RELAXIN STIMULATES THE GROWTH OF THE MAMMARY</td>
<td></td>
</tr>
<tr>
<td>PARENCHYMA</td>
<td>40</td>
</tr>
<tr>
<td>1.2. RELAXIN STIMULATES THE GROWTH OF THE MAMMARY</td>
<td></td>
</tr>
<tr>
<td>STROMA</td>
<td>41</td>
</tr>
<tr>
<td>2. THE MAMMARY GLAND AS A SITE OF RELAXIN SYNTHESIS</td>
<td>43</td>
</tr>
</tbody>
</table>
VI. STATEMENT OF THE PROBLEMS

1. THE SEQUENCE ANALYSIS OF A GUINEA PIG RELAXIN.................................45
2. AN ANALYSIS OF THE RELAXIN GENE(S) IN THE GUINEA PIG GENOME.............48
3. THE TRANSCRIPTION OF A RELAXIN GENE IN THE GUINEA PIG ENDOMETRIUM DURING THE REPRODUCTIVE CYCLE.................................48

CHAPTER TWO. MATERIALS AND METHODS

I. MATERIALS AND MISCELLANEOUS CHEMICALS..............52
II. SOLUTIONS AND BUFFERS........................................56
III. METHODS

A. AN ATTEMPT TO CLONE AND SEQUENCE A RELAXIN GENE TRANSCRIBED IN THE MAMMARY GLAND OF THE GUINEA PIG

1. PREPARATION AND LABELING THE RAT PREPRORELAXIN cDNA PROBE

1.1. INTRODUCTION OF PLASMID (pRLX 9) INTO E. coli NM522.................................62
1.2. ISOLATION, PURIFICATION AND CONCENTRATION OF pRLX 9..............................63
1.3. DIGESTION OF pRLX 9 WITH EcoRI AND HindIII........................................65
1.4. RESOLUTION AND RECOVERY OF THE RAT PREPRORELAXIN cDNA INSERT FROM AN AGAROSE GEL....................................................66
1.5. RADIOACTIVELY LABELING THE RAT PREPRORELAXIN cDNA PROBE BY USING RANDOM PRIMER DNA LABELING METHOD..............................67

2. SCREENING THE LACTATING GUINEA PIG MAMMARY GLAND LIBRARY BY USING RAT PREPRORELAXIN cDNA AS A MOLECULAR PROBE
2.1. PLATING AND TRANSFERRING THE LAMBDA gt 11 LIBRARY.................................68
2.2. HYBRIDIZATION WITH THE RADIOACTIVE RAT PREPRORELAXIN cDNA PROBE.................69

3. PURIFICATION OF PHAGE PARTICLES CONTAINING GUINEA PIG RELAXIN INSERTS......................71

4. GROWING UP RELAXIN SPECIFIC CLONES AND MAKING HIGH-TITER STOCKS.........................72

5. PREPARATION OF DNA TEMPLATE FOR SEQUENCING BY POLYMERASE CHAIN REACTION (PCR)............73

6. DIDEOXYNUCLEOTIDE CHAIN TERMINATION SEQUENCING
6.1. SEQUENCING REACTION........................................74
6.2. DENATURING POLYACRYLAMIDE SEQUENCING GELS........................................75
6.3. DRYING THE SEQUENCING GELS AND EXPOSURE TO AN X-RAY FILM..............................77

7. COMPUTER-ASSISTED SEQUENCING DATA ANALYSIS....78
B. PARTIAL SEQUENCE ANALYSIS OF A RELAXIN GENE
    TRANSCRIBED IN THE GUINEA PIG ENDOMETRIUM
    BY PCR
1. PREPARATION OF POLY(A)+ RNA FROM THE LATE
    PREGNANT GUINEA PIG ENDOMETRIUM..............79
2. SINGLE-STRANDED cDNA SYNTHESIS.................80
3. PRIMER DESIGN.................................81
4. DOUBLE-STRANDED AMPLIFICATION OF RELAXIN
    SPECIFIC FRAGMENTS.............................82
5. SOUTHERN VERIFICATION OF DOUBLE-STRANDED
    PCR PRODUCTS
5.1. PREPARATION OF A FULL-LENGTH PIG
        PREPROREALXIN cDNA PROBE....................83
5.2. SOUTHERN BLOTTING AND HYBRIDIZATION.........83
6. ISOLATION OF THE RELAXIN SPECIFIC FRAGMENT
    FROM A NUSIEV AGAROSE GEL......................84
7. PREPARATION OF SINGLE-STRANDED SEQUENCING
    TEMPLATES BY AN ASYMMETRIC PCR STRATEGY......85
8. DIDEOXY SEQUENCING AND COMPUTER-ASSISTED
    SEQUENCING DATA ANALYSIS.......................86
C. AN ANALYSIS OF THE RELAXIN GENE(S) IN THE
    GUINEA PIG GENOME
1. PREPARATION OF GENOMIC DNA FROM GUINEA PIG
    LIVER............................................86
2. DIGESTION OF GENOMIC DNAs WITH RESTRICTION
    ENDONULCEASES.................................87
3. SOUTHERN BLOTTING AND HYBRIDIZATION.............87

D. THE TRANSCRIPTION OF A RELAXIN GENE IN THE
   GUINEA PIG ENDOMETRIUM DURING THE REPRODUCTIVE
   CYCLE
   1. RADIOACTIVELY LABELING 3' TERMINI OF RELAXIN
      SPECIFIC PROBES WITH TdT AND ALPHA
      P-32 dATP........................................89
   2. POLY(A)+ RNA EXTRACTION........................90
   3. NORTHERN BLOTTING AND HYBRIDIZATION.........91

CHAPTER THREE. RESULTS
   I. AN ATTEMPT TO CLONE AND SEQUENCE A RELAXIN
      GENE TRANSCRIBED IN THE MAMMARY GLAND OF THE
      GUINEA PIG........................................94
   II. PARTIAL SEQUENCE ANALYSIS OF A RELAXIN GENE
       TRANSCRIBED IN THE GUINEA PIG ENDOMETRIUM BY
       PCR.............................................107
   III. AN ANALYSIS OF THE RELAXIN GENE(S) IN THE
        GUINEA PIG GENOME..............................120
   IV. THE TRANSCRIPTION OF A RELAXIN GENE IN
       THE GUINEA PIG ENDOMETRIUM DURING THE
       REPRODUCTIVE CYCLE.............................123

CHAPTER FOUR. DISCUSSION
   I. AN ATTEMPT TO CLONE AND SEQUENCE A RELAXIN GENE
      TRANSCRIBED IN THE MAMMARY GLAND OF THE GUINEA
      PIG.............................................128
II. PARTIAL SEQUENCE ANALYSIS OF A RELAXIN GENE
TRANSCRIBED IN THE GUINEA PIG ENDOMETRIUM
BY PCR.................................................................134

III. AN ANALYSIS OF THE RELAXIN GENE(S) IN THE GUINEA
PIG GENOME............................................................142

IV. THE TRANSCRIPTION OF A RELAXIN GENE IN THE
GUINEA PIG ENDOMETRIUM DURING THE REPRODUCTIVE
CYCLE.................................................................146

V. SUMMARY AND CONCLUSION..................................151

APPENDIX. AMPLIFICATION OF THE 3'END OF THE GUINEA PIG
ENDOMETRIAL RELAXIN GENE BY A RACE-PCR
STRATEGY..............................................................153

BIBLIOGRAPHY........................................................169
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Classes of eukaryotic DNAs.</td>
<td>3</td>
</tr>
<tr>
<td>2. The homology between amino acid sequences of insulin in different species</td>
<td>20</td>
</tr>
<tr>
<td>3. The homology between amino acid sequences of IGF-I in different species</td>
<td>21</td>
</tr>
<tr>
<td>4. The homology between amino acid sequences of IGF-II in different species</td>
<td>22</td>
</tr>
<tr>
<td>5. The homology between amino acid sequences of relaxin in different species</td>
<td>23</td>
</tr>
<tr>
<td>6. Summary of characteristics of highly purified relaxin preparations</td>
<td>25</td>
</tr>
<tr>
<td>7. Homologies of the relaxin genes in different species</td>
<td>31</td>
</tr>
<tr>
<td>8. Immunostaining pattern of relaxin in the guinea pig endometrium</td>
<td>37</td>
</tr>
<tr>
<td>9. The effects of relaxin on the mammary growth</td>
<td>39</td>
</tr>
<tr>
<td>10. Comparison of the immunostaining patterns of relaxin in the guinea pig mammary gland and endometrium</td>
<td>44</td>
</tr>
<tr>
<td>11. Primers used in amplifying and sequencing of the guinea pig preprorelaxin by PCR</td>
<td>82</td>
</tr>
<tr>
<td>12. The size of the insert of seven positive clones</td>
<td>99</td>
</tr>
<tr>
<td>13. Comparison of the partial nucleotide sequence and its predicted amino acid sequence of guinea pig relaxin gene with the corresponding region of this hormone in different species</td>
<td>137</td>
</tr>
<tr>
<td>14. Semi-quantitative evaluation of expression of the relaxin gene in the guinea pig endometrium during the reproductive cycle in comparison to relaxin peptide detected</td>
<td>150</td>
</tr>
<tr>
<td>15. Oligonucleotide primers used to amplify the 3'end of the guinea pig relaxin gene</td>
<td>154</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Schematic representation of a typical protein coding transcription unit in eukaryotes</td>
<td>5</td>
</tr>
<tr>
<td>2. Schematic representation of a typical eukaryotic gene promoter</td>
<td>7</td>
</tr>
<tr>
<td>3. Flow chart of the expression of a eukaryotic gene (human relaxin gene)</td>
<td>11</td>
</tr>
<tr>
<td>4. Schematic representation of the mechanism of the splicing of a mRNA precursor</td>
<td>15</td>
</tr>
<tr>
<td>5. The overall homologies of insulin family members at the amino acid level</td>
<td>18</td>
</tr>
<tr>
<td>6. Comparison of amino acid and mRNA sequences of pig and rat preprorelaxin</td>
<td>27</td>
</tr>
<tr>
<td>7. Amino acid sequences of relaxin in different species</td>
<td>33</td>
</tr>
<tr>
<td>8. A strategy for sequencing a guinea pig relaxin gene</td>
<td>46</td>
</tr>
<tr>
<td>9. A strategy for the determination of the number of relaxin gene(s) in the guinea pig genome</td>
<td>49</td>
</tr>
<tr>
<td>10. The rat preprorelaxin cDNA insert identified by the double digestion of EcoRI and HindIII</td>
<td>95</td>
</tr>
<tr>
<td>11. Guinea pig relaxin specific clones identified by the P-32 labeled rat preprorelaxin cDNA probe</td>
<td>97</td>
</tr>
<tr>
<td>12. Partial nucleotide sequence (350 bps) of the guinea pig relaxin positive clone #1</td>
<td>100</td>
</tr>
<tr>
<td>13. Comparison of the 9 amino acid segment with the end of the A chain of relaxins in different species</td>
<td>103</td>
</tr>
<tr>
<td>14. Comparison of the 5'end of the 350 bps fragment (a total of 27 bps) with the mRNA sequence of pig and rat relaxins</td>
<td>105</td>
</tr>
<tr>
<td>15. Relaxin specific fragments amplified from the cDNA pool prepared from the endometrial tissue of late pregnant guinea pig using PCR</td>
<td>108</td>
</tr>
</tbody>
</table>
16. The double-stranded PCR products shown in figure 13 reamplified by asymmetric PCR..........................110

17. The dideoxy sequencing of the single-stranded PCR products showed in figure 14..........................112

18. Comparison of the nucleotide sequence of the guinea pig preprorelaxin fragment (226 bps) with the corresponding region of the pig preprorelaxin........114

19. Partial nucleotide sequence (226 bps) and the predicated amino acid sequence of the relaxin gene transcribed in the endometrium of the late pregnant guinea pig..................................................116

20. Comparison of the amino acid sequence of the guinea pig preprorelaxin fragment (74 amino acid residues) with the corresponding region of pig, rat, human H1 and human H2 preprorelaxin.................................118

21. Southern analysis of human and guinea pig genomic DNAs digested with three different restriction endonucleases.................................121

22. Northern analysis of the endometrial tissues of the guinea pig during the reproductive cycle........125

23. The probable location of the 350 bps fragment in the guinea pig relaxin gene.................................131

24. Schematic representation of the RACE-PCR protocol....160

25. Amplification of the 3' end of the guinea pig relaxin gene by a RACE-PCR strategy......................163

26. Southern verification of the RACE-PCR products with a P-32 labeled internal oligonucleotide probe...165

27 The nucleotide sequence and its derived amino acid sequence of the prorelaxin gene transcribed in the endometrium of the late pregnant guinea pig......167
CHAPTER ONE

INTRODUCTION AND BACKGROUND
I. THE EUKARYOTIC GENOME AND THE PROTEIN-CODING GENE: STRUCTURE AND EXPRESSION

1. The eukaryotic genome and the protein-coding gene

The eukaryotic genome is much more complex than that of the prokaryote. In addition to functional genes coding for proteins and RNAs, the eukaryotic genome contains many long DNA stretches that seem to be superfluous. This "extra" DNA consists of introns, pseudogenes, repetitious sequences and unclassified spacer DNA between the functional genes. It has been estimated that the protein-coding portion of the eukaryotic genome makes up only as low as 1% of total RNA of the cell (Darnell et al., 1986). The basic components of a eukaryotic genome are summarized in Table 1. So far, the function of most of the "extra" DNA in the eukaryotic genome is not clear. Introns are present in protein-coding genes in all of eukaryotic genomes, and are removed from protein-coding DNA sequences during mRNA processing. Pseudogenes were formally functional, but have lost some critical nucleotide sequences and have stop codons in all three reading frames (Oliver and Ward, 1985; Darnell et al., 1986). Some of the repetitious DNA may serve a useful function, such as satellite DNA and Alu sequences, but it is probably vastly more than is required (Alberts et al., 1986).

Finally, even if there are some unidentified functional genes in spacer DNA, much of this spacer DNA have been suggested to serve only to connect genes and complete the
integrity of chromosomes (Darnell et al., 1986).

Table 1
Classes of eukaryotic DNA

<table>
<thead>
<tr>
<th>Protein-coding genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solitary</td>
</tr>
<tr>
<td>Duplicated (most duplicates are no longer exact)</td>
</tr>
</tbody>
</table>

| RNA-coding genes (most are tandemly duplicated) |
| Repetitious DNA |
| Simple sequences |
| Dispersed intermediate repeats |
| (short, 150-300 bp; long, 5-7 kb) |

| Mobile genetic elements |
| Transposons |
| Reverse transcription copies |

| Unclassified spacer or connecting DNA |

(From: Darnell et al., 1986)

The concept of the gene has changed due to the development of molecular biology. According to the current molecular definition, a gene consists of all the DNA sequences necessary to produce a single RNA molecule or polypeptide product (Darnell et al., 1986).

The structure of the eukaryotic protein-coding gene has been shown to be fairly consistent (Figure 1). It includes
a promoter region, followed by a DNA sequence that is copied into a primary transcript. The structure of the eukaryotic promoter is more complex than the prokaryotic ones and is summarized in Figure 2. In nearly all of the eukaryotic promoters, a consensus sequence, TATAAAT, is identified about 25 bp upstream from the transcription start site. This TATA box is flanked by high G+C sequences. Another common sequence, GG(T/C)CAATCT (the CAAT box), is also found in many eukaryotic promoters in the -75 region. Some eukaryotic promoters need a third element which has a consensus sequence GGGCGG (the GC box). It is believed that these common sequences in the promoter region are sites of binding of transcription factors (Freifelder,1987). RNA polymerase II is strictly dependent on the presence of these bound transcription factors to initiate the mRNA synthesis and in addition, the TATA box probably involves the determination of the initiation site (Freifelder, 1987).
Figure 1

Schematic representation of a typical protein-coding transcription unit in eukaryotes. (From Darnell, 1986)
Flanking sequences.

Initiation site

Transcription unit

Termination site or region

Untranslated Poly A site

Intron

Untranslated

Exon 1

Exon 2

Promoter elements

Cap site

ATG (start of protein-coding sequence)

-500 to -100 (TATA)

5' 3'
Figure 2

Schematic representation of a typical eukaryotic gene promoter. Only the TATA box is represented in almost all promoters. The CAAT box occurs much less frequently, and the GC box has only occasionally been observed. An upstream site is very common but is not considered to be part of the promoter. (From Freilfelder, 1987).
8

rnRNA,

100 base pairs

Upstream

GC

CAAT

TATA

mRNA

3'

5'
The protein-coding region of a eukaryotic gene may be segmented by one or more introns, as showed in Figure 1. The first exon contains a 5'-untranslated region and the final exon includes a 3'-untranslated region extending to a poly A addition site. The function of the DNA sequences that flank such a transcription unit is not clear. Many of these DNA stretches can not encode a peptide because they have multiple translation stop codons in all three reading frames. Upstream elements are other types of DNA sequences which can influence the transcription activity of a protein-coding gene greatly when they are occupied by specific factors. A predominant feature of these upstream elements is that they are usually far away from the initiation site (several hundred nucleotide pairs) and their precise position is not important for their functioning. Among these upstream elements, an enhancer is a most unusual one: a) it can stimulate the rate of transcription nearly 100-fold, b) it may not only be moved to various distances from the promoter region of a gene but can be inverted or even moved downstream from the promoter, without loss of enhancing activity (Darnell, 1986).

2. The expression of a protein-coding gene

The expression of a protein-coding gene in eukaryotes is controlled at several levels: a) transcription of a specific gene to form a primary transcription product, b) processing the primary transcription product and transporting the mature
product (mRNA) into the cytoplasm, c) translation of the mRNA into a primary protein product, and d) posttranslational modification of the primary protein product to form a final protein molecule which has a specific biological activity (Figure 3).

Gene expression starts by the enzymatic synthesis of an RNA molecule that is a copy of one strand of a DNA corresponding to a specific gene.
Figure 3

Flow chart of the expression of a eukaryotic gene (Human relaxin gene). (From Kemp and Niall, 1984)
DNA

Transcribed pre-mRNA

Processing

mRNA

Translation

Preprorelaxin

Posttranslational processing

Prorelaxin

Processing

Ser Leu

(Lys)

Gln Cys

(Arg)

Human relaxin(s)
In eukaryotes, some proteins are essential for the initiation of gene expression. These proteins bind to the promoter region of a protein-coding gene and enable RNA polymerase II to reach the initiation site. The exact mechanism of the transcription initiation and termination of a eukaryotic protein-coding gene is not clear at present (Freifelder, 1987). Regulation of eukaryotic gene expression has been shown by steroid hormones and growth factors (Darnell, 1986).

The product of transcription is an RNA molecule called the primary transcript. An eukaryotic primary transcript varies greatly in size, from shorter than 2 kb to longer than 30 kb (Darnell, 1990). It is believed that a primary transcript has to be processed in the nucleus by: a) the addition of a 5' end cap which plays a role in the translation initiation, b) the addition of a 3'end of poly A tail which has an effect on stabilizing mRNA molecule and by c) the modification of most primary transcripts by splicing. In this latter process, one or more intervening RNA sequences (introns) are removed and the remaining pieces (exons) are joined to form a mature messenger RNA. It is estimated that, in mammalian cells, over 70 % of a primary transcript molecule is intron and is removed during the splicing (Darnell, 1986). The mechanism of splicing of a primary transcript is unclear at the moment. A consensus sequence can be found at the ends of a typical intron: (5')GU...AG(3'). It has been suggested
that the exon-intron junctions are brought together for
correct cutting by a small ribonucleoprotein complex
containing U 1 RNA, a 165 bp nuclear RNA (Freifelder, 1987).
The U 1 RNA has been found to have a continuous sequence that
is highly homologous to the intron sequences at most splicing
junctions (Freifelder, 1987). A current model of the splicing
of a primary transcript hypothesized by Albert (1982) is shown
in Figure 4.

Once completed, mature messenger RNA quickly enters the
cytoplasm through the nuclear pores and directs the synthesis
of a protein. In some cases, the protein product must undergo
a series of posttranslational modifications to a functional
product having a specific biological activity.
Figure 4

Schematic representation of the mechanism of the splicing of a mRNA precursor. (From Alberts, 1982)
RNA corresponding to intron
excised intron sequence

primary RNA transcript
5' ≈ AGGUAAAGU ≈ 3'

RNA
5' ≈ AGGUAAAGU ≈ 3'
cut and rejoin at donor junction

RNA corresponding to intron
spliced messenger RNA with intron sequence removed

RNA
protein
small ribonucleoprotein particle [U1]

RNA
5' ≈ A G G ≈ 3'
cut and rejoin at acceptor junction
II. THE INSULIN FAMILY OF HORMONES

The insulin family of hormones is composed of several members: insulin, insulin-like growth factor (IGF) I and II, relaxin and nerve growth factor. It is believed that these peptide hormones have arisen by a mechanism of gene duplication (Bolander, 1989). In general, the members of the insulin family share a basic molecular structure. The precursor of mature hormones is a single chain polypeptide of four portions, signal peptide, b chain, connecting peptide and c chain. During the posttranslational modification, the signal peptide and connecting peptide are removed. The remaining portions of the molecule, b chain and a chain, are assembled by two interchain disulfide bounds and one intrachain disulfide bridge to form a functional hormone.

The overall homology of insulin family members at the amino acid level is summarized in Figure 5. The homology of insulins in different species is summarized in Table 2. The homology of IGF-I in different species is summarized in Table 3. The homology of IGF-II in different species is summarized in Table 4. The homology of relaxins in different species is presented in Table 5.
Figure 5

The overall homologies of insulin family members at amino acid level.
IGF-I \rightarrow 64\% \Rightarrow \text{Insulin} \Rightarrow 50\% \Rightarrow \text{IGF-II}

\text{IGF-I} \leftarrow 50\% \Rightarrow \text{Insulin} \Rightarrow 50\% \Rightarrow \text{IGF-II}

20-25\% \downarrow \quad 20-25\% \downarrow \quad 20-25\% \downarrow

\text{Relaxin} \quad \text{Relaxin} \quad \text{Relaxin}
Table 2
The homology between amino acid sequences of insulin in different species

<table>
<thead>
<tr>
<th></th>
<th>G. Pig</th>
<th>R &amp; M1</th>
<th>R &amp; M2</th>
<th>Pig</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. Pig</td>
<td>100</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>R &amp; M1</td>
<td>67</td>
<td>100</td>
<td>96</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>R &amp; M2</td>
<td>67</td>
<td>96</td>
<td>100</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>Pig</td>
<td>67</td>
<td>92</td>
<td>92</td>
<td>100</td>
<td>98</td>
</tr>
</tbody>
</table>

Note: G. Pig = guinea pig, R = rat, M = mouse

Homology = %
Table 3

The homology (%) between amino acid sequences of IGF-I in different species

<table>
<thead>
<tr>
<th></th>
<th>Guinea Pig</th>
<th>Mouse</th>
<th>Rat</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guinea Pig</td>
<td>100</td>
<td>94</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>Mouse</td>
<td>94</td>
<td>100</td>
<td>99</td>
<td>94</td>
</tr>
<tr>
<td>Rat</td>
<td>96</td>
<td>99</td>
<td>100</td>
<td>96</td>
</tr>
</tbody>
</table>
Table 4

The homology (%) between amino acid sequences of IGF-II in different species

<table>
<thead>
<tr>
<th></th>
<th>Guinea pig</th>
<th>Rat</th>
<th>Mouse</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guinea pig</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Rat</td>
<td>NA</td>
<td>100</td>
<td>97</td>
<td>94</td>
</tr>
<tr>
<td>Mouse</td>
<td>NA</td>
<td>97</td>
<td>100</td>
<td>91</td>
</tr>
</tbody>
</table>

Note: NA = not available.
Table 5
The homology (%) between amino acid sequences of relaxin in different species

<table>
<thead>
<tr>
<th></th>
<th>Pig</th>
<th>Rat</th>
<th>Shark</th>
<th>Human 1</th>
<th>Human 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pig</td>
<td>100</td>
<td>40</td>
<td>40</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Rat</td>
<td>40</td>
<td>100</td>
<td>40</td>
<td>48</td>
<td>44</td>
</tr>
<tr>
<td>Shark</td>
<td>40</td>
<td>40</td>
<td>100</td>
<td>42</td>
<td>44</td>
</tr>
<tr>
<td>Human 1</td>
<td>46</td>
<td>48</td>
<td>42</td>
<td>100</td>
<td>77</td>
</tr>
</tbody>
</table>
III. THE RELAXIN GENES AND RELAXINS IN DIFFERENT SPECIES

The attitude towards relaxin changed after the techniques for isolating highly purified native hormone began to be available in the mid 1970's. Between 1974 and 1981, relaxin was isolated from a range of reproductive tissues obtained from different species. The characteristics of these preparations are summarized in Table 6. A further impetus has been the application of molecular techniques to the elucidation of the structure of relaxin genes from different species.

1. The relaxin genes

Studies of hormone specific genes are of value because they provide information on the structure of the hormone precursor and the hormone itself. At present, information on the relaxin genes is limited to the complete analysis of cDNA sequences of pig (Haley et al.,1982), rat (Hudson et al., 1981), rhesus monkey (Crawford et al.,1989), and the genes encoding for human relaxin H 1 and H 2 (Hudson et al., 1983 and 1984).
Table 6

Summary of characteristics of highly purified relaxin preparations

<table>
<thead>
<tr>
<th>Species (source)</th>
<th>Yield (mg/kg f.t.*)</th>
<th>Molecular weight</th>
<th>Isoelectric point</th>
<th>Bioactivity** (GPU/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porcine ovaries</td>
<td>38 (CMB) 6,300</td>
<td>10.6</td>
<td>2,5003,000(a)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>34 (CMA) 6,300</td>
<td>10.7</td>
<td>2,500-3,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36 (CMA') 6,300</td>
<td>10.8</td>
<td>2,500-3,000</td>
<td></td>
</tr>
<tr>
<td>Rat ovaries</td>
<td>140 (CM1) 6,000</td>
<td>7.6</td>
<td>CM1 equivalent to CM2 and relaxin(a)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>140 (CM2) 6,000</td>
<td>9.4</td>
<td>less than pig relaxin(c)</td>
<td></td>
</tr>
<tr>
<td>Shark ovaries</td>
<td>2-5 6,000</td>
<td>-</td>
<td>less than pig relaxin(c)</td>
<td></td>
</tr>
<tr>
<td>Rabbit placenta</td>
<td>10 7,200 6.8</td>
<td>23(b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horse placenta</td>
<td>1.5 (R1) 5.600</td>
<td>-</td>
<td>28(a)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2 (R2) -</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.4 (R3) -</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

* f.t. = fresh tissue

**
(a). Mouse interpubic ligament bioassay (Steinertz et al., 1960).
(b). In vitro mouse uterus bioassay (Kroc et al., 1959; Larkin et al., 1981).
(c). Guinea pig pubic symphysis palpation bioassay and in vitro guinea pig uterus bioassay (Steinz et al., 1969)

(Modified from: Sherwood, 1988)
The isolation of highly purified relaxin from the pig (Sherwood and O'Bryne, 1974) and rat (Sherwood, 1979) allowed the elucidation of their amino acid sequences and provided the basis for the synthesis of oligodeoxyribonucleotide primers. These in turn, were used to make specific cDNA probes for isolating relaxin specific clones from a varies of cDNA libraries on the reasonable assumption of conservation of gene structure in the different species. The sequence analysis of relaxin specific clones thus provided the structures of both the pig and rat relaxin genes (Figure 6). The coding region of the pig relaxin gene contains approximately 540 nucleotides coding for 180 amino acids. Of these, only 150 nucleotides correspond to 50 amino acids of the A and B chains of the relaxin molecule. The rest of the nucleotides are used to code a signal peptide of 24 amino acids and a large connecting peptide of 104 amino acids. The divergence of the amino acid sequence between pig and rat relaxin was well known before the elucidation of the cDNA sequences. However, at the nucleotide sequence level, the gene structures tend to be conserved. The nucleotide sequence homologies between pig and rat prorelaxin for the B, C and A chains are 63.0%, 66.8%, and 64.2% respectively. Like most eukaryotic gene families, the 3'end-untranslated regions of pig and rat relaxin gene show a considerably higher sequence divergence than the coding region.
Figure 6

Comparison of amino acid and mRNA sequences of pig (Top) and rat preprorelaxin (Bottom). The conserved sequences are boxed. (from Haley et al., 1982)
There are no significant homologies between 3'-untranslated regions of pig and rat relaxin genes, except for 10 bases following the termination codon and for the 25 bases around the poly A signal AAUAAUA sites. (Figure 6). The poly A signal sequence AAUAAUA is exactly the same in the pig and rat relaxin genes.

The limited availability of human ovaries prevented the isolation of human relaxin and direct amino acid sequence determination by conventional techniques. However, in a study of the pig and rat relaxin cDNA, a highly homologous region (71% at the nucleotide level) was identified in the portion of the C peptide (corresponding to amino acid 45-94 of the C peptide), and was presumed to be conserved in the relaxin genes of different species. A pig relaxin cDNA probe (147bp) constructed from this restriction fragment was used to screen a human genomic library (Hudson et al., 1983). Two relaxin specific lambda clones were identified on the basis of strong hybridization to this probe. Extensive sequence analysis proved that one of these two positive clones (lambda H7) contained the complete coding sequence for a relaxin polypeptide called human relaxin H1 (Hudson, 1983). The other clone was an incomplete version of the first. Southern analysis of human genomic DNA digests by a using specific probe corresponding to the H1 relaxin gene revealed that the human genome might contain a second relaxin gene (Hudson et al., 1984). A subsequent study strongly suggested that these
two genes are non-allelic and arose by gene duplication (Crawford et al., 1984). Rescreening of the genomic DNA library using the H1 relaxin gene specific probe confirmed the existence of the H2 relaxin gene. A study of a cDNA library prepared from ovarian tissue from a pregnant woman confirmed that the H2 relaxin gene is selectively transcribed in the human ovary during pregnancy (Hudson et al., 1984). However, these authors failed to demonstrate the transcription of H1 gene in the human ovary and other reproductive tissues. It was concluded therefore that the H1 gene was probably a pseudogene, or belonged to the human growth gene family whose site of tissue specific transcription remains unresolved (Seeburg, 1982). More recently, the H1 relaxin gene has been shown in our laboratory to be a functional gene transcribed in decidua, trophoblast and prostate (Hansell et al., 1991).

A radioactive human relaxin cDNA probe was used to screen a cDNA library prepared from the ovary of a pregnant rhesus monkey, and several positive clones containing an initiating methionine codon and a poly A tail were isolated (Crawford et al., 1989). The cDNA sequence analysis and amino acid sequence prediction demonstrated that the structure of the rhesus monkey relaxin is not strongly conserved when compared with the relaxin genes of the human (Crawford et al., 1989). Southern analysis of rhesus monkey genomic DNA digests by using a rhesus monkey relaxin specific probe showed that there
is only one relaxin gene in the rhesus monkey genome (Crawford et al, 1989).

The homologies of relaxin genes in different species are summarized in Table 7.

<table>
<thead>
<tr>
<th></th>
<th>pig</th>
<th>rat</th>
<th>human H1</th>
<th>human H2</th>
<th>rhesus monkey</th>
</tr>
</thead>
<tbody>
<tr>
<td>pig</td>
<td>100</td>
<td>63.1</td>
<td>62.3</td>
<td>73.5</td>
<td>72.2</td>
</tr>
<tr>
<td>rat</td>
<td>63.1</td>
<td>100</td>
<td>57.3</td>
<td>69.1</td>
<td>62.2</td>
</tr>
<tr>
<td>human H1</td>
<td>62.3</td>
<td>57.3</td>
<td>100</td>
<td>90.3</td>
<td>88.7</td>
</tr>
<tr>
<td>human H2</td>
<td>73.5</td>
<td>69.1</td>
<td>90.3</td>
<td>100</td>
<td>92.5</td>
</tr>
</tbody>
</table>

2. Relaxins

The amino acid sequence of relaxin has diverged considerably during evolution. Only 40 % to 48 % of the amino acid residues are conserved in pig, rat, shark and human relaxins (Sherwood, 1988). Among these conserved amino acid residues (Figure 7) are the six cystine residues (A-6, A-7, A-11, A-20, B-7 and B-19) which form the insulin-like disulfide bridge structure, two glycine residues (B-8 and B-20) which provide unique torsion angles for chain folding (Sherwood, 1988), three arginine residues (B-9, B-13 and
A-13) and three lysine residues (B-5, A-5 and A-14) which have been suggested to involve receptor binding (Bedarker et al., 1982; Bedarker et al., 1977, Dodson et al., 1982).

A possible three-dimensional structure for relaxin can be obtained by using coordinates derived from crystalline insulin. Despite strenuous attempts using sub-gram amount of porcine relaxin prepared in our laboratory, relaxin has defied crystallization and its three-dimensional structure is unknown. Based on the comparison of the amino acid sequences between pig relaxin and pig insulin and the three-dimensional structure of crystalline insulin, it seems that these two hormones share some structural features: a hydrophobic core, three right-handed helical segments which are located at N-terminus and C-terminus of the A chain and the mid-region of the B chain (Sherwood, 1988; Schwabe et al., 1990).
Figure 7

Amino acid sequences of relaxin in different species. Residues common to the five sequences are boxed, and those in positions comparable to those that contribute to the hydrophobic core of insulin are underlined. Numbers of residues postulated to be involved with the binding of relaxin to its receptor are starred. (From: Sherwood, 1988)
RELAXIN B CHAINS

<table>
<thead>
<tr>
<th></th>
<th>-7</th>
<th>-6</th>
<th>-5</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</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>
</tr>
</thead>
<tbody>
<tr>
<td>PIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S</td>
<td>T</td>
<td>N</td>
<td>D</td>
<td>F</td>
<td>I</td>
<td>K</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>R</td>
<td>E</td>
<td>L</td>
</tr>
<tr>
<td>RAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>V</td>
<td>S</td>
<td>E</td>
<td>E</td>
<td>W</td>
<td>M</td>
<td>D</td>
<td>Q</td>
<td>V</td>
<td>I</td>
<td>Q</td>
<td>C</td>
</tr>
<tr>
<td>HUMAN 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L</td>
<td>W</td>
<td>L</td>
<td>D</td>
<td>D</td>
<td>V</td>
<td>I</td>
<td>K</td>
<td>L</td>
<td>C</td>
<td>G</td>
<td>R</td>
<td>E</td>
</tr>
<tr>
<td>HUMAN 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td>S</td>
<td>W</td>
<td>M</td>
<td>E</td>
<td>E</td>
<td>V</td>
<td>I</td>
<td>K</td>
<td>L</td>
<td>C</td>
<td>G</td>
<td>R</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIG</td>
<td>L</td>
<td>W</td>
<td>V</td>
<td>E</td>
<td>I</td>
<td>C</td>
<td>G</td>
<td>S</td>
<td>V</td>
<td>S</td>
<td>W</td>
<td>G</td>
<td>R</td>
<td>T</td>
<td>A</td>
<td>L</td>
</tr>
<tr>
<td>RAT</td>
<td>A</td>
<td>W</td>
<td>I</td>
<td>E</td>
<td>V</td>
<td>C</td>
<td>G</td>
<td>A</td>
<td>S</td>
<td>V</td>
<td>G</td>
<td>R</td>
<td>L</td>
<td>A</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>HUMAN 1</td>
<td>A</td>
<td>Q</td>
<td>I</td>
<td>A</td>
<td>I</td>
<td>C</td>
<td>G</td>
<td>M</td>
<td>S</td>
<td>T</td>
<td>W</td>
<td>S</td>
<td>K</td>
<td>R</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>HUMAN 2</td>
<td>A</td>
<td>Q</td>
<td>I</td>
<td>A</td>
<td>I</td>
<td>C</td>
<td>G</td>
<td>M</td>
<td>S</td>
<td>T</td>
<td>W</td>
<td>S</td>
<td>K</td>
<td>R</td>
<td>S</td>
<td>L</td>
</tr>
</tbody>
</table>

RELAXIN A CHAINS

<table>
<thead>
<tr>
<th></th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</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>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HUMAN 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HUMAN 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIG</td>
<td>I</td>
<td>A</td>
<td>R</td>
<td>L</td>
<td>C</td>
</tr>
<tr>
<td>RAT</td>
<td>I</td>
<td>A</td>
<td>K</td>
<td>L</td>
<td>C</td>
</tr>
<tr>
<td>HUMAN 1</td>
<td>L</td>
<td>A</td>
<td>K</td>
<td>Y</td>
<td>C</td>
</tr>
<tr>
<td>HUMAN 2</td>
<td>L</td>
<td>A</td>
<td>R</td>
<td>F</td>
<td>C</td>
</tr>
</tbody>
</table>

A=Ala, C=Cys, D=Asp, E=Glu, F=Phe, G=Gly, H=His, I=ile, K=Lys
L=Leu, M=Met, N=Asn, P=Pro, Q=Gln, R=Arg, S=Ser, T=Thr, V=Val
W=Trp, Y=Tyr,
IV. RELAXIN AND THE GUINEA PIG ENDOMETRIUM

The guinea pig uterus consists essentially of myometrium and endometrium. The latter is lined by simple columnar epithelium. The epithelium of the endometrium projects into the underlying stroma to form branched, coiled, tubular glands which are known as the endometrial glands (Wagner, 1976).

The guinea pig uterus as a major source of relaxin was first suggested by Hisaw and Zarrow in the 1940's (Hisaw, et al., 1942, Hisaw, et al., 1944, Zarrow, et al., 1948). It was noted that pelvic relaxation did not occur if hysterectomy was performed at the time of ovariectomy.

Some four decades later, the cellular source of relaxin in the guinea pig uterus was demonstrated unequivocally. Employing an antibody to purified porcine relaxin and the peroxidase-antiperoxidase method, relaxin immunostaining was shown over only the cells of the endometrial glands of the late pregnant guinea pig (Pardo and Larkin, 1980). A study extending this to endometrial relaxin throughout pregnancy in the guinea pig, showed that the endometrial relaxin was detectable by immunostaining as early as in mid pregnancy (day 30 of gestation), and maximal in late pregnancy (day 60 of gestation). Following parturition on day 65, endometrial relaxin level fell rapidly, and was undetectable by day 3 of lactation. (Pardo et al., 1982). A subsequent immunocytochemical study at the electron microscopy level showed that the endometrial gland cells were the site of
accumulation of large (500 to 1,000 nm diameter), dense membrane-bounded granules which contained relaxin (Pardo et al., 1984). More recently, relaxin has been shown in the endometrial gland cells of the cyclic guinea pig, estrous cycle day 9 and 14 by a modified immunostaining method (Bryant-Greenwood et al., 1991). Table 8 summarizes the immunostaining data of relaxin in the guinea pig uterus.

Relaxin-like bioactivity in the serum of pregnant guinea pigs was detected by the guinea pig pubic symphysis palpation assay by Zarrow (1948). Relaxin-like immunoactivity in serum was reported by O'Byrne et al. (1976) and more convincingly by Boyd et al. (1981). In addition the latter reported the presence of relaxin in lactation (Boyd et al. 1981). The question arose, whether relaxin was either produced locally or simply sequestered from other non-uterine sources. Unequivocal evidence that guinea pig endometrial relaxin was produced locally was first obtained in our laboratory (Tashima et al., 1988) by Northern analysis with two radioactive synthetic oligonucleotide probes (48 mer) corresponding to the C peptide region of pig prorelaxin.

The endometrial glands have classically been assumed to have some effect on nourishment of the embryo prior to implantation (Finn, 1977). Identification of relaxin in guinea pig endometrial gland cells throughout pregnancy and at parturition suggest multiple functions of the glands: local maintenance of the uterus in a quiescent status during
pregnancy (Porter et al., 1972), lengthening of the interpubic ligament (Wahl et al., 1977), cervical dilation prior to parturition (Sherwood, 1988), and increasing plasminogen and collagenolytic activity in the fetal membrane at birth (Koay et al., 1986).

Table 8

Immunostaining pattern of relaxin in the guinea pig endometrium

<table>
<thead>
<tr>
<th>Reproductive State</th>
<th>Staining Intensity*</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrous Cycle (day 9, day 14)</td>
<td>++</td>
<td>Bryant-Greenwood et al., 1991.</td>
</tr>
<tr>
<td>Late pregnant (day 60, day 63)</td>
<td>+++++</td>
<td>Pardo et al., 1982</td>
</tr>
<tr>
<td>Lactation (day 5)</td>
<td>+++</td>
<td>Bryant-Greenwood et al., 1991.</td>
</tr>
<tr>
<td>(day 21)</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>(day 28)</td>
<td>++</td>
<td></td>
</tr>
</tbody>
</table>

* Arbitrary scale
V. RELAXIN AND THE MAMMARY GLAND

1. The action of Relaxin on the mammary gland

Hisaw's interest in the modification of the female birth canal before parturition led to the discovery of relaxin (Hisaw, 1926). The effects of relaxin on organs other than the mammary gland, such as the interpubic ligament, the uterus and the cervix have been extensively investigated and relatively well elucidated (Sherwood, 1988). On the other hand, there have been only sporadic reports of the mammary gland as a target organ for this hormone and the mechanism by which relaxin promotes mammary growth is poorly understood. The first evidence of the mammotrophic effect of relaxin was shown in early studies performed on laboratory mammals (Hamolsky and Sparrow, 1945; Trentin, 1951; Garretta and Talmage, 1952; Smith, 1954; Wada and Turner, 1959). When highly purified relaxin became available, investigations on the effects of relaxin on the growth of mammary gland have been reported (Bigazzi and Bani, 1985, 1986a, 1986b). The effects of the relaxin on the mammary gland in several species are summarized in Table 9.
<table>
<thead>
<tr>
<th>Species</th>
<th>effects</th>
<th>co-factors</th>
<th>authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>G. pig</td>
<td>promotes general growth</td>
<td>E*</td>
<td>Gareet et al., 1952</td>
</tr>
<tr>
<td></td>
<td>promotes RNA and DNA synthesis.</td>
<td>E</td>
<td>Sheffield et al., 1984</td>
</tr>
<tr>
<td>Rat</td>
<td>promotes general growth and lobulation</td>
<td>E+P*</td>
<td>Hamolsky et al., 1945</td>
</tr>
<tr>
<td></td>
<td>causes lobulo-alveolar growth</td>
<td>E &amp;</td>
<td>Smith et al., 1953</td>
</tr>
<tr>
<td></td>
<td>promotes RNA and DNA synthesis</td>
<td>STH*</td>
<td>Harness et al., 1977</td>
</tr>
<tr>
<td>Mouse</td>
<td>promotes lobulo-alveolar growth and DNA synthesis</td>
<td>E</td>
<td>Wada et al., 1959</td>
</tr>
<tr>
<td></td>
<td>enhances the growth of the mammary epithelium</td>
<td>E</td>
<td>Bani et al., 1986a</td>
</tr>
<tr>
<td></td>
<td>stimulates proliferation and differentiation of myoepithelial cells</td>
<td>E</td>
<td>Bani et al., 1985</td>
</tr>
<tr>
<td></td>
<td>induces hypertrophy of adipose cells</td>
<td>-</td>
<td>Bani et al., 1986b</td>
</tr>
<tr>
<td></td>
<td>induces hyperplasia of adipose cells</td>
<td>E</td>
<td>Bani et al., 1986b</td>
</tr>
<tr>
<td>Rabbit</td>
<td>promotes duct growth</td>
<td>E</td>
<td>Garrett et al., 1952</td>
</tr>
</tbody>
</table>

* E= estrogen; P= progesterone; STH= somatotropin
1.1. **Relaxin stimulates the growth of the mammary parenchyma**

The mammary gland of the ovariectomized mouse with no hormonal treatment shows all of the features of a resting mammary gland. The parenchyma of each gland consists of a major duct which gives rise to a system of branches. The duct system of the resting mammary gland shows few branches ending in slender or club shaped twigs and is lined by a single layered epithelium.

Bani et al. (1986) reported that in mice treated with estrogen for 7 days, the histological picture and the ultrastructure of the duct system were similar to the mammary gland of untreated animals, but in contrast mice treated with estrogen plus relaxin, the duct system of the mammary gland showed marked changes. The major ducts gave rise to numerous branches and the distal ducts actively proliferated thus causing further branching and elongation of all of the ductal tree. A true lobule-alveolar organization of the gland took place, due to the extension of the main ducts through the connective septa of the fat pad. By electron microscopy, Bani et al. (1985) observed that, in estrogen plus relaxin treated mice, fairly numerous undifferentiated cells between the basal pole of the mature epithelial lumenal cells and the apices of the myoepithelial cells. Some of these cells were undergoing mitosis. In estrogen plus relaxin treated mice, all the stages of development and differentiation of myoepithelial cell could be observed more frequently than that seen in
animals treated with estrogen only (Bani et al., 1985). These observations suggest that relaxin greatly stimulated or facilitated the de novo formation of myoepithelial cells and that estrogen priming was essential for this phenomenon (Bani et al., 1985).

1.2 Relaxin stimulates the growth of the mammary stroma

In addition to the enhancing effects of relaxin on the growth of the mammary ductal epithelial cells and the de novo formation of myoepithelial cells in the mammary gland, it is also noted that this hormone can exert an influence on the differentiation of the mammary stroma.

Bianchi et al. (1986b) found that in the resting mammary gland of ovariectomized mice, there were a few poorly branched ducts covered by thin connective tissue septa containing adipose cells, mainly of the monovacuolated type. But, after 7 days of estrogen exposure alone, the stroma of the mammary gland showed more numerous blood vessels and connective tissue cells of various types, adipocytes formed larger clusters, and among them the multivacuolated adipocytes were seen more frequently. In estrogen plus relaxin treated mice, Bianchi et al. (1986b) observed that the stromal area was greatly increased, the adipocytes were very numerous and tightly packed, and the areas of adipose tissue exceeded the edges of the areola and exhibited conspicuous thickness. Moreover, the monovacuolated adipocytes appeared very large and the multivacuolated ones could be seen easily near the end buds.
and the branching points of the ducts (Bianchi et al., 1986b). It has noted that relaxin alone induced hypertrophy of the adipose cells, whereas to obtain hyperplasia of adipose cells, as well as hypertrophy, estrogen priming was essential (Bani et al., 1986b).

Fibroblastic cells are another major component of the mammary stroma. Sheffield and Anderson (1984) reported that relaxin plus estradiol increases DNA and RNA synthesis in mammary fibroblasts which were isolated from guinea pig mammary gland and grown in vitro. McMurtry and co-workers (1978) observed that radioactive labeled relaxin was bound to fibroblasts isolated from human skin and the mouse pubic symphysis and was mitogenic to these cells. In addition to stimulating nucleotide synthesis in fibroblasts, relaxin alone and/or synergizing with other hormones was shown to enhance protein synthesis in the mammary stroma, especially the collagen accumulation (Sheffield et al., 1984). These investigations on the effects of relaxin suggest that this hormone may play a direct role in vivo in connective tissue remodeling, which would involve both the breakdown and resynthesis of collagen and other connective tissue components and hence link to and facilitate parenchymal and stromal growth.
2. The mammary gland as a site of relaxin synthesis

In the pregnant guinea pig, the major source of relaxin is the endometrial gland cells (Pardo et al. 1980; Tashima et al. 1988). The production of relaxin was also demonstrated in the mammary gland of the pregnant guinea pig as judged by immunocytochemistry and Northern analysis (Peaker et al., 1989). Using the same techniques these authors reported for the first time that mammary tissue contained relaxin. A light, uniform staining for relaxin appeared in the cytoplasm of all mammary epithelial cells in the cyclic animals (day 14). As mammary gland growth progressed during pregnancy, the staining for relaxin became non-uniform and less intense. However, as the mammary gland became fully differentiated in lactation, all of the epithelial cells showed intense and uniform staining for relaxin throughout the cytoplasm. Mammary synthesis of relaxin, rather than sequestration from the circulation, was suggested because of the persistence of the mammary staining for relaxin in animals from which the uterus, a major source of relaxin in this species, had been removed (Peaker et al., 1989). The local production of relaxin in the mammary gland was further confirmed by Northern analysis of mammary tissue from the lactating guinea pig (Perker et al., 1989).

It has been shown that there was a distinct qualitative difference in the relaxin concentration of the mammary gland and endometrium of the guinea pig (Peaker et al. 1989; Larkin
et al. 1986). In the mammary gland, the relaxin level was minimal in late pregnancy and reached a maximal level during lactation, whereas the endometrial relaxin was maximal in the last week of gestation and declined greatly during lactation.

Table 10 summarizes the immunostaining patterns of relaxin in these two organs under different physiological conditions.

Table 10
Comparison of the immunostaining patterns of relaxin in the guinea pig mammary gland and endometrium

<table>
<thead>
<tr>
<th></th>
<th>Mammary gland*</th>
<th>Endometrium*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrous cycle</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>(day 14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid pregnancy</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>(day 35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late pregnancy</td>
<td>+</td>
<td>++++</td>
</tr>
<tr>
<td>(day 63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(day 5)</td>
<td>++++</td>
<td>+++</td>
</tr>
<tr>
<td>(day 21)</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>(day 28)</td>
<td>+</td>
<td>++</td>
</tr>
</tbody>
</table>

* Staining intensity was based on an arbitrary scale.
VI. STATEMENT OF THE PROBLEMS

The aims of this dissertation are: a) to obtain the nucleotide sequence and derived amino-acid sequence of the guinea pig preprorelaxin. b) to show how many relaxin gene(s) are present in the guinea pig genome. c) to study the transcription of the relaxin gene in the guinea pig endometrium during the reproductive cycle.

Accordingly the approaches used are outlined separately below.

1. Sequence analysis of a guinea pig relaxin gene

The structure of the guinea pig preprorelaxin(s), either at the amino acid or the nucleotide sequence level is not known. Therefore, as an important first step in the investigation, I have studied the nucleotide sequence of a relaxin gene transcribed in the pregnant guinea pig by using a combination of molecular cloning, polymerase chain reaction and DNA sequencing techniques (Figure 8).
Figure 8

A strategy for sequencing a guinea pig relaxin gene
Screen a bacteriophage cDNA library with a relaxin specific prob

Identify the relaxin specific clones

Purify the relaxin specific clones

Make a single-stranded cDNA pool from mRNA

Screen the cDNA pool and amplify the relaxin specific fragment by PCR

Prepare of sequencing templates by PCR

Dideoxy sequencing

Computer-assisted sequencing data analysis
2. **An analysis of the relaxin gene(s) in the guinea pig genome**

It is not known how many relaxin genes exist in the guinea pig genome. The relaxin synthesised in the guinea pig endometrium and mammary gland in pregnancy and lactation (Peaker *et al.* 1989) may be derived from a single gene or multiple relaxin genes. To clarify this point, I have investigated the number of relaxin gene(s) in the guinea pig genome by a strategy outlined in Figure 9.

3. **The transcription of a relaxin gene in the guinea pig endometrium during the reproductive cycle**

Relaxin immunoactivity was demonstrated in the guinea pig endometrium under different physiological conditions using an antiserum to pig relaxin (Table 8). The endometrial synthesis of relaxin was confirmed by Northern analysis of endometrium collected from the late pregnant guinea pig (Tashima *et al.*, 1988). To extend these observations, I have studied the transcription of the relaxin gene by Northern analysis of the guinea pig endometrium collected at different stages of their reproductive cycle: the estrous cycle, day 35 of gestation, day 63 of gestation and day 6 of lactation.
A strategy for the determination of the number of relaxin gene(s) in the guinea pig genome. The crucial point of selection of restriction endonucleases is that the relaxin specific probe will not be cut by the enzymes, so that the corresponding region, which binds with the probe, of the relaxin gene(s) in the genome will not be digested. This assures that the numbers of hybridized bands shown on X-ray films equals the number of the relaxin gene(s) in the genome. Column A is the situation for the genome which contains one relaxin gene. Column B is the situation for the genome having two relaxin genes.
**Column A.**

Single relaxin gene genome

- relaxin gene
- Completely digested with restriction endonucleases
- Southern blotting
- Hybridized with guinea pig relaxin cDNA probe
- Enzymes: 1 2 3

**Column B.**

Multiple relaxin gene genome

- relaxin gene
- Completely digested with restriction endonuclease
- Southern blotting
- Hybridized with guinea pig relaxin cDNA probe
- Enzymes: 1 2 3
CHAPTER TWO

MATERIALS AND METHODS
1. **Animal and tissues**

   Non-pregnant adult guinea pigs and pregnant guinea pigs (Dunkin Hartely) were purchased from the Laboratory Animal Service, University of Hawaii at Manoa. The guinea pigs were maintained on laboratory guinea pig chow (Ralston-Purina, St. Louis, MO), green vegetables and water. The photoperiod were 12 hrs light and 12 hrs dark.

   The non-pregnant adult guinea pig was primed with estrogen(10 ug/day) and progesterone(2 mg/day) for two weeks. On the 14 day of the artificial estrous cycle, the animal was killed in a CO tank. Uterus and liver tissue were removed quickly and cut into small pieces. The tissues were frozen immediately in liquid N₂ and stored at -80 C.

   The pregnant guinea pigs were killed in an CO₂ tank on day 35 of gestation (mid-pregnancy), day 63 of gestation (late-pregnancy), and day 6 of lactation, two in each group. The uterus was quickly removed and the endometrium scraped from the myometrium, and cut into small pieces. The tissues were frozen in liquid N₂ and stored at -80 C.

2. **Enzymes**

   The following enzymes were purchased from United States Biochemical Company: HindIII (50 iu/ul), Reverse Transcriptase (15.0 iu/ul), Terminal Deoxynucleotidyl Transferase (17.0 iu/ul), Sequenase version 2.0 (13.0 iu/ul). The EcoRI (20,000 iu/ul) and PstI (40,000 iu/ul) were purchased from New England
Biolabs. Ribonuclease A and Proteinase (22 iu/mg protein) were from Sigma Chemical Company. Ampli Taq (5 iu/ul) was from Perkin-Elmer Cetus, and Klenow Fragment (3.0 iu/ul) from BRL Life Technologies Company.

3. **Elutip-D column**
   It was purchased from Scheicher & Schuell Inc.

4. **Genomic DNA**
   Human genomic DNA (1 ug/ul) was purchased from CLONTECH. Guinea pig genomic DNA was prepared in our laboratory and purchased from CLONTECH.

5. **Host cells**
   E coli NM522 was kindly provided by Dr. Dorsey Stuart and used as a host cell for plasmid RLX 9 (pRLX 9), which contains a rat preprorelaxin cDNA insert. E coli Y 1090 was purchased from CLONTECH and used as a host cell for lambda gt 11 recombinants.

6. **Lactating guinea pig mammary gland cDNA library**
   This lambda gt 11 cDNA library was obtained as a gift from Dr. R. K. Craig of Imperial College, London. It was constructed in an special way to remove most of the casein mRNA, a very high copy message in the lactating mammary gland. The library contains 0.36 million recombinants and has a phage concentration of $10^9$ pfu/ml.
7. **Magnagraph Nytran membrane, 3 mm Whatman paper and Sequencing paper**

They were purchased from Micro Separations Inc., from Fisher Scientific and from BIO-RAD Laboratories respectively.

8. **Nensorb 20 column**

It was purchased from E.I. DuPont de Nemours Company.

9. **Oligonucleotides**

Human preprorelaxin B chain 48 mer and pig preprorelaxin C peptide I were offered by Drs. W. Scott Young III and Michael J. Brownstein, National Institute of Mental Health. All primers for the PCR were made by the Biotech-Facility, University of Hawaii at Manoa. Oligo d(T)12-18 was purchased from Pharmacy Company.

10. **Pig preprorelaxin cDNA probe**

Plasmid PR 308-6 (pPR 308-6) was kindly provided by Dr. Simon Kwok of the University of the Kansas Medical Center. It contains a cDNA insert encoding the full-length pig preprorelaxin. This cDNA insert was excised by EcoRI digestion and used as a relaxin specific probe.

11. **Rat preprorelaxin cDNA probe**

Plasmid RLX-9 (pRLX 9) was a gift from Dr. Melvyn S. Soloff of Medical College of Ohio, Toledo, OH. It was constructed by inserting a full length rat relaxin cDNA molecule into the polylinker site of the vector pUC8. The insert was excised by a EcoRI and HindIII double digestion.
12. **Radioactive nucleotides**

All radioactive materials were purchased from E.I.Dupont de Nemours Company: deoxyadenosine 5'-triphosphate tetra (triethylammonium) salt, alpha-P32, 3000ci/mmol. and deoxyadenosine 5'-(alpha-thio) triphosphate, S-35, 1000-1500 ci (37.0-55.5 TBq)/mmol.

13. **Random priming labeling kits**

They were purchased from United States Biochemical Company and BRL Life Technologies Company.

14. **Reagent kit for DNA sequencing with sequenase**

It was purchased from United States Biochemical Company.

15. **Miscellaneous chemicals**

The chemicals that purchased from Fisher Scientific Company: Ammonium acetate, Disodium ethylenediamine tetraacetate(EDTA), Ethidium bromide, NH₄ acetate, NaH₂PO₄, NaOH, CaCl₂, KOAC, MgCl₂, KCL, Ammonium sulfate, Sodium sulfate, Sodium citrate-7 H₂O, Diethylpyrocarbonate(DEPC), Chloroform, Isoamylalchol, Ethanol, Isopropanol, Methanol, Glycerol, Glucose, Dithiothreitol, Gelatin, Formamide, Dimethyl sulfoxide(DMSO).

The chemicals purchased from Sigma Chemical Company: Trizma base, Trizma hydrochloride, Sodium dodecyl sulfate(SDS), Bovine serum albumin(BSA), Agarose(for LB plate), Mineral oil, 2-mercaptoethanol, Sigmacote, Salmon testes DNA (sonicated, denatured, 10 mg/ml), TMED (N,N,N,N-tetramethylethylene).
The chemicals purchased from Pharmacia Company: Dextran sulfate, Sephadax G-25 and G-50.

The chemical purchased from BIO-RAD Laboratories: Ultra pure DNA grade agarose.

The chemicals purchased from Difco Laboratories: Tryptone, Yeast extract.

The chemical purchased from Promega Company: RNAsin.

The chemicals purchased from BRL life technologies company: DNA and RNA molecular weight marker.

16. Materials for photograph

X-ray films were from FUJI Company, Polaroid 55 film (positive/negative 4 x 5 cm instant sheet film) were from Polaroid Company, Developer and Fixer solution were from KODAK Company.

II. SOLUTIONS AND BUFFERS

1. Acetate buffer (2 M, pH 4.2), 100 ml
   73.6 ml of 2 M acetic acid, 26.4 ml of 2 M sodium acetate.

2. 40 % acrylamide stock solution
   38% acrylamide, 2 % bis-acrylamide. Stored in cool and dark.

3. 6 % acrylamide gel solution
   Measured 460 g of urea, 100 ml of 10 x TBE buffer, 150 ml of 40 % acrylamide stock solution. Added double distilled water and brought to a final volume of 1000
ml. Stored in cool and dark for 2 - 4 weeks.

4. **Ampli Taq enzyme buffer**
   670 mM Tris-HCl pH 8.8, 30 mM MgCl$_2$, 166 mM(NH$_4$)$_2$SO$_4$.
The concentration of MgCl$_2$ were optimized according different primers.

5. **Calcium chloride solution**
   60 mM CaCl$_2$, 15 % glycerol, 10 mM PIPES pH 7.0.

6. **DEPC treated dd water**
   Added 0.5 ml of diethylpyrocarbonate into 500 ml double distilled water. Shook well and left at room temperature overnight. Autoclaved and stored at room temperature.

7. **Denhardt's solution, x 10**
   10 g Ficoll 400, 10 g polyvinylpyrrolidone, 10 g bovine serum albumin (Pentax Fraction V). Filtered and stored at -20 C in 25 ml aliquots.

8. **EDTA (ethylenediamine tetraacetic acid) 0.5 M**
   Dissolved 18.6 g Na$_2$EDTA.2H$_2$O in 700 ml dd water.
   Adjusted pH to 8.0 with 10 M NaOH (about 50 ml).
   Added dd water to 1 liter.

9. **Eluting buffer (for mRNA extraction)**
   10 mM Tris-HCl (pH 7.5), 1.0 mM EDTA, 0.05 % SDS.

10. **Ethidium bromide (10 mg/ml)**
    Dissolved 0.2 g ethidium bromide in 20 ml double distilled water. Mixed well and stored at 4 C in dark.
11. Guanidinium thiocyanate stock solution
   Dissolved 250 g guanidinium thiocyanate in 293 ml of DEPC treated water, added 17.6 ml of 0.75 M Na citrate (pH 7.0) and 26.4 ml of 10 % Sarcosyl.

12. Guanidinium solution D
   Mixed 576 ul of 2-mercaptoethanol and 80 ml of guanidinium thiocyanate stock solution.

13. Glyoxal loading buffer
   50 % glycerol, 10 mM Na$_2$H$_2$PO$_4$ (pH 7.0),
   0.4 % bromphenol blue.

14. GTE solution
   50 mM glucose, 25 mM Tris-HCl(pH 8.0), 10 mM Na EDTA.

15. Hybridization solution (for Southern analysis)
   6 x SSPE, 1 % SDS, 50 ug/ml sonicated herring sperm DNA.

16. Hybridization solution 50 ml (for cloning)
   24 ml of formamide, 12 ml of 20 x SSC, 0.5 ml of 2 M Tris-HCl(pH 7.6), 0.5 ml of 100 X Denhardt's solution, 10 ml of 50 % dextran sulfate, 0.5 ml of 10 % SDS, 3.5 ml of double distilled water.

17. High-stringency wash buffer (for cloning)
   0.2 X SSC, 0.1 % SDS.

18. Hybridization solution (for Northern analysis)
   50 % formamide, 2.5 x Denhardt's reagent, 200 ug/ml denatured and fragmented salmon sperm DNA, 0.1 % SDS, 5 x SSPE.
19. "5 M KOAC 1.5 ml
   90 ul of 5 M KOAC, 42.3 ul of dd water, 13.7 ul of Glacial HOAC.

20. LB medium per liter
   10 g of tryptone, 5 g of yeast extract, 5 g of NaCl,
   1 ml of 1 N NaOH.

21. LB/ampicillin plate
   Measured 10 g of tryptone, 5 g of yeast extract,
   5 g of NaCl, dissolved in double distilled water.
   Added 1 ml of 1 N NaOH and 15 g agar. Mixed and brought up to 1000 ml. Autoclaved. When the agar containing medium cooled to 50 C, added ampicillin (50 ug/ml medium). Poured the medium onto sterile disposable petri dishes and allowed to solidify.
   Dried plates in a laminar flow hood for 30 min with the lids off. Stored dry LB/ampicillin plates at 4 C for no more than two weeks.

22. Lambda dilution buffer
   10 mM MgCl₂, 10 mM CaCl₂.

23. Loading buffer x 10 (for Southern analysis)
   20 % Ficoll 400, 0.1M EDTA (pH 8.0), 1 % Sodium dodecyl sulfate, 0.25 % bromphenol blue,
   0.25 % xylene cyanol.

24. Loading buffer x 1 (for mRNA extraction)
   0.5 M NaCl, 1 mM EDTA, 0.1 % SDS, 10 mM Tris-HCl (pH 7.5).
25. **Low-stringency wash buffer**
   2 X SSC, 0.1 % SDS.

26. **Proteinase K digestion buffer**
   100 mM NaCl, 10 mM Tris-HCl (pH 8.0), 25 mM EDTA (pH 8.0), 0.5 % SDS, 0.1 mg/ml proteinase K.

27. **Prehybridization solution (for Southern analysis)**
   6 x SSPE, 1 % SDS, 10 x Denhardt's solution,
   50 ug/ml sonicated herring sperm DNA.

28. **Prehybridization solution (for Northern analysis)**
   50 % formamide, 5 x Denhardt's reagent, 5 x SSPE, 0.1 % SDS, 200 ug/ml denatured and fragmented salmon sperm DNA.

29. **Restriction endonuclease buffers x 10**
   100 mM Tris-HCl pH 7.5, 100 mM MgCl₂,
   10 mM dithiothreitol, 1 mg/ml bovine serum albumin or gelatin. The concentration of NaCl depends upon the restriction endonuclease: 0.5 mM (for low salt buffer), 1.0 mM (for medium salt buffer), 1.5 mM (for high salt buffer).

30. **20 X SSC per liter**
   175 g NaCl(3 M), 88 g sodium citrate-7 H₂O(0.3M),
   adjusted pH to 7.0 with 1 M HCl.

31. **SSPE 20 x, two liters**
   420.77 g of NaCl, 55.2 g of NaH₂PO₄, 80.0 ml of 0.5 M EDTA, adjusted pH to 7.7 and Autoclave.
32. **Suspension medium**
   5.8 g NaCl, 2 g MgSO₄·7H₂O, 50 ml 1 M Tris-HCl (pH 7.5) 0.1 % gelatin (Difco).

33. **Sephadex G-25 (or G-50) suspension**
   Measured 10 g of Sephadex G-25 (or G-50). Placed it into a 125 ml culture flask and added 100 ml of TE buffer (pH 8.0). Left the flask on a bench overnight. Decanted TE buffer, replaced with fresh TE buffer and stored the Sephadex G-25 (or G-50) suspension at 4°C.

34. **TAE buffer x 1**
   0.04 M Tris acetate, 0.001 M EDTA.

35. **Tailing buffer x 5**
   500 mM Potassium Cacodylate (pH 7.2), 10 mM CoCl₂, 1 mM DDT.

36. **TBE buffer x 10 per liter**
   108 g of Tris base, 55 g of boric acid, 40 ml of 0.5 EDTA pH 8.0.

37. **TE buffer pH 7.6 or 8.0**
   10 mM Tris-HCl (pH 7.6 or 8.0), 1 mM EDTA (pH 8.0).

38. **Tris-HCl buffer x 1 M**
   Dissolved 121 g Trizma base in 800 ml dd water. Adjusted to desired pH with concentrated HCl. Mixed and added double distilled water to 1 liter.
III. METHODS

A. AN ATTEMPT TO CLONE AND SEQUENCE OF A RELAXIN GENE TRANScribed IN THE MAMMARY GLAND OF GUINEA PIG

1. Preparation and labeling the rat preprorelaxin cDNA probe

1.1. Introduction of plasmid (pRLX 9) into E. coli NM522

A single colony of E. coli NM 522 was inoculated into 5 ml of LB medium and grown overnight at 37 C in a shaking water bath. The overnight culture of bacterial cells (0.4 ml) was pipetted into 40 ml of LB medium in a sterile 1000 ml culture flask. The flask was incubated in a 37 C shaking water bath with good aeration, to an OD 600 of 0.4. The resultant culture was aliquoted into four of the 25 ml prechilled sterile Corex tubes and incubated on ice for 10 min. The bacterial cells were centrifuged for 7 min at 3000 rpm, 4 C in a SS-34 rotor. The supernatants were decanted. The pellets were resuspended in 10 ml of the ice-cold CaCl₂ solution. The cells were centrifuged for 5 min at 2500 rpm, 4 C. The supernatants were discarded, and the cell pellets were resuspended in 10 ml of the ice-cold CaCl₂ solution and incubated on ice for another 30 min. The cells were recentrifuged once more. The supernatants were removed, and each pellet was resuspended in 2 ml of ice-cold CaCl₂ solution. The cell suspensions were left on ice overnight. The resultant competent cells were dispensed into sterile Eppendorf tubes (100 ul of each aliquot), mixed with 1/4 volume of 70 % sterile glycerol and stored frozen at -80 C.
An aliquot of pRLX 9 (10 ug) in a volume of 10 ul was added into a 15 ml sterile disposable culture tube and placed on ice. The competent cells were rapidly thawed by warming between the hands and 100 ul of these cells dispensed immediately into the tube containing pRLX 9. The competent plasmid mixture was mixed gently and incubated on ice for 10 min. The cells/plasmid mixture was heated at 37 C for 5 min (heat shock). One milliliter of LB medium was added to the tube containing the mixture and the tube incubated in a 37 C shaking water bath for 1 hr. The transformation culture was plated on a LB/ampicillin plate. The plate was inverted and incubated in a 37 C incubator. Bacterial colonies appeared in 12–16 hr.

1.2. Isolation, purification and concentration of pRLX 9

A single colony of pRLX 9 transformed E. coli NM522 was inoculated into 40 ml of LB medium in a sterile 1000 ml culture flask. The flask was placed in a 37 C shaking water bath overnight. The overnight culture was dispensed into four 15 ml Corex glass tubes. The cells were centrifuged for 10 min, at 4,000 rpm, at 4 C in a SS-34 rotor. The supernatants were decanted and each of the cell pellets resuspended in 1 ml of GTE solution. The cell suspensions were transferred into a 15 ml Corex glass tube and let sit on a bench for 5 min. One hundred microliter of lysozyme (10 mg/ml) was added to the tube, the content was mixed well and incubated on ice for 15 min. Two milliliter of 0.2 N freshly prepared NaOH/1%
SDS was pipetted into the tube. The contents were mixed gently but thoroughly and incubated on ice for another 5 min. A prechilled "5M" KOAC solution (0 C, 1.5 ml) was added to the cell suspension. The mixture was vortexed vigorously and let sit on ice for 5 min. The tube containing the mixture was centrifuged for 5 min, 10,000 rpm, at 4 C in a SS-34 rotor.

A volume of supernatant (4.5 ml) was transferred to a fresh 15 ml Corex glass tube and 9 ml of cold ethanol (-20 C) added to the tube. The tube was mixed and incubated on ice for 5 min and the cells centrifuged for 5 min, 10,000 rpm at 4 C. The supernatant was decanted and the pellet was drained well. The pellet was resuspended in 250 ul of TE buffer (pH 8.0). The solution was transferred into an fresh Eppendorf tube and an equal volume of 5 M LiCl added. The tube was incubated on ice for 10 min, and than centrifuged for 5 min, at 13,000 rpm at 4 C. The supernatant was decanted into a fresh Eppendorf tube and 1 ml of cold ethanol (-20 C) added. The contents were mixed well and incubated on ice for 5 min. The tube was centrifuged for 5 min, at 13,000 rpm, at 4 C. The supernatant was discarded, and the pellet was dissolved in 50 ul of TE buffer (pH 8.0). A volume of 7.5 M NH₄OAC (25 ul) and cold ethanol (150 ul) were added, mixed well and incubated at -20 C for 3 hr. The tube was centrifuged for 5 min, 13,000 rpm, at 4 C. The supernatant was removed and the pellet was briefly dried in a vacuum. Finally, the pellet
(plasmid DNA) was dissolved in 50 ul of TE buffer (pH 8.0) and stored at -20 C.

Residual RNA was removed by DNase free-RNase digestion (10ul of 10 mg/ml, 1 hr, at 37 C), and followed by phenol/chloroform extraction and NH4OAC/ethanol precipitation. The yield of pRLX 9 was quantitated by spectrophotometric measurements.

1.3. Digestion of pRLX 9 with EcoRI and HindIII

The rat relaxin cDNA insert was released from the pRLX 9 by digestion with EcoRI and HindIII. The reaction was carried out in a two-step procedure. A volume of pRLX 9 (4 ug) and 2 ul of 10 x EcoRI buffer were pipetted into a clean microcentrifuge tube and double distilled water added to bring the reaction volume to a final volume of 20 ul. Twenty units of EcoRI (5 iu/ug DNA) was pipetted into the tube, mixed well and incubated at 37 C for 2 hr. After phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation, the linearized pRLX 9 was resuspended in 20 ul of 1 x HindIII buffer. Twenty units of HindIII (5 iu/ug DNA) was added, the reaction mixture mixed well and incubated in a 37 C water bath for another 2 hr. The digestion reaction was stopped by adding 0.5 ul of 0.5 M EDTA. The reaction mixture was stored at 4 C for further isolation by agarose gel electrophoresis.
1.4. Resolution and recovery of the rat preprorelaxin cDNA insert from an agarose gel

Two microliter of 10 x loading buffer was mixed with the reaction mixture containing the rat preprorelaxin cDNA insert. The mixture was loaded in the wells of an agarose gel (1.5 %). The HindIII digests of lambda DNA was used as a DNA molecular weight marker. Agarose gel electrophoresis was carried out at 60 V for 2 hr in 1 x TAE buffer. When the electrophoresis was complete, the gel was stained by soaking it in a diluted solution of ethidium bromide (0.5 ug/ml in water) for 30 min and agitating gently. The gel was rinsed with double distilled water, and then visualized by an UV transilluminator and photographed. Two distinct bands were seen: one was at approximately 1 kb containing the rat preprorelaxin cDNA insert, the other was at 2.7 kb corresponding to the pUC 8 vector. The portion of the gel containing the 1 kb band was physically removed from the remainder of the gel. The agarose slice was placed into a dialysis tubing filled with 1 x TAE and subjected to the electroelution at 48 V for 2 hr, followed by running at 100 V for 30 sec in a reversed polarity condition. TAE buffer containing the rat preprorelaxin cDNA was recovered from the dialysis tubing. The rat preprorelaxin cDNA was concentrated and purified using a Elutip-D column, followed by ethanol precipitation, and resuspended in 50 ul of TE buffer (pH 8.0). The rat preprorelaxin cDNA was quantitated by
spectrophotometric measurements. The rat preprorelaxin cDNA was aliquoted and stored at -20 C.

1.5. **Radioactively labeling the rat preprorelaxin cDNA probe by using random primer DNA labeling method**

The rat preprorelaxin cDNA (25 ng in a volume of 5 ul) was transferred into an Eppendorf tube. The cDNA was denatured by heating for 5 min, at 95 C and chilled on ice. To the tube containing the rat preprorelaxin cDNA, the following were added: 3 ul of dNTPs mixture (dATP:dGTP:dTTP = 1:1:1), 2 ul of reaction mixture (random hexanucleotides in 10 x reaction buffer), 5 ul of 50 uCi alpha p-32 dATP, 4 ul of double distilled water and 1 ul (2 iu) of Klenow enzyme. The reaction mixture was gently mixed and microcentrifuged. The tube containing the reaction mixture was incubated in a 37 C water bath for 1 hr. The labeling reaction was terminated by adding 2 ul of 0.2 M EDTA (pH 8.0). The P-32 labeled rat preprorelaxin cDNA was purified using a NENSORB™ 20 cartridge which removed unincorporated radioactive nucleotides, proteins, salt and other low molecular weight materials. After counting and calculating the specific activity, the P-32 labeled rat preprorelaxin cDNA was used as a hybridization probe directly.
2. Screening the lactating guinea pig mammary gland cDNA library by using the rat preprorelaxin cDNA as a molecular probe

2.1. Plating and transferring the lambda gt 11 library

E. coli Y 1090 was streaked out on a LB plate and the plate incubated overnight at 37 C. A single isolated bacterial colony from the plate was inoculated into 5 ml of LB medium containing 0.2 % maltose and 10 mM MgSO(4). The bacteria were grown into saturation at 37 C with good aeration. A certain amount of phage particles (2000 pfu/82 mm plate) was pipetted into a sterile disposable culture tube and mixed with 0.2 ml of the saturated bacterial culture. The tube was place at room temperature for 20 min, and then moved into a 37 C incubator for 15 min. Melted top agarose (2.5 ml) at 45 C was mixed with the contents of the tube and immediately poured onto a prewarmed LB plate(82 mm). The top agarose was quickly spread over the entire surface of the plate by tilting it gently. After hardening, the plate was inverted and incubated at 37 C until plaques covered the plate, but were not confluent (about 8 hr).

The plate was chilled at 4 C for 1 hr to harden the agarose. Nytran membrane was numbered with a soft pencil. Using a forceps, the membrane was applied face down (pencil marker side up) on the cold LB plate bearing phage particles, taking care to avoid bubbles. The membrane was left on the plate for 10 min to allow transfer of phage particles to the
membrane. During this transfer period, the membrane was marked by stabbing a 20-G sterile needle through the membrane into agar at 3 asymmetric points around the edge of the plate. When the transfer was complete, the membrane was carefully peeled off from the plate with a blunt, flat forceps. The membrane was placed face up on a piece of paper towel and allowed to dry at room temperature for 10 min. Another replica could be made from the plate. Two replicas were treated in the following procedure to denatured phage DNAs and immobilize them on the Nytran membrane.

A piece of Whatman 3 MM paper (46 x 57) was placed on a bench and saturated with DNA denaturing solution (0.2 M NaOH / 1.5 M NaCl). The membrane bearing phage DNAs was applied on the paper, DNA side faced up, for 2 min. The membrane, with DNA side faced up, was transferred onto another piece of Whatman paper saturated with neutralizing solution (0.4 M Tris-HCl pH 7.6 plus 2 X SSC) for 2 min, and then to the third piece of Whatman paper saturated with 2 x SSC for 2 min. Phage DNA was fixed onto the membrane by baking at 80 C for 2 hrs. The baked membrane was wrapped in two pieces of Whatman paper and stored at 4 C in a tightly-sealed container.

2.2. Hybridization with the P-32 labeled rat preprorelaxin cDNA probe

The membrane bearing phage DNA was placed on the top of 5 ml hybridization solution containing 50 % formamide and the
solution allowed to seep through the membrane. Other membranes were wetted in the same way, producing a stack of wet membranes (10 membranes/stack). The stack of wetted membranes was placed into a suitable sized plastic bag. The bag was filled with enough hybridization solution to evenly cover the membranes. The bag containing the membrane was then sealed by heat. The total volume of hybridization solution used to wet and cover the membranes was noted so that the amount of radioactive probe could be calculated. Membranes were prehybridized by placing the bag in a 42 C shaking water bath for 6 hr. Before prehybridization was complete, an appropriate amount of P-32 labeled cDNA probe was pipetted into an Eppendorf tube containing 2 mg sonicated herring sperm DNA. The tube was placed in boiling water for 10 min, and then cooled directly on ice. After mixing with 2 ml of hybridization solution, the boiled, cold radioactive probe was added to the bag containing membranes by usbuffer at 42 C (15 min x 2). After checking the radioactive background of the membranes, membranes were resealed separately in plastic bags. The bags were marked with radioactive ink to assist in alignment, and then autoradiographed at -20 C for 48 hr. When a relaxin specific signal was identified on the film, the film was aligned with the membrane and then compared with the original lambda library plate. The position of the labeled cDNA on the film allowed the correct plaque on the plate to be selected. Phage particles containing guinea pig
relaxin-like sequence were picked up and released into a suspension medium, and then subjected to further purification.

3. Purification of phage particles containing guinea pig relaxin inserts

A sterile Pasteur pipette was inserted into the top agarose of the primary plate to cut a circular plug of the gel containing phage particles of interest. The gel plug was blown into 0.5 ml of suspension solution with one drop of chloroform. The phage particles were released into the solution at 4 C overnight. As soon as the titer of the phage particle was known, a secondary plate was made with a phage density less than 100 phage particles per plate and grown at 37 C for 8-12 hr. The phage particles were transferred onto a Nytran membrane. The membrane was processed, hybridized, washed and exposed to X-ray films using the same procedure for screening the primary plates. When the positive plaques on the secondary plates were identified, the plaque which gave the strongest hybridization signal was picked and placed into 0.5 ml of suspension solution with one drop of chloroform. This secondary phage stock was replated on a tertiary plate and screened as described above. The tertiary phage stock was evaluated by further plating and hybridization. When all plaques on the plate were positive, a final phage stock was made from one of the plaque on the last
around phage stock. The five plaques on the secondary plates were identified, the plaque which gave the strongest hybridization signal was picked and placed into 0.5 ml of suspension solution with one drop of chloroform. This secondary phage stock was replated on a tertiary plate and screened as described above. The tertiary phage stock was evaluated by further plating and hybridization. When all plaques on the plate were positive, a final phage stock was made from one of the plaque on the last around phage stock. The final phage stock was stored at 4 °C in suspension solution with a few drops of chloroform.

4. Growing up relaxin specific clones and making high-titer stocks

As soon as the titer of relaxin specific clones had been determined, high-titer stocks of these pure, positive clones were made by a plate lysis procedure. In order to make DNA isolated from the recombinants suitable for enzymatic manipulation, agar was replaced by agarose in the LB plate and top agar recipes. A fresh culture of E. Coli Y 1090 (OD 600 = 0.4) was infected with a suitable amount of the phage recombinants obtained from the last round purification and plated to a 150 mm fresh LB plate. The plate was incubated at 37 °C for 8 to 12 hr until 90 % to 100 % of the lawn of E. Coli Y 1090 was lysed. After pipetting 3 ml of suspension medium and 3 drops of chloroform on the plate, the plate was
placed at 4 C overnight for phage release. A high-titer stock of phage recombinant containing guinea pig relaxin gene sequence was obtained and stored at 4 C with a few drops of chloroform.

5. Preparation of DNA template for sequencing by polymerase chain reaction (PCR)

To make a double-stranded PCR product, a reaction mixture was prepared in a clean PCR tube as following: 10 ul of 10 x Taq buffer, 10 ul of 8 mM dNTPs (2 mM each of dATP, dGTP, dTTP, dCTP), 5 ul of each of two primer flanking the EcoRI cloning site (25 pmol each), 5 ul of phage stock, 0.125 AmpliTaq enzyme, and double distilled water to a final volume of 100 ul. In a blank control of reaction mixture, the phage stock was replaced with 5 ul of double distilled water. The cap of the tubes were replaced tightly and the reaction mixtures were centrifuged. The PCR reaction was run for 35 cycles in a DNA thermal cycler with the cycle profile: 95 C for 1 min, 55 C for 1 min and 72 C for 3 min. When the PCR reaction was complete, the PCR products were microcentrifuged. The PCR products were analyzed by an agarose gel (1 %) electrophoresis.

To make a single-stranded sequencing template, an asymmetric PCR was performed. A reaction mixture was prepared in a clean PCR tube as following: 10 ul of 10 x Taq buffer, 10 ul of 8 mM dNTPS (dATP, dGTP, dTTP and dCTP), 5 ul
of one primer (25 pmol) used in double-stranded PCR, 5 ul of a diluted (1:100) another lambda gt 11 primer (limiting primer, 10 uM), 1 ul of double-stranded PCR products, 0.125 ul of Tag enzyme, and double distilled water was added to make a final volume of 100 ul. The PCR reaction was carried out for 40 cycles in the DNA thermal cycler with the cycle condition: 95 C for 1 min, 55 C for 1 min and 72 C for 3 min. The single-stranded PCR products was concentrated and purified with a Centricon-100. The centriconed single-stranded sequencing DNA templates were stored at -20 C.

6. Dideoxynucleotide chain termination sequencing

6.1. Sequencing reaction

The centriconed single-stranded DNA template (7 ul) was transferred to a clean microcentrifuge tube containing 1 ul of 10 uM sequencing primer (the limiting primer used in asymmetric PCR) and 2 ul of 5 x sequencing buffer. This annealing mixture was heated at 95 C for 5 min and immediately cooled at -20 C for 3 min. After centrifuging the annealing mixture to the bottom of the tube, the tube was placed into a dry-bath block which had been chilled to 0 C. The block was left on a bench, allowing it warm up to room temperature.

The 2.5 ul of each ddNTPs (ddATP, ddGTP, ddCTP and ddTTP) were pipetted into four microcentrifuge tubes. The tubes were labeled as G, A, T, C, respectively, and were capped
tightly to prevent evaporation. The tubes were then prewarmed in a 37°C water bath before beginning the labeling reaction.

The labeling reaction mixture was prepared in a clean microcentrifuge tube on ice: 1.0 μl of 0.1 M DTT, 1.7 μl of 1x TE buffer (pH 7.6), 0.5 μl of S-35 dATP, and 0.3 μl of Sequenase Version 2.0. The annealing mixture was mixed with 5.5 μl of the radioactive labeling reaction mixture, and then incubated in a 20°C water bath for 5 min. When the labeling reaction was complete, 3.5 μl of the labeling mixture was transferred into each of the tubes containing dd NTPs. The termination mixtures were mixed well and incubated at 37°C for 5 min. The termination reactions were stopped by adding 4 μl of stop solution. The sequencing reaction samples were stored at -20°C.

6.2. Denaturing polyacrylamide sequencing gels

A set of glass gel plates (Bio-rad Lab) was cleaned carefully with detergent and double distilled water. The plates were dried completely with 95% ethanol. The gel side of the plates was coated with a thin film of Sigmacote. Two strips of 0.4 mm spacer (Bio-rad Lab) were placed between and on the sides of the glass gel plates. The side and bottom of the plates were taped tightly and clamped with binding clamps.

In clean beaker, 100 ml of 6% acrylamide solution was mixed thoroughly with 660 μl of 10% ammonium persulfate and
66 ul of TEMD. The notch side of the glass plates was raised to a certain angle (60 - 70 degree), and the acrylamide solution was slowly poured into the plates along one side. The angle of the plates and the rate of flowing were adjusted to avoid bubbles. When the gel pouring was complete, the plates were laid on a stopper to maintain an angle of 45 degrees. A shark-tooth comb (Bio-red Lab) was inserted immediately upside down for 1/4 inch into the gel at the notch side of the plates. The gel was left on the bench for 1 hr before setting it up in a gel apparatus. After making sure that the polymerization was complete, the comb, the clamps and the tape were removed from the plates. The surface of the plates were carefully cleaned to get rid of spilled urea/acrylamide solution. The glass gel plates were set up in the gel apparatus by following the instructions. The top reservoir of the gel apparatus was filled with 0.5 x TBE buffer (1500 ml), and the bottom with 0.5 x TBE/1 M sodium acetate (2250 ml). The top of the gel was rinsed with TBE buffer in the top reservoir by using a Pasteur pipette to remove excess urea and fragments of polyacrylamide. The shark-tooth comb was then inserted right-side up, with points just barely sticking into the gel to make wells. The gel was prewarmed at 1700 V for 20 min. Before loading the sequencing samples, the wells were rinsed thoroughly with TBE buffer with the power off, and the samples denatured at 92 C for 2 min. The gel was loaded in the order of "G->A->T->C",
with 3.5 ul of each samples per lane. The sequencing gel was run at 1700 V, 75 W. Running time depended on the sequence data of interest, for 3 hr, 5 hr, 12 hr or even longer.

6.3. Drying sequencing gels and exposure to an X-ray film

When the gel run was complete, the power was turned off and the gel plate was removed from the support. The gel plate was laid flat on the bench and pried apart by using a plastic spatula. A piece of sequencing paper was applied on the gel, and the gel was transferred to the paper by gently peeling the paper/gel from the glass plate. The paper with gel side up was immersed in 5% methanol plus 5% acetic acid for 30 min to remove urea from the gel. The fixative solution was then aspirated, so as not to disturb the gel. The paper with gel side up was transferred on a gel dryer (Bio-rad Lab) and covered with a piece of plastic wrap. Bubbles under the plastic wrap were removed carefully. The gel was dried for 1-2 hr at 80°C on the gel dryer. When the drying was complete, the plastic wrap was peeled off carefully. The dried gel was exposed to an X-ray film with an intensifying screen at room temperature for 48 hr. After development and drying, the sequencing data was read from the bottom to the top of the X-ray film.
7. **Computer-assisted sequencing data analysis**

A VAX computer and the GCG sequence analysis software package was made available to the Biotechnology Program by Dr F.C. Greenwood. The GCG package consists of over a hundred programs which are categorized by function into 15 groups. Some of the programs were used in sequencing data analysis. They are described in brief as follows:

**SeqED**

Enters and modifies sequences.

**Gap**

Compares two nucleotide sequences and gives percentage of homology. Makes an optimal alignment between the two sequences by inserting gaps to maximize the number of matches.

**Map**

Displays both strands of a DNA sequence with a restriction map and possible protein translations. This program can also be used to translate a DNA sequence only in open reading frames.

**Fast A**

Conducts a search for similarity between a query sequence and any group sequences. It answers the question: what sequences in the database are similar to the input sequence.
B. **PARTIAL SEQUENCE ANALYSIS OF A RELAXIN GENE TRANSCRIBED IN THE GUINEA PIG ENDOMETRIUM USING A PCR**

1. **Preparation of poly(A)+ RNA from the late pregnant guinea pig endometrium**

   Eight grams of the frozen tissue was used to prepare poly (A)+ RNA. The frozen tissue was wrapped in a piece of aluminum foil and placed on an lead block which had been chilled to -20 C. The frozen tissue was smashed with a prechilled hammer. The powdered tissue was placed immediately into a beaker containing 80 ml of guanidinium solution D and homogenized under a fume hood. The content of the beaker was transferred to eight 15 ml sterile Corex glass tubes, 10 ml of the solution in each. Into each tube, the following were added: 1 ml of 2 M acetate buffer, 10 ml of water saturated phenol and 2 ml of chloroform/isoamyl alcohol (49:1). The solutions were mixed thoroughly by vortexing and incubated on ice for 15 min, and then centrifuged for 20 min, at 10,000 rpm, at 4 C in a SS-34 rotor. The supernatants were transferred into eight fresh corex tubes containing 10 ml of cold isopropanol. The solutions in the tubes were mixed thoroughly and the tubes were placed at -20 C for 3 hr to precipitate RNA. The total RNAs were pelleted by centrifuging for 20 min, at 10,000 rpm, 4 C in a SS-34 rotor. The RNA pellets were washed with 75 % prechilled ethanol, and the pellets combined into one tube during washing. The total RNA in the tube was pelleted by
centrifuging for 20 min, at 10,000 rpm, 4 C. The supernatant was then removed, and the pellet was dried in a vacuum for 15 min. The dried RNA pellet was dissolved in 3 ml of DEPC treated double distilled water. The yield of total RNA was determined at the spectrophotometer. The RNA sample was then mixed with equal volume of 2 x loading buffer and denatured by heating at 65 C for 5 min. The denatured RNA sample was chilled on ice and loaded onto an oligo dT cellulose column which had been washed and equilibrated with 1 x loading buffer. The column was then washed thoroughly with 1 x loading buffer to remove any residual protein and ribosomal RNA. The poly(A)+ RNA was then eluted with eluting buffer in a volume of 4 ml. The concentration of the poly (A)+ RNA was determined at the spectrophotometer and aliquoted into Eppendorf tubes, 30 ug in each tube. The poly(A)+ RNA in the tube was mixed well with a 1/10 volume of 3 M sodium acetate and 2.2 x volume of 100 % ethanol and incubated at -20 C overnight. The poly (A)+ RNA in the tube was pelleted by centrifuging for 30 min, at 1 3,000 rpm, at 4 C in a microcentrifuge. The supernatant was replaced by 1 ml of 75 % ethanol. The poly(A)+ RNA samples were stored at -80 C.

2. Single-stranded cDNA synthesis

The poly(A)+ RNA (1 ul of 1 ug/ul) was transferred into a screw cap microcentrifuge tube. The tube was capped tightly and then heated in a 92 C water bath for 3 min. When
the denaturation was complete, the tube containing poly(A)+ RNA was quenched on ice and the following added: 0.5 ul of RNAsin, 2 ul of 10 x PCR buffer, 1 ul of random hexamers, 8 ul of each 10 mM dNTPs (dATP, dCTP, dGTP and dTTP), 2 ul of reverse transcriptase and 6 ul of DEPC treated double distilled water. The reaction mixture was mixed gently and incubated at 42 C for 60 min. The reaction was then stopped by heating the mixture at 95 C for 5 min, chilled on ice and diluted with double distilled water to a final volume of 500 ul. The single-stranded cDNA product was aliquoted and stored at -20 C.

3. **Primer design**

The mRNA sequence of pig preprorelaxin was used for PCR primer design. Four sets of PCR primers corresponding to different regions of pig preprorelaxin mRNA were selected by a PCR Primer Selection Computer Program (EPICENTER Software, Table 11). PCR primers were made by the Biotechnology facility.
Table 11

Primers used in amplifying and sequencing of the guinea pig relaxin gene by PCR

<table>
<thead>
<tr>
<th>Number</th>
<th>flank</th>
<th>source</th>
<th>location</th>
<th>size</th>
<th>distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>5'end</td>
<td>pig</td>
<td>-16...-10 aa</td>
<td>20 mer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'end</td>
<td>pig</td>
<td>72.....79 aa</td>
<td>23 mer</td>
<td>241 bps</td>
</tr>
<tr>
<td>II</td>
<td>5'end</td>
<td>pig</td>
<td>-11...-6 aa</td>
<td>17 mer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'end</td>
<td>pig</td>
<td>70.....76 aa</td>
<td>17 mer</td>
<td>225 bps</td>
</tr>
<tr>
<td>III</td>
<td>5'end</td>
<td>pig</td>
<td>13.....19 aa</td>
<td>19 mer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'end</td>
<td>pig</td>
<td>142..147 aa</td>
<td>18 mer</td>
<td>363 bps</td>
</tr>
<tr>
<td>IV</td>
<td>5'end</td>
<td>pig</td>
<td>49.....55 aa</td>
<td>18 mer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'end</td>
<td>pig</td>
<td>142...147 aa</td>
<td>18 mer</td>
<td>255 bps</td>
</tr>
</tbody>
</table>

4. Double-stranded amplification of relaxin specific fragments

Several PCR primers were used in attempts to amplify relaxin specific fragments from a cDNA "pool" prepared from late pregnant guinea pig endometrium. The following additions into a clean PCR tube were performed on ice: 10 ul of 10 x PCR buffer, 10 ul of 10 x 8 mM dNTPs (2 mM each of
dATP, dGTP, dCTP and dTTP), 10 ul of cDNA template, 5'end and 3'end PCR primers (25 pmol each), 60 ul of double distilled water and 0.125 ul of Ampli Taq enzyme. For a blank control, cDNA template was replaced by 10 ul of double distilled water. The reaction mixture was centrifuged, and the surface of the mixture was covered with two drops of mineral oil. The PCR reaction was run for 40 cycles in a DNA thermal cycler (Perkin-Elmer Cetus) with a cycle profile: 94 C for 1 min, 45 C for 1 min and 72 C for 3 min.

5. Southern verification of double-stranded PCR products

5.1. Preparation of a full-length pig preprorelaxin cDNA probe

The 830 bps insert of pPR 308-6 was excised with EcoRI and recovered by electroelution from the agarose gel, purified and concentrated by an Elutip-D column. After a final ethanol precipitation, the full-length pig preprorelaxin cDNAs were resuspended in sterile double distilled water and used as a diagnostic probe (See chapter two, section III, B.1 for the technical details).

5.2 Southern blotting and hybridization

The double-stranded PCR products and DNA molecular weight marker were mixed with 10 x loading buffer and loaded onto a 2 % agarose gel. The gel was run at 100 V for 1 hr, and stained with a diluted solution of ethidium bromide (1 ug/ml) for 15 min. The stained gel was visualized by using an UV
light source and photographed. The double-stranded PCR products were denatured in situ by soaking the gel in 100 ml of 1 M NaCl/0.5 M NaOH (15 min x 2). After neutralizing with 1 M NH₄OAC (15 min x 2), the PCR products were transferred onto a Nytran membrane by a capillary method in 10 x SSPE overnight. The PCR products on the Nytran membrane were immobilized by baking the membrane at 80°C for 2 hr. The membrane bearing PCR products was stored in a tightly-sealed container at 4°C.

The full-length pig preprorelaxin cDNA probe was P32-labeled by a random priming labeling method, and the radioactive cDNA probe was hybridized with the PCR products fixed on the Nytran membrane (See chapter two, section III, A.3 for the technical details).

6. Isolation of the relaxin specific fragment from a Nusiev agarose gel

A Nusiev agarose gel (2 %) was prepared carefully in a cool room. A total of 20 ul double-stranded PCR products was loaded on the gel and run in the cool room at 80 V for 2 hr. When the electrophoresis was complete, the gel was carefully removed from the gel box, and stained with a diluted solution of ethidium bromide (1 ug/ml). The specific PCR products was visualized by a UV light source and photographed. The gel slice containing the specific PCR products of interest was physically removed with a new blade. The extra agarose was
trimmed off. The gel slice was then transferred to a clean PCR tube, melted by incubating the tube in a 65 C water bath for 5 min, diluted with double distilled water (1:10 or 1:100), and used as template for further amplification.

7. Preparation of single-stranded sequencing templates by an asymmetric PCR strategy

Into a clean PCR tube, the following additions were performed on ice: 10 ul of 10 x PCR buffer, 10 ul of 10 x dNTPs (2 mM each of dGTP, dATP, dTTP and dCTP), 50 pmol of one PCR primer(5 ul of a 10 uM solution), 1 pmol of the limiting primer(5 ul of a 1:50 diluted 10 uM solution), 5 ul of double-stranded PCR products, 65 ul of double distilled water, and 0.125 ul Ampli Taq enzyme. The reaction mixture was spun down to the bottom of the tube. The surface of the mixture was covered with two drops of mineral oil. The PCR reaction was then run for 40 cycles in an automated thermocycler with the cycle profile: 94 C for 1 min, 55 C for 1 min and 72 C for 3 min.

After the PCR reaction was complete, single-stranded PCR products were monitored on a TBE agarose gel (2 %) and photographed. The single-stranded PCR products was concentrated and purified with a Centricon-30 (Amico). The centriconed single-stranded sequencing DNA template was then stored at -20 C.
8. **Dideoxy sequencing and computer-assisted sequencing data analysis**

The centriconed single-stranded sequencing DNA template was sequenced with sequenase version 2.0. The limiting PCR primer was used as a sequencing primer. The technical details of the dideoxy sequencing have been described in Chapter two, section III, B.6. Sequencing data were analysed by GCG sequence analysis software package (Wisconsin Package). SEQed, Gap, Map and Fast programs were used.

C. **AN ANALYSIS OF THE RELAXIN GENE(S) IN THE GUINEA PIG GENOME**

1. **Preparation of genomic DNA from guinea pig liver**

   The frozen guinea pig liver tissue (1 ug) was used to prepare guinea pig genomic DNA. The tissue was crushed on a prechilled lead block (-20 C) with a hammer to a fine powder, and immediately transferred into a culture flask containing 10 ml of proteinase K digestion buffer. The flask was covered tightly and incubated in a shaking water bath at 50 C for 18 hr. When the enzymatic digestion was complete, the digest was distributed into two Corex tubes. The contents of two tubes were mixed with equal volume of phenol/chloroform/isoamyl alcohol. The mixtures were vortexed, and then centrifuged for 10 min, at 10,000 rpm, at 4 C in a SS-34 rotor. The supernatants were transferred into two fresh
Corex tubes, and mixed with a 1/2 volume of 7.5 M ammonium acetate and 2 (original) volume of 100 % ethanol. At this point, DNA formed an immediate stringy precipitate. DNA was pelleted by centrifuging for 2 min, at 5,000 rpm, 4 C in the SS-34 rotor. The pellets were rinsed with 70 % ethanol, and then air dried. The dried DNA pellets were dissolved in TE buffer (pH 7.5). The genomic DNA sample was aliquoted and stored at -20 C.

Human genomic DNA prepared from human placenta was used as a control.

2. Digestion of genomic DNAs with restriction endonucleases

Guinea pig, and human genomic DNA were completely digested with three restriction endonucleases (HindIII, PstI and EcoRI) in suitable buffers. The following additions into three microcentrifuge tubes were performed on ice: 12 ug of genomic DNA, 2 ul of 10 x restriction endonuclease buffer, 60 iu restriction endonuclease (5 iu/ug DNA), and double distilled water to a final volume of 20 ul. The mixture was mixed gently, but thoroughly, and then incubated at 37 C for 2 hr. The reactions were stopped by addition of 0.5 ul of 0.5 M EDTA and stored at 4 C.

3. Southern blotting and hybridization

The genomic DNA digests and the DNA molecular weight marker were mixed with 10 x loading buffer and loaded onto a
1 % agarose gel. Gel electrophoresis was carried out at 48 V for 2 hr. After which, the gel was removed from gel box and stained with a diluted solution of ethidium bromide (1 ug/ml) for 15 min. The stained gel was visualized by using an UV light source and photographed. The gel was depurinated in 100 ml of 0.25 N HCl for 8 min at room temperature with gentle agitation. After a brief rinse in double distilled water, the gel was soaked in 100 ml of 1.0 M NaCl/0.5 M NaOH (15 min x 2), and then neutralized with 1 M NH₄Cl/0.5 M NaOH (pH 7.0, 15 min x 2). The genomic DNA digests in the gel were transferred onto a Nytran membrane by a capillary method in 10 x SSPE overnight. The transfer efficiency was checked by restaining the gel with ethidium bromide. When the transfer procedure was complete, Nytran membrane was rinsed in 2 x SSPE and was baked at 80 C for 2 hr. Nytran membrane bearing genomic DNA digests was stored in a tight sealed container at 4 C. The membrane bearing guinea pig genomic DNA digests was prehybridized in a 42 C shaking water bath for 3 hr. When prehybridization was complete, the bag was opened, and the prehybridization solution was squeezed out completely. The genomic DNA digests on the membrane were then hybridized with a P-32 labeled guinea pig preprorelaxin cDNA probe in a 60 C shaking water bath for 15-20 hr (labeled by a random priming method. See chapter three, section II for details of the guinea pig preprorelaxin cDNA probe).
The Nytran membrane bearing human genomic DNA digests was hybridized with a P-32 labeled human relaxin B chain 48 mer oligonucleotides (labeled by the Tdt method) under the same hybridization condition.

When hybridization was complete, the membranes were washed in 200 ml of the following solutions at room temperature: a) in 2 x SSPE/0.1 % SDS for 5 min, b) in 2 x SSPE/0.1 % SDS for 15 min, c) in 0.5 x SSPE/0.1 % SDS for 15 min, d) in 0.1 x SSPE/0.1 % SDS for 15 min. A final washing was carried out in 0.1 x SSPE / 1 % SDS for 30 min at 42 C. The membranes were then sealed in a fresh sealable bag and exposed to an X-ray film at -20 C. Exposure time was determined by checking radioactive emission after the final washing.

D. THE TRANSCRIPTION OF A RELAXIN GENE IN THE GUINEA PIG ENDOMETRIUM DURING THE REPRODUCTIVE CYCLE

1. Radioactively labeling 3' termini of relaxin specific probes with TdT and alpha P 32-dATP

An amount of guinea pig preprorelaxin cDNA probe (5pmol in a volume of 5 ul) was transferred to a clean microcentrifuge tube. The following additions were performed on ice: 10 ul of 5 x tailing buffer, 15 ul of 50 pM P-32 dATP, 15.5 ul of double distilled water and 4.5 ul of TdT enzyme. The reaction mixture was gently mixed, centrifuged and incubated
at 37 C for 15 min. The labeling reaction was stopped by adding 0.5 ul of 0.5 M EDTA.

The radioactive probes were separated from unincorporated dNTPs by a spin column method. The bottom of a 3 ml disposable syringe was plugged with glass wool and filled with 3 ml of swollen Sephadex G-25 suspension (Sephadex G-50 was used for probes larger than 100 bps). The resin-filled syringe was placed into a 15 ml disposable centrifuge tube and spun for 3 min at # 4 of the speed scale in a clinical centrifuge to pack the resin in the column. The radioactive reaction mixture was diluted with TE buffer to a final volume of 100 ul, and loaded onto the center of the gel bed. The syringe was then transferred into a fresh 15 ml disposable centrifuge tube and spun for 5 min at a # 6 of the speed scale of the clinical centrifuge. The P-32 labeled guinea pig preprorelaxin cDNA was in the bottom of the centrifuge tube. After counting and calculating the specific activity, it was used as a radioactive probe.

2. Poly(A)+ RNA extraction

Guinea pig endometria were collected at different stages of the reproductive cycle: day 14 of the artificial estrous cycle (estrogen and progesterone primed, see chapter two, section I for details), day 35 of gestation, day 63 of gestation and day 6 of lactation. The total RNA was extracted by a guanidinium thiocyanate method. Poly(A)+ RNA
was isolated from the total RNA using an oligo-dT cellulose column (see chapter two, section III, B.1 for technical details).

3. Northern blotting and hybridization

Poly(A)+ RNA samples were microcentrifuged at 13,000rpm at 4°C for 30 min. The supernatants were decanted and the pellets were dried briefly in a vacuum. Poly(A)+ RNA samples were resuspended in 7.2 ul of DEPC treated double distilled water. To each of 7.2 ul poly(A)+ RNA sample, the following additions were performed: 6.0 ul of DMSO, 3.0 ul of 6 M Glyoxal and 1.8 ul of sodium phosphate buffer (pH 6.5). The mixtures were incubated at 50°C for 1 hr to denatured poly A+ RNA samples. A RNA molecular weight marker was denatured in the same way. When the denaturation was complete, the poly (A)+ RNA samples were cooled on ice and mixed with 2 ul of loading buffer, and loaded immediately on 1.4 % agarose gel. The denaturing gel electrophoresis was run at 60 V for 90 min in 0.01 M sodium phosphate buffer. After the electrophoresis was complete, the gel was removed from the gel box and stained in a diluted solution of ethidium bromide for 30 min at room temperature. The poly(A)+ RNA samples were visualized using a UV light source and photographed. The poly(A)+ RNA samples were transferred onto a Nytran membrane by a capillary method in 10 X SSPE overnight. When the transfer procedure was complete, the membrane was rinsed in 5 x SSPE,
and baked at 80 C for 2 hr. The membrane was stored at 4 C in a sealed box. The gel was restained with ethidium bromide to check transfer efficiency.

The membrane bearing poly(A)+ RNA samples was prehybridized (with 50 % formamide) at 42 C for 3 hr. The poly (A)+ RNA samples on the membrane were then hybridized with a P-32 labeled guinea pig preprorelaxin cDNA probe at 42 C for 20 hr with 50 % formamide (1-5 x 10^6 cpm/ml hybridization solution, see chapter three, section II for details of the probe).

When the hybridization was complete, the membrane was washed twice in 200 ml of 6 x SSPE/0.1 % SDS for 15 min at room temperature, and then twice in 200 ml of 1 x SSPE/0.1 % SDS at 37 C. The membrane was sealed in a fresh sealable bag and was exposed to an X-ray film at -20 C. Exposure time was determined by checking radioactive emission after washing.
CHAPTER THREE

RESULTS
I. AN ATTEMPT TO CLONE AND SEQUENCE A RELAXIN GENE
TRANSCRIBED IN THE MAMMARY GLAND OF THE GUINEA PIG

It has been shown that a rat preprorelaxin cDNA probe produces a stronger hybridization signal in the Northern analysis of poly(A)+ RNA prepared from the guinea pig mammary tissue than a 48 mer pig relaxin oligonucleotide probe (Tashima, unpublished data). Therefore, a full length rat preprorelaxin cDNA probe was prepared and was used to screen a cDNA library constructed from the mammary tissue of the lactating guinea pig.

Plasmid RLX 9 (pRLX 9) was constructed by inserting a full length rat preprorelaxin cDNA molecule into the polylinker site of the vector pUC 8. The cDNA insert was released by a double digestion of the plasmid DNA with EcoRI and HindIII and resolved by electrophoresis on an agarose gel (1.5 %). The portion of the gel containing the 1 kb insert cDNA (Figure 10) was excised from the gel. The rat relaxin cDNA was then recovered from the gel slice by electroelution and further purified and concentrated with an Elutip-D column.

A total of 20,000 lambda gt 11 recombinants were screened with the P-32 labeled rat preprorelaxin cDNA probe. Seven positive clones were identified from their strong hybridization signals on the X -ray films (Figure 11).
Figure 10

The rat preprorelaxin cDNA insert was released by a double digestion of EcoRI and HindIII. The insert cDNA was resolved by electrophoresis on an agarose gel (1.5 %). One band at a position of 1 kb represented the rat preprorelaxin cDNA (arrow) and the other, which size was 2.7 kb, corresponded the pUC 8 vector. HindIII digestion of the lambda DNA was used as a DNA molecular weight marker.

Lane A: DNA molecular weight marker.
Lane B,C,D,E: rat preprorelaxin cDNA insert and its vector.
Figure 11

Guinea pig relaxin specific clone (arrow) identified by the P-32 labeled rat preprorelaxin cDNA probe.
The positive clones were purified by a serial dilution of the phage particles, replating and hybridization until all plaques on the plate were positive. High-titer stocks of the purified positive clones were made and used for preparation of sequencing template.

The cDNA insert of the positive clones were obtained by a PCR strategy. The inserts were uniformly small, with one exception, as shown in Table 12.

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>#2</td>
<td>#3</td>
<td>#4</td>
<td>#5</td>
<td>#6</td>
<td>#7</td>
</tr>
<tr>
<td>566bp</td>
<td>66bp</td>
<td>96bp</td>
<td>31bp</td>
<td>216bp</td>
<td>111bp</td>
<td>96bp</td>
</tr>
</tbody>
</table>

The cDNA insert of the largest positive clone #1. was reamplified by an asymmetric PCR strategy and the single-stranded PCR product was sequenced by the dideoxy sequencing method. A total of 350 bps readable sequence data was obtained (Figure 12).
Partial nucleotide sequence (350 bps) of the guinea pig relaxin positive clone #1.
gprlx.seq  Length: 350  April 19, 1990  20:30  Check: 1832

  1  TTTATTTTAT TGGTCCACC AACCTGTGTAAG GCCCTATTAC CCCACTCCTT
  51  TACCCCCACT ACTCAGCTAT CAATAAATCT AGCTATAGCC ATCCACTATG
 101  AGCAGCAACT GTAATCTTAG GTTTCGGTAT AAAATAAAAG CATCCTAGCC
 151  ATTTCCAGCA GGAATGAGGG AATAGCCCCC GGACTGTAGT AACTTGCA
 201  TACTAATGCCT AATTGTAACA TCATGTTATT CATCTAGCA CTTAATAGCC
 251  TAGTGCTACG CCTAATATAC GCGGCCAGAG CTGTAGACCC TAAGAGGGCA
 301  ACWGTGTAH CTATAGACCT CTAGGCTCTA TAACTACTAG CTACTTCCTGG
Analysis by the "Fast A" program of Wisconsin Package against EMBL(European Molecular Biology Laboratory)database, showed that the homology of this 350 bp fragment to the mRNA sequence of pig prepro-relaxin was "probably significant". When compared with the mRNA sequence of preprorelaxin of different species (pig, rat, human H 1 and human H2), the average homology was 42.2 %.

This 350 bps fragment was analyzed by the "Map" program of Wisconsin Package and there were no "open reading frames" to be identified.

The 5'end of this 350 bps fragment (a total of 27 bps) was shown to encode a 9 amino acid residues segment which was ended by a cys-TAG (stop codon) structure. This structure was thought to be the end of the A chain of guinea pig relaxin but the remaining portion of this 9 amino acid segment was quite different from the corresponding region of the A chain in other relaxins (Figure 13). Therefore this 350 bps fragment appears to be in the 3' untranslated portion of guinea pig relaxin gene because analysis of the 5'end of this 350 bps fragment (a total of 27 bps) by the "Gap" program of Wisconsin Package showed that it had a significant homology (55.6 %) with a region of the 3'-untranslated portion of mRNA of pig and rat relaxins (Figure 14).
Figure 13

Comparison of the 9 amino acid segment predicted from the 5' end of the 350 bps fragment with the end of the A chain of relaxins in different species. The amino acid residues which may contribute to the receptor binding were underlined.
Guinea pig: N'-- Phe Ile Leu Leu Gly Pro Pro Thr Cys --C'
Pig: N'-- Ile Arg Lys Asp Ile Ala Arg Ler Cys --C'
Rat: N'-- Thr Arg Arg Ser Ile Ala Lys Leu Cys --C'
Human H 1: N'-- Thr Lys Arg Ser Leu Ala Lys Tyr Cys --C'
Human H 2: N'-- Thr Lys Arg Ser Leu Ala Arg Phe Cys --C'
Figure 14

Comparison of the 5' end of the 350 bps fragment (a total of 27 bps) with the mRNA sequence of pig and rat relaxins.
Guinea pig:

1 ....TTTATTTTATGGGTCCACCAACCTGT................. 27

Pig:

651 TTAGGGTTTTTCTTTATGTGTAAGAAAATGTCCTTCATTTATGTGAT 700

Guinea pig:

1 ....TTTATTTTATTGGGTCCACCAACCTGT................. 27

Rat:

680 CTTTTCTTATTTATTAGGATCTGAGAAACCATAATTCCACCTTGATTGT 730
II. PARTIAL SEQUENCE ANALYSIS OF A RELAXIN GENE TRANSCRIBED IN THE GUINEA PIG ENDOMETRIUM BY PCR

Poly(A)+ RNA isolated from the endometrial tissue of late pregnant guinea pig was copied into cDNA in vitro. The resultant cDNA "pool" was screened using a PCR strategy. Four sets of the synthetic oligonucleotide primers corresponding to different regions of the mRNA sequence of pig preprorelaxin were designed with a PCR primer selection computer program (EPICENTER Software). One set (Set I) of these relaxin specific primers (see Chapter two, Table 11) amplified a relaxin specific fragment from the cDNA "pool" (Figure 15). After purification on a 2 % Nusiev agarose gel, this double-stranded relaxin specific fragment was reamplified using an asymmetric PCR strategy (Figure 16). The resultant single-stranded PCR products were centrifuged and sequenced by the dideoxy sequencing method. A total of 227 bps were read from three sequencing X-ray films (Figure 17). Computer-assisted sequence data analysis of this relaxin specific fragment (226 bps) showed that it had a significant similarity to the mRNA sequence of pig preprorelaxin (Figure 18). This relaxin specific fragment (226 bps) had been predicted to encode the intact B chain of guinea pig preprorelaxin and the partial amino acid sequences of the signal peptide and connecting peptide of this hormone (Figure 19). The comparison of this guinea pig preprorelaxin fragment with that of other species is presented in Figure 20.
Figure 15

Relaxin specific fragments amplified from the cDNA "pool" prepared from the endometrial tissue of late pregnant guinea pig using PCR.

The nucleotide sequence of the 5' terminal primer was from the signal peptide region of the pig preprorelaxin mRNA (-16 to -10 amino acid residue). The nucleotide sequence of the 3' terminal primer was from the connecting peptide region of the pig preprorelaxin mRNA (72 to 79 amino acid residue).

The PCR products were resolved by electrophoresis on an agarose gel (2 %). The gel was stained with a diluted solution of ethidium bromide (1 ug/ml).

Lane A: DNA molecular weight marker.
Lane F: Reaction mixture without cDNA template (PCR Blank)
Lane B, C, D, E, G: Relaxin specific fragments.
The double-stranded PCR products showed in Figure 13 was reamplified by asymmetric PCR. The 5' terminal primer used in the double-stranded amplification served as a limiting primer.

The single-stranded PCR products (arrow) run more slowly than the double-stranded PCR products.

Lane A: DNA molecular weigh marker.
Lane B,C,D,E,F,G,H,I: Relaxin specific fragments.
Lane J : Reaction mixture without cDNA template (PCR Blank)
Figure 17

The dideoxy sequencing of the single-stranded PCR products shown in Figure 14.

The samples were loaded in duplicates (1, 2, 3). The gradient sequencing gel (6%) was run 6 hr. The sequence data was read from the bottom to the top, and in an order, G -> A -> T -> C.
Comparison of the nucleotide sequence of the guinea pig preprorelaxin fragment (226 bps) with the corresponding region of the pig preprorelaxin.

The comparison was carried out by the Gap program of Wisconsin Package. One gap was inserted in order to make a maximal match between two molecules.
<table>
<thead>
<tr>
<th></th>
<th>GCCAGGCTTCTGGAACTGTATCTCTCTGGGTTTTTGGACAAAGTTATTAAG 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>..CAACTTCCCCAGAGAAATCCGAGCCAGAGTACGAACGATTATTAAG 96</td>
</tr>
<tr>
<td></td>
<td>GTATGTGGACGTATTTAGCTGGAAATCAAGATAGATATTTGTGGCAAGAT 100</td>
</tr>
<tr>
<td>97</td>
<td>GCATGCCGGCCAGAAATAGCTCCGGTCTGTGGGCTGGAGATCTGTGGCTCCTGGT 146</td>
</tr>
<tr>
<td>101</td>
<td>TCTGTTGGGAAGATATGACTACCGGCCAAGAGAAACAGCGCATTCTGGGAT 150</td>
</tr>
<tr>
<td>147</td>
<td>CTCTGGGAAAGAACTGCTCAGCGCCAGCTGGAA 193</td>
</tr>
<tr>
<td>151</td>
<td>CTGGACAATCCGCAAAACTCATGCCATCCTCTATCAACAAAGCAG..GTA 198</td>
</tr>
<tr>
<td>194</td>
<td>CTGGACCCCCGCGCAAAACCATGCCATCTCCATCCACAAAGATGCAGAA 243</td>
</tr>
<tr>
<td>199</td>
<td>GACTCCCTAAATATGTGGGAAATCCATTGC. 227</td>
</tr>
<tr>
<td>244</td>
<td>ATCTAAAGATGGTTGGAATTTGTTCCT</td>
</tr>
</tbody>
</table>
Figure 19

Partial nucleotide sequence (226 bps) and the predicted amino acid sequence of the relaxin gene transcribed in the endometrium of the late pregnant guinea pig.

Possible signal peptide/B chain cleavage site and B chain/C peptide boundaries are shown by downward-pointing arrows.
Signal peptide

\[ \text{CAG GCT TCT GGA ACT GTA TCT CCT GGG TTT TTG GAC AAA GTT ATT} \]
\[ \text{Gln Ala Ser Gly Thr Val Ser Pro Gly Phe Leu Asp Lys Val ile} \]

\[ \text{AAG GTA TGT GGA CGT GAT TTA GTC CGA ATC AAG ATA GAT ATT TGT} \]
\[ \text{Lys Val Cys Gly Arg Asp Leu Val Arg ile Lys ile Asp ile Cys} \]

\[ \text{GGC AAG ATT CTG TTG GGA GAT ATG ACT ACC GGC CAA GAG AAA CAG} \]
\[ \text{Gly Lys ile Leu Leu Gly Asp Met Thr Thr Gly Gln Glu Lys Gln} \]

\[ \text{CGC ATT CTG GGA TCT GGA CAA TCC GCA GAA ATC ATG CCA TCC TCT} \]
\[ \text{Arg ile Leu Gly Ser Gly Gln Ser Ala Glu ile Met Pro Ser Ser} \]

\[ \text{ATC AAC AAA GAG GTA GAC TCC CTA AAT ATG TTG GAA TCC ATT} \]
\[ \text{ile Asn Lys Glu Val Asp Ser Leu Asn Met Leu Glu Ser ile} \]
Figure 20

Comparison of the amino acid sequence of the guinea pig preprorelaxin fragment (74 amino acid residues) with the corresponding region of pig, rat, human H1 and human H2 preprorelaxins.

The sequence has been aligned to maximize homology. Amino acid identities are boxed. The putative cleavage sites are indicated. Amino acid residues are numbered from the N terminus of the pig preprorelaxin B Chain.
<table>
<thead>
<tr>
<th></th>
<th>signal peptide</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-20</td>
</tr>
<tr>
<td>Pig</td>
<td>MPRLFSYLLGVLWLLSLSQLPRESIP</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>MSSRLLLLQGLGFOWLFLSQQPCAR</td>
<td></td>
</tr>
<tr>
<td>Human H1</td>
<td>MPRLFLLFECLLLNQFSRAVA</td>
<td></td>
</tr>
<tr>
<td>Human H2</td>
<td>MPRLFFHLLGVCCLLLNQFSRAVA</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>B chain</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G. pig</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>Pig</td>
<td>GQSTNKACGRERLVRVLWVEICGS</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>EEWMDOVIQVCGGRGYARAWIEVCAG</td>
<td></td>
</tr>
<tr>
<td>Human H1</td>
<td>AKWKDDVIKLCGRELVRAIQIAICGM</td>
<td></td>
</tr>
<tr>
<td>Human H2</td>
<td>DSWMEEEVIKLCGRELVRAIQIAICGM</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C chain</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G. pig</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
</tr>
<tr>
<td>Pig</td>
<td>VSWGRTLSTLEEPQLETGP</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>SVGRLALSQEEPEAPLARQATAEV</td>
<td></td>
</tr>
<tr>
<td>Human H1</td>
<td>STWSKRSLSQEADA</td>
<td></td>
</tr>
<tr>
<td>Human H2</td>
<td>STWSKRSLSQEADA</td>
<td></td>
</tr>
</tbody>
</table>

|     |                   | 50 |
| Pig | PSSINKEVDMLESI |   |
| Rat | PSSFIDAEILKMMLEFVPNLPQEL |   |
| Human H1| PSSFINDTEDTIIMLEFIANLPPEL |   |
| Human H2| PSSFINDTEDTIINMSFVANLPQEL |   |

A: Alanine  C: Cysteine  D: Aspartic  E: Glutamic acid
F: Phenylalanine  G: Glycine  H: Histidine  I: Isoleucine
K: Lysine  L: Leucine  M: Methionine  N: Asparagine
P: Proline  Q: Glutamine  R: Arginine  S: Serine
T: Threonine  V: Valine  W: Tryptophan  Y: Tyrosine
III. AN ANALYSIS OF THE RELAXIN GENE(S) IN THE GUINEA PIG GENOME

The number of the relaxin genes in the guinea pig genome was determined by a strategy shown in Figure 9. A crucial element in this project was the selection of suitable probes for hybridization. The target sequences in the genomic DNA to which the probes hybridized should not be cleaved by restriction endonucleases used in the experiments. It is assured that the numbers of the hybridizing bands equal to the numbers of the relaxin genes in the genome. A human relaxin B chain 48 mer and a guinea pig preprorelaxin cDNA probe were analyzed by the Map program in the Wisconsin Package (See Chapter three, Section II for details of the guinea pig preprorelaxin cDNA probe). Three restriction endonucleases, HindIII, PstI and EcoRI were selected to digest the guinea pig genomic DNA. Southern analysis of the human genomic DNA with the human relaxin B chain 48 mer probe clearly confirmed the published data of Hudson, et al., 1983 that there are two relaxin genes in the human genome (Figure 21 A). Using the same method, with the guinea pig relaxin cDNA probe it was strongly suggested that the same situation pertains in the guinea pig genome: at least two relaxin genes (Figure 21 B) are present and sufficiently similar to hybridize to the known sequence obtained in section II.
Figure 21

Southern analysis of human and guinea pig genomic DNAs digested with three different restriction endonucleases.

Autoradiography A shows that there are two hybridizing bands in each lane (arrows): evidence that there are two relaxin genes in the human genome.
Lane 1: Human genomic DNA digested with HindIII.
Lane 2: Human genomic DNA digested with PstI.
Lane 3: Human genomic DNA digested with EcoRI.

Autoradiography B shows that there are two hybridizing bands in each lane (arrows): evidence that there are two relaxin existing in the guinea pig genome.
Lane 4: Guinea pig genomic DNA digested with HindIII.
Lane 5: Guinea pig genomic DNA digested with PstI.
Lane 6: Guinea pig genomic DNA digested with EcoRI.
IV. THE TRANSCRIPTION OF A RELAXIN GENE IN THE GUINEA PIG ENDOMETRIUM DURING THE REPRODUCTIVE CYCLE

Poly(A)+ RNAs were isolated from the endometria of guinea pigs at different times of the reproductive cycle: day 14 of the artificial estrous cycle (primed with estrogen and progesterone, see chapter two, section I for details), day 35 of gestation, day 63 of gestation and day 6 of lactation. The poly(A)+ RNA preparations were glyoxal denatured, fractionated on agarose gels (1.7 %) and transferred onto Nytran membranes by the conventional method. An RNA ladder (0.16-1.17 kb) was used as a molecular weight marker. The poly(A)+ RNAs on the Nytran membrane were hybridized with the guinea pig preprorelaxin cDNA probe obtained in section II. This cDNA probe was prepared by a PCR strategy and confirmed to be guinea pig relaxin gene specific by the dideoxy sequencing analysis (See Chapter two, section II for details). Northern analyses of guinea pig endometria unequivocally demonstrated the transcription of the relaxin gene in this tissue (Figure 22). The results were: 1) the mRNA transcripts had a molecular size of 1.0 kb, consistent with the published data (Tashima, et al., 1988), 2) the mRNA transcripts were detected in mid pregnancy (day 35 of gestation), 3) more of the transcripts were detected during the last week of gestation (day 63 of gestation, and 4) there was no evidence for the transcription of the relaxin gene during lactation (day 6 of lactation), interestingly, there
appeared to be abundant transcripts in the guinea pig endometrium during the estrous cycle (Figure 22).
Northern analysis of the endometrial tissues of guinea pigs during the reproductive cycle.

This autoradiography shows the hybridization of poly(A)+ RNAs isolated from guinea pig endometria with a P-32 labeled cDNA probe, which was prepared using a PCR strategy and confirmed to be guinea pig relaxin gene specific by the dideoxy sequencing analysis.

The guinea pig endometria used for poly(A)+ RNA isolation were collected at different stages of their reproductive cycle:

Lane A: Cyclic (animals were primed with estrogen progesterone for 14 days).

Lane B: Mid pregnant (day 35, gestation).

Lane C: Late pregnant (day 63, gestation).

Lane D: Lactating (day 6, lactation).
CHAPTER FOUR

DISCUSSION
I. AN ATTEMPT TO CLONE AND SEQUENCE A RELAXIN GENE
TRANScribed IN THE MAMMARY GLAND OF THE GUINEA PIG

A study using immunocytochemistry and Northern analysis demonstrated the production of relaxin in the mammary gland of the lactating guinea pig (Peaker et al., 1989). This study used heterologous probes (antiserum raised against the pig relaxin and oligonucleotide probes designed from the pig relaxin mRNA sequence). It was not possible to quantitate the transcriptional activity of the relaxin gene in this tissue. It is known that there are abundant transcripts for "milk proteins" in the mammary gland of the lactating guinea pig. In order to increase the chances of picking up a clone representing guinea pig relaxin, the cDNA library made from the lactating guinea pig mammary tissue was constructed to enrich the abundance of relaxin transcripts in the mRNA preparation. According to information provided by Dr R. Craig, cytoplasmic poly(A)+ RNAs were first isolated from the nuclear poly(A)+ RNAs which might include pre-poly(A)+ RNAs together with their introns. The cytoplasmic poly(A)+ RNAs were copied into cDNAs and constructed into plasmids. The cDNA library was then screened with different poly(A)+ RNA probes, some of them encoding casein proteins. The inserts of the non-casein recombinants were isolated by nick translation and subjected to Northern analysis. Depending on the stringency of the hybridization conditions, recombinants corresponding to the low copy messenger RNAs were pooled and
used as a cDNA library. This cDNA library from Dr R. Craig was used for the first approach in this dissertation to clone the guinea pig relaxin gene.

In this study, seven positive clones were identified and purified with a P-32 labeled full-length rat preprorelaxin cDNA probe. The cDNA inserts were isolated using a PCR method. It is known that the average size of the coding region of preprorelaxins in different species is about 540 bp (Hudson et al., 1981; Haley et al., 1982). Among these positive clones, only clone # 1 contained a cDNA insert which was potentially large enough to cover the whole coding region of a preprorelaxin molecule. Therefore, it was selected for dideoxy sequencing. A total of 350 bps were read from several sequencing X-ray films. Computer-assisted sequence data analysis of this 350 bp fragment showed a probably significant homology to the mRNA of the pig preprorelaxin and an average homology of 42.2 % compared to the mRNA sequence of preprorelaxins in other species. Further investigation suggested that the cDNA insert of the positive clone # 1 could be a 5'end truncated guinea pig preprorelaxin molecule based on its failure to hybridize with the pig preprorelaxin B chain 48 mer oligonucleotide probe (amino acid residue 11-26) and the 42.2 % homology at the nucleotide level which was lower than expected (60 -70 %), but which agreed with the observations that the 3'portion of the preprorelaxins are quite different between species (Haley et al., 982).
The "Map" analysis of this 350 bps fragment by Wisconsin Package showed a 9 amino acid segment at its 5' terminus. This 9 amino acid segment had a common feature of the end of the A chain of relaxins in other species, i.e., cys-TAG (Figure 13), but the remaining portion of this 9 amino acid segment was very different from the corresponding region of the other relaxins (Figure 13). The nucleotide sequence of this 9 amino acids segment was then shown to have a 55.6 % homology with a region of the 3' untranslated portion of pig and rat relaxins (Figure 14). The probable location of this 350 bps fragment was the 3' untranslated region of guinea pig relaxin gene based on the evidence shown above (Figure 23). The other positive clones had even smaller cDNA inserts (from 31 to 216 bps) and none of them hybridized with a pig preprorelaxin C peptide 48 mer oligonucleotide probe. It seemed unwise to continue screening this cDNA library for a full-length guinea pig preprorelaxin molecule.

Construction of a comprehensive cDNA library is still not easy, particularly for rare species of mRNA. In general, it needs high quality preparations of mRNA and cDNA. The higher the concentration of the sequences of interest in the starting mRNA, the easier the task of isolating relevant cDNA clones becomes. The mRNA for relaxin had been enriched before construction of the cDNA library used in this study, but there was no information about the quality of mRNA used as the template for synthesis of the first strand of cDNA.
Figure 23

The probable location of the 350 bps fragment (clone #1) in guinea pig relaxin gene molecule.

SP: signal peptide
B: B chain
C: C chain
A: A chain
Rat relaxin cDNA probe (800 bps)

5'   ______|_______|________|________|_________|   3' tail   3'  

5'   ______|_______|________|________|________|_______|   3' tail   3'  

The 350bps fragment
Moreover, information about the synthesis of cDNA and ligation of the cDNA into the vector was not available. It was apparent that this cDNA library contained shorter inserts than desired, suggesting that the synthesis of the second strand of cDNA might be inefficient (Sambrook et al., 1989). The insert of the positive clone # 1 could not be excised from the lambda gt 11 vector by EcoRI digestion. This problem is almost always caused by inefficiency in adding synthetic linkers to the ends of the double-stranded cDNA (Sambrook et al., 1989).

Other reasons that might account for the failure to identify a cDNA clone representing the sequence of guinea pig relaxin might be the vector used to construct this cDNA library and the probe used to screen it. It has been noted that passage of a cDNA library constructed in lambda gt 11 as is the case here, could result in loss of some classes of recombinants since some recombinants might express low levels of fusion proteins that were either toxic to the host cell or which suppress growth of the infecting phage particles (Sambrook et al., 1989). It was recommended that immunological probes be used to screen a cDNA library constructed in lambda gt 11 (Sambrook et al., 1989).

In summary a classical approach to the problem was not successful, yielding a partial sequence of the 3'end tail of guinea pig relaxin gene. Rather than remake a fresh library, a new strategy was attempted with the advent of PCR.
II. PARTIAL SEQUENCE ANALYSIS OF A RELAXIN GENE TRANSCRIBED IN THE GUINEA PIG ENDOMETRIUM BY PCR

The concentration of the transcripts for relaxin in the starting mRNA preparation is crucial for any attempt to sequence the gene for this peptide hormone. Northern analysis of the guinea pig mammary gland and endometrium confirmed the production of relaxin in these two tissues (Tashima et al., 1988, Peaker et al., 1989), but the transcriptional level of the relaxin gene in the late pregnant guinea pig endometrium was much higher than that in the mammary gland of this species during lactation (Peaker et al., 1989). Both show heavy immunostaining for relaxin but under different conditions: the mammary gland stain requires a much longer incubation with the same concentration of antiserum (Larkin et al., 1986, Peaker et al., 1989). Therefore, endometrial tissue was collected from the guinea pig in late pregnancy to clone and sequence the relaxin gene in this species by a new approach.

It was assumed, a priori, that the nucleotide sequence of guinea pig preprorelaxin is probably similar to that of pig preprorelaxin and it is possible to identify areas which are more or less conserved.

PCR has become a useful tool developed in recent years for molecular biology. It is a very sensitive method for singling out a single DNA molecule from a complex mixture and amplifying it a million-fold in vitro. On the other hand,
PCR is a simple technique as described by Mullis et al., 1986 that "it requires no more than a test tube, a few simple reagents and a source of heat". In practice, however, this technique is relatively complicated and, as yet, incompletely understood. Specifically, it is used to amplify part of a gene whose sequence is not available, either at the nucleotide level or at the amino acid level.

In this approach, four sets of oligonucleotide primers designed from the mRNA sequence of the pig preprorelaxin were used to screen a cDNA pool prepared from the endometrium of the late pregnant guinea pig using PCR. The PCR conditions were carefully optimized. All of these sets of primers amplified putative relaxin-related fragments as judged by the size and by Southern verification with a full length pig preprorelaxin cDNA probe. However, only one set of the oligonucleotide primers (Set I) amplified a DNA segment which was proved by dideoxy sequencing to be guinea pig preprorelaxin: a good illustration of one of the limits of the PCR technique. Although the sequence of the guinea pig relaxin gene was not known, all relaxin genes thus far identified have a number of diagnostic nucleotides in some parts of the molecule.

In the dideoxy sequencing of the PCR products amplified by the oligonucleotide primer Set I, a total of 226 bps were read from three X-ray films. Computer-assisted sequence data analysis of this 226 bps fragment showed that it had a
significant similarity to the mRNA sequence of pig preprorelaxin (Figure 18). The homology of this 226 bps fragment with the corresponding region of the mRNA sequence of preprorelaxin in other species is presented in Table 13.

Further analysis of this 226 bps fragment by the Map program of the Wisconsin Package confirmed that it encodes the entire B chain of the guinea pig preprorelaxin (32 amino acid residues), the C-terminus of the signal peptide (8 amino acid residues), and the N-terminus of the connecting peptide of this hormone (34 amino acid residues). The nucleotide sequence and the predicted amino acid sequence of this fragment has been presented in Figure 19. The comparison of the guinea pig preprorelaxin with that of other species, at the amino acid level, is presented in Figure 20. The comparison of this guinea pig preprorelaxin fragment with the corresponding region of this hormone in other species is presented in Table 13.
Table 13

Comparison of the partial nucleotide sequence and its predicted amino acid sequence of guinea pig relaxin gene with the corresponding region of this hormone in other species

<table>
<thead>
<tr>
<th></th>
<th>Guinea pig relaxin</th>
<th>Pig relaxin</th>
<th>58.8</th>
<th>54.8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Rat relaxin</td>
<td>56.6</td>
<td>61.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Human H 1</td>
<td>61.4</td>
<td>57.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Human H 2</td>
<td>59.6</td>
<td>58.9</td>
</tr>
</tbody>
</table>
The core sequence of the guinea pig preprorelaxin B chain contains all the essential elements for biological activity (Figure 20): two cystine residues (B 10 and B 21) contribute to form an insulin-like disulfide bridge structure (Sherwood, 1988), two glycine residues (B 11 and B 22) provide unique torsion angles for the chain folding (Sherwood, 1988), one lysine and two arginine residues (B 8, B 12 and B 16) have been suggested to be involved with the binding of relaxin to its receptor (Bedarker et al., 1982; Blundell et al., 1983; Dodson et al., 1982). The nonpolar amino acid residues that contribute to the hydrophobic core of insulin (Bedarker et al., 1982; Bedarker et al., 1977) are also found at the right positions of the B chain of guinea pig preprorelaxin: valine (B 9), leucine (B 14), valine (B 15), isoleucine (B 17), and isoleucine (B 21). Most of conserved amino acid residues found in the signal peptide and connecting peptide regions of preprorelaxins in species (pig, rat, Human H 1 and Human H 2) are also represented in the corresponding positions of the guinea pig preprorelaxin (Figure 20).

The processing sites in the guinea pig preprorelaxin have to be deduced from a comparison to the homologous structures determined in other preprorelaxin species. The cleavage by a membrane peptidase, after a residue with a small side chain (alanine or glycine), is common to the processing of many signal peptides (Steiner et al., 1980). The cleavage of the signal peptide to generate a guinea pig prorelaxin probably
occurs at the pro(-1)-gly(1) peptide bond (Figure 20). The hydrophobic end of the signal peptide (pro residue) may hold it in the membrane of the endoplasmic reticulum, such that it can be cut at the gly residue to release the prorelaxin molecule into the endoplasmic reticulum for further modification.

Subsequent processing steps are not easy to define in the guinea pig preprorelaxin molecule. The gly-gln structure seems to be a reasonable candidate for further processing of the guinea pig prorelaxin. This gly-glu structure is adjacent to the B-C chain junction of prorelaxins in other species, and the same structure has been suggested to be the putative processing point of the signal peptide-B chain of pig prorelaxin (Figure 20, Sherwood, 1988).

Fidelity is a major factor to be considered in any product of PCR which gives rise to two types of artifacts: point mutation and mosaic allele. The latter is generated by the in vitro recombination between different PCR products (Gyllensten, 1989). It is estimated that the error rate in PCR is about 0.25 % in a 30 cycle amplification (Saiki et al., 1988b), that is, two random PCR products amplified from 100-1000 ng of genomic DNA after 30 cycles might be expected to differ once every 400-4000 bp (Saiki et al., 1989). The error frequency appears to increase in the presence of higher concentrations of dNTPs, Mg++, and Ampli Taq enzyme (Innis, et al., 1990). The homology between the target sequence (to
be amplified) and that of the PCR primer and the annealing temperature are also critical in reducing the rate of misincorporation but the misincorporation in PCR is inevitable. However, for practical purposes, such error frequency is not a problem when the PCR products are used as hybridization probes or as templates for direct DNA sequencing (Sambrook et al., 1989). In direct sequencing, these PCR artifacts will not be visible against the consensus sequence on the gel (Gyllensten, 1989).

Several useful pieces of technical information were obtained from this study. The oligonucleotide primers designed from different regions of the mRNA sequence of pig preprorelaxin should be tested individually. The oligonucleotide primer Set I shared partial nucleotide sequence with the oligonucleotide primer Set II (Table 11), nevertheless, the primer Set I amplified a segment of guinea pig relaxin gene, but primer Set II picked up a relaxin-related molecule from the same cDNA pool. Based on the error frequency of PCR (1/400-4000), there is an optimal size of the amplified DNA segment, like the 300 bps in this study. Finally, the sequencing data should not be taken at face value. It was fortunate that in the case of sequencing of the guinea pig relaxin gene, the cystine residues at the right position in the molecule were diagnostic of a relaxin together with other amino acid residues necessary for the biological activity of this hormone: those for chain folding, for the binding to its
receptor, and for participation in the formation of the hydrophobic core of this peptide hormone.

The partial nucleotide sequence and the corresponding predicted amino acid sequence of guinea pig preprorelaxin has been elucidated in this section of the study. This is an important step for the studying of the relaxin gene(s) in the guinea pig genome (the third part of this thesis) and for the detection of its mRNA transcripts in the endometrium under different physiological conditions (the fourth part of this dissertation study). Even without the completing sequence of guinea pig relaxin gene, this segment is a significant portion of the gene/peptide. It could be used: as a species-specific probe to rescreen the lactating mammary gland cDNA library; to do in situ hybridization of mammary gland and endometrium; to design a partial sequence for making a specific antibody, and to conduct studies of the developmental regulation of relaxin genes assumed to be present in the mammary gland and endometrium in this species.

The sequencing of the remaining part of the guinea pig endometrial relaxin gene is being undertaken using a RACE-PCR strategy (Appendix).
III. AN ANALYSIS OF THE RELAXIN GENE(S) IN THE GUINEA PIG GENOME

Human genomic DNA has been shown to possess at least two relaxin genes. Hudson et al (1983) digested human genomic DNA prepared from the placenta with several restriction endonucleases and subjected them to a Southern analysis. In their study, two hybridizing bands were identified from each of the human genomic digests using a P-32 labeled human relaxin cDNA probe. One of the bands, 2.25 kb in the EcoRI digestion, represented a relaxin gene (H1) present in the human genome. Another weaker hybridizing band, 7 kb in the EcoRI digestion, suggested the presence of a second relaxin gene (H2) in the human genome. This presumption was confirmed by rescreening a human genomic DNA library with a human relaxin specific cDNA probe and sequencing.

In this dissertation study, human genomic DNA was reinvestigated and used as a positive control and it was confirmed that there are two relaxin genes in the human genome. The sizes of the restriction fragments in the EcoRI digestion (2.25 and 7 kb) were exactly as shown in the published data of Crawford et al., 1984.

Neither the nucleotide sequence nor amino acid sequence of guinea pig preprorelaxin is known and probing the guinea pig genome with heterologous relaxin probes could have led to misinterpretations. In this study, a partial nucleotide sequence of guinea pig preprorelaxin was first elucidated.
Although this is a partial sequence, it allowed an analysis of the relaxin gene(s) in the guinea pig genome. By employing the same strategy used by Crawford et al. (1984), the guinea pig genome was shown to possess two relaxin genes. The results were identical when guinea pig genomic DNA was prepared from guinea pig liver or purchased commercially. It is interesting that the restriction digestion patterns for human and guinea pig genomic DNAs are very similar to each other (Figure 21 A and B).

The genomes of all non-primate mammals so far studied contain one relaxin gene (rat, Hudson et al. 1981; pig, Haley et al. 1989). Two Old World monkeys, rhesus monkey and baboon, were also found to possess one relaxin gene (Crawford et al. 1989). Because there are two relaxin genes in the human genome, it was concluded that the duplication of the relaxin gene occurred exclusively during primate evolution (Crawford et al. 1989). However, the evidence presented here that there may be two relaxin genes in the guinea pig genome gives an exception to this conclusion. It reopens this issue: when the duplication of the relaxin gene occurred during evolution. More studies need be carried out with different species in order to make a generalizations.

Another interesting but not clear question in the relaxin field is why some species need different relaxin genes when others apparently have only one. Both human relaxin genes (H1 and H2) have been found to be transcribed in vivo. The
H2 relaxin gene appears to be selectively transcribed in human ovaries collected during early pregnancy or just prior to parturition (Hudson et al. 1984), and in the corpus luteum of the menstrual cycle (Ivell et al., 1989). Recently, the transcription of the H1 relaxin gene has been confirmed in our laboratory. By combining PCR, restriction mapping and sequencing techniques, the mRNA transcripts of the H1 relaxin gene were identified in addition to those of the H2 relaxin gene in the human decidua, trophoblast and prostate (Hansell et al. 1991), most recently extended to the preovulatory follicle (Lee et al., 1991). The transcription of H1 and H2 relaxin genes in the decidua, trophoblast and prostate, which all stain immunohistochemically for relaxin, suggested that both peptides may be produced locally with differing biological activities mediated by different receptors. It is premature to speculate on their local physiological roles in vivo at present, since there are no samples of native H1 and H2 relaxins available for extensive biological studies, and little is known of the nature and distribution of receptors for the human relaxins.

In the previous section of this dissertation, a relaxin gene was shown to be transcribed in the guinea pig endometrium during late pregnancy using PCR and direct sequencing. In this section of the dissertation, another relaxin gene has been suggested existing in the guinea pig genome. This relaxin gene was shown a possible structural diversification
with the endometrial relaxin gene as judged by the weaker hybridizing intensity with the endometrial relaxin cDNA probe (Figure 21). Whether this guinea pig relaxin gene is expressed and whether the expression is tissue specific will not be known until it is cloned and sequenced, or the mRNA transcripts of this gene be identified. At present, it can not exclude the possibility that this relaxin gene is a pseudogene. A pseudogene is usually to have mutation in its nucleotide sequence, either in the promoter or intron splicing region which inhibit transcription of processing to a functional molecule. Another possibility if that this relaxin gene may be a functional gene. If it is the case, the structural difference from the endometrial relaxin gene, as judged by the hybridizing intensity, implied that the protein product of this relaxin gene may play a differing physiological role mediated by a different receptor. The similarity between the guinea pig and human genome presented here offers an excellent opportunity to further clarify the physiological significance of two different relaxin genes in vivo.
IV. THE TRANSCRIPTION OF A RELAXIN GENE IN THE GUINEA PIG ENDOMETRIUM DURING THE REPRODUCTIVE CYCLE

It has been known for over 40 years that the uterus of the guinea pig is a source of relaxin in this species (Hisaw et al., 1944; Zarrow, 1947, 1948).

With an antiserum against pig relaxin and the PAP method (peroxidase-antiperoxidase), Pardo et al. (1980) first demonstrated that the endometrial gland cells of the pregnant guinea pig contain relaxin. By using the same method, Pardo et al. (1982) extended this study and demonstrated the localization of relaxin in the guinea pig endometrium under different physiological conditions. It was shown that relaxin was not detectable in the uterus of day 15 pregnant guinea pigs. Only weak relaxin immunoactivity was detected in a few endometrial glands in day 30 pregnant animals. As the pregnancy advanced, the immunoactivity of the endometrial relaxin increased. In the last week of gestation, all cells of all endometrial glands showed an uniform and dark immunostaining of relaxin. Following parturition, endometrial relaxin staining fell off rapidly, and was almost undetectable by day 5 of lactation. Using antiserum against pig relaxin and protein A colloid gold method and a technique for identification of carbohydrate, Larkin and Renegar (1986) reported that there were two types of dense granules (carbohydrate-containing and relaxin-containing granules) in the guinea pig endometrial cells. The carbohydrate-containing
granules were identified in the estrous cycle and throughout
the pregnancy. However, the relaxin-containing granules
appeared in the cytoplasm of the endometrial cells at day 45
of gestation, and reached a maximal number before
parturition. There was a remarkable transition from the
carbohydrate-containing granules to the relaxin-containing
granules in endometrial cells during middle to late stages of
gestation. There was no evidence that the relaxin-containing
granules were present in the endometrial cells in lactating
animals. Recently, by increasing the sensitivity of the
immunostaining, Bryant-Greenwood et al. (1991) successfully
demonstrated the presence of the relaxin in the endometrial
gland cells of the cyclic guinea pig (day 9 and 14 of the
estrous cycle).

A recent study carried out in this laboratory (Tashima et
al., 1988) showed that relaxin was indeed produced in the
guinea pig endometrium and was not sequestered from another
source. In this study, poly(A)+ RNAs prepared from the late
pregnant guinea pig endometrium were blotted on a Nytran
membrane and hybridized with three pig relaxin specific
oligonucleotide probes (48 mer). Two probes corresponding to
the C peptide region of the pig preprorelaxin identified the
presence of the mRNA transcripts of guinea pig preprorelaxin.
The size of the preprorelaxin mRNA in this species was
determined as approximately 1.0 kb. However, little informa-
tion could be obtained from heterologous probes due to
limitations in sensitivity. Therefore, in order to understand the physiological role of uterine and the putative mammary relaxins, homologous probes are needed.

In the previous section (Section II) of this dissertation, the partial nucleotide sequence of the guinea pig preprorelaxin was successfully elucidated. P-32 labeled guinea pig relaxin cDNA probe clearly demonstrated the transcription of the guinea pig endometrial relaxin gene in different stages of the reproductive cycle. The transcription activity of this gene was shown to be minimal in the middle of pregnancy, to be maximal in the last week of the gestation, appearing in the artificial estrous cycle, and disappearing after parturition. This gene transcription pattern is consistent with the immunostaining pattern reported previously (Pardo et al., 1982; Larkin et al., 1986; Bryant-Greenwood et al., 1991; Table 14).

In this study, there were more mRNA transcripts of relaxin in the cyclic stage than in mid-gestation. This is because the tissue was collected from estrogen-progesterone primed animals. It has been known that both these steroid hormones can stimulate the synthesis of relaxin (Zarrow, 1948). It appears therefore that the 5'upstream regulatory segments of the endometrial relaxin gene contain estrogen and progesterone receptor response elements. It will be of interest to determine whether the mammary gland relaxin gene has the same elements.
Using a homologous pig relaxin radioimmunoassay, plasma relaxin immunoactivity was detected in guinea pigs during lactation (Boyd et al., 1981). A uniform and intense immunostaining for relaxin was observed in the mammary gland of lactating guinea pigs when incubation with antiserum was significantly extended. This observation is in agreement with the lower transcriptional level of relaxin gene relative to the late pregnant guinea pig endometrium (Peaker et al., 1989). In this study, the mRNA transcript of relaxin gene was not detected in the endometrium of lactating (day 6) guinea pigs by a guinea pig preprorelaxin cDNA probe. Thus, the guinea pig uterus, a main source of relaxin in this species, does not appear to be a source of relaxin in this physiological status. This suggests that some extrauterine source might be responsible for the plasma relaxin and mammary immunostaining relaxin during lactation. The mammary gland and ovaries are possible sources. It is known that in the guinea pig a new estrous cycle begins within 10 hr postpartum (Sisk, 1976). Immunoreactive relaxin has been identified in tissue extracts of ovaries during the estrous cycle and throughout pregnancy (Nagao et al., 1981).

In summary, the results suggest that an endometrial relaxin is synthesized locally throughout pregnancy, and estrogen and progesterone are able to enhance the transcription of this relaxin gene, and that extrauterine relaxin(s) may play a role in these and other physiological stages.
<table>
<thead>
<tr>
<th>Reproductive status</th>
<th>Immunostaining* (relaxin)</th>
<th>Northern analysis** (relaxin mRNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrous cycle</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>(E + P primed)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid pregnancy</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>(day 35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late pregnancy</td>
<td>++</td>
<td>++++</td>
</tr>
<tr>
<td>(day 63)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactation</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>(day 5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Staining intensity was based on an arbitrary scale. In middle-pregnancy and day 5 of lactation, the staining intensity was highly variable. Some endometrial cells stain darkly, interspersed with non-staining cells. Variation from gland to gland in number of cells staining.

(Source: Bryant-Greenwood, 1991)

** Based on an arbitrary scale.

*** Animals were primed with estrogen (10 ug/day) and progesterone (2 mg/day) for 14 days.
V. SUMMARY AND CONCLUSION

The main results and significant findings of this dissertation are presented as follows.

1. **Successful elucidation of a partial nucleotide sequence and derived amino acid sequence of guinea pig endometrial preprorelaxin.**

   By employing molecular cloning, polymerase chain reaction and dideoxy sequencing techniques, the structure of a guinea pig relaxin gene transcribed in the late pregnant endometrium was partially determined. It encodes the intact B chain of this hormone plus part of its signal peptide and connecting peptide region. It is the first report regarding the guinea pig relaxin gene sequence. The partial elucidation of the guinea pig relaxin gene sequence has made possible a study of the guinea pig genome for relaxin genes and an investigation of the transcription of this gene in the endometrium during the reproductive cycle. It has also made possible a strategy to complete the sequence, as outlined in the Appendix.

2. **The guinea pig genome contains two relaxin genes**

   The genome of all non-primate mammals, so far studied, contains one relaxin gene, including two Old World primates, rhesus monkey and baboon. The human genome has been shown to contain two different relaxin genes. It is generally believed that the duplication of the relaxin gene occurred during primate evolution. However, Southern analysis of
genomic DNA digest with the species-specific relaxin cDNA probe presents an exception to this hypothesis: the guinea pig, a non-primate mammal, appears to contain two relaxin genes in its genome. The similarity between human and guinea pig genome provides an opportunity to investigate the multiple physiological roles of relaxin in vivo.

3. The guinea pig endometrial relaxin gene turns-on during the estrous cycle, throughout pregnancy, and then turns-off after parturition.

The transcription of the guinea pig endometrial relaxin gene during some portion of the reproductive cycle was confirmed by Northern analysis with the species-specific relaxin cDNA probe. The transcription activity of this relaxin gene was minimal in the mid pregnancy, but maximal during the last week of the gestation. It turned on in the estrous cycle (in estrogen and progesterone primed animals), but was off after parturition. The transcription pattern of this gene is congruent with its immunostaining pattern and plasma levels.
APPENDIX: AMPLIFICATION OF THE 3' END OF THE GUINEA PIG ENDOMETRIAL RELAXIN GENE BY A RACE-PCR STRATEGY

I. MATERIALS AND SOLUTIONS

1. Animal

A late pregnant guinea pig (Dunkin Hartely) was purchased through the Laboratory Animal Service, University of Hawaii at Manoa. The guinea pig was killed in an CO₂ tank. The uterus was quickly removed and the endometrium was scraped from the myometrium. The tissue was cut into small pieces and frozen in liquid N₂ and stored at -80 C.

2. Oligonucleotide primers

Three oligonucleotide primers were synthesized by the Biotechnology facility (Table 15).

3. Reverse transcriptase (15.0 iu/ul)

It was purchased from United States Biochemical Company.

4. Terminal deoxynucleotidyl transferase (17.0 iu/ul)

It was purchased from United States Biochemical Company.

5. Ampli Tag (5 iu/ul)

It was from Perkin-Elmer Cetus.

6. Enzymes

The following enzymes were purchased from United States Biochemical: DNA polymerase I, T4 polynucleotide kinase, Sml and Bacterial alkaline phosphatase.

7. GENE CLEAN kit

It was purchased from BIO 101, Inc.
8. M13 cloning kit
   It was purchased from New England BIOLABS.

9. 10 x RTC buffer
   500 mM Tris-HCl (pH 8.15) at 41 C, 60 mM MgCl(2),
   400 mM KCl, 10 mM DTT, each dNTP at 10 mM
   ammonium sulfate.

10. 10 x PCR buffer
    670 mM Tris-HCl (pH 8.8 at room temperature), 67 mM
    MgCl(2), 1.7 mg/ml of bovine serum albumin, 166 mM

11. TE buffer
    10 mM Tris-HCl(pH 7.6), 1 mM EDTA.

The source of other chemicals have been listed in
Chapter Two, section I.
Table 15
Oligonucleotide primers used to amplify the 3' end of the guinea pig relaxin gene

<table>
<thead>
<tr>
<th></th>
<th>Oligonucleotide primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Guinea pig relaxin: 5'-TCTGGGATCTGGACAATC-3' specific 18 mer*</td>
</tr>
<tr>
<td>2.</td>
<td>Poly T adapter: 5'-GACTCGAGTCGACATCGATTTTTTTTTTTTTTTTTTTTTTTTTTTT-3'</td>
</tr>
<tr>
<td>3.</td>
<td>Adapter-17 mer: 5'-GACTCGAGTCGACATCG-3'</td>
</tr>
</tbody>
</table>

* See Chapter three, section III for the details of the guinea pig preprorelaxin sequence.
II. METHODS

1. Single-stranded cDNA synthesis

Poly(A)+ RNA was isolated from late pregnant guinea pig endometrial tissue by a classical method (see chapter two, section III, B 1 for details). The poly(A)+ RNA (1 ug) was mixed with 16.5 ul of double distilled water and denatured for 3 min at 65 C. The tube containing heated RNA was quenched on ice, and the following added: 2 ul of 10 x RTC buffer, 10 units of RNASin (0.286 ul), 0.5 ug of dT(17)-adapter primer (0.5 ul), and 10 units of reverse transcriptase (0.67 ul). The reaction mixture was incubated at 42 C for 1 hr, and then at 52 C for 30 min. The reaction mixture was diluted with TE buffer to a final volume of 1 ml. The cDNA "pool" was stored at 4 C.

2. Amplification of the 3' end of the guinea pig endometrial relaxin gene by a RACE-PCR strategy

A reaction cocktail was assembled in a clean PCR tube on ice: 5 ul of 10 x PCR buffer, 5 ul of DMSO, 5 ul of 10 ul x dNTPs (15 mM each), 30 ul of double distilled water, 1 ul of adapter primer (25 pmol/ul), 1 ul of guinea pig relaxin C chain 18 mer (25 pmol/ul), and 5 ul of guinea pig endometrial cDNA. The reaction mixture was heated in a DNA thermal cycler (Perkin-Elmer Cetus) at 95 C for 5 min. The mixture was then cooled to 72 C, and 2.5 units of Ampli Taq enzyme was added. The surface of the reaction mixture was covered
with 30 ul of prewarmed mineral oil (72 C). Templates and primers were annealed at 55 C for 5 min, and the enzyme reaction was extended at 72 C for 40 min. The PCR reaction was continued in the DNA thermal cycler for 40 cycles with a cycle profile: 95 C for 40 seconds, 55 C for 1 min and 72 C for 3 min. A final extension of the PCR reaction was carried out at 72 C for 15 min.

3. **Southern verification of RACE-PCR products with a relaxin specific internal oligonucleotide probe**

RACE-PCR products were mix with 10 x loading buffer, and were loaded onto a 2 % agarose gel. The gel electrophoresis was carried out at 100 V for 1 hr. When complete, the gel was stained with a diluted solution of ethidium bromide (1 ug/ml) for 20 min. The ethidium bromide bonded RACE-PCR products were visualized using an UV light source and photographed. The RACE-PCR products were denatured in situ by soaking the gel in 100 ml of 1 M NaCl/0.5 M NaOH (15 min x 2), and then neutralized with 1 M NH₄ acetate(15 min x 2). The RACE-PCR products were transferred onto a Nytran membrane by a capillary method in 10 x SSPE overnight. The Nytran membrane was baked at 80 C for 2 hr to immobilize the RACE-PCR products. The RACE-PCR products fixed on the Nytran membrane were hybridized with a P 32-labeled pig preprorelaxin oligonucleotide probe (C peptide #1, 48 mer).
After posthybridization washing, the membrane was exposed to an X-ray film for 10 min.

4. **Subcloning and sequencing**

RACE-PCR products were purified by a double gene clean method. The ends of the RACE-PCR product were modified with DNA polymerase I (United States Biochemical) and T4 polynucleotide kinase (United States Biochemical). M13mp18 RF DNAs were digested with SamI and dephosphorylated with Bacterial alkaline phosphatase. The ligation of RACE-PCR product with M13mp18 vector was carried out at 16 C for 24 hr. E coli JM101 competent cells were infected with the ligation reaction mixture. The guinea pig relaxin specific clones were identified with a P-32 labeled pig preprorelaxin oligonucleotide probe (C peptide #1, 48 mer). Single-stranded DNA templates were prepared from isolated guinea pig relaxin clones and sequence with Sequenase version 2.0.
III. RESULTS AND DISCUSSION

The RACE (rapid amplification of cDNA ends) method produces cDNAs by using PCR to amplify copies of the region between a single point in the mRNA transcript and the one end of it. A prerequisite for the RACE protocol is availability of a short stretch of sequence from the gene of interest. In brief, for rapid amplification of the 3' end of the cDNA, mRNA is reverse transcribed using a "hybrid" primer consisting of oligo(dT)(17 residues) linked to a unique oligonucleotide ("adapter") primer. Amplification is subsequently performed using the adapter primer, which binds to each cDNA at its 3' end, and a primer specific to the gene of interest. The RACE-PCR protocol is illustrated in more detail in Figure 24.

In the previous part (chapter two, section II) of this study, a partial nucleotide sequence of a guinea pig relaxin gene was elucidated. It is the starting point for using the RACE strategy to generate cDNAs which overlap with the previous sequencing data of the guinea pig relaxin gene. As the first step of this study, oligonucleotide primers were designed and synthesized. The gene specific primer corresponds to the beginning of the C peptide (C 6-C 12) of the guinea pig preprorelaxin. By combining with the "adapter" primer, this guinea pig relaxin specific primer amplified an approximately 600 bps fragment from a cDNA pool prepared from late pregnant guinea pig endometrium (Figure 25). Southern
hybridization of this 600 bps fragment (RACE-PCR products) with a radioactive internal oligonucleotide probe (pig preprorelaxin C peptide #1 48 mer) confirmed that it is relaxin specific and has an overlap with the previous sequencing data of the guinea pig relaxin gene (Figure 26).

The guinea pig relaxin specific RACE-PCR product was subcloned into M13mp18 vector. Three positive clones were purified and subjected to dideoxy sequencing Using a M13 universal sequencing primer (-40). A total of 333 bps were read from sequencing of three isolated positive clones. Computer-assisted sequence data analysis of this 333 bps guinea pig relaxin specific fragment showed that it had an overlap with the 3' end of the previous guinea pig relaxin sequence data. This 333 bps guinea pig relaxin specific fragment had been predicted to encode the rest part of the guinea pig endometrial relaxin molecule. The nucleotide sequence and its derived amino acid sequence of the guinea pig endometrial relaxin gene is presented in Figure 27.
Schematic representation of the RACE-PCR protocol: rapid amplification of the 3' end of cDNAs. (From: Frohman, 1990)
RACE: 3' end

mRNA

```plaintext
--------------------------
|                            |
```

cDNA
```

1. Denature
2. Animate primer (3'amp)
3. Extend
4. Denature
5. Animate primer 5'30
6. Extend
7. Denature
8. Animate primers: 5'amp and 3'30
9. Extend
10. Copy

Up to 10^6 copies of cDNA
Figure 25

Amplification of the 3'end of the guinea pig relaxin gene by a RACE-PCR strategy. The size of the main bands were about 600 bps, and were consistent with the remaining portion of the relaxin gene transcribed in the late pregnant guinea pig endometrium.

The samples were loaded as duplicates (Lane 1, 2 and 3). The reaction mixture without cDNA template (PCR blank) was loaded in lane 4.
Southern verification of the RACE-PCR products with a P-32 labeled internal oligonucleotide probe. The sequence of the probe was from the C peptide region of pig preprorelaxin. It was one amino acid residue down from the 3' end of the relaxin gene specific primer used in the RACE protocol and had 69% homology with the corresponding region of the guinea pig preprorelaxin.

Lane A: DNA molecular weight marker
Lane B, C and D: three duplicates of guinea pig relaxin fragment (RACE-PCR products)
Figure 27

The nucleotide sequence and its derived amino acid sequence of the prorelaxin gene transcribed in the endometrium of the late pregnant guinea pig

Possible cleavage sites of the connecting peptide are shown by downward-pointing arrows.
Signal peptide
CAG GCT TCT GGA ACT GTA TCT CCT GGG TTT TTG GAC AAA GTT ATT
Gln Ala Ser Gly Thr Val Ser Pro Gly Phe Leu Asp Lys Val Ile

AAG GTA TGT GGA CGT GAT TTA GTC CGA ATC AAG ATA GAT ATT TGT
Lys Val Cys Gly Arg Asp Leu Val Arg Ile Lys Ile Asp Ile Cys

GGC AAG ATT CTG TTG GGA GAT ATG ACT ACC GCC CAA GAG AAA CAG
Gly Lys Ile Leu Leu Gly Asp Met Thr Thr Gly Gin Glu Lys Gin

CGC ATT CTG GGA TCT GGA CAA TCC GCA GAA ATC ATG CCA TCC TCT
Arg Ile Leu Gly Ser Gly Gin Ser Ala Glu Ile Met Pro Ser Ser

ATC AAC AAA GAG GTA GAC TCC CTA AAT ATG TTG GAA TCC ATT GCT
Ile Asn Lys Glu Asp Ser Leu Asn Met Leu Glu Ser Ile Ala

AAT TTG CCA GAA GAG CTG AGG GCA ATG CTG GAG AAA CAG CCA
Asn Leu Pro Glu Glu Leu Arg Ala Met Leu Pro Glu Lys Gin Pro

TCA TCA CAG CTA CAA CAA TAT GTA CCT GCA TTA AAG AAT TCA
Ser Ser Pro Gln Leu Gln Gin Tyr Val Pro Ala Leu Lys Asn Ser

AAT GTT GCC GTT AAA GAA CTT AAT AAA ATT ATT CGT GGT AGA CAA
Asn Val Ala Val Lys Glu Leu Asn Lys Ile Ile Arg Gly Arg Gin

GAA GAA GCA GAA GAC AAC AGC CAT TCA TTA TTA AAA GAT TTC AAC
Glu Glu Ala Glu Asp Asn Ser His Ser Leu Leu Lys Phe Asn Leu

TTG AAT ATT TAT TCA CCA AAA AAA CGA CAG CTG GAT ATG ACA GTG
Asn Ile Tyr Ser Pro Lys Lys Arg Gin Leu Asp Met Thr Val Ser

AGT GAG TGT TGC CAA GTT GGC TGT ACT AGA AGA TTT ATT GCA AAC
Glu Cys Cys Gin Val Val Gly Cys Thr Arg Arg Phe Ile Ala Asn

TCA TGC TGA gatgagttatctgtgac---3'
Ser Cys End
BIBLIOGRAPHY


Pardo, RJ., Larkin, LH. and Fields, PA. (1980):


