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.

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.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Partial Structure and Characterization of the Mouse Oxytocin Receptor Gene

Caterina Russo

A Thesis

in

The Department

of

Chemistry and Biochemisrty

Presented in Partial Fulfilment of the Requirements
For the Degree of Masters of Science at
Concordia University
Montreal, Quebec, Canada

July 2001

ABSTRACT

Partial Structure and Characterization of the Mouse Oxytocin Receptor Gene

Caterina Russo

Oxytocin (OT) is a hypophyseal nonapeptide which exerts a wide spectrum of central and peripheral effects. One of its well-known functions is its involvement in uterine contractions during labor. These functions are mediated by specific oxytocin receptors (OTRs). In order to study the molecular mechanisms underlying the OT/OTR system, I have isolated and characterized 2 identical recombinant phage λ-DASH II genomic clones containing the mouse OTR gene using sequence information available from the human and rat OTR genes. Restriction enzyme mapping of the phage showed that the OTR gene spans > 20 kbp. A 6.0 kbp *Sst* I fragment was subcloned and sequenced. Comparison of this sequence to that of the published rat OTR gene sequence confirmed that the fragment was indeed the mouse OTR. The predicted amino acid sequence is 94% and 97% identical to the human and rat OTR sequences, respectively. The transcription start site was mapped by 5' rapid amplification of cDNA ends (RACE). The major start site was mapped at 307 bp upstream of the ATG codon. The mouse OTR lacks an apparent TATA or CCAAT box as is the case for the rat and human OTR genes. The mouse OTR gene contains three exons and two introns. The first intron lies within the 5'-UT region and is 97 bp long. The second intron interrupts the region encoding the sixth and seventh transmembrane domains and is at least 12 kbp long. The OTR gene is highly expressed at parturition and gives rise to at least two transcripts of 5.0 and 3.6 kbp, respectively.

ACKNOWLEDGEMENTS

The λ -DASH II mouse genomic library (Balb/c strain) was kindly provided by Dr. A. Peterson of the Molecular Oncology Deparment at the Royal Victoria Hospital.

I would like to express my deep appreciation and gratitude to Dr. H. H. Zingg for his guidance, help, understanding, and the opportunity to expand my realms of knowledge. Without him, this thesis would not exist.

I would like to thank Dr. J. Turnbull for having agreed to take me on as a student and for her support, guidance, and pep talks. Cheers, Jo!

I would like to thank Dr. B. Turcotte and Dr. B. Rehfuss for their advice during the course of this work.

I thank Dr. F. Rozen for having given me the confidence and independence to devise and explore techniques that were crucial for my work. You're a great coach!

I wish to express special thanks to all my colleagues, past and especially the present ones. Your kind words, honesty, faith in me, and moral support will never be forgotten. This thesis would not have been completed without your help. Know that you can also count on me!

To my husband, Tony, and my two adorable gems, Adrian and Julian. You have been my rock and my strength. Thank you for your love and patience.

To my Mother and Father who have made so many sacrifices over the years, so that I would have the opportunity to follow my dreams. You have always been there to help me. Thank you, and know that among my many dreams, one was to make you as proud of me as I am proud to be your daughter.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS

OT	Oxytocin
OTR	Oxytocin Receptor
LH	Leutinizing Hormone
AVP	Arginine Vasopressin
NP I	Neurophysin I
NP II	Nuerophysin II
VPR	Vasopressin Receptor
PLC	Phosholipase C
cAMP	Cyclic Adenosine Monophosphate
TMD	Transmembrane Domain
PI	Phosphoinositide
IP3	Inositol Triphosphate
PKC	Protein Kinase C
CNS	Central Nervous System
PG	Prostaglandins
PAF	Platelet-activating Factor
ET-1	Endothelin-1
Il-1	Interleukin-1
TNF	Tumor Necrosis Factor
IL-6	Interleukin-6
IL-8	Interleukin-8
LB	Luria-Bertani Medium

SM	Sodium Magnesium Buffer
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PFU	Plaque Forming Units
PIPES	Piperazine-1,4-bis[2-ethanesulfonic acid]
SDS	Sodium Dodecyl Sulphate
TAE	Tris-Acetate EDTA
DEPC	Diethylpyrocarbonate
CPM	Counts Per Minute
EDTA	Ethylenediaminetetra-acetate
TE	Tris-Ethylenediaminetetra-acetate
dATP	Deoxyadenosine 5`- Triphosphate
dTTP	Deoxythymidine 5`-Triphosphate
dCTP	Deoxycitidine 5`-Triphosphate
dGTP	Deoxyguanosine 5`-Triphosphate
X-gal	5-Bromo-4-Chloro-3-Indolyl-β-D-Galactoside

PREFACE

The following is a description of the partial cloning and characterization of the mouse oxytocin receptor gene. It is a partial cloning due to the fact that the last exon and 3'-untranslated region of the gene were not fully cloned or sequenced. It should also be noted that in 1996, a Japanese group led by Kubota published the mouse oxytocin receptor gene in Molecular and Cellular Endocrinology 124: 25-32, which was when I was also working on the project. Therefore, my work is not unique. However, there are, nonetheless, some differences in the results with respect to the transcription initiation start site, the structure of the gene, the deduced amino acid sequence, and the expression pattern of the receptor. My work will be described in detail followed by a discussion comparing my results and the results obtained by Kubota's group.

INTRODUCTION

1. History

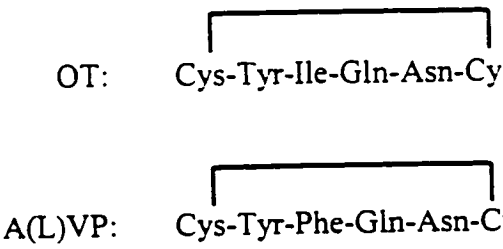
Oxytocin is a hypothalamic nonapeptide, which acts both as a hormone and a neurotransmitter. Extracts from the posterior pituitary were first noted to have "oxytocic" effects at the turn of this century (1). Nearly 50 years later, Du Vigneaud et al. (2) sequenced the endogenous 9-amino acid peptide and synthesized oxytocin. Oxytocin was the first peptide hormone to have its structure determined and the first hormone to be chemically synthesized in a biologically active form (2). Oxytocin is synthesized in the hypothalamus together with the related molecule, arginine vasopressin. Both hormones

are synthesized as part of precursor molecules which contain, in addition to the nonapeptides, a moiety called neurophysin (neurophysin I for OT and neurophysin II for vasopressin). Molecular confirmation of the oxytocin sequence came in the early 1980s upon the cloning of cDNAs and genes encoding the oxytocin-neurophysin preprohormone (3). The cDNA and genes encoding the arginine -vasopressin-neurophysin II preprohormone (4) were cloned just prior to the oxytocin neurophysin preprohormone.

Comparison of the cDNA sequences of the OT-Np I and AVP-Np II precursors reveals a striking homology of 197 nucleotides. This region codes precisely for the highly conserved central portion of the neurophysins. After careful analysis of oxytocin, AVP and their related neurophysins in different species, it was suggested that the hormones emerged by gene duplication from a common ancestor about 450 million years ago (5). Although vasopressin-oxytocin peptides were first characterized as neurohypophyseal hormones in mammals, in recent years, the phylogenetic distribution of this family of peptides has been greatly expanded.

1.1 The ligands: OT and VP

At present, vasopressin-oxytocin peptides have been characterized in over 50 vertebrate species (6). The hormone superfamily comprises more than a dozen variants, mostly nonapeptides with several highly conserved residues including two cysteine residues at positions 1 and 6. These two residues form a disulfide bridge and confer upon the molecules a cyclic conformation. As it can be seen below, vasopressin and oxytocin differ from one another only by two residues:



Although both oxytocin and vasopressin are very similar in structure, they are quite different in their functions. Vasopressin (VP) is mainly involved in hydroosmotic regulation while oxytocin (OT) assumes functions that are essential for mammalian reproduction. The well-characterized 'classical' functions of OT include uterine contractions during parturition and milk ejection during lactation (7). In addition, OT also functions to stimulate endometrial prostaglandin F2 (PGF2) production, pituitary prolactin and LH secretion (8), luteolysis, sperm transport, and natriuresis (9). As a neurotransmitter, OT mediates sexual affiliative and maternal behavior as well as food and salt appetite-suppressing effects (10). OT is the strongest uterotonic substance known and is used widely to induce labor (11). OT is synthesized in the hypothalamic paraventricular (PVN) and supraoptic nuclei (SON) during late embryogenesis (11) and is released into the circulation via the posterior pituitary. OT is also synthesized in the corpus luteum (12), uterus (13), placenta (14), amnion (14), and testis (15). In general, the hormone is released from its site of synthesis by a regulated pathway and in the form of pulses (16), though in some local sites of synthesis, e.g., decidua, oxytocin may be produced via a constitutive, nonpulsatile route.

1.2 Receptors

Oxytocin and vasopressin exert their multiple physiological actions by a family of four different receptors (7). These receptors have been categorized based on the second messenger systems to which they are coupled. Vasopressin receptors (VPRs) have been subdivided into two main classes: V1 and V2 receptors. The V1-type receptor has two members, V1a and V1b. Both V1a and V1b receptors are coupled to phospholipase C (PLC), (7) resulting in an increase in intracellular calcium via stimulation of phosphoinositol (PI) turnover. Activation of the V2 receptor leads to stimulation of adenylate cyclase and results in an increase in intracellular cyclic adenosine monophosphate (cAMP). Oxytocin, on the other hand, exerts its multiple hormonal and neural effects via a single receptor, OTR, coupled to $G\alpha_{q/11}$ and possibly Gi (17, 18). As shown in the table below, oxytocin is a relatively specific ligand for the OTR, but vasopressin is more promiscuous and seems to bind to VPRs as well as to OTRs with high affinity.

Ligand/receptor interactions of the vasopressin/oxytocin system.

Ligand	Receptor	Second messenger	Typical locations
AVP	V1a	IP_3, Ca++	Vessels, liver, CNS, platelets, uterus, kidney, adrenals
	V1b	IP_3, Ca++	Pituitary corticotrophs
	V2	cAMP	Kidney: collecting duct
OT	OT	IP_3, Ca++	Uterus, mammary gland, CNS, kidney, thymus, ovary, testes, pituitary lactotrophs

The signaling mechanisms of vasopressin and oxytocin receptors involve distinct receptor-associated G-proteins. Cloning and sequencing of the cDNAs encoding the V2 receptors (19, 20, 21), the V1a receptors (22, 23, 24), the V1b (25) and the OT receptors

of different species (12, 21, 26, 27, 28, 53) reveal that they are all members of the large family of G-protein-linked membrane receptors.

This superfamily of G-protein-linked membrane receptors consists of a protein molecule that spans the membrane seven times and, therefore, contains seven helical transmembrane domains (29). Topologically, these receptors contain four extracellular domains (E1-E4) and four cytoplasmic domains (C1-C4) which include an extracellular N-terminal domain (E1) and an intracellular C-terminal domain (C4). The transmembrane domains are connected by intracellular and extracellular loops. An alignment of the sequences of the different receptor subfamilies already cloned and sequenced reveals certain common features. In particular, some highly conserved residues encompassing TMD 2 (FQVLPQ), E1 (GPD) and E2 (DCWA and PWG) are unique to the vasopressin/oxytocin receptor family. These residues are quite divergent in other members of the G-protein-linked receptor family. It has been suggested that perhaps these conserved residues may form part of the recognition site for the specific, highly related ligands, oxytocin and vasopressin (11). The strongest divergence occurs in the N- and C-terminal domains as well as in the third cytoplasmic loop, C3. The strong divergence in C3 may reflect differences in G-protein coupling. The cysteine residues in E2 and E3, are present in most G-protein-linked receptors. These two cysteine residues may mediate the formation of a disulfide bond, which is thought to be important for ligand binding. Also highly conserved in the C-termini of all vasopressin and oxytocin receptors are two adjacent cysteine residues. These residues represent a prenylation site by which the C-terminus is properly anchored into the plasma membrane.

Recently, our laboratory has engaged in collaborative studies in order to map the sites of OT/OTR interaction. Based on other ligand/receptor interactions studied and preliminary results from site-directed mutagenesis studies, it seems that peptide ligands, such as oxytocin, bind to the outermost parts of the receptors. These areas include extracellular loops and the tops of transmembrane domains but, they do not generally reach deep down into the transmembrane regions. This confirms the requirement for conserved residues that is seen in the extracellular loops and some of the transmembrane domains in the vasopressin/oxytocin receptor family (unpublished data).

1.3 Signaling

The OTR is coupled to a signaling pathway that leads to phosphoinositide (PI) hydrolysis through phospholipase C (PLC). This, in turn, invokes an increase in intracellular concentrations of two important second messengers: calcium and inositol triphosphate (IP3). The rise in cytoplasmic calcium via intracellular calcium pools causes a wave of pulses within the cell. The impact of this effect is quite spectacular since calcium ultimately leads to the activation of Protein Kinase C (PKC) and, in myometrial cells, calcium is responsible for the stimulation of myosin light chain kinases. These kinases, once activated, are likely involved in initiating the contraction process, culminating in **LABOR**.

The oxytocin receptor is expressed in over twelve different cell types, including uterine myometrium and epithelium, mammary gland myoepithelium, pituitary lactotrophs, certain renal epithelial cells, specific CNS neurons, astrocytes, and T-cell lymphocytes. The oxytocin receptor possesses an outstanding feature which is lacking

even in its closest relatives, the vasopressin receptors: the OTR is expressed in a highly regulated manner, in a stage- and tissue-specific fashion. This is clearly illustrated by the dramatic 10-100-fold upregulation of the uterine OTR immediately prior to parturition, followed by an equally rapid downregulation (30). In contrast, the mammary gland OTR is upregulated during the lactation period, when the uterine OTR expression dips. It is this tissue- and stage-specific regulation of the OTR that results in a switching of targets and enables circulating OT to assume a dual role: to impart uterine contractions during labor and milk ejection during lactation.

1.4 Function

The concentration of oxytocin in maternal plasma during spontaneous parturition does not change significantly before and during labor (31, 32). However, the number of oxytocin binding sites in the pregnant uterus increases dramatically at parturition (30, 33). The same phenomenon is paralleled in the rat and other mammals. In the rat, the sensitivity of uterine smooth muscle to oxytocin increases very abruptly near the onset of labor. This increased sensitivity is ascribed to a marked upregulation of oxytocin receptor concentration (27). It has been postulated that this dramatic upregulation of oxytocin receptors at term represents a trigger for parturition (30), implying that the effectiveness of oxytocin depends on the expression of its receptor. Recently, a group of investigators has provided additional evidence that it is not just the timing and magnitude of oxytocin receptor gene expression that may provoke parturition, but also the fraction of active vs spare receptors (18). In other words, it is not enough for oxytocin to bind to its receptor. The established increased sensitivity of the myometrium to oxytocin could be the result of

both upregulation of oxytocin receptors and an increase in the fraction of receptors coupled to signal transduction components, one of which is G*q*. In fact, the group led by Soloff has demonstrated that in the rat myometrium, virtually all of the oxytocin receptors are found coupled to G-proteins and effector enzymes (18).

Although evidence accumulated over the years substantiate the fact that oxytocin may be the strongest uterotonin partaking in parturition, other uterotonins and cytokines may be involved in this complex event. Some of these include prostaglandins (PGs), leukotrienes, platelet-activating factor (PAF), endothelin-1 (ET-1), interleukin1 (IL-1), tumor necrosis factor (TNF), interleukin-6 (IL-6), and interleukin-8 (IL-8). Some of these components are normal constituents of amniotic fluid . There is a trigger which sets off a series of cascading events culminating in the final result, labor. This trigger probably involves oxytocin receptor upregulation plus coupling to G-proteins. By the same token, it should be kept in mind that the slightest turn of events (i.e., premature OTR upregulation) could offset this delicate balance. The end result is detrimental: premature labor. Premature labor is a leading cause of perinatal mortality and morbidity.

The importance of understanding mechanisms involved in the regulation of cell- and stage-specific expression of the oxytocin receptor cannot be understated. Our laboratory has undertaken this task. Specifically, the questions are directed towards unveiling the essential role of the oxytocin receptor in mammalian reproductive physiology. One way of attaining this goal is to look at the effects of abolishing the oxytocin receptor gene altogether. The method used to accomplish this feat is by constructing a knock-out mouse for the receptor. In order to do so, one must first possess the mouse oxytocin receptor gene sequence. Mice will then be engineered so that the

LacZ gene replaces the OTR coding sequence and is, therefore, under the control of the endogenous OTR promotor. The LacZ gene will be used as a marker to map the precise developmental expression pattern of the oxytocin receptor gene through galactosidase staining. This will shed some light on the physiological relevance of the oxytocin/oxytocin receptor system with regards to the onset of parturition, the timing of parturition, the viability of the progeny, maternal behavior, and whether lactation can ensue.

2. MATERIALS AND METHODS

2.1 Chemicals

All chemicals used were of reagent grade and were purchased from various companies.

2.2 Animals

Non-pregnant and untimed pregnant Balb/c female mice were obtained from Charles River Laboratories (St. Constant, Quebec, Canada). They were housed and fed in a temperature- and humidity-controlled environment with lights on from 0700 h to 1900h. Animals were killed by decapitation under light ether anesthesia as they were giving birth, or after one day lactation. Control (non-pregnant) Balb/c mice were cycling females weighing 20-30 grams. They were sacrificed in the same manner as the parturient and lactating mice. The procedures were approved by the Bioethics Committee of the Royal Victoria Hospital Research Institute.

2.3 Genomic Library

A λ DASH II genomic library pertaining to the Balb/c mouse strain was a generous gift from Dr. A. Peterson (McGill University, Dept. of Molecular Oncology, Royal Victoria Hospital, Montreal, Quebec). The library was contructed using Balb/c mouse genomic DNA which was partially digested with *Sau*3A I and subcloned into the *Bam* HI site of the λ DASH II phage arms.

2.4 Library Screening

In order to isolate recombinant phage containing the mouse OTR gene, the following strategy was adopted: A genomic phage library was subdivided into 40 pools, each containing 2.5 x 10^5 plaque-forming units (pfu). The library was plated on 150 mm plates in the following manner: On day 1, an NZY plate (100 mm) (5 g NaCl, 2 g $MgSO_4 \cdot 7H_2O$, 5 g yeast extract, 10 g casein hydrolysate) was streaked with LE392 bacterial host cells. On day 2, a single colony was chosen from the streaked plate and a fresh overnight culture was prepared. The colony was grown in 10 mL LB medium (5 g NaCl, 5 g yeast extract, 10 g Bactotryptone) supplemented with 0.2% maltose and 10 mM $MgSO_4$ in an incubator at 30°C with gentle shaking. On day 3, the cells were pelleted by centrifugation in a Centra CL2 Thermo IEC rotor at 350 x g for 10 minutes at 4°C and resuspended in sterile 10 mM $MgSO_4$ until an OD_{600} of 1 mL equivalent to 0.5 was obtained. Each phage pool containing 2.5 x 10^5 pfu was added to SM buffer (per liter: 5.8 g NaCl, 2 g $MgSO_4 \cdot 7H_2O$, 50 mL 1M Tris (pH 7.5), 5 mL 2% gelatin) for a final total volume of 100 μL. Added to this was 600 μL of LE392 bacterial host cells. This mixture was then incubated at 37°C for 15 minutes after which 7 mL of prewarmed Top Agar at 48°C (same as NZY plates with 0.7% agarose added) was added, mixed and poured over prewarmed 150 mm NZY plates. The plates were incubated overnight at 37°C. The next day, following incubation, 10 mL of SM buffer was added to each plate. The plates were incubated at room temperature for 2 hours with gentle agitation. The SM buffer containing phage lysates was then collected from each plate and transferred to 15 mL polypropylene (pp) tubes. A few drops of chloroform were added to each tube. The

tubes were then centrifuged for 10 minutes at 4°C in a Sorvall RC-5B Superspeed centrifuge using the SS-34 rotor at a speed of 11 951 x g. The supernatant was carefully decanted into a fresh 15 mL pp tube and recentrifuged again at the same speed and temperature for an additional two minutes. This process was repeated in order to completely pellet bacterial cells. The first round of screening was done using the polymerase chain reaction method (PCR) to screen for the presence of phage containing inserts homologous to the rat OTR. Ten μL of each phage lysate from the 40 pools was boiled for 10 minutes. Each PCR reaction contained 10 μL of phage lysate, 50 picomoles (pmols. of the forward primer, OTR-F_1 .(5'ATGGAGGGCGCGCTCGCAGCCAACTG 3') 50 pmols of the reverse primer, OTR-R_3. (5'CAGGTCGGCGATGCTTAGGTGCTTCAT3') 200 mM Tris (pH 8.4), 0.5 mM KCl, 2 mM $MgCl_2$, 0.4 mM of each of dATP, dTTP, dGTP, dCTP and 1 unit of TAQ DNA Polymerase (Invitrogen Canada Inc., CAT #18038-042). This mixture was denatured for 1 minute at 94°C, annealed at 70°C for 1 minute and extended at 72°C for 1 minute. This cycle was repeated 35 times. The reactions were allowed a final 10 minute extension at 72°C at the end of the 35 cycles. Ten μL of each PCR reaction was then analyzed on a 1.2% agarose gel (with 1% ethidium bromide). A band size of 257 base pairs (bp) corresponding to the first 257 nucleotides (nt) of the OTR coding region was expected. Two positive phage clones (6,30) were identified from the primary screening. A phage titer was performed in order to obtain 5 x 10^3 pfu per 150 mm plate. Each positive clone was subdivided into 20 pools, each containing 5 x 10^3 pfu and replated. Phage DNA was then purified by the polyethylene glycol precipitation method, as follows.

2.5 Preparation of phage DNA

In the next step, phage DNA was prepared by the polyethylene glycol (PEG) method. Ten mL of SM buffer was added to each plate. The plates were incubated at room temperature with gentle agitation for two hours. The supernatant was collected in 15 mL pp tubes. A few drops of chloroform were added to the tubes which were then centrifuged in a Sorvall SC-5B Superspeed centrifuge using an SS-34 rotor at a speed of 11 951 x g at 4°C for 10 minutes. The supernatant was centrifuged again for an additional two minutes to remove any remaining bacterial debris. A mixture of RNAse A/DNAse I was added to the supernatant to a final concentration of 1 mg/mL. This was incubated at 37°C for 30 minutes. One volume of 20% PEG/2 M NaCl in SM buffer was added and vortexed. This was incubated for 60 minutes on ice water. The DNA was pelleted by centrifugation for 20 minutes, at 4°C, at a speed of 11 951 x g in the SS-34 rotor. The tubes were carefully drained for a few minutes, upside-down. The resulting pellet was resuspended in 500 μL SM buffer containing 5 μL of 10% SDS and 5 μL of 0.5 M EDTA, and transferred to an Eppendorf tube. This was vortexed and incubated at 68°C for 15 minutes. In order to eliminate proteins, three organic phase extractions were performed: a phenol, phenol/chloroform and a chloroform extraction. Finally, phage DNA was precipitated with the addition of 1 volume of isopropanol and stored at -20°C overnight. The following day, the phage DNA was collected by centrifugation in an IEC Micromax table top centrifuge at 4°C, at a speed of 21 000 x g for 15 minutes. After drying, each pellet was resuspended in 100 μL H_2O. Five μL of DNA from each subpool of each positive clone (i.e., two positive clones, twenty subpools, therefore a total of forty

reactions) was analyzed by polymerase chain reaction using the same conditions as in the initial screening. Four positive pools (6-16,6-17,30-4, 30-17) were then identified, each of which was replated. Five 150 mm plates containing $2.5x10^3$ pfu per plate from each positive pool were prepared. This time, the pools were screened by the method of Benton and Davis(34). Details of this procedure were as follows: primary and replica phage lifts were made from each plate using Hybond-N hybridization nylon transfer membranes (Amersham Pharmacia Biotech, Montreal, Canada, cat. # rpn.132N) in order to rule out any false positives. The phage were then denatured in a solution containing 1.5 M NaCl and 0.5 M NaOH for two minutes. They were then neutralized by submerging in a mixture of 1.5 M NaCl, 0.5 M Tris-Cl (pH 8.0), for 5 minutes, and finally rinsed for 30 seconds by submerging in 0.2 M Tris-Cl (pH 7.5)/2x SSC. All these steps were done at room temperature. The membranes were finally blotted briefly on Whatman 3 mm filters. The DNA was then crosslinked to the filters using an autocrosslink setting on a Stratalinker TM UV crosslinker or, alternatively, the DNA was baked at 80°C for one and one half to two hours. These membranes were then analyzed by Southern blotting.

2.6 Southern blot analysis

The membranes were pre-hybridized in a solution for a double stranded probe (2x PIPES, 50% de-ionized formamide, 0.5% SDS, 100 μg per mL denatured, sonicated salmon sperm DNA) at 37°C for two hours. The probe used for analysis of the membranes was a mouse specific probe. The mouse-specific probe was prepared by polymerase chain reaction, using mouse Balb/c genomic DNA (2 μg) as a template and the rat specific

primer pair, F1/R3, also giving rise to a 257 bp fragment corresponding to the first 257 nt of the coding region. The resulting fragment was purified and labeled using the Random Primed DNA Labeling Kit (Roche Diagnostics Corporation, Laval, Quebec). According to the manufacturer's instructions, 25 nanograms of the DNA (F1/R3, 257 bp double stranded fragment) was denatured by heating for 10 minutes at 100°C and subsequently cooled on ice. To this DNA, was added 0.5 mM dATP, 0.5 mM dGTP, 0.5 mM dTTP, 2 μg of hexanucleotide mixture, 5 μL of 50 μCi [α^{32}P] dCTP, (3 000 Ci/millimole, aqueous solution) and 1 μL of Klenow enzyme (2 units per μL, in glycerol, 50% (volume/volume)). The final reaction volume was 20 μL. This mixture was incubated for 30 minutes at 37°C. The reaction was terminated by the addition of 2 μL of 0.5 M EDTA or heating at 65°C for 10 minutes. The volume of the probe was brought up to 100 μL with TE (pH 8.0). Elimination of unincorporated deoxyribonucleoside triphosphates was performed by chromatography on a Sephadex G-50 1 cc column. The randomly labeled probe was first denatured at 100°C for ten minutes, then added to a final concentration of 1 x 10^6 cpm per mL. The membranes were then incubated overnight at 37°C. The following day, the membranes were washed in 0.1x SSC/0.1% SDS for 15 minutes at 42°C. This was repeated twice. After washing, the membranes were blotted dry on Whatman 3 mm paper, placed between two sheets of Saran wrap, and exposed overnight at -70°C in a cassette with intensifying screens. The films were then exposed and developed.

2.7 Screening of plates and purification of positive phages

The screening by Southern blotting resulted in four putative positive clones from the subclone 6-17 and three putative positive clones from subclone 30-17. A plug was picked from each of these subclones using the wide end of a 5¾ inch Pasteur pipette. To this plug was added 1 mL of SM buffer and a few drops of chloroform and it was stored in a 5 mL polypropylene tube at 4°C. These plugs can be kept tightly sealed at 4°C for two to three years. In the final round of screening, the positive clones were plated in order to obtain anywhere from 1 to 300 pfu per 150 mm plate. Five plates were plated for each positive clone. Phage lifts were made again from each of these plates in duplicates. They were prehybridized and hybridized using the same solutions, probe, and conditions as the previous round of screening. The membranes were washed at 42°C, three times for 15 minutes in 0.1% SDS/0.1x SSC, exposed overnight, and developed as before. All of the resulting individual positive plugs were picked with a Pasteur pipette and stored in 1 mL of SM buffer containing a few drops of chloroform in 5 mL polypropylene tubes, sealed, and kept at 4°C. After calculating the phage titer from each of these plugs, two small plates (100 mm diameter) per plug were prepared. After an overnight incubation at 37°C, the plates were checked for complete lysis. If they were completely lysed, 5 mL of SM buffer was added. The plates were then kept at room temperature with gentle shaking for 2 hours. The supernatant was collected and phage DNA was prepared by the polyethylene glycol precipitation method as described previously. One-thirtieth of the phage DNA was analyzed on a 0.8% agarose gel containing 1% ethidium bromide. Twelve μL of the DNA solution was then used for restriction enzyme mapping.

2.8 Restriction Enzyme Mapping of the Phage DNA

Each phage DNA clone was digested with 10 units of the following restriction enzymes: *Eco*R I, *Pst* I, *Sst* I, *Xba* I, and *Kpn* I. These restriction enzyme digests were then analyzed by electrophoresis on a 0.8% agarose gel containing 1% ethidium bromide and 1x TAE running buffer. A Southern blot of the phage digests was prepared by transferring the DNA onto a hybond-N nylon membrane (Amersham Pharmacia Biotech, Canada). The DNA was denatured for 30 minutes with gentle shaking at room temperature in 0.5 M NaOH, 0.5 M NaCl. It was neutralized in 1.5 M NaCl, 0.5 M Tris-Cl (pH 8.0) for 15 minutes with gentle shaking at room temperature and, finally, the DNA was transferred overnight to the nylon membrane in 0.25 M NaOH, 1.5 M NaCl by capillary method (35). The DNA was crosslinked to the membrane using a Stratalinker TM UV Crosslinker. The Southern blot was then probed with the random primer labeled pure mouse F1/R3 257 bp fragment. The Southern blot was prehybridized and hybridized at 42°C in 2x PIPES, 50% de-ionized formamide, 0.5% SDS and 100 μg per mL of denatured, sonicated, salmon sperm DNA. The probe was prepared as before using the Random Primed DNA Labelling kit (Roche Diagnostics Corporation, Laval, Quebec). Probe was added to a final concentration of $1x10^6$ cpm per mL. The following day, the Southern membrane was washed twice for 15 minutes at 50°C in 0.1x SSC/0.1% SDS. The Southern was exposed for two minutes at -70°C in a cassette with intensifying screens.

In order to estimate the size of the phage insert, double digests were performed on the phage DNA (*Eco*R I/*Bgl* II, *Eco*R I/*Kpn* I, *Bgl* II/*Kpn* I and *Bgl* II). A Southern

blot of these phage digests was prepared and probed with the same probe, utilizing the same conditions as the single digests. After the analysis of the two phage Southern blots, it was concluded that the phage insert was at least 25 kbp in length. A 6 kbp *Sst* I fragment was subcloned in the pBluescript KS vector (Stratagene). The subcloned fragment was then sequenced manually, on both strands, by the dideoxy chain-termination method (Amersham Pharmacia Biotech, Montreal, Quebec, Cat # 27168201 and ^{35}S dATP, 1250 Ci/mmol) using primers corresponding to flanking vector sequences or internal mouse DNA sequences (i.e., primers F1, R_3). Additional primers were synthesized according to the sequences obtained from each of these sequencing reactions for the length of each of the DNA strands. The sequencing reactions were analyzed on 6% denaturing, polyacrylamide gels (19:1 Bis/Acrylamide).

2.9 Preparation of Mouse Genomic DNA

One gram of Balb/c mouse liver tissue (with gall bladder removed) was ground into a fine powder using a prechilled mortar and pestle. Liquid nitrogen was added to facilitate this. The powder tissue was resuspended in 6 mL digestion buffer (100 mM NaCl, 10 mM Tris-Cl (pH 8.0), 25 mM EDTA (pH 8.0), 0.5% SDS, 0.1 mg/mL proteinase K) and incubated with shaking at 50 °C for 18 hours in a 50 mL pp Falcon tube. The sample was then extracted three times with an equal volume of phenol/chloroform/isoamyl alcohol. The phases were separated each time by centrifugation in a Centra CL2 Thermo IEC centrifuge at a speed of 1 700 x g for 10 minutes. The aqueous (top) layer was transferred to a new tube to which was added ½ volume of 7.5 M ammonium acetate and

two (original, i.e., 6 mL) volumes of 100% ethanol. The stringy DNA precipitate was recovered by centrifugation using the same rotor and speed as the extraction steps, except for the duration, which was 2 minutes here. The DNA pellet was rinsed with 70% ethanol and air-dried. The DNA pellet was then resuspended in TE buffer and its concentration was estimated by dosing at 260 nm.

2.10 Isolation of part of Exon 3

The remaining codons of the gene corresponding to the seventh transmembrane domain and C-terminal domain, were isolated by PCR. Three μg of Balb/c mouse genomic DNA was used as the template. The forward primer was 3printron (5' AGTGGTACCTGGCATGGTTGTTGCTGCGGC 3'), which flanks the 3' region of the 12 kbp intron 2 of the rat OTR gene sequence, and the reverse primer, Bam4-Rev14 (5' TGATGGCTGAGTGACCCGATCAT 3'), corresponds to the 3' untranslated region of the rat OTR gene. The PCR reaction was carried out as previously described except that an annealing temperature of 55°C was used. The resulting fragment was subcloned in the pCRII vector (Invitrogen Canada Inc.) and sequenced as described earlier.

2.11 Isolation of RNA from uterine tissues

Total RNA was prepared from uterine tissues of Balb/c female parturient, one day lactating, and non-parturient mice. The RNA was prepared according to the manufacturer's instructions for the Trizol™ Reagent (Invitrogen Canada Inc.). Approximately 940 mg of uterine tissues from parturient and one day lactating mice as

well as one gram of non-parturient uterine tissues were homogenized in 10 mL of Trizol reagent using a Tekmar tissuemizer (1 mL Trizol per 52 to 100 milligrams of tissue). The homogenized tissues were then incubated for 5 minutes at room temperature to permit the complete dissociation of nucleoprotein complexes. A chloroform extraction was then performed by adding 200 μL of chloroform per mL of Trizol reagent. The tubes were then hand-shaken vigorously for 15 seconds and incubated for 2 to 3 minutes at room temperature. The samples were then centrifuged at 12 000 x g, for 15 minutes, at 4^{0}C, in an IEC Micromax centrifuge. The upper aqueous phase which contained the RNA was then transferred to a fresh Eppendorf tube. The RNA of each sample was precipitated by the addition of 500 μL of isopropanol per mL of Trizol reagent used. The samples were incubated at room temperature for 10 minutes followed by a centrifugation at 12 000 x g for 10 minutes at 4^{0}C. The resulting RNA pellet was then washed with RNAse-free 75% ethanol. The pellet was mixed by vortexing with the 75% ethanol and centrifuged at 7 500 x g for 5 minutes at 4^{0}C. The pellet was air dried and resuspended in RNAse-free water. The concentration was then calculated by measuring the absorbance at a wavelength of 260nm. The samples were then stored at -80^{0}C for further use.

2.12 Preparation of poly (A)$^+$ mRNA

Preparation of poly (A)$^+$mRNA was performed using oligo dT cellulose resin beads (Roche Diagnostics) in 1.5 mL Eppendorf tubes. The resins were prepared in such a way as to allow a maximum fixation of 4 μg of poly (A)$^+$mRNA/μg of oligo dT cellulose. The appropriate amount of oligo dT was weighed out (calculate 2-4% of total mRNA

equals poly (A)$^+$mRNA recovered at the end) and equilibrated with three washes in water. Then three subsequent washes were performed in a solution containing 50 mM Tris-HCl (pH 7.4), 500 mM NaCl, 10 mM EDTA, and 0.5% SDS. Each time this buffer was added, the oligo dT was centrifuged in a tabletop centrifuge until maximum speed was reached (21 000 x g), and then stopped. The supernatant was aspirated and the wash was repeated. The total RNA volume was brought up to 500 μL with DEPC-treated H_2O. This mixture was incubated for five minutes at 85^{0}C. To this, was added one volume of 100 mM Tris-HCl (pH 7.4), 1 M NaCl, 20 mM EDTA, and 1% SDS. This RNA preparation was added to the resin and mixed gently (not vortexed). Binding of poly (A)$^+$mRNA to the resin was performed at room temperature on a rotating platform. The following day, the Eppendorf tubes were centrifuged for two minutes at 500 x g in an IEC Micromax centrifuge. The supernatant was then aspirated. The resin containing the poly (A)$^+$mRNA was washed three times with 1 mL of 50 mM Tris-HCl, (pH 7.4), 500 mM NaCl, 10 mM EDTA, and 0.5% SDS. The efficiency of these washes was maximized by allowing the poly (A)$^+$mRNA to incubate at room temperature for ten minutes while on a rotating platform. In between washes, the Eppendorf tubes were centrifuged to reach maximum speed (21 000 x g), then stopped. In each case, the supernatant was carefully aspirated. The resin was then washed three times with 1 mL of 50 mM Tris-HCl (pH 7.4), 0 .1 M NaCl, and 10 mM EDTA. Again this buffer was added to the resin, allowed to mix for ten minutes on a rotating platform, and the samples were then given a quick centrifugation (reach maximum speed, stop centrifuge). The supernatant was also carefully aspirated. The poly (A)$^+$ mRNA was finally diluted with

two volumes of 250 μL of DEPC-treated H_2O. To ensure that there were no traces of oligo DT cellulose in the final poly $(A)^+$mRNA preparation, the samples were centrifuged after each addition of DEPC-treated H_2O, and the supernatant was transferred to a fresh Eppendorf tube (add DEPC-treated H_2O, shake 5 minutes, quick centrifugation). The 250 μL of DEPC-treated H_2O eluates were pooled and poly $(A)^+$ mRNA was precipitated by the addition of 2.5 volumes of ethanol and 1/10 of a volume of 3M sodium acetate. The samples were then incubated overnight at -20^0C. The following day, the poly $(A)^+$mRNA was collected by centrifugation at maximum speed (21 000 x g) for 15 minutes at 4^0C. The pellet was rinsed in 70% RNAse-free ethanol and air-dried. The resulting poly $(A)^+$mRNA was then resuspended in DEPC-treated H_2O and the concentration of poly $(A)^+$mRNA was estimated from the absorbance at 260nm.

2.13 Northern blot analysis.

A Northern blot of the poly $(A)^+$mRNA from each of the 3 tissues (parturient, 1 day lactation, non parturient) was prepared. All apparati used for the preparation of the Northern blot were rinsed with 10% SDS, then DEPC-treated water. A 1% RNAse-free agarose gel was prepared in the following manner: for each 111 mL of agarose preparation, 27 mL of 37- 41 % formaldehyde, and 6 mL of 0.5 M phosphate buffer (pH 7.2) were added. The running buffer contained 180 mL of 37% formaldehyde and 40 mL of 0.5 M phosphate buffer (pH 7.2). This mixture was brought to one liter with DEPC-treated water. Ten μg of poly $(A)^+$mRNA of each sample was loaded per well in

the agarose gel. Five μg of RNA ladder was used as a marker and prepared according to the manufacturer's instructions (Invitrogen Canada Inc.). All RNA samples and the ladder were then incubated 5 to 10 minutes at 65°C before loading onto the gel. The samples were electrophoresed at 150 volts for 5 to 6 hours with a cooling system at 4°C. At the end of run, the gel was rinsed in water, with gentle shaking, for 15 minutes at room temperature, in order to remove excess formaldehyde. The gel was then gently shaken for 15 minutes at room temperature in 20x SSC (175.3 g of NaCl/88.2 g sodium citrate per liter of water). RNA was then transferred with 20x SSC onto Hybond-N nylon membrane (Amersham Pharmacia Biotech) by capillary method (35). The Northern blot was probed with the same 257 bp PCR fragment corresponding to the first 257 bp of the mouse OTR coding region (primer pair F1/R3). The probe was prepared using the Random Primed DNA Labeling Kit (Roche Diagnostics). After labeling of the probe, unincorporated radioactivity was separated from incorporated radioactivity on a Sephadex G-50 column as in the Southern blot analysis. The Northern blot was hybridized for 16 hours in 50% formamide, 5x SSC, 10x Denhardt's solution, 0.5% SDS, 10 mM sodium phosphate buffer, (pH 7.0), and 100 μg/mL denatured, fragmented, salmon sperm DNA. Probe ($1x10^6$ CPM/mL) was added to the hybridization solution. This was done at 42° C. The following day, the Northern was washed once for 15 minutes with 5x SSPE (175.3 g of NaCl, 27.6 g of $NaH_2PO_4 \cdot H_2O$ and 7.4 g of EDTA in 1L of water)/0.1% SDS at 65° C, once for 15 minutes with 2x SSPE/0.1% SDS at 65°C, and once for 15 minutes with 1x SSPE/0.1% SDS at 65°C. The Northern was then exposed using Kodak XAR film in a cassette with intensifying screens for 48 hours at -70°C.

2.14 5' Rapid Amplification of cDNA Ends (5'-RACE).

5'-RACE was used in order to determine the position of the transcription start site. The first step was to prepare cDNA. Four μg of poly (A)$^+$mRNA was used as a template. To this was added 2 μg of random hexamers ($pd(N)_6$) and water to a final volume of 10 μL. This mixture was annealed at 70°C and slow cooled to 37°C. Then, 1.5 mM of dNTPs, 0.5 units RNasin (Invitrogen Canada, Inc.), 15 mM DTT, 5x RT (reverse transcription) buffer, 400 units of M-MLV (Moloney Murine Leukemia Virus) reverse transcriptase, and water to 20 μL were added. This was incubated at 37°C for one hour followed by heat inactivation at 90°C for 5 minutes. The resulting cDNA was then G-tailed at the 3' end: ten μL was incubated for 30 minutes at 37°C with 50 mM dGTP, 5x Tailing Buffer (Invitrogen Canada Inc.), 15 units of TdT (terminal deoxynucleotidyl transferase, Invitrogen Canada Inc.), and water to 25 μL. The G-tailed cDNA was then amplified. A cocktail containing 2 μL of G-tailed (or untailed) cDNA, 5 pmol of anchorC_{14} primer (5' GCGGCCGCTCGAGTCGACATCGATC_{14} 3'), 45 pmol of anchor primer, 50 pmol of MRACE 1 primer (complementary to: 5' ACGTGCCACACTTTAAAGAGCC3'), PCR buffer, dNTPs, water, and 5 units of Taq DNA Polymerase was prepared. This was denatured at 94°C for 1 minute, annealed at 55 °C for 1 minute, and extended for 2 minutes at 72°C . This cycle was repeated 35 times. Two μL of the resulting PCR fragment was reamplified in order to obtain a more specific band. In this PCR reaction, 50 pmol of the anchor primer, and MRACE 2 primer (complementary to: 5' ATAGACACCTGGACTGAGCG 3') were used. All conditions remained the same except for the annealing temperature, which was raised to

60°C. One μL from this PCR reaction was used for subcloning into the PCRII vector. Positive clones were then sequenced as described earlier.

3. RESULTS

3.1 Screening and Restriction Enzyme Mapping of the Phage Insert.

Screening of a mouse Balb/c genomic library with rat-derived primers followed by restriction enzyme mapping and Southern blotting revealed a phage insert of at least 25 kbp (Fig. 1). An *Sst I* fragment was subcloned into pBluescript and sequenced. Dissection of the sequence showed that it contained 1.3 kbp of 5'-untranslated region, the first 306 codons, and approximately 3.5 kbp of intronic sequence (Fig. 2). Comparison of this sequence with that of the rat OTR sequence confirmed it to be the OTR. The first 306 codons represent the first six transmembrane domains. The rat OTR is interrupted by an intron that is at least 12 kbp in length between the sixth and the seventh transmembrane domains. Since comparison of the mouse sequence with that of the rat revealed striking similarities, it was logical for one to assume the same of the mouse OTR. Since the aim of this project was to obtain 5' flanking region and the full coding region, a different approach was taken to complete the remaining codons of the seventh transmembrane domain, including the stop codon, and part of the 3'-untranslated region of the receptor. Therefore, instead of rescreening the library again, the last codons were amplified by the polymerase chain reaction method. This PCR was performed using mouse genomic DNA as a template and rat specific primers flanking the 5'-end and the 3'-end of the rat exon 3 (3prim intron/Bam 4-Rev 14). The resulting PCR fragment was subcloned into the pCRII vector and sequenced. This fragment contained intronic sequence (from the 12 kbp intron), the last 83 codons, and some 3' - untranslated sequence.

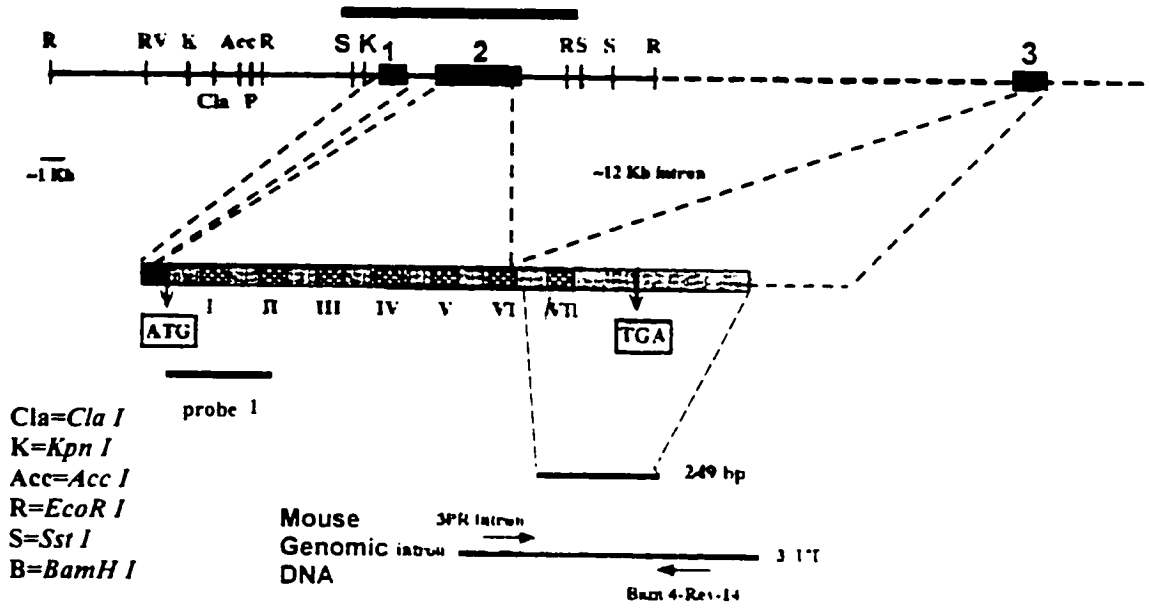
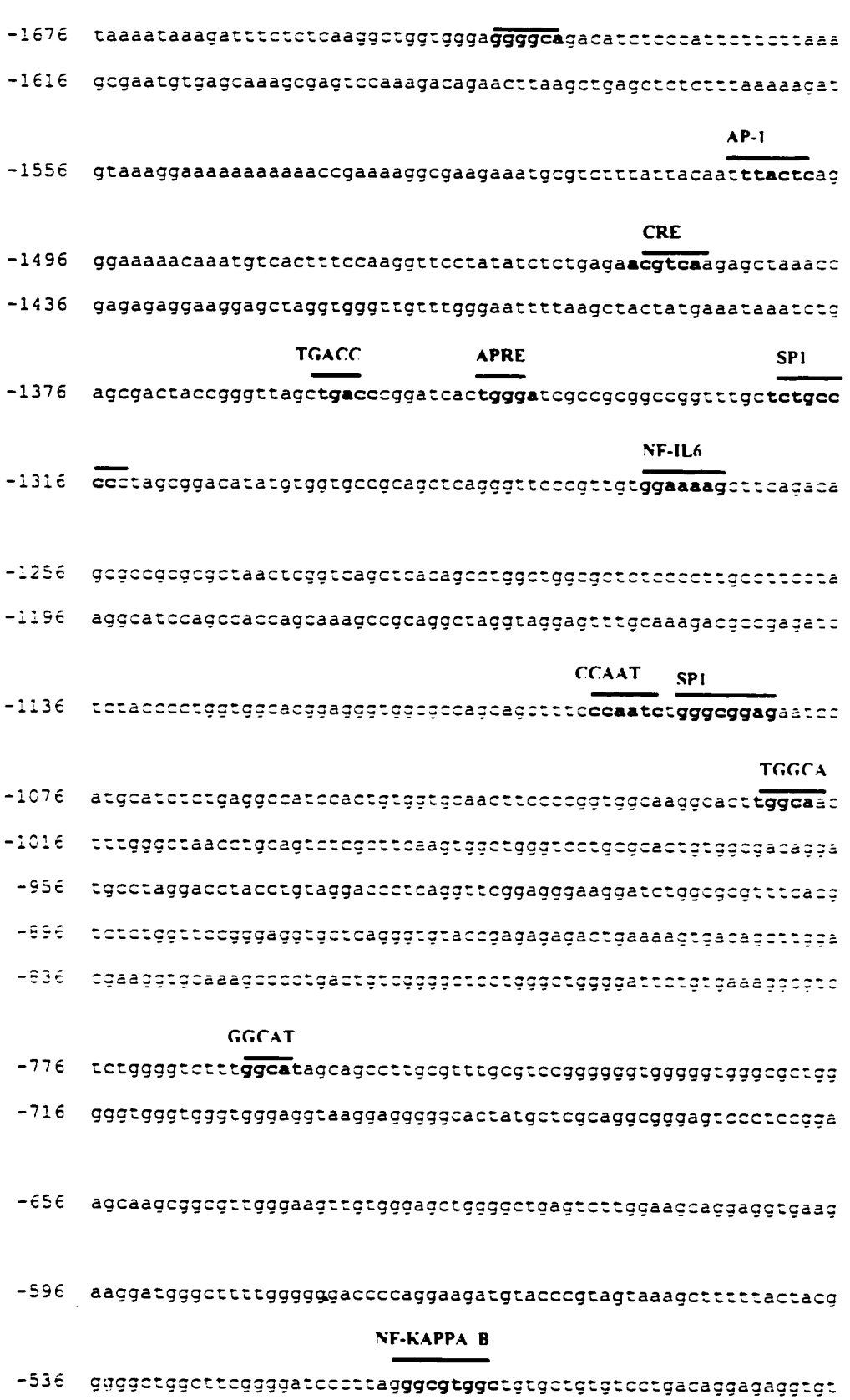
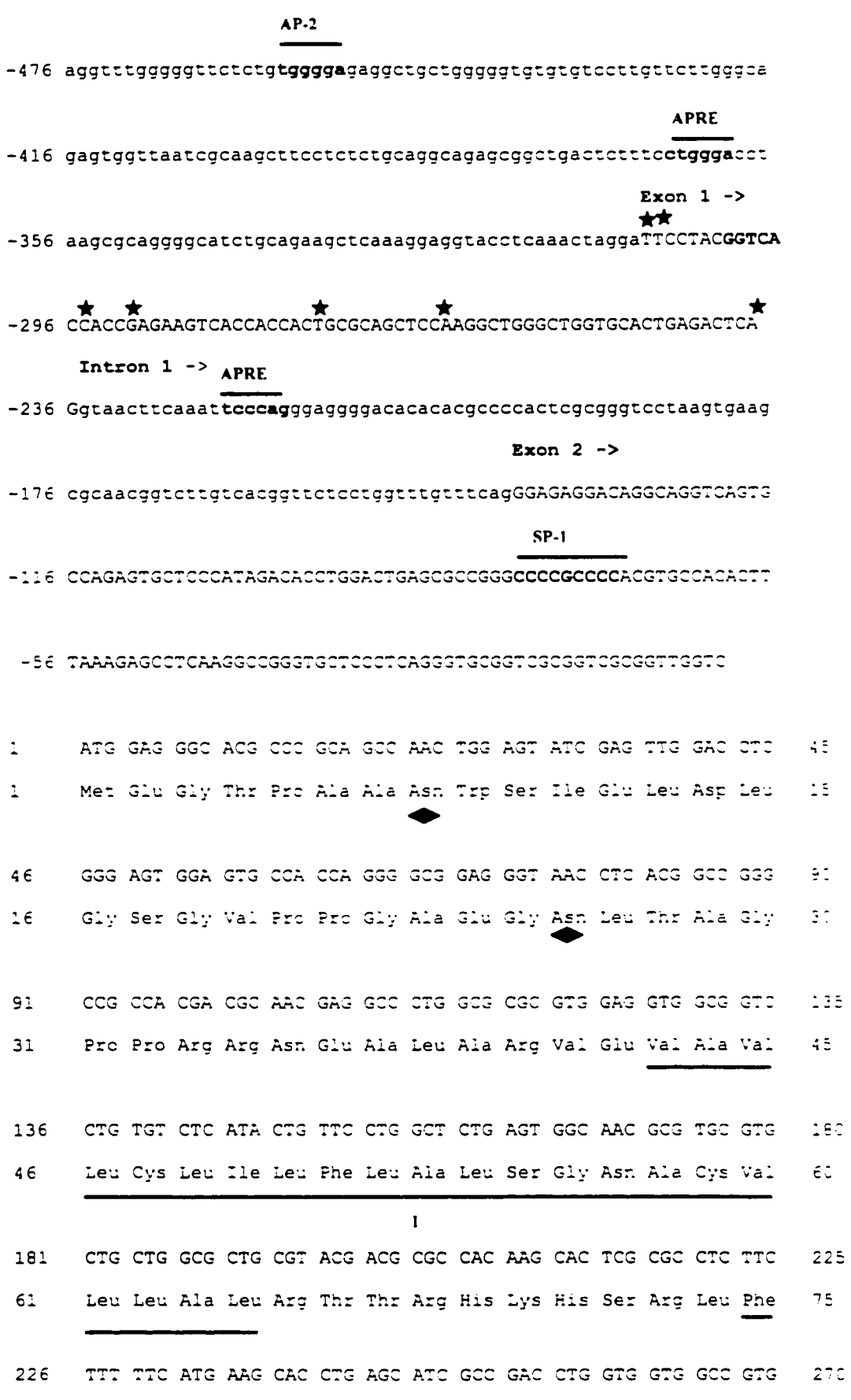


Figure 1: Structure and Restriction Enzyme Map of the Mouse OTR GENE

A Balb/c mouse genomic library was screened with a rat- derived 257 bp probe (probe 1). Restriction enzyme mapping and Southern blotting revealed the phage insert to be at least 25 kbp in length. Exons 1-3 are indicated by heavy bars. The 6.0 kbp *Sst I* fragment contains the first 306 codons corresponding to transmembrane domains I-VI. The sequence of a part of exon 3 was obtained by PCR using rat-specific primers (3PR intron/ Bam4-Rev14) and mouse genomic DNA as a template. Sequencing analysis of this fragment revealed the last 83 codons of the receptor. However, the end of exon 3 was not fully cloned and sequenced and is therefore represented by a dashed line in the figure.


SP-1
-1676 taaaataaagatttctctcaaggctggtgggaggggcagacatctcccattcttcttaaa
-1616 gcgaatgtgagcaaagcgagtccaaagacagaacttaagctgagctctctttaaaaagat
AP-1
-1556 gtaaaggaaaaaaaaaaaccgaaaaggcgaagaaatgcgtctttattacaatttactcag
CRE
-1496 ggaaaaacaaatgtcactttccaaggttcctatatctctgagaacgtcaagagctaaacc
-1436 gagagaggaaggagctaggtgggttgtttgggaattttaagctactatgaaataaatctg
TGACC
APRE
SP1
-1376 agcgactaccgggttagctgacccggatcactgggatcgccgcggccggtttgctctgcc
NF-IL6
-1316 ccctagcggacatatgtggtgccgcagctcagggttcccgttgtggaaaagcttcagaca
-1256 gcgccgcgcgctaactcggtcagctcacagcctggctggcgctctcccccttgccttccta
-1196 aggcatccagccaccagcaaagccgcaggctaggtaggagtttgcaaagacgccgagatc
CCAAT
SP1
-1136 tctacccctggtggcacggagggtggcgccagcagctttccccaatctgggcggagaatcc
TGGCA
-1076 atgcatctctgaggccatccactgtggtgcaacttccccggtggcaaggcacttggcaatc
-1016 tttgggctaacctgcagtctcgcttcaagtggctgggtcctgcgcactgtggcgacagga
-956 tgcctaggacctacctgtaggaccctcaggttcggagggaaggatctggcgcgtttcacg
-896 tctctggttccgggaggtgctcagggtgtaccgagagagactgaaaagtgacagcttgga
-836 cgaaggtgcaaagccccctgactgtcgggggctcctgggctggggattctgtgaaaggcgtc
GGCAT
-776 tctggggtctttggcatagcagccttgcgtttgcgtccgggggggtgggggtgggcgctgg
-716 gggtgggtgggtgggaggtaaggaggggggcactatgctcgcaggcgggagtccctccgga
-656 agcaagcggcgttgggaagttgtgggagctggggctgagtcttggaagcaggaggtgaag
-596 aaggatgggcttttgggggggaccccaggaagatgtacccgtagtaaagctttttactacg
NF-KAPPA B
-536 ggggctggcttcggggatcccttagggcgtggctgtgctgtgtcctgacaggagaggtgt


AP-2
-476 aggtttgggggttctctgtgggga gaggctgctgggggtgtgtgtccttgttcttgggca
APRE
-416 gagtggttaatcgcaagcttcctctctgcaggcagagcggctgactctttcctggga cct
Exon 1 ->
-356 aagcgcaggggcatctgcagaagctcaaaggaggtacctcaaactaggaTTCCTACGGTCA
-296 CCACCGAGAAGTCACCACCACTGCGCAGCTCCAAGGCTGGGCTGGTGCACTGAGACTCA
Intron 1 ->
APRE
-236 Ggtaacttcaaattcccagggaggggacacacacgccccactcgcgggtcctaagtgaag
Exon 2 ->
-176 cgcaacggtcttgtcacggttctcctggtttgtttcagGGAGAGGACAGGCAGGTCAGTG
SP-1
-116 CCAGAGTGCTCCCATAGACACCTGGACTGAGCGCCGGGCCCCGCCCCACGTGCCACACTT
-56 TAAAGAGCCTCAAGGCCGGGTGCTCCCTCAGGGTGCGGTCGCGGTCGCGGTTGGTC
1 ATG GAG GGC ACG CCC GCA GCC AAC TGG AGT ATC GAG TTG GAC CTC 45
1 Met Glu Gly Thr Pro Ala Ala Asn Trp Ser Ile Glu Leu Asp Leu 15
46 GGG AGT GGA GTG CCA CCA GGG GCG GAG GGT AAC CTC ACG GCC GGG 91
16 Gly Ser Gly Val Pro Pro Gly Ala Glu Gly Asn Leu Thr Ala Gly 31
91 CCG CCA CGA CGC AAC GAG GCC CTG GCG CGC GTG GAG GTG GCG GTC 135
31 Pro Pro Arg Arg Asn Glu Ala Leu Ala Arg Val Glu Val Ala Val 45
136 CTG TGT CTC ATA CTG TTC CTG GCT CTG AGT GGC AAC GCG TGC GTG 180
46 Leu Cys Leu Ile Leu Phe Leu Ala Leu Ser Gly Asn Ala Cys Val 60
I
181 CTG CTG GCG CTG CGT ACG ACG CGC CAC AAG CAC TCG CGC CTC TTC 225
61 Leu Leu Ala Leu Arg Thr Thr Arg His Lys His Ser Arg Leu Phe 75
226 TTT TTC ATG AAG CAC CTG AGC ATC GCC GAC CTG GTG GTG GCC GTG 270

76 Phe Phe Met Lys His Leu Ser Ile Ala Asp Leu Val Val Ala Val

II

271 TTC CAG GTT CTC CCG CAG CTG CTG TGG GAC ATC ACC TTC CGC TTC

91 Phe Gln Val Leu Pro Gln Leu Leu Trp Asp Ile Thr Phe Arg Phe

316 TAC GGG CCC GAC CTG CTG TGT CGT CTG GTC AAA TAC TTG CAG GTG

106 Tyr Gly Pro Asp Leu Leu Cys Arg Leu Val Lys Tyr Leu Gln Val

361 GTG GGC ATG TTC GCC TCC ACC TAC CTG CTG TTG CTG ATG TCG CTC

121 Val Gly Met Phe Ala Ser Thr Tyr Leu Leu Leu Leu Met Ser Leu

III

406 GAC CGC TGC CTG GCC ATC TGC CAG CCG CTG CGC TCA CTG CGC CGC

136 Asp Arg Cys Leu Ala Ile Cys Gln Pro Leu Arg Ser Leu Arg Arg

451 CGA ACC GAC CGC CTG GCG GTG CTG GCG ACG TGG CTC GGC TGC CTG

151 Arg Thr Asp Arg Leu Ala Val Leu Ala Thr Trp Leu Gly Cys Leu

IV

496 GTG GCC AGC GTG CCG CAG GTG CAC ATT TTC TCG CTG CGC GAA GTG

166 Val Ala Ser Val Pro Gln Val His Ile Phe Ser Leu Arg Glu Val

541 GCG GAC GGC GTC TTC GAT TGC TGG GCG GTC TTC ATC CAG CCC TGG

181 Ala Asp Gly Val Phe Asp Cys Trp Ala Val Phe Ile Gln Pro Trp

586 GGA CCC AAG GCC TAC GTC ACG TGG ATC ACG CTC GCC GTC TAC ATT

196 Gly Pro Lys Ala Tyr Val Thr Trp Ile Thr Leu Ala Val Tyr Ile

V

631 GTA CCG GTC ATC GTG CTG GCC GCC TGC TAT GGT CTC ATC AGC TTC

211 Val Pro Val Ile Val Leu Ala Ala Cys Tyr Gly Leu Ile Ser Phe

676 AAG ATC TGG CAG AAT CTG CGA CTC AAG ACG GCA GCC GCG GCG GCA

226 Lys Ile Trp Gln Asn Leu Arg Leu Lys Thr Ala Ala Ala Ala Ala

721 GCT GCC GAG GGG AGT GAC GCA GCC GGT GGA GCT GGC CGT GCG GCG

241 Ala Ala Glu Gly Ser Asp Ala Ala Gly Gly Ala Gly Arg Ala Ala

766 TTG GCA CGG GTC AGT AGT GTC AAG CTT ATC TCC AAG GCC AAA ATC

256 Leu Ala Arg Val Ser Ser Val Lys Leu Ile Ser Lys Ala Lys Ile

811 CGC ACA GTG AAG ATG ACC TTC ATC ATT GTT CTG GCC TTC ATC GTG

271 Arg Thr Val Lys Met Thr Phe Ile Ile Val Leu Ala Phe Ile Val

VI

856 TGC TGG ACG CCT TTC TTC TTC GTG CAG ATG TGG AGC GTC TGG GAC

286 Cys Trp Thr Pro Phe Phe Phe Val Gln Met Trp Ser Val Trp Asp

901 GTC AAT GCG CCC AAA GAA G

301 Val Asn Ala Pro Lys Glu

Intron 2 ->

302 gtagtgagggtggagaggaggagggtggcagggccaggccctacttccactatctctgga

303 ctcccagggttctagggtctttctggttccacatcatcatcagtctctggtggcactcca

304 ctcccgaggtccatcagtctctggactttgaggaggtcacttctctgtgtctccaggact

305 acagcatagaaatacccacgagagtgcaaactcttgccagaaggcaggcaagaactccag

306 tgcaggcctggggcaactttcagtgtgaaaattcttattgaggaggggggaaactgaggca

307 ggcgactgtctttcctacagagagacagaatggaagtgggttttcatcctgaccacaggt

308 ctcagg...../..... **>12 kb intron**

agtgtgctggaattcggcttagtggtacctggcatggttgttgctgcggcaaccatgggc

309 cgagtcgggaaggtctgggatgaccccagagtgacacctgttctcttttcctcccgcccc

310 gccccacag

Exon 3 ->

CT TCT GCC TTC ATC ATT GCC ATG CTC TTG GCC AGC CTC AAC AGC

Ala Ser Ala Phe Ile Ile Ala Met Leu Leu Ala Ser Leu Asn Ser

TGC TGC AAC CCA TGG ATC TAC ATG CTC TTC ACG GGC CAT CTC TTC

Cys Cys Asn Pro Trp Ile Tyr Met Leu Phe Thr Gly His Leu Phe

VII

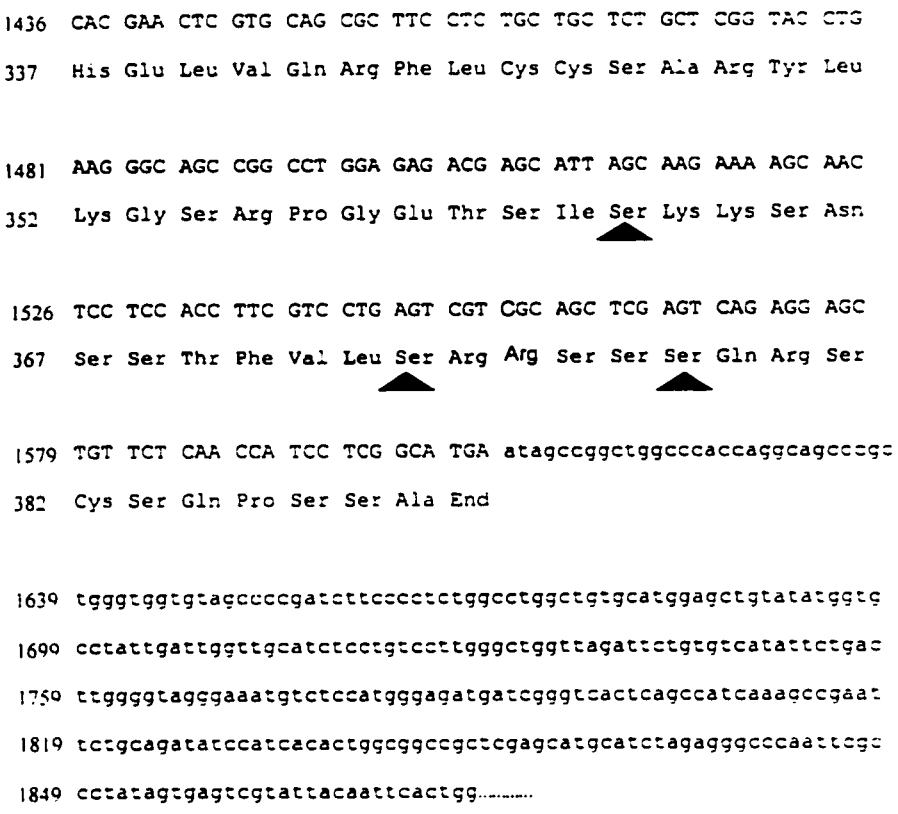


Figure 2: Nucleotide and deduced amino acid sequences of the mouse OTR gene.

Potential binding sites for known transcription factors are indicated. APRE, acute-phase response element; GGTCA or TGACC, half of classical estrogen response element palindrome; Stars (★), transcriptional initiation sites. The predicted transmembrane domains are underlined. (◆) Potential glycosylation site; (▲) potential phosphorylation site.

3.2 Analysis of the coding region

Figure 3 is a representation of the predicted amino acid sequence of the mouse OTR compared to that of the rat and human OTR sequences. It can be seen that the sequences are highly conserved between the three species. Specifically, the mouse and rat OTR are 97% identical. The mouse and human OTR are 94% identical. The seven transmembrane α-helices are highly conserved among the GPCR family members. In this case, the residues in the cytoplasmic and extracellular loops are also highly conserved. The NH_2-terminal extracellular loop is where the three species are most divergent. The two arrows point to the cysteine residues found in all three species, which are thought to form a disulfide bridge linking E2 to E3. Another common feature is the two adjacent cysteine residues within the C-terminal domain. These are palmitoylated and allow proper anchoring of the cytoplasmic tail in the lipid bylayer.

3.3 Determination of the Transcription Start Sites by 5'- End RACE.

Two different amplification schemes were used, each utilizing two nested primer pairs and reverse transcribed mRNA as a template. The expected band size of the cDNA was about 230 bp. Subcloning this fragment and subsequent sequencing led to the determination of the transcription start site. In figure 4, the Anchor C-14 primer sequenceis shown followed immediately by the deoxyribonucleotide indicative of the start of transcription (please see arrow). The figure is a representation of the screening of 10 different PCR subclones. It should be noted that not all subclones ended with the same base as the start site. At the bottom part of the figure is a scheme of the primers

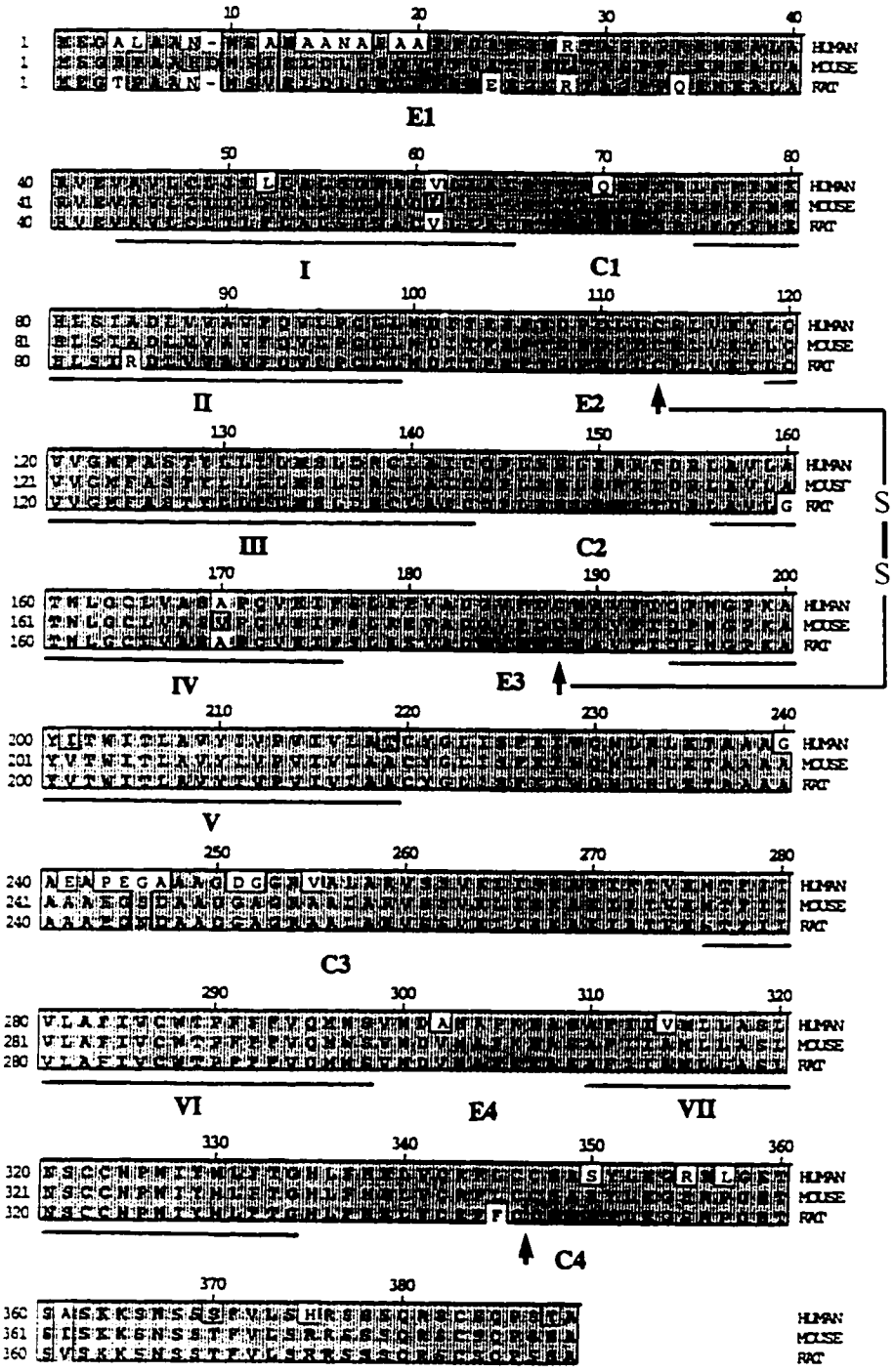
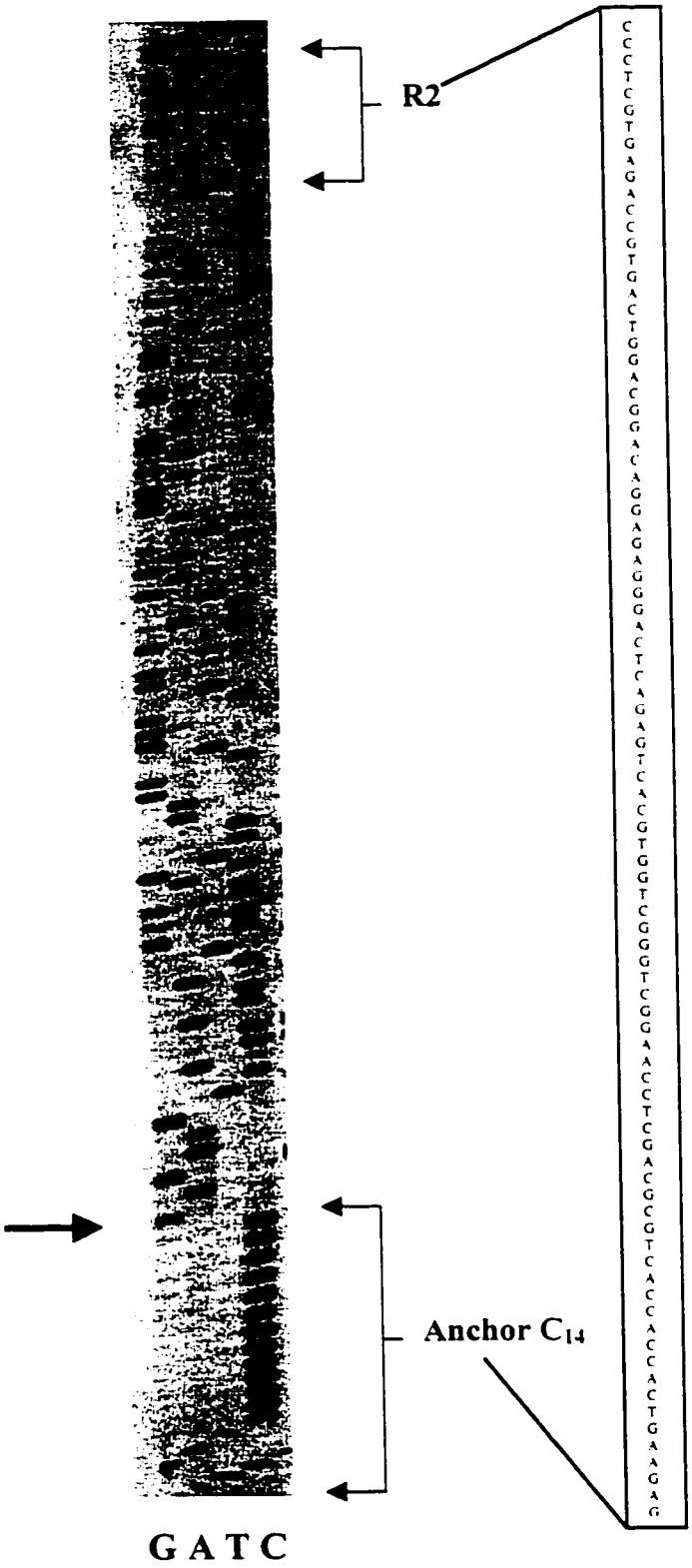


Figure 3: Human, Mouse, and Rat translated OTR sequences.

The transmembrane domains are represented by I-VII. The cytoplasmic (C1-C4) and extracellular (E1-E4) are also shown. The two cysteines in E2 and E3, which are indicated by the arrows, may form a disulfide bridge. Meanwhile, the two adjacent cysteines in C4, are thought to form a prenylation site necessary for anchoring the protein into the plasma membrane.

a)



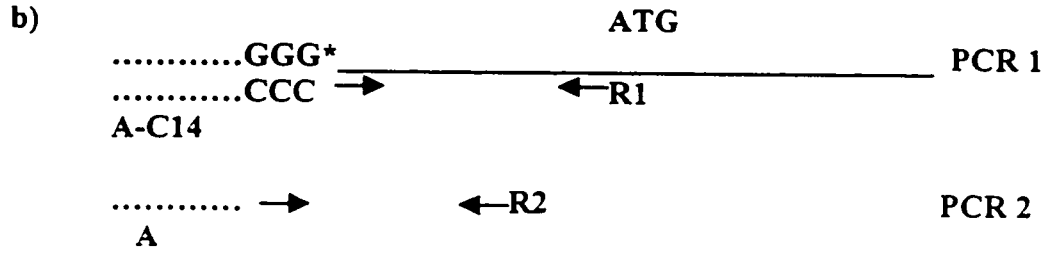


Figure 4: Determination of transcriptional start sites by 5'-END RACE

A) A representation of the dideoxy sequencing reactions of RACE product (sense strand). The sequences corresponding to the upstream primer (anchor plus C_{14} tail) and the downstream primer (R2) , are in between brackets. The gene sequence corresponding to one of the RACE clones (A/R2) is shown in the box. The sequence is read from the bottom going towards the top (i.e., 5' to 3' direction), starting at the transcription start site, which in this case is at position –291 (indicated by left arrow).

B) RACE strategy: The first amplification step, PCR 1, used G-tailed cDNA, a downstream primer, R1, which is 40 bp upstream of the translation start site and an upstream primer (anchor plus C_{14} tail). The second step, PCR 2, used a nested primer, R2, which is 75 bp upstream of the translation start site and the anchor sequence alone as primers. The asterisk represents the transcription start site.

used for the two PCR reactions and the resulting fragments that were subcloned and sequenced. In all of the RACE cDNA clones, nt -139 to -235 were missing. Analysis of this fragment revealed that it is flanked by sequences matching exon/intron boundary consensus sequences (i.e., gt/ag) (36). Therefore, the RACE subclones obtained all have an intron that is consistently spliced out in the uterus. It can be concluded that there is a 97 bp intron in the noncoding region of the mouse oxytocin receptor. Comparison of the sequences obtained from the 5' RACE experiment and genomic sequences reveal that the first exon comprises the noncoding region of the receptor. Exon two contains the coding sequences up to codon 306, which is interrupted by a large intron (at least 12 kpb, data not shown).

3.4 Expression of the OTR gene

Northern blot analysis of the OTR mRNA extracted from uteri of Balb/c mice shows two transcripts: a major one estimated at 5.0 kbp, and a smaller, less prominent one at 3.6 kbp (Fig. 5). These two transcripts may arise as a consequence of alternative use of different polyadenylation sites. The OTR is not expressed in the non-pregnant uterus. However, expression of the gene remains quite elevated in the one -day lactating uterine sample.

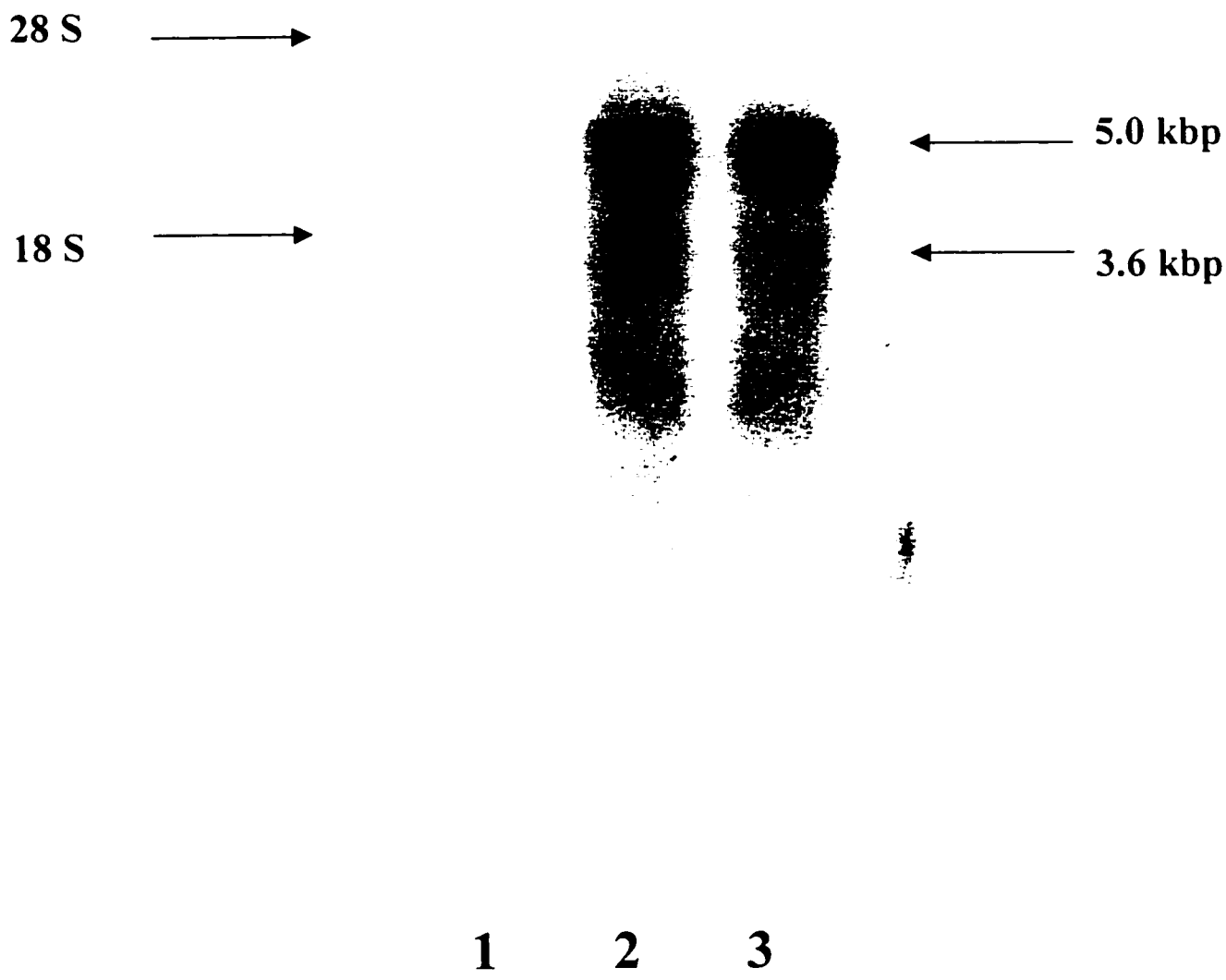


Figure 3: Northern Blot Analysis of OTR mRNA

Each lane contains 10 mg of poly(A)$^+$ RNA extracted from Balb/c mice. Lanes 1, 2, and 3 contain RNA from uterine tissues of non-pregnant, parturient and one day lactating mice, respectively. Two transcripts were detected one at 5.0 kbp and one at 3.6 kbp. The migration positions of 28 S and 18 S rRNA are indicated at left. Size estimates based on RNA size markers are indicated at right.

4. DISCUSSION

The neurohypophyseal hormone, oxytocin, while merely a nonapeptide, has a wide variety of neuronal and physiological functions. Some of these functions are still poorly understood. In particular, its important role in reproduction is highlighted by the possibility that an imbalance in the OT/OTR system may increase the incidence of neonatal morbidity and mortality. It is the strongest uterotonin known to date. Yet, there lies a mystery as to how it imparts its actions. A direct physiological assessment of the important role of the oxytocin/oxytocin receptor system in an organism can be made for example, by knocking out the receptor. One can then attempt to answer the following questions: Can parturition ensue? If parturition is normal, is it delayed? Can the mother lactate? How does the mother behave towards its siblings? The model system used for the construction of the oxytocin receptor knockout is the mouse. Embryonic stem cells derived from the early mouse embryo can be maintained indefinitely in culture in an undifferentiated state. This makes it possible to introduce a genetic alteration into cells in culture. This alteration can then be transferred to the mouse germline for the study of its effects on the intact animal. This allows a targeted mutation to be inserted into an animal's genome without altering the rest of the genetic material. In order to introduce a targeted mutation of the oxytocin receptor, its sequence in the mouse is needed. Specifically, the sequence of at least 2 kbp of both the 5' flanking and intronic regions are required. The coding sequence of the oxytocin receptor (up to codon 306) is replaced by the coding sequence of the LacZ gene and the Neomycin cassette. Therefore, the final construct would consist of the LacZ gene/Neomycin cassette flanked on either side by the mouse oxytocin receptor 5' flanking sequences (at the 5'end) and intronic sequences (at

the 3' end). This construct is transfected into mouse embryonic stem cells by electroporation. The stem cells that have incorporated the construct are then injected into mouse blastocysts, which are reimplanted into a surrogate mother's uterus. Mice are then mated and screened for germline transmission. Most of the endogenous oxytocin receptor coding sequence should be replaced by the LacZ gene/Neomycin cassette through homologous recombination. The expression pattern of the oxytocin receptor in mice can then be analyzed by X-gal staining due to the LacZ gene.

In order to obtain the mouse oxytocin receptor gene sequence, a mouse genomic library was screened. The polymerase chain reaction method was adopted for the first two rounds of screening since rat primers were readily available and the technique yields much faster and more precise results than the conventional mode of screening (34). Another advantage gained by using the polymerase chain reaction method, is that the radioactivity usually used in the screening steps was not required. Kubota *et al.* also screened a mouse genomic library using more conventional techniques (35). The probe I used for screening the library was a rat OTR-derived one (F1/R3), while Kubota's group used human OTR cDNA-derived probes. This difference does not change the outcome of the results with respect to the gene sequence pooled out from the library since the rat and human cDNA sequences are 93% identical. The only difference may be in the length of the gene sequence that is pooled out. A longer probe may in the end hybridize to a longer region found in the library, resulting in a larger fragment that can be subcloned and analyzed. Since the sequences of the mouse OTR gene obtained by Kubota's group and myself were quite similar, it can be concluded that both approaches were worthy.

4.1 Library screening and restriction enzyme mapping of the phage insert.

Screening of a mouse Balb/C genomic library with a 257-bp DNA fragment encoding the N-terminal part of the rat OTR gene homologue (primers (F1/R3)) resulted in the isolation of 2 identical phage clones (16, 30). Restriction enzyme mapping and Southern blotting revealed a phage insert, which was at least 25 kbp long. An *Sst I* fragment was subcloned into the plasmid pBluescript and sequenced. Sequence analysis, and comparison to the rat OTR sequence revealed that the 6.0 kbp *Sst I* fragment contained 1.3 kbp of 5'-flanking region, the first 306 codons corresponding to transmembrane domains 1-6 of the receptor, and approximately 3.5 kbp of intronic sequence. Comparison of this sequence with the rat OTR disclosed the sixth and seventh transmembrane domains to be interrupted by an intron that was at least 12 kbp. This is also seen in the human OTR gene except that the intron is at least 17 kbp long. Nucleotide sequence identity and amino acid alignment surrounding the splice site (i.e., gt/ag) indicate that the >12 kbp intron in the human, rat, and mouse OTR genes are located in homologous positions and are in the same phase. What this means is that the intron interrupts the coding region exactly at the same place in all three species. The same is true for the human V2 vasopressin receptor gene. The remaining 83 codons encoding the seventh transmembrane domain were obtained by PCR method. Rat-specific primers flanking the 5' and 3'ends of the last exon and mouse genomic DNA as a template were used for the reaction (3Printron/Bam 4 Rev-14). After sequence analysis of this extended 249 nt fragment, it was concluded that it was part of the last exon of the mouse OTR gene. This was in contrast to how Kubota's group obtained the last exon. They applied the 3' RACE method instead, where poly $(A)^+$ RNA was the template. The

only resultant clone was subcloned and sequenced. Analysis of this subclone revealed that it contained the C terminus and the entire 3' untranslated region, including polyadenylation signals, of the mouse OTR gene. In both cases, the coding sequence was identical, except for the base at position 1550. Kubota's sequence contains a T, whereas my sequence contains a C. This affects the triplet coding for either cysteine or arginine (please refer to analysis of the coding region in section 4.2).

The OTR gene is present as a single copy in the human genome and was mapped to the gene locus 3p25-3p26.2 (37,38,39).

4.2 Analysis of the Coding Region

The predicted amino acid sequences of the rat and mouse OTR are 97% identical, whereas the human and mouse OTR are 94% identical. There are two consensus sequences for an N-linked glycosylation site (codons 8 and 27, N-X-S/T consensus motif), although the human, porcine, sheep, and bovine OTR genes contain three potential glycosylation sites. This post-translational modification may play a role in receptor biosynthesis and internalization. The coding region also contains numerous potential consensus sequences for phosphorylation sites located in the third intracellular loop and C-terminal tail. For example, Ser/Thr-X-Arg-Lys is the consensus sequence for the phosphorylation by protein kinase C, an important downstream effector enzyme, which has been shown to be coupled to the oxytocin receptor. These phosphorylation sites are thought to play a role in the desensitization, and possibly internalization of the receptor. Another conserved feature is the presence of two cysteine residues in the second and third extracellular loops (E2 and E3), which are thought to form a disulfide

bridge. The role of this potential bridge has not been established yet. There are two well- conserved adjacent cysteines in the C-terminal tail, which form a potential site for prenylation. Prenylation of the C-tail may be important for proper anchoring of the receptor within the lipid bilayer. Therefore, according to sequence analysis and comparison, the 6.0 kbp *Sst I* fragment, is the OTR.

A major difference between my predicted protein sequence and the protein sequence published by Kubota *et al.*, is at codon 375. The rat and human oxytocin receptors, as well as my sequence, all contain an arginine residue at that postion, whereas a cysteine residue has replaced the arginine in the Kubota *et al.* publication. The arginine residue may very well be the *bona fide* residue since it is conserved in the other two species. In the sequences, the codon triplet for the arginine residue is CGC, while the codon triplet for the cysteine residue is TGC. It can be seen that they differ only in the first base of the triplet codon. If indeed an error was made in resolving the sequence by the Kubota group, it may have been due to a stretch of compressed bases on their sequencing gel. However, it seems unlikely that codon 375 would correspond to a cysteine for the following reasons: First of all, arginine is basic whereas cysteine is polar and uncharged. Secondly, a cysteine at that position may cause the formation of a disulfide bridge in an inappropriate location, in turn affecting the overall structure of the receptor. And, since the C-tail is so close to the third cytoplasmic loop, which is thought to be involved in G-protein coupling, a change in the receptor's conformation may have drastic effects on this coupling mechanism. All these postulations suggest that codon 375 codes for arginine in the mouse.

Some investigators have performed site-directed mutagenesis studies to gain a better understanding of which regions or specific amino acid residues may be involved in the formation of the ligand- binding pocket. They have also looked at which regions of the receptor are important for its selectivity profile. Postina et al. (40) demonstrated that the amino terminus of the OTR takes part in hormone binding and probably interacts with the hydrophobic, leucyl residue in position 8 of oxytocin, which lies in the acyclic part of oxytocin. Sequence comparisons of the OTR family showed that the first and second extracellular loops are highly conserved. The first and second extracellular loops and the extracellular regions of the transmembrane domains are believed to form the ligand-binding pocket for oxytocin. The NH_2-terminal domain and the first extracellular loop interact with the linear C-terminal tripeptidic part of oxytocin, whereas the second extracellular loop of the OTR may interact with the cyclic part of the hormone. The OTR has a weak ligand selectivity profile. Other hormones with the same cyclic part will bind with the same affinity as the OT hormone. On the other hand, ligands that differ in the linear tripeptidic part will bind to a much lower extent. Therefore, the cyclic part is what allows OTR to bind OT with higher affinity than other cyclic hormones. Chini *et al.* has shown through the site-directed mutagenesis of vasopressin receptors that when Tyr 115 in the first extracellular loop of the V1a receptor was replaced by Phe, the residue naturally occurring in the OTR, a shift in the pharmacological profile of the V1a receptor towards that of OTR occurred (41). The residues within the transmembrane domains, which form the hydrophobic α- helices are also highly conserved. These residues, along with those forming the extracellular domains are thought to be involved in ligand binding. The residues that are the least conserved lie within the third intracellular loop.

This is the region thought to be responsible for signal transduction via interaction with G-proteins, as is the case for many other G-protein coupled receptors. Oxytocin receptors are coupled to the $G_{q/11}\alpha$ class of GTP binding proteins. This class of proteins, together with $G\beta\gamma$, stimulate the activity of phospholipase C-β isoforms. Two important second messengers are generated: inositol triphosphate and 1,2-diacylglycerol. The former triggers Ca^{2+} release from intracellular Ca^{2+} stores, whereas the latter stimulates protein kinase C, which in turn phosphorylates other target proteins. Calcium is an important cofactor involved in many cellular processes, which include the formation of complexes with calmodulin. The calcium-calmodulin complexes trigger the activation of myosin light-chains, which in turn initiates smooth muscle contractions, as seen in the uterus and myometrium during labor.

4.3 *In VIVO* Expression of the OTR Gene

As shown by Northern blot analysis, the OTR gene is strongly expressed in the mouse uterus compared to nonpregnant uterus at the moment of parturition and remains elevated at one day lactation. The largest transcript detected was estimated at 5.0 kbp. A smaller, less prominent band at 3.6 kbp is also seen in the uterus of both parturient and one day lactating mice. These bands were detected with a mouse-specific probe encompassing the first 257 nucleotides of the coding region. The large 5.0 kpb transcript is also detected by Kubota's group. However, they were unable to detect the smaller sized transcript. Although different methods were used to isolate total RNA, (Kubota's group used guanidine thiocyanate followed by ultracentrifugation, whereas I used the Trizol

method) both methods yield comparable types, amount and quality of RNA. The possibility that the 3.6 kbp is the product of RNA degradation cannot be ruled out. On the other hand, several transcripts were detected in other species (bovine and rat OTR genes). This difference could be explained through alternative use of different polyadenylation sites, degradation, or specific cleavage of the mRNA. And since in my case, the full-length OTR gene was not isolated and sequenced, one can only speculate, that, as in the rat and bovine OTR genes, (26,28) the mouse OTR transcripts seen may differ with respect to the lengths of their 3' untranslated regions.

In this study and in Kubota's publication, it was confirmed that there is a marked increase of OTR mRNA expression in the mouse uterus at the onset of parturition. This seems to be a consistent pattern seen in the several species studied so far (rat, porcine, sheep, bovine). However, in my case, the levels of OTR expression remained elevated in the uterus of a one day lactating mouse. This is contrary to the results obtained by Kubota's group and other publications, (30) where levels are supposed to dip dramatically. There may be three possible explanations for this observation. First of all, it is difficult to precisely delineate the exact end of parturition and beginning of lactation. Secondly, suckling may impart contractions of the uterine smooth muscle as seen during labor. In humans, lactation does invoke uterine contractions. It is these contractions that aid the uterus to regain its prepartum size. Therefore, suckling-induced contractions may require the sustained elevation of OTR expression in the uterus seen in the one day lactating mouse in my experiments. Thirdly, there may be a great variability between the mice sacrificed for extraction of the uterus. One mouse may have experienced many pregnancies compared to another. No one has really looked at the effects that several

gestations can have on the levels of expression of the OTR in the uterus. It may take longer for the levels of OTR expression to return to normal in the uterus of an animal that is experiencing its second or third gestation period. Therefore, further studies are required to determine if the elevated levels of OTR expression seen in my Northern in the uterus of the one day lactating mouse are an artifact.

4.4 Analysis of the 5' UTR

Transcriptional start sites were mapped by 5`-end RACE. Two different two-step amplification schemes were used, each utilizing a nested primer pair. In four of the clones analyzed, the 5`-end of the cDNA was located at nt -307. However, six additional cDNA clones were identified ending at nt -305, -295, -291, -285, -274, and -238. In all of the cDNA clones obtained by RACE, nt –139 to –235 were missing. This finding, together with the fact that this segment is flanked by sequences matching the exon/intron junction consensus sequences (36), indicates that this segment corresponds to an intron that is spliced out in the uterus. However, in the rat OTR, this intron is not consistently spliced out (26). My RACE results seem to considerably contradict results found in Kubota's publication. The transcription start site was determined in both cases by 5`-RACE. Kubota's group also added a primer extension analysis to confirm their results. There are at least two other methods to determine the transcription start site: nuclease S1 protection assay and ribonuclease protection assay. Because RNA is easily degraded, it is more reliable when determining the transcription start site if two independent methods are utilised. The transcription start site in Kubota's publication is located approximately 800 bp upstream of the one I determined. A major drawback when using these

techniques could be the quality of RNA. If the RNA is degraded, the result is an underestimation of the transcription start site due to the formation of a truncated product, such as the cDNA used 5'-RACE. In 5'-RACE, the first step, the reverse transcription of poly $(A)^+$ RNA, is the crucial step. If the poly $(A)^+$ RNA is degraded, the resulting cDNA will be truncated. Another problem encountered during this step is that the reverse trancriptase has difficulty in transcribing RNA templates that have a high degree of secondary structure in their 5' regions. This will also lead to the generation of truncated cDNAs. One way to circumvent this is to increase the temperature of the reverse transcription. In the protocol used by Kubota's group, the reverse transcription was carried out at 42°C compared to 37°C used in my experiment, therefore the probability of a truncated cDNA is higher in my case. However, even though Kubota's group synthesized a longer cDNA fragment (compared to mine), it was not the full length, since it did not contain the transcription start site as determined by primer extension. It should be noted that the transcription start site I obtained is located at almost the same position as the one determined in the rat OTR.

The determination of the transcription start site, as delineated by Kubota's group, when compared to the genomic clone, would suggest that the mouse OTR gene consists of four exons and three introns, as in the human OTR gene. Whereas in my case, I concluded that the mouse OTR gene consists of three exons and two introns, as in the rat OTR gene. Since the mouse is closer phylogenetically to the rat than human, the OTR gene structure is probably more similar in these two species. This difference is very important when it comes time to engineer the mouse OTR knockout. If my conclusion is correct, and the mouse is engineered according to the sequence from the Kubota *et al.*

paper, the knockout would be missing around 800 bases of the 5' flanking region, which may have important regulatory elements that may affect the strength of the promotor. Therefore, it would be an advantage to use my 5' flanking region to engineer the knockout.

The exact start and end of the first intron in my sequence, which corresponds to the second intron in Kubota's sequence, differ. In my case, the sequences found at the junctions of intron 1 with exons 1 and 2 both conform to the consensus sequence of exon/intron boundary splice sites (36). However, the sequences found by Kubota et al. at the junctions between their intron 1 with exons 1 and 2, and between their intron 2 with exons 2 and 3 do not correspond to consensus splice sites. In both cases, my intron 1, which corresponds to their intron 2, are approximately the same size.

4.5 Analysis of the Promoter Region

The transcriptional start site of many genes is usually preceded upstream by a TATA or CCAAT homologous sequence. However, the region encompassing the transcriptional start site I isolated, does not contain any of these sequences. Although most genes contain a TATA-Box, which has been shown to initiate transcription through its presence in the vicinity of the transcription start site, the initiator function of TATA-less genes has been suggested to be carried out by other transcription initiator sequences (42).

The proximal promoter region contains APREs (acute phase response elements) known to be involved in mediating the effects of interleukins (ILs): (I) the T(G/T) NNGNNAA (G/T) motif, recognized by NF-IL6; (ii) the CTGGGA motif, recognized by the acute-phase response factor (APRF); and (iii) the TGGGCA motif, an element recognized by NF-1-related factor that is associated with IL-1-responsive promoter. An

NF-IL6 consensus sequence is present upstream of the main Cap site. The second type of APRE (CTGGGA), occurs twice. This motif represents the binding site for APRF (or STAT 3), a member of the STAT family of transcription factors (43). The proximal promoter region also contains TCF-1 consensus sequences. TCF-1 is a T-cell specific nucleofactor (44) that plays an important role in the establishment of the mature T-cell phenotype. Recent findings, (45), have pointed out the involvement of OT- mediated signaling in T-cell education at the thymus, and ensure further studies concerning the role of OT-OTR physiology in the immune system of the thymus.

The presence of APREs in the OTR gene promoter suggests an acute induction of OTR expression at the onset of labor, a phenomenon mechanistically similar to the fast induction of acute-phase response genes. The uterus is also populated by macrophages and other specific lymphocytes (46,47). Interleukin-1β released from macrophages stimulates the production and release of interleukin-6 by uterine stromal cells. Interleukin-1 is a central pathophysiological mediator of infection-induced premature delivery and preterm delivery can be prevented by an interleukin-1 agonist in mice (48). Therefore, the importance of these acute phase-response elements in the promoter region cannot be disregarded. It can be postulated that under physiological as well as pathophysiological conditions, inflammatory cytokines are important inducers of labor and do so via a mechanism involving cytokine-induced transcriptional activation of the OTR gene.

It has been shown that estrogens induce a strong upregulation of OT binding and OTR mRNA expression in the brain and uterus of castrated rats. (49,50,51,52). There are four half- palindromic motifs of ERE in the rat OTR promoter. In the case of the chicken

ovalbumin gene, activation by estrogen is mediated by several widely spaced half-palindromes (TGACC or GGTCA) (54). Therefore, a potential direct effect of estrogens on OTR gene expression cannot be excluded. As well, the sequence published by Kubota *et al.* contains about 800 extra bases in the 5' flanking region. At the very beginning of their sequence, there lies a classical ERE. Our laboratory has assembled several promoter deletion constructs of the rat OTR in an attempt to define important upstream elements that may regulate the activity of the promoter region. The effects of estrogen were tested in these promoter constructs, which were linked uptream of the luciferase gene in the pGL2 BASIC reporter plasmid. The addition of estrogen (100μM) showed at best only a two-fold induction when compared to the naked plasmid. (Data not shown) Another group of investigators led by Dorsa (55) examined a different segment of the rat OTR promoter, which contains a full ERE, and obtained similar results. It can be proposed that there is one or more undefined accessory protein(s), that may have an impact on estrogen- induced OTR transcription, *in vivo* but not *in vitro*, due to the results obtained with the experiments carried out by our laboratory and by the group led by Dorsa *et al.* The function of the ERE in the mouse OTR promoter remains to be elucidated, however *in vitro* experiments may not be the optimal system to study its role.

5. Conclusion

Within the last decade, the OTR structure has been elucidated in many species and many advances have been made in the field. Cloning the mouse OTR has helped in gaining a better understanding of the OT/OTR system. The mouse OTR gene sequence is quite homologous to other species, therefore, its mechanism of action may also be similar in

these species. However, further questions remain to be answered, these include: 1) what are the molecular mechanisms involved in OTR gene expression and 2) what triggers the onset of parturition under physiological and pathophysiological conditions. The availability of the mouse OTR gene sequence now opens the way for a series of studies utilising gene-targeting techniques and the rodent model system, which has practical advantages over other mammalian species. Only time will tell how many of these questions will be answered.

References

1. Dale, H. H. On some physiological actions of ergot. J Physiol (Lond) 34: 163-206. 1906.

2. Du Vigneaud ,V, Ressler, C, and Trippett S . The sequence of amino acids in oxytocin, with a proposal for the structure of oxytocin. J Biol Chem 205: 949-957, 1953.

3. Land H,Grez M., Ruppert S, Schmale H, Rehbein M., Richter D, and Schütz G. Deduced amino acid sequence from the bovine oxytocin-neurophysin I precursor cDNA.Nature 302: 342-344, 1983.

4. Land H, Schütz G, Schmale H, and Richter D. Nucleotide sequence of cloned cDNA encoding bovine arginine vasopressin-neurophysin II precursor. Nature 295: 299-303 , 1982.

5. Acher R. Molecular evolution of biologically active polypeptides. Proc. R. Soc.: B210, 21-43, 1980.

6. Acher R, Chauvet J, and Chauvet MT. Man and the chimaera. Selective versus neutral oxytocin evolution. Adv Exp Med Biol 395: 615-627, 1995.

7. Zingg H H. Vasopressin and oxytocin receptors in: Sheppard MC, Franklyn JA (eds) Bailliere's Clinical Endocrinology and Metabolism. Bailliere Tindall, London, U.K., 75-76, 1996.

8. Samson WK, Lumpkin MD, McCann SM. Evidence for a physiological role for oxytocin in the control of prolactin secretion. Endocrinology 119: 554-560, 1986.

9. Verbalis JG, Mangione MP, Stricher EM. Oxytocin produces natriuresis in rats at physiological plasma concentration. Endocrinology 128: 1317-1322, 1991.

10. Insel TR. Oxytocin-a neuropeptide for affiliation: evidence from behavioral, receptor autoradiographic and comparative studies. Psychonueroendocrinology 17: 3-35, 1992.

11. Gimpl G and Fahrenholz F. The oxytocin receptor system; Structure, function, and regulation. Physiological Reviews 81, No. 2, 629-683, 2001.

12. Riley PR, Flint APF, Abayasekara DRE, and Stewart HJ. Structure and expression of an ovine endometrial oxytocin receptor cDNA. J of Mole Endo, 15: 195-202, 1995.

13. Lefebre DL, GIAID A, Bennett H, Lariviere R, and Zingg H H. Oxytocin gene expression in rat uterus. Science 256: 1553-1555, 1992.

14. Lefebre DL, Lariviere R, and Zingg HH. Rat amnion: a novel site of oxytocin production. Biol Reprod 48: 632-639, 1993.

15. Bathgate RAD and Sernia C. Characterization and localization of oxytocin receptors in the rat testis. Journal of Endocrinology, 141 : 343-352, 1994.

16. Gainer H, and Wray S. Cellular and molecular biology of oxytocin and vasopressin. In Physiology of Reproduction. E. Knobil and J.D. Neill, eds. (Raven Press, New York), 1099-1130,1994.

17. Ku CY, Qian A, Wen Y, Anwer K, Sanborn BM. Oxytocin stimulates myometrial guanisine triphosphatase and phospholipase-c activities via coupling to G $alpha_{q/11}$ Endocrinology 136 : 1509-1515, 1995.

18. Strakova Z, Soloff MS. Coupling of oxytocin receptor to G proteins in rat myometrium during labor: G*i* receptor interaction. Am J Physiol, 272: E870-E876, 1997.

19. Seibold A, Brabet P, Rosenthal W, Birnbaumer M. Structure and chromosomal localization of the human antidiuretic hormone receptor gene. American J of Human Genetics, 51 : 1078-1083, 1992.

20. Lolait SJ, AM OC, Mc Bride OW, Konig M, Morel A, Brownstein MJ. Cloning and characterization of a vasopressin V2 receptor and possible link to nephrogenic diabetes insipidus. Nature 357 : 336-339, 1992.

21. Gorbulev V. Molecular cloning and functional characterization of V2 [8-Lysine] vasopressin and oxytocin receptors from a pig kidney cell line. Eur J Biochem 215: 1-7,1993.

22. Thibonnier M.Molecular cloning, sequencing, and functional expression of a cDNA encoding the human V1a vasopressin receptor. J Biol Chem 269: 3304-3310, 1994.

23. Hirasawa A. Cloning, functional expression and tissue distribution of human cDNA for the vascular-type vasopressin receptor. Biochem Biophy Res Commn 203: 72-79, 1994.

24. De Keyzer Y. Cloning and characterization of the human V3 pituitary vasopressin receptor. FEBS Lett. 365: 215-220, 1994.

25. Sugimoto T. Molecular cloning and functional expression of a cDNA encoding the human V1b vasopressin receptor. J Biol Chem 269: 27088-27092, 1994.

26. Rozen F, Russo C, Banville D, Zingg HH. Structure, characterization and expression of the rat oxytocin receptor gene. Proc Natl Acad Sci USA 92: 200-204, 1995.

27. Kimura T, Tanizawa O, Mori K, Brownstein MJ, Okayama H. Structure and expression of a human oxytocin receptor. Nature 356 : 526-529, 1992.

28. Bathgate R, Rust W, Balvers M, Hartung S, Morley S, Ivell R. Structure and expression of the bovine oxytocin receptor gene. DNA & Cell Biology 14: 1037-1048, 1995.

29. Baldwin JM. The probable arrangement of the helices in G protein-coupled receptors. EMBO Journal 12: 1693-1703, 1993.

30. Soloff MS, Alexandrova M, Fernstrom Mj. Oxytocin receptors: triggers for parturition and lactation. Science 204: 1313-1315, 1979.

31. Dawood MY, Yeidorkala K, Trevedi D, and Fuchs F. Oxytocin in maternal circulation and amniotic fluid during pregnancy.J of Clin Endocrinol Metab. 49: 429-434, 1979.

32. Leake RD, Weitzman RE, Glantz TH, and Fisher DH. Plasma oxytocin concentrations in men, nonpregnant women, and pregnant women before and during spontaneous labor.J of Clin Endocrinol Metab.53: 730-733, 1981.

33. Fuchs AR, Fuchs F, Husslein P, and Soloff MS. Oxytocin receptors in the human uterus during pregnancy and parturition.Am J Obstet Gynecol. 150: 734-741.1984.

34.Benton W.D., and Davis R.W. Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196:180-182. 1977.

35.Shambrook J, Fritsch E F, and Maniatis T. Molecular Cloning. A Laboratory Manual. Second Edition. 9.34-9.47. 1989.

36. Mount SM. A catalogue of splice junction sequences. Nucleic Acids Res. 10.459-472, 1982.

37. Inoue T, Kimura T, Azuma C, Inazawa J, Takemura M, Kikuchi T. Kubota Y. Ogita K, and Saji F. Structural Organization of the human oxytocin receptor Gene. J Biol Chem 269: 32451-32456, 1994.

38. Michelini S, Urbanek M, Dean M, and Goldman D. Polymorphism and genetic mapping of the human oxytocin receptor gene on chromosome 3. Am J Genet 60: 183-187, 1995.

39. Simmons-CF, Clancy Te, Quan R, and Knoll JH. The oxytocin receptor gene (OXTR) localizes to human chromosome 3p25 by fluorescence in situ hybridization and PCR analysis of somatic cell hybrids. Genomics 26: 623-625, 1995.

40. Postina R, Kojro E, and Fahrenholz F. Separate agonist and peptide antagonist binding sites of the oxytocin receptor defined by their transfer intothe V2 vasopressin receptor. J Biol Chem 271: 31593-31601, 1996.

41. Chini B, Mouillac B, Ala Y, Balestre M-N, Trumpp-Kallmeyer S, Hoflack J, Elands J, Hibert M, Manning M, Jard S, and Barberis C.Tyr115 is the key residue for determining agonist selectivity in the V1a vasopressin receptor.EMBO J. 14, 2176-2182, 1995.

42. Wobbe CR, and Struhl K. Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro. Mol Cell Biol 10 : 3859-3867, 1990.

43. Akira S, Nishio Y, Inoue M, Wang X.-J, Wei S, Matsusaka T, Yoshida K, Sudo T, Naruto M, and Kishimoto T. Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediat signaling pathway. Cell 77: 63-71, 1994.

44. Oosterwegel M, van de Wetering M, Dooijes D, Klomp L, Winoto A, Georgopoulos K, Meijlink F, and Clevers H. Cloning of murine TCF-1, a T cell-specific transcription factor interacting with functional motifs in the CD3 and T Cell receptor enhancers. J Exp Med 73: 1133-1142, 1991.

45. Greenen V, Vandersmissen E, Martins H, Goxe B, Kecha O, Legros JJ, Lefebre PJ, Benhida A, Rentier-Delrue F, and Martial JA. Cellular and molecular aspects of Thymic T-cell education to neurohypophysial principles. In: Neurohypophysis. Recent progress of Vasopressin and Oxytocin (Saito T, Kurokawa K, and Yoshida S, Eds) pp. 309-319, Elsevier, Amsterdam, 1995.

46. Hunt JS. Immunologically relevant cells in the uterus.Biol Reprod 50: 461-466, 1994.

47. Vince GS, Starkey PM, Jackson MC, Sargent IL, and Redman CWG. Flow cytometric characterisation of cell populations in human pregnancy decidua and isolation of decidual macrophage. J Immunol Methods 132: 181-189, 1990.

48. Romero R, and Tartakovsky B. The natural interleukin-1 receptor antagonist prevents interleukin-1-induced preterm delivery in mice. Am J Obstet Gynecol 167: 1041-1045, 1992.

49. Fuchs A-R. in Oxytocin: Clinical and laboratory studies, eds. Amico J, and Robinson AG. (Elsevier, New York), pp. 207-235, 1985.

50. McCarthy MM, Kow LM, and Pfaff DW. Speculations concerning the physiological significance of central oxytocin in maternal behavior.Ann N.Y. Acad Sci 652: 70-82, 1992.

51. Breton C, Pechoux C, Morel G, and Zingg, HH. Oxytocin receptor messenger

ribonucleic acid: characterization, regulation, and cellular localization in the rat pituitary gland. Endocrinology 136: 2928-2936, 1995.

52. Larcher A, Neculcea J, Breton C, Arslan A, Rozen F, Russo C, and Zingg HH. Oxytocin receptor gene expression in the rat uterus during pregnancy and the Estrous cycle and in response to gonadal steroid treatment. Endocrinology 136: 5350-5356, 1995.

53. Kubota Y, Tadashi K, Kazumasa H, Hashimoto K, Tokugawa KN, Azuma C. Saji F, and Murata Y. Structure and expression of the mouse oxytocin receptor gene. Mol and Cell Endo 124: 25-32, 1996.

54. Kato S, Tora L, Yamauchi J,Masushige S, Bellard M, and Chambon P. A far upstream estrogen response element of the ovalbumin gene contains several half-palindromic 5'-TGACC-3' motifs acting synergistically. Cell 68,731-742, 1992.

55.Bale TL and Dorsa DM. Cloning, novel promoter sequence, and estrogen regulation of a rat oxytocin receptor gene. Endocrinology 138(3):1158-8. 1997.