Molecular Cloning, Characterisation, and Gene Targeting Vector Design for the Murine Preprotachykinin-A (PPT-A) and Neurokinin-1 Receptor (NK1-R) Genes

A thesis submitted for the degree of
Master of Science at the
University of Glasgow

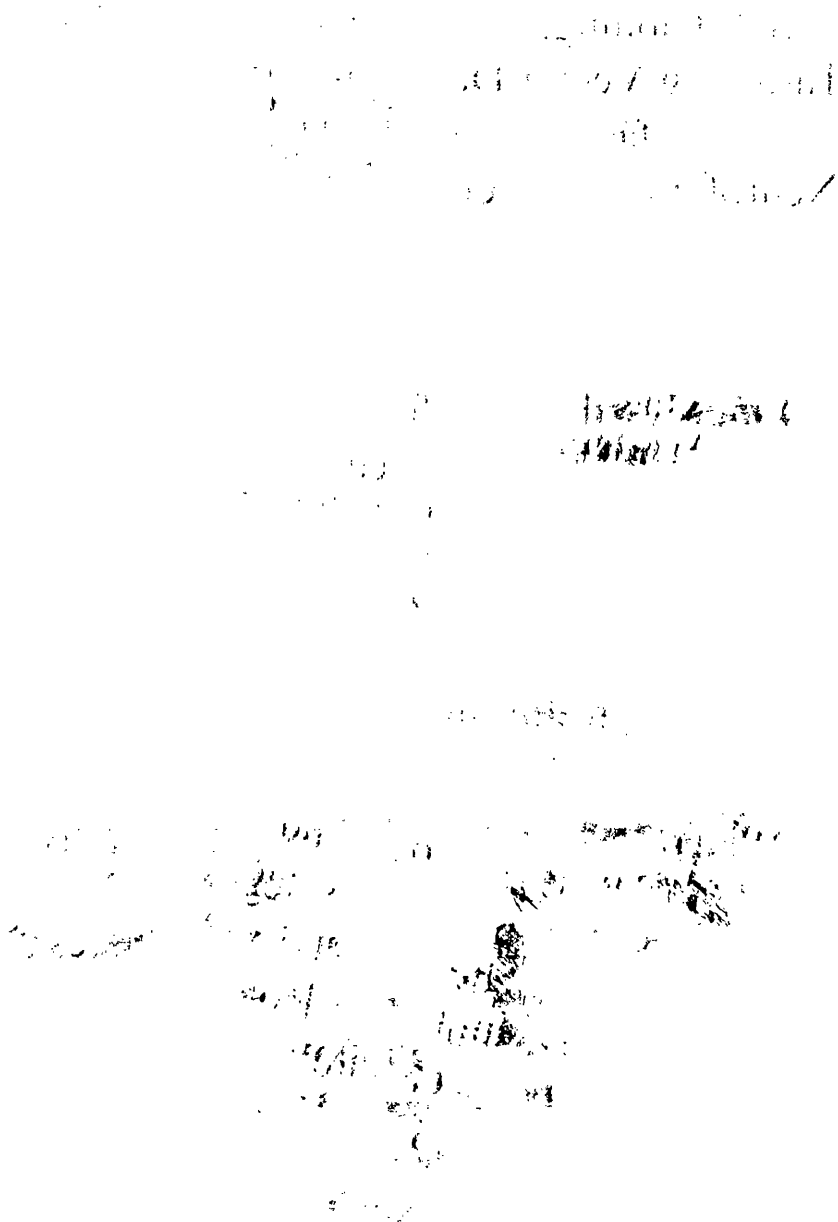
By

Colin Smith MacKay

Robertson Institute of Biotechnology
Division of Molecular Genetics
Institute of Biomedical and Life Sciences
Anderson College
The University
Dunbarton Road
Glasgow
G11 6NU

1997

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346



The research reported in this thesis is my own original work except where otherwise stated and has not been submitted for any other degree.

Colin Smith MacKay, 1997.

Dedicated to the lasting memory
of Annetta Smith and Alastair MacKay.

ACKNOWLEDGEMENTS

I would like to thank my employer, Amersham Pharmacia Biotech, for its support and understanding while I have been writing this thesis.

I would like to sincerely thank my project supervisor, Professor R. Wayne Davies for giving me the opportunity to achieve.

Special thanks to Ed Gallagher and Denise Donald for their willingness to help and for allowing me to learn from their expertise. Thanks also to Armand Savioz, Lynne Loughlan, Vibs Rowell, and Thora Glencorse for their help in the lab and Yiming Song for his help with the vectors. Thanks also go to John Cameron, Martin McBride and Diane Sinclair for their help with this thesis. Thanks to everyone in the lab, and indeed the department, who have contributed in one way or another over the years, and who have made my time in the department enjoyable.

Thanks to Dr. Nakanishi, Kyoto University, for the rat NK-1R cDNA.

Many thanks to my family for their persistent encouragement.

And thank you Maria.

This CASE studentship was funded by the Science and Engineering Research Council and sponsored by Glaxo (now GlaxoWellcome) Research and Development Ltd. I acknowledge their support.

ABSTRACT

The neuropeptide substance P, recently re-named neurokinin-1 and abbreviated to NK-1, is a member of a family of neuropeptides called the tachykinins. Those neuropeptides are distributed throughout both the mammalian central and peripheral nervous systems and have likely roles in a myriad of cellular functions ranging from nociception to development. Neurokinin-1 has been localised to several interesting regions of the brain, regions that show phenotypic changes in line with the progression of clinical neurodegenerative diseases such as Huntingdons disease for example. NK-1 is encoded by the preprotachykinin-A (PPT-A) gene. The neurokinin-1 receptor (NK-1R) gene encodes the G-protein-coupled receptor for NK-1. In order to investigate the proposed involvement of the PPT-A and NK-1R genes in diseases of the nervous system and other cellular activities, it is important to firstly understand the normal function and regulation of those genes at the molecular level.

This project sets out to examine the normal function and regulation of the PPT-A and NK-1R genes by molecularly cloning the murine PPT-A and NK-1R genes and provisionally characterising those genes. Gene targeting vectors for both genes will then be described. Vectors were designed for use in homologous recombination experiments, which is the first step required in order to generate transgenic mice bearing specifically altered forms of those genes.

This objective was pursued by firstly cloning the murine NK-1 precursor (PPT-A) gene from a genomic bacteriophage library. A 423 bp probe was generated by polymerase chain reaction (PCR) from mouse genomic DNA template and was subcloned and sequenced. Optimal hybridisation conditions were then determined and the probe used to isolate lambda bacteriophage clones from a library. Similarly, positively-hybridising clones of the NK-1R (receptor) gene were isolated from another bacteriophage library using an 865 bp rat NK-1R cDNA-derived probe.

Secondly, the PPT-A bacteriophage clones were analysed by restriction mapping, PCR and by hybridisation to exon-specific probes. The exons of the murine PPT-A gene were then subcloned, restriction mapped and positioned relative to each other on linear restriction maps. Some mouse PPT-A sequence was obtained and aligned with other species to determine the extent of identity and homology. A similar approach was used for the NK-1R bacteriophage clones.

Thirdly, gene targeting vectors for use in homologous recombination experiments in mouse embryonic stem cells were then constructed for each of the genes. Multiple cloning and subcloning steps were required to do this. Several endonuclease recognition sequences were altered by ligation of DNA linker sequences. A *neomycin* (G418) resistance gene with its own promoter was introduced into PPT-A exon 3, or in the the other vector, the *neomycin* resistance gene was introduced in place of a deleted 2.4 kb genomic sequence containing PPT-A exons 3 and 4. For the NK-1R vector, exons 3 and 4 were deleted and replaced by the *neo* resistance gene. In each vector, a thymidine kinase (TK) counter-selection gene from *herpes simplex* I virus (HSV1) was also inserted.

ABBREVIATIONS

ATP	adenosine 5'-triphosphate
APS	ammonium persulphate
ap	ampicillin
BSA	bovine serum albumin
bp	base pairs
cDNA	complememtary deoxyribonucleic acid
CIP	calf intestinal alkaline phosphatase
Ci	curie
cm	centimetre
cpm	counts per minute
dpm	disintegrations per minute
D	dalton
DMSO	dimethylsulphoxide
DEPC	diethyl pyrocarbonate
DTT	dithiothreitol
DMF	dimethylformamide
DNAase	deoxyribonuclease
dNTP	deoxynucleotide 5'-triphosphate
dATP	deoxyadenosine 5'-triphosphate
dGTP	deoxyguanosine 5'-triphosphate
dTTP	deoxythymidine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
ddNTP	dideoxynucleotide 5'-triphosphate
ddATP	dideoxyadenosine 5'-triphosphate
ddGTP	dideoxyguanosine 5'-triphosphate
ddTTP	dideoxythymidine 5'-triphosphate
ddCTP	dideoxycytidine 5'-triphosphate
ssDNA	single stranded deoxyribonucleic acid
dH_2O	distilled water
EDTA	ethylene diamine tetra-acetic acid (disodium salt)
EtBr	ethidium bromide
EtOH	ethanol
g	centrifugal force equivalent to gravitational acceleration
IPTG	isopropylthio-β -D-galactoside

HCl	hydrochloric acid
hrs	hours
kDa	kilodalton
kg	kilogramme
kb	kilobases
l	litre
LMP	low melting point (agarose)
MOPS	morpholinopropanesulphonic acid
mA	milliamp
mg	milligramme
ml	millilitre
mM	millimolar
mmol	millimole
M	molar
mins	minutes
Mr	molecular weight
ng	nanogramme
NK-1	neurokinin-1 protein
NK-1R	neurokinin-1 receptor gene
NZCYM	casein, amino acid, yeast extract media
nt	nucleotide
oligo	oligodeoxyribonucleotide
OD	optical density
ORF	open reading frame
°C	degrees centigrade
PPT-A	preprotachykinin-A gene
pmol	picomole
p.f.u.	plaque forming units
PCR	polymerase chain reaction
p.s.i.	ponds per square inch
poly(A)+	polyadenylated
PEG	polyethylene glycol
RNAase	ribonuclease
RF	ribonuclease free
RNA	ribonucleic acid
mRNA	messenger ribonucleic acid
rRNA	ribosomal ribonucleic acid
SAP	Shrimp Alkaline Phosphatase

TFB	transformation buffer
tRNA	transfer ribonucleic acid
rpm	revolutions per minute
RT	room temperature
SPR	substance P receptor gene
SSC	standard saline citrate
sec.	second
SAP	shrimp alkaline phosphatase
sp. act.	specific activity
SDS	sodium dodecyl sulphate
Tet	tetracycline
TEMED	NNN-N'-tetraethylethylenediamine
Tris	tris (hydroxymethyl) methylamine
temp.	temperature
µ	units
µCi	microcurie
µl	microlitre
µM	micromolar
µm	micron
UV	ultraviolet
vol	volume
V	volt
v/v	volume by volume
w/v	weight by volume
W	watt
Xgal	5-bromo-4-chloro-3-indoyl-β –D –galactoside
%	percentage

TABLE OF CONTENTS

CHAPTER 2: MATERIALS AND METHODS

LIST OF FIGURES AND TABLES

CHAPTER 1

INTRODUCTION

The objective of this chapter will be to provide an overview of the background information and literature which both prompted and acted as the platform on which the experimental investigations detailed in this thesis were built. Information which was both pertinent and available when this research project was first designed will be described below, as will the influential research findings which were published while the project was ongoing and which subsequently impacted on the progress of the research.

The overview sets the scene by briefly introducing the modes of communication in the mammalian nervous system and then focuses on one important mode in particular, namely the neuropeptides as a signalling system. To be more specific, the tachykinin family of neuropeptides are of specific interest. The discovery and the history to date of the tachykinin neuropeptide neurokinin-1 (NK-1, formerly substance P) is then detailed, as is the equivalent chronology of the G-protein linked, membrane-bound, NK-1 receptor (formerly substance P receptor). The basis for the study of this neuropeptide, and its corresponding receptor, will then be described with respect to the clinical relevence and proposed therapeutic value of such research. Areas where further research would be useful and the principle objective of the research detailed in this text is also discussed. The field of modern homologous gene targeting technology, and in particular the approach adopted for the research presented below, will also be explored.

1.1. COMMUNICATION IN THE MAMMALIAN NERVOUS SYSTEM

The brain is without doubt the most complex and least understood mammalian organ. At the point of maturity it weighs approximately 1.4 Kg and is composed of 180 billion cells, of which 50 billion are directly engaged in information processing. In addition to the scale, the complexity of this organ is difficult to comprehend. Each of those 50 billion cells is estimated to recieve up to 15,000 physical connections from other cells. Communication between those cells is therefore fundamental for the multicellular organism to function. Such communication is central to the coordination of many cellular functions including cell metabolism, cell growth and differentiation. In mammalian nervous systems, specific responses are illicited by highly specialised messenger molecules. The neuro-physiologist Otto Loewi in 1921 was the first scientist to present evidence to support the concept that biochemical substances are central to communication between nerves, muscles and glands. This evidence ultimately provided the foundation for the later identification of one precise mode of intercellular communication. By stimulating the vagus nerve to a frogs heart, he noted that the rate of the heart declined. If the slowing heart was simultaneously rinsed in a solution which was collected and later poured over another frogs heart, a decline in the rate of heartbeat of the second heart could be induced. A substance, which he named *'Vagusstoff'* (vagus substance) and which was later identified as acetylcholine, had been released from the first heart into solution and had been subsequently responsible for the slowing of the second heart. Conversely, by stimulating the heart to increase its rate of beating, Loewi managed to identify the molecule epinephrine which up-regulates the mammalian heartbeat. He therefore proposed that neurons communicate with each other and with other tissues by the release of biochemical neurotransmiters.

A neurotransmitter is defined as a chemical which is found in the terminal of a neuron, is released from the terminal when the neuron fires, and when incubated in the presence of innervated tissue, is capable of inducing the same response in the tissue as the response generated by neural stimulation. Additionally, there must be a

mechanism in the area of the synaptic cleft to inactivate the transmitter either chemically or by way of uptake of the neurotransmitter. There are many subtypes of neuortransmitter. Several amino acids act in this way, including gamma-aminobutyric acid (GABA) to which there are a family of receptors, several antagonists and a number of known agonists such as baclofen (detailed in section 1.7.3). The 'classical' (or canonical) chemical neurotransmitters include acetylcholine which can act as a ligand for both muscarinic and nicotinic receptors. The nicotinic receptors are known to form transmembrane ion channels, whereas the muscarinic receptors act through G-proteins (GTP-binding proteins) to activate a variety of second messenger systems (Nathanson *et al*, 1987). Several biogenic amines, including dopamine and serotonin, have been studied and found to act as neurotransmitters. One known agonist to serotonin receptors is the psycho-active narcotic LSD. An antagonist to the receptor for the amino acid transmitter molecule glycine is strychnine.

A number of peptides are known to be neuroactive and are collectively known as the neuropeptide family of transmitters. The neuropeptides are diverse and include neurokinin-1 which is the focus of the research presented in this text. Tachykinins in general and NK-1 specifically are discussed in depth below.

1.2 NEUROPEPTIDES

1.2.1 The History of Neuropeptides

Neuropeptides are biologically active peptides which are produced by neurons and which can illicit responses proximally or distally within the mammalian nervous system and other tissues. Of the many cell types that comprise the nervous system, peptidergic neurons are the nerve cells which are specialized to produce and secrete neuropeptides. The list of known neuropeptides is ever increasing and currently stands at between 100 and 1000 taking

account of processed variants of the known peptides (Davies *et al*, 1994). Neuropeptides which mediate long-range communication between neurons are traditionally referred to as 'hormones'. Those which mediate short-range communication and occur in membrane-bound vesicles are termed 'transmitters' (Burbach *et al*, 1992). A broad range of neuropeptides exist in the brain partly as a consequence of molecular processing such as alternative splicing of primary gene transcripts and differential precursor processing. The known neuropeptides are sub-divided into groupings in line with their common sites of either release or action. For example, there are the brain-gut peptides which can induce specific responses in the gastro-intestinal tract, amongst other tissues, and include glucagon, substance P, neurotensin, leucine enkephalin and vasoactive intestinal peptide (VIP). Also there are the hypothalamic-releasing hormones such as thyrotropin-releasing hormone (TRH) and somatostatin. Another group comprising the pituitary peptides such as β-endorphin and adrenocorticotropin, are secreted by the pituitary gland. Several other neuropeptides have been identified but as yet do not lie within any of the existing sub-categories. The other neuropeptides include angiotensin II, bradykinin, vasopressin and oxytocin (Kandel *et al*, 1981). Neuropeptide Y is one neuropeptide which is researched intensively. This 36 amino acid neuropeptide is found in high levels in the central and peripheral nervous systems. There is growing evidence indicating that this neuropeptide acts as a neurotransmitter and neuromodulator in several regions of the brain (Morris, 1992).

1.2.2 The Roles of Neuropeptides

1.2.2.1 Neuropeptides and Nociception

It has been proposed that NK-1 has a role in nociception (transduction of pain). An understanding of the molecular basis of the activation of peripheral nociceptive sensory neurons is necessary to allow pharmacological intervention in the treatment of pain. One research approach to this is the generation of hybrid cell lines. Wood *et al* (1990) generated hybrid cell lines by fusing

neonatal rat dorsal root ganglia neurons with mouse neuroblastoma N18Tg2 cells. The resulting hybrids exhibited sensory neuron-like properties which are not displayed by the parental neuroblastoma cells. One such property is the synthesis of NK-1. These novel cell lines provide one means to study the mechanisms of nociceptor activation and the regulation of expression of sensory neuron-specific markers such as the neuropeptide NK-1. This model may subsequently permit a clearer understanding of NK-1 synthesis and its role in nociception.

1.2.2.2 Neuropeptides and Behaviour

Neuropeptides are known to play a vital role in both the control and expression of behaviour in many organisms (Tublitz *et al*, 1991). For example, in insects, moulting is initiated by ecloisin hormone (Reynolds *et al*, 1983), and in lobsters defensive postures are triggered by proctolin (Siwicki *et al*, 1987). In rats, drinking is activated by angiotensin II (Fitzsimmons *et al*, 1978), maternal behaviour is enhanced by oxytocin (Fahrbach *et al*, 1986), and feeding is inhibited by cholecystokinin (Forman *et al*, 1977). Neuropeptides are believed to influence a range of other more complex, long-term behavioural changes such as learning, arousal, memory, emotional state and schizophrenia (Gilles *et al*, 1979; McKelvy *et al*, 1986). The lack of knowledge regarding the exact nature of those influences has prompted an upsurge of research in the field of neuropeptides. The molecular approach to neuropeptide research (section 1.7.4) should help to identify those influences.

1.2.2.3 Modulation of the Immune System

The regulation of the response of the immune system has traditionally been thought to be controlled by interactions between immunocompetent cells. Such cells were believed to express surface determinants, secrete regulatory molecules and be capable of recognising the signals emanating from other cells of the immune system. An alternative system has subsequently been observed. The nervous system, which had previously been considered to have no involvement in immunocompetence, may have a direct influence on

immunological responses (Payan *et al*, 1984). An early example of this was presented by Webber *et al* (1970) who demonstrated that stimulation of the post-ganglionic sympathetic nerves attached to bone marrow elevated the level of leucocyte production. Neuropeptides were implicated in the regulation of the local immune response as determined by tissue inflammation. Neuropeptides were shown to be transported to the distal nerve endings of unmyelinated dorsal root C-fibres and released upon stimulation of the axons. As an example of the involvement of neuropeptides as a modulator of the immune system, the neuropeptide NK-1, when incubated *in vitro* with human T-lymphocytes, stimulates DNA and protein synthesis as determined by ^{3}H-thymidine and ^{3}H-leucine incorporation, respectively (Payan, 1983). Furthermore, a particular subset of blood T-lymphocytes have been shown to carry specific receptors for NK-1 (Payan *et al*, 1983). This was determined using a fluorescent conjugate of NK-1 in addition to radiolabelled NK-1. Somatostatin is another neuropeptide associated with modulation of the immune system. Both NK-1 and somatostatin elicit neurocytotoxic histamine release from rat peritoneal mast cells *in vitro* which is as rapid and complete as that evoked by conventional, IgE-dependent stimulation (Foreman *et al*, 1983). Both of these neuropeptides have also been associated with 'neurogenic inflammation'. This is elicited by substances which are released from the peripheral terminals of sensory nerve fibres (Gamse *et al*, 1980). The electrical stimulation of these peripheral primary afferent neurons activates their neuropeptide-rich terminals. Their release, in turn, results in a local increase in vasodilation and vascular permeability to large molecules and leukocytes. Taking account of the observations presented above, Payan *et al* (1984) proposed a model linking the sensory neuropeptides and the effector pathway of the immune system. He hypothesised that a noxious stimuli of sufficient intensity could cause retrograde impulses in neuropeptide-containing C-fibres and this would cause the release of NK-1 and somatostatin from peripheral terminals as an axon reflex. NK-1 would result in vasodilation and the increased permeability of the local microcirculation as well as activation of mast cells to release histamine and leukotactic peptides which trigger established response cascades within the immune system.

Recent experimental evidence (Smith *et al*, 1994) has confirmed the inter-relationships between the central nervous, the neuroendocrine and the immune system, where there is extensive overlapping usage of neurotransmitters, hormones and receptors by these systems. Chambers *et al* (1993) has reviewed at length the role of both neuropeptides and other neurotransmitters in the modulation of immune function. Chambers summarises the published findings with respect to the several neuropeptides which have been shown to act on immune function by way of inhibiting or enhancing B-cell and/or T-cell function; where NK-1 has been shown to enhance T-cell function. The occurrence of interleukin receptors on neural cells demonstrates the bidirectionality of the nervous system-immune system relationship. The evidence suggests that the immune and nervous systems can operate in conjunction in the best interest of the host. However, the exact nature of this relationship has yet to be defined.

1.2.2.4 Neuropeptides and Development

There are a large number of published findings which support the involvement of neuropeptides in mammalian development. Neurokinin-1 is an example of such a neuropeptide and is discussed in depth in section 1.4. As an example, McGregor *et al* (1995) reported the transient expression of the NK-1 gene in the developing rat endocrine pancreas and noted that this can be regulated in the RINm5F cell line model system. Expresssion of NK-1 precursor therefore varies during development.

1.2.3 Activation and Inactivation of Neuropeptides

The production of biologically active peptides and proteins in cells and tissues requires not only biosynthesis but also post-translational modification. These modifications include proteolysis (the hydrolysis of peptide bonds), glycosylation, phosphorylation and methylation. Proteolysis is the most general of all covalent modifications occuring *in vivo* (Wold *et al*, 1981). 'Limited proteolysis' is the cleavage of selected peptide bonds, as opposed to

the non-selective degradation of a whole peptide (Linderstrom-Lang *et al*, 1949). Limited proteolysis has particular relevence to neuropeptides as these peptides are often activated by such cleavage. These proteolytic events are programmed so as to occur in specific cellular compartments where the appropriate enzymes are both present and active. The rough endoplasmic reticulum (RER) is the site of synthesis for the full neuropeptide gene product. This is the largest form of the peptide and is referred to as the pre-pro-protein (For example, pre-pro-tachykinin). The "pre" refers to a 15-30 amino acid 'signal sequence' which is rich in hydrophobic residues and is located at the N-terminus of the newly synthesized peptide. As the peptide emerges from the RER the 'pre' sequence is cleaved leaving a pro-protein. The pro-protein is often packaged into secretory granules in the golgi aparatus of the cell (Blobel and Dobberstein, 1975). The first proteolytic cleavage of the pro-protein can occur either in the golgi aparatus or when the pro-protein has been packaged into secretory granules. Biosynthesis and axonal transport studies of the neuropeptide vasopressin in the rat hypothalamo-neurophysial system have indicated that pro-vasopressin is processed in the secretory granules during axonal transport (Grainer *et al*, 1977). It is likely that NK-1 is processed in the same way. Modern molecular cloning technology has permitted the identification and determination of the primary structures and amino acid sequences of neuropeptide pro-proteins. This, in turn, has allowed the identification of amino acid motifs which are common to groups of pro-proteins and which are believed to act as the recognition sequences which are crucial to the proteolytic processing steps (Nakanishi *et al*, 1979). The majority of known pro-proteins have pairs of basic amino acids flanking the cleavage site, for example, the lysine and arginine amino acid pairing for pro-vasopressin (Land *et al*, 1982). The common amino acid sequences of the tachykinin family of neuropeptides are detailed in Section 1.5. Before the intact pro-protein is cleaved, other post-translational modifications such as glycosylation and the formation of disulphide bridges often takes place. Conversely, prior to cleavage, processes such as C-terminal amidation and N-terminal acetylation are believed to occur (Czicki *et al*, 1977, Phelps *et al*, 1980, England *et al*, 1980). Several methods for the identification of neuropeptide

processing products, including radioimmunoassay, have been presented by Harmar *et al* (1986).

Some neuropeptide pro-proteins can generate more than one biologically active peptide. Pro-opiocortin can be converted into β-endorphin and three other active peptides (Eipper *et al*, 1980). Similarly, pro-enkephalin contains six 'Met-enkephalin' sequences and one 'Leu-enkephalin' (Gubler *et al*, 1982). The most pertinent example of this 'common precursor' or 'multivalent' production of several active peptides from one original pre-pro-protein is NK-1. This, and several other neuropeptides including substance K, are derived from the single PPT-A gene (see Section 1.3.4). Intermediates in proteolytic processing of tachykinins precursors have been detected by Kream *et al* (1985) and the possibility cannot be excluded that one or more of these intermediates is an active peptide in its own right. Other mechanisms can also account for the generation of multiple peptides in this way such as control of transcriptional processing. Those mechanisms are discussed below.

Neuropeptide pro-proteins typically undergo two types of proteolytic activity in order to generate an active peptide. Firstly, an enzyme is required to recognise the paired amino acid basic residues (as detailed above), and secondly, a carboxypeptidase-B-like enzyme is required to remove the basic residues from the C-terminus of the peptide. Further post-translational modification can occur after cleavage, commonly C or N-terminal modifications such as the sulphation of tyrosine in cholecystokinin (CCK) (Anastasi *et al*, 1968).

An early and pertinent example of the cleavage of neuropeptides by tissue-specific enzymes is presented by Blumberg *et al* (1980). This publication documents the cleavage of the proline-glycine bond in bovine NK-1 (11 amino acids in total: *H-Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH2*) into the N-terminal tetrapeptide *H-Arg-Pro-Lys-Pro-OH* and the heptapeptide *H-Gly-Gln-Phe-Phe-Gly-Leu-Met-NH2*. The heptapeptide has been shown to be as active as uncleaved NK-1 itself in a number of bioassays. The tetrapeptide has been found to stimulate the formation of cAMP and neurite

extension in neuroblastoma cells as well as to enhance the phagocytotic activity of macrophage. The intrinsically distinct activities of these cleavage products suggested that NK-1 has at least a dual role and at that early stage hinted at the multifunctionality of NK-1 which was to be clarified by further research.

Peptide neurotransmitters are believed to be inactivated by enzymes in the synaptic cleft, e.g., hydrolysis by a peptidase. Alternative proposed methods of peptide transmitter removal include the uptake of the peptide by neighbouring cells and the internalisation of the whole ligand-receptor complex.

1.3 THE TACHYKININ NEUROPEPTIDE FAMILY

1.3.1 Discovery and History

In 1931, two pioneering pharmacologists, Gaddum and von Euler, discovered a hypotensive and spasmogenic activity in acid ethanol extracts of equine intestine which they found to be different from the neurotransmitter acetyl choline. It was during early experiments to map the distribution of acetyl choline that this new activity was first noted. Not only was it found to be capable of causing the contraction of the rabbit jejunum in the absence of atropine, but this research indicated that it was proteinaceous in nature (von Euler and Gaddum, 1931). This new activity was named *substance P* (P for preparation). This neuropeptide was subsequently re-named neurokinin-1 (NK-1, see Section 1.3.2). In 1949, Erspamer, whose career would later increasingly focus on neuropeptide research, identified an activity with hypotensive and spasmogenic properties similar to those previously associated with NK-1. He noted this while studying the posterior salivary glands of the mediterranean octopus *Elidone moschata*. This activity, christened *eledoisin*, was identified as a polypeptide. Significantly, eledoisin was found in octopus salivary gland at a higher concentration than

that of NK-1 in mammalian brain tissue. 1450 kg of live octopus was required as starting material in order to purify enough of the peptide to permit sequence analysis. In 1962, eledoisin was the first tachykinin peptide to be sequenced (Erspamer *et al*, 1962). Fellow amphibian peptides *physalaemin* and *phyllomedusin* were subsequently identified, purified and sequenced by the same group (Erspamer *et al*, 1964, Anastasi *et al*, 1970). These active peptides demonstrated a characteristic fast onset of action on tissues of the gut (compared to the slower acting *bradykinins*) and were consequently named the *tachykinins* (TK). There is a high degree of functional diversity within the tachykinin peptide family. Regulation of both peptide synthesis and peptide-receptor interaction has contributed to this functional diversity (Nakanishi *et al*, 1987). Study of the peptides has also highlighted a common carboxy-terminal amino acid sequence; **Phe-X-Gly-Leu-Met-NH_2.** This common sequence is characteristic of the tachykinin family of neuropeptides where X is either a hydrophobic or aromatic amino acid residue (Erspamer *et al*, 1981). In the tachykinin neuropeptide neurokinin-1, the residue X is phenylalanine, whereas in substance K, residue X is valine (Kimura *et al*, 1983; Maggio *et al*, 1983; Nawa *et al*, 1983).

The biological properties of the tachykinins were studied using synthetic peptide analogues (Section 1.7.3.2). The conserved amino acid sequence Phe-X-Gly-Leu-Met-NH_2 was found to be critical. Prior to its amino acid sequence being conclusively determined (by Chang *et al* in 1971), Erspamer predicted that the sequence of the mammalian TK neurokinin-1 would include this conserved amino acid motif which was common to the non-mammalian peptides (Erspamer *et al*, 1971). The subsequent discovery of other mammalian and non-mammalian TKs has indicated that members of this peptide family probably exist in all phyla from coelenterates to vertebrates.

Figure 1.1 : Amino Acid Sequences of the Tachykinin Peptides.

Mammalian Tachykinin Peptide Sequences

Substance P (neurokinin-1)	Arg-Pro-Lys-Pro-Gln-PHE-Phe-GLY-LEU-MET-NH 2
Substance K (neurokinin-A)	His-Lys-Thr-Asp-Ser-PHE-Val-GLY-LEU-MET-NH 2
Neuromedin K (neurokinin-B)	Asp-Met-His-Asp-Phe-PHE-Val-GLY-LEU-MET-NH 2
Neuropeptide-K	Asp-Ala-Asp-Ser-SerIle-Glu-Lys-Gln-Val- -Ala-Leu-Leu-Lys-Ala-Leu-Tyr-Gly-His- -Gly-Gln-Ile-Ser-His-Lys-Arg-His-Lys- -Thr-Asp-Ser-PHE-Val-GLY-LEU-MET-NH 2
Neuropeptide- γ	Asp-Ala-Gly-His-Gly-Gln-Ile-Ser-His-Lys-Arg-His- -Lys-Thr-Asp-Ser-PHE-Val-GLY-LEU-MET-NH 2
Neurokinin-A(3-10)	Thr-Asp-Ser-PHE-Val-GLY-LEU-MET-NH_2

Non-Mammalian Tachykinin Peptide Sequences

Scyliorhinin I	Ala-Lys-Phe-Asp-Lys-PHE-Tyr-GLY-LEU-MET-NH 2
Physalaemin	pGlu-Ala-Asp-Pro-Asn-PHE-Tyr-GLY-LEU-MET-NH_2
Lys-Physalaemin	pGlu-Ala-Asp-Pro-Lys-PHE-Tyr-GLY-LEU-MET-NH_2
Uperolein	pGlu-Pro-Asp-Pro-Asn-Lys-PHE-Tyr-GLY-LEU-MET-NH 2
Hylambatin	Asp-Pro-Pro-Asp-Pro-Asp-Arg- -PHE-Tyr-GLY-Met-MET-NH_2
Entero-Hylambatin	Asp-Pro-Pro-Asp-Pro-Asp-Arg-PHE-Tyr-GLY-Met-MET-NH_2
Phyllomedusin	pGlu-Asn-Pro-Asn-Arg-PHE-Ile-GLY-LEU-MET-NH 2
Scyliorhinin II	Ser-Pro-Ser-Asn-Ser-Lys-Cys-Pro-Asp-Gly-Pro-Asp- Cys=PHE-Val-GLY-LEU-MET-NH 2
Glu-Pro-Kassinin	Asp-Glu-Pro-Lys-Pro-Asp-Gln-PHE-Val-GLY-LEU-MET-NH 2
Kassinin	Asp-Val-Pro-Lys-Pro-Asp-Gln-PHE-Val-GLY-LEU-MET-NH 2
Entero-Kassinin	H-Asp-Glu-Pro-Asn-Pro-Asp-Gln-PHE-Ile-GLY-LEU-MET-NH_2
Eledoisin	pGlu-Pro-Ser-Lys-Asp-Ala-PHE-Ile-GLY-LEU-MET-NH 2
Canonical tachykinin C-terminal sequence	- - **PHE-X-GLY-LEU-MET-NH** 2

Figure 1.1: Source references: For substance P (NK-1), see Chang *et al*, (1971), for substance K and neuromedin K, see Kimura *et al*, (1983), for scylorhinin, see Conlon *et al*, (1986), for phylaemin, see Erspamer *et al*, (1964), for uperolein, see Anastasia *et al*, (1975), for hylambatin, see Yasuhara *et al*, (1981), for phyllomedusin, see Anastasi *et al*, (1970), and for eledoisin, see Erspamer *et al*, (1962), for neuropeptide-K, neuropeptide-γ and neurokinin-A[3-10], see Helke *et al*, (1990).

1.3.2 Nomenclature

There have been several attempts to reach agreement on the use of a single name for each of the tachykinin peptides. Several identical peptides have been cited in publications under more than one name. For example neuromedin K is also referred to as neurokinin B. The International Tachykinin Nomenclature Committee, chaired by R. Michael Snider of Pfizer Inc., presented recommendations regarding tachykinin nomenclature (Henry *et al*, 1989). Unfortunately, the recommendations have not been adopted unanimously. The gene at the heart of the research presented in this text, PPT-A, encodes several neuropeptides including neurokinin-1 (NK-1) which was originally referred to as substance P. The preprotachykinin-A (PPT-A) gene has been referred to as the 'SP' or 'substance P gene' in the past.

The tachykinins exhibit different pharmacological properties and on this basis this, were divided into two sub-groups (Lee *et al*, 1982). One group, composed of substance P (NK-1) and physalaemin, was named the SP-P sub-group, whereas the other, comprising kassinin and eledoisin, was designated the SP-E group. Physalaemin and kassinin were used as the key discriminatory compounds, respectively. This crude system of distinguishing the tachykinins served a purpose at the time but has subsequently been found to be over-simplistic. It has been superceded and is no longer used. Substance P is now referred to as neurokinin-1 (NK-1).

1.3.3 Biosynthesis

As discussed above, neuropeptides are synthesized as precursor proteins in the cell body of neurons. They are transported by an anterograde action to the nerve terminals where they are enzymatically cleaved into their active form, stored in specific vesicles (different vesicles from when they are stored as co-transmitters) and released when the neuron is stimulated. This contrasts with the classical neurotransmitters which are synthesized in the nerve terminals from where they are released

directly. The mammalian tachykinins, being neurokinin-1, substance K and neuromedin K are encoded by two genes, namely, the preprotachykinin-A (PPT-A) gene and preprotachykinin-B (PPT-B) gene.

1.3.4 The Preprotachykinin-A (PPT-A) Gene

The preprotachykinin-A (or PPT-A) gene, encodes several mammalian tachykinin neuropeptides. Neurokinin-1, neurokinin A, neuropeptide K, and neuropeptide γ are each encoded from this gene by way of tissue-specific, alternative post-transcriptional (alternative RNA splicing) and post-translational pathways (Carter *et al*, 1990). Each of the PPT-A gene-encoded peptides are biologically active (Regoli *et al*, 1989). Of these, NK-1 and substance K are the most extensively characterised. Alternative splicing of the primary PPT-A RNA transcripts generates three mRNAs which, in turn, generate several distinct precursor proteins. McDonald *et al* (1989) presented results which demonstrated that multiple tachykinins are produced and secreted as a consequence of post-translational processing of the three preprotachykinin-A precursor proteins, α-, β-, and γ-preprotachykinin. The transcripts encoded by the NK-1 gene are named α–, β– and γ–preprotachykinin mRNA. The α transcript encodes neurokinin-1 (SP) only, whereas the β-transcript encodes neurokinin-1, neurokinin-A (formerly SK), neuropeptide K and neurokinin A(3-10), which is a post-translational variant of neurokinin A. The γ–transcript encodes neurokinin-1, neurokinin-A, neuropeptide-γ, and neurokinin A(3-10) (Helke *et al*, 1990, Kawaguchi *et al*, 1986). In rat and human PPT-A expressing tissue, the β- and γ–transcripts together constitute 99% of the total transcripts derived from the gene. The α–transcript is therefore relatively rare. Tissue-specific generation of the bovine α and β-transcripts was demonstrated originally by Nawa *et al* (1984). The neurokinin-1 peptide is encoded by exon 3 of the PPT-A gene whereas neurokinin-A is encoded by exon 6 (Nawa *et al*, 1984). Differential post-translational processing of the neurokinin-A portions of the β- and γ–PPT-A mRNAs results in the generation of different forms of neurokinin-A; neuropeptide K, which is derived

from exons 3 to 6 of the β-transcript, and neuropeptide-γ which is derived from exons 3, 4 and 5 of the γ-PPT-A mRNA. Aqueous hybridisation and S1 nuclease protection assay techniques were used to examine both the RNA transcripts and the neuropeptides which are encoded by the PPT-A gene (Krause *et al*, 1989).

Figure 1.2: Transcription and Transcript Splicing of the Rat PPT-A Gene

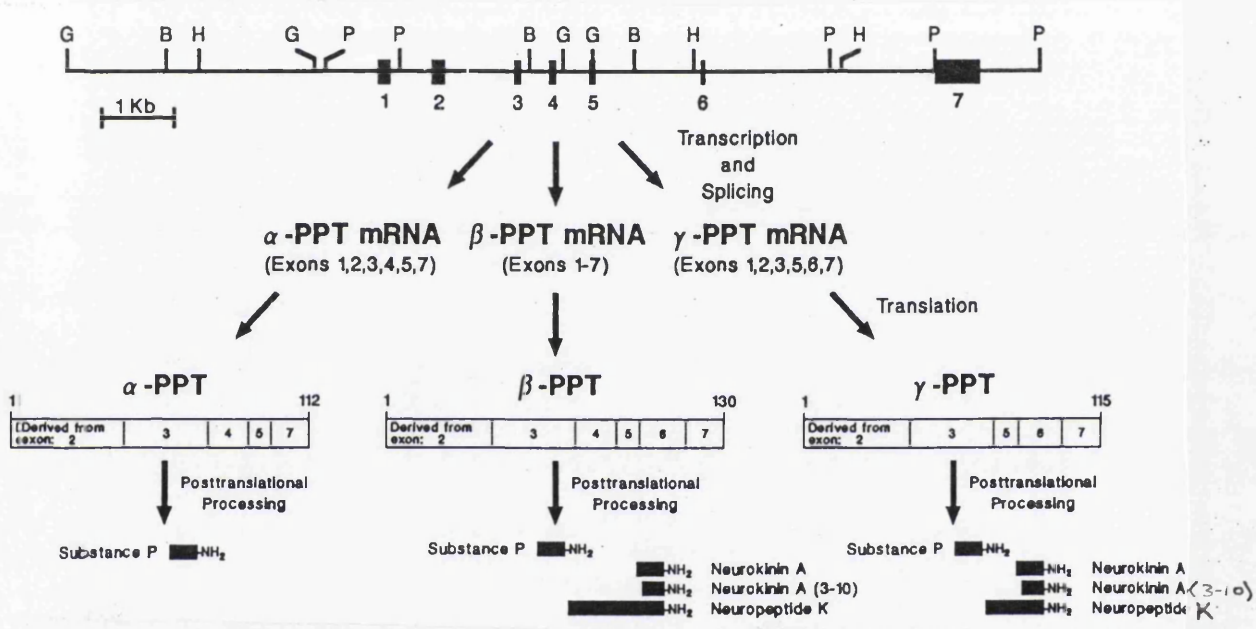


Figure 1.2: A schematic representation of the transcription and transcript splicing of the rat PPT-A gene primary transcript. Furthermore, the translation and post-translational processing of the α-, β- and γ-PPT precursor proteins is also represented (Helke *et al*, 1990). The rat NK-1 gene is depicted at the top and the exons are numbered 1 through to 7. Transcription of the gene and splicing of the primary transcript generates the α-, β- and γ-PPT mRNAs which are translated with concomitant signal peptide cleavage. The number above each PPT precursor represents the number of amino acids in each peptide. The abbreviations used are: H, *HinD* III; G, *Bgl* II; P, *Pst* I; B, *Bam*H I; PPT, preprotachykinin. Schematic taken from Helke *et al*, 1990.

As described in the neuropeptide section above, neuropeptides are often activated by cleavage of a precursor protein. The primary structure of two forms of bovine brain neurokinin-1 precursor protein were identified by Nawa *et al* (1983). The bovine PPT-A gene is 8.4 kb in length and comprises 7 exons and 6 introns (Nakanishi *et al*, 1986).

cDNAs encoding the preprotachykinin-A and preprotachykinin-B protein precursors respectively, were isolated from a mouse brain cDNA library by Kako *et al* (1993). The cDNAs isolated represented the β- and γ–PPT-A transcripts. The mouse neurokinin-1 precursor protein was shown to consist of 130 amino acids which were highly conserved with their rat, human and bovine equivalents, with 99%, 93% and 92% amino acid identity respectively. The neurokinin-1 and neurokinin-A sequences were found to be identical between these species. The mouse preprotachykinin-B precursor protein was shown to comprise 116 amino acids and have 93% and 78% amino acid sequence identity with equivalent rat and bovine sequences respectively. Similarly, the sequence of neurokinin B was also identical for those species.

The nucleotide sequence of the cDNA which encodes the human neurokinin-1 precursor protein, β-preprotachykinin, was determined by Harmar *et al* (1986). The β-PPT polypeptide is 129 amino acids in length and encompasses regions which encode neurokinin-1 and neurokinin A. The same research group was later to identify a fourth NK-1 gene mRNA variant. In addition to a mRNA containing all 7 exons of the gene (β-PPT) and two other alternatively spliced forms lacking exon 4 (γ-PPT) or exon 6 (α-PPT), a fourth variant was isolated which lacked both exons 4 and 6. A PCR based approach was employed to amplify cDNAs which were then sequenced. This novel transcript was named δ-PPT and is likely to encode a novel precursor polypeptide (Harmar *et al*, 1990).

1.3.5 Regulation of the PPT-A Gene

The rat PPT-A gene has been shown to be regulated by a number of stimuli including growth factors (Kessler *et al*, 1981), conditioned

media (Nawa *et al*, 1990), cocaine (Hurd *et al*, 1992), inflammation (Noguchi *et al*, 1992), and factors present in innervated target tissue (Barakat-Walter *et al*, 1991). Expression of the PPT-A gene and the level of tachykinin biosynthesis appears to be under homeostatic control. This is clearly the case in both the peripheral and central nervous systems where expression can be induced or repressed by modulation as a result of extracellular signals (Walker *et al*, 1991). Tachykinin expression in the anterior pituitary gland has been shown to be under hormonal control (Jonassen *et al*, 1987).

Mulderry *et al* (1993) used deletion analysis to study the rat PPT-A promoter and the regulation of expression. Adult rat sensory neurons in culture were transfected by microinjection of plasmids containing genomic DNA sequences linked to a *lacZ* (B-galactosidase) reporter gene. Deletion analysis results indicated that a 5'-flanking sequence, 865 bp upstream of the rat PPT-A gene transcription initiation site was responsible for directing high levels of expression of the reporter gene.

Mulderry *et al* (1993) also found that the majority of the sequences which are important to the regulation of this gene in rat dorsal root ganglia (DRG) neurons are located between 47 and 865 nucleotides upsteam of the start of the gene. DNAase 1 footprinting analysis studies of the rat PPT-A gene promoter were conducted by Mendelson *et al* (1995a). This technique facilitates the identfication of regulatory sequences by digesting DNA with DNAase I to leave only the DNA sequences to which a regulatory peptide is directly bound. Such sequences are protected from the nuclease and therefore remain intact. Two footprints were generated which permitted the identification of E-box consensus sequences (CANNTG) which are recognised by dimeric protein complexes. Such complexes are composed of members of the basic helix-loop-helix (bHLH) and basic helix-loop-helix leucine zipper (bHLH-zip) families of proteins (Murre *et al*, 1989). One was found to span nucleotides -67 to -47 while the other spanned -177 to -155. The function of these complexes have not been defined. The functional E-box motif in the proximal rat PPT-A promoter spans nucleotides -67 to -47. This was later mutated by insertion of a 10 bp DNA sequence. Using

cultured PC12 cells, this element was shown to be likely to function as an important *cis*-regulatory domain in this promoter (Paterson *et al*, 1995). Purine-rich motifs, which may serve as binding sites for AP-1, Sp1 or other zinc finger-containing proteins, were also identified upstream of the rat PPT-A gene. Several other footprints were identified and could potentially be regulatory sequences for the rat PPT-A gene. The elements identified are likely to be involved in both tissue-specific and stimulus-dependent regulation of the gene (Mendelson *et al*, 1995a; Paterson *et al*, 1995).

The rat PPT-A gene was isolated from a genomic library by Carter *et al* (1990). One major and two minor transcription initiation sites were identified upstream of the rat PPT-A gene. Comparison of the rat and the bovine PPT-A gene sequences demonstrated highly conserved regions throughout the entire coding region and within the 5' flanking region.

The cell lines RINm5F and RINr1046-38 currently offer one means of investigating which factors and mechanisms are involved in the regulation of the NK-1 gene *in vitro*. Using RINm5F cells, McGregor *et al* (1995) reported the transient expression of the PPT-A gene in the developing rat endocrine pancreas. This approach could be extended to the mouse PPT-A gene.

1.3.6 Preprotachykinin-B (PPT-B) Gene

Neurokinin-B, which was previously known as neuromedin K, is encoded by the fifth exon of the preprotachykinin-B (PPT-B) gene and has been isolated from porcine spinal cord. In order to do this, a bioassay which determined the tachykinin-like effect of a sample on guinea-pig ileum smooth muscle contraction was employed (Kangawa *et al*, 1984). Later, a cDNA encoding the 116 amino acid precursor of neurokinin-B was isolated from a rat cerebral cortex cDNA library (Bonner *et al*, 1987).

The diversity of the mammalian tachykinin system may have evolved as a result of mechanisms such as gene duplication and differential

expression of duplicated genes. An insight into this was gained as a consequence of cloning the bovine preprotachykinin-B (PPT-B) gene. Molecular cloning and sequence analysis techniques were utilised to identify cDNA and genomic DNA clones of the mammalian neuropeptide, neuromedin K (Kotani *et al*, 1986). Neuromedin K is a 126 amino acid peptide with a sequence which includes a signal peptide sequence and is encoded by the PPT-B gene. This gene encodes two mRNAs which differ in the 5' untranslated region. The gene organisation of the PPT-B gene was found to closely resemble, in terms of the mRNA transcripts, that of the PPT-A gene, which encodes the precursor for substance P, substance K and other neuropetides (see above). The genetic resemblence suggests that the two preprotachykinin genes may be derived from a common ancestor gene. Using *in situ* hybridisation techniques, Nakanishis' research team have demonstrated that the expression patterns of the mRNAs of the PPT-A and PPT-B genes are different (Kotani *et al*, 1986).

1.3.7 Regulation of the Preprotachykinin Genes

Synthesis of the precursor proteins from the preprotachykinin genes can be regulated in several ways including transcriptional activation, splicing of primary transcripts, translational efficiency and post-translational processing. Transcription can be activated by the binding of *trans*-acting factors to *cis* regulatory sequences in the promoter or enhancer regions of the gene, and therefore allowing RNA polymerase II to transcribe the gene. Alternative splicing of primary transcripts can vary the synthesis of pro-peptides, often in a development-dependent or tissue-specific manner. For example, bovine α-PPT-A mRNAs are found in high relative concentrations in the bovine brain tissue whereas the β-PPT-A mRNAs are abundant in bovine intestinal and thyroid tissues (Nawa *et al*, 1983). Similarly, species-specific splicing has been observed in rat and human tissues and the relative levels of the α, β and γ-PPT-A mRNA transcripts determined. In rat, the γ transcript is the most abundant, whereas in human tissues, the β-transcript is prevalent (Helke *et al*, 1990). For example, in the rat striatum, PPT-A mRNA from post-natal day 2 to day 5, levels do marginally fluctuate, (Haverstick *et al*,1989).

1.3.8 Roles of Tachykinins

Tachykinins are distributed throughout the peripheral and central nervous systems. Within the CNS, it has been suggested that tachykinins are involved in dopaminergic activation and dopamine release in addition to sleep control, Alzheimer's disease and neuroprotection. With respect to the PNS, tachykinins are thought to be involved in skin hypersensetivity and modulation of the immune response. Mammalian tachykinins have been associated with sensory and motor function as well as immunologic, inflammatory, cardiovascular, respiratory and gastrointestinal functions. Additionally, they may be trophic and mitogenic factors.

The extent of tachykinin research is vast and several hundred related papers are published every year. Most of that research is inevitably focussed on substance P (NK-1) which was the first neuropeptide to be identified, and is arguably much the best studied (Maggio *et al*, 1988). The tachykinin neuropeptides are widely distributed and are active in the central and peripheral nervous systems and evoke numerous responses in numerous tissues. They are known to evoke responses in neurons and act as secretagogues (promoting secretion of other substances). They are also potent vasocontractors and can induce the contraction of smooth muscle either directly (or indirectly via the release of classical neurotransmitters). Of particular interest to the neurobiologist is the apparent involvement of tachykinins in the nigro-striatal pathway and primary sensory afferent transmission. The role of tachykinins as mediators between the nervous and immune systems has also generated much research interest (Section 1.2.2.3).

1.3.8.1 Asthma

In 1992, a study by Ichinose using a tachykinin receptor antagonist called FK-224, demonstrated that such antagonists may have a use in the treatment of asthma. This was demonstrated by the reversal of broncho-constriction which had been induced by the inhalation of bradykinin in asthmatic patients (Ichinose, 1992).

1.3.8.2 Inflammation and Arthritis

Rheumatoid arthritis can be initiated by an auto-immune reaction (Guard *et al*, 1991). The sensory and sympathetic nervous system is important in the maintainance and progression of rheumatoid arthiritis. Support for this is provided by the observation that sensory nociceptive fibres (which activate sympathetic fibres via spinal connections) demonstrate neurokinin-1 immunoreactivity. Tachykinin antagonists may provide useful local treatment for rheumatoid arthritis. Tachykinin peptides may have some involvement with neurogenic inflammation (Guard *et al*, 1991) and may therefore be involved in the processes of arthritis.

1.3.8.3 Development

Previous studies have demonstrated that immature cortical neurons can express neuroactive substances which are not found in the adult cortex; neurokinin-1 is one such substance (Zhang *et al*, 1994). Immunocytochemistry and *in situ* hybridisation (histochemistry) have been used to investigate the distribution of neurokinin-1 and neuromedin K in various cortical areas of the rat cerebrum during development (Zhang *et al*, 1994). Studies were performed at several prenatal and postnatal stages in order to identify a population of transitory neurons equivalent to those that appear to play a role during the development of the cerebral cortex in mice. The number of neurokinin-1 mRNA-positive neurons and neurokinin-1-immunoreactive cells was found to decrease from postnatal day 10 and 15 onward, respectively. The identification of transient populations of tachykinin-expressing neurons has provided further evidence of the role of tachykinins in mammalian neurobiological development.

The floor plate is a ventral mid-line structure which is crucial to the organisation of the developing vertebrate CNS and may also be involved in regulation of neuronal differentiation by guiding developing axons. Neurokinin-1 is one of the first transmitters to be expressed in the developing spinal cord. The distribution of the neurokinin-1 receptor in the vertebrate spinal cord was examined by *in situ* hybridisation and immunocytochemistry (Heath *et al*, 1995). Neurokinin-1 and other peptide modulators may contribute to the

fine tuning of developmental processes in the embryonic CNS. Elsewhere, (see section 1.4.3) results have indicated that the PPT-A gene for example, which encodes several tachykinins, is actively expressed in the pancreas during development, but not in the adult pancreas (McGregor *et al*, 1995).

1.3.8.4 Neurotransmission

The proposed role of the tachykinins in neurotransmission is supported by the evidence indicating a role for these peptides in nociception (Section 1.2.2.1). Furthermore, the tachykinin neuropeptides are present in nerve terminals packaged in vesicles which are encased by membranes which can fuse with the membranes of other cells. Evidence for the role of tachykinins in the regulation of neurotransmission is presented by Maggio *et al* (1988). The role of tachykinins in the perception of pain presented by Helke *et al* (1990) also supports this.

1.3.8.5 Sleep Regulation

The locus coeruleus region of the mamalian brain plays an integral role in the control of sleep and alertness (Kolb *et al*; 1989). Tachykinins, including neurokinin-1, have been shown to be potent activators of locus coeruleus neurons and it has been proposed that they are likely to be involved in sleep/alertness control. NK-1 also activates other neuronal cell types and therefore the precise involvement of this neuropeptide with sleep has yet to be clarified. Further research is required.

1.3.8.6 Immunoregulation

The CNS is believed to regulate the immune system, but to an unknown extent (see section 1.2.2.3). Tachykinins are thought to be significant mediators of this. Neurokinin-1 and neurokinin-A stimulate T-cell proliferation and tachykinin receptors have been localised on lymphocytes and other cells of the immune system. Tachykinins may also mediate local inflammation as they have been

shown to stimulate histamine secretion from mast cells (Ljundahl *et al*, 1978) and polymorphonuclear leucocytes (Bar-Shavit *et al*, 1980), in addition to mediating local vasodilation in response to injury.

1.4 SUBSTANCE P

1.4.1 The Discovery and History of Substance P (NK-1)

As detailed in the tachykinin section above, while investigating the tissue distribution of acetylcholine, Von Euler and Gaddum (1931) discovered a spasmogenic and hypotensive activity in acid ethanol extracts of equine intestine which they named substance P and which has formed the bedrock of neuropeptide research. Early experiments indicated that this activity was proteinaceous in nature rather than a classical, acetyl choline-like neurotransmitter (Gaddum and Schild, 1934). Of the tachykinin, brain-gut peptides, substance P has the longest scientific history and was the first active compound extracted from neural tissue which was found to be proteinaceous. Substance P has since been re-named neurokinin-1. As detailed above, neurokinin-1 and neurokinin A are the major peptides expressed by the PPT-A gene and they are produced predominantly by neuronal cells of the central and peripheral nervous systems (Krause *et al*, 1989).

For thirty years after the discovery of neurokinin-1, research focused on the pharmacological activities, the chemical properties and the tissue distribution of this tachykinin (Pernow *et al*,1983). However, with the benefit of hindsight, the variable purity of neurokinin-1 isolates used during that period calls into question the accuracy of the results presented and published at that time. Findings made prior to the complete purification and amino acid sequencing of neurokinin-1 (Chang *et al*, 1971) included the assertion that neurokinin-1 was identical in brain and gut, unevenly distributed in the nervous system, and active in the gastrointestinal

tracts of many species. These observations have been subsequently validated. Of the several known peptidergic neurotransmitters/ neuromodulators, neurokinin-1 is one of the best-characterised (Nakanishi *et al*, 1986).

1.4.2 Distribution of Neurokinin-1

1.4.2.1 Central Nervous System

As a consequence of the common carboxy terminal amino acid sequence, antisera raised against one particular tachykinin peptide and used in immunocytochemical studies, may cross-react with one or more of the other tachykinins (Milner *et al*, 1988). This, in addition to the relatively recent discovery of the peptides NPγ and NKA(3-10) expressed from the same PPT-A gene, has cast a doubt over the specificity of early radioimmunoassay and immunocytochemical studies. The distinct regional and cellular distribution of each neuropeptide has yet to be ascertained accurately. However, to date, the highest concentration of neurokinin-1 in any region of the brain, is found in the substantia nigra where it is concentrated in nerve endings of afferent neurons in the striatum, and is subsequently released (Duffy *et al*, 1975). This also supports the role of neurokinin-1 as a neurotransmitter.

Neurokinin-1 can be isolated from the brains of all vertebrates ranging from fish to mammals, including man. The concentration of neurokinin-1 appears to be lower in species with highly differentiated brains; in increasing order, man, cat, rabbit, guinea-pig, rat, duck, pigeon, and chicken (Grabner *et al*, 1959). Generally neurokinin-1 concentration is higher in grey matter than white matter. The highest brain concentrations have been found in the substantia nigra, hypothalamus, globus palladus, caudate nucleus, putamen and the central grey matter (Pernow *et al*, 1983). In the spinal cord, the dorsal half contains more neurokinin-1 than the ventral. In the peripheral nervous system, levels are low relative to the central nervous system yet are highest in the spinal ganglia and sympathetic trunk of the peripheral system (Pernow *et al*, 1953).

Neurokinin-1 has also been localised to the gastrointestinal tract of several species (Euler *et al*, 1977). The distribution of neurokinin-1 has been determined traditionally using radioimmunoassay and immunocytochemical techniques (Section 1.7.3). However distribution has been determined most accurately using the *in situ* hybridisation technique, where, in the rat central nervous system for example, distinct patterns of cells producing PPT-A mRNA have been mapped (Warden *et al*, 1984). Such cells are prominent in laminae I and II of the spinal cord and in the thoracic intermediolateral cell column. Similarly, the sensory neurons of the dorsal root ganglia and the medullary raphe neurons which project to the spinal cord, contain relatively high levels of PPT-A mRNA (Warden *et al*, 1984) (see figure 1.3).

A subpopulation of primary afferent neurons have been shown to be positively immunoreactive for neurokinin-1. In addition to several other neuropeptides including vasoactive intestinal polypeptide (VIP) and calcitonin gene related polypeptide (CGRP), neurokinin-1 has been localised to late differentiating (LD) neural crest cells in culture (Matsumoto *et al*, 1994). Conversely, a population of early differentiating (ED) neural crest cells was found to lack this neurokinin-1 immunoreactivity.

The organisation of the extrapyramidal motor system includes the reciprocal neuronal connections between the corpus striatum (caudate-putamen) and the substantia nigra. Distinct defects in the extrapyramidal system have been shown to result in Huntingdon's and Parkinson's disease. Microdissected brain tissue has been examined and high concentrations of neurokinin-1 have been demonstrated in both the caudate nucleus and the zona reticulata of the substantia nigra (Brownstein *et al*, 1976, Davies *et al*, 1976). The immmunocytochemical experiments of Ljungdahl *et al* (1978) showed that the neurokinin-1-immunoreactive neuronal cell bodies in the caudate-putamen are the source of the neurokinin-1-immunoreactive nerve terminals in the substantia nigra and the globus pallidus. This suggests that neurokinin-1 may have some role to play in Huntington's and/or Parkinson's disease because those regions are known to demonstrate altered morphology in patients

with the clinical symptoms of those diseases. It has also been noted that neurokinin-1-immunoreactive neurons in the substantia nigra appear to regulate the nigral dopaminergic neurons which project into the striatum. Similarly, expression of the PPT-A gene in the striatum is regulated by those nigrostriatal dopaminergic cells (Bannon *et al*, 1987).

Figure 1.3 :Distribution of Tachykinin (TK) Immuno-Reactivity in the Rat.

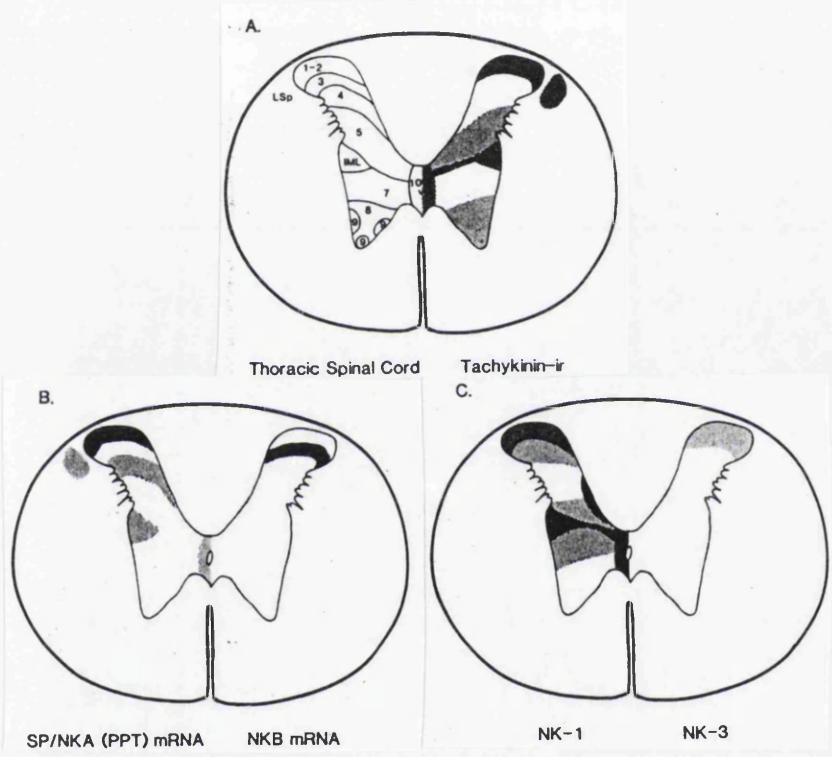


Figure 1.3: Representation of a) laminar divisions of the thoracic spinal cord in rat (left side) and the distribution and relative intensity of tachykinin (TK) immuno-reactivity in the rat. TK immunoreactivity was assessed using a monoclonal antibody to neurokinin-1 which cross-reacts extensively with NK-A (and NK-A related peptides) and NK-B. b) Distribution and relative density of cell bodies containing PPT mRNA (left side) or NK-B precursor mRNA (Warden *et al*, 1988). c) Distribution and relative density of NK-1 binding sites (left side, labelled with [^{125}I]BH-SP) and of NK-3 binding sites (labelled with [^{125}I]BH-eledoisin, Mantyh *et al*, 1989). IML, intermediolateral cell column; LSp, lateral spinal nucleus; NK-A, neurokinin-A; SP, substance P; PPT, preprotachykinin; NK-B, neurokinin-B; (Helke *et al*, 1990).

Figure 1.4 : Distribution of NK-1 in Rat and Human Tissues

	Rat				Human	
	Kanazawa, 1976 ($n = 4$)	Brownstein, 1976 ($n = 20$)	Douglas, 1982 ($n = 6$)	Gale, 1978 ($n = 10$–20)	Emson, 1980 ($n = 3$–13)	Cooper, 19[illegible] ($n = 5$–10)
Telencephalon						
Somatosensory cortex	19 ± 2*					
Visual cortex	23 ± 4					
Olfactory bulb	62 ± 11	20 ± 5				41 ± 8
Olfactory tubercle	300 ± 14					
Hippocampus	37 ± 4	340 ± 45			104 ± 29	
Amygdala	382 ± 86			340 ± 60	25 ± 12	26 ± 3
Nucleus caudatus	247 ± 28			370 ± 80	138 ± 14	113 ± 12
Putamen				330 ± 60	112 ± 29	81 ± 11
Globus pallidus	332 ± 45			1800 ± 330	877 ± 253	518 ± 15
Septum	405 ± 44	116 ± 14			89 ± 23	
Lateral septal nucleus		360 ± 20				
Internal capsule	137 ± 21					
Retina	66 ± 14					
Optic nerve	65 ± 14					0.6 ± 0.[illegible]
Diencephalon						
Hypothalamus (medial)	626 ± 64	206 ± 37				122 ± 22
Hypothalamus (lateral)						135 ± 20
Habenula	377 ± 101					
Mammillary body	207 ± 8					83 ± 14
Pituitary stalk and median eminence						134 ± 21
Posterior pituitary	489 ± 24					
Anterior pituitary	71 ± 2	20 ± 5				
Thalamic nucleus	215 ± 68					35 ± 8
Globus pallidus	332 ± 45			1800 ± 330	877 ± 253	518 ± 151
Lateral geniculate body	82 ± 12	90 ± 13				37 ± 10
Mesencephalon						
Substantia nigra	1725 ± 106		2560 ± 195			
Zona compacta		296 ± 44		4720 ± 480	1264 ± 239	
Zona reticularis		1138 ± 136		4290 ± 490	1535 ± 170	922 ± 156
Zona lateralis		298 ± 39				
Periaqueductal central grey			614 ± 37			130 ± 21
Interpeduncular nucleus	599 ± 63	590 ± 129	994 ± 149		83 ± 25	
Red nucleus			58 ± 12			76 ± 19
Interferior colliculi			94 ± 18			234 ± 19
Pons						
Locus coeruleus			332 ± 50			199 ± 32
Nuclei parabrachiales			546 ± 124			
Motor nucleus V			145 ± 29			
Dorsal raphe nucleus			222 ± 25			
Cerebellum						
Cortex	10 ± 4	2 ± 1	12 ± 1	20 ± 6		1 ± 0.1
Nuclei			29 ± 5			1 ± 0.1
Medulla oblongata						
Nucleus cuneatus			320 ± 32			
Nucleus tractus spinalis V			383 ± 125			
Nucleus tractus solitarii			436 ± 83			
Motor nucleus XII			217 ± 26			
Area postrema			167 ± 36			114 ± 47
Nucleus reticularis gigantocellularis			95 ± 11			
Inferior olive			127 ± 15			8 ± 2
Nucleus raphe magnus			229 ± 23			71 ± 15
Medulla spinalis						
Dorsal horn	1070 ± 160					
Dorsal column	129 ± 29					
Ventral horn	134 ± 33					

Figure 1.4: Regional distribution of neurokinin-1 in the tissues of the rat and human nervous systems (taken from Pernow review, 1983). All values have been converted to pmol/g of wet tissue. Values are given as "mean+/- SEM". Authors of original publications are given in the column headings. Determined by radioimmuno assay.

1.4.2.2 The Peripheral Nervous System

Neurokinin-1 is localised to the peripheral as well as the central nervous system and two neuronal networks in particular. Both the primary sensory neurons and the neurons intrinsic to the gastrointestinal tract contain neurokinin-1-immunoreactive cells. The central role of neurokinin-1 in the contraction of the smooth muscle of the gastrointestinal tract via the peripheral nervous system is well documented (Bartho *et al*, 1985) and is summarised below. Neurokinin-1 is also localised to numerous cell bodies including the trigeminal and the spinal ganglia (Cuello *et al*, 1978).

1.4.2.3 Peripheral tissues

Neurokinin-1-containing fibres have been localised to most peripheral tissues and organs ranging from the human skin (Dalsgaard *et al*, 1983) to the taste buds of the cat (Lunberg *et al*, 1979). Neurokinin-1-containing nerve cells and fibres are also present in all parts of the peripheral visual system, and in nerves which govern motility in the gastrointestinal tract (Pernow, 1983).

1.4.3 Apparent Functional Roles of Neurokinin-1

The excitatory capability of neurokinin-1 has been determined and this may well be the primary role of this peptide in the adult mammal. Further to von Euler and Gaddums' observations (1933), Konishi *et al* (1974) showed that hypothalamic neurokinin-1 caused excitatory pulses in spinal motor neurones of newborn rats. As further evidence of its capabilities as a neurotransmitter, Matsuto *et al* (1984) also demonstrated that neurokinin-1 is capable of depolarising spinal motorneurons. In the central nervous system, neurokinin-1 and neurokinin A function as activating molecules in synaptic transmission (neuromodulators), whereas in the peripheral nervous system they act as neuro-effectors in the paracrine regulation of a variety of neuronal cells (Maggio *et al*, 1988). There are similarities in the distribution of neurokinin-1 and neurotensin in the anterior pituitary gland where these peptides may act as

paracrine regulators (Aronin *et al*, 1986). These peptides appear to function as paracrine neuro-effectors of peripheral tissues such as gastrointestinal smooth muscle (Bartho *et al*, 1985). Neurokinin-1 acts on postganglionic cholinergic neurons in the gut to cause acetylcholine release which in turn causes smooth muscle contraction. Neurokinin-1 depolarizes neurons and in this respect is 200-900 times more potent than L-glutamate (Bartho *et al*, 1985). Several non-neuronal cell types (such as gastrointestinal muscle cells) also contain these peptides (Ericsson *et al*, 1990). The function underlying the occurence of these peptides in several tissues, and the physiological importance of their presence in those tissues, remains to be resolved. As an example, tachykinin-immunoreactive neuronal cells are connected to the islets of Langerhans (Sharkey *et al*, 1984) and the anterior pituitary gland (Ju *et al*, 1989) yet the function of such connections is not clear. In the mature spinal cord neurokinin-1 is released from primary afferent terminals in the dorsal horn and has been implicated in the transmission of nociceptive sensory information (Willis *et al*, 1991). Evidence has also been presented indicating that neurokinin-1 has a role in pain. At the cellular level, neurokinin-1 has been localised to the cell bodies of sensory ganglia and the spinal trigeminal nucleus system (Cuello *et al*, 1978) and therefore provides a distributional support for the involvement of neurokinin-1 in nociception. From a neuro-physiological standpoint, evidence includes the observation that neurokinin-1 applied iontophoretically to single units located in the dorsal horn of the spinal cord selectively excites those cells (Henry *et al*, 1977). There is also a large volume of pharmacological evidence. This includes the use of capsaicin which inhibits the axonal flow of neurokinin-1 when applied locally (Gamse *et al*, 1982).

Neurokinin-1 also acts in the cortex where it has a powerful stimulatory effect on cortical Betz neurons (Phyllis *et al*, 1974). Early studies proposed a relationship between acetyl choline and peptides such as neurokinin-1 where neurokinin-1 was thought to have a presynaptic, acetylcholine-releasing action (Phyllis *et al*, 1974). This has subsequently been corroborated. In the hypothalamus, a calcium-dependent, potassium-induced release of

neurokinin-1 has been demonstrated. This may have some role in the firing of hypothalamic neurons (Ogata *et al*, 1981). Release of neurokinin-1 both *in vivo* and *in vitro* in the striato-nigral pathway has been demonstrated, as has the excitatory and inhibitory effects of neurokinin-1 at this locus. The effects of administering neurokinin-1 to laboratory animals for example are multiple. The complexity of these actions may stem from the relationship between the tachykinin peptides and the classical neurotransmitters. For example, iontophoretic administration of neurokinin-1 to single neurons in the substantia nigra induces a prolonged increase in neuroactivity (Davies *et al*, 1976), and bilateral injections of neurokinin-1 into the rat substantia nigra increases basal locomotive activity (Kelley *et al*, 1978).

Results suggests that there is a significant relationship between the dopaminergic nigro-striatal and neurokinin-1 striato-nigral pathways and that there may be a role for neurokinin-1 as a dopamine modulator. This is indicated by results which show that neurokinin-1 afferents enhance the firing rate of dopamine neurons (Hanson *et al*, 1981). GABA, which is also found at high levels in the substantia nigra has been attributed a transmitter role in inhibitory striato-nigral neurons although a clear anatomical distinction between GABA and neurokinin-1 pathways in this area has been demonstrated (Brownstein *et al*, 1976). The proposed functional significance of neurokinin-1 in the substantia nigra is further supported by the low levels of neurokinin-1 observed in this area of the brain of patients with Huntingdon's chorea. This disease is typified by a marked atrophy in the basal ganglia (Gale *et al*, 1978). The role of neurokinin-1 in this remains unclear.

Application of neurokinin-1 to the dorsal surface of the rabbit medulla elicits a considerable enhancement in respiratory rate and amplitude (Yamamoto *et al*, 1981). This suggests that the stimulatory effect of neurokinin-1 on ventilation is exerted on structures located in the lower brain stem. Neurokinin-1 also has a wide distribution and an important role in vasodilation of the circulatory system as well as a role in airway hyper-reactivity (Lundberg, 1983). It is also a potent stimulator of salivation in

several animals but not in others. The sialagogic properties of neurokinin-1 are well documented (Pernow, 1983). NK-1 also has a role in kidney function and water homeostasis (Mills *et al*, 1974).

Neurokinin-1 and other neuropeptides have been found to frequently coexist where high concentrations of both molecules have been localised to the same cell. This has occured in both central and peripheral neurons. The precise reasoning for this co-existence, based on the assumption that there is a mutual relationship, is not clear. The work carried out by Lundberg *et al* (1982) based on a population of neurons containing acetylcholine and vasoactive intestinal peptide immunoreactivities, suggests that one phsysiological function of the neuropeptides such as neurokinin-1 may be to potentiate the effect of co-existing classical neurotransmitters.

Further evidence suggesting the involvement of neurokinin-1 in the development of the central nervous system is presented by De Felipe *et al* (1995). The chemotropism of developing axons is well established. The results of De Felipes' research indicate a role for neurokinin-1 in regulating the release of chemo-attractants from the cells of the embryonic rat floor plate and subsequently the layout of the developing neuronal pathways.

1.4.4 Biosynthesis and Degradation

The biosynthesis of neuropeptides and the activation of those peptides is detailed in Sections 1.3.3 and 1.4.4. Using ^{35}S-methionine and ^{3}H-proline in incorporation studies, the biosynthesis of neurokinin-1 in dorsal root ganglia has been researched (Harmar *et al*, 1981). This peptide is synthesized ribosomally and is transported to nerve terminals by an anterograde movement. In the sciatic nerve for example, transport has been determined to be about 1mm per hour. Faster-moving neurokinin-1 travels in a peripheral direction along axons at about 120mm per day (Harmar *et al*, 1981). Demecoline, which inhibits the axonal transport of vesicles (including those containing neurotransmitters), causes accumulation

of neurokinin-1 in the ganglion. Neurokinin-1 is synthesized in the cell bodies of the dorsal root ganglia from where it migrates along axons to the dorsal horn of the spinal cord and the substantia gelatinosa of the medulla. It is also transported peripherally to the sensory neurons and the nerve terminals situated in peripheral tissues (Takahashi *et al*, 1975). Neurokinin-1 is released from nerve terminals upon electrical stimulation and when evoked by potassium (Otsuka *et al*, 1976).

Many tissues and organs have the capacity to degrade neurokinin-1 (Berger *et al*, 1979) and the location of this activity may be indicative of the sites where the neurokinin-1 ligand is released. The capacity to degrade neurokinin-1 correlates strongly with the levels of neurokinin-1 present in that tissue, for example in the gastrointestinal tract (Lee *et al*, 1982). The inactivation of neuropeptides such as NK-1 is detailed further in section 1.2.3.

1.4.5 Localisation of Neurokinin-1

Neurokinin-1 has been measured by radioimmunoassay using raised antiserum and monoclonal neurokinin-1 antibodies. Highly specific antibodies to certain portions of the neurokinin-1 peptide have been generated. For example, an antiserum directed specifically to the N-terminal part of the molecule was shown to have only 0.1% and 0.01% cross-reactivity to the SP^{2-11} and SP^{3-11} peptides, respectively, which lack the N-terminal sequence common to all tachykinins (Lee *et al*, 1980). Immunocytochemical techniques have also been successful used in the study of the cellular distribution of neurokinin-1. Indirect immunofluorecence techniques using conjugated, species-specific antibodies have been used to study neurokinin-1 levels in formalin-fixed tissue sections. The *in situ* hybridisation approach is the method currently favoured for determining the location of neurokinin-1. Warden *et al* (1988) used 48-base synthetic oligodeoxynucleotide probes radioactively labeled with ^{35}S to examine the cell labelling pattern for neurokinin-1 and neurokinin-B mRNA in the rat CNS. This study demonstrated the

sensitivity and specificity of *in situ* hybridisation histochemistry for mapping peptidergic neurons.

Pharmacological studies of the NK-1 receptor has been facilitated by the development of chemical antagonists. Antagonists such as baclofen (Saito *et al*, 1975) and other neurokinin-1 analogues have been used in an attempt to further understand the physiological role of the neurokinin-1 receptor. These chemicals are used in conjunction with the assays described above (Section 1.7.3).

1.5 THE TACHYKININ RECEPTOR FAMILY

Communication between cells is critical for a multicellular organism to function. The bovine, rat and human substance K receptors have been cloned and characterised, as have the rat neurokinin-1 and neuromedin K receptors [Masu *et al*, (1987), Sasai *et al*, (1989), Gerard *et al*, (1990), Yokota *et al*, (1989), Hershey *et al*, (1990)]. Many signal molecules use transmembrane signaling systems to illicit a response in a target cell. Such a system requires at least two components; the extracellular part which binds the ligand (namely, the receptor) and secondly, the intracellular part which continues the transduction cascade (namely, the internal effector molecule). Receptors have subsequently been classified as either ligand-gated ion channels (such as the nicotinic acetylcholine receptor), receptors with an intrinsic intracellular enzymatic domain (such as the insulin receptor), and receptors that couple to a guanine nucleotide-binding regulatory protein (G-protein) which in turn activates an effector molecule like an enzyme (for example the β-adrenoceptor, Burbach *et al*, 1992). The neuropeptide receptors, including the tachykinin receptors fall into the latter, G-protein coupled membrane receptor category. G-protein coupled receptors are detailed in section 1.5.4.

The effects illicited by the tachykinin neuropeptides are mediated by those membrane receptors. Pharmacological, biochemical and molecular cloning studies have demonstrated that there are three

distinct mammalian tachykinin receptors, NK1, NK2 and NK3. They are members of the G-protein-coupled receptor super-family, each of which possess seven putative transmembrane domains (Lefkowitz *et al*, 1989). The nomenclature NK1, NK-2 and NK-3 for the three tachykinin receptors replaced the previous system whereby those receptors were known as SP-P, SP-E and SP-N respectively (Iverson, 1990). An early report indicating three distinct types of TK receptor in rat peripheral tissues was presented by Maggi *et al* (1987). (TK = Tachykinin).

The tachykinin receptors are present at low concentrations as cellular components and are tightly embedded in the plasma membrane. The neurokinin-1 receptor was solubilised from bovine brain using detergents (Nakata *et al*, 1988). The high-affinity binding of neurokinin-1 to solubilised protein was inhibited by GTP, suggesting a G-protein related receptor coupling (Too *et al*, 1988). Study of the mammalian receptors was undertaken by Harada *et al* (1987) using receptor mRNA injection into *Xenopus* oocytes. The expressed receptor was examined electrophysiologically using ligand binding. These studies indicated that there are at least two distinct mRNAs encoding the neurokinin-1 and substance K receptors (Harada *et al*, 1987). Based on these experiments and others, the signal transduction mechanism for activation of the tachykinin receptors was proposed (Harada *et al*, 1987, Nakanishi *et al*, 1990); The tachykinin receptor expressed in an oocyte acts through a G-protein to cause production of IP_3 (inositol triphosohate) which in turn elevates cytoplasmic calcium ion levels. This calcium then activates a calcium-dependent chloride channel in the plasma membrane.

1.5.1 Primary Receptor Structure

The rat neurokinin-1, substance K and neuromedin K receptors consist of 407, 390 and 452 amino acids respectively. These receptors have seven hydrophobic segments, each consisting of about 20-25 uncharged amino acid residues, and share a significant sequence similarity with G-protein coupled receptors (Dixon 1988,

O'Dowd *et al*, 1989, Bonner *et al*, 1989). The amino terminal and carboxy terminal regions of the three receptors also show a pattern similar to the G-protein coupled receptors. Each of the receptors possess potential N-glycosylation sites at the amino terminal and multiple threonine and serine residues as possible phosphorylation sites at the carboxy terminal. Subsequently, all three tachykinin receptors are members of the G-protein family of G-protein-coupled receptors and possess seven transmembrane domains with extracellular amino termini and cytoplasmic carboxy termini (Nakanishi, 1989) (see figure 1.5). The core sequences covering the seven putative transmembrane domains are highly conserved (54-66%) among the three tachykinin receptors (Shigemoto *et al*, 1990). Furthermore, the β_2-adrenergic receptor contains two cysteine residues in the first and second extracellular loops which are believed to form a disulphide linkage between the loops (Dixon *et al*, 1987). Another cysteine residue, which is located immediately adjacent to transmembrane segment VII, has been shown to be palmitoylated. It can therefore act as the membrane-embedded anchor for this receptor. The three cysteine residues are conserved in the three tachykinin receptors.

1.5.2 Receptor Specificity

The combination of the cross-reactivity of the receptors with other TKs and the low abundance of the receptors in cell membranes has hindered their analysis using classical tools such as antisera, antagonists and radioligands (Maggio *et al*, 1988). The isolation of receptor gene cDNA clones has allowed them to be studied individually. From receptor ligand-binding studies, the Ki values of the receptors has been determined (see figure 1.5). The neurokinin-A and neuromedin K receptors are, in terms of ligand/agonist specificity, more closely related to each other than to the NK-1R. This contrasts with the higher overall homology between the neurokinin-1 and neuromedin K receptors at the amino acid level. Marked species-specific differences in NK-1R pharmacology have been noted. For example, the antagonist CP96345 has been shown to be two orders of magnitude less potent on NK-1 receptors in rat and

mouse than in human (Snider *et al*, 1991). The converse is observed for the antagonist RP67580 (Garrett *et al*, 1991). Furthermore, the three TK receptors show desensitisation in response to repeated application of agonists, but to different degrees (Shigemoto *et al*, 1990). The pharmacological approach to the study of neuropeptides and their receptors is discussed in greater depth later in this text.

Figure 1.5 : The Ki Dissociation Values for Three Rat TK Receptors

1. For the Rat NK-1 Receptor

Ki for neurokinin-1 was 1.5×10^{-10} M.
Ki for neurokinin-A was 1.7×10^{-8} M.
Ki for neuromedin K was 5.3×10^{-8} M.

2. For the Rat NK-2 Receptor

Ki for neurokinin-A was 3.0×10^{-10} M.
Ki for neuromedin K was 5.6×10^{-9} M.
Ki for neurokinin-1 was 1.0×10^{-7} M.

3. For the Rat NK-3 Receptor

Ki for neuromedin K was 4.2×10^{-10} M.
Ki for neurokinin-A was 2.9×10^{-8} M.
Ki for neurokinin-1was 1.9×10^{-7} M.

Figure 1.5: The Ki dissociation values as determined for the the three rat tachykinin receptors with respect to three of the potential tachykinin ligands; neurokinin-1, neurokinin-A and neuromedin K (taken from Shigamoto *et al*, 1990). The relative potencies of the ligands is indicated by the Ki values given.

Figure 1.6 : Transmembrane Topology of the Tachykinin Receptors

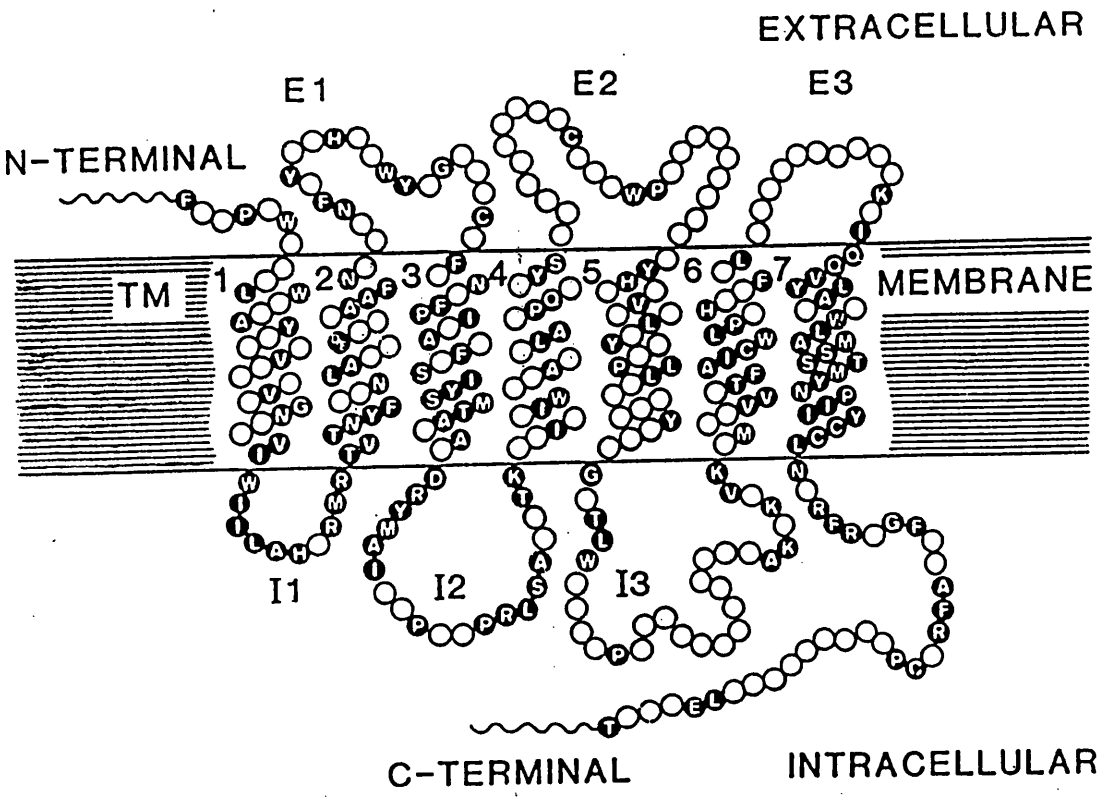


Figure 1.6: Generalised transmembrane topology of the tachykinin receptors as prototypes of G-protein-coupled receptors. The model spans the regions of the NK-1, NK-2 and NK-3 receptors of the rat from the first to the last identical amino acid in the three receptors. Consequently, most of the N- and C-terminal regions are not shown (taken from Shigamoto *et al*, 1990). Amino acids in black are identical in all three receptors. The nomenclature of the transmembrane domains (TM1-7), intracellular (I1-3) and extracellular loops (E1-E3) is indicated (Burbach *et al*, 1992).

Figure 1.7 : Linear Map of the Five Rat NK-1 Receptor Gene Exons

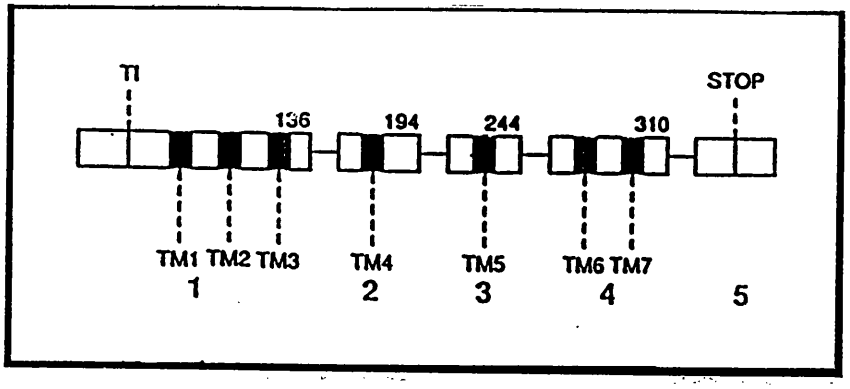


Figure 1.7: Linear map of the five exons of the rat NK-1 receptor and the corresponding trransmembrane segments encoded by those exons (taken from Watson *et al*, 1994; G-protein linked receptor handbook).

Figure 1.8 : Structural Model of the Human NK-1 Receptor

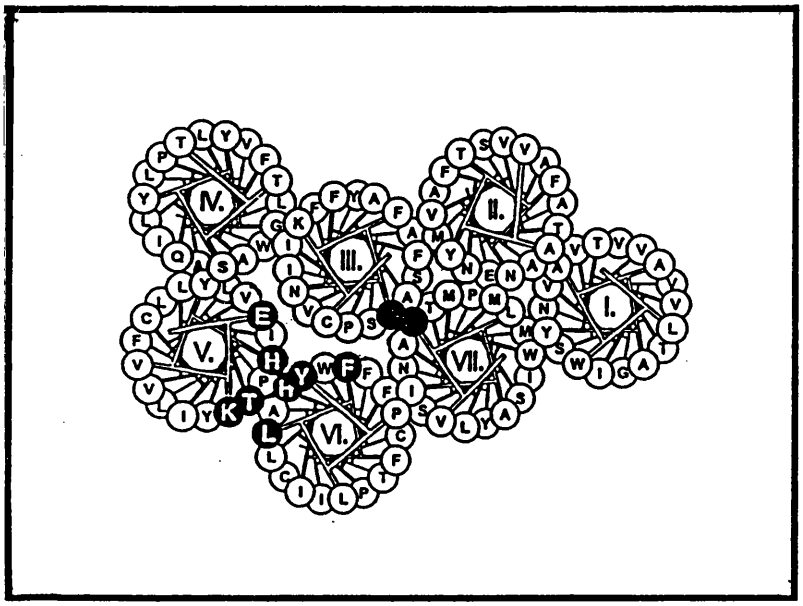


Figure 1.8: Helical-wheel representation of the structural model of the human NK-1 receptor. An anti-clockwise orientation of the helical wheels is shown in an outside-inwards view of the receptor. The putative binding site for the non-peptide antagonist CP96, 345, as determined by mutational analysis, is indicated in white on black sysmbols (taken from Elling *et al*, 1995).

Tachykinin family

1.5.3 Tissue Distribution

Using the corresponding cDNA as a hybridisation probe, the tissue distribution of the three tachykinin receptor mRNAs in rat was determined (Tsuchida *et al*, 1990). The NK-1R (neurokinin-1 receptor) mRNA was found to be distributed widely in both the central nervous system (CNS) and in peripheral tissues. In the CNS, it is expressed at high levels in the hypothalamus, the striatum, the spinal cord and the olfactory bulb. Of the peripheral tissues, it is highly expressed in the submandibular gland, the urinary bladder and the small and large intestines. The NK-3 receptor (NK-3R) mRNA is also expressed in both the CNS and in peripheral tissues, but is expressed at its highest levels in the cerebellum and the hypothalamus. The NK-2 receptor (NK-2R) mRNA however, appears to be expressed principally in peripheral tissues such as the stomach,

the urinary bladder and the adrenal glands. The tissue distribution of the tachykinin receptor mRNAs is therefore different in CNS and peripheral tissues (Tsuchida *et al*, 1990). Quantitative receptor autoradiography using radiolabelled tachykinins was used to localise and characterise tachykinin binding sites in rat CNS and peripheral tissues (Mantyh *et al*, 1989). Furthermore, the distribution of the mRNA for the precursor tachykinin peptides (neurokinin-A for example, Maggio *et al*, 1984) does not mirror the mRNA distribution of the corresponding tachykinin receptor (neurokinin-A for example, Sasai *et al*, 1989). The NK-1 receptor was also shown to be expressed in a subset of floor plate cells in the embryonic rat spinal cord which suggests a potential role for substance P in spinal cord development (Heath *et al*, 1995). The rat substantia nigra contains the highest concentration of neurokinin-1 in the brain but has no or few NK-1 receptors. This is an example of the non-correlation between the density of neurokinin-1 binding and neurokinin-1 innervation (Mantyh *et al*, 1984, Sivam *et al*, 1992). Whitty *et al* (1995) quantitated, localised and investigated the regulation of the NK-1, NK-2 and NK-3 receptors in the rat substantia nigra. Relative levels of receptor were found to be NK-3 > NK-1 >>> NK-2. The evidence presented by Whitty *et al*(1995) indicates that neurokinin-1 released from striatonigral nerve terminals may act on nigral cells through NK-1 receptors, whereas NK-A (which is a co-transmitter of neurokinin-1) may act preferentially on the same dopaminergic neurons through NK-2 receptors.

Stoessl *et al*, (1994) generated 6-hydroxydopamine lesions in the rat brain in order to study the localisation of striatal and nigral tachykinin receptors. The results indictate that both NK-1 and NK-3 receptors within the striatum are predominantly post-synaptic with respect to dopamine neurons, whereas nigral NK-3 receptors are located on dopaminergic neurons in this region of the brain.

Using multiple sequence alignment, a high degree of homology at the amino acid level, ranging up to 90%, was found *within* families of G-protein linked receptors such as the tachykinin receptors. However, homology *between* families was found to be as low as 20% (Attwood *et al*, 1991).

1.5.4 G-Protein-Coupled Receptors

G-proteins are guanosine nucleotide-binding regulatory proteins which are often coupled to membrane receptors. G-proteins function as intermediates in the transmembrane signal transduction pathway. They mediate the actions of many extracellular messengers and stimuli ranging from photons and small molecules (ie. catecholamines) to large molecules such as glycopeptide hormones (Burbach *et al*, 1992). They are found in many organisms from man to viruses. In figure 1.9 there is a list of 16 (courtesy of Burbach, 1992) different G-protein coupled receptors for neuropeptides or small peptide hormones. This signal transduction system comprises three components; the receptor, the G-protein and the effector molecule. The role of G-proteins in transmembrane signaling was first indicated by an absolute guanosine triphosphate (GTP) requirement for the hormonal activation of adenyl cyclase (Rodbell *et al*, 1971). There are many such receptors including those for a number of protein, biogenic amines, polypeptide hormones and several neurotransmitters. The rhodopsin receptor and the β-adrenergic receptors (for agonists such as epinephrine) are the best characterised members of the G-protein coupled family of receptors.
G-proteins are heteromers and comprise an α, β and a γ-subunit (in order of descending molecular weight). The α-subunits differ from one member of this family to another whereas the β– and γ-subunits are common to some of the receptors. The G-proteins function cyclically: the system is activated when GTP binds to the α-subunit of the G-protein. Hydrolysis of that GTP to GDP and Pi either initiates or is responsible for deactivation of the system. Dissociation of GDP appears to be rate limiting (Gilman *et al*, 1987). This system cycles each time the receptor is stimulated. All G-protein coupled receptors show a high amino acid homology particularly between the seven transmembrane segments which span the cell membrane as has been shown for the rhodopsin receptor (Hargrave *et al*, 1982). The transmembrane domains are believed to be formed by α-helices of hydrophobic amino acids (Kyte *et al*, 1982). Cysteine amino acid residues are often palmitoylated and provide membrane anchors for the receptors. Serine residues are proposed to be targets for phosphorylation by protein kinases.

Ligand binding is believed to take place in a pocket formed by the seven transmembrane domains together with the extracellular loops (Tota *et al*, 1990). A three dimensional model of the G-protein coupled receptors was proposed by Findlay and Eliopoulos (1990).

Many neurotransmitter receptors have been identified, the majority of which belong to this G-protein coupled receptor superfamily. It has become evident that there are now more than two distinct classes within the G-protein coupled peptide receptors. One is typified by the NK-2 receptor which is closely related to receptors for the biogenic amines and thyroid stimulating hormones. The common feature of this class is the strictly conserved amino acid sequences and motifs. Another class, the 'secretin-like' G-protein coupled receptors have no significant sequence homologies with the neurokinin-like receptors yet have their own common amino acid motifs (Burbach *et al*, 1992).

Figure 1.9 : Cloned Neuropeptide Receptors and Cloning Strategy Used.

Receptor	Species	Cloning Strategy	Reference
Angiotensin A1	Man	Homology screening	Furuta *et al*, 1992
Angiotensin A1	Man	COS expression	Murphy *et al*, 1991
Angiotensin A1	Rat	Homology screening	Iwai *et al*, 1991
Bradykinin B2	Rat	Oocyte expression	McEashern *et al*,1991
Calcitonin	Pig	COS expression	Lin *et al*, 1991
Endothgelin (ETa)	Man	Protein purification	Nakamuta *et al*, 1991
Endothgelin (ETa)	Man	Homology screening	Sakamoto *et al*, 1991
Endothgelin (ETa)	Man	Homology screening	Ogawa *et al*, 1991
Endothgelin (ETa)	Rat	COS expression	Sakurai *et al*, 1990
Endothgelin (ETa)	Cow	Protein purification	Saito *et al*, 1991
Neuromedin B	Man	Homology screening	Corjay *et al*, 1991
Neuromedin B	Rat	Homology screening	Wada *et al*, 1991
Neuromedin K	Rat	Oocyte expression	Shigemoto *et al*, 1990
Neuromedin K	Man	PCR	Buell *et al*, 1992
Neuromedin K	Man	Homology screening	Takahashi *et al*, 1992
Neuropeptide Y	Bovine	PCR	Rimland *et al*, 1991
Neuropeptide Y	Fruit fly	PCR	Li *et al*, 192
Neurotensin	Rat	Oocyte expression	Tanaka *et al*, 1990
Substance K	Man	Homology screening	Gerard *et al*, 1990
Substance K	Rat	Homology screening	Sasai *et al*, 1989
Substance K	Cow	Oocyte expression	Masu *et al*, 1987
Substance P (NK-1)	Man	PCR	Takeda *et al*, 1991
Substance P (NK-1)	Man	Homology screening	Takahashi *et al*, 1992
Substance P (NK-1)	Rat	PCR	Hershey *et al*, 1990
Substance P (NK-1)	Rat	Oocyte expression	Yokota *et al*, 1989
Tachykinin (type 1)	Fruit fly	Oocyte expression	Li *et al*, 1992
Tachykinin (type 2)	Fruit fly	PCR	Monnier *et al*, 1992

Figure 1.9: A list of cloned and identified neuropeptide receptors. The author who published each receptor and the cloning strategy used to clone each receptor is also given (taken from Burbach *et al*, 1992).

1.5.5 The Neurokinin-1 (NK-1) Receptor

The NK-1 receptor (NK-1R), which was previously referred to as the substance P receptor (SPR), is discussed in depth in (section 1.6).

1.5.6 The NK-2 Receptor: Substance K Receptor

The distribution of this receptor is limited to the peripheral tissues such as the smooth muscle of the gastrointestinal, respiratory and urogenital systems and there is little evidence of its expression in the central nervous system. Neurokinin-2 receptors exhibit an affinity for neurokinin-A that is 10 and 1000 times higher than for neurokinin-B and neurokinin-1, respectively (Whitty *et al*, 1995, see Section 1.5.2). The human NK-2 receptor gene has been localised to chromosome 10 and encodes a 398 amino acid protein with a molecular weight of 44,425 Daltons. This receptor is palmitoylated at cysteine residues Cys324 and Cys325, is glycosylated at asparagine residues 11 and 19 and has a disulphide bond between Cys106 and Cys325 residues.

Neurokinin-A (substance K) is the most potent mammalian tachykinin ligand to the NK-2 receptor and can induce receptor activation at low nanomolar concentrations (Watson *et al*, 1994). More than one sub-type of the rat NK-2 receptor have been demonstrated using [^{125}I]His-neurokinin-A (Brown *et al*, 1992). This is probably a consequence of this receptor being activated by more than one ligand (see Figure 1.5). Neurokinin-A is found at high concentrations in the substantia nigra. When microiontophoresed on to single neuronal units of the rat substantia nigra, NK-2 excites approximately 50% of both dopaminergic and non-dopaminergic neuron (Innis *et al*, 1985). This degree of response corroborates the reported high concentration of NK-2 receptors in the substantia nigra.

The identification of genomic DNA clones for the murine substance K receptor was reported by Sundelin *et al* (1992). The coding sequence of this gene was found to be more than 13 Kb including the five

distinct exons. As reported for the mouse neurokinin-1 receptor (Sundelin *et al*, 1992, above) the genomic clone DNA was used to isolate corresponding cDNA clones using the polymerase chain reaction method. The deduced amino acid sequence was found to be 94% identical to the rat and 85% identical to the human sequence.

A cDNA clone for the bovine substance K receptor was isolated by combining molecular cloning technology and an electro-physiological assay in *Xenopus* oocytes (Masu *et al*, 1987). The DNA and amino acid sequence analysis of this clone demonstrated that the substance K receptor is a 384 amino acid polypeptide. It contains seven hydrophobic domains which are believed to be membrane-spanning domains indicating that the bovine substance K receptor is also a G-protein linked receptor (Masu *et al*, 1987).

A cDNA of the rat substance K receptor has been isolated, also by molecular cloning (Sasai *et al*, 1989). From the 2,979 bp cloned rat substance K receptor cDNA sequence, the sequence of a 390 amino acid polypeptide was deduced. It was found to comprise seven hydrophobic motifs and was shown to share a significant similarity with other G-protein linked receptors (O'Dowd *et al*, 1989). This similarity is particularly strong in the transmembrane segments III and VII. RNA blot hybridisation studies demonstrated that the rat substance K receptor is encoded by two mRNA transcripts which differ at the 5' end of the 5'-untranslated sequence. Substance K receptor transcripts were shown to occur in the gastrointestinal tract as is characteristic of the tachykinin (brain-gut) peptides.

Xiao-Jiang *et al* (1991) screened a *Drosophila melanogaster* embryonic cDNA library at low stringency using bovine NK-2 receptor-derived probes. cDNA clones encoding a *Drosophila* receptor for tachykinin-like peptides were identified. The cDNAs were expressed in *Xenopus* oocytes and analysed using Northern, polymerase chain reaction and *in situ* hybridisation techniques. Peak expression of this receptor gene was noted at the later stages of embryogenesis and coincided with periods of significant neuronal development in *Drosophila* fruitflies.

1.5.7 The NK-3 Receptor

The neuropeptide NK-3 acts preferentially on the NK-3 receptor (NK-3R). NK-3 receptors have been identified in guinea pig cerebral cortex and in brain slices of rat and mouse (Guard *et al*, 1990, Dietl *et al*, 1991, Bergstrom *et al*, 1987). The human NK-3 receptor was cloned, characterised at the molecular level and its distribution determined by Buell *et al* (1992). This research group chose to do this by screening a human genomic bacteriophage library in order to isolate and subclone the exons encoding the translated portion of the NK-3 transcript. The cloned sequence allowed the design of human NK-3 specific PCR primers and subsequently the amplification of NK-3 cDNA. Further study demonstrated that this receptor is expressed in many areas of the human brain and selectively in some peripheral tissues. The human NK-3 receptor encodes a 465 amino acid protein with a molecular weight of 52,201 Daltons. This receptor is palmitoylated at cysteine residue 374, is glycosylated at asparagine residues 23, 50 and 73, and a disulphide bond exists between residues Cys158 and Cys233. Neurokinin-3 receptors are the least characterised of the tachykinin receptors.

cDNA clones for the rat neuromedin K (neurokin-B) receptor were isolated from a brain cDNA library by Shigemoto *et al* (1990). mRNA transcribed from this cDNA was injected into *Xenopus laevis* oocytes whereupon the expressed receptor was studied electrophysiologically. The rank order of ligand affinity for the receptor was neuromedin K (NKB) > substance K (NKA) > neurokinin-1 (NK1). This rat receptor was found to be a 452 amino acid peptide which is another G-protein-coupled receptor (Shigemoto *et al*, 1990).

1.6 THE NEUROKININ-1 (NK-1R) RECEPTOR

Recent technological progress in the fields of molecular biology and molecular pharamacology indicates that there are multiple receptors and receptor subtypes for neurotransmitters and neuropeptides. To recapitulate, three tachykinin receptors have been identified,

namely the neurokinin-1 (NK-1), NK-2 and NK-3 receptors, and are described above. The three receptor gene cDNAs were cloned from mammalian tissue in 1987 (Masu *et al*, 1987, Nakanishi *et al*, 1987). These receptors are pharmacologically distinct (Watson *et al*, 1987), but when they are added at a high enough concentration, each of the endogenous mammalian tachykinin neuropeptides can elicit a response from each of the three receptors. This is most likely to be a consequence of the common carboxy terminal amino acid sequence which is characteristic of the tachykinin peptides. The mammalian NK-1 receptor has been localised to both peripheral and brain tissues. Of the mammalian tachykinins which are potential ligands for this receptor, neurokinin-1 has the highest affinity with a Kd = 0.5-1.0. The rat NK-1 receptor cDNA has been cloned (Yokota *et al*, 1989). In the same year, Nakaniski and his research team in Kyoto, Japan, cloned the full rat neurokinin-1 (NK-1R) receptor gene. They found that the rat neurokinin-1, substance K and neuromedin K receptors displayed a high degree of homology in their amino acid sequences (Masu *et al*, 1989). Overall percentage identity was found to be between 47 and 58%. From the cDNA, the amino acid sequence was determined to total 407. Seven putative transmembrane domains have been identified and a high sequence similarity to other G protein-coupled receptors was shown to exist (Yokota *et al*, 1989).

Binding of the ligand to the NK-1 receptor activates it and consequently results in the hydrolysis of inositol phospholipid as demonstrated in several tissues such as the rat brain (Mantyh *et al*, 1984). The activation of the NK-1 receptor is typical of the G-protein coupled receptors (detailed in section 1.5.4. Appell *et al* (1992) used three classes of NK-1 receptor antagonist to demonstrate the species-specific difference in the human, guinea-pig, and rat receptors; heterosteroids, cyanines and modified peptides. Key amino acid changes or altered lipid environments are the most likely explanation for the differences noted. Regulation of NK-1 receptor synthesis most likely occurs at the transcriptional level. Expression of the the PPT-A and NK-1R genes occur at the same developmental stage (see *in situ* results, chapter 6) but not at the same location.

1.6.1 Distribution of the NK-1 Receptor

The NK-1 receptor is expressed by neurons and glial cells in the mammalian CNS and in smooth muscle cells, endothelial cells, fibroblasts, and many circulating immune and inflammatory cells. In the spinal cord, the number of NK-1 binding sites is highest in dorsal horn laminae I and II, the intermediolateral cell column and in lamina X (Charlton *et al*, 1985). Ventral horn regions demonstrate moderate levels of NK-1R. In the dorsal horn, NK-1-binding sites are located postsynaptic to sensory nerve terminals. However, in the intermediolateral cell column, the NK-1-binding sites are located on the autonomic and somatic motor neurons (Helke *et al*, 1986). There is a low density of NK-1 binding sites in the substantia nigra (SN) in contrast to the high relative density in the caudate-putamen (Buck *et al*, 1986). The low density in the SN contrasts with the number of neurokinin-1-immunoreactive terminals in this region. However, NK-1 receptor density varies under certain conditions. For example, the normal rabbit optic nerve lacks NK-1 receptors. However, 99 days after transection, this nerve contains the highest density of NK-1 receptors found in the whole rabbit forebrain. As the optic nerve is in effect a glial scar at this stage, glial cells *in vivo* appear to preferentially and transiently express high levels of NK-1 receptor (Mantyh *et al*, 1989). In rat, relatively high levels of NK-1R mRNA were found in the urinary bladder and the sublingual salivary gland, moderate levels in the submandibular salivary gland, hippocampus, midbrain, and olfactory bulb, and lower levels in the remainder of the CNS and the alimentary canal (Hershey *et al*, 1991).

Figure 1.10: Distribution of Rat mRNA Encoding the NK-1 Receptor.

RNA source	pg mRNA/25µg total RNA (mean +/-1 standard error)
Central Nervous System	
Whole brain	130 +/- 10
Olfactory bulb	230 +/- 10
Cortex	135 +/- 10
Striatum	265+/- 10
Hypothalamus	190 +/- 10
Hippocampus	260 +/- 10
Midbrain	240 +/- 10
Cerebellum	20 +/- 10
Medulla-pons	210 +/- 10
Spinal cord	120 +/- 10
Retina	115 +/- 10
Other Tissues	
Sublingual gland	355+/- 10
Submandibular gland	270 +/- 10
Parotid gland	50 +/- 10
Stomach	25 +/- 5
Duodenum	75 +/- 5
Jejunum	110 +/- 5
Ileum	165 +/- 5
Colon	80 +/- 5
Liver	Not detectable
Spleen	Not detectable
Urinary bladder	400 +/-20
Testis	Not detectable
Lacrimal gland	20 +/-5
Trigeminal ganglion	Not detectable
Thyroid	Not detectable
Anterior pituitary	Not detectable
Muscle	Not detectable

Figure 1.10: Tissue distribution of mRNA encoding the rat NK-1 receptor. Quantification made use of a mRNA standard (taken from Hershey *et al*, 1991).

Endoglycosidase F was used to biochemically determine and characterise that there are two forms of NK-1R in the rat submaxillary gland which differ in the length of their respective carboxy termini (Kage *et al*, 1992). Fong *et al* (1992) have reported the existence of a shorter form of the human NK-1R which is generated by alternative splicing and has a truncated carboxy terminal which lacks seven amino acids. Neurokinin-1 has an affinity for this form of the receptor which is reduced approximately 10-fold.

1.6.2 Organisation of the NK-1R Gene

1.6.2.1 The Human NK-1R Gene

The molecular cloning of the human neurokinin-1 receptor (NK-1R) gene was published in 1992 (Takahashi *et al*, 1992). The gene comprises 5 exons spanning more than 60 kb and is a 407 amino acid G-protein coupled receptor. Clones were isolated from a human genomic bacteriophage library using sequences derived from the rat SPR cDNA (plasmid prTKR2, Shigemoto *et al*, 1990). Segments of the same plasmid were used as a hybridisation probe used to identify the mouse genomic NK-1R clones described in this thesis. The human genomic DNA clones which were identified did not overlap with one another indicating that the gene spans more than 60 kb. Exons 1 to 5 of the human NK-1R gene are 977, 196, 151, 197 and approximately 3300 bp respectively. The exons are interspersed by introns of >22 kb, >26 kb, >7 kb and approximately 2.2 kb. RNA blot hybridisation analysis was used to examine the 5' sequence of this gene and a TATA box was found to be located 24-27 bp upstream of the putative transcription initiation site. The TATA box is involved in the polymerase nucleotide selection which initiates transcription at a specific site. Furthermore, several promoter, enhancer and regulatory DNA elements were also identified in the 5'-flanking region including a cyclic AMP responsive element (TGATGTCT) and an AP-1 sequence.

The human NK-1R gene is located on chromosome 2. This 407 amino acid receptor has a molecular weight of 46,248 Daltons. Post-translational modifications of this receptor include glycosylation sites on two amino acids; Asn14 and Asn18. Furthermore, a disulphide bond is believed to form between the cysteine residues at positions 105 and 180 and palmitoylation of the receptor occurs at cysteine residues 322 and 333 thus providing membrane-inserted anchoring side-chains for this receptor (Takeda *et al*, 1991). The gene organisation and primary structure of the human NK-1 receptor (and the neuromedin K receptor) has been presented by Takahashi *et al* (1991). They demonstrate that it is a G-protein coupled receptor encoded from 5 exons extending over 60 kb. RNA blot hybridisation and primer-extension analyses were used in Takahashi's study of the human NK-1 receptor gene.

1.6.2.2 The Rat NK-1R Gene

Firstly, the cDNA for the rat NK-1R gene was cloned by Hershey *et al* (1990). Genomic clones for the rat NK-1R gene were then cloned and the expression pattern of the gene analysed, again by Hershey *et al* (1991). Again the gene was shown to be G-protein coupled and to comprise 5 exons extending over 45 kb. Exon 1 was found to encode the entire 5' untranslated region and the coding region through to the end of the third transmembrane domain. Exon 2 contains the second intracellular region, the fourth transmembrane domain and the second intracellular region. Exon 3 encodes the fifth transmembrane domain and the third intracellular region. Exon 4 contains the sixth transmembrane domain, the third intracellular region and the seventh transmembrane domain. Exon 5 encodes the cytoplasmic carboxyl region and the entire 3' untranslated region. The exons are 965, 195, 151, 197 and 2010 bp in length respectively. The transcription initiation site was also determined using solution hybridisation-nuclease protection experiments.

1.6.2.3 The Murine NK-1R Gene

Mouse genomic DNA clones of the NK-1R gene were isolated from a murine cosmid library in 1992 (Sundelin *et al*, 1992) using

heterologous probes composed of short sequences derived from the rat NK-1R gene cDNA (Hershey, 1990). The gene was found to be more than 30 kb in length. The same research group then isolated the corresponding cDNA for the mouse neurokinin-1 receptor gene using polymerase chain reaction methodology. The isolated cDNAs were transcribed *in vitro* and microinjected and expressed in *Xenopus* oocytes. Ligand binding studies of the newly expressed receptor confirmed its pharmacological identity. The amino acid sequence of the mouse neurokinin-1 receptor gene was determined to be 99% identical to the rat sequence.

1.6.3 Structure and Proposed Binding Sites of the NK-1 Receptor

Masu *et al* (1987) was the first to report the primary structure of a tachykinin receptor. This receptor was the neurokinin-1 receptor (Masu *et al*, 1987). The precise site where the ligand binds to the G-protein component of the NK-1 receptor has not yet been identified conclusively; however there has been considerable mutational analysis of the NK-1 receptor (Fong *et al*, 1993, Gether *et al*, 1994) and NK-2 (Gether *et al*, 1993) receptors from which binding sites have been proposed. It has been indicated that the non-peptide antagonists act through residues located between the 7 transmembrane domains, whereas natural peptide agonists bind mainly to residues scattered in the exterior part of the receptor (Fong *et al*, 1992, Rosenkilde *et al*, 1994). Mutational analysis published by Elling *et al* (1995) demonstrated the conversion of a NK-1 receptor binding site for the antagonist CP36,345 to a metal-ion binding site by the systematic introduction of histidyl residues (see figure 1.8).

1.7 APPROACHES TO STUDYING NEUROPEPTIDE FUNCTION

The brain of an individual will differ from the brain of another in the same way as their faces will vary, yet there are structures of the

brain that are common to all individuals. Indeed, those structures appear to be common to all mammals. The anatomist Lorente de No, after examining a mouse brain with a microscope, commented that its fine structure was little different from that of the human brain, except in proportion. Since many features of the brain are common and many animal and human behaviours are similar, it is possible to learn about the human brain by studying the brains of other animals and vice versa.

There are several approaches to studying brain function. The *comparative* approach is used to obtain clues to brain function by correlating structure and behaviour. The key finding is derived from this approach which allows mammals to be distinguished from other animals by the thickness of the outer layer of brain tissue; the neocortex. This structure is particularly large in humans. The developmental, or *ontogenetic* approach describes the changes in brain structure during the development of an individual. In this way, newly developing structures can be correlated to emerging behaviours and the younger brain can be used as a model of the adult brain. The *cytoarchitectonic* approach is based on describing the architecture of cells; their size, shape, structure, distribution and connections. This utilises the property whereby certain dyes have an affinity to specific parts of the nerve cell. The advent of the electron microscope accelerated research of this nature. Another approach is the traditional *pharmacological* approach which utilises the binding specificities of receptors for ligands. In this way an insight can be gained into the role and function of the receptors by the use of agonists and antagonists to that receptor (This approach is discussed in Section 1.7.3). The most recent approach to the study of brain function is the *molecular and biochemical* approach where an understanding of the biochemical organisation of the brain is pursued. It is clear that certain clusters of nerve cells synthesise, contain and can sectrete biochemical substances that are central to intercellular communication. The normal activity of those cells can be related to behaviour and conversely, the abnormalities in the functioning of cells at particular loci can be related to abnormal behaviour. Modern molecular biology has provided the latest set of tools that allow the neuroscientist to

adopt this biochemical approach. Molecular gene cloning, manipulation of nucleic acids and the means to introduce specific, targeted genetic mutations have fuelled the ongoing surge in molecular and biochemical neurobiological research.

1.7.1 Strategies for Cloning Neuropeptide Precursor and Receptor Genes

There are two general approaches to cloning neuropeptide receptors. The first is ligand based, where the starting point is a ligand from which the DNAs can be identified. The second approach is DNA based where the DNA is the starting point. The ligand-based approach includes the classical receptor purification method where a bound ligand identifies a receptor which is purified and its amino acid sequence determined. From this, oligodeoxynucleotides can then be designed as hybridisation probes for screening genomic or cDNA libraries in order to isolate clones. The human endothelin ETB receptor has been cloned in this way (Nakamuta *et al*, 1991). Another ligand-based approach includes the functional expression of cDNAs, where mRNA transcribed *in vitro* from a cDNA library are expressed in *Xenopus* oocytes. Detection of an expressed receptor is based on the electrophysiological response of the oocyte to the peptide ligand as measured by voltage clamp. The bovine neurokinin-A receptor was cloned in this way (Masu *et al*, 1987) and this approach was used in an early study of tachykinin receptors NK-1 and NK-2 (Harada *et al*, 1987). Another ligand-based method involves the use of mammalian COS cells rather than *Zenopus* oocytes. Receptors are functionally expressed in COS-1 or COS-7 cells after transfection of cDNA libraries. COS cells provide high levels of expression (Gluzman *et al*, 1981). This allows identification of a particular cDNA in an expressed library.

The DNA-based approaches to cloning neuropeptide precursor and receptor genes includes the use of a cDNA derived from one species to identify clones from a library derived from a different species. This approach was used in the research detailed in this thesis to clone the mouse NK-1 presursor and receptor genes. A segment of the rat NK-1 receptor cDNA was used as a heterologous hybridisation

probe to screen mouse genomic DNA libraries in order to identify DNA clones. This approach was used successfully and is detailed in depth in this text. This approach can also facilitate the identification of new, as opposed to known, receptors based on nucleotide homology. Another DNA-based, molecular approach uses degenerate oligodeoxynucleotide PCR (DOP-PCR) methodology, where degenerate primers to highly conserved regions such as the membrane spanning domains of G-proteins have been used to amplify related sequences from cDNAs (Frohman *et al*, 1988). Similarly, oligodeoxynucleotides derived in this way from conserved regions can be radioactively labeled and used as probes to screen genomic DNA or cDNA libraries.

1.7.2 Neuroantibodies and Immunocytochemistry

An alternative and complementary approach to the study of functional and developmental processes in the mouse central nervous system, which involves neuroantibodies, has been presented by Picciolo *et al* (1991). This is based on the local expression of specific monoclonal antibodies by cells of the central nervous system. This paper describes research which uses the anti-neurokinin-1 rat monoclonal antibody NC1/34HL. It was expressed in specific neuronal cell subpopulations. This antibody specifically recognises and binds the carboxy terminal of neurokinin-1, preventing it from binding to its receptor. This direct interference with the neuro-cellular communication pathway (by sequestering the neurokinin-1 ligand) is analogous to the gene targeted deletion of neurokinin-1 expression. In both cases, the normal action of the neuropreptide is likely to be altered. Results obtained in this way will complement, along with the classical pharmacological results, the molecular gene targeting approach to the study of the functional role of neurokinin-1. This approach is likely to be applied to other tachykinin neuropeptides in the future.

Antibodies have also been used to localise neuropeptide receptors in cell lines and tissues. The NK-1 receptor has been studied in two ways, each based on this antibody technique (Vigna *et al*, 1994).

Firstly, antisera were raised against the 15 amino acids that comprise the intracellular carboxyl terminus of the NK-1 receptor. The second approach involved the generation of antibodies raised against a chimeric NK-1 receptor. The chimeric receptor construct was genetically engineered (by polymerase chain reaction) to tag a hydrophilic 'flag' peptide (DYKDDDDK) onto the extracellular amino terminus of the NK-1 receptor which was expressed in rat kidney epithelial cells (Vigna *et al*, 1994).

1.7.3 The Pharmacological Approach

The traditional pharmacological approach to studying active ligands and their corresponding receptors involved the use of agonists and antagonists as tools. Pharmacological research in the field of neurobiology is extensive, as testified by the annual literature on the subject. It also serves to complement and corroborate the results determined using the molecular approach (or vice versa). The molecular approach is detailed elsewhere in this text. The pharmacological approach generates results which implicate certain peptides with certain biological processes.

1.7.3.1 Agonists and Antagonists

These molecules can be biochemical molecules or proteinaceous ligands such as the neuropeptides. Radiolabelled ligands ('radioligands') were used initially in the study of tachykinin receptor binding, the first report being that of Nakata *et al* (1978) using tritiated neurokinin-1 (^{3}H-NK-1). The potencies of various agonists have been determined in both pharmacological studies and in radioligand binding assays. The archetypal NK-1 assay is to determine the degree of contraction induced in the longitudinal muscle of the guinea-pig ileum in the presence of atropine. Both natural and synthetic tachykinin agonists and antagonists have been used in this type of research. Selective peptide antagonists for the NK-1 and NK-2 receptors have been developed. For example, to elucidate the role of tachykinin receptor subtypes in sensory nervous system processing, selective agonists to the NK-1 and NK-3

receptors were administered intrathecally. The time taken for conscious rats to respond to noxious radiant heat stimuli was measured. As the applied dose of the agonist, namely [Pro^9,$Met(O_2)^{11}$]NK-1 in this case, increased, the reaction time decreased proportionally (Laneuville *et al*, 1988). Further nociceptive responses including compulsive biting, scratching and licking of the hindlimbs were also induced by a NK-1 selective agonist (Papir-Kricheli *et al*, 1987). These experiments mimic the nociceptive effects of tachykinins when released from primary sensory nerve terminals of the dorsal horn and demonstrate the pharmacological approach and the study of pharmacologically-induced behavioural changes. A hexapeptide agonist named septide [$pGlu^6$, Pro^9]substance $P(SP)_{6-11}$ has been shown to be as potent as neurokinin-1 in eliciting smooth muscle contraction in several *in vitro* preparations (Pradier *et al*, 1994).

Pharmaceutical companies such as GlaxoWellcome have developed and published details of novel tachykinin antagonists (McKnight *et al*, 1991, Ward, 1990). Studies using one antagonist indicate the existence of NK-2a and NK-2b receptor subtypes (Maggi *et al*, 1992). More recently, potent and highly selective non-peptide antagonists have been identified for the NK-1 receptor (Garrett *et al*, 1991; Cascieri *et al*, 1992). Research has suggested, with the benefit of the non-peptide antagonists, that pharmacologically the NK-1 and NK-2 receptors may be capable of further sub-classification in the future. Neurokinin-1 antagonists include baclofen which is known to effectively antagonise the depolarising action of neurokinin-1 on spinal motoneurons (Saito *et al*, 1975). Baclofen is known to block the excitatory activity of neurokinin-1 but is not a specific neurokinin-1 antagonist.

There has been a surge in pharmacological study of the neuropeptides, and in particular the tachykinin receptors in the 1990s. This interest stems from the development of an array of selective, high affinity, non-peptide antagonists which pharmacolgists can use. The activity and distribution of the binding sites in the central nervous system for a non-peptide neurokinin-1 antagonist (CP-66,345) has been documented by McLean *et al* (1991). The cDNAs for the NK-1, NK-2 and NK-3 receptors have been cloned

and functionally characterised. Subsequently, chimeric and point-mutated receptors have been used to further elucidate the receptor structure-function relationship with regard to binding domains for peptide agonists and non-peptide antagonists, and agonist-stimulated second messenger responses (Nakanishi *et al*, 1993).

Antagonists are of greater importance, pharmacologically speaking, than agonists principally because antagonists bind to the recognition site of the receptor but lack intrinsic biological activity, whereas agonists bind and activate receptors. Antagonists can be used to distinguish between receptors. The peptidergic TK antagonists were developed by modification of the amino acid sequence.

1.7.3.2 Peptide and Non-Peptide Analogues

Another approach to peptide ligand research uses peptide analogues. The principle of peptide analogues is to modify a peptide in order to reduce the agonistic activity and to subsequently generate an effective inhibitor molecule. The rationale for the design of a series of cyclic peptide analogues to the tachykinin receptors is described by McKnight *et al* (1991). Several NK-1 analogues have been synthesized. Modification of the pharmacologically critical carboxy terminal amino acid sequence of the NK-1 peptide can generate a peptide with weak, or even no agonistic effect. For example, [D-Arg^{1}, D-Pro^{2}, D-Trp^{7,9}, Leu^{11}]NK-1 which is a particularly potent antagonists (Russell *et al*, 1983). The various neurokinin-1 antagonists have been pharmacologically tested on several tissues (see Pernow review, 1983).

1.7.4 The Molecular Approach to the Study of Neuropeptide Function

Genetics has traditionally been seen as the most effective way to study the molecular intricacies of complex biological systems. The basic tenet that genotype correlates directly to phenotype has been examined from both the phenotypic and genotypic perspectives using standard genetics and more recently, reverse genetics, respectively.

1.7.4.1 Reverse Genetics

The traditional approach to genetical analysis has been phenotype focussed, from where the corresponding genotype is pursued. This standard genetic analysis approach starts by observing discrete variations in well-defined aspects of phenotype before classifying mutations into functional units (genes) using complementation tests. Genes are then located on a genetic map using recombinational mapping. The genetic map can be related to the physical location of the gene in the DNA molecule using the distribution function of local recombination rates. Therefore, when the gene in question is identified, a direct link between alterations in the gene and its product and an effect on the biology of the organism can be determined. Organisms such as *Caenorhabditis elegans* and *Drosophila melanogaster* which have relatively small genomes and in a practical sense are easy to use, have been the favoured models for doing this. The rate-limiting component of this approach is the physical isolation of the DNA corresponding to a genetic locus.

In contrast, an alternative approach which guarantees the establishment of a correlation between altered genotype and altered phenotype is known as 'reverse genetics' (see figure 1.11). This approach starts with a cloned DNA sequence which is used to generate animals (such as mice) with mutated forms of a particular gene.

Figure 1.11 : Reverse Genetics and Standard Genetics

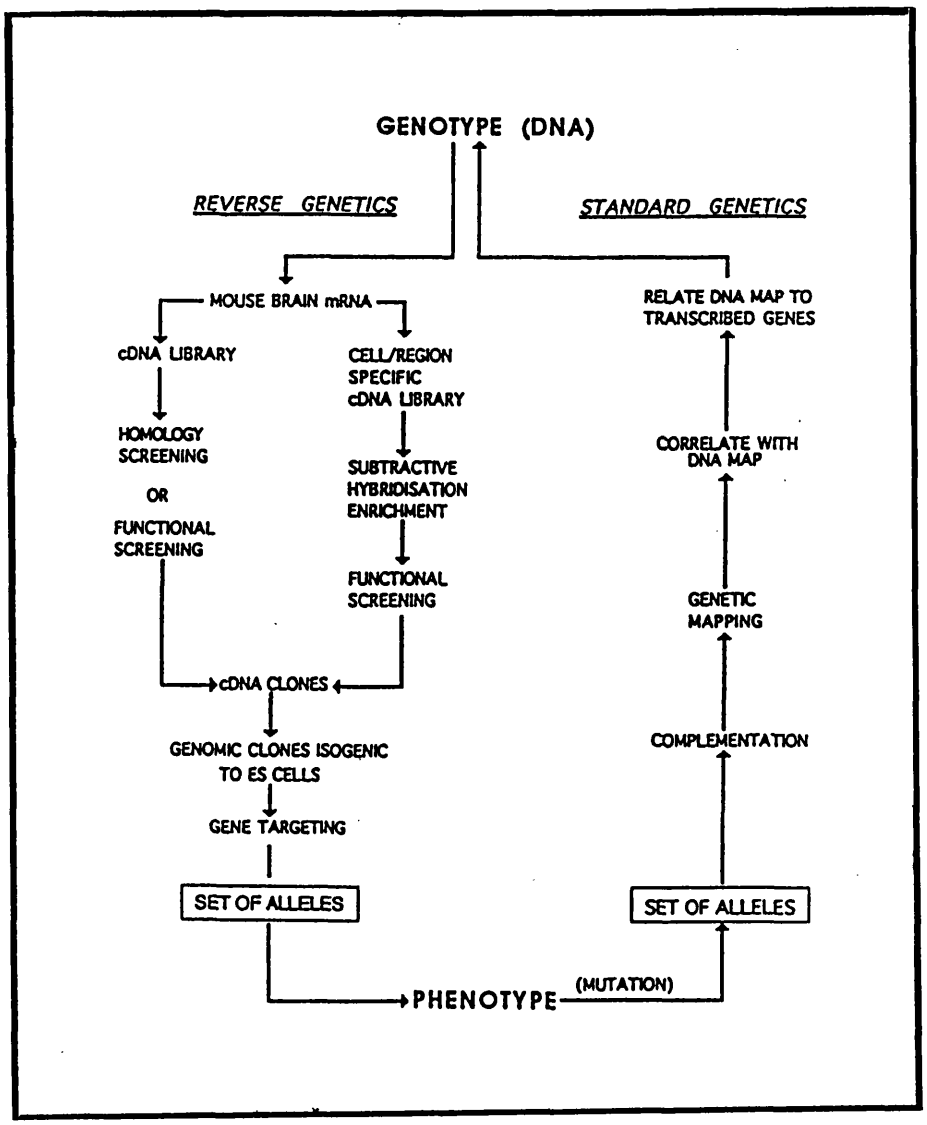


Figure 1.11: Reverse mouse genetics and its relationship to standard mouse genetics. Two different approaches to determining the phenotype/genotype relationship (taken from Davies *et al*, 1994).

1.7.4.2 Gene Targeting

The number of genes cloned in the last decade has increased rapidly, yet the function of many of those genes remains unknown. The recent development of gene targeting technology has provided the means to scrutinise genes and to study gene function in the living animal (review: Hooper *et al*, 1993).

Reverse genetics utilises gene targeting techniques in order to generate specific changes in genotype such as a loss-of-function mutations (which are referred to as gene 'knock-outs'). In essence, a gene targeting experiment involves the transfection of cells (embryonic stem cells in this instance) in tissue culture with a linearised, cloned sequence of isogenic genomic DNA which has been specifically modified (such as by the insertion of a marker gene into the exonic sequence of the cloned gene). The generation of mutant mice lacking the PPT-A or NK-1R gene is the objective. Gene targeting is used to 'knock out' one allele of the gene in ES cells. In the event of successfully generating targeted ES cell lines, those lines would be used to create chimaeric mice which, if germ line transmissible, would produce homozygous mutant mice for the PPT-A or NK-1R genes. Homologous recombination is required for the successful introduction of the modified sequence into the recipient, target genomic loci. This experimental approach was used successfully for the first time in mouse cells in 1985 (Lin *et al*, 1985) and in human cells in the same year (Smithies *et al*, 1985). This technique has been applied to embryonic stem (ES) cells resulting in the production of mutant mouse cell lines in tissue culture. ES cells are derived from the inner cell mass of 3.5 day post-coitum mouse embryos (at the blastocyst stage) and can be cultured *in vitro* while retaining totipotency. Targeted ES cell lines are injected into C57BL/6 blastocysts for example and are re-introduced into the uterus of a pseudo-pregnant mouse foster-mother, where they develop, forming a proportion of the tissues of the resulting chimaeric mouse. Any induced mutation should therefore enter the germ-line of the chimaera and the germ-line of any offspring transgenic mice of that chimaera. Therefore, a

strategy for the introduction of mutations in any mouse gene which has been cloned, does exist.

The frequency of the homologous recombination events between the target genomic DNA locus and the targeting fragment ('targeting vector') will vary from gene to gene. There is also a degree of background and non-specific recombination. The use of isogenic DNA helps to reduce this, as does the length of the targeting fragment used; a length of 6 kb in total with at least 1 Kb either side of the modified sequence, is recommended for a typical replacement vector (Davies *et al*, 1994). Screening recombinants for successful integration is performed using polymerase chain reaction (PCR) or Southern blotting. Several non-homologous integration events can occur including end extension repaired fragments (Aratani *et al*, 1992). For this scenario, the incoming fragment is extended by copying chromosomal sequences which are not included in the targeting vector. To circumvent the PCR analysis anomalies which can also arise at the screening stage as a result of this, Southern blotting is the definitive screen for targeted recombinants. In a targeting experiment, typified by Mansour *et al* (1988), a replacement vector and a positive-negative selection protocol resulted in a homologous to heterologous recombination event ratio of between 1:10 and 1:100 for most genes (as determined by G418 resistance).

There are two established gene targeting vector strategies; the replacement vector and the insertion vectors strategies (Thomas *et al*, 1987). The insertional strategy encompasses the integrative and the promoter-trap vectors (Davies *et al*, 1994) (see Figure 1.12). The replacement vector strategy is the most widely used approach and is the one which has been adopted in this study to target the cloned murine PPT-A and NK-1R genes described in this text. A replacement vector consists of a sequence of DNA homology interrupted within its exonic sequence by a selectable marker, such as a neomycin cassette which would typically comprise a promoter and polyadenylation sequences thus creating a functional neomycin phosphotransferase gene 'cassette'. In addition to being mutagenic in its own right, this cassette confers positive selection for the

genomic integration of the vector into ES cell chromosomal DNA (using G418 selection). For this, typically a minimum of 6 kb of genomic DNA is interrupted by the introduction of the selectable marker gene and the vector is linearized outside the region of homology. In order to minimise the high number of background recombinations which can arise from replacement recombination strategies, a counter-selectable marker gene is included and is typically positioned outside the region of homology. In most cases, the thymidinine kinase (TK) gene from the *herpes simplex I* virus (HSV1) is used for this purpose. This particular counter-selection marker is included in the mouse PPT-A and NK-1R gene targeting vectors detailed in this text. The replacement vector strategy is believed to work by homologously integrating the vector DNA by way of two reciprocal crossover events occuring between the two arms of homologous genomic sequence in the vector and the equivalent sequences in the endogenous gene. The exact mechanism has not been defined, but Ellis *et al*, (1989) proposed another possible mechanism where a single strand exchange within one of the regions of homology would occur. This would be followed by branch migration and resolution within the second arm of homology to complete integration up to the point where homology ends; this is typically where the TK counter-selection gene is located. Therefore, successful integration by this mechanism will result in the exclusion of the TK gene in the ES cell chromosomal DNA. However, in the case of random, non-homologous recombination events, the TK gene will be integrated and will thus permit the researcher to distinguish between precise and random recombination events. This selection is posssible using a toxic nucleotide analogue such as gancyclovir or FIAU (1-[2-deoxy-2-fluro-β-D-arabinofuranosyl]-5-iodouracil). These biochemicals act as substrates for the thymidine kinase gene in the vector but not in the cellular DNA. The system whereby selection *for* integration and *against* non-homologous integration is possible simultaneously is known as the positive-negative selection strategy (PNS) and the enrichment for targeted clones was reported to be up to 2000-fold for this technique (Mansour *et al*, 1988), although this figure is realistically nearer 20. In the case of the integrational vector strategy, various forms of engineered neomycin cassette can be used such as enhancer-less

cassettes, (Jasin *et al*, 1988), a neomycin cassette lacking a poly-adenylation signal, (Donehowe *et al*, 1992), or a promoter-less neomycin cassette (Jeannotte *et al*, 1991). This type of vector requires the chromosomal integration site to provide the sequences necessary in order to generate a functional neomycin gene/cassette. Stated another way, as a measure of successful recombination, the expression of the selectable marker (and/or a visible marker gene) is dependent upon integration into a transcribed gene (Skarnes *et al*, 1992). In effect, the marker gene is fused upstream of a splice acceptor site so that mRNA splicing leads to the production of a fusion protein. Therefore, in embryonic stem cells, gene-trap vectors of this kind generate mutations which can be examined in resultant transgenic mouse lines. For a promoter-less construct, approximately 1% of random integrants provide promoter sequences for neomycin activation. Therefore, this approach results in an enrichment for homologous integrations of approximately 100-fold.

The insertional vector strategy involves the placement of a selectable marker gene outside the region of genomic DNA homology in a given vector and for that vector to be linearized within the region of homology prior to ES cell transfection (Hasty *et al*, 1991). This generates a recombinational hotspot which results in an integration frequency 5-12 fold higher than that of the replacement vector strategy and results in the complete insertion of the vector into the target locus, thus creating a duplication of the region of homology at the integration site. This is resolved during later cell multiplication. An in-frame fusion is typically made of the *neo* coding region to the exon of the gene to be targeted. This form of mutational disruption of a gene will typically generate a null allele. As a result of this strategy, few random, non-homologous recombinations occur (also see Section 7. 4). Insertional vectors can be used to introduce specific subtle mutations into the target gene as demonstrated by the 'Hit and Run' and 'In-Out' vectors (Hasty *et al*, 1991, and Valancius *et al*, 1991, respectively). Insertional vectors use a two step recombinational approach to do this. The first step involves selecting those cells which have integrated the vector as determined by neomycin gene expression. The second step involves applying negative selection against the thymidine kinase

selectable marker. This can result in excision of the vector along with one copy of the duplicated sequence so that the desired mutation is retained within the genetic locus (This involves intrachromosomal recombination). (Promoterless strategy can also be used with replacement vectors.)

In contrast to the relatively crude loss-of-function mutations, gene targeting can also be used to introduce subtle, specific mutations as detailed above. This allows the study of the *in vivo* structure-function relationships of a particular protein or *in vivo* analysis of functional elements of promoters. Specific, single base alterations can be used to study the importance of particular sequences within a gene. The vectors described in this text are designed to generate whole gene knock-outs, while providing the foundation for the design of subtle mutations in the future.

There are several ways of introducing altered DNA sequences into the germ-line of a cell line in addition to plasmid or cosmid targeting vectors. Yeast artificial chromosome (YAC) clones, for example, have been used successfully to produce transgenic mice, both by DNA injection into fertilised eggs (Schedl *et al*, 1993) and transformation of ES cells (Strauss *et al*, 1993)

targetting efficiency than replacement vectors (see higher Thomas et al, 1992).

targetting efficiency than replacement vectors (see Thomas et al, 1992).

Figure 1.12 : Gene Targeting using Mouse Embryonic Stem Cells

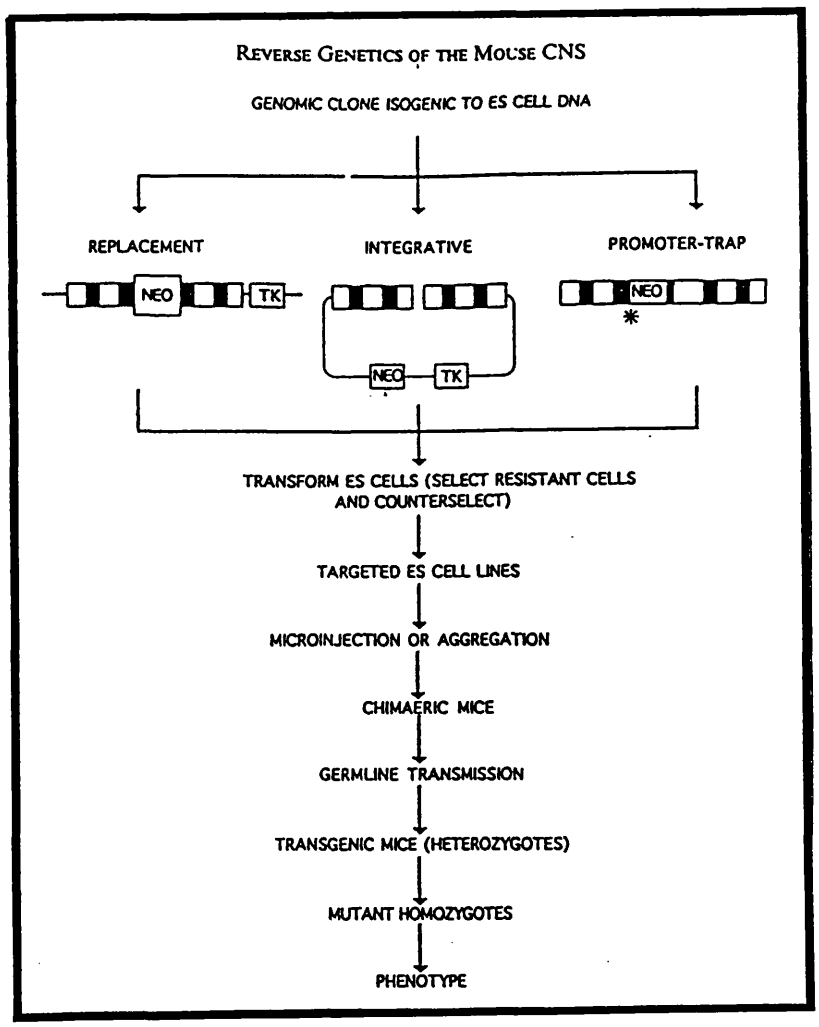


Figure 1.12: Gene targeting approaches using the mouse embryonic stem cell system. The principles and procedures of the replacement, integrative and promoter-trap approaches are detailed. NEO; neomycin phosphotransferase gene conferring G418 resistance, TK; herpes simplex virus thymidine kinase gene conferring sensitivity to gancyclovir or FIAU. Exons are shaded, introns are not shaded. The asterisk indicates that in promoter-traps the *neo* gene is fused in-frame with an exon of the gene to be targeted and lacks a promoter and translation start signal, whereas in replacement fragments it is out of frame and has normal regulatory and translation elements (Davies *et al*, 1994, Hooper *et al*, 1993).

1.7.4.3 Gene Targeting and Clinical Neuroscience

In the field of mammalian and human neuroscience, the predominant analytical approach remains the pharmacological one (as detailed above). The potential of either standard or reverse genetical analysis as a means of examining neural gene function has been on the whole neglected but has gained increased attention over the last few years.

When using standard genetics to study the central nervous system, one problematic area is the identification of a mutation at the phenotypic level. Several behavioural tests exist such as the elevated T-maze which is used to indicate the degree of anxiety. Subsequently, many of the observed mutations of the mouse are behavioural (Sidman *et al*, 1979). When successful, the molecular approach, however, will always generate a phenotypic marker. This can be reinforced in some instances by the occurence of some kind of behavioural manifestation. The molecular approach can therefore provide an excellent tool in the pursuit of an understanding of the role of specific molecules, and the genes encoding those molecules, in neurological conditions including neurodegenerative diseases.

1.8 THE AIMS OF THIS RESEARCH

The aims of the research presented in this thesis are as follows:

1) To use the published rat and bovine PPT-A and NK-1R receptor gene sequences to design suitable heterologous DNA probes in order to facilitate the identification of bacteriophage clones of those genes from a murine genomic DNA library.

2) To identify and isolate clones for both the murine PPT-A (substance P) precursor gene and the murine NK-1 receptor (NK-1R) gene from that library.

3) From the identified bacteriophage clones, selected plasmid subclones would be generated. The sequence, and the genomic organisation of the genes contained within that sequence, would subsequently be characterised by restriction enzyme mapping, DNA sequencing and exon hybridisation mapping.

4) The plasmid subclones could also be used as the foundation for the design and generation of gene targeting vectors for both genes for use in homologous recombination experiments. Those vectors would be subsequently used in embryonic stem (ES) cell tissue culture transfection experiments in order to generate transgenic mice with altered, preferably halted, PPT-A or NK-1R gene expression.

Those targeting experiments would be undertaken as a means of defining the functional role of both genes and the peptides they encode in normal neural function. This in turn would provide a basis for the elucidation of the proposed role of the tachykinin neuropeptide signalling systems in neuronal disorders and in particular neurodegenerative disease.

CHAPTER 2

MATERIALS AND METHODS

INTRODUCTION

This chapter contains the general methods and procedures used in the experiments which were the basis of this research. They are detailed below and are sub-divided into four sections: (1) microbiological techniques and standard media, (2) bacterial strains, bacteriophage, plasmids and vectors, (3) manipulation of bacteriophage and (4) isolation and manipulation of nucleic acids.

MICROBIOLOGICAL TECHNIQUES AND STANDARD MEDIA

2.1 *E. COLI* GROWTH CONDITIONS

E. coli strains were grown as liquid cultures in L-broth inoculated with a single colony in the presence of the appropriate antibiotic at 50 μg per ml. *E. coli* strain LE392 was grown without antibiotic present. The volume of L-broth inoculated depended on the quantity of plasmid required. Routinely, 3 ml and 100 ml cultures were used for small (mini) and medium (midi) plasmid preparations respectively. In order to maximise aeration of the culture, the volume of the flask used was typically five times that of the culture. Cultures were incubated at 37°C in an orbital shaker at 250 rpm overnight or for shorter periods depending on the intended use of the cultures. *E. coli* strains were also propagated on petri dish plates containing 25 ml L-agar and the appropriate antibiotic, and were incubated at 37°C overnight.

2.2 STANDARD MEDIA FOR THE PROPAGATION OF *E.COLI*

L-Broth (LB): 10 g tryptone, 5 g NaCl, 5 g yeast extract, 20 mg thymine, made up to 1 litre with distilled water and the pH adjusted to 7.0 with 0.1M NaOH.

L-Agar (LA): As for L-broth without glucose and with agar added to 15 g per litre.

2xYT Medium: 10 g yeast extract, 16 g bactotryptone, 5 g NaCl, made up to 1 litre with distilled water.

S.O.C. Transformation Recovery Medium: 2 g bactotryptone, 0.5 g yeast extract, 1 ml 1M NaCl, 0.25 ml 1M KCl, added to 97 ml of distilled water and autoclaved; supplemented with 1 ml 2M glucose and 1 ml of 1M $MgCl_2$ /1M $MgSO_4$. Filter sterilized.

NZCYM Broth: 5 g NaCl, 5 g bacto-yeast extract, 10 g NZ amine, 1 g casamino acids, 2 g $MgSO_4.7H_2O$, added to 950 ml distilled water, shaken until dissolved then the pH adjusted to 7.0 with 5M NaOH. Sterilised by autoclaving.

NZCYM Agar (top and bottom): As for NZCYM broth except with agar added to 15 g per litre for bottom agar and to 7 g per litre for top agar.

Terrific Broth: 12 g bactotryptone, 24 g bacto-yeast extract, 4 ml glycerol dissolved in 900 ml distilled water. Autoclaved for 15 min. at 15 p.s.i. When cooled to <60°C, add 100 ml of sterile 0.17M KH_2PO_4, 0.72M $KHPO_4$ solution.

2.3 STERILISATION

Glassware and growth media were sterilised by autoclaving at 120°C, 15 p.s.i. for 15 min. Buffers and supplements were heated to

105°C, 15 p.s.i. for 10 min. Heat-labile solutions were sterilised by filtration through Nalgene 0.22 μm pore membranes.

2.4 INTRODUCTION OF DNA INTO E. COLI

Plasmids were introduced into *E. coli* XL1-Blue, TG1, TG2, or nova blue strains by genetic transformation using either rubidium chloride transformation or electroporation. The transformation efficiencies varied according to both the strain and the method used. Typically, efficiencies were in the range of 10^6-10^7 colonies μg^{-1} plasmid DNA.

2.4.1 Rubidium Chloride Transformations

A 2.5 ml overnight culture of recipient *E. coli* was grown in 2xYT media. The culture was diluted 1:100 into L-broth (100 ml typically) and incubated at 37°C, 250 rpm for 2 hrs, until OD_{600} was 0.4. Cells were harvested by centrifugation at 5000xg for 5 min at 4°C, then resuspended in 1 ml ice-cold TFB (see solutions) and placed on ice for 15 min. 37 μl of DMF (100% dimethyl formamide) was then added. After 5 min on ice, 37 μl of dilute ß-mercaptoethanol solution (27 μl in 500 μl dH_2O) was added and a further 10 min incubation on ice followed. Finally, another 37 μl of DMF was added and, after 5 min on ice, the cells were considered competent and ready for use in transformations.

200 μl aliquots of competent cells were added to a maximum of 10 μl of plasmid DNA solution (1-100ng) or ligation mix in a 1.5 ml microfuge tube. The mix was placed on ice for 45 min then heat-shocked at 42°C for 3 min (or 40 sec for nova blue cells). 200 μl of SOC media was immediately added and the culture was incubated at 37°C, 250 rpm for 1hr. One tenth of the culture was plated on L-agar plates containing ampicillin antibiotic and incubated at 37°C overnight.

2.4.2 Transformations by Electroporation

A 2.5 ml overnight culture of recipient *E. coli* culture was grown in 2xYT media. The culture was diluted 1:100 in fresh L-broth and incubated for 3 hrs at 37°C until OD_{600} was 0.6 at which point the culture was chilled on ice for 15 min. The cells were harvested by centrifugation at 4000xg for 15 min at 4°C. Cells were resuspended in 1 litre of ice-cold water and centrifuged as above. Cells were then resuspended in 0.5 litres of ice-cold water and re-centrifuged. Finally cells were resuspended in 3 ml of 10% glycerol to a concentration of $3x10^{10}$ cells per ml. The electro-competent cells can be aliquoted and stored at -70°C for up to 6 months or used immediately. For transformation of competent cells by electroporation a Bio-Rad Gene Pulser™ was used. The apparatus was set at a capacitance of 25 μF and a voltage of 2.5 kV. The 'Pulse Controller Unit' was set at 200 Ω. To 40 μl of competent cells, 1 μl of plasmid DNA or 5 μl of ligation mix was added and incubated on ice for 15 min. The mix was then transferred to a 0.2 cm electroporation cuvette (pre-chilled). One pulse was then applied generating 12.5 kV/cm with a time constant ideally of 4/5 msec. Immediately after transformation, 1 ml of SOC media was added and the mix was incubated at 37°C, 250 rpm for 1 hr. 50 μl is then plated on antibiotic-containing L-agar plates and incubated at 37°C overnight.

2.5 ANTIBIOTICS, INDICATORS AND INDUCERS

Figure 2.1 : Antibiotics, indicators and inducers used .

Supplement	Concentration	Stock concentration
Ampicillin	50 μgml^{-1}	100mgml^{-1} in distilled water.
IPTG	20mM	20mgml^{-1} in distilled water.
Tetracyclin	10 μgml^{-1}	10mgml^{-1} in dilute HCl (1:1000).
X-Gal	80 μgml^{-1}	40mgml^{-1} in DMF.

Figure 2.1: Antibiotics, indicators and inducers used in the research described in this thesis.

The indicator X-Gal (5-bromo-4-chloro-3-indolyl-ß-galacto-sidide), and the inducer IPTG (Isopropyl-β-D-thiogalacto-pyranoside) stocks were stored at -20°C, and were used in conjunction with several plasmid cloning vectors (such as pBluescript from Stratagene) to identify *E. coli* strains carrying recombinant vectors containing DNA sequence inserted into the multiple cloning sites of that vector. Recombinants containing inserts appear white, whereas those lacking inserts appear blue. This selection is based on the principle of gene disruption by the inserted DNA fragment which subsequently permits the synthesis of bacterial colonies which lack the blue colour. Visual screening of plates can therefore permit identification of recombinants. The antibiotics listed above were stored at -20°C.

BACTERIAL STRAINS, BACTERIOPHAGE, PLASMIDS AND VECTORS

2.6 BACTERIAL STRAINS

All bacterial strains are derivatives of *E. coli* K-12.

Figure 2.2 : Bacterial Strains Used.

Strain	Genotype	Ref./Source
XL1-Blue	supE44 hsdR17 recA endA gyrA46 lac recA1 thi-1	Bullock *et al,* 1987
LE392	supE44 supF58 hsdR514 galT22 metB1 galK22 lacY1 trpR55	Sambrook *et al,* 1988
TG2	supEthiΔ(lac-proAB)hsdΔ5 Δ(srl-recA)306::Tn10(tet^r)	Sambrook *et al,* 1988
DH5α	F$^-$ϕ80d lacZ ΔM15 recA1 endA1 gyrA96 thi-1 hsdR17 ($r_K^- m_K^+$) supE44 relA1 deoR Δ(lacZYA-argF) U169	Promega Corp.
TG1	supE thi Δ(lac-proAB) hsdΔ5 F'[traD36proAB+lacLlacZ ΔM15]	Stratagene Inc.
Nova Blue	Δ(mcrBC-mrr)($r_B^- m_B^-$) Δ(gpt-proA)62 leu supE44 ara14 galK2l acY1 mtl-1 rpsL20(Str^r) xly-5 recA13[F$^+$::Tn1000(Tc^s)]	Novagen Ltd.

Figure 2.2: Bacterial strains used in the research presented in this text.

2.7 BACTERIOPHAGE LAMBDA VECTORS

Three different mouse genomic bacteriophage lambda libraries were screened for either positively-hybridising PPT-A or SPR gene clones. The vector was different for each library.

Figure 2.3 : Bacteriophage Vectors and Clones Used.

Library	Vector	Probe	Clones Isolated	Library Ref./Source
Glaxo	λFIX	PPT-A	λPPT1-λPPT5	Courtesy Glaxo plc.
Cambridge	λ2001	PPT-A	λPPT6-λPPT10	Smith, pers. comm.
Toronto	λDASH	SPR	λSPR1-λSPR16	Rossant, pers comm.

Figure 2.3 : A list of the bacteriophage vectors used to generate the libraries, the sources of those libraries and a list of the clones isolated.

2.8 PLASMID VECTORS AND SUBCLONES

Figure 2.4: Plasmid Vectors and Subclones Generated.

Plasmid	Description	Ref./Source
pBluescriptII	3kb pUC19-derived phagemid	Short *et al*, 1988
pBluescriptII	3kb pUC19-derived phagemid cloning and sequencing vector: pKS$^+$ and pSK$^-$ versions used.	Short *et al*, 1988
pGEM® 5Zf+	3kb cloning and sequencing vector with multiple features including blue/white selection.	Promega Corp.
pDK101	pGEM 5Zf+ derived cloning vector designed for PCR product cloning.	Kovalic *et al*, 1991.

(continued)

(continued)

Plasmid	Description	Ref./Source
pSP9	9kb *Nco* I-*Nco* I fragment of λ PPT9 in pGEM5.	Chapter 4
pSP7	7kb *Nco* I-*Nco* I fragment of λ PPT9 in pGEM5.	Chapter 4
pSP10.6	10.6kb *Sst* I-*Sst* I fragment of λPPT9 in pKS+	Chapter 5
pKS+.Bst	pKS+ vector with *Bst* EII site introduced in *Kpn* I site by linker insertion.	Chapter 5
pSP7.2	7.2kb *Bst* EII-*Bst* EII fragment of pSP10.6 in pKS+.Bst.	Chapter 5
pSP7.2 II	As pSP7.2 minus 0.9kb *Mlu* I-*Mlu* I fragment.	Chapter 5
pSP7.2 III	pSP7.2 II with 2.7kb *Mlu* I-*Mlu* I fragment of pSP900V containing *neo* gene in PPT-A exon 3.	Chapter 5
pSP7.2 IV	As pSP7.2 III with 2.7kb *Mlu* I-*Mlu* I fragment of pSP900V in opposite orientation	Chapter 5
pSP900	0.9kb *Mlu* I-*Mlu* I fragment of pSP9 cloned in pSPORT.	Chapter 5
pSP900 II	0.9kb *Mlu* I-*Mlu* I fragment of pSP9 cloned in pIBI24.	Chapter 5
pSP900 III	As pSP900 II with polylinker *Pst* I site lost by *Sal* I-*Xho* I digestion and re-ligation.	Chapter 5

(continued)

(continued)

Plasmid	Description	Ref./Source
pSP900 IV	As pSP900 III with *Eco* RI site introduced in *Pst* I site by linker.	Chapter 5
pSP900 V	As pSP900 IV with 1.8kb *Eco* RI-*Eco* RI *neo* gene fragment in same 5'-3' orientation as PPT-A gene.	Chapter 5
pSP900 VI	As pSP900 V with *neo* in opposite orientation to PPT-A gene.	Chapter 5
pSPR5	5kb *Bam* HI-*Bam* HI fragment of λSPR12 in pKS+.	Chapter 4
pSPR8	8kb *Bam* HI-*Bam* HI fragment of λSPR13 in pKS+.	Chapter 4
pSPR8.5	8.3kb *Sst* I- *Eco* RI fragment of λSPR12 in pKS+.	Chapter 4
pPPTNeo	Blunt-ended 1.8kb *Eco* RI-*Eco* RI *neo* gene cloned into blunt-ended pSP7.2 II *Mlu* I site in same orientation as PPT-A gene.	Chapter 5
pPPTNeo II	As pPPTNeo with *neo* in opposite orientation to PPT-A gene.	Chapter 5
pIBI24	2.9kb pEMBL-derived cloning and sequencing vector.	Dente et al, 1983
prTKR2	1.3kb rat SPR gene cDNA in pSK+.	Yokota et al, 1989
pCOL1	0.5kb mouse genomic PPT-A PCR product in pDK101 vector.	Chapter 3

(continued)

(continued)

Plasmid	Description	Ref./Source
pCOL2	0.5kb rat genomic PPT-A PCR product in pDK101vector.	Chapter 3
pSKNeo	pSK⁻ containing 1.8kb *Eco* RI-*Eco* RI *neo* gene from pPNT.	Donald, 1994 (pers. comm.)
pGEM11.TK	2.9kb *Eco* RI-*Hind* III *thymidine kinase* gene cloned in pGEM11.	Galagher,1994(pers. comm.)
pSPR1.9 I	1.9kb *Bam* HI-*Bst* EII fragment of pSPR8 in pKS+.Bst vector.	Chapter 5
pSPR1.9 II	As pSPR1.9 I but with polylinker *Hind*III site deleted by blunt-ended re-ligation.	Chapter 5
pSPR2.1	2.1kb *Bam* HI-*Hind* III fragment of pSPR8 in pKS+.	Chapter 5
pKS.TK	2.9kb *Kpn* I-*Not* I *thymidine kinase* gene cloned in pKS+.	Chapter 5
pSPORT®1	4kb cloning and sequencing vector containing polylinker *Mlu* I site.	BRL Ltd.

Figure 2.4 (continued): Plasmid vectors used and subclones generated during the research described in this text.

2.8.1 Plasmid Subclone Nomenclature

A system was selected to name the plasmid subclones. For plasmid pSP9, the 'SP' refers to the NK-1 gene (which encodes substance P) from which the subclone was derived and the '9' refers to the size (in Kb) of the plasmid insert. This nomenclature system was extended to the receptor gene NK-1R (formerly the SPR gene) as well as NK-1 plasmid subclones. For example, a 5 kb subcloned fragment

of the NK-1R gene was named pSPR5 where the 'SPR' refers to the old name for the NK-1R gene from which the subclone was derived and the '5' refers to the size of the insert. The "R" distinguishes the receptor from the precursor gene subclones. If two subclones were generated which contained DNA inserts of the same size and were derived from the same gene, they were given the suffix a, b and so on, in lower case. (For example pSPR5a and pSPR5b). This further suffix was not required for the subclones detailed in this text.

MANIPULATION OF BACTERIOPHAGE

2.9 BACTERIOPHAGE PLATING CELLS

The *E. coli* strain LE392 was used in order to propogate bacteriophage lambda in liquid lysates, and for the generation of plated lawns of bacteria on which phage plaques were grown. The cells required for both methods were prepared by inoculating 2.5 ml of L-broth (supplemented with 1 ml of 20% maltose per ml of culture) with a single bacterial colony and incubating overnight at 37°C, 250 rpm without antibiotic. The overnight culture was diluted 1:100 in fresh L-broth and incubated for 3 hrs at 37°C until an OD_{600} of 0.5 was reached. Cells were then centrifuged at 4000 x g for 10 min at room temperature. The pellet was resuspended in 10 mM $MgSO_4$ to an OD_{600} of 2.0 and stored at 4°C. Plating cells were used for several weeks, but for optimum plating efficiencies, fresh cells were prepared on each occasion.

2.10 GROWTH OF BACTERIOPHAGE LIQUID LYSATES

Between 5 and 15 µl of high titre (approximately 10^{10} p.f.u. μl^{-1}) bacteriophage stock was added to 100 µl of LE392 plating cells, and incubated at 37°C, 250 rpm for 20 min to allow the phage to adsorb to the bacteria. 50 ml of pre-warmed (37°C) NZCYM broth was then added and the culture was incubated overnight in a 37°C shaker.

Successful bacterial lysis and a high phage content is indicated by a clear culture containing clumps of aggregated lysed bacterial debris. Phage DNA can then be extracted from this lysate.

2.11 PLATING BACTERIOPHAGE ON NZCYM AGAR PLATES

In order to screen a bacteriophage genomic library, the phage have to be plated on NZCYM agar plates at a concentration which ensures that distinct plaques are visible. (Plaque sizes vary depending on the vector used.) To purify a particular identified phage, the phage are again plated on NZCYM agar plates but at a lower density to allow a single plaque to be removed without contamination by other neighbouring phage.

Bacteriohage were added to LE392 plating cells and incubated at 37°C for 20 min to allow the phage to adsorb to the bacteria. The volume added depended on the titre of the phage stock and the number of plaques required on the plate. For small 82mm diameter round plates, 100 p.f.u were added. The cell/phage mix was then added to liquid NZCYM top agar (cooled to 45°C) and then poured onto an NZCYM bottom agar plate (previously poured on a level surface and dried). Volumes were also dependent on the size of plate used, as described below.

Figure 2.5 : Plating Cells and Agar Volumes.

	Plate Size		
	82mm (Diam.)	10mm x 10mm	245mm x 245mm
Plating cells (ml)	0.2	0.5	2.0
Top Agar (ml)	3.0	7.0	30.0

Figure 2.5 : Volumes of plating cells and top agar used per NZCYM plate.

Once the plates have set on a level surface they are inverted (to minimise moisture smearing the plaques), and incubated at 37°C overnight. Before taking nylon membrane lifts from the plates (for

hybridisation purposes), the plates were cooled to 4°C to prevent the top agar layer adhering to the membrane and preventing further duplicate lifts being taken.

2.12 TITRE OF BACTERIOPHAGE

To titre a particular lambda bacteriophage stock, serial dilutions of the stock were made by diluting 100 µl in 900 µl of SM phage buffer to produce a 10^{-1} dilution. That solution was similarly diluted to produce a 10^{-2} dilution and so on in sequence to at least a 10^{-6} dilution. 10 µl of each dilution was then plated on NZCYM (82mm diam.) plates as described above and incubated at 37°C overnight. The following day, the number of plaques on each plate was counted and the titre calculated in plaque forming units (pfu). For example, 500 plaques counted on the 10^{-6} dilution plate indicates a phage stock titre of $5x10^{10}$ $pfu.ml^{-1}$.

2.13 STORAGE OF BACTERIOPHAGE

Bacteriophage lysates were stored at 4°C in the presence of a drop of chloroform (50µl), which was added to prevent the growth of any surviving bacteria. For long-term storage of bacteriophage, DMSO was added to a 7% concentration (v/v), mixed gently and placed in dry ice or liquid nitrogen for 20 min before being transferred to a -70°C freezer for storage. Phage stored under these conditions will remain viable indefinitely.

2.14 TRANSFER OF DNA TO HYBOND-N™ MEMBRANES

To screen lambda phage plaques for probe homology, duplicate lifts were taken from the phage plates whereby the DNAs to be screened were transferred to a filter membrane such as Amersham Hybond-N™ which was subsequently hybridised to probes in aqueous solution. The membrane retains the spatial orientation of the plaques relative

to one another so that positively hybridised plaques can be located and isolated later.

Lambda phage NZCYM plates grown overnight were chilled at 4°C for 1 hr to prevent the plaque-covered top agar layer sticking to the membrane (leaving only the bottom agar layer when subsequent lifts are made). A piece of membrane was cut which was slightly smaller than the dimensions of the plate. The membrane was then placed on the plate for 30 sec during which time holes were made in three asymmetrical locations using a needle to pierce the membrane and the NZCYM layers. (Marking the holes in the media with India Ink after the first membrane lift allows the piercing of holes in subsequent filters to be localised to the same relative position as for the first lift.) After 30 sec the membrane was gently removed and placed plaque-side up on 3MM paper soaked in denaturing solution for 7 min. The membrane was then transferred to 3MM paper soaked in neutralising solution for 7 min. The filters were washed twice by submersion in 2X SSC buffer for 7 min then laid flat to air-dry. The DNA was then permanently fixed to the membrane by ultra violet cross-linking (using the auto-cross-link function on a Stratagene 'Stratalinker™' apparatus) or by baking at 80°C for 1 hr minimum. The membrane which was used to make the second duplicate lift from the phage plate was placed on the plate for 1 min rather than 30 sec. The third and fourth lifts were placed on the plate for 2 and 3 min respectively then fixed as described above.

2.15 ISOLATION OF DNA FROM BACTERIOPHAGE LYSATES

200 ml of bacteriophage lysate was centrifuged at 10,000 x g for 10 min to pellet any remaining bacterial debris. The cleared lysate was then incubated at 37°C with 100 $\mu g ml^{-1}$ DNAase and 100 $\mu g ml^{-1}$ RNAase for 1 hr to degrade bacterial DNA and RNA. The phage were precipitated by adding 50 ml of 30% PEG (polyethylene glycol) and incubated on ice for 1 hr. Centrifugation at 12,000xg for 10 min pelleted the phage which were subsequently resuspended in 10 ml SM

buffer. (This concentrated, cleared phage stock can be stored at 4°C for several months.) 0.5 ml of this stock was incubated with EDTA (to 20 mM) and proteinase K (to 50 $\mu g ml^{-1}$) at 37°C for 1 hr to break open the phage heads and release the phage DNA. The volume was increased to 1 ml with dH_2O then phenol and chloroform extracted. The DNA was ethanol precipitated but not in the presence of one tenth volume of 3M sodium acetate as is standard practice, as salt carry-over can sometimes hinder digestion of the DNA. The DNA was resuspended in 100 µl of dH_2O. For subsequent restriction enzyme digestions, 3 µl of this DNA was used per reaction.

ISOLATION AND MANIPULATION OF NUCLEIC ACIDS

2.16 GENOMIC DNA EXTRACTION FROM MAMMALIAN TISSUE

Genomic DNA was prepared from frozen mouse liver tissue previously dissected from fresh cadavers and stored at -70°C. Approximately 1 g of starting tissue was placed in a mortar and kept immersed in liquid nitrogen while ground to a fine powder with a pestle. The powder was then dispersed in 10 ml of extraction buffer (see solutions section) and incubated at 37°C for 1 hr. Proteinase K enzyme was then added to a concentration of 100 $\mu g.ml^{-1}$, and incubated at 37°C for 3 hrs or overnight. An equal volume of phenol was then added and the two phases were mixed by slow and gentle end-over-end inversion for 10 min followed by 1 hr on a rolling shaker. The mix was then centrifuged at 5000xg for 15 min at room temperature. The upper aqueous phase was then gently removed with a cut-tip pipette (so as not to physically shear the DNA) and phenol extracted again twice. Extraction with an equal volume of chloroform was followed by precipitation of the DNA by addition of 0.2 volumes of 10M ammonium acetate and 2 volumes of 100% ethanol. On inversion, the DNA visibly precipitated as gelatinous strands which were removed using a hook-shaped pasteur pipette with its end sealed using a bunsen burner. The DNA was then

centrifuged at 10,000xg for 15 min and washed twice in 70% ethanol. The DNA pellet was then briefly air-dryed (over-drying limits resuspension) and resuspended in 1 ml of distilled water by rotation on a rolling shaker overnight. The OD_{260} and OD_{280} are measured so that the DNA concentration and degree of protein contamination of the resuspension can be determined using a simple equation (Sambrook *et al*, 1988). The DNA was then ready for restriction enzyme digestion and other manipulations.

2.17 ISOLATION OF DNA FROM *E. COLI* CULTURES

Several methods were used for the isolation of plasmid DNA from *E coli* cultures depending on the quality and quantity of DNA required. The methods used are listed below.

2.17.1 Alkaline Lysis Method

The alkaline lysis method (Birnboim and Doly, 1979) was the used during the early stages of this project to isolate plasmid from small (1.5ml) or large (50-200ml) culture volumes of *E. coli.*

Reagents: Birnboim Doly I (BDI); 50 mM glucose, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA. Immediately prior to use, lysozyme was added to a working concentration of 2 $mg.ml^{-1}$. Birnboim Doly II (BDII); 0.2 M NaOH, 1% (w/v) SDS, stored in a non-glass container. Birnboim Doly III (BDIII); 5 M KOAc (pH 4.8). Phenol: redistilled phenol buffered with 0.5 M Tris-HCl (pH 8.0), 0.1% (w/v) 8-hydroxyquinoline. DNAase-free RNAase A: Pancreatic RNAase dissolved in dH_2O to a concentration of 10 $mgml^{-1}$, heated to 100°C for 15 min and cooled to room temperature. RNAase was stored in aliquots at -20°C.

2.17.1.1 Small Scale Plasmid Prep

Plasmid DNA was isolated from 1.5 ml of overnight *E. coli* culture. Cells were pelleted by centrifugation in a microfuge tube at 14,000xg for 30 sec, and resuspended in 100 µl of BDIII (with lysozyme) by pipetting. Incubation at 37°C for 15 min was followed by addition of 200 µl BDIII and end-over-end inversion for 1 min. 150 µl of pre-chilled BDIII was then added and the mix placed on ice for 5 min. Most of the cell debris and chromosomal material was pelled by centrifugation at 14,000xg, 4°C for 10 min. The supernatant was then extracted with an equal volume of phenol then similarly with chloroform. The DNA was then precipitated with 0.1 volume of sodium acetate (pH 5.2) and 2 volumes of 100% ethanol and pelleted by centrifugation at 14,000xg for 30 min. The pellet was washed twice in 70% ethanol and air-dried for 5min before resuspension in 50µl dH_2O. Typically, 20 µg of plasmid was isolated per 1.5 ml of starting culture of a quality suitable for most *in vitro* procedures.

2.17.1.2 Large Scale Plasmid Prep

200 ml of overnight stationary phase culture was pelleted in a 250 ml polypropylene tube by centrifugation at 8000xg for 10 min. Cells were resuspended using a cut-tip pipette in 7.5 ml of BDI solution and incubated at room temperature to allow the lysozyme to break open the bacteria. The cell suspension was then placed in a 30 ml polypropylene tube containing 7.5 ml of BDII and placed on ice for 5 min and occasionally inverted. 7.5 ml of BDIII was then added and another 5 min incubation on ice follows. The bulk of the bacterial and chromosomal debris was then pelleted at 15,000xg for 25 min at 4°C. The 22.5 ml of supernatant was then split into two clean 30 ml tubes containing 19 ml of 100% isopropanol and after a 15 min incubation at room temperature centrifuged at 12,000 x g for 20 min. The DNA was further purified by equilibrium-density centrifugation on a caesium chloride-ethidium bromide (CsCl.EtBr) gradient. The nucleic acid pellet was resuspended in 3.9 ml of TE and added to 7.4 g of CsCl dissolved in 4 ml of TE to a density of 1.58 gml^{-1}. 300 µl of ethidium bromide (10 $mgml^{-1}$) was added,

mixed and the solution was left in the dark for 1 hr. The solution was then ultra-centrifuged at 289,000xg for 16 hrs at 20°C in a Beckman Ti70 angled rotor/L8 Beckman ultra-centrifuge. The solution was then illuminated in the dark with a hand-held ultra-violet lamp to permit visualisation of two distinct DNA bands. The upper band constitutes chromosomal and relaxed form plasmid DNA, whereas the lower band constitutes supercoiled plasmid. The lower band was removed by puncturing the tube below the level of the band with a 1 ml syringe and gently withdrawing the fluid. The ethidium bromide was then removed by repeated extraction with water-saturated N-butanol (butan-1-ol). Three volumes of dH_2O was then added prior to DNA precipitation with 6 volumes of 100% ethanol. The DNA was pelleted at 12,000xg for 30 min at 4°C, washed twice with 70% ethanol, and resuspended in 1 ml of dH_2O. The DNA is suitable for most *in vitro* manipulations.

2.17.2 Quick Plasmid Preps

Technological advances during this project provided faster methods for the isolation of plasmid DNA which dispensed with the use of toxic reagents such as phenol. These methods are also based on the alkaline lysis methods (Birnboim and Doly, 1979). Specifically, Promega 'Magic Minipreps™' replaced the small-scale protocol, and Qiagen 'Midi-Columns™' were used in place of the time-consuming large-scale protocol detailed above.

2.17.2.1 Promega Magic Minipreps ™

1 ml of culture was pelleted at 12,000xg for 30 sec. Cells were resuspended in 200 µl of BDI (RNAase added to 100 $\mu g ml^{-1}$), then 200 µl of BDII was added and the suspension left at room temperature for 5 min. 200 µl of BDIII was added and the tube was inverted before centrifugation at 12,000xg for 5 min. The supernatant was mixed with 1 ml of Magic Miniprep Resin™ which specifically binds plasmid but not chromosomal DNA due to the pH-specific binding properties of the resin. Using a 2.5 ml syringe, the mix was pushed through a Promega mini-column which retains the

resin-plasmid complex but not the other unwanted mix components. The column was washed with 2 ml of wash buffer (see solutions) and pulse centrifuged for 20 sec in a 1.5 ml microfuge tube. 50 µl of dH_2O was added and the column pulse centrifuged to elute the DNA while leaving the resin particles in the column. The typical yield of 10 µg of DNA has minimal RNA contamination and is suitable for restriction but the quality can occasionally limit its use for other *in vitro* manipulations.

2.17.2.2 Qiagen™ Midi-Columns

This protocol is designed for isolation of as much as 100 µg of plasmid from 30-150 ml of culture. 100ml of culture was pelleted by centrifugation at 12,000xg for 5 min and resuspended in 4 ml of P1 buffer (for P-buffers, see solutions) containing RNAase (100 $\mu g.ml^{-1}$). 4 ml of P2 buffer was added and mixed gently then incubated at room temperature for 5 min. The same volume of P3 buffer was added next and the solution mixed then centrifuged at 15,000xg for 30 min at 4°C. A second spin ensures removal of particulate material. The supernatant was then applied to a Qiagen Tip-100™ (previously equilibrated with 3 ml of QBT buffer; for Q-buffers, see Solutions). The column functions by the same principle as the Promega mini-columns whereby the plasmid DNA is temporarily bound to the Qiagen column in a pH-specific manner. When all of the supernatant had entered the column, it was washed with 10 ml of QC buffer. The plasmid DNA was then eluted with 5 ml of QF buffer, and precipitated with 0.7 volumes of isopropanol. The precipitate was then pelleted at 15,000xg for 30 min at room temperature. After two 70% ethanol washes, the pellet was air-dried for 5 min and resuspended in 1 ml of dH_2O.

2.18 QUANTIFICATION OF NUCLEIC ACIDS

2.18.1 Double Stranded DNA

The nucleic acid concentration and the degree of protein contamination was determined by spectrophotometry using a Milton Roy Spectronic 1201 spectrophotometer. Measurements at OD_{260} and OD_{280} were taken. An $OD_{260}=1$ for double stranded DNA indicates a concentration of 50 $\mu g\mu l^{-1}$. A more precise, empirical value can be determined by multiplying the OD_{260} value by 33 (multiplication factor for double stranded DNA), and by the dilution factor (of the solution measured), to give a DNA concentration in $\mu g\mu l^{-1}$. For a given DNA solution, the value obtained by dividing the OD_{260} reading by the OD_{280} reading indicates the extent of protein contamination. A value of >1.75 indicates protein contamination is negligible, where a value of 1.8 indicates a pure DNA solution. The presence of phenol can also lower these values (Sambrook *et al*, 1988).

2.18.2 Single Stranded Oligodeoxynucleotides

The DNA concentration of single-stranded oligodeoxynucleotides used as hybridisation probes, or to make DNA linkers, was determined by measurement of OD_{260}. This value was multiplied by 33 (multiplication factor for single-stranded DNA) and divided by the volume in µl of stock solution diluted for measurement. The DNA concentration calculated is in $\mu g\mu l^{-1}$. Alternatively, the concentration in $pmolesml^{-1}$ can be calculated by dividing the OD_{260} reading by [the number of nucleotides in the deoxyoligonucleotide multiplied by 0.01].

2.19 PRECIPITATION OF NUCLEIC ACIDS

2.19.1 Double Stranded DNA

Double stranded DNA was precipitated by addition of 1/10 volume of 3M sodium acetate and 2 volumes of 100% ethanol followed by a 30 min (minimum) incubation on ice. Alternatively DNA in water or TE was precipitated by addition of 0.7 volumes of isopropanol, and incubated at room temperature for 10 min (minimum). The DNA was pelleted by centrifugation at 12,000xg for 15 min, then washed twice with 70% ethanol, air dried for 5 min and resuspended in appropriate buffer/dH_2O.

2.19.2 Single Stranded Oligodeoxynucleotides

Single stranded oligodeoxynucleotides were precipitated with 1/10 volume of 5.2M NaAc and 2 volumes of 100% ethanol, incubated at -20°C overnight or at -70°C for 1 hr. Precipitation was followed by centrifugation at 12,000xg for 30 min at 4°C, or at room temperature in the case of isopropanol precipitations. The DNA pellet was then washed twice with 70% ethanol and air-dried for 5 min before being resuspended in the appropriate buffer/dH_2O.

2.20 DNA DIGESTION WITH RESTRICTION ENDONUCLEASES

Restriction enzymes were supplied by Bethesda Research Laboratories (BRL), Promega Corp. or New England Biolabs Ltd. Digestion of plasmid or phage DNA was typically carried out in a 10µl volume, constituting 1 µl DNA (0.2-2.0 µg), 1 µl of appropriate 10X (BRL REact™ for example) restriction enzyme buffer, 1µl of restriction enzyme (1-10 unitsμg^{-1} DNA), and dH_2O to a final volume of 10 µl. For double or triple restriction enzyme digests where the buffers were incompatible, the Pharmacia One-Phor-All™ buffer system was used. The reaction was incubated at 37°C for 1-2 hrs. Digestion of genomic DNA required longer incubation of up to 5 hrs.

The reaction was halted either by addition of gel loading buffer, heat-inactivation appropriate to the enzyme being used, or by phenol/chloroform extraction if subsequent digestion with a different enzyme/buffer was to follow. Digestion of larger amounts of DNA required each reaction component to be scaled up accordingly. Gel electrophoresis of samples was used to determine when digestion was complete in such cases.

2.21 LIGATION OF RESTRICTED DNA FRAGMENTS

Two methods were used for ligation of DNA fragments. Firstly, conventional ligations requiring the addition of all reaction components, and secondly, the more convenient Pharmacia Ready-To-Go™ ligation kit.

2.21.1 Conventional Ligations

Ligation of selected DNA fragments into plasmid vectors used a reaction volume of 20 µl with a total DNA concentration of 6 $mgml^{-1}$. The molar ratio of vector fragment to insert fragment ranged from 1:2 to 1:10 depending on the fragment sizes and the state of de/phosphorylation of the fragments. Vector and insert DNAs were mixed with 1/10 volume of 10X ligation buffer (Pharmacia), ATP to a final concentration of 1 mM, and 1 unit of T4 DNA ligase/μg^{-1} DNA. The volume was made up to 20 µl with dH_2O, and the reaction incubated at 16°C overnight or room temperature for 4 hrs.

2.21.2 Ready-To-Go™ Ligation Kit

This protocol simply required the dilution of insert and vector DNA fragments in dH_2O to a final volume of 20 µl. The solution was added to a Ready-To-Go tube (Pharmacia) which contained dessicated ligation mix (ligase, buffer and ATP), and left at room temperature to allow re-hydration of the ligation mix for 5 min. After gentle pipetting to ensure complete resuspension, and brief

centrifugation to remove any bubbles, the tube was placed at 16°C overnight. Ligation efficiencies were found to be improved considerably by the addition of 2 µl of 1unitμl^{-1} T4 DNA ligase enzyme.

2.22 DEPHOSPHORYLATION OF RESTRICTED DNA FRAGMENTS

Removal of the terminal 5' phosphate group from restriction enzyme digested double-stranded DNA fragments prevents ligation of two otherwise compatible fragments. This reduces re-ligation of vector DNA which, under ligation conditions, competes with the preferred insert/vector ligation reaction. Vector re-ligation can reduce the number of successful vector recombinants containing insert. The dephosphorylation reaction was performed by Calf Intestinal Alkaline phosphatase (CIP), or Shrimp Alkaline Phosphatase (SAP). The former was used originally but phenol/chloroform extraction was required to remove it completely. This enzyme was superseded by SAP because SAP is irreversibly heat-inactivated (in Tris buffers, pH 8.0-8.5) by incubation at 65°C for 15 min.

2.22.1 Calf Intestinal Alkaline Phosphatase (CIP)

CIP enzyme is supplied by Boehringer Mannheim Ltd. Five pmoles of restriction enzyme digested DNA with protruding phosphorylated 5' ends (approximately 7 µg of 5kb DNA fragment), can be dephosphorylated by 0.1units of CIP. The DNA and the enzyme are incubated in 1/10 volume of 10X CIP buffer at 37°C for 30 min. Phenol/chloroform extraction is followed by precipitation and resuspension in dH_2O.

2.22.2 Shrimp Alkaline Phosphatase (SAP)

SAP enzyme is supplied by Amersham USB. To dephosphorylate 1 pmole of DNA termini (2.5 µg of 3 Kb fragment), different amounts of

SAP are required depending on the kind of termini. The minimum amounts of enzyme required are listed below.

Figure 2.6 : Units of SAP used per Reaction.

Terminus	Units of Phosphatase (SAP)
5'-Protruding	0.1 unit
Blunt	0.2 unit
5'-Recessed	0.5 unit

Figure 2.6 : Units of SAP used in dephosphorylation reactions.

The DNA and SAP enzyme are incubated in 1/10 volume of 10X SAP buffer at 37°C for 30 min then heat-inactivated. The dephosphorylated DNA is then ready for ligation.

2.23 ORGANIC SOLVENT EXTRACTION

Proteins were removed from nucleic acid solutions by phenol, phenol/chloroform, and chloroform extractions. Chloroform was supplied by Prolabo Ltd., and phenol by Aldrich Chemicals Ltd. Prior to extraction, the volume of the nucleic acid solution was sometimes increased with dH_2O to limit the amount of nucleic acid inevitably lost during extraction. Re-distilled phenol was buffered with 500 mM Tris.HCl (pH 8.0) and contained 0.1% (w/v) 8-hydroxyquinoline. Phenol/chloroform constituted 50 parts phenol, 49 parts chloroform and 1 part isoamyl alcohol.

Typically, an equal volume of phenol was added to the nucleic acid solution and mixed using a vortex mixer. The organic and aqueous phases were separated by centrifugation at 12,000xg for 5 min. This could be repeated several times until no protein precipitate was visible at the interface. The upper aqueous phase was removed and extracted with an equal volume of chloroform in the same way. The upper aqueous phase was again removed and precipitated with

either ethanol or isopropanol. Precipitation also removed any lingering solvent in the nucleic acid soluton.

2.24 PREPARATION OF RADIOACTIVELY-LABELLED DNA PROBES

Probes were labelled with radioactive nucleotides by either random priming using Klenow enzyme, resulting in incorporation of [α^{32}P]-labelled nucleotides into a newly synthesized complementary DNA strand, or by 5'-end labelling using T4 polynucleotide kinase, where a [γ^{32}P] ATP molecule donates a radioactive phosphate group which replaces the existing phosphate at the 5'-terminal ends of a DNA strand. Modifying enzymes were supplied by Promega Corp.

2.24.1 Random Primer Extension DNA Labelling Method

This method is based on the procedure of Feinberg and Vogelstein (1983 and 1984). A Boehringer Mannheim random priming kit was used to set up the following reaction.

- 25-50 ng of DNA fragment in a maximum volume of 10 µl was denatured at 95°C for 10 min then cooled immediately on ice.
- 5 µl (50 µCi) of [α^{32}P] dCTP (3000 Ci.mmol^{-1}) was added.
- 2 µl of reaction mixture containing random hexanucleotides and 10X reaction buffer was added.
- 1 µl of each unlabelled nucleotide, (in this case dATP, dGTP, and dTTP), from a 0.5 mM stock was added.
- 1 µl of Klenow enzyme from a 2 unitsμl^{-1} stock was added.
- dH_2O was added to a total volume of 20 µl.
- The reaction was then incubated at 37°C for 30 min then halted by incubation at 65°C for 10 min. The unincorporated dNTPs were removed, by Sephadex-G50 column chromatography.

2.24.2 Sephadex G-50 Column Chromatography

To the random primer extension labelling reaction, 5 µl of 50 mg.ml^{-1} dextran blue dye and 5 µl of 50 mgml^{-1} phenol red dye was added. The sample was then loaded on to a Sephadex-G50 (Pharmacia) column, 4 cm long and 0.5 cm in diameter, pre-equilibrated with column buffer (10 mM Tris.HCl, 100 mM NaCl, 1 mM EDTA, pH 7.5), or a Pharmacia Nick-Column™. The labelled DNA and the template DNA was co-eluted with the dextran blue dye, whereas the unincorporated nucleotides were retained. The labelled DNA sample was then denatured by incubation at 95°C for 10 min then chilled on ice immediately.

2.24.3 5'-End Labeling of Oligodeoxynucleotides

Oligodeoxynucleotide probes were radioactively-labeled with [γ^{32}P] using T4 polynucleotide kinase. The labelling reaction comprised:

- 50 ng of oligodeoxynucleotide.
- 1 µl of 10X kinase buffer.
- 50 µCi (5 µl) of [γ^{32}P] ATP (4500 Ci.mmol^{-1}).
- 1 µl (10 units) T4 polynucleotide kinase.
- dH_2O to 10 µl.
- The reaction was incubated at 37°C for 30 min.

The probe was used in this form and not treated by chromatography to remove unincorporated radioactively-labeled nucleotides.

2.25 AGAROSE GEL ELECTROPHORESIS

Horizontal neutral agarose gels were used to visualise and separate DNA. Typically, analytical gels were 0.5% (w/v) agarose in 1X TBE. Higher percentage gels were required to visualise smaller DNA bands such as PCR products. When recovering DNA from agarose gels 1x TAE buffer was used instead of 1x TBE.

2.25.1 Mini Gel Electrophoresis

Pharmacia GNA-100 gel kits were used for rapid visualisation of DNA restriction digests and to determine the extent of DNA recovery after precipitation. 1.5 g of standard grade agarose was added to 300 ml of 1X TBE (or TAE), and heated to boiling and cooled to 60°C. To this gel stock EtBr was added to a concentration of 200 $ngml^{-1}$. 30 ml of 60°C agarose was then poured into a 105 x 75 mm gel caster with an 11 well comb and allowed to set. When set the gel was placed in the gel tank containing 0.5 litres of 1X TBE. Mini gels were operated at 2-10 Vcm^{-1}, typically 75 V for 45 min or longer depending on the band sizes and the resolution required. DNA bands were visualised on a 302 nm UVP Inc. 'Chromato-Vue' transilluminator

2.25.2 Large Gel Electrophoresis

Large gels were required to resolve large DNA bands, to accurately size bands, and to electrophorese samples for Southern analysis. A BRL Ltd. 'Horizon 11.14' gel kit was used where 100 ml of agarose was poured into a 140 x 105 mm gel caster with a 14 well comb and allowed to set. Genomic and lambda DNA digest samples were heated to 65°C for 5 min prior to loading. Gels were run for 6 hrs at 40 V, or for genomic digests, overnight at a lower voltage in a lower percentage gel.

2.26 GEL PHOTOGRAPHY AND IMAGING

Gels containing EtBr were visualised on a transilluminator and either photographed (with Kodak Wratten filters) using Polaroid type 67 film or 'imaged' with a Mitsubishi video copy processor printed on Mitsubishi high density thermal imaging paper for video printers.

2.27 RECOVERY OF DNA FROM AGAROSE GELS

Restricton enzyme digested DNA fragments were recovered from agarose gels using Isogene™, Costar Spin-X™ columns, NA45 paper or by low melting point agarose organic extraction. The Isogene™ kit method was used for purification of PCR products. Costar Spin-X columns were originally used for the recovery of larger DNA bands for cloning purposes. However the recovery yield was found to sometimes be relatively low (as experience by this reseacher) and this method was superseded by the NA45 paper recovery method. This in turn was replaced by the low melting point agarose organic extraction method, as this method was found to give the highest yield of good quality DNA for use in ligations.

2.27.1 Isogene™ Method

The Isogene™ DNA purification and extraction kit was supplied by Perkin Elmer Cetus Ltd. DNA was bound to a finely particulate slurry in the presence of sodium iodide. These conditions favour the solubilisation of agarose and proteins and this facilitates extraction of DNA. Bound DNA is eluted with dH_2O. A weighed gel slice was placed in a 1.5 ml microfuge tube containing 2 volumes of sodium iodide solution and incubated at 60°C for 10 min to solubilise the agarose. The tube was then vortexed briefly and incubated for another 5 min at 60°C. An appropriate volume of binding slurry (dependent on the volume of the gel slice, but approximately 100 µl in 1 ml of solubilised agarose solution) was added and mixed end-over-end at room temperature for 10 min. The supernatant was removed after a 10 sec centrifugation at 12000 x g. The pellet was resuspended and briefly centrifuged at 12000 x g for 10 sec three times in wash buffer (1 M Tris.HCl [pH 7.5], 1 M NaCl, 100 mM EDTA). After the final brief centrifugation, the pellet was resuspended in 20 µl dH_2O, pulse centrifuged again and the supernatant removed and retained. The elution step was repeated generating another 20 µl of DNA eluate which was pooled with the first. The sample was checked by electrophoresis then used.

2.27.2 Spin-X™ Columns

The columns were supplied by Costar Ltd. The gel slice was placed in the column and incubated at -20°C for 10 min. The column was then centrifuged at 10,000 x g for 10 min. The top section of the column containing the residual agarose was removed leaving the lower section which contains the eluted DNA solution. The eluate can then be used immediately for ligation if it is concentrated enough, or ethanol precipitated to concentrate the DNA and remove salts. The DNA was resuspended in dH_2O.

2.27.3 NA45 Paper Method

NA45 paper was purchased from Schleicher and Schuell Inc. This paper is a 0.45 µm DEAE membrane used for filtration and separation. The method is based on the temporary adhesion of DNA fragments to the paper at low salt concentrations and elution of the DNA under high salt/high temperature conditions. Firstly, the paper itself has to be activated by washing in 10 mM EDTA for 10 min and in 0.5 M NaOH for 5 min. The paper was then washed repeatedly in dH_2O and stored at 4°C until required (up to several months). To isolate a particular DNA band, a linear incision was made in the gel in front of the band (with respect to direction of migration), and a piece of activated paper inserted. The gel was run for a further 10 min at 75 V. The paper (with the DNA on the surface) was then removed and washed three times in 500 µl of low salt NET buffer. 200 µl of high salt NET was next added and the paper/buffer incubated at 68°C for 15 min. The eluate was removed and another 200 µl of high salt NET was added and the paper re-incubated for 30 min. The eluates were then pooled, ethanol precipitated and the DNA resuspended in dH_2O.

2.27.4 Organic Extraction from L.M.P. Agarose Gels

A weighed slice from a low melting point (L.M.P) gel was placed in a 1.5 ml microfuge tube containing 3 volumes of buffer (1 mM EDTA

[pH 8.0], 20 mM Tris.HCl [pH 8.0]). The solution was incubated at 65°C for 5 min, vortexed and cooled slowly to room temperature. Phenol and chloroform extractions followed to remove the agarose. The upper aqueous phase was ethanol precipitated and resuspended in dH_2O.

2.28 SOUTHERN TRANSFER ANALYSIS

Southern transfer analysis methodology is based on the procedure of Southern (1975). This technique was used to probe genomic DNA digests to optimise hybridisation conditions prior to screening mouse genomic lambda libraries. It was also used to probe restriction enzyme digested lambda clones with oligodeoxynucleotides to locate exons.

Once the DNA had been resolved by electrophoresis and photographed alongside a ruler (so that positive bands could be identified later), the DNA was transferred by capillary transfer to Hybond-N™ membrane under alkaline conditions. One corner of the gel was trimmed for orientation purposes. The gel was then immersed in 0.25 M HCl for 15 min, in denaturing solution (see solutions) for 45 min, and neutralising solution (see solutions) for 30 min. The gel was equilibrated for 5 min in 10X SSC. The transfer of the DNA was initiated by placing the gel on a 3MM paper wick (supported by an inverted gel former) with both ends in a tank of 10X SSC placed underneath. Above the gel, a piece of Hybond-N™ membrane (with a corresponding corner trimmed) was placed. Buffer was drawn through the wick, the gel, and the membrane by placing a few centimetres of absorbent paper towels and a weight on top. After blotting overnight, the membrane was rinsed in 2X SSC to remove any agarose and baked at 80°C for a minimum of 1 hr. It was then ready for probing.

2.29 AQUEOUS HYBRIDISATION CONDITIONS

2.29.1 Hybridisation Using Random Primed Probes

Prior to addition of labelled homologous probe DNA, the filters to be probed were pre-hybridised in hybridisation buffer using a Techne™ hybridisation oven. The buffer was filtered through a 0.4 µm Nalgene filter before the addition of salmon sperm DNA. Typically 1 ml of buffer is added per cm^2 of membrane (usually 40 ml). The filters were pre-hybridised at 60°C for 3 hrs in a Techne™ hybridisation flask. Labelled, denatured probe DNA was added to 7 ml of hybridisation buffer and added to the pre-hybridised filters and incubated at 60°C for 16 hrs, or overnight.

After hybridisation, the filters were washed in 0.5 litres of 1X SSC, 0.1% (w/v) SDS at 60°C for 30 min. The washing stringency was then increased by washing the filters in 0.2X SSC, 0.1% (w/v) SDS at 60°C for 30 min. The filters were then heat-sealed individually in plastic bags and subjected to autoradiography. Should the filters need to be re-probed, the filters are stripped (if they have been kept moist since hybridisation) by the addition of 0.5 litres of boiling 0.1% SDS solution and placing the submerged filters on a shaking platform until the solution cools to room temperature.

2.29.2 Hybridisation Using Oligodeoxynucleotide Probes

Filters were pre-hybridised in 20 ml of oligodeoxynucleotide hybridisation buffer at 45°C typically for 1 hr. The labelled oligodeoxynucleotide probe was added to 7 ml of hybridisation buffer and placed in a Techne™ hybridisation flask at 45°C for a minimum of 4 hrs. The filters were then washed once in 4X SET (see solutions), 0.1% SDS for 5 min at a temperature 5°C below the temperature at which the hybridisation had been performed, then secondly washed in the same buffer for 5 min at the same

temperature at which the hybridisation had been performed. Filters were then sealed and subjected to autoradiography as above. Oligodeoxynucleotide-probed filters are stripped by addition of 0.1% SDS solution, heated to 80°C and allowed to cool to room temperature on a shaking table.

2.30 MANUAL DNA SEQUENCING

Manual DNA sequencing was carried out by the dideoxy termination method (Sanger *et al*, 1975), using double-stranded template. A Sequenase™ kit supplied by United States Biochemical Corp. (USB) was used. Template was prepared with 'Magic Minipreps™'. Buffers and mixes used below supplied by USB Corp.

Annealing: 1 µl of 1 M NaOH and 1 µl of 10 ngμl^{-1} primer was added to 5 µl (3/4 µg) of template DNA and incubated at 37°C for 10 min to denature the template. 1 µl of 1 M HCl and 2 µl of 5X sequenase buffer was added (to neutralise the reaction) and incubated at 37°C for 10 min. The reaction was then chilled on ice for up to 2 hrs.

Labelling: To the chilled reaction, 1 µl of 0.1 M DTT, 2 µl of dilute labelling mix (diluted 1:5 in dH_2O), 0.5 µl of [α^{35}S]-dATP (1000 Ci/mmol), and 2 µl of dilute Sequenase enzyme (diluted 1:8 in enzyme dilution buffer) was added. The reaction was then incubated at room temperature for 2-5 min.

Termination: 3.5 µl aliquots of the labelling reaction were added to 2.5 µl aliquots of the four pre-warmed dideoxy nucleotide termination mixes (ddGTP, ddATP, ddTTP, and ddCTP at 37°C). These reactions were incubated at 37°C for exactly 5 min then halted by the addition of 5 µl Stop Mix (which contains loading buffer, see Amersham Sequenase version 2.0 squencing kit protocol).

2.30.1 Acrylamide Sequencing Gels

2.5 µl of each of the four sequencing reactions per template was loaded on an 8% acrylamide/7M urea sequencing gel, prepared for use with a IBI-supplied 'Base-Runner™' apparatus. The samples were heated to 80°C for 5 min prior to loading. The electrophoresis was powered by a Pharmacia LKB multidrive powerpack operating at 2500 V, 60 W, and 20 mA. When resolved, the gel was transferred to 3MM paper, vacuum dried (using a Hoeffer S.E. Ltd. gel drier), then incubated with autoradiographic film for between one and seven days. Sequence was then read directly from the autoradiograph when illuminated on an IBI digitizer lightbox.

2.31 AUTOMATED DNA SEQUENCING

2.31.1 Preparation of Template

A 5 ml *E. coli* culture was grown overnight in terrific broth with the appropriate antibiotic. Two 1.5 ml aliquots were pelleted and resuspended together in 200 µl of BDI buffer. 200 µl of BDII buffer was added, mixed and incubated on ice for 5 min. 300µl of BDIII was next added, mixed and incubated on ice for 5min. Cellular debris was removed by centrifugation at room temperature for 10 min at 12,000 x g. RNAase was added to a concentration of 20 $\mu g ml^{-1}$ and incubated at 37°C for 20 min. The solution was then extracted with chloroform twice, ispropanol precipitated, washed in 70% ethanol and vacuum dried for 3 min. The pellet was dissolved in 32 µl dH_2O then precipitated with 8 µl of 4 M NaCl and 40 µl of 13% PEG. The DNA was again pelleted at 12,000 x g for 20 min at 4°C, washed in 70% ethanol, vacuum dried for 3 min then resuspended in 20 µ l dH_2O.

2.31.2 Cycle Sequencing

This sequencing method is based on the incorporation of fluorescent dye-labelled terminator nucleotides into the synthesized DNA. All

four termination reactions are performed in the same reaction tube (0.5 ml thin-walled reaction tube). DNA cycle sequencing was performed on a Perkin Elmer Cetus Ltd. GeneAmp 9600 PCR system. The sequencing reagents were supplied as a *Taq* DyeDeoxy™ Terminator cycle sequencing kit by Applied Biosystems Inc. (Refer to Applied Biosystems Inc. automated sequencing training manual for detailed cycle sequencing reaction reagents used and their concentrations). The optimised cycle sequencing programme is detailed below.

The reaction was first incubated at 96°C for 2 min prior to thermal cycling.

- Rapid thermal ramp to 96°C.
- 96°C for 10 sec.
- Rapid thermal ramp to 50°C.
- 50°C for 5 sec.
- Rapid thermal ramp to 60°C.
- 60°C for 4 min.

This cycle was repeated 25 times, then the reactions were ramped to, and retained at, 4°C. The extension products then undergo spin column purification prior to loading on a sequencing gel using a NBL Centriflex™ gel filtration cartridge. Firstly the column is prepared by centrifugation at 1500 xg for 2 min to pack the column and to expel storage buffer into a centrifugation tube. The column is transferred to a fresh tube and the cycle sequencing sample is loaded onto the column. The column is centrifuged at 1500 xg for 4 min and the eluate is precipitated with 65 µl of ethanol/3 M sodium acetate (12:1 v/v) and pelleted by centrifugation at 11000 x g for 5 min. After a 70% ethanol wash, the pellet is air-dried for 10 min. The DNA is resuspended in 4 µl of loading buffer (supplied with cycle sequencing kit).

2.31.3 Automated Sequencing Gel Electrophoresis and Analysis

Cycle sequencing reactions were resolved on a 6% acrylamide/urea gel. The sequencing machine used was an Applied Biosystems Inc. 373 system where 32 samples, each comprising the four termination reactions, could be electrophoresed simultaneously. The collected data was analysed using Genejockey II software. The collected data is presented as a graphical representation of the fluorescence of each of the four fluorescent nucleotides incorporated into the DNA template during cycle sequencing. The software package allows translation of the graphical information into a readable G,C,A or T format. Sequence can then be easily manipulated at the computer terminal. Quality sequence can then be selected for further analysis and immediate restriction mapping. The selected sequence can also be used to identify analogous or similar sequences by way of database searches.

2.32 OLIGODEOXYNUCLEOTIDE MANIPULATIONS

Oligodeoxynucleotides of 20-30 bases were used as probes in aqueous hybridisations or were annealed to generate linkers which were cloned as a means of introducing restriction sites into plasmids. Deoxyoligonucleotides were synthesized on an Applied Biosystems DNA synthesizer or commercially by Cruachem Ltd. The commercially synthesized oligodeoxy-nucleotides were supplied as quantified, previously deprotected DNA pellets.

2.32.1 Deprotection of Oligodeoxynucleotides

The oligodeoxynucleotides synthesized using the departmental machine had to be deprotected by removing the glass beads (to which the DNA is bound) from the synthesis column and placing them in 1

ml of 30% aqueous ammonia hydroxide in a nunc™ tube. This solution was incubated at room temperature for 2 hrs. The beads were then pelleted and the supernatant removed and kept. To the supernatant, another 1 ml of 30% aqueous ammonia hydroxide was added and incubated at 50°C overnight. This deprotected oligo in ammonia hydroxide solution was then stored at -20°C until the oligodeoxynucleotide was required whereupon the solution was ethanol precipitated, centrifuged and resuspended in dH_2O

2.32.2 Annealing Oligodeoxynucleotides to Generate Linkers

To introduce a particular restriction enzyme recognition sequence at a chosen locus in a DNA sequence, a pair of oligodeoxy-nucleotides of approximately 25 bases were designed and synthesized. The oligodeoxynucleotides were annealed to generate a double stranded linker sequence for sub-cloning. The oligodeoxynucleotides were complementary to each other for most of their sequence except for a few bases at the ends which would protrude when the oligodeoxynucleotides were annealed. The protruding bases were specifically designed to create double stranded termini which were compatable with termini generated by a particular restriction enzyme. Ligation of linker to a selected, restricted DNA fragment was therefore possible. The oligodeoxynucleotides, when annealed, were designed so that a particular restriction enzyme recognition sequence was included between the linker termini. The introduced restriction site could then be used for further fragment sub-cloning.

The deoxyoligonucleotides were annealed by mixing 1 µg of each DNA in a total volume of 100 µl and incubating for 10 min in a water bath at 70°C. The reaction was then allowed to cool to room temperature slowly by simply turning off the water bath. The reaction tube was left in the water bath overnight, so that the reaction would cool at the same slow pace as the water. The annealed oligodeoxynucleotide linkers were then phosphorylated with T4 polynucleotide kinase prior to ligation (see section 2.22).

2.33 AUTORADIOGRAPHY

Autoradiography was performed in metal cassettes (medical chest X-ray type). The films used were Fuji medical X-ray film, and Amersham Hyperfilm™ (for sequencing and *in situ* studies). The intensity of the autoradiographic images was enhanced when the cassettes were placed at -70°C to expose in the presence of DuPont Cronex Lightening Plus intensifying screens. Amersham Hyperfilms were developed by hand and the regular Fuji films using an X-ograph Compact-X2 developing machine.

2.34 POLYMERASE CHAIN REACTIONS

Polymerase chain reactions were performed on genomic DNA, bacteriophage clones and plasmid DNA. A Perkin Elmer Cetus 'DNA Thermal Cycler' was used. The reaction constituted:

- 200-500 ng template DNA (denatured at 95°C for 5 min).
- 50 ng of each oligodeoxynucleotide primer.
- 2.5 µ l Promega 10X *Taq* buffer (final 1X conc.).
- 0.5 µl 10 mM dNTP nucleotide mix.
- 0.5 µl Promega *Taq* polymerase (10u/µl).
- dH_20 to 25 µl.
- overlay with 20 µl of mineral oil to prevent evapouration.

The primers (18-25 bases) were designed so that their melting temperatures were roughly equivalent and in the range 55-65°C and their G+C content was ideally 50%. The first test PCR reactions were performed at low stringency conditions. The stringency is influenced by the annealing time and temperature, the magnesium ion concentration, the primer to template DNA ratio, the reaction extension time and the number of cycles. When a PCR product band of the expected size appeared (determined by gel electrophoresis), the reaction was optimised to reduce spurious background bands.

The previously described stringency parameters were altered to achieve this.

2.35 PREPARATION OF RNAase-FREE (R.F.) SOLUTIONS

For *in situ* hybridisation experiments, RNAase-free solutions were required. Diethyl pyrocarbonate (DEPC) is used to remove any RNAase from water, salt buffers, and the other solutions used in *in situ* hybridisation buffer. Under air extraction conditions, 0.5 ml of 100% DEPC (w/v) is added per litre of the solution which is to be treated and the mix is shaken vigorously. The bottle of solution

undergoing DEPC treatment is then left overnight in an operating fume hood with the bottles screw-cap loosened slightly. After autoclaving the solution is considered RNAase-free and ready for use.

Appendix I: Solutions

Please find below a list of the buffers which were made (rather than purchased commercially), and which were used in the experiments detailed in this text. The composition of those buffers and the pH values (where relevent) are given.

(1) Mini-prep column wash buffer. 200mM NaCl, 20mM Tris, 5mM EDTA. Mixed 1:1 with 100% ethanol prior to use.

(2) Birnboim-Doly solution P1 (BDI). (Lysis buffer) 50mM Glucose, 10mM EDTA, 25mM Tris (pH 8).

(3) Birnboim-Doly solution P2 (BDII). (Denaturing buffer) 0.2M NaOH, 1% SDS. This solution is freshly made prior to use.

(4) Birnboim-Doly solution P3 (BDIII). (Neutralising buffer) 5M KAc. For 100ml; 60ml 5M KAc, 11.5ml glacial acetic acid, 28.5ml dH_2O. Stored at 4°C.

(5) Denaturing solution (Agarose gels). 1.5M NaCl, 0.5M NaOH. Stored at room temperature.

(6) Neutralising solution (Agarose gels). 3M NaCl, 0.5M Tris (pH 7.6).

(7) 20X SSC. 3M NaCl, 300mM Na_3 Citrate, pH to 7.0.

(8) Pre- and Hybridisation buffer. 5X SSC, 0.1% SDS, 5X Denhardts solution, 100μg/ml denatured salmon/herring sperm DNA, 5% dextran sulphate.

(9) 50X Denhardts solution. 5g Ficoll, 5g polyvinylpyrolidone, 5g BSA (bovine serum albumin), made up 500ml with dH_2O.

(10) Deoxyoligonucleotide hybridisation buffer. For 50mls, 500µl 10% SDS, 2.5g dextran sulphate, 5ml 50X Denhardts solution, 10ml 20X SET, 1ml 1M phosphate buffer (pH 6.8).

(11)

(12) 20X SET buffer. To 1.5l dH_2O, add 350.64g NaCl, 96.88g Tris, 14.89g EDTA, then dH_2O to 2 litres and pH to 8.0 with HCl.

(13) Low salt NET buffer. 100mM NaCl, 0.1mM EDTA, 20mM Tris (pH 8.0).

(14) High salt NET buffer. 1M NaCl, 0.1mM EDTA, 20mM Tris (pH 8.0).

(15) Transformation buffer (TFB). 100mM rubidium Chloride, 10mM MES, 45mM manganese chloride, 10mM calcium chloride, 3mM hexaminocobaltic acid.

(16) QBT buffer. 750mM NaCl, 50mM MOPS, 15% ethanol, pH7.0, 0.15% triton X-100.

(17) QF buffer. 1M NaCl, 50mM MOPS, 15% ethanol, pH 7.0.

(18) QC buffer. 1.25M NaI, 50mM MOPS, 15% ethanol, pH 7.0.

(19) 50X TAE. 242.5g Tris, 18g sodium acetate, 18g $Na_2EDTA.2H_20$ made up to 1 litre with dH_20 , pH 8.3.

(20) Extraction buffer. 10mM Tris-HCl (pH 8.0), 0.1M EDTA (pH 8.0), 0.5% SDS, 20µgml^{-1} boiled RNAase A.

(21) TE buffer. 10mM Tris-HCL, 1mM EDTA, pH 8.0.

(22) 10X Kinase buffer. 0.5M Tris-HCl (pH 7.6), 100mM magnesium chloride, 50mM DTT, 1mM spermidine, 1mM EDTA.

(23) 10X MOPS. 42g morpholinopropane sulphonic acid, 6.8g sodium acetate, 3.7g Na_2EDTA (pH 8.0) made up to 1litre with distilled water.

(24) Agarose gel loading buffer. 0.025% (w/v) bromophenol blue, 0.025 (w/v) xylene cyanol, 25% (w/v) ficoll, 0.5% (w/v) SDS, 50mM EDTA.

CHAPTER 3

RESULTS: CLONING

INTRODUCTION

This chapter describes the steps taken to clone the mouse preprotachykinin-A (PPT-A) gene and the neurokinin-1 receptor (NK-1R) gene which was formerly known as the substance P receptor (SPR) gene. The isolation of bacteriophage clones containing either PPT-A or NK-1R genomic sequence from a bacteriophage lambda library is described. The design of the DNA probes, the optimisation of the hybridisation conditions and the result of the library hybridisations are described and illustrated. This chapter is divided into two sections, the first dealing with the PPT-A gene and the second with the NK-1R gene.

CLONING THE MURINE PPT-A GENE

3.1 PPT-A PROBE DESIGN FOR GENOMIC LIBRARY SCREENS

The identification of positively-hybridising PPT-A bacteriophage clones from genomic lambda libraries required a suitable homologous sequence of genomic PPT-A DNA for use as a probe. No subcloned PPT-A genomic or cDNA sequence from either mouse or rat was available to allow the selection of a suitable probe. A polymerase chain reaction (PCR) based approach was adopted to generate a probe using mouse genomic DNA as the PCR template. Genomic DNA was prepared as described in section 2.16. From the

published rat genomic DNA sequence (Carter *et al*, 1990), PCR primers were selected to amplify an exon-spanning sequence of the mouse PPT-A gene. Primers were designed to ideally be between 18 and 28 bases in length, have a guanine/cytosine content of 50%, and to hybridise to loci separated by less than 1500 bp on the genomic template. Primers were selected which hybridised to highly conserved genomic regions, preferably within exons, and which would not anneal to each other to form 'primer dimers'. PCR products of a size within this range are synthesised routinely and were considered ideal in length for use as library hybridisation probes. Three different pairs of PCR primers were tried at low through to high PCR stringency conditions (see Section 3.1.4). These primers failed to generate PCR products of the predicted size. This was calculated from the published rat genomic DNA sequence (Carter *et al*, 1990). There is approximately 95% homology between rat and mouse species. The possibility therefore existed that, when mouse genomic DNA was used as template, the size of the subsequent PCR product could have been marginally different from that generated using the same primers but with rat genomic DNA as template. No bands of even an approximately similar size were generated using mouse DNA template. The fourth pair of primers to be tried were successful and generated a rat PCR product of the expected size (514 bp including the primers). The successful PCR primers were called CS6 and CS7.

3.1.1 Polymerase Chain Reaction Primer CS6

Primer CS6 is a 24-mer with a guanine/cytosine to adenine/thymine (GC:AT) ratio of 10:14, and a melting temperature (Tm) of 68°C. The Tm of PCR primers were calculated by summing the number of bases. A value of 4°C was allocated to each guanine or cytosine residue and 2°C to each adenine or thymine residue. The primer Tm can be calculated precisely using a formula (Erlich, 1989). Primer CS6 hybridises to the centre of PPT-A exon 3 (97 bp in rat), and to the sense strand, priming 5'-3' (see Figure 3.1). The sequence of primer CS6 is identical to the rat genomic sequence and differs from the bovine sequence by 1 nucleotide (see Figure 3.1).

This single base degeneracy does not alter the amino acid (leucine) translated from this codon in exon 3 of the rat (and cow) PPT-A genes. As a result of the expected high degree of homology with the corresponding mouse sequence, this primer was a strong candidate for successful PCR amplification using mouse DNA template.

Figure 3.1: Primer CS6 Sequence Homologies

(a) Primer CS6: Rat Genomic DNA and Amino Acid Sequence.

```
5'  CCTCAGCAGTTCTTTGGATTAATG  3'
    ProGlnGlnPhePheGlyLeuMet
```

(b) Corresponding Bovine DNA and Amino Acid Sequence.

```
5'  CCTCAGCAGTTCTTTGGATTGATG  3'
    ProGlnGlnPhePheGlyLeuMet
```

(c)

```
      1         10        20        30        40        50
5'tctagGAGGAAATGCCCGAGCCCTTTGAGCATCTTCTGCAGAGAATCGCTCGGAGA
   |||||+|||||||||+|||||||||+|||||||||+|||||||||+|||||||||+
   agatcCTCCTTTACGGGCTCGGGAAACTCGTAGAAGACGTCTCTTAGCGAGCCTCT
        Î Start of rat PPT-A exon 3.

            60  CS6 Primer      80        90  95
   CCCAAGCCTCAGCAGTTCTTTGGATTGATGGGCAAACGGATGCTGgtgagat  3'
   |||||||||+|||||||||+|||||||||+|||||||||+|||||||||+||
   GGGTTCGGAGTCCTCAAGAAACCTAACTACCCGTTTGCCTACGACcactcta
```

Figure 3.1: (a) Rat genomic DNA sequence and corresponding amino acid sequence of PCR primer CS6. (b) Bovine genomic DNA sequence and corresponding amino acid sequence of PCR primer CS6. (c) DNA sequence of the 95 bp rat PPT-A exon 3. Bold and underlined sequence denotes the site where PCR primer CS6 hybridises. Capital letters denote exonic sequence and small case letters denote intronic sequence.

3.1.2 Polymerise Chain Reaction Primer CS7

Like primer CS6, primer CS7 was a 24-mer but with a GC:AT ratio of 11:13 and therefore a melting temperature of 70°C. This primer hybridises to the 3' end of PPT-A exon 4 (which is 45 bp in rat) and to the anti-sense DNA strand, priming 3'-5' (as shown in Figure 3.2 below). The sequence of this primer has 100% identity in both rat (Carter *et al*, 1990) and cow (Kotani *et al*, 1986) genomic DNA. This primer was also a good candidate for success ful PCR of the PPT-A gene with mouse genomic DNA template.

Figure 3.2 : Primer CS7 Sequence Homologies.

(a) Primer CS7: Rat Genomic DNA and Amino Acid Sequence.

```
3'  CACCGGGACAATTTCCGGGAAATA  5'
    ValAlaLeuLeuLysAlaLeuTyr
```

(b) Corresponding Bovine Genomic DNA and Amino Acid Sequence.

```
3'  CACCGGGACAATTTCCGGGAAATA  5'
    ValAlaLeuLeuLysAlaLeuTyr
```

(c)

```
5'tcagATTCCTCAATTGAAAAGCAAGTGGCCCTGTTAAAGGCCCTTTATGgtaa 3'
  ||||+|||||||+|||||||+|||||||+|||||||+|||||||
3'agtcTAAGGAGTTAACTTTTGCTTCACCGGGACAATTTCCGGGAAATACcatt 5'
    Î                        Primer CS7
    Start of rat PPT-A exon 4.
```

Figure 3.2: (a) Rat genomic DNA sequence and corresponding amino acid sequence of PCR primer CS7. (b) Bovine genomic DNA sequence and corresponding amino acid sequence of PCR primer CS7. (c) DNA sequence of the 45 bp rat PPT-A exon 4. Bold and underlined sequence denotes the site where PCR primer CS7 hybridises. Capital letters denote exonic sequence and small case letters denote intronic sequence.

3.1.3 Genomic Template Polymerase Chain Reactions

Polymerase chain reaction conditions are described in section 2.34. PCRs using primers CS6 and CS7 were conducted using rat genomic DNA as template in a bid to determine, in the first instance, whether or not the primers would successfully anneal to and consequently generate a PCR product of the expected size from template which was known to be 100% homologous.

At the annealing temperature of 64°C, the primers were successful and generated a single PCR product band of approximately 514 bp as visualised by electrophoresis on a 1% agarose gel. Each of the PCR reactions contained 500 ng of genomic DNA as template, and 25 ng of each primer. Control reactions included (a) reactions minus template and (b) reactions with only one or the other primer (as shown in Figure 3.3).

The successful rat genomic PCR reaction was repeated under the same conditions using mouse and then human genomic DNA as template. In both cases a single PCR product band of the expected size (approximately 514 bp) was generated. The same experimental controls were included as for the original rat genomic PCR.

3.1.4 Optimisation of Genomic PCRs

Non-specific PCR products are responsible for the background smear visualised when a sample of the PCR reaction is electrophoresed on an agarose gel. To reduce the amplification of non-specific PCR products, the stringency of the reaction was increased. By increasing the annealing temperature of the reaction incrementally by 1 °C, an optimal temperature of 68°C was found where minimal non-specific products were synthesized. An increase or decrease in $[Mg^{2+}]$ concentration, genomic DNA template concentration or primer concentration did not affect the degree of background amplification significantly for this particular PCR.

Figure 3.3 : Genomic DNA PCR Reactions

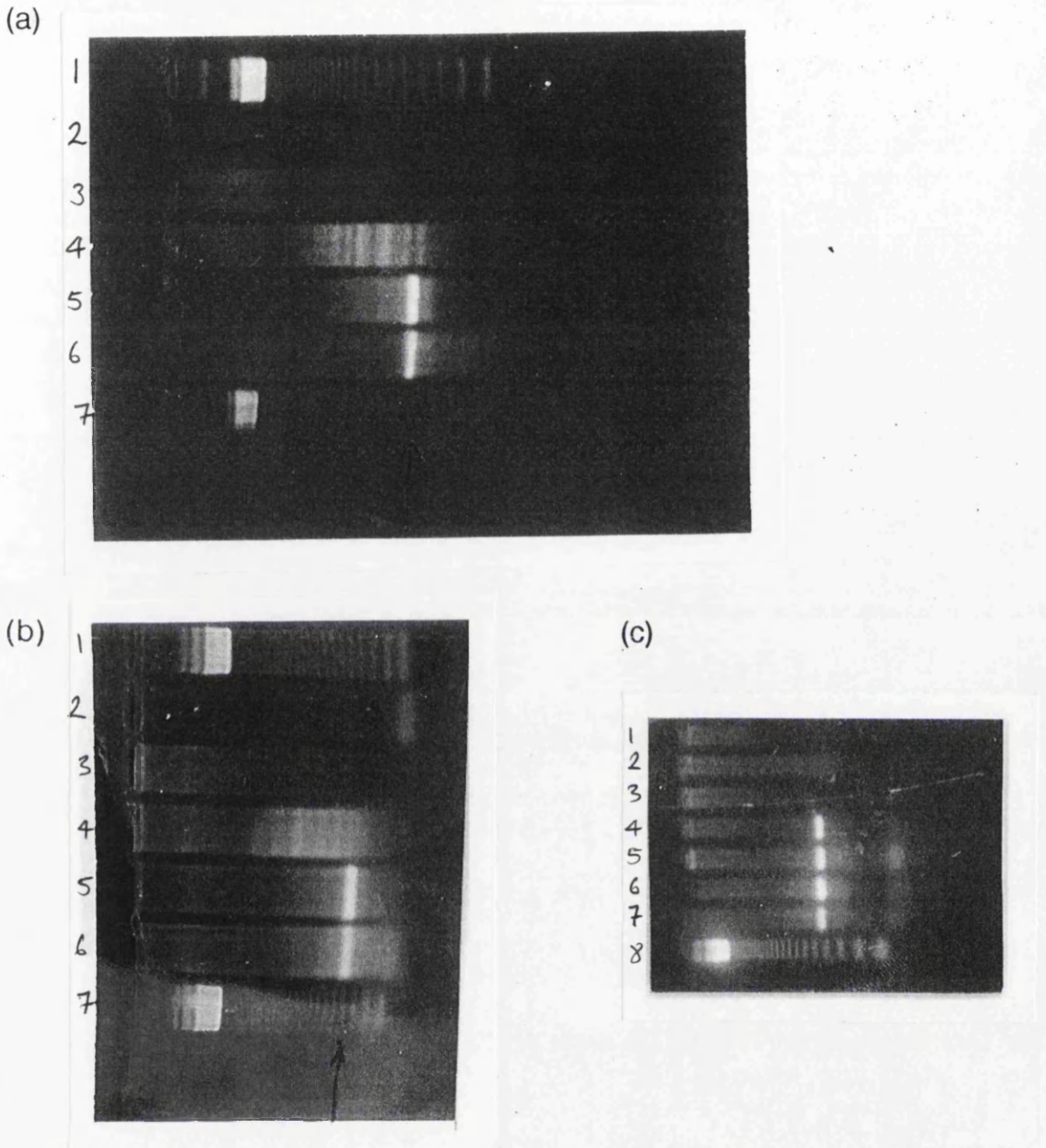


Figure 3.3: Rat, mouse and human genomic DNA PCR with primers CS6 and CS7.

(a) Rat genomic DNA template. Lane 1; 123 bp ladder, Lane 2; no template DNA, Lane 3; primer CS6 only, Lane 4; primer CS7 only, Lane 5; test with 0.5 μg template DNA, Lane 6; test with 1 μg template DNA. Lane7; 123 bp ladder.

(b) Mouse genomic DNA template. Lane 1; 123 bp ladder, Lane 2; no template DNA, Lane 3; primer CS6 only, Lane 4; primer CS7 only, Lane 5; test with 0.5 μg template DNA, Lane 6; test with 1 μg template DNA., Lane 7;123 bp ladder.

(c) Human genomic DNA templates. Lane 1; no template DNA, Lane 2; primer CS6 only, Lane 3; primer CS7 only, Lane 4; other test sample, Lane 5; other test sample, Lane 6; other sample, Lane 7; 0.5 μg template DNA, Lane 8; 123bp ladder.

The successfully optimised mouse genomic PCR was repeated on a larger scale and electrophoresed.

The 514 bp band was excised and the DNA recovered as described in section 2.27.1. The recovered 514 bp PCR product was used as template for a further PCR reaction. Using this template, as opposed to genomic DNA, helped to reduce the number of non-specific loci to which the primers could anneal. This increased the proportion of specific, expected PCR product relative to the non-specifc product which was synthesized utilising loci outwith the PPT-A gene sequence. The number of background PCR products was thus negligible while the 'purity' of the resulting PCR product was increased significantly. Under optimal PCR conditions, mouse, rat and human genomic DNA template continued to produce single bands of the predicted size.

Figure 3.4: Optimised Mouse Genomic DNA PCR Reactions

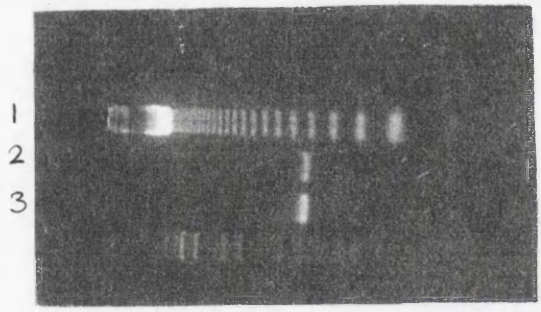


Figure 3.4: Optimised mouse genomic DNA template PCR (primers CS6 and CS7).
Lane 1; 123 bp ladder, Lane 2; 1 µl of PCR reaction (100 ng of mouse genomic DNA template in 20 µl volume), Lane 3; 3 µl of PCR reaction (100 ng of mouse genomic DNA template in 20 µl volume).

A PCR product of predicted size does not conclusively demonstrate the authenticity of the PCR product. In order to verify that the PCR product was indeed an amplified PPT-A sequence, it was deemed necessary to subclone the PCR product and use DNA sequencing and endonuclease restriction analysis to ensure the product was from the correct amplified genomic location and therefore suitable for use as a hybridisation probe for the isolation of PPT-A-specific clones from a bacteriophage library. It was expected that the mouse PCR product sequence and restriction profile should resemble the equivalent published rat and bovine data.

3.1.5 Subcloning the Genomic PCR Products

Genomic PCR products were cloned into the plasmid vector pDK101 which is a modified version of the pGEM5Zf(+) plasmid (Promega Corp.). This vector is also known as 'T-vector' (Kovalic, 1991) because of the single thymine residue which protrudes from the DNA strand at each end of the plasmid when it is linearized with the restriction endonuclease *Xcm* I (supplied by New England Biolabs Ltd). PCR products often have a single, protruding adenine residue gratuitously added to each end of the product by the *Taq* polymerase modifying enzyme used in polymerase chain reactions (but not always, Hu *et al*, 1993). The respective complementary overhanging thymine and adenine residues under DNA ligation conditions promote ligation of the PCR product to the plasmid vector and therefore the generation of recombinant DNA molecules (subclones).

3.1.5.1 Preparation of DNA Fragments for Ligation

The pDK101 plasmid was grown under ampicillin selection, and the DNA was isolated using the Qiagen Midi-Prep System as described in section 2.17.2.2. The vector plasmid was then linearized with *Xcm* I and electrophoresed on a 0.7% TAE/low melting point agarose gel. The restricted DNA fragment was recovered from the agarose slice using the Isogene™ method (section 2.27.1). The linear vector was dephosphorylated using calf intestinal alkaline phosphatase (section 2.22.1), and was subsequently used in ligation reactions after the

phosphatase enzyme was removed by phenol/chloroform extraction (section 2.23). If the phosphatase is not removed, the enzyme would proceed to dephosphorylate the insert fragment in the ligation reaction and therefore drastically reduce the frequency of successful ligation of insert to vector.

The insert fragments for ligation into the prepared vector, (the mouse and rat PCR products), were electrophoresed, excised and recovered in the same way as for the pDK101 vector (detailed above). However they were not dephosphorylated. The purified PCR products were synthesized by using previously band-isolated PCR product as template, as described above, to reduce the probability of cloning PCR artefacts.

3.1.5.2 Ligation of PCR Product into pDK101

The conventional method (as opposed to a ligation kit) was used to ligate the mouse and rat PCR products to pDK101 vector plasmid (section 2.21.1), where 100 ng of vector and 100 ng of insert (or multiples thereof) were added to each ligation in a total volume of 10 µl. An insert to vector molar ratio of 3:1 was considered likely to generate the highest number of successful recombinants. A different PCR product (*etl*-1, from Dr. P. Sheils, University of Glasgow), which had previously been cloned successfully was used as a control ligation insert. Controls lacking ligase enzyme were also included.

3.1.5.3 Transformation of PCR Product Ligations

The ligations were phenol/chloroform extracted before transformation because of the relatively high salt concentration in the ligation mix. This can reduce the transformation efficiency by causing the electroporation equipment to 'arc' (similar to short-circuiting). The ligations were transformed by electroporation (section 2.4.2) with competent *E. coli* XL1-Blue cells and plated on amp^+/tet^+ agar plates. Transformation controls included: uncut plasmid only, cells plus linearized vector only, and a competent cells only transformation (plated on amp^-/tet^- agar plates). Insert-

containing recombinant clones were identified by blue/white visual selection with the aid of the indicator X-gal and the inducer IPTG. The positive and negative controls gave results as expected, and the test ligation/transformation plates displayed both blue and white colonies. Ten single white colonies were selected for further investigation from the mouse and rat PCR product/pDK101 ligation test plates. The colonies were streaked on selective plates and incubated at 37°C overnight before single colonies were again selected. The plasmid DNA was prepared from these cultures and used as template for manual sequencing.

3.1.6 Analysis of PCR Product Subclones

To determine the authenticity of the PCR product sequences, preliminary restriction fragment size analysis was conducted, followed by PCR analysis and DNA sequencing.

3.1.6.1 Restriction Fragment Size Determination

As a preliminary authenticity check, the restriction endonuclease *NcoI* was used to release the insert DNA fragment from the vector. The insert size was compared to that of the original PCR product by electrophoresis. For each of the two PCR product ligations, one recombinant plasmid subclone shown to contain an insert of the correct predicted size was isolated. They were named pCOL1 and pCOL2 and contained the mouse and rat PCR products respectively.

3.1.6.2 Polymerase Chain Reaction Analysis

The subclone plasmids containing putative PCR products were further investigated to determine their integrity. The recombinant plasmids were used as templates for PCR reactions. The primers, CS6 and CS7, and the reaction conditions used were the same as those used originally to generate the PCR products prior to subcloning. Plasmids pCOL1 and pCOL2 each generated single PCR bands of approximately 514 bp as expected.

3.1.6.3 Sequencing Analysis of Cloned PCR Products

Plasmid pCOL1 was prepared as a double stranded DNA sequencing template (section 2.17.1.1). The alignment of the 248 bp of sequence obtained from the cloned mouse PPT-A PCR product (pCOL1) was aligned with the published PPT-A rat sequence. The percentage identity was found to be 87.9% and is shown below.

Figure 3.5: pCOL1 Sequence Aligned with Rat PPT-A Sequence.

```
  1 GATCTATCAGGCTACCTGGTCTGCATGCTTGCTCCTTCCCGGAGAACCCA 50
    | ||||||| |||||| ||+|||||||| +|||||||||| |||| ||||+
306 GGTCTATCAGGCTACCCGGTCTGCATGCCTGCTCCTTCCTGGAGCACCCA 356

 51 AATGTCTCTTGTACTGGGAGTATAGATTTACACGCCTAAAGGCTTTTGAT 100
    | ||||||| |||| | ||+|||| ||| +|| |||||| +| ||||||| +
357 AGTGTCTCTTGTACCGCGAGTATAATTTTACATGCCTAAAGACTTTTGAT 406

101 TTCCCAGAAGTCTTCGATCTACTTCACTAACTCTAGATCGTGGGGGGAGG 150
    |||| |||| +|||||||| +|| |||| +| |||| ||+ |||||||| +
406 TCCCTGAAGTCTTCGATCCACTACACTTGCACTAGCTCAAAGTGGAGAGG 456

151 GAGGGGGGATCCATTTCTCTTGCTTCACTGCACCAAAGTGATGCAAATTG 200
    ||||||||| +|||||||  |+|||||||  +|||||| ||+| ·|||||| +
457 GAGGGGGGATCCATTTCCCTTGCTTCACAGCACCAAGGTGACCCAAATTG 506

201 GAAACTAACCTTAGCTAAACAGTCCTCTGACTAAAGATCAAAATTACA 248
    ||||||||| +|||||||||| +| ||| |||+|||||||||+ |||||||
507 GAAACTAACCTTAGCTAAACACTCCCCTGACTAAAGATCCCAATTACA 554
```

Figure 3.5: Alignment of 248 bp of plasmid pCOL1 sequence (upper) with published rat NK-1 sequence (lower). Plasmid pCOL1 comprises the mouse PCR product (Section 3.1.6.2) subcloned into vector pDK101. The published rat NK-1 sequence was retrieved from the EMBL database (accession number m34160) and is defined as containing the sequence of rat PPT-A exons 3 and 4. NB. The PCR primers used to generate the (subcloned) PCR product (primers CS6 and CS7) were designed to anneal to exons 3 and 4 respectively. The (inter-species) percentage identity between the published rat and the PCR-cloned mouse sequence is 87.9% over this short stretch of sequence.

Figure 3.6: pCOL2 Sequence Aligned with Rat PPT-A Sequence.

```
   1........CTCAGCAGTTCTTTGGATTAATGGGCAAGCGGGATGCTGGTG 42
            |||||||||||||||||||||||||||| ||||||||||||
  51 ACCCAAGCCTCAGCAGTTCTTTGGATTAATGGGCAAACGGGATGCTGGTG 100

  43 AGATGAGCAATAGTCACTAGATCTATCAGGCTACCTGGTCTGCATGCTTG 92
     |||||||||||||||||| |||||||||||||| |||||||||| ||
 101 AGATGAGCAATAGTCACTAGGTCTATCAGGCTACCCGGTCTGCATGCTTG 150

  93 CTCCTTCCTGGAGAACCCAAATGTCTCTTGTACTGGGAGTATAGATTTAC 142
     |||||||||||| |||||| |||||||||||| | ||||||| |||||
 151 CTCCTTCCTGGAGCACCCAAGTGTCTCTTGTACCGGCAGTATAATTTTAG 200

 143 ACGCCTAAAGGCTTTTGATTCCCAGAAGTCTTCGACT..............  179
     | |||||||| |||||||||||| ||||||||||
 201 ATGCCTAAAGACTTTTGATTCCCTGAAGTCTTCGATCCACTAACACTTGC 250
```

Figure 3.6: Alignment of 179 bp of sequence from plasmid pCOL2 (upper sequence) with the published rat PPT-A sequence (lower sequence). Plasmid pCOL2 comprises the rat PCR product (Section 3.1.6.2) subcloned into vector pDK101. The published rat PPT-A sequence was retrieved from the EMBL database (accession number m34160) and is defined as containing the sequence of rat PPT-A exons 3 and 4. NB. The PCR primers used to generate the (subcloned) PCR product (primers CS6 and CS7) were designed to anneal to exons 3 and 4 respectively. The percentage identity between the published rat and the PCR-cloned rat sequence is 91.1% over this short stretch of sequence.

The PCR, sequencing and restriction analysis of the subclones pCOL1 and pCOL2 did verify their expected authenticity. Subclones pCOL1 and pCOL2 contained, respectively, the subcloned mouse and rat exon 3/4-spanning PPT-A PCR products of the expected size and sequence identity.

3.2 TEST HYBRIDISATIONS WITH PPT-A PROBE

Prior to using the above 514 bp PCR-derived PPT-A probe in bacteriophage library hybridisation experiments, it was necessary to determine the optimal hybridisation conditions for this particular probe. Southern analysis was employed to do this, where restricted mouse genomic DNA was transferred by capillary blot to a nylon membrane (Hybond-N) which was subsequently probed using the PCR-derived probe sequence. The stringency of the hybridisation conditions, such as the hybridisation temperature, the salt concentration of the hybridisation buffer, the washing temperature and the salt concentration of the wash buffer, was increased incrementally. Optimal hybridisation conditions were visually determined when there was minimal background hybridisation signal and when a strong positive hybridisation signal was visible on autoradiographed, probed membranes.

3.2.1 Radioactively-Labelling the PCR-Derived PPT-A Probe

DNA hybridisation probes are typically radioactively-labelled by the random primer extension method (section 2.24.1) using random oligodeoxynucleotides (hexamers) as primers. In order to label PCR products, the primers used in the original PCR reaction can be used as specific rather than random primers ('specific priming'). Alternatively, where a PCR product has been subcloned, it can be used as template in a 'hot PCR'. In such cases, the combined dNTP mix in the PCR reaction is replaced by three unlabelled nucleotides ('cold' dATP, dGTP and dTTP for example) and one 'hot' labelled nucleotide (γ-dCTP for example). This results in simultaneous incorporation of the radioactive label and PCR synthesis of the probe DNA. Band purification of the 'hot' labelled PCR product follows or a sample of the PCR reaction can be denatured and used directly as a probe for hybridisations. For the following NK-1 Southern analysis experiment and the primary screen of the Glaxo library (section 3.3.3), the hot PCR method was used, where pCOL1 plasmid was used as template. The 'specific priming' labelling method was favoured

for the secondary screen of the Glaxo library and was used to label the PCR-derived probe for all subsequent hybridisations.

3.2.2 Southern Hybridisations with PPT-A Probe

A restricted genomic DNA membrane was prepared for use in test hybridisations. To do this 10 µg of mouse and rat genomic DNA was restricted with either *EcoR* I and *Pst* I. Samples of each reaction were checked by electrophoresis to ensure complete digestion as demonstrated by the appearance of 'satellite' type bands and a DNA 'smear' in the restricted sample lanes (see Figure 3.7 below). When fully digested, the remaining reaction was electrophoresed to ensure adequate separation of the restricted DNA bands over a gel distance of 16 cm. The gel was then capillary blotted overnight and the DNA fixed to the membrane by ultra-violet cross-linking. The lanes from the gel were marked on the membrane to allow alignment and sizing of bands later. The membrane was prehybridised in a hybridisation oven with 15 ml of hybridisation buffer for 6 hrs at 60 °C.

25 µl of mouse genomic PCR reaction, containing 125 ng of DNA was used as probe after denaturation. The probe was added to 6 ml hybridisation buffer and incubated with the membrane for 14 hrs. There was 20 ng of probe DNA per ml of hybridisation buffer, with approximately 1.8x 10^9 dpm.µg^{-1} and incorporation of 70%. After incubation, the filter was washed then exposed to autoradiographic film for 24 hrs. This film displayed significant non-specific background hybridisation. However, strong positively hybridising bands could be seen, where three bands were visible in the *Pst*I digested mouse and rat genomic DNA lanes. The wash stringency was then increased to 1.5 x SSC and then to 1 x SSC, while the temperature and the duration of the washes remained the same. The membrane was again exposed to autoradiographic film for 24 hrs. The developed film displayed a considerable reduction in background signal. The number of positively hybridising bands in the mouse and rat *Pst*I digested lanes was also reduced from three to

one (see Figure 3.8). The PCR-derived PPT-A probe was shown to hybridise specifically to mouse (and rat) genomic DNA under these hybridisation and washing conditions. These optimised hybridisation conditions were deemed satisfactory to permit the use of this probe in bacteriophage library screening experiments.

Figure 3.7 : Electrophoresed, Digested Mouse and Rat Genomic DNA.

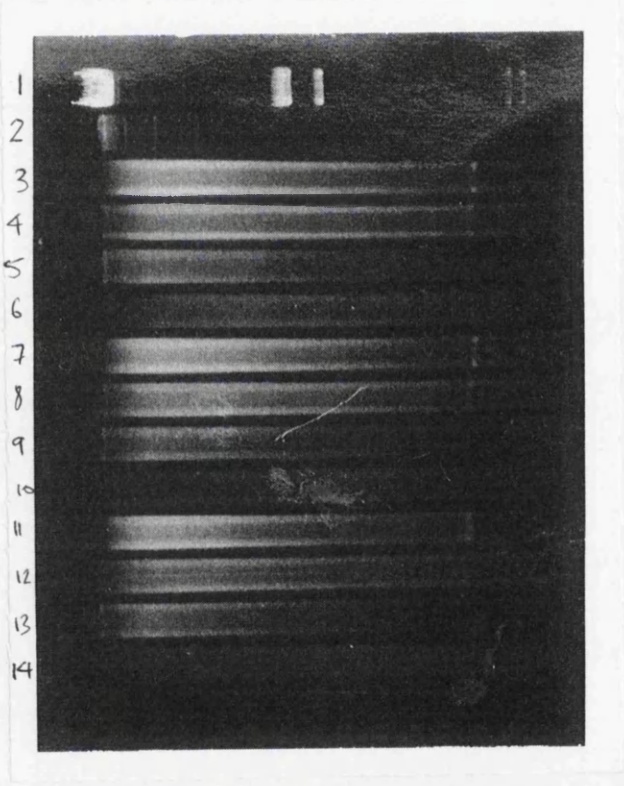
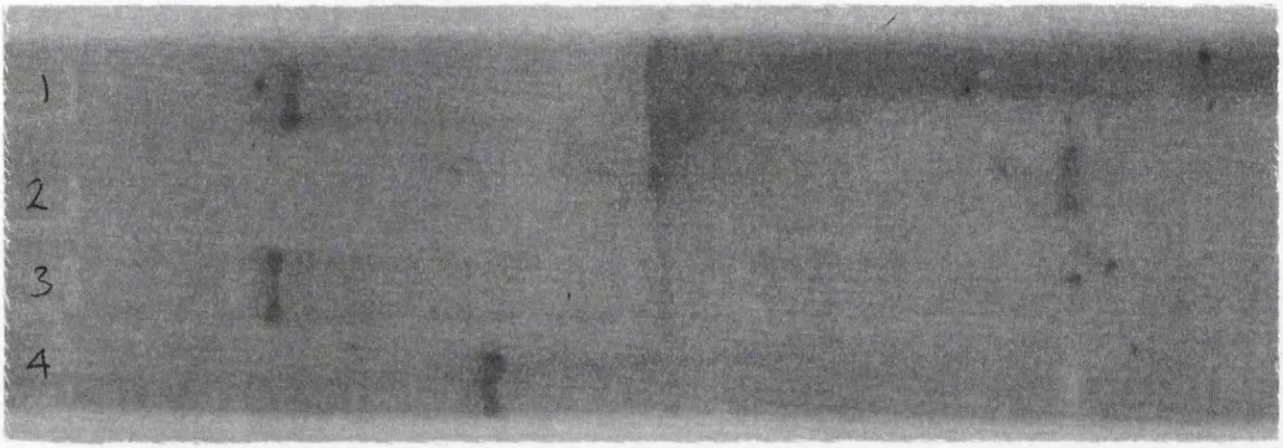


Figure 3.7: Electrophoresed, endonuclease digested mouse and rat genomic DNA. Lane 1; lambda/*Hin*D III marker, Lane 2; incompletely digested marker, Lane 3; 10 µg rat genomic DNA digested with *Eco*R I, Lane 4; 10 µg rat genomic DNA digested with *Pst* I, Lane 5; 10 µg mouse genomic DNA digested with *Eco*R I, Lane 6; 10 µg mouse genomic DNA digested with *Pst* I. (Lanes 3, 4, 5 and 6 are repeated in lanes 7, 8, 9 and 10, and also lanes 11, 12, 13 and 14 respectively).

Figure 3.8 : Autoradiographs of Southern Blotted Mouse and Rat Genomic DNA

(a)



(b)

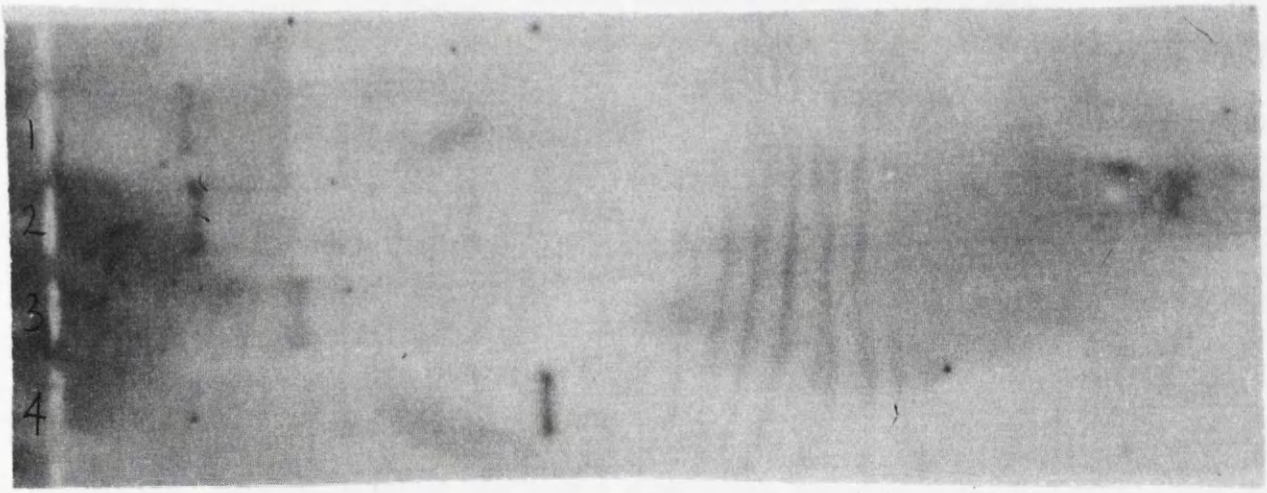


Figure 3.8: Autoradiographs of *EcoR* I and *Pst* I-digested, Southern blotted mouse and rat genomic DNA (see Section 2.28). (a) This Southern filter was hybridised with the mouse PPT-A exon 3 and exon 4-specific PCR product probe radioactively labelled internally by the 'hot PCR' procedure (as described above). This Southern filter comprises 4 lanes and is a direct capillary blot of the digested DNAs photographed and shown in figure 3.7 (above). (b) This Southern filter was used as a control in the NK-1 primary bacteriophage library hybridisation which is detailed later in this text (see Section 3.3.3.1). Lane 1; 10 μg mouse genomic DNA digested with *Eco*R I, Lane 2; 10 μg mouse genomic DNA digested with *Pst* I , Lane 3; 10 μg rat genomic DNA digested with *Eco* R I, Lane 4; 10 μg rat genomic DNA digested with *Pst* I.

3.3 ISOLATION OF PPT-A GENOMIC LIBRARY CLONES

3.3.1 The Construction and Source of the Genomic Bacteriophage Libraries

Genomic mouse preprotachykinin-A clones were isolated from two separate bacteriophage lambda libraries. Both libraries were constructed from strain SVJ129 mouse genomic DNA so that the cloned DNA would be compatible for use in homologous recombination experiments using SVJ129 mouse strain embryonic stem (ES) cells. A higher targeting efficiency will be found when DNA from the same strain as the ES cells into which it is transfected, is used. This is due to the high levels of intra-strain DNA homology. The first library to be screened by aqueous hybridisation and successfully to yield PPT-A clones was a library supplied courtesy of Dr. Francois Conquet, Head of the Transgenics Unit at the Glaxo Institute of Molecular Biology in Geneva, Switzerland. This library was constructed so as to include restricted mouse genomic DNA inserts of between 9 kb and 20 kb. The genomic DNA was restricted with BamHI prior to size selection and vector/insert ligation. The bacteriophage vector was λFIX (Sambrook *et al,* 1988). The second library to be screened successfully was supplied courtesy of Dr. Andrew Smith at the MRC Centre in Cambridge. This library was also constructed using SVJ129 strain mouse genomic DNA, with insert fragments of between 9 Kb and 20 Kb cloned into the lambda vector λ2001 (Sambrook *et al,* 1988). Libraries were supplied as high titre liquid bacteriophage suspensions ($>10^{10}$ pfuμ l^{-1}).

3.3.2 Titre Calculation, Plating and Storage of Libraries

The concentration of bacteriophage in the original library master stock was determined by titration. Having accurately determined this, one million bacteriophage were plated as described in Section 2.12. The quantity of genomic DNA in one million library

bacteriophage inserts was at least equivalent to one whole mouse genome. The probability of cloning any selected mouse genomic sequence from this number of bacteriophage is therefore high. One hundred thousand (10^5) bacteriophage were plated on each of ten 245mm x 245mm plates. Each batch of 10^5 phage were mixed with 2ml of LE392 cells, incubated at 37°C for 30 min, and added to 30 ml of pre-warmed NZCYM top agar. This mixture was poured on to a plate and allowed to set prior to incubation overnight at 37°C. Four replica lifts using Hybond-N™ membrane (see section 2.14) were taken from each plate (two for immediate use, and two for storage). When lifts had been taken, the plates were overlayed with 3MM paper soaked in 25% (v/v) glycerol, wrapped in parafilm and stored at -70°C.

3.3.3 PPT-A Screen of the Glaxo Library

3.3.3.1 Primary PPT-A Screen of the Glaxo Library

Two of the four identical replica filters taken from each of the ten bacteriophage plates were pre-hybridised for 4 hrs with 10 ml of hybridisation buffer in revolving hybridisation flasks (see section 2.29). The filters were divided so that one from each pair of replica membranes was placed in each of the two flasks. Sufficient probe was labelled (and denatured) to generate 10 ml of hybridisation buffer containing 100 ng of labelled probe per ml. 5 ml of probe/buffer solution was then added to each of the two flasks and hybridised. A series of mouse genomic DNA digests which had been Southern blotted onto Hybond-N membrane and had previously been probed successfully were used as a positive control. 1µl of 100 $ng.\mu l^{-1}$ denatured, wild-type bacteriophage lambda DNA was spotted on to a piece of Hybond-N membrane, as was 1µl of a 10^{-1}, 10^{-2} and a 10^{-3} dilution of this DNA, and then fixed to the membrane by UV cross-linking. Two identical filters like this were made, one for each flask. These 'dot-blot' filters were used as the negative control for this experiment, and were placed in the flasks along with the replica and the positive control filters. After hybridisation, the filters were washed as

described in section 3.2.2, and placed at -70°C with autoradiographic film for 3 days. On development, the films indicated that 6 of the 10 filters (numbers 2, 5, 6, 7, 8 and 10) displayed clear positive signals duplicated at the same location on the equivalent partner filter (hybridised in a separate flask).

Figur 3.9: PPT-A Bacteriophage Clones Isolated from Primary Glaxo Library Screen

Clone Number	Pair of Library Filters	Number of Duplicate Clones	Strength of Signal
1	No. 2	1	Medium
2	No. 5	1	Strong
3	No. 6	1	Faint
4	No. 7	1	Very Strong
5	No. 8	1	Medium
6 &7	No. 10	2	Both Faint

Figure 3.9: The number and the strength of the hybridisation signals generated by PPT-A bacteriophage clones isolated from the Glaxo library in the primary screen.

In total, seven distinct (duplicated) locations on six different master plates demonstrated positive PPT-A hybridisation signals due to this primary library screen. Using the orientation marks on the filters and the autoradiographs, the corresponding areas on the master plates, from which the hybridisation signals were emanating, were located. A zone of approximately 1 cm^2, encompassing the proposed source of the signal, was then cut from the master plate using a razor blade. This area was estimated to contain 50 distinct bacteriophage plaques (each with potentially a different insert), any of which could be the single clone containing the PPT-A sequence. This isolate, which contained frozen NZCYM medium, bacterial cells and bacteriophage, was placed in 1 ml of phage buffer and incubated at 4 °C overnight to resuspend the bacteriophage in solution. The low temperature minimised bacterial growth.

3.3.3.2 Secondary PPT-A Screen of the Glaxo Library

The resuspended bacteriophage from the seven master plate isolates were titred as before. An appropriate volume of the bacteriophage suspension was plated on 125 mm x 125 mm NZCYM plates so as to generate two hundred phage plaques per plate. Two replica filter lifts were taken from each plate. The filters were probed independently in the presence of positive and negative control filters (Genomic DNA Southern blots and non-homologous plasmid dot-blots, respectively). Of the seven primary clones isolated and screened, only five displayed duplicate hybridisation signals at the secondary stage. The signals varied in strength and some non-specific background signal was noted. Positive signals emanating from the same location on duplicate filters were selected for further, tertiary screening and were isolated.

3.3.3.3 Tertiary PPT-A Screen of the Glaxo Library

The clones isolated from the secondary screen were resuspended and titred. The 5 clones were plated again, but at a lower density to ensure a phage-free zone of about 0.5 cm between each plaque. For each clone, 100 plaques were grown on a plate. The hybrid-isation was repeated as before. All five were successfully identified in the tertiary screeen and isolated accordingly.

3.3.3.4 Quaternary PPT-A Screen of the Glaxo Library

A quaternary hybridisation was conducted to ensure the purity of the isolated clones. Regarding the density of plaques on the tertiary plates, the proximity of the plaques to each other was thought, with hindsight, to be too close to guarantee the isolation of single plaques. This was also a concern because of possible lateral diffusion of bacteriophage through the medium from one plaque to another. In order to ensure the purity of an isolated plaque, the five isolated putative clones from the tertiary screen were re-plated at the very low density of twenty plaques per (small square) plate. Duplicate lifts from each plate were probed as before. Positive, duplicate hybridisation signals were generated by all five clones. On

one of the five quaternary phage plates, every one of the twenty plaques on the plate produced a strong signal (see Figure 3.11 below). A single plaque at least 2 cm from the nearest neighbouring plaque was isolated from each plate. The isolated clones were named λPPT1, λPPT2, λPPT3, λPPT4 and λPPT5.

To ensure that the five purified clones were retained indefinitely, two 4°C (liquid) and two -70°C (frozen) stocks were made for each clone (see section 2.13). The original phage 'plug' isolated from the quaternary plate was also stored at 4°C in the presence of a small amount of chloroform should the re-isolation of the clone become necessary. The original phage 'plugs' isolated from the primary, secondary and tertiary plates described previously were stored at 4°C in the same way.

Figure 3.10 : Duplicate Filters from Tertiary PPT-A Screen of Glaxo Library.

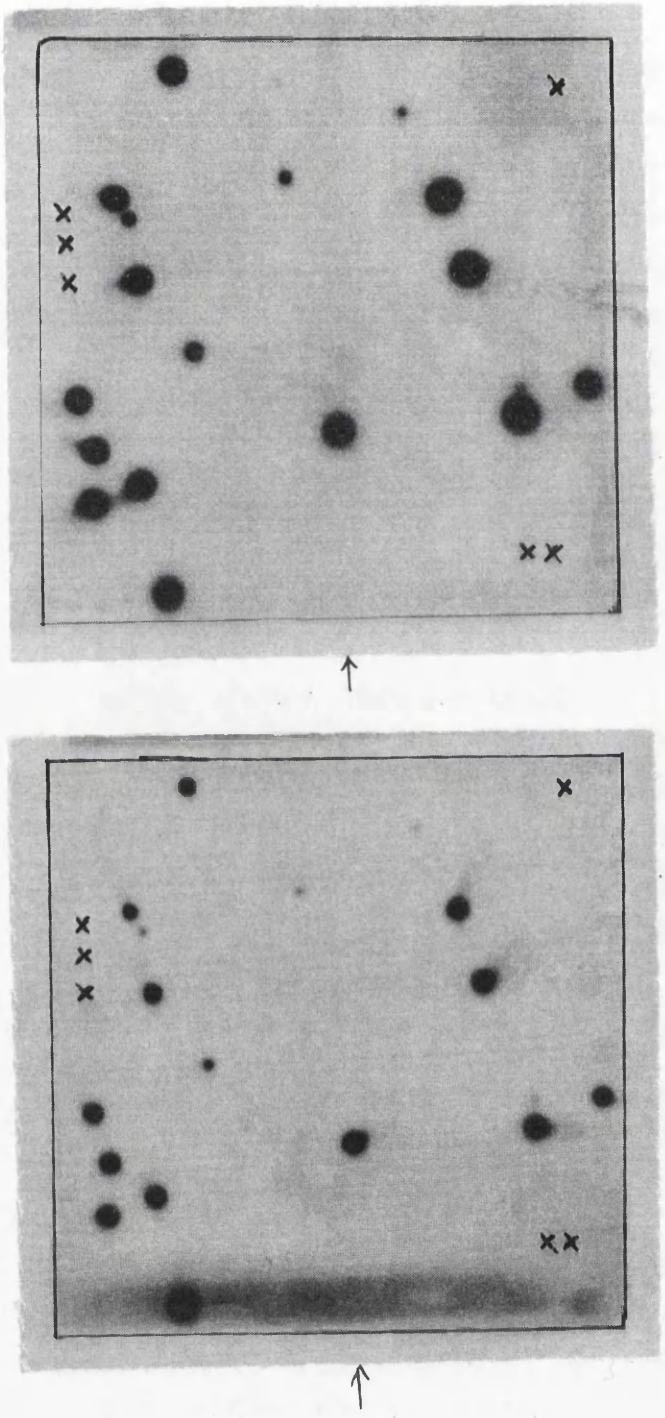


Figure 3.10: A pair of duplicate filters from the tertiary hybridisation of the candidate bacteriophage clones (from the Glaxo library) probed with the PPT-A-specific probe. A single duplicate hybridisation signal is located in the centre of the filters. This plaque was isolated and was purified further by way of a quaternary hybridisation screen (see figure 3.11). 2 day exposure at -70°C.

3.3.4 PPT-A Screen of the Cambridge Library

Each plaque isolated from the Glaxo library at the quaternary stage was used to generate a 100 ml bacteriophage lysate containing only one type of bacteriophage. From the lysate, DNA was prepared for use in the mouse PPT-A gene characterisation experiments described in chapter 4. (The characterisation of the NK-1R gene is also described in chapter 4.) However, at this stage, it suffices to say that early characterisation experiments indicated that of the five clones isolated, none contained exon 1 or the rest of the 5' region of the mouse PPT-A gene upstream of exon 3. Exons 3 through to 7 were however included in one or other of the isolated phage clones. To characterise the mouse PPT-A gene adequately, it was deemed necessary to clone the whole gene and certainly all of the seven (expected) exons.

To remedy this, another library, the Cambridge library described in section 3.3.1, was screened with the same PCR-derived PPT-A probe used for the Glaxo library screens. The probe was labelled by the 'specific priming' method (section 3.2.1). The Cambridge library was titred, plated and stored in the same way as the Glaxo library (see section 3.3.2). Twenty pairs of duplicate filter lifts were made from the Cambridge library master plates (small square plates). They were used in the following hybridisations.

3.3.4.1 Primary PPT-A Screen of the Cambridge Library

The twenty pairs of Cambridge library filters were separated from their partner filters and prehybridised independently for 6 hrs at 60°C in two flasks. Labelled probe was denatured and added to 10 ml of hybridisation buffer to a concentration of 100 ng of labelled probe per ml. 5 ml of probe/buffer solution was then added to each of the two flasks and hybridised. A series of mouse genomic Southern filters were used as positive controls. The filters were washed as before and incubated at -70°C with autoradiographic film for three days. From the twenty filter pairs, eleven candidate clones were identified which demonstrated hybridisation signals emanating from the same approximate location

on both filters. The clones were isolated from the master plates and titred as before.

Figure 3.11: PPT-A Clones Isolated from Primary Cambridge Library Screen

Clone Number	Pair of Library Filters	Number of Duplicate Clones	Strength of Signal
1 &2	No. 8	2	Both faint
3 &4	No.11	2	Strong. Faint
5	No.12	1	Strong
6	No.15	1	Medium
7 &8	No.17	2	Strong. Faint
9, 10 &11	No.19	3	2 Strong. Faint

Figure 3.11: The number and strength of signal generated by the PPT-A bacteriophage clones isolated from the Cambridge library in the primary screen.

3.3.4.2 Secondary PPT-A Screen of the Cambridge Library

The secondary screen of the Cambridge library was similar to that of the Glaxo library (prehybridise for 4 hours, hybridise for 10 hours). The filters were washed and incubated with autoradiographic film for 4 days. Of the 11 primary candidate clones, 6 were successfully identified at the secondary stage.

3.3.4.3 Tertiary PPT-A Screen of the Cambridge Library

Five of the 6 clones from the secondary screen generated strong, duplicate tertiary hybridisation signals. Single plaques were isolated and master stocks made for each (section 2.13). The 5 clones isolated were named λPPT6 through to λPPT10. The 5 clones isolated from the Glaxo library together with the five isolated from the Cambridge library were used in the characterisation of the mouse PPT-A gene as described in Chapter 4 and to generate targeting vectors as described in Chapter 5.

Figure 3.12 : Duplicate Filters from the Primary, Secondary and Tertiary PPT-A Screens of the Cambridge Library

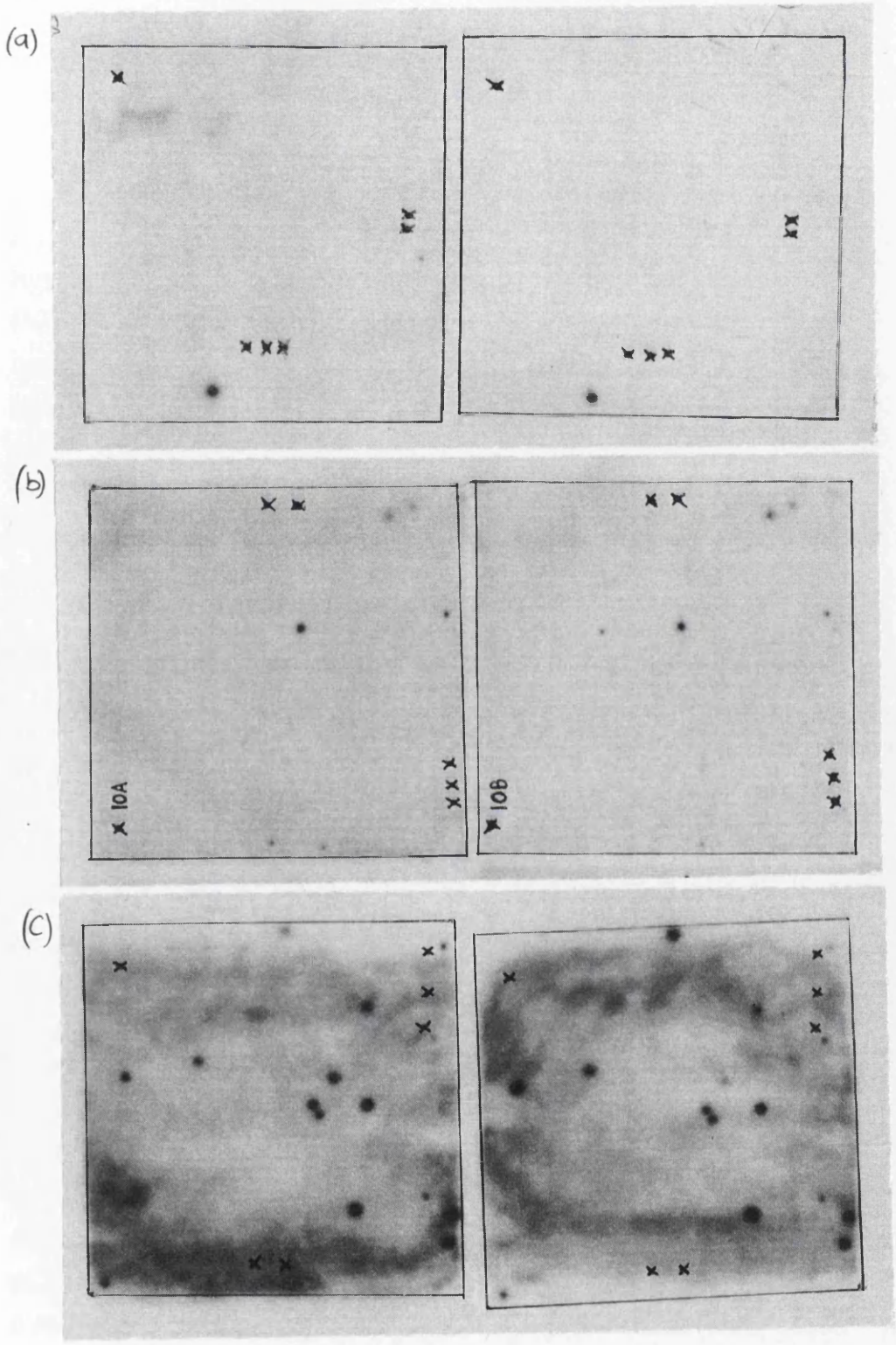


Figure 3.12: Duplicate filters from (a) primary, (b) secondary and (c) tertiary hybridisations of the Cambridge library with the PPT-A probe. Exposure times: for the primary, 4 days; secondary, 2 days; tertiary, 24 hrs.

CLONING THE MURINE NK-1R GENE

3.4 DESIGN OF NK-1R PROBE FOR GENOMIC LIBRARY SCREENS

In order to isolate clones containing genomic NK-1R sequence from a bacteriophage lambda library, a suitable probe for use in hybridisation experiments was required. The previously described, PCR-based approach used to generate the NK-1 probe was not necessary because a suitable NK-1R cDNA cloned from another species (rat) was available. The NK-1R probe was derived from a plasmid named prTKR2 which was supplied courtesy of Dr. Shigetada Nakanishi at the University of Kyoto, Japan. The 6.4 Kb prTKR2 plasmid constituted a 3408 bp rat NK-1R gene cDNA insert cloned into the 2.9 Kb plasmid vector pBluescript SK(+) (Yokoto *et al*, 1989).

3.4.1 Sequence analysis of the Rat NK-1R cDNA Sub-clone prTKR2

The first step in choosing a particular probe, based on the cDNA sequence and the endonuclease restriction profile of that sequence, was to determine the authenticity of the prTKR2 insert by sequence analysis. The plasmid was prepared for manual sequencing as a double stranded template as described in section 2.30. Using the chain termination sequencing method (Sanger, 1975) and the 17 nucleotide deoxyoligonucleotide primer T7 which anneals to the polylinker of pBluescript SK(+), 251 bp of sequence was obtained from prTKR2 and is shown below. The sequence was then processed by a 'FASTA' search of the GENEMBL database. The prTKR2 sequence obtained demonstrated a 98.4% identity to the published rat NK-1R cDNA sequence. This high percentage identity was demonstrated over a 255 bp sequence where the four mis-matched nucleotides in the prTKR2 sequence was attributed to sequencing error. That sequence and its alignment with the published sequence is shown below. The authenticity of the prTKR2 plasmid was verified by the sequence analysis.

Figure 3.13 : prTKR2 Sequence Aligned with Rat NK-1R cDNA.

```
                                                    10        20        30
prTKR2/T7                                          CTGGCAGGTGAGAATGCTGTTCCTTTCCGG
                                                   |||||||||||||||||||||||||||| |
Ratspr       ACTTTGCATCCTGAGATGGCTTCAGGACCCCTGGCAGGTGAGAATGCTGTTCCTTTCC-G
           1780       1790       1800       1810       1820       1830

                    40         50         60         70         80         90
prTKR2/T7    TCTCAGATGTGACTGGAAGCGAAAGCTTCATGTGACAT+CGTGAACATCCTCATGGTTCA
             ||||||||+|||||||||+|||||||||+||||||||||||||||||+||N|||||||+|
Ratspr       TCTCAGATGTGACTGGAAGCGAAAGCTTCATGTGACATCCGTGAACATCCTCATGGTTCA
           1840       1850       1860       1870       1880       1890

                  100        110        120        130        140        150
prTKR2/T7    GATAATTGGATGCACAAAAGCTCTATTCATCGTGTCAGAAAAAACAAGGCNACCCACCAC
             ||||||||+|||||||||+|||||||||+|||||||||+|||||||||+|:|||||||+|
Ratspr       GATAATTGGATGCACAAAAGCTCTATTCATCGTGTCAGAAAAAACAAGGCCACCCACCAC
           1900       1910       1920       1930       1940       1950

                  160        170        180        190        200
prTKR2/T     CAGGATTTATAACTTCTGCCAGCCTTGTTGAATCAAACATTTTGAGGACATCGGCAT--C
             ||||||||+|||||||||+|||||||||+|||||||||+|||||||||+|||||||| |
Ratspr       CAGGATTTATAACTTCTGCCAGCCTTGTTGAATCAAACATTTTGAGGACATCGGCATACC
           1960       1970       1980       1990       2000       2010

           210        220        230        240        250
prTKR2/T7    TGGAGAAGGCAAGATTTATACAGAAGTAATATAAACCAGAAA
             ||||||||||+|||||||||+|||||||||+|||||||||+|
Ratspr       TGGAGAAGGCAAGATTTATACAGAAGTAATATAAACCAGAAAAAAAGAAAAGGAAAAGAA

           2020       2030       2040       2050       2060       2070
```

Figure 3.13: 251 bp of sequence from rat cDNA plasmid subclone prTKR2 obtained using the T7 sequencing primer by manual sequencing. Alignment of that sequence with the published rat NK-1R cDNA sequence stored in the GENEMBL database. Letters in bold denote comparative differences most likely due to sequencing artefacts. Sequence identity was found to be 98.4%.

3.4.2 Restriction Endonuclease 'Mapplot' of prTKR2

A fragment of the 3.4 kb prTKR2 insert was selected for use as the NK-1R probe. The chosen fragment was selected so as to comply with several criteria. The probe fragment came from the protein coding region of the (rat NK-1R) cDNA. Within this region, the genomic sequence is most highly conserved between species. Since the effectiveness of the probe is dependent on the degree of homology between the probe and the target sequence, this is an important factor to consider when selecting a probe sequence. Since the probe for the library hybridisation was to be labelled using the random primer extension DNA labelling method, the fragment should be within the recommended size range (see section 2.24.1). The probe fragment was conveniently released from the rest of the cDNA insert by restriction with only one endonuclease. To help in the selection of a suitable probe fragment, a linear 'Mapplot' of the full rat NK-1R cDNA sequence (GENEMBL access number J09057.Gb_Ro) was generated with the help of GCG software for the Viglen computer. This produced a linear representation of the cDNA sequence onto which the relative cleavage sites of several restriction endonucleases (individually) were added. It became clear by visually scanning the restriction patterns which endonuclease generated the most appropriate probe fragment.

3.4.3 Selection of the NK-1R Probe Fragment

From the rat cDNA mapplot, several candidate probe fragments were considered. The fragment finally chosen was a 965 bp *Bsg* I-*Bsg* I fragment which originates wholly from within the protein coding region of the rat NK-1R cDNA. The restriction endonuclease *Bsg* I has a seven base pair recognition sequence, and cleaves eukaryotic DNA relatively infrequently. This enzyme cleaves the 3408 bp rat NK-1R cDNA only five times. The 965 bp fragment was the largest generated by this restriction where the next largest fragment was 350 bp. The chosen 965 bp fragment did comply with the aforementioned criteria. It was therefore considered appropriate

for use as the NK-1R probe in hybridisations to identify mouse NK-1R clones from a bacteriophage lambda genomic library.

3.5 PREPARATION OF NK-1R PROBE FOR GENOMIC LIBRARY SCREENS

3.5.1 Purification of the NK-1R Probe Fragment.

The *Bsg* I restriction pattern of plasmid prTKR2 predicted by the mapplot was confirmed by digestion. The restriction reaction was electrophoresed on a 1% agarose gel and the band sizes were determined. When the restriction pattern generated by the small scale reaction was deemed satisfactory, the digestion was scaled up so that 20 µg of prTKR2 was restricted in one reaction. The large scale reaction was electrophoresed

The 965 bp fragment was isolated from the gel with a scalpel and the DNA recovered from the gel
method (see section 2.27.2). After precipitation, the recovered DNA was resuspended in dH_2O to a concentration of 50 ng.µl as determined by a spectrophotometer and was ready for labelling.

3.5.2 Radioactively-Labelling the NK-1R Probe

The purified 965 bp NK-1R probe fragment was labelled using the random primer extension method (section 2.24.1). Random deoxyoligonucleotides were used as primers where the DNA was labelled to a specific activity of 1.8 x 10^9 dpm/µg. Prior to use, the labelled probe was denatured

3.5.3 Test Hybridisations with NK-1R Probe

As a means of determining the proficiency of the labelled NK-1R probe to hybridise to NK-1R sequences and differentiate between those and non-NK-1R sequences, a test hybridisation was conducted. This experiment tested whether or not the hybridisation conditions which had been optimised previously for the NK-1 library screen (section 3.3.2) would be suitable for the NK-1R probe. DNA which was known to contain NK-1R sequence (prTKR2 plasmid, above) and DNA which was known *not* to contain NK-1R sequence (plasmid pCOL1, section 3.1.6.1) was spotted on to a Hybond-N™ nylon membrane. 1 µl samples from each of a series of DNA dilutions, ranging from 100 ng/µl down to 0.1 ng/µl, were denatured and spotted onto the membrane and fixed by baking (a "dot blot"). The probe was 100 % homologous to, and was therefore expected to hybridise to, the prTKR2 DNA whereas the pCOL1 DNA acted as the negative control.

The dot blot was pre-hybridised for 4 hours. 50 ng of the 965 bp prTKR2/*Bsg* I DNA fragment was labelled as described above. The probe was added to 8 ml of hybridisation buffer to a concentration of 50 ng.ml. The hybridisation was incubated at 60 °C and washed at the same stringency as for the NK-1 library screens described previously (section 3.2). The membrane was incubated with autoradiographic film for 2 days. When developed, the film clearly demonstrated that the probe positively hybridised to the DNA containing NK-1R sequence down to the 0.1 ng/µl level, but not to the DNA lacking NK-1R sequence at any level. The hybridisation therefore demonstrated that the labelled NK-1R probe could distinguish specifically between NK-1R and non-NK-1R sequences under these hybridisation conditions. It was not necessary to optimise a set of hybridisation and washing conditions specifically for the NK-1R probe. Consequently, the hybridisation and washing conditions used to screen the Glaxo and Cambridge libraries for NK-1 clones (sections 3.3.3 and 3.3.4) were used in the following hybridisations undertaken to identify NK-1R clones from the Toronto library.

3.6 ISOLATION OF NK-1R GENOMIC LIBRARY CLONES

3.6.1 Source and Construction of NK-1R Probed Library

Genomic mouse neurokinin-1 receptor clones were isolated from a bacteriophage lambda library. The library was constructed from strain SVJ129 mouse genomic DNA so that the cloned DNA would be compatible for use in homologous recombination experiments using SVJ129 mouse strain embryonic stem (ES) cells. The library which was screened by aqueous hybridisation and successfully yielded NK-1R clones was supplied courtesy of Dr. Alexandra Joyner, Toronto, Canada. This library was constructed so as to include restricted mouse genomic DNA inserts of between 9 Kb and 20 Kb. The genomic DNA was partially restricted with Sau3A1 prior to size selection and vector/insert ligation into the bacteriophage vector was λDASH. The 'Toronto' library was supplied as a high titre liquid bacteriophage suspension.

3.6.2 Titre Calculation, Plating and Storage of Toronto Library

The concentration of bacteriophage in the original Toronto library master stock was determined by titration, as previously described for the libraries used to isolate NK-1 precursor clones (section 3.3.2). One million bacteriophage were plated on 20 small square NZCYM agar plates (50,000 per plate). Four replica lifts, using Hybond-N™ membrane, were then taken from the each plate.

3.6.3 NK-1R Screen of the Toronto Library

3.6.3.1 Primary NK-1R Screen of the Toronto Library

A pair of identical replica filters were taken from each of the twenty bacteriophage library plates. Six of these pairs were probed initially. They were pre-hybridised for 4 hrs (see section 2.29). The filter pairs were split and probed independently. Labelled and denatured NK-1R probe was added to a probe concentration of 100 ng per ml. 5 ml of probe/buffer solution was then added to each of the two flasks and hybridised. A series of 'dot-blotted' prTKR2 and pCOL1 plasmid dilutions were used as positive and negative controls respectively for this hybridisation. After hybridisation, the filters were washed as described in section 3.2.2, and incubated with autoradiographic film. On developing the film, strong and weak hybridisation signals, all of which were duplicated on their partner filter, were identified. The signal from non-specific background hybridisation was minimal. The areas on the six library master plates from which the duplicate signals were emanating were located, isolated, resuspended and stored as described previously (section 3.3.3.1). In order to identify as many NK-1R candidate clones as possible, the 14 remaining pairs of Toronto library filters which had not been probed, were probed in exactly the same way as for the initial six pairs of filters.

In total 186 candidate bacteriophage NK-1R clones were identified as a consequence of the primary screen. Despite identical exposure times, the strength of the hybridisation signals generated by the numerous clones varied considerably. This can be seen below on the autoradiograph (See Fig 3.14). All 186 clones were isolated from the twenty master plates and stored as 4 °C stocks. On average, each plate displayed 8 duplicate candidate clones.

3.6.3.2 Secondary NK-1R Screen of the Toronto Library

Of the 186 candidate clones identified at the primary stage, 20 were selected for further analysis and were screened again by aqueous hybridisation. One clone was selected from each of the 20 master plates. Clones which generated both strong and weak hybridisation signals were chosen so as to represent the range in the strength of signal noted at the primary stage. Since the NK-1R probe was derived from the rat NK-1R cDNA, the probe may well have hybridised to bacteriophage clones containing one or more mouse NK-1R exon. Theoretically the greater the number of mouse NK-1R exons contained in a bacteriophage insert, the greater the amount of target sequence present for the probe to hybridise to. Therefore the number of exons may account for the variation in signal strength. The approach of selecting a range of clones was adopted to ensure that as many exons as possible were contained in the 20 clones screened at the secondary stage.

The 20 clones were titred, plated and replica membrane lifts were taken from the bacteriophage plates as described (Section 3.3.3.1). The bacteriophage were plated at the density of 100 plaques per plate. The pre-hybridisation and the probe labelling were the same as for the Toronto library primary screen (above). The hybridisation was incubated at 60 °C for 18 hrs. The filters were washed and incubated with autoradiographic film for 2 days. Of the 20 clones probed at the primary level, 19 were identified at the secondary stage and underwent the tertiary hybridisation.

3.6.3.3 Tertiary NK-1R Screen of the Toronto Library

The 19 isolated clones were resuspended, titred and plated again but at the lower density of 50 plaques per plate. The NK-1R hybridisation was repeated as before. The exposure time was 4 days. Of the 19 NK-1R clones isolated from the secondary screen, 14 were identified by the tertiary hybridisation. The clones were isolated from the master plates and 4 °C and -70 °C stocks were made for each of the bacteriophage clones (which were named λSPR1 through to λ SPR14).

Figure 3.14 : Filters from Toronto Library/NK-1R Probe Hybridisations.

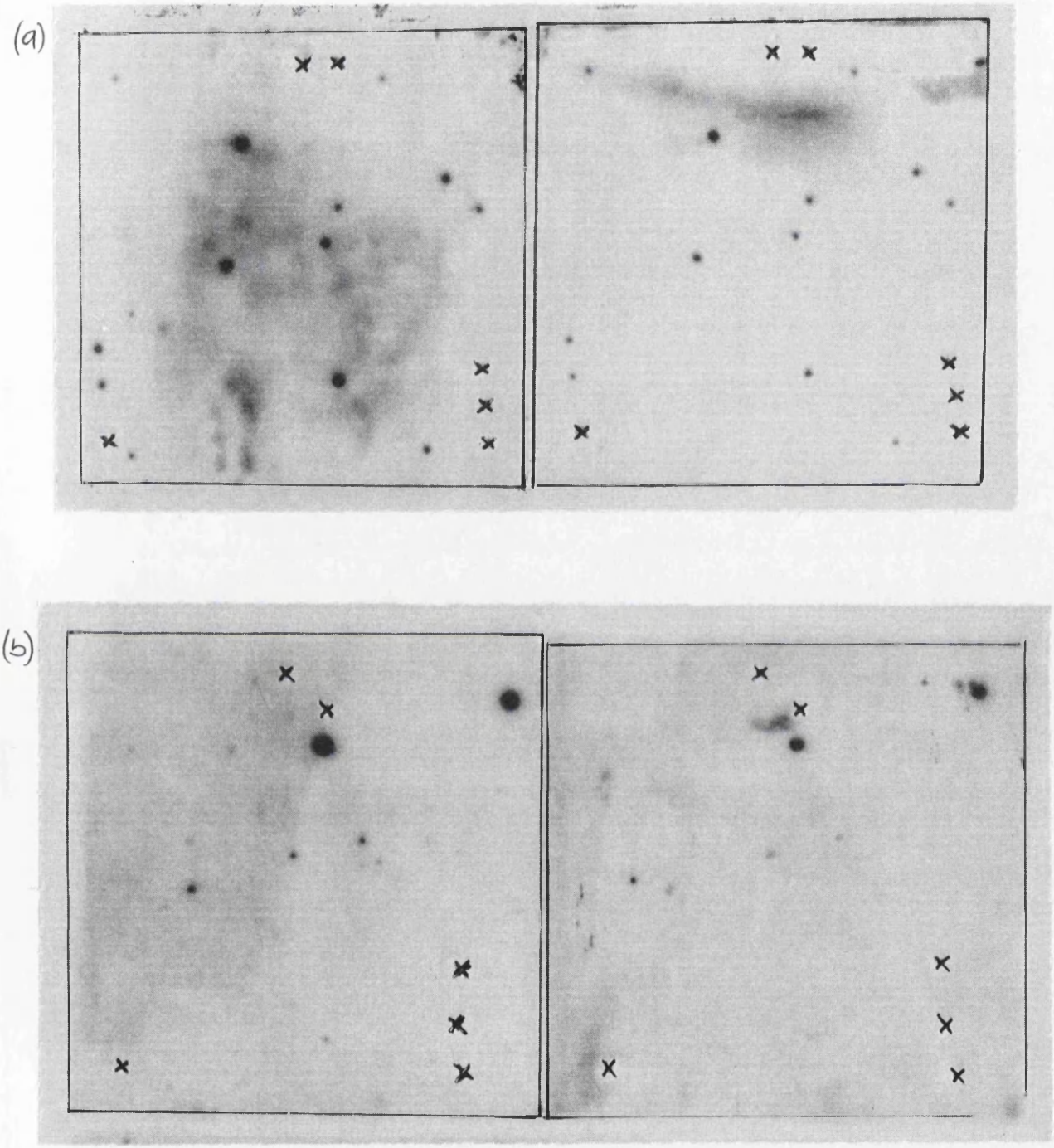


Figure 3.14: Pairs of filters from (a) primary hybridisation, and (b) secondary hybridisation of the Toronto bacteriophage library with the NK-1R-specific probe. The duplicate hybridisations signals are clear on the partner filters.

3.7 CLONING THE MOUSE NK-1R EXON 2

Preliminary characterisation of the NK-1R bacteriophage clones, as described in depth in Chapter 4, indicated that each of them contained at least one NK-1R exon. However, not one of the 14 positive NK-1R clones contained the mouse NK-1R exon 2. In rat, the NK-1R gene has 5 exons, where exons 3, 4 and 5 are all to be found within a 3 kb genomic region while exon 2 is separated from exons 1 and 3 by introns of approximately 15 and 23 kb respectively (Hershey, 1991). Bearing in mind that there is on average 95% identity between the rat and mouse species at the genetic level, it was thought likely that the mouse NK-1R gene also had 5 exons and that the mouse exon 2 would be flanked by introns of a comparable size to that in the rat. This was later confirmed to be the case (Sundelin *et al*, 1992), and at the time was considered a reasonable explanation why none of the 14 mouse NK-1R bacteriophage clones contained exon 2.

For the mouse NK-1R gene, it was considered important to clone the mouse NK-1R exon 2 because while the the other four were subcloned, this exon remained only in a bacteriophage clone. The exon 2-specific oligodeoxynucleotide probe (CS10) shown above was therefore used to screen the 186 NK-1R candidate clones which had been isolated from the Toronto library as a result of the primary library screen which used the prTKR2/*Bsg* I 965 bp rat cDNA fragment as probe. Of the 186 isolated bacteriophage plaques, 20 had been screened at the secondary stage and had been shown to lack NK-1R exon 2 and were not screened again. Of the remaining 166 primary clones, 38 were chosen randomly to be further screened for mouse NK-1R exon 2 sequence.

3.7.1 NK-1R Exon 2 Oligodeoxynucleotide Probe Design

All 14 clones lacked exon 2 as determined by hybridisation using a oligodeoxynucleotide probe known to be 100% homologous to the mouse NK-1R exon 2 (see Figure 3.12). This probe (CS10) was used in the NK-1R exon 2-specific hybridisations described here. The

design of this probe was based on the known mouse exon 2 sequence which was obtained from Sundelin *et al,* (1992).

Figure 3.15 : Sequence of Murine NK-1R cDNA from which Oligodeoxynucleotide CS10 was Designed.

(a) **Primer CS10**

```
5'ATACATGGCCATCATCCACCCTCTTCAGCCCCGGCT 3' (contd.to 194 bp)
  |||||||||+||||||||||||||||||||||||||
  TATGTACCGGTAGTAGGTGGGAGAAGTCGGGGCCGA
  Î
  Start of mouse SPR exon 2 sequence.
```

Figure 3.15: (a) 37 bp of the mouse NK-1R exon 2 (194 bp) sequence indicating the sequence from which the oligodeoxynucleotide CS10 was derived. CS10 was used as a probe to identify genomic sequences containing the mouse NK-1R exon 2. The sequence of primer CS10 is shown in bold.

3.7.2 Hybridisations using the Exon 2-Specific Probe

All of the hybridisations which used the deoxyoligonucleotide CS10 as the probe were conducted in accordance with the protocol described in section 2.29.2. The hybridisation temperature was 45 °C and the probe was radioactively-labelled using polynucleotide kinase (section 2.24.3).

The 186 plaques which had been isolated from the library master plates after the primary screen had each been stored, in the interim, in 1 ml of bacteriophage buffer (see solutions) at 4 °C. Under these conditions for several weeks, the bacteriophage had resuspended. Each of the 38 randomly selected primary clones were plated on small square NZCYM plates and incubated at 37°C overnight. The next day, the plates were chilled at 4°C before replica filter lifts were taken from each.

3.7.3 Re-Probing the Primary Toronto Library Screen Clones

The replica lift filters taken from the 38 re-plated primary screen clones were pre-hybridised for 1 hr and hybridised for 3 hrs. This hybridisation was equivalent to a secondary screen, as all 38 clones had been screened previously all be it with the cDNA-derived, rather than the oligodeoxynucleotide probe. The washed filters were incubated with autoradiographic film for 4 days. Of the 38 clones re-plated and re-screened with the exon 2-specific probe, only 2 (clones number 17 and 32) generated duplicate signals and were analysed further.

For the final, tertiary screen, clones 17 and 32 were plaque isolated from the secondary plates, resuspended, titred and re-plated at low plaque density (50 plaques per plate) as described before. Replica membrane lifts were made which were probed with the exon 2-specific probe again. The filters were washed and incubated with film overnight. Both clones were successfully detected at the tertiary stage. Single plaques were isolated

The two clones containing mouse NK-1R exon 2 sequence were called λSPR15 and λSPR16.

The 14 clones isolated from the Toronto library by the original cDNA-derived probe, together with the two exon 2-containing clones which were isolated later from the same library, were used in the characterisation of the mouse NK-1R gene as described in Chapter 4. (They were named λSPR1 through to λSPR16). They were also used in the design and construction of homologous recombination targeting vectors for the mouse NK-1R gene as described in Chapter 5.

The following text, chapter 4, describes the characterisation of the murine NK-1 and NK-1R genes.

Figure 3.16 : Linear Alignment of PPT-A Clones Showing the Overlap and the Exons Contained in Each Clone.

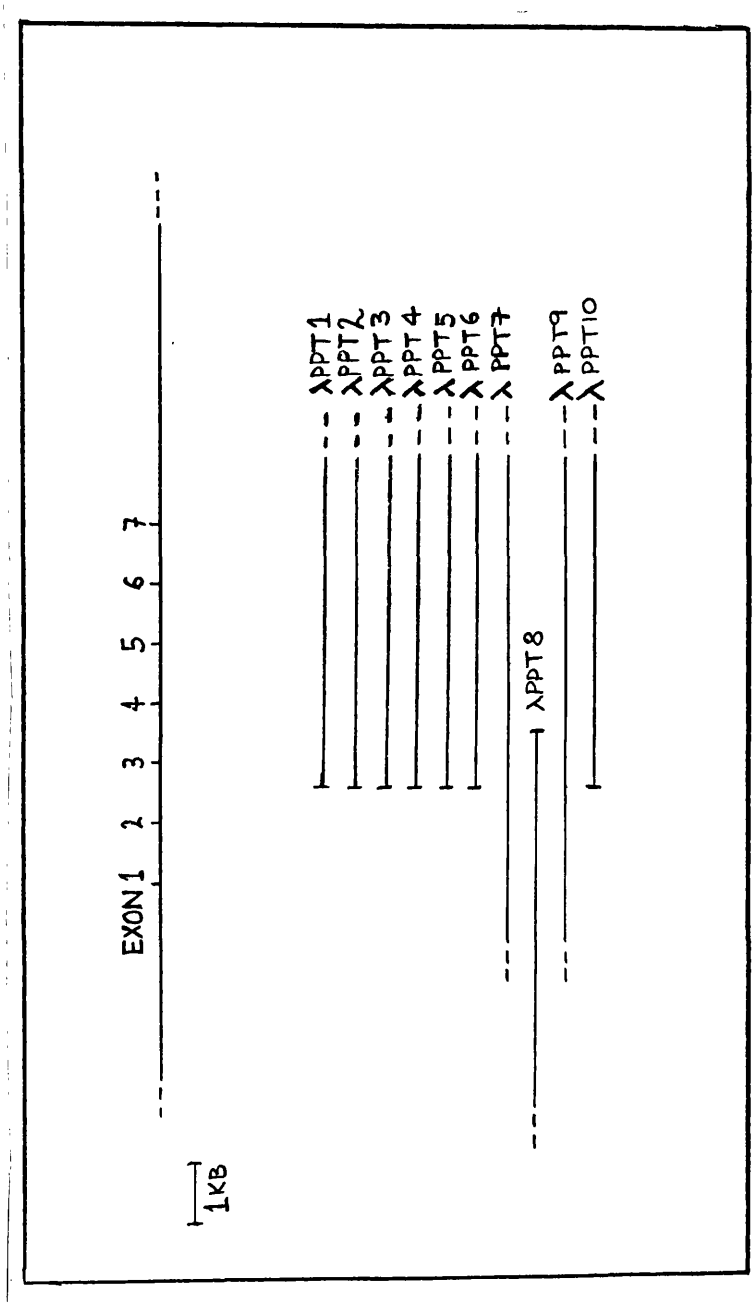


Figure 3.16: Linear schematic alignment of the NK-1 precursor bacteriophage clones to demonstrate the extent of overlap and the exons contained in each clone. The bacteriophage clones are represented relative to the publihed genomic DNA organisation of the rat NK-1 gene (shown at the top, Carter *et al*, 1990).

Figure 3.17 : Linear Alignment of NK-1R Clones Showing the Overlap and the Exons Contained in Each Clone.

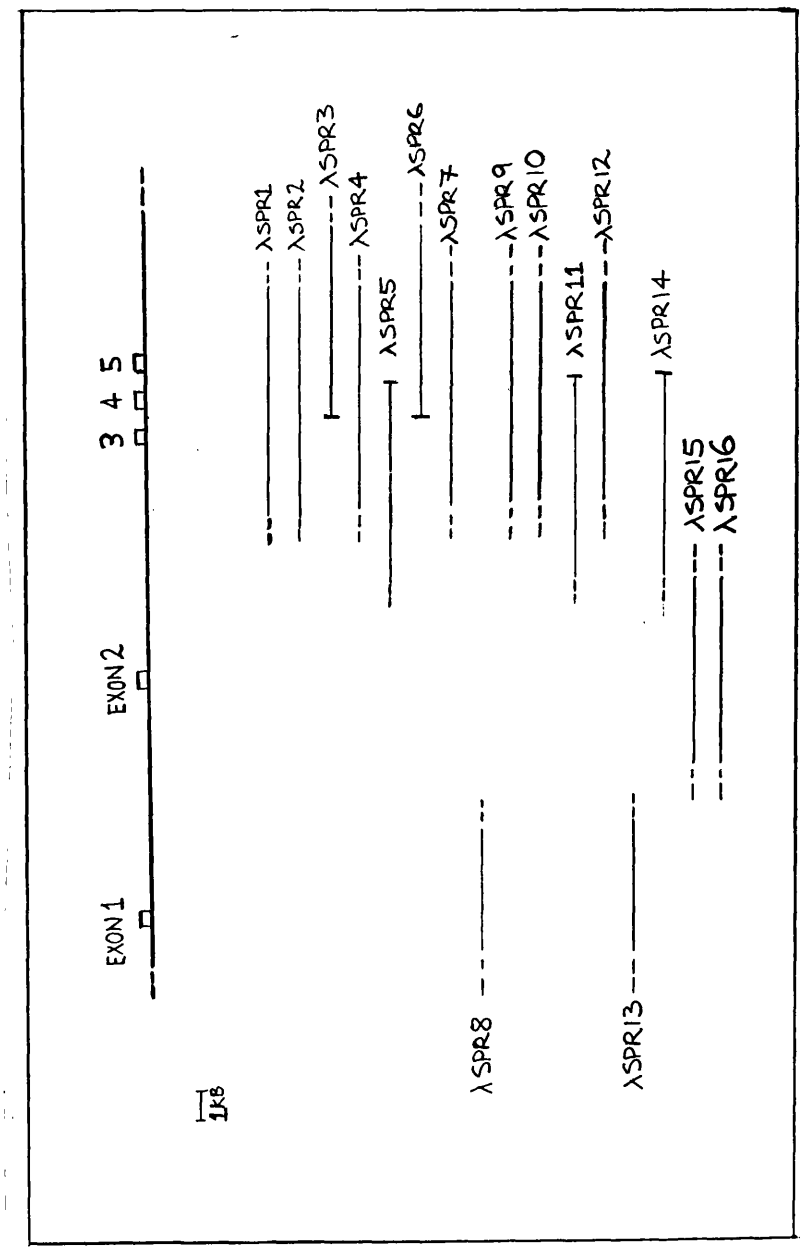


Figure 3.17: Linear alignment of the NK-1R receptor bacteriophage clones to demonstrate the extent of overlap and the exons contained in each clone. The bacteriophage are represented relative to the published genomic DNA organisation of the rat NK-1R gene (shown at the top of the figure, Sundelin *et al*, 1992).

CHAPTER 4

RESULTS: CHARACTERISATION

INTRODUCTION

The goal of this chapter is to describe the characterisation of the genomic structures of the mouse PPT-A and NK-1R genes. The purpose of this was to establish the authenticity of the genomic DNAs cloned in Chapter 3 and which were believed to contain the mouse PPT-A and NK-1R genes. To do this, the positively-hybridising mouse PPT-A and NK-1R bacteriophage clones were analysed in several ways. Based on that analysis, the generation of exon-containing plasmid subclones is described. The aim of generating such subclones was to facilitate the molecular characterisation of the mouse PPT-A and NK-1R genes to as full a degree as possible. Using those subclones, the relative genomic location of the exons and some exonic sequence was determined in addition to the generation of a map of the restriction sites within each subclone. This procedure is described for each gene individually, starting with the PPT-A precursor gene. The results detailed in this chapter were subsequently used in the design of the gene targeting vectors which are described in Chapter 5 and are required for homologous recombination experiments.

4.1 ANALYSIS OF THE PPT-A/GLAXO LIBRARY CLONES

The five positive PPT-A bacteriophage clones which had been isolated from the Glaxo genomic DNA library were named λPPT1 through to λPPT5. As briefly mentioned in Chapter 3 (Section

3.3.3.4), these bacteriophage clones were analysed initially so as to determine how many of the seven (putative) PPT-A exons were contained within the insert DNA of each clone. This analysis was conducted by the polymerase chain reaction and by aqueous hybridisation using an exon-specific oligodeoxynucleotide probe.

4.1.1 PCR Analysis of Clones Clones PPT1/2/3/4 & 5

In an attempt to authenticate the existence of mouse PPT-A genomic DNA sequence within the insert sequence of the isolated bacteriophage clones λPPT1/2/3/4 and 5, polymerase chain reactions were performed. Bacteriophage DNA was used as template and was prepared as described in Section 2.1.5. The PCR reaction conditions used to synthesize the original probe fragment (for PPT-A library hybridisations, Section 3.1) were also used in the clone analysis PCRs. Therefore, successful generation of a 514 bp PCR product from both the bacteriophage clone template and the mouse genomic DNA template (Section 3.1.3) would indicate that the clone contained genomic mouse PPT-A sequence. 100 ng of bacteriophage DNA template was used per reaction. 5 µl of each reaction was electrophoresed . Each of the five bacteriophage clones successfully displayed a single PCR band of 514 bp in length. Primers CS6 and CS7 were designed to hybridise to PPT-A exons 3 and 4 respectively. Therefore, the successful PCR analysis indicated that each of the five clones contained at least mouse PPT-A exons 3 and 4.

4.1.2 Restriction Analysis of PPT-A/Glaxo Clones

Bacteriophage DNA was digested as described in Section 2.20. Clones λPPT1/2/3/4 and λPPT5 were digested with the endonuclease *Bam*H I for 2 hrs and electrophoresed . The resulting restriction profile was identical for clones λPPT3 and λPPT4. A different but consistent restriction pattern was generated by clones λPPT1, λPPT2 and λPPT5. Each of these two subsets of PPT-A clones were found to contain the same genomic DNA insert.

This demonstrated that only two distinct PPT-A clones had been isolated from the Glaxo library.

4.1.3 Location of Exons 3 and 4 in PPT2 and PPT4

In addition to the PCR analysis (Section 4.1.1), further evidence that the Glaxo clones λPPT2 and λPPT4 contained PPT-A exons 3 and 4 was provided by hybridisation analysis. A series of digestions were conducted on both λPPT2 and λPPT4 DNA (representing the two distinct clones) using the endonucleases *Sal* I, *Bgl* II, *HinD* III, *BamH* I, *Not* I, *Kpn* I and *EcoR* V. The reactions were electrophoresed and capillary blotted
The DNA was fixed to the membrane by UV cross-linking. 200 ng of the original PCR-derived mouse PPT-A probe fragment (Section 3.1.3) was radioactively-labeled by 'specific priming' (Section 2.2.4). The membrane was probed and washed in the same way as for the library hybridisations described previously (Section 3.2). A DNA dot-blot was used as the control. An exposure of only 3 hrs was required to produce a clear and distinct positive image on the film as a consequence of the very strong signal generated (see Figures 4.1 and 4.2 below). The result reiterated the difference between the genomic inserts in λPPT2 and λPPT4 as demonstrated by restriction analysis (above). The existence of common sequence in both clones was demonstrated by the positive signal which emanated from the 1.4 kb *HinD* III/*Sal* I DNA fragment of both λPPT2 and λPPT4. This in turn implied that the clones had overlapping genomic DNA inserts.

Figure 4.1: Digestion and Hybridisation of PPT-A Bacteriophage Clone PPT2

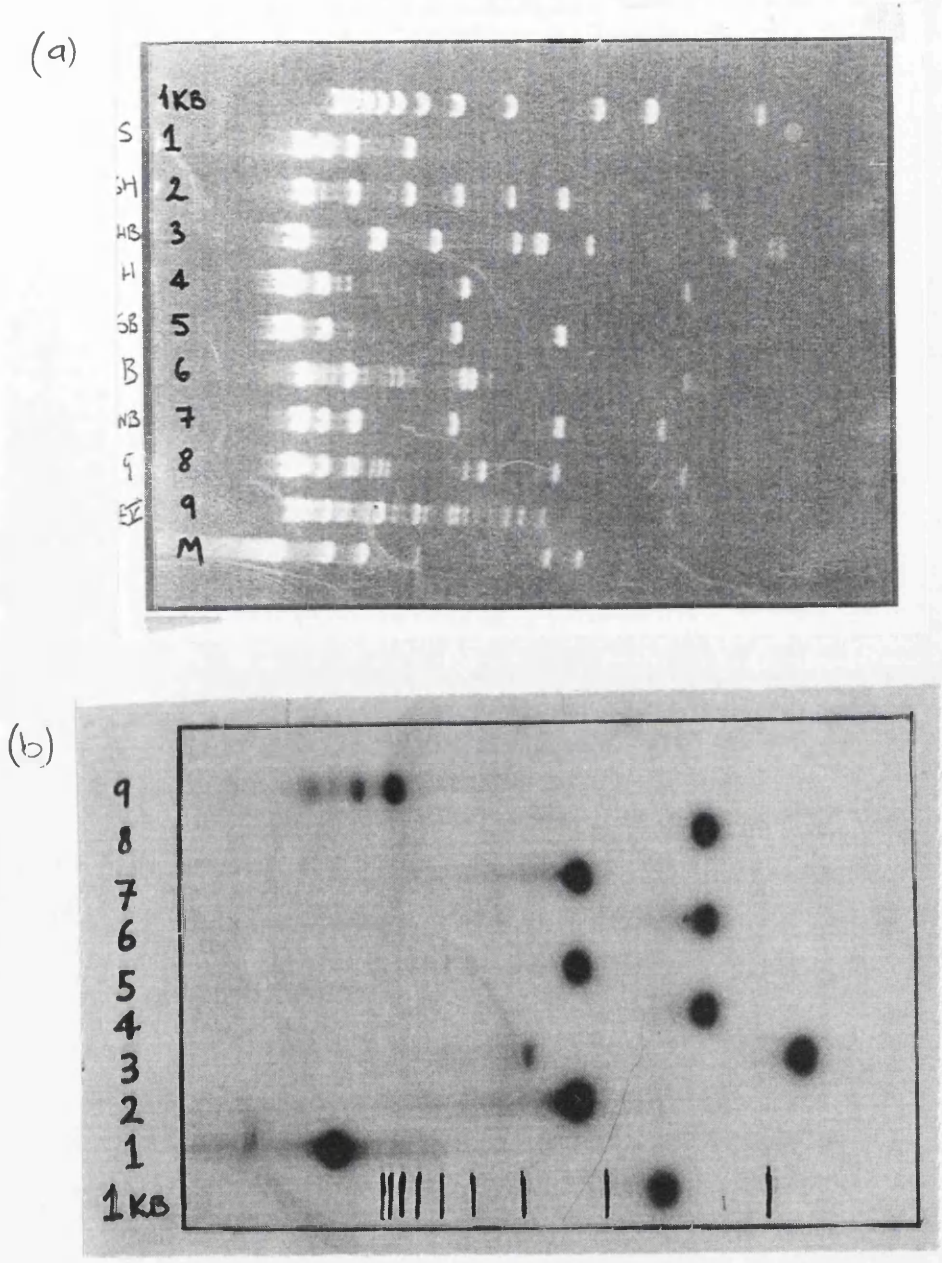


Figure 4.1: (a) 1 µg of clone λPPT2 digested with several endonucleases and electrophoresed on a 0.7% agarose/TBE gel. Lane 1, *Sal* I; lane 2, *HinD* III/*Sal* I; lane 3, *Bgl* II; lane 4, *Bam*H I; lane 5, *HinD* III; lane 6, *Bam*H I/*Not* I; lane 7, *HinD* III/*Kpn* I; lane 8, *HinD* III/*Bam*H I; and lane 9, *Eco*R V. The '1KB' lane is 1 kb DNA ladder marker lane. Please ignore lane 'M' which is another marker lane containing partiallydigested DNA. (b) the corresponding autoradiograph for the above λPPT2 DNA, digested, capillary blotted and probed with the PPT-A exon 3 and 4-specific probe (see Section 4.1.3).

Figure 4.2: Digestion and Hybridisation of PPT-A Bacteriophage Clone PPT4

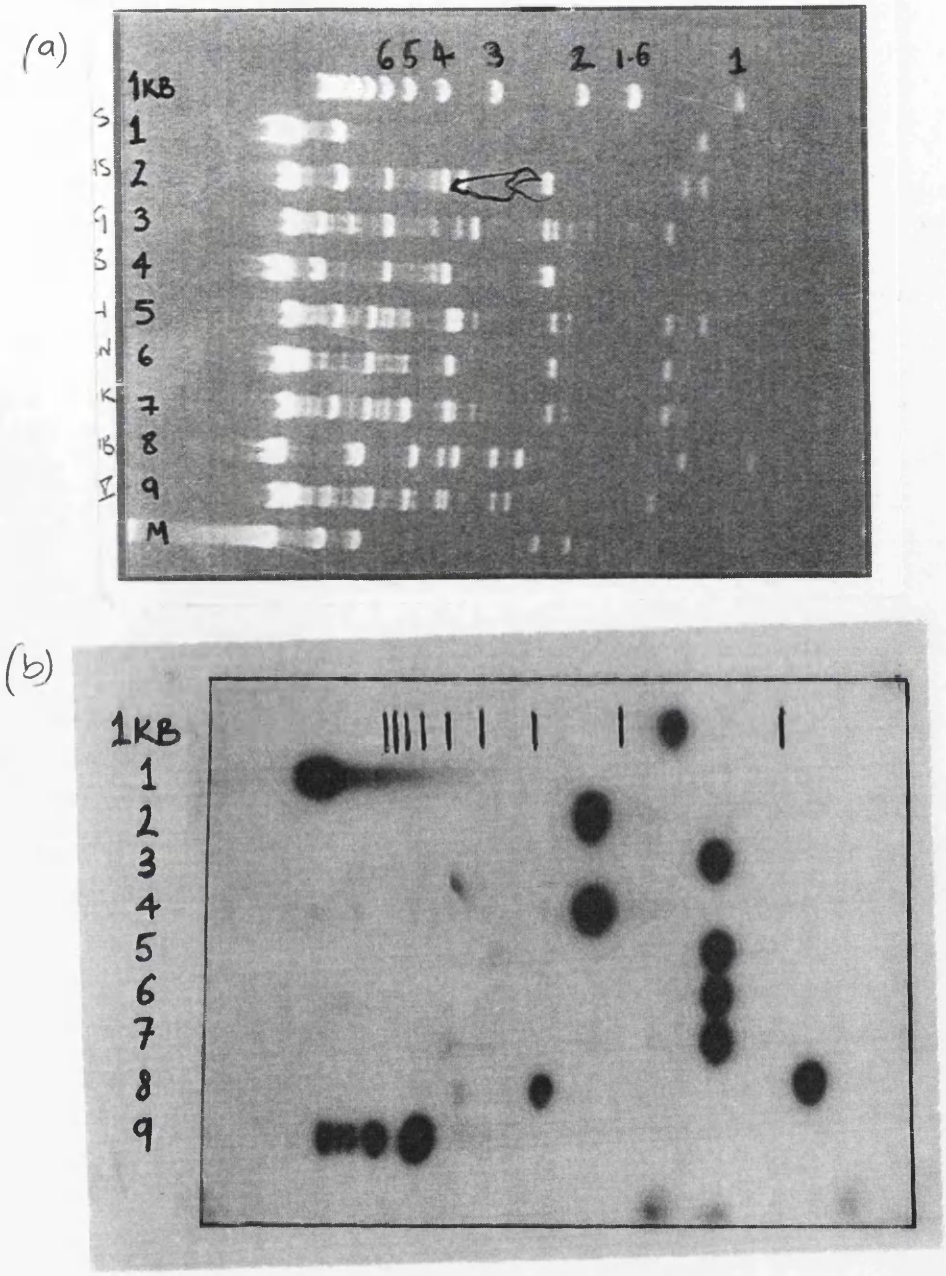


Figure 4.2: (a) 1 µg of PPT-A bacteriophage clone λPPT4 digested with several endonucleases and electrophoresed on a 0.7% agarose/TBE gel. Lane 1, *Sal*I; lane 2, *HinD* III/*Sal* I; lane 3, *HinD* III/*Bam*H I; lane 4, *HinD* III; lane 5 with *Sal* I /*Bam*H I; lane 6, *Bam*H I; lane 7, *Bam*H I/*Not* I; lane 8, *Bgl* II; and lane 9, *Eco*R V. The '1KB' lane is 1 kb DNA ladder marker lane. Please ignore lane 'M'.

(b) the corresponding autoradiograph for the above λPPT4 DNA, digested, capillary blotted and probed with the PPT-A exon 3 and 4-specific probe (see Section 4.1.3).

4.1.4 Location of 5' and 3' Regions of Clones PPT2 and PPT4

Having confirmed that only two distinct PPT-A bacteriophage clones had been isolated from the Glaxo library (i.e., λPPT2 and λPPT4), and that both clones contained exons 3 and 4, it was considered necessary to determine if either clone contained the 5' and/or the 3' end of the mouse PPT-A gene. In order to do this, oligodeoxynucleotides specific to the first and last PPT-A exons, namely exons 1 and 7, were designed and synthesized based on the published rat genomic sequence (Carter, 1990).

4.1.4.1 Oligodeoxynucleotide Hybridisation to Exon 1

The oligodeoxynucleotide CS8 was designed to hybridise to the central region of the 100 bp rat exon 1. This region of the rat exon 1 sequence, and the equivalent bovine sequence, is not translated (see below). The bovine differs from the rat sequence by three (interspersed) nucleotides. Oligodeoxynucleotide CS8 was a 19-mer with a GC:AT ratio of 13:6 and a melting temperature of 64°C.

Figure 4.3: Oligodeoxynucleotide CS8 aligned to Rat PPT-A Genomic Sequence

```
5'   1         10        20        30        40
  actAGAGCGCACTCGGACCAGCTCCACTCCAGCACGCCGGCGGAGGAGAGCG
  |||+|||||||||+|||||||||+|||||||||+|||||||||+||||||||
  tgaTCTCGCGTGAGCCTGGTCGAGGTGAGGTCGTGCGGCCGCCTCCTCTCGC

  50    Probe CS8    70        80        90            3'
AGGAGCGCCCAGCAAGTGCGCACCTGCGGAGCATCACCGGGTCCGACCGCAgtg
||+|||||||||+|||||||||+|||||||||+|||||||||+|||||||||+|
TCCTCGCGGGTCGTTCACGCGTGGACGCCTCGTAGTGGCCCAGGCTGGCGTcac
```

Figure 4.3: 100 bp of rat PPT-A exon 1 genomic sequence (Carter, 1990). Exon 1 specific oligodeoxynucleotide probe CS8 is shown as bold, underlined text. Capital letters denote exonic sequence and lower case letters denote non-exonic sequence. Rat PPT-A exon 1 is not translated.

λPPT2 and λPPT4 DNA was restricted with several endonucleases (*Sal* I, *Nco* I, *BstE* II, *BamH* I), electrophoresed, capillary blotted and fixed to Hybond membrane and probed as described previously (section 3.7). A DNA dot-blot was used as the control for this experiment. Plasmids containing the rat (Nakanishi, 1986) and bovine (Harmar, pers. comm.) PPT-A cDNAs were used as the positive controls, whereas pBluescript and prTKR2 (Section 3.5) DNA was used as the negative controls. 1 ng of each DNA was denatured and spotted on to the membrane and immobilised. 50 ng of oligodeoxynucleotide CS8 was radio-labelled by the polynucleotide kinase method and added to hybridisation buffer to a concentration of 10 ng per ml. The hybridisation reaction was incubated at 45°C after which the membrane was washed and incubated with autoradiographic film overnight. The autoradiograph for the control dot-blot membrane indicated that specific hybridisation had worked. However, the test membrane bearing restricted λPPT2 and λPPT4 DNA did not produce any positive signal. This indicated that bacteriophage clones λPPT2 and λPPT4, which earlier had been shown to contain mouse PPT-A genomic sequence, lacked exon 1 and the adjacent 5' region of the gene.

4.1.4.2 Oligodeoxynucleotide Hybridisation to Exon 7

To ascertain whether clone λPPT2 and/or clone λPPT4 contained the 3' region of the PPT-A gene, the same approach to that used to locate the 5' end of the gene was adopted (Section 4.1.4.1). Based on the rat PPT-A sequence (Carter *et al*, 1990) an oligodeoxynucleotide, which was designed to hybridise to the (translated) 5' end of the 593 bp exon 7, was synthesized. This oligodeoxynucleotide was named CS3. Because the sequence of CS3 was derived from the translated part of the rat exon 7 it was considered most likely to be conserved between rodent species (rat to mouse) and therefore be of use as a homologous oligodeoxynucleotide probe to the mouse exon 7 sequence.

Figure 4.4: Oligodeoxynucleotide CS3 Relative to Rat PPT-A Genomic Sequence

```
5'      1     Probe CS3    20         30         40
   tgcagTGGCTTATGAAAGAAGCGCAATGCAGAACTACGAAAGAAGCGTAAATAAA
   |||||+|||||||||+|||||||||+|||||||||+|||||||||+|||||||||
   acgtcACCGAATACTTTCTTCGCGTTACGTCTTGATGGTTTCTTCGCATTTATTT
       ValAlaTyrGluArgSerAlaMetGlnAsnTyrGluArgArgLysEnd

   50       60        70        3'
   CCGTGTAACGCACTATCTATTCATCTC...(exon continued to 593 bp)
   +|||||||+|||||||||+|||||||
   GGCACATTGCGTGATAGATAAGTAGAG
```

Figure 4.4: The first 77 bp of the 593 bp rat PPT-A exon 7 genomic DNA sequence (Carter *et al*, 1990). Oligodeoxynucleotide CS3 relative to that exonic sequence is shown in underlined, bold text. Exonic sequence is shown in capital letters whereas intronic sequence is shown in lower case text. Amino acids translated from this exon are shown in italics. Bold italics denote the amino acids translated from within the CS3 sequence. Translated sequence is highly conserved between rat and mouse.

Probe fragment CS3 constituted 21 nucleotides, had a GC:AT ratio of 10:11, and a melting temperature of 62°C. The hybridisation experiment which used probe CS3 was conducted under the same conditions as for probe CS8 (Section 4.1.4.1). Clone λPPT2 and λPPT4 DNA was restricted with the endonucleases *Kpn* I, *Nco* I and *Sal* I, electrophoresed simultaneously on a 0.5% agarose gel and blotted on to Hybond-N membrane. The control DNA dot-blot was the same as in the CS8 hybridisation detailed above. The test membrane was probed with radioactively-labelled (exon 7-specific) probe CS3, washed in 0.1 x SET buffer and autoradiographed. The developed film displayed strong positive hybridisation signals emanating from fragments of both PPT-A bacteriophage clones. The positive and negative controls for this experiment worked. The result indicated that each of the PPT-A bacteriophage clones, λPPT2 and λPPT4, isolated from the Glaxo library lacked the 5' region of the gene (including exon 1) but did comprise the 3' region of the gene (including exon 7).

Figure 4.5: Clones λPPT2 and λPPT4 Hybridised with an Exon 7-Specific Probe

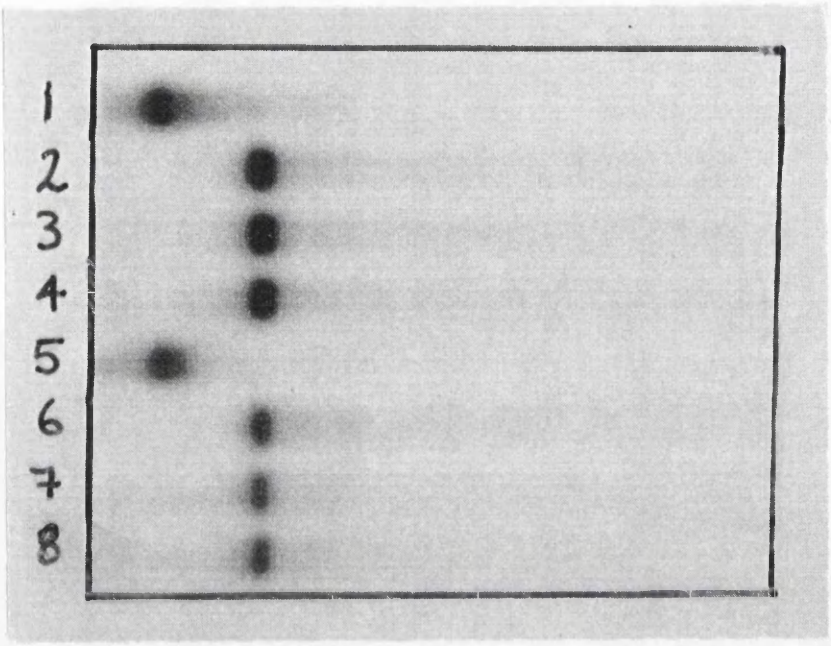


Figure 4.5: Autoradiograph of endonuclease digested λPPT2 (lanes 1-4) and λPPT4 (lanes 5-8) electrophoresed, capillary blotted and hybridised with exon 7-specific oligodeoxynucleotide probe CS3. Lanes 1 and 5; *Kpn* I, lanes 2 and 6; *Kpn* I /*Not* I, lanes 3 and 7; *Not* I, lanes 4 and 8; *Sal* I/ *Not* I.

4.2 ANALYSIS OF THE PPT-A/CAMBRIDGE LIBRARY CLONES

With respect to both gene characterisation and design of a gene targeting vector, it was considered necessary to isolate bacteriophage clones containing the 5' region of the PPT-A gene. On account of this a different library, the Cambridge genomic DNA library, was screened (Section 3.3.4) and the five isolated clones were named λPPT6 through to λPPT10. The same PCR-derived, exon 3/4-specific probe as was used to screen the Glaxo library was used to screen the Cambridge library. Therefore clones λPPT6 through to λPPT10 were analysed to determine which, if any, contained exon 1 and the adjacent 5' region of the mouse PPT-A gene.

4.2.1 Oligodeoxynucleotide Hybridisation to Exon 1

Bacteriophage clones λPPT6 to 10 were analysed by aqueous hybridisation. In order that the clones could be screened rapidly, they were analysed in an unusual but effective way which did not require the preparation of bacteriophage DNA or the bacteriophage clone lysates from which DNA is prepared.

After the tertiary library hybridisation described in section 3.3.4.3 a single bacteriophage plaque was isolated from the library master plate for each of the five identified PPT-A clones. Each plaque was incubated in 1 ml of phage buffer to allow the bacteriophage to resuspend. A small square NZCYM media plate was overlayed with NZCYM top agar mixed with LE392 plating cells and allowed to set (Section 2.11). (This differed from the standard procedure in that the top agar/plating cell mix lacked bacteriophage). For each of the five PPT-A/Cambridge library clones, 2 µl of the single-plaque resuspension was spotted onto the NZCYM plate, prepared as described above, leaving several centimetres between each 'spot'. The plate was then incubated at 37 °C overnight. The following day, as expected, the plate displayed five distinct plaques of approximately 1 cm diameter representing bacteriophage clones λPPT6 to 10.

Duplicate replica membrane lifts were made from the plate (Section 2.14). The membranes were then hybridised using the same exon 1-specific oligodeoxynucleotide probe as used to screen clones λPPT1 to 5 in Section 4.1.4.1. A dot-blot membrane was used as the control for this experiment. A plasmid which contained the bovine PPT-A cDNA was used as the positive control and pBluescript was used as the negative. The hybridisation was successful and allowed the identification of clones containing exon 1. Clones λPPT6, λPPT7, λPPT9 and λPPT10 were shown to contain exon 1 while only clone λPPT8 was shown to lack this exon.

4.2.2 PCR Analysis of Clones PPT6/7/8/9 and 10

Clones λPPT6/7/8/9 and 10 were anlysed by polymerase chain reaction in the same way and under the same reaction conditions as clones λPPT1 to 5 (Section 4.1.1). The analysis was equally conclusive in demonstrating that clones λPPT6 to 10 contained exons 3 and 4 because a single 514 bp PCR product was generated from the bacteriophage clone DNA template using primers CS6 and CS7. Further confirmation that clones λPPT6 to 10 contained exon 3 was provided by successful hybridisation of the exon 3-specific probe CS6 (Section 3.1.1) to each of the clones.

4.2.3 Oligodeoxynucleotide Hybridisation to Exon 7

DNA from clones λPPT6/7/8/9 and 10 was digested with the endonucleases *Nco* I and *Sal* I, electrophoresed, blotted and probed with oligodeoxynucleotide CS3 (as in Section 4.1.4.2). The same controls were successfully included. Hybridisation indicated that PPT-A clones λPPT6, λPPT7, λPPT9 and λPPT10 did indeed contain exon 7 and the 3' end of the mouse PPT-A gene. λPPT8 did not.

4.2.4 Occurence of Exons in PPT-A Bacteriophage Clones PPT1 Through to PPT10 (Figure 4.5)

Figure 4.6 : Exons Contained in PPT-A Bacteriophage Clones λPPT1 to λPPT10

	Exon						
Clone	1	2	3	4	5	6	7
λPPT1			√	√	√	√	√
λPPT2			√	√	√	√	√
λPPT3			√	√	√	√	√
λPPT4			√	√	√	√	√
λPPT5			√	√	√	√	√
λPPT6			√	√	√	√	√
λPPT7	√	√	√	√	√	√	√
λPPT8	√	√	√				
λPPT9	√	√	√	√	√	√	√
λPPT10			√	√	√	√	√

Figure 4.6: Table detailing the exons contained in PPT-A bacteriophage clones λPPT1 to λPPT10 as determined by positive aqueous hybridisation using oligodeoxynucleotides specific to each exon. The sequence of the oligodeoxynucleotide probes was based on the published rat genomic (Carter *et al*, 1990) and mouse cDNA (Kako *et al*, 1993) sequences.

4.3 SUBCLONING THE MOUSE PPT-A GENE

The rat PPT-A gene has been shown to contain 7 exons and be 8 kb in size (Carter *et al*, 1990). The mouse PPT-A gene was predicted to be of similar genomic organisation due to the high rat/mouse species homology. DNA probes specific to exons 1, 3, 4 and 7 had been shown to hybridise to clone λPPT9. On that basis, it was considered likely that this clone, with a genomic DNA insert of between 10 and 20 kb, would contain the whole mouse PPT-A gene. Clone λPPT9 was

therefore selected for further analysis and subcloning. λPPT9 DNA was digested with several endonucleases in order to identify one which would digest the bacteriophage lambda (vector) DNA infrequently while digesting the insert DNA at a few loci so as to generate fragments of between 5 and 10 kb. Insert fragments of this size were considered ideal for subcloning. Endonucleases with recognition sequences of six or eight nucleotides, such as *Not* I, *Nco* I and *Sal* I, were tested. The restricted DNAs were electrophoresed, transferred to membrane and probed with exon 1 and exon 7-specific oligodeoxynucleotide probes (as described in Sections 4.1.4.1 and 4.1.4.2 respectively). The endonuclease *Nco* I was chosen because it generated insert-derived DNA fragments of 1.8, 4.3, 7.2, and 9.5 kb. Of these fragments, the 7.2 kb fragment contained exon 1 and the 9.5 kb fragment contained exon 7. As these two DNA fragments contained the 5' and 3' regions of the mouse PPT-A gene respectively they were selected for subcloning. (It was later demonstrated that the 7.2 and 9.5 kb fragments were adjacent with respect to the genomic organisation of the gene and between them contained all seven of the mouse PPT-A exons).

4.3.1 Shotgun Ligation; Subclone pSP9

5 μg of λPPT9 DNA was restricted with 10 units of *Nco* I because this digestion had previously been shown to generate a 9 kb fragment containing PPT-A exons 1, 2, 3 and 4 (see Section 4.1.2). A sample of the reaction was electrophoresed on an agarose gel to confirm that restriction of the DNA was complete. On completion, the endonuclease was removed from the reaction by organic extraction with phenol and chloroform (Section 2.23). A plasmid, namely pGEM®5Zf(+), was selected as a suitable subcloning vector because of the *Nco* I site located in the multiple cloning site. 5 μg of pGEM®5Zf(+) vector DNA was restricted with *Nco* I. Complete restriction of the vector DNA was also determined by gel electrophoresis. The buffer concentration of the reaction was altered to facilitate the dephosphorylation of the restricted vector DNA with the modifying enzyme CIP (section 2.21.1). After incubation, the CIP and *Nco* I enzymes were removed from the

reaction by extraction with phenol and chloroform. 'Conventional' ligation reactions were then performed as described in section 2.21.1. The ligations were transformed into *E. coli* strain DH10B cells by the heat-shock method (section 2.4.1). Transformants containing plasmids with λPPT9-derived genomic DNA inserts were distinguished by blue/white selection (see Section 2.4). White colonies were picked, grown under ampicillin selection and plasmid DNA was prepared from each bacterial culture by the alkaline lysis method detailed in section 2.17.1. 1 μg of subclone DNA was restricted with *Nco* I and electrophoresed alongside DNA marker (1kb ladder) so that the size of the insert could be determined visually. One such plasmid subclone was found to contain the 9 kb λPPT9/*Nco* I DNA fragment and therefore exon 1 of the mouse PPT-A gene. This subclone was named pSP9 (see Section 2.8.1 for nomenclature criteria).

4.3.2 Subclone pSP7

The white colonies obtained from the λPPT9/pGEM®5Zf(+) shotgun ligation transformation (Section 4.3.1) were fully analysed. None of the subclones had the 7.2 kb λPPT9/*Nco* I fragment as an insert. This DNA fragment had to be subcloned into the vector pGEM®5Zf(+) specifically rather than by the shotgun ligation method. λPPT9 DNA was restricted with *Nco* I and electrophoresed on a 0.5% agarose/TAE gel. The 7.2 kb DNA fragment was band-isolated from the gel and the DNA recovered using a (section 2.27.2). The purified 7.2 kb DNA fragment was used in ligations which were, in turn, transformed into competent *E. coli* strain DH10B. White colonies were selected, cultured and analysed by digestion with *Nco* I. The plasmid subclone with the 7.2 kb λPPT9/*Nco* I fragment as an insert was named pSP7.

4.4 RESTRICTION MAPPING THE PPT-A SUBCLONES pSP9 AND pSP7

PPT-A plasmid subclones pSP9 and pSP7 were restriction endonuclease mapped in the same way. The vector DNA was the same for both subclones. The full restriction map of the vector DNA (pGEM®5Zf+) was known (Promega Corp. catalogue). Each subclone was digested, in the first instance, with each of the 30 standard endonucleases kept in the laboratory. From the restriction pattern generated, the number of restriction sites in the insert was determined for each enzyme. Only endonucleases which were found to cut the insert DNA less than four times were more extensively mapped. In order to do this, a unique restriction site in the multiple cloning site of the vector DNA, and which did not cut the insert DNA, was located; for example *EcoR* V. The insert restriction sites were then located relative to that known polylinker site. When the newly located restriction sites for the several test enzymes were plotted, a comprehensive linear restriction map for each subclone was constructed.

Figure 4.7: Restriction Maps of Plasmid Subclones pSP9 and pSP7

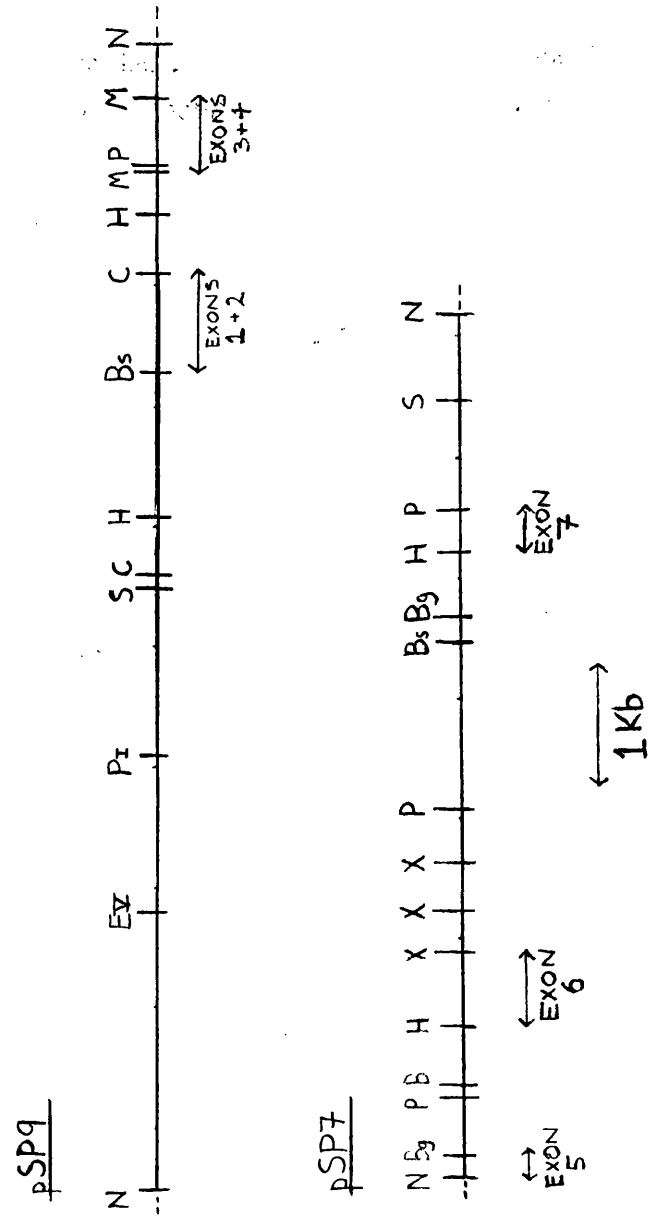


Figure 4.7: Linear map of the restriction sites located within the cloned inserts of plasmid subclones pSP7 and pSP9. Exons were localised to the smallest known DNA fragment by hybridisation using exon-specific oligodeoxynucleotides. C; *Cla* I, Ev; *EcoR* V, P; *Pst* I, H; *HinD* III, M; *Mlu* I, Bg; *Bgl* II, P1; *Pvu* I, B; *BamH* I, X; *Xba* I, N; *Nco* I, S; *Sst* I, Bs; *BstE* II.

4.5 LOCATING THE MOUSE PPT-A EXONS

In order to construct the restriction maps of the plasmid subclones pSP9 and pSP7 (Section 4.4), the single and double endonuclease reactions were electrophoresed on 0.7% agarose gels (up to 10 reactions per gel). The gels were photographed so that the DNA fragments could be accurately sized and a map made. Instead of discarding the gels, they were Southern blotted onto nylon membrane and stored. The membranes were later probed by aqueous hybridisation using radioactively-labelled oligodeoxy nucleotides specific to each of the seven exons in order to define which DNA fragment of which subclone contained which exon. All hybridisations detailed in Sections 4.5.1 to 4.5.6 were performed at 45 °C and by the protocol described in Section 2.29.2 and with DNA dot-blot membranes as positive and negative controls. In each case, the controls indicated that the hybridisations were both successful and specific.

N.B. By this stage of the project, the mouse PPT-A gene was known to contain seven exons as demonstrated by Kako (1993). This publication detailed the mouse PPT-A cDNA sequence which, rather than the rat PPT-A genomic sequence (Carter 1990), was subsequently used in the design of deoxyoligonucleotide probes specific to mouse PPT-A exons.

4.5.1 Mouse PPT-A Exon 1

As a result of restriction mapping subclone pSP9, the DNA fragments were electrophoresed and blotted onto membrane. Each membrane bore between 5 and 10 blotted pSP9 DNA restrictions reactions. The reactions were typically doubles where two endonucleases had been incubated simultaneously. Eleven such membranes were probed with the exon 1-specific oligodeoxynucleotide CS8 (Section 4.1.4.1). The probe hybridised to a 2.7 kb *Nco* I/*BstE* II fragment (Figure 4.7, lane 3) and a 2.7 kb *Cla* I/ *Cla* I fragment (not shown). Thus, the mouse PPT-A exon 1 was localised to the 800 bp *Cla* I/*BstE* II fragment of subclone pSP9: the overlap of the two fragments which hybridised positively.

Figure 4.8: Location of PPT-A Exon 1 in Plasmid Subclone pSP9

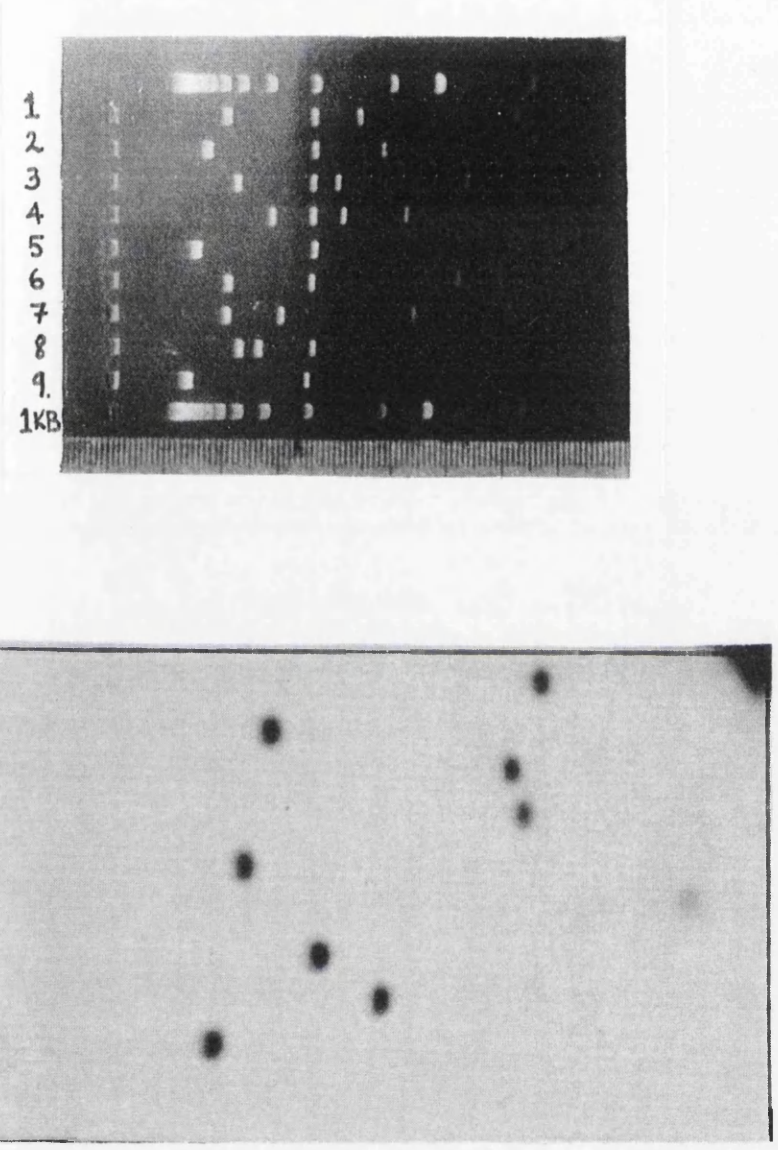


Figure 4.8: Upper image: Photograph of plasmid subclone pSP9 digested with several endonucleases and electrophoresed on a 0.7% agarose/TBE gel. Lower image: This gel was capillary blotted onto Hybond-N and hybridised with the PPT-A exon 1-specific probe CS8 to generate the autoradiograph shown beneath. Note the 2.7 kb *Bst*E II/*Nco* I fragment in lane 3 of both images. DNA marker used was 1 kb LADDER.

4.5.2 Mouse PPT-A Exon 2

The same 11 membranes as were probed in Section 4.5.1 were probed with an exon 2-specific probe. The sequence of this oligodeoxynucleotide, named CS14, was based on the mouse PPT-A cDNA sequence published by Kako (1993). Oligodeoxynucleotide CS14 was a 20-mer which hybridised to the 5' end of the mouse exon 2 sequence, had a GC:AT ratio of 10:10 and a melting temperature of 60 °C.

Figure 4.9: Sequence of Exon 2-Specific Oligodeoxynucleotide CS14

5' 3'

ATGAAAATCCTGTGGCCGTG

Figure 4.9: Sequence of the oligodeoxynucleotide CS14 as used in hybridisations to locate mouse PPT-A exon 2 in subclone pSP9. Sequence based on the published mouse PPT-A cDNA sequence of Kako *et al* (1993). According to the published mouse sequence, the first nucleotide of the above oligodeoxynucleotide sequence is the 12th nucleotide from the 5' end of the mouse exon 2 sequence. The mouse PPT-A exon 2 comprises 135 bp in total.

The CS14 probe hybridised to the 2.7 kb *Nco* I/*BstE* II fragment and the 2.7 kb *Cla* I/*Cla* I fragment of pSP9. Therefore, the mouse PPT-A exon 2 was, like exon 1, localised to the same 800 bp *Cla* I /*BstE* II fragment of subclone pSP9.

4.5.3 Mouse PPT-A Exons 3 and 4

Oligodeoxynuclotides CS6 and CS7 (Section 3.1.1 and 3.1.2) were used to locate exon 3 and exon 4 respectively. The 11 filters bearing restricted pSP9 DNA were stripped and re-hybridised using these probes. Both probes hybridised to a 900 bp *Mlu* I/*Mlu* I DNA fragment of subclone pSP9. This fragment was later subcloned as part of the

gene targeting vector construction (Chapter 5). Both probes also hybridised to an 800 bp *Mlu* I/*Pst* I fragment. The *Pst* I restriction site was also found to be of importance during vector construction (for insertion of the neomycin resistance gene; see Section 5.1.3.2).

4.5.4 Mouse PPT-A Exon 5

Four membranes bearing capillary blotted, restricted subclone pSP7 DNA were used in the localisation of exon 5. An exon 5-specific oligodeoxynucleotide, named CS15, was synthesized based on the mouse PPT-A cDNA sequence (Kako *et al*, 1993) and used in aqueous hybridisations. CS15 had a GC:AT ratio of 10:10 and a T_m of 60°C.

Figure 4.10: Sequence of Exon 5-Specific Oligoeoxynucleotide CS15

5' 3'

ACATGGCCAGATCTCTCACA

Figure 4.10: Sequence of oligodeoxynucleotide CS15 which was used in the hybridisation to localise the mouse PPT-A exon 5 in subclone pSP7. This sequence is based on the published mouse PPT-A cDNA sequence of Kako *et al* (1993). According to the published mouse sequence, the first nucleotide of the above oligodeoxynucleotide sequence is the 1st nucleotide at the 5' end of the mouse exon 5 sequence. The mouse PPT-A exon 5 comprises only 24 bp in total.

Probe CS15 was found to hybridise to a 200 bp *Nco* I/*Bgl* II fragment of subclone pSP7.

4.5.5 Mouse PPT-A Exon 6

The four pSP7 membranes which had been probed previously were stripped and re-hybridised with an exon 6-specific oligodeoxynucleotide probe. The probe, which was a 20-mer with a

melting temperature of 64°C, was derived from the mouse cDNA sequence (Kako *et al*, 1993) and was named CS16.

Figure 4.11: Sequence of Exon 6-Specific Oligodeoxynucleotide CS16

5' 3'

ACATGGCCAGATCTCTCACA

Figure 4.11: Sequence of oligodeoxynucleotide CS16. CS16 was used in the hybridisation to localise the mouse PPT-A exon 6 in subclone pSP7. The sequence of this oligodeoxynucleotide is based on the published mouse PPT-A cDNA sequence (Kako *et al*, 1993). According to the published mouse sequence, the first nucleotide of the above CS16 oligodeoxynucleotide sequence is the 3rd nucleotide at the 5' end of the mouse exon 6 sequence. The mouse PPT-A exon 6 comprises 53 bp in total. CS16 lies within the translated sequence of exon 6 from which the neuropeptide substance K (neurokinin-A) is encoded (Kako *et al*, 1993).

Probe CS16 was found to hybridise positively to both the 3.2 kb *HinD* III/*BamH* I and the 3.8 kb *HinD* III/*HinD* III fragment. Mouse PPT-A exon 6 was further localised to the 600 bp *HinD* III/*Xba* I fragment of subclone pSP7.

4.5.6 Mouse PPT-A Exon 7

The pSP7 membranes were stripped and re-hybridised with the exon 7-specific oligodeoxynucleotide CS3 (Section 4.1.4.2). This probe hybridises to the 5' end of exon 7. The probe selectively hybridised to 2.5 kb *Pst*I fragment, the 2.7 kb *BamH* I/*Nco* I (shown below) and the 2.2 kb *HinD* III fragment. The 5' end of the mouse PPT-A exon 7 was therefore localised to the 350 bp *HinD* III/*Pst* I fragment of subclone pSP7.

4.6 SEQUENCE OF THE MOUSE PPT-A GENE

Once purified, the DNA from the bacteriophage clones isolated from the Glaxo and Cambridge genomic libraries (see Sections 3.3.3 and 3.3.4, respectively) were analysed, by endonuclease digestion (Section 4.1.2) and also by PCR analysis (see Section 4.11). This, however, did not provide sufficient evidence to confirm that the cloned DNAs did indeed contain mouse PPT-A sequence; further evidence was required to authenticate this.

In order to do this, DNA sequencing was performed using the plasmid subclones derived from the isolated bacteriophage, as templates. A number of sequencing primers were used to do this. The short oligodeoxynucleotide primers (of between 18 and 22 bp in length) which had been used in earlier experiments as PCR primers (see Section 3.1.3) were also used as sequencing primers. Similarly, using the published rat PPT-A sequence (Carter *et al*, 1990), further PPT-A exon-specific sequencing primers were designed, synthesized and used. Furthermore, standard plasmid vector primers such as T7 primer, which anneals to the multiple cloning site sequence of pBluescript KS+, were used (also T3 and SP6 primers, see Stratagene catalogue).

Both manual and automated sequencing was performed using the methodologies described in Sections 2.30 and 2.31. Novel DNA sequence was obtained and is presented in this text. In the alignments shown below, novel sequence is referred to as "query sequence". The query sequence was aligned with the DNA entries which have been published and are held on computerised DNA databases (such as that of the European Molecular Biology Laboratory; GenEMBL). In the alignments shown below, this sequence is referred to as "database sequence". For each of the query sequences analysed in this way, a number of alignments were generated with different degrees of sequence identity. Alignments representing several different intronic and exonic regions of the mouse PPT-A gene (for example the 5' upstream region) were obtained using this approach. Only the closest

("best-fit") alignment between the query sequence and the database sequence was selected and is given below (although other alignments with lower percentage identities were also generated). The origin of the database sequence, the accession number and the percentage sequence identity are given for each alignment. Additionally, the primer used to generate the query sequence is specified for each alignment.

4.6.1 Murine PPT-A Gene 5' Flanking Sequence

The plasmid subclone pSP9 (see Section 4.3.1) was used as sequencing template in combination with the T7 sequencing primer which anneals to the T7 RNA polymerase transcription initiation site in the plasmid vector pBluescript KS+ (both primer and vector from Stratagene). 83 bp of the mouse genomic DNA sequence was found to align with database sequence flanking the 5' region of the rat PPT-A gene. That database sequence is derived fom the 4247 bp sequence entered onto the database from the species *Rattus norvegicus* and which has the accession number gb|L07328|RATPPT5FL. The alignment of those sequences, as shown below, demonstrates a 93% sequence identity between the query and database sequences. This is in line with predicted mouse to rat species DNA sequence identities.

Figure 4.12: Alignment of Murine and Rat PPT-A Gene Sequence Upstream of Exon 1 in the 5' Flanking Region

```
A: Query:     83 CTTCCTCTGTCTCCTACTCTCTTCAGAAATCGAACATGAAAATCCTCGTGGCCGTGGCGG 24
                 | | ||||||||||||||||||| |||||||| |||||||||||||||||||| |||||||
B: Data :   3855 CCTTCTCTGTCTCCTACTCTCTCCAGAAATCCAACATGAAAATCCTCGTGGCGGTGGCGG 3914

   Query:     23 TCTTTTTTCTCGTTTCCACTCAA 1
                 |||||||||||||||||||||||
   Data:    3915 TCTTTTTTCTCGTTTCCACTCAA 3937
```

Figure 4.12: Analysis of the murine PPT-A sequence upstream of exon 1 which was generated during this research; aligned with the corresponding published rat PPT-A gene sequence (GenEMBL database, see Section 4.6.1 above). 93% identity shown.

4.6.2 Murine PPT-A Gene Exons 1 and 2 Sequence

Plasmid subclone pSP7.2 (Section 5.1.2) was used as sequencing template with the sequencing primer T7. 130 bp of the generated mouse genomic sequence was found to align with the rat PPT-A gene exon 1 and exon 2 database sequence. The rat sequence is derived from the 1757 bp *Rattus norvegicus* sequence database entry (accession number gb|M34159|RATPPTA1). The figure below demonstrates that 120 of 133 deoxynucleotides align exactly giving a 90% sequence identity.

Figure 4.13: Alignment of Murine and Rat PPT-A Exon 1 and Exon 2 Sequence

```
A: Query:    405 GTCTTCAGGCTTGAGAGTGTGGGTCAGTGGGTAGGGGACTGGGACGCTGAGACTCAAGGA 346
                 | ||||||||||  |||||||||||||||||||||||| |||||||| |||||  ||| ||
B: Data:    1108 GCCTTCAGGCTTCGGAGTGTGGGTCAGTGGGTAGGGGGCTGGGACGTTGAGAGGCAAAGA 1167

   Query:    345 GAGGAGGATTTGAGGCTCTTTGGCGCCTCAGGAGCCTTCTTAAAGGGTTTGGAGAAGTGT 286
                 |||||||| |||||||||||||| | |||| |||||||||||||||||||||||| |||
   Data:    1168 GAGGAGGACTTGAGGCTCTTTGGCACGTCAGTAGCCTTCTTAAAGGGTTTGGAGAAATGT 1227

   Query:    285 TCGACCTGGTGGG 273
                 |||||||||||||
   Data:    1228 TCGACCTGGTGGG 1240
```

Figure 4.13: Analysis of the murine PPT-A exon 1 and exon 2 sequence; aligned sequence from subclone pSP7.2 (generated with T7 primer) and the corresponding published rat PPT-A gene sequence (GenEMBL database, see Section 4.6.3 above). 90% identity shown.

4.6.3 Murine PPT-A Gene Exon 3 Sequence

The plasmid subclone pSP7.2 (section 5.1.2) was used as sequencing template with the primer CS6 (Section 3.1.1). 35 bp of the mouse genomic sequence was found to align with the bovine PPT-A gene exon 3 database sequence (197 bp *B. taurus* DNA fragment, accession number emb|X01396|BTPPT3). The figure below shows the plus strand to plus strand alignment of those sequences where 29 of 35 nucleotides align precisely giving a 82% sequence identity.

Figure 4.14: Alignment of Murine and Bovine PPT-A Exon 3 Sequence

```
A: Query:      5 CACGCGGGATGCTGGTGAGATGAGCAATAGTCACT 39
                 ||  ||||||||||||||||||| ||| | ||| ||
B: Data:     134 CAAACGGGATGCTGGTGAGATGGGCAGTCGTCCCT 168
```

Figure 4.14: Analysis of the murine PPT-A exon 3 sequence; aligned sequence from subclone pSP7.2 (generated with CS6 primer) and the equivalent published bovine PPT-A gene sequence (see Section 4.6.3 above). 82% identity shown.

4.6.4 Murine PPT-A Gene Exon 3 and Exon 4 Sequence

The sequencing template was plasmid subclone pSP7.2 and the primer was CS6 (Section 3.1.1). 150 bp of the query mouse genomic DNA sequence was found to align with the rat PPT-A gene exon 3 and exon 4 database sequence (875 bp *Rattus norvegicus* DNA sequence, with the accesion number gb|M34160|RATPPTA2). The figure below shows the plus strand to plus strand alignment where 132 of 150 nucleotides align to give an 88% sequence identity.

Figure 4.15: Alignment of Murine and Rat PPT-A Exon 3 and Exon 4 Sequence

```
A: Query:      5 CACGCGGGATGCTGGTGAGATGAGCAATAGTCACTAGATCTATCAGGCTACCTGGTCTGA 64
                 ||  |||||||||||||||||||||||||||||||||||||| |||||||||||||| ||||||
B: Dat:      270 CAAACGGGATGCTGGTGAGATGAGCAATAGTCACTAGGTCTATCAGGCTACCCGGTCTGC 329

   Query:     65 ATGCTTGTTCCTTCCTGGAGAACCCAAATGTCTCTTGTACTGGGAGTATAGATTTACACG 124
                 ||||  ||  |||||||||||| |||||| |||||||||||| | |||||||  |||||| |
   Dat:      330 ATGCCTGCTCCTTCCTGGAGCACCCAAGTGTCTCTTGTACCGCGAGTATAATTTTACATG 389

   Query:    125 CCTAAAGGCTTYGGATTCCCAGAAGTCTTC 154
                 |||||||  |||  |||||||  |||||||||
   Dat:      390 CCTAAAGACTTTTGATTCCCTGAAGTCTTC 419
```

Figure 4.15: Analysis of the murine PPT-A exon 3 and exon 4 sequence; aligned sequence from subclone pSP7.2 (generated with CS6 primer) and the equivalent published rat PPT-A gene sequence (see Section 4.6.4 above). 88% identity shown.

4.6.5 Murine PPT-A Gene Exon 7 Sequence

Plasmid subclone pSP10.6 (Section 5.1.1) was used as a sequencing template along with T3 primer (Stratagene). 129 bp of mouse genomic sequence was found to align with the rat PPT-A gene exon 7 database sequence (779 bp *Rattus norvegicus* DNA sequence, accession number gb|M34162|RATPPTA4). The figure below demonstrates the plus strand to minus strand alignment where 112 of 119 nucleotides align precisely giving an 94% sequence identity between mouse and rat in the region of exon 7.

Figure 4.16: Alignment of Murine and Rat PPT-A Exon 7 Sequence

```
A: Query:    249 TGCAGTGGCTTATGAAAGAAGCGCGATGCAGAACTACGAAAGAAGACGTAAATAAACCCC 190
                 |||||||||||||||||||||||| |||||||||||||||||||| ||||||||||||
B: Data:       2 TGCAGTGGCTTATGAAAGAAGCGCAATGCAGAACTACGAAAGAAGGCGTAAATAAACCCT 61

   Query:    189 GTAACGCACTATCTATTCATCTTCATCTGTGTCAGTGAGCAGTGAACGGTAAAATAAAA 131
                 ||||||||||||||||||||||| |||||||||| | ||||||||| ||||||||||||
   Data:      62 GTAACGCACTATCTATTCATCTCCATCTGTGTCCGCGAGCAGTGAGCGGTAAAATAAAA 120
```

Figure 4.16: Analysis of the murine PPT-A exon 7 sequence; aligned sequence from subclone pSP10.6 (using primer T3) and the equivalent published rat PPT-A gene sequence (see Section 4.6.5 above). 94% identity shown.

4.7 ANALYSIS OF THE NK-1R/TORONTO LIBRARY CLONES

Sixteen positively-hybridising NK-1R bacteriophage clones, named λSPR1 through to λSPR16, were identified as a result of the hybridisation screen of the Toronto bacteriophage library (Section 3.6.3). Of these, clones λSPR15 and λSPR16 were known to contain mouse NK-1R exon 2 genomic sequence (Section 3.7). In order to determine which bacteriophage clone contained which of the remaining NK-1R exons, the clone DNA was analysed by aqueous hybridisation using exon specific oligodeoxynucleotides. Endonuclease restriction mapping was also used as an indicator of which bacteriophage clones contained which exons by means of determining common restriction site patterns.

4.7.1 Restriction Analysis of Clones SPR1 to 14

DNA was prepared from bacteriophage clones λSPR1 to λSPR14 (Section 2.15). 2 μg of each DNA was digested with the endonuclease *BamH* I and electrophoresed on a 0.5% agarose gel. The gel was capillary blotted onto nylon membrane where the DNA was immobilised. Initial restriction analysis of the clones demonstrated that there were four subsets of NK-1R bacteriophage clones which had identical inserts, namely clones λSPR1 and 4, clones λSPR3 and 6, λSPR5,11 and 14, as well as clones λSPR9 and 12. It was found that of the 14 clones isolated from the Toronto library (as a consequence of the first library screen) there were only 9 which had distinct genomic DNA inserts. Clone λSPR15 was known to contain the same (exon 2-specific) insert as λSPR16. When this is taken into account, there were a grand total of 10 distinct NK-1R bacteriophage clones.

4.7.2 Oligodeoxynucleotide Probes for NK-1R Exons

N.B. By this stage of the project the mouse NK-1R gene, like the rat NK-1R gene, was known to comprise 5 exons as demonstrated by Sundelin et al *(1992). The sequence of each exon is detailed in that publication.*

The following oligodeoxynucleotides were used as probes to localise each of the exons in the NK-1R bacteriophage clones. Probes CS9, CS11, CS12 and CS13 were used to probe exons 1, 3, 4 and 5 respectively. The sequence of the oligodeoxynucleotides was based on the published mouse NK-1R sequence (Sundelin *et al*, 1992) where the first (5') nucleotide of each oligodeoxynucleotide sequence corresponds to the first (5') nucleotide of each mouse NK-1R exon respectively. The probe used to localise exon 2 is detailed in section 3.7.1. The oligodeoxynucleotides CS9, CS11, CS12 and CS13 have melting temperatures of 58, 60, 56, and 64 °C respectively.

Figure 4.17: Sequence of Oligodeoxynucleotides CS9, CS11, CS12 and CS13 Homologous to Mouse NK-1R Exons 1,3,4 and 5

	5'	3'
(a) CS9:	ATGGATAACGTCCTTCCTGT	
(b) CS11:	GTACCACATCTGTGTGACCT	
(c) CS12:	GTGGTCAAAATGATGATCGT	
(d) CS13:	GTTCCGTCTGGGCTTCAAGC	

Figure 4.17: Sequence of oligodeoxynucleotides CS9, CS11, CS12 and CS13. The oligodeoxynucleotides hybridise to mouse NK-1R exons 1, 3, 4 and 5 respectively. The sequence of exon 2-specific oligodeoxynucleotide CS10 is detailed in Section 3.7.1. Each of the oligo deoxynucleotides is shown in the 5'-3' orientation.

4.7.3 Exonic Content of Clones SPR1 to SPR14

A nylon membrane bearing λSPR1-14 clone DNA restricted with *BamH* I (Section 4.8.1) was probed with oligodeoxynucleotides CS9, CS10, CS11, CS12 and CS13 in turn. The hybridisations were performed as described in section 2.29.2. The positive and negative controls in each case indicated that hybridisation had been successful. In addition to determining the number of exons in each NK-1R clone, the hybridisations served to identify DNA fragments of a suitable size for plasmid subcloning.

Exon 1-specific probe CS9 was found to hybridise to clones λSPR8 and λSPR13 only. A common 8.5 kb *BamH* I/*BamH* I fragment from each of these clones was identified as containing the mouse NK-1R exon 1 DNA sequence. The 8.5 kb *BamH* I/*BamH* I fragment of λSPR13 was selected for further analysis and was subcloned.

The exon 2-specific probe (CS10) failed to hybridise to any of the 14 λSPR clones. However, the controls indicated that the hybridisation had worked. It was therefore concluded that none of the clones contained the mouse NK-1R exon 2 sequence. As a result of this finding the bacteriophage clones isolated as a result of the primary screen of the Toronto library were re-hybridised but with probe CS10 (Section 3.7). Two bacteriophage plaques which positively hybridised for exon 2 sequence were successfully isolated in this way and were named λ SPR15 and λ SPR16.

The exon 3-specific oligodeoxynucleotide CS11 was found to hybridise to 10 clones of which 6 were distinct, namely λSPR1/2/5/7/9 and λSPR10. A positive signal emanated from a 5.1 kb DNA fragment of these clones. The 5.1 kb *BamH*I fragment of λSPR12 was subcloned and analysed further.

Exon 4-specific probe CS12 was found to hybridise to 12 of the 14 clones. Of those, only 7 clones were distinct (namely clones λSPR1/2/3/5/7/9 and 10). NK-1R exon 4 was also localised to 5.1 kb *BamH* I/*BamH* I fragment of λSPR12 as exon 3 (above).

Exon 5-specific oligodeoxynucleotide CS13 hybridised to the same seven clones as for exon 4. Exon 5 was also localised to the 5.1 kb *BamH* I/*BamH* I fragment of λSPR12 which was subcloned and analysed in depth later in this text.

Figure 4.18: Mouse NK-1R Exons Contained in the Isolated Bacteriophage Clones

	Exon 1	Exon 2	Exon 3	Exon 4	Exon 5
λSPR 1 *			√	√	√
λSPR2			√	√	√
λSPR3 **				√	√
λSPR4 *			√	√	√
λSPR5 °			√	√	
λSPR6 **				√	√
λSPR7			√	√	√
λSPR8	√				
λSPR9 •			√	√	√
λSPR10			√	√	√
λSPR11 °			√	√	
λSPR12 •			√	√	√
λSPR13	√				
λSPR14 °			√	√	
λSPR15 +		√			
λSPR16 +		√			

Figure 4.18: The NK-1R exons contained in the NK-1R phage clones isolated from the Toronto library by aqueous hybridisation using oligodeoxynucleotide probes. Clones with common superscript symbols were shown by digestion to contain common inserts.

4.8 SUBCLONING THE MOUSE NK-1R GENE

The above table (4.12) demonstrates that no single bacteriophage clone contained all five mouse NK-1R exons. Therefore, more than one DNA fragment had to be subcloned to allow a more detailed characterisation of the mouse NK-1R gene.

4.8.1 Subclone pSPR8.5

The 8.5 kb *BamH* I/*BamH* I fragment of λSPR13 was shown by hybridisation to contain NK-1R exon 1 sequence. In order to subclone this fragment, 10 μg of λSPR13 DNA was digested with *BamH*I and electrophoresed on a TAE buffer/agarose gel. The 8.5 kb DNA fragment was recovered from the gel using a Spin-X column (Section 2.27.2). The plasmid vector pBluescriptII KS(+) was prepared in the same way. The isolated DNA fragments were used in ligation reactions. The ligations were transformed into *E. coli* strain TG1 bacterial cells (which demonstrate high transformation efficiencies). White colonies were picked and cultured. DNA was prepared from each culture and restricted with *BamH* I to determine the insert size. A plasmid with an 8.5 kb insert was identified. The subclone with the 8.5 kb λSPR13/*BamH* I fragment as an insert was named pSPR8.5.

4.8.2 Subcloning the NK-1R Exon 2: pSPR6.8

Bacteriophage clones containing NK-1R exon 2 sequence (λSPR15 and λSPR16) were isolated from the Toronto library (Section 3.7) using an exon 2-specific oligodeoxynucleotide probe (CS10). In order to select a suitable fragment for subcloning, DNA was prepared from bacteriophage λSPR15 and λSPR16 lysates and incubated separately with the endonucleases *BamH* I, *Nco* I and *EcoR* V. The restriction reactions were electrophoresed capillary blotted The membrane was hybridised with probe CS10 as detailed in section 3.7.2 (a control DNA dot-blot was

included and was successful). The resulting autoradiograph confirmed the existence of a common genomic DNA insert in both λSPR15 and λSPR16 (as indicated earlier by restriction analysis). The autoradiograph also demonstrated that a 6.8 kb *Nco* I/*Nco* I fragment of both λSPR15 and 16 hybridised positively for the mouse NK-1R exon 2. This fragment was subcloned into the polylinker *Nco* I site of pGEM5Zf(+) to generate the subclone pSPR6.8. This subclone was restriction mapped accordingly. Due to time restraints, the hybridisation experiment to localise the mouse NK-1R exon 2 within subclone pSPR6.8 was not conducted and remains outstanding.

4.8.3 Subclone pSPR5.1 (Exons 3, 4 and 5)

The 5.1 kb *BamH* I/*BamH* I fragment of λSPR12 was shown to contain SPR sequence from exons 3, 4 and 5 (Section 4.8.3). This DNA fragment was subcloned in the same way, and into the same vector, as the 8.5 kb fragment (described in Section 4.9.1). The plasmid subclone with the 5.1 kb λSPR12 *BamH* I/*BamH* I fragment as an insert was named pSPR5.1.

4.9 Restriction Mapping the NK-1R Subclones pSPR8.5 and pSPR5.1

The mouse NK-1R gene subclones, plasmids pSPR8.5, pSPR6.8 and pSPR5.1, were restriction mapped in the same way as the NK-1 plasmid subclones (Section 4.4) where known restriction sites in the vector polylinker were used to orientate restriction sites in the insert. In order to generate a restriction map of each subclone, restriction reactions were electrophoresed

Fragment

sizes were then translated into a linear map. The DNA was then capillary blotted from the gel on to nylon membrane and immobilised. The membranes were stored at 4°C for later use.

4.10 LOCATING THE MOUSE NK-1R EXONS

The Southern blotted membranes generated in section 4.10 and bearing subclone DNA were probed using oligodeoxynucleotide probes specific to the NK-1R exons (as in section 4.8.2). The probes were labelled using polynucleotide kinase (section 2.24.3).

4.10.1 Mouse NK-1R Exon 1

Membrane-bound restricted pSPR8.5 DNA was hybridised with labelled oligodeoxynucleotide CS8. The probe hybridised to the 1.5 kb *Sst* I/*Sst* I and the 1.7 kb *BamH* I/*Pst* I fragments of this subclone. Therefore, mouse NK-1R exon 1 was localised to the 250 bp *Sst* I/*Pst* I fragment of pSPR8.5.

4.10.2 Mouse NK-1R Exon 3

When hybridised with an exon 3-specific probe (CS11), several restricted DNA fragments of subclone pSPR5.1 generated a positive signal. From this, NK-1R exon 3 was localised to the 600 bp *Pvu* II/*Ava* II fragment of subclone pSPR5.1.

4.10.3 Mouse NK-1R Exon 4

The restricted pSPR5.1 DNA membranes were stripped and re-hybridised with oligodeoxynucleotide CS12. Consequently, mouse NK-1R exon 4 was localised to the 400 bp *Bgl* II/*Ava* II fragment of subclone pSPR5.

4.10.4 Mouse NK-1R Exon 5

To locate mouse NK-1R exon 5, the same hybridisation as for exon 4 (above) was performed, but using the oligodeoxynucleotide CS13 as the probe. Exon 5 was localised to the 500 bp *Ava* II/*EcoR* I fragment of subclone pSPR5.1 (lane 2, Figure 4.14, below). To further confirm this, the probe was shown to hybridise to both the 700 bp *EcoR* I/*EcoR* I (lane 4) and the 1.6 kb *Ava* II/*Xho* I fragment (lane 8).

Figure 4.19: Restriction Map of NK-1R clones pSPR5.1, pSPR8.5 & pSPR6.8.

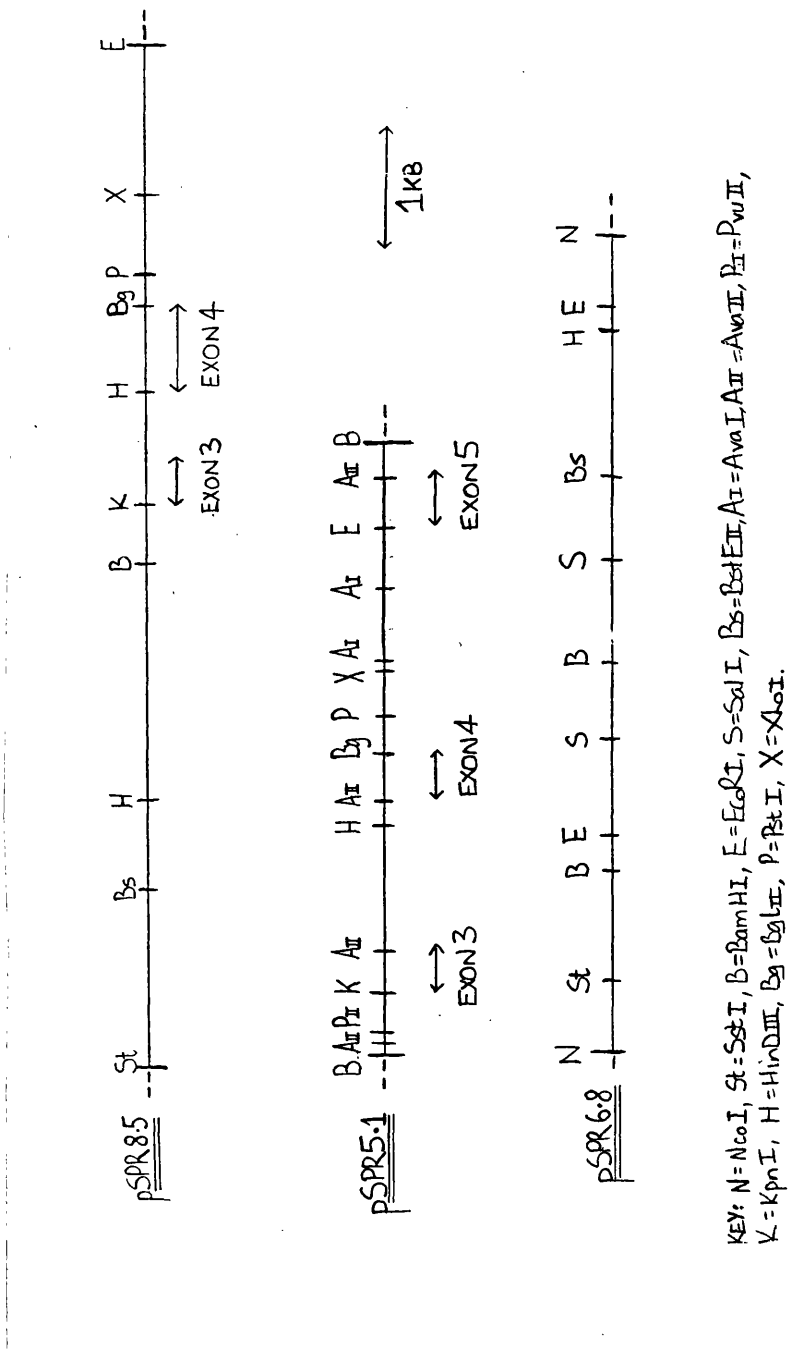


Figure 4.19: Linear representations of the endonuclease recognition sites contained with the NK-1R plasmid subclones pSPR5.1, pSPR8.5 and pSPR6.8.

Figure 4.20 : Subclone pSPR5.1Hybridised with Exon-5-Specific Probe CS13

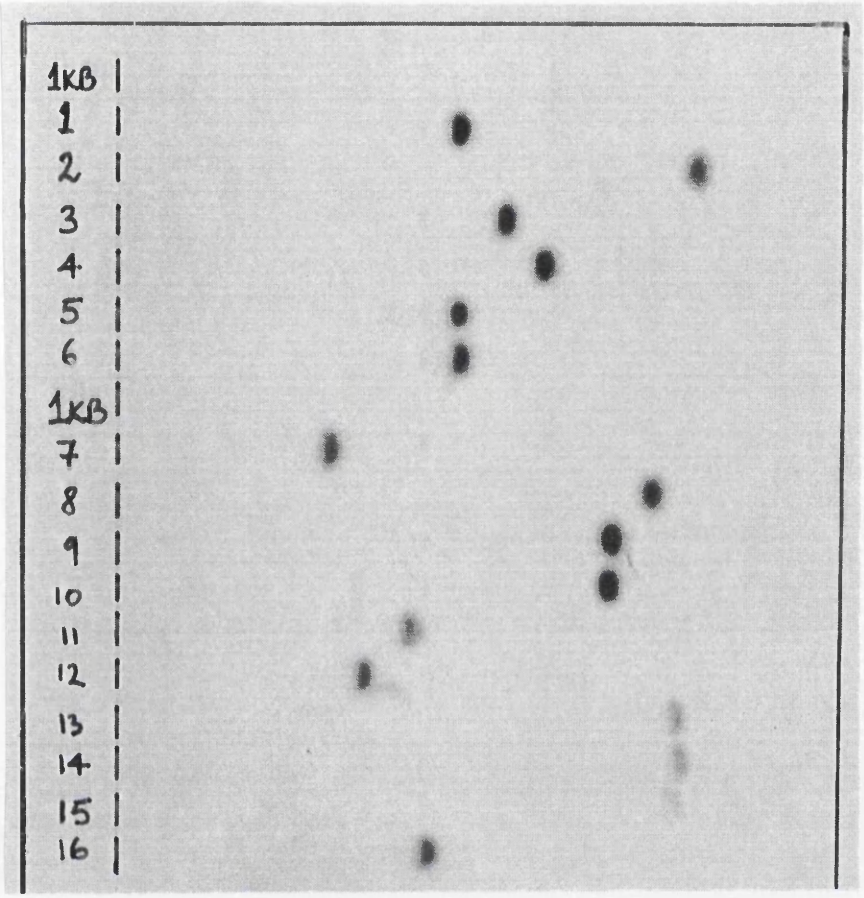


Figure 4.20: Autoradiograph of 0.5 µg of subclone pSPR5.1 DNA digested with endonucleases, electrophoresed, capillary blotted and hybridised with exon-5-specific probe CS13. Lanes '1KB' contain 1 kb DNA ladder. Lane 1; *Ava* II/*Sst* I, lane 2; *Ava* II/*Eco*R I, lane 3; *Ava* II/*Pst* I, lane 4; *Ava* II/*Xho* I, lane 5; *Ava* II/*Sst* I, lane 6; *Ava* II, lane 7; *Nco* I, lane 8; *Nco* I/*Eco*R I, lane 9; *Nco* I /*Pst* I, lane 10; *Nco* I /*Xho* I, lane 11; *Nco* I /*Sst* I, lane 12; *Stu* I, lane 13; *Stu* I /*Eco*R I, lane 14; *Stu* I /*Pst* I, lane 15; *Stu* I/ *Xho* I, lane 16; *Stu* I/ *Sst* I.

CHAPTER 5

RESULTS: GENE TARGETING VECTORS

INTRODUCTION

Chapter 5 describes the steps taken to construct gene targeting vectors for the cloned mouse PPT-A and NK-1R genes. The vectors were constructed to facilitate homologous recombination experiments in tissue cultured mouse embryonic stem (ES) cells with a view to generating transgenic mice lacking the functional form of either the PPT-A or the NK-1R gene. Two gene targeting vectors were made for the PPT-A gene whereas only one NK-1R construct was completed. The genes are dealt with in turn where the PPT-A gene is detailed first and is followed by the NK-1R gene.

GENE TARGETING VECTORS FOR THE MOUSE PPT-A GENE

5.1 DESIGN AND CONSTRUCTION OF PPT-A GENE TARGETING VECTOR No. 1

When designing gene targeting vectors, several basic criteria are taken into account. These important criteria are believed to have a direct and significant bearing on the ability of a particular vector to promote specific, targeted DNA recombination events. One such criteria refers to the size of the two genomic DNA fragments which

flank the neomycin resistance gene and which constitute regions of homology between the genomic target DNA and the genomic DNA contained in the targeting vector. When the PPT-A vectors were designed, in order to get a workable targeting frequency, the minimum combined length of those two flanking fragments which was thought to be in the region of 6 kb. The recommended ideal length is approximately 10 kb. Above 15 kb, the targeting frequency remains roughly the same. Additonally, the higher the degree of conservation within the flanking fragments, the higher the targeting frequency. In order to comply with this criterion, a single DNA fragment containing all seven mouse PPT-A exons was subcloned. The targeting efficiency has also been shown to be higher when isogenic DNA is used, where the genomic DNA in the vector is derived fom the same strain of mouse as the ES cells in to which they are to be transfected.

5.1.1 Subclones pSP10.6

Plasmid subclones pSP7 and pSP9 (Section 4.3) were generated by subcloning 7 kb and 9 kb *Nco* I-*Nco* I restricted DNA fragments of bacteriophage clone λPPT9 respectively. From the restriction maps of these subclones, it was noted that each contained a single *Sst* I restriction site estimated to be separated by 10.6 kb of genomic DNA (based on the assumption, later proved correct, that the two cloned inserts were contiguous genomic sequences). Restriction analysis of λPPT9 confirmed this size estimate. One *Sst* I site was found to lie upstream of exon 1 while the other was downstream of exon 7. This 10.6 kb *Sst* I/*Sst* I fragment was later shown (by restriction map alignment with subclones pSP7 and pSP9, see Figure 5.1) to contain all 7 mouse PPT-A exons and was subcloned into the vector pBluescript KS$^+$. This subclone was named pSP10.6, the restriction map for which is shown below.

Figure 5.1 : Alignment of PPT-A Subclones SP9, SP7, SP10.6 and pSP7.2

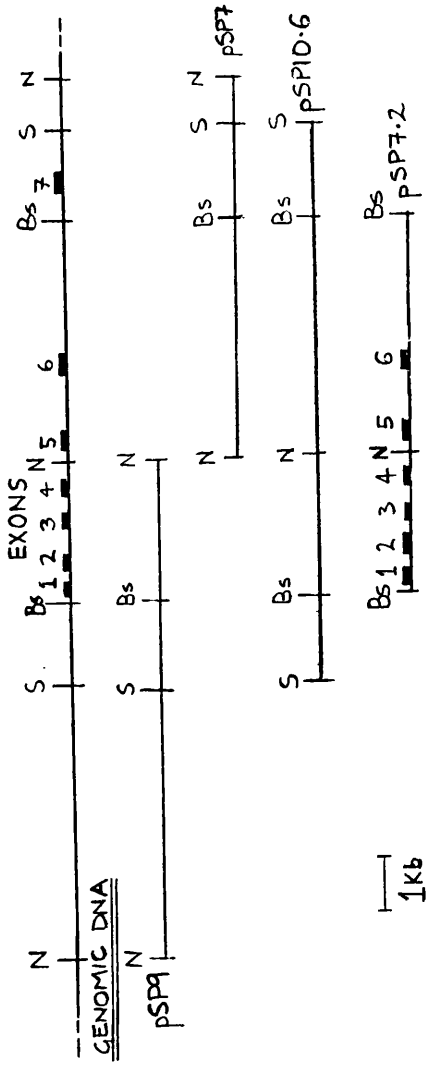


Figure 5.1: Representation of the PPT-A plasmid subclones pSP9, pSP7, pSP10.6 and pSP7.2 aligned to indicate the extent of insert overlap relative to the genomic organisation of the mouse PPT-A gene. See figure 4.6 for detailed restriction maps of plasmid subclones pSP7 and pSP9. N; *Nco* I, S; *Sst* I, Bs; *Bst* E II.

5.1.2 Subclones pSP7.2 and pSP7.2 II

A smaller fragment of the pSP10.6 subclone was selected for further subcloning on the basis of the restriction sites that would be available for the insertion of the *neo* and thymidine kinase resistance genes. These genes are inserted at a later stage in the vector construction. From the restriction map of pSP10.6, two *BstE* II sites were noted lying immediately upstream of exon 1 and upstream of exon 7. A *BstE* II digestion of pSP10.6 generated a 7.2 kb fragment which lacked the PPT-A exon 7 but was of a appropriate size and sequence for use in a targeting vector. In order to subclone this 7.2 kb *BstE* II/*BstE* II fragment of subclone pSP10.6, a suitable plasmid vector with a *BstE* II site in the multiple cloning site (MCS) was required. No such vector was commercially available so one was constructed by introducing a *BstE* II restriction site into the MCS of pBluescriptKS+.

To do this, two oligodeoxynucleotides were designed so as to anneal to each other and in so doing, generate a *BstE* II recognition sequence flanked by overhanging *Kpn* I cohesive termini. This would result in the formation of a short double stranded DNA linker sequence suitable for cloning into the *Kpn* I site of pBluescriptKS+. The sequence of the oligodeoxynucleotides, and the linker they combine to create, is shown below (Figure 5.2).

Oligodeoxynucleotides CS19 and CS20 were synthesized by Cruachem Ltd. and supplied as quantified, precipitated DNA pellets. Both DNAs were resuspended, annealed and phosphorylated as described in section 2.32.2. Conversely the pBluescript KS+ DNA was restricted with *Kpn* I, The linearized DNA was then dephosphorylated by shrimp alkaline phosphatase (SAP, Section 2.22.2) which was later heat inactivated. The linker DNA was ligated to the linearized pBluescript DNA. The ligation reactions included a range of vector to insert DNA ratios (from 1:1 to 1:10) and the appropriate positive and negative controls which were demonstrated by way of transformation to have worked. A 2 µl sample of the ligation was transformed into *E. coli* (TG1 strain) and

bacterial clones were selected, cultured and analysed by restriction with *BstE* II. This allowed identification of a plasmid which was linearized by *BstE* II and therefore contained the newly introduced *BstE* II restriction site (the short linker sequence). This recombinant plasmid was named pKS+.BstE II.

Plasmid pKS+.BstEII was used as the recipient DNA for the 7.2 kb *BstE* II-*BstE* II fragment of plasmid pSP10.6 (Section 5.1.1) which was cloned to generate the subclone named pSP7.2 (version 1). The authenticity of this subclone was verified by restriction analysis and alignment of its restriction map with that of subclones pSP10.6, pSP9 and pSP7.

At a later stage, and as a means of increasing the ligation efficiency of a troublesome ligation, plasmid pSP7.2 was modified to create a derivative subclone named pSP7.2 II. Plasmid pSP7.2 II differed from the original plasmid pSP7.2 (version 1) in that the 900 bp *Mlu* I/*Mlu* I restriction fragment (which contained exons 3 and 4) was removed from pSP7.2 (version 1) by *Mlu* I digestion. The *Mlu* I-digested DNA was then ligated to itself and transformed into *E. coli* TG1 cells. One re-ligated plasmid transformant lacking the 900 bp *Mlu* I-*Mlu* I fragment was isolated and named pSP7.2 II (version 2).

Figure 5.2: Sequence of Oligodeoxynucleotides CS19 and CS20

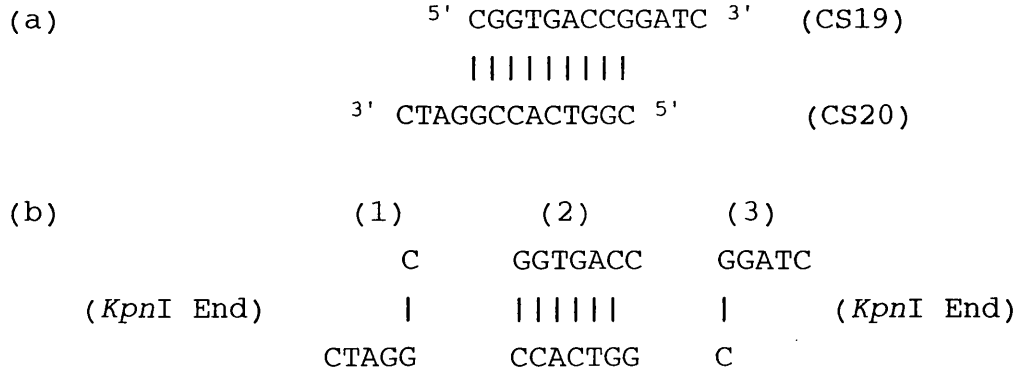


Figure 5.2: Sequence of the oligodeoxyonucleotides CS19 and CS20. Part (a) shows the oligodeoxynucleotides annealed to each other, and part (b) shows the linker diagramatically sub-divided to show the component sections; (1) a *Kpn* I cohesive terminal, (2) a *BstE* II recognition sequence, and (3) a*Kpn* I cohesive terminal.

Figure 5.3: Restriction Map of Subclones pSP7.2, pSP7.2 II, pSP7.2 III and pSP7.2 IV

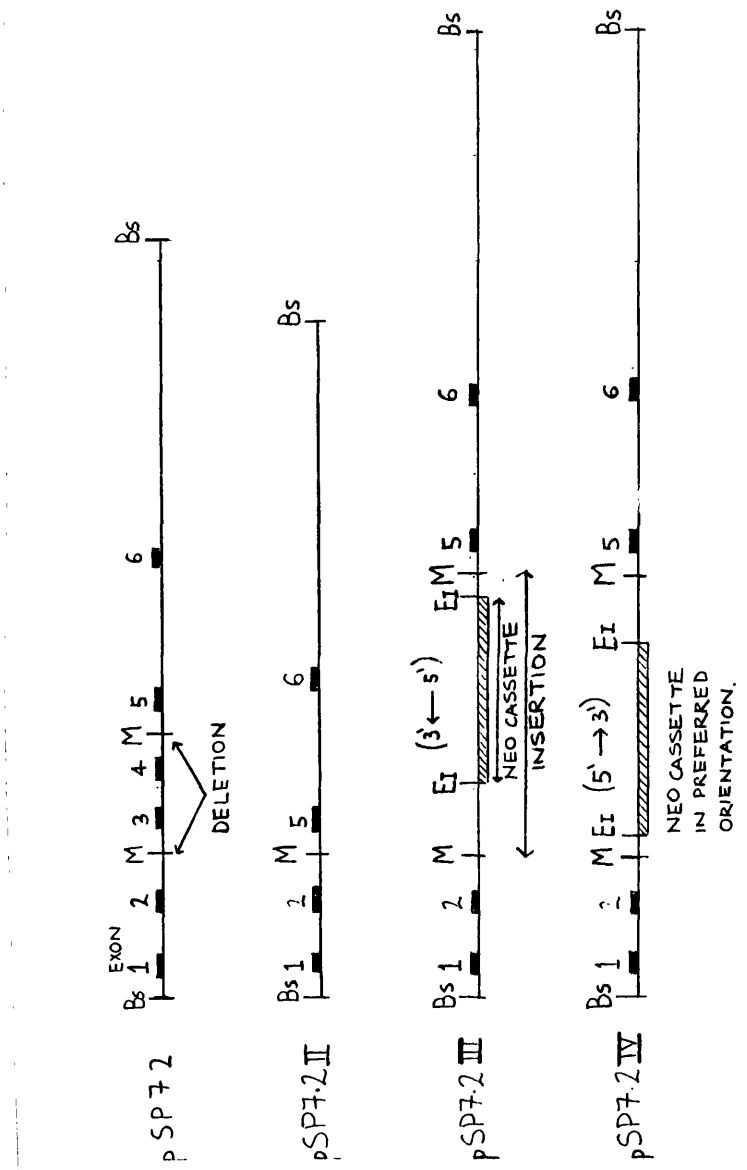


Figure 5.3: Linear representations of the endonuclease recognition sites in the insert of the PPT-A subclones pSP7.2, pSP7.2 II (for both, see Section 5.1.2), pSP7.2 III and pSP7.2 IV (for both, see sections 5.1.4). M; *Mlu* I, Bs; *Bst* E II, E1; *Eco* R I.

5.1.3 Cloning the Neomycin Resistance Gene into PPT-A Targeting Vector No. 1

5.1.3.1 Subcloning PPT-A Exons 3 and 4: Plasmid pSP900

As detailed in Chapter 1, the PPT-A gene encodes several neuropeptides. The *neo* gene was inserted into the targeting vector in order to disrupt the transcription of the gene and to prevent (in a targeted mouse) the subsequent translation of the neuropeptide NK-1 specifically. The selection of a partcular genetic location for the insertion of the *neo* gene within the targeting vector was therefore important. An ideal insertion site to do this lay within the NK-1-encoding exon 3 of the PPT-A gene. On examination of the restriction map of subclone pSP9 (containing PPT-A exons 1, 2, 3 and 4), two 'unique' *Mlu* I restriction sites were noted which were shown to flank exons 3 and 4. The *Mlu* I sites were found to be 900 bp from each other and were located in introns 2 and 4. The 900 bp *Mlu* I-*Mlu* I fragment of subclone pSP9 was subcloned into the 2.9 kb plasmid vector pIBI24 (Dente *et al*, 1983) to generate the subclone which was named pSP900 (version 1, see Figure 5.4). Using the published mouse PPT-A cDNA sequence and the University of Winconsin GCG sequence analysis computer software, a restriction map of exons 3 and 4 was assembled. The authenticity of the insert was confirmed both by digestion and by PCR analysis using the primers specific to exons 3 and 4, primers CS6 and CS7 respectively (as before, see Section 3.1).

5.1.3.2 Deletion of the Polylinker *Pst* I Site: pSP900 II

A *Pst* I site in exon 3 was noted and confirmed (by restriction enzyme digestion) to be unique within the *insert* of plasmid pSP900. Based on the published mouse cDNA sequence (Kako *et al*, 1993) this *Pst* I site was shown to be located 35 bp from the 5' end of the 97 bp exon 3 and 13 bp immediately upstream of the short exonic sequence which is translated to produce the 11 amino acid neuropeptide NK-1. This site was chosen as the insertion site for the *neo* gene because the upstream insertion location would prevent the transcription of the downstream NK-1-encoding sequence. However, the multiple

cloning site (MCS) of plasmid pSP900 (derived from the vector pIBI24) also contained a *Pst* I recognition sequence. This second *Pst* I site meant that the one located within exon 3 could not be considered unique and as such would complicate the cloning of the *neo* gene into the exonic *Pst* I site. To ensure this site remained unique, the polylinker *Pst* I site of subclone pSP900 was deleted. To do this, another two unique restriction sites, namely *Sal* I and *Xho* I sites, were used. DNA cleaved by *Sal* I generates cohesive termini which are compatible with those generated by *Xho* I. Those termini can therefore be ligated to each other. In this instance, this approach would result in the removal of the problematic *Pst* I site lying between the *Sal* I and *Xho* I sites in the pSP900 MCS. This approach was successfully used to generate the plasmid subclone pSP900 II (version 2; i.e., a modified second version of the original plasmid subclone pSP900). Plasmid pSP900 II can therefore be linearised by incubation with *Pst* I (see Figure 5.4).

5.1.3.3 Deletion of the Polylinker *EcoR* I Site: pSP900 III

Although the unique *Pst* I site located in PPT-A exon 3 (which was introduced into plasmid pSP900 II) was chosen as the site for insertion of the *neo* gene, the 1.8 kb *neo* cassette itself was only available when released from the *EcoR* I site in the polylinker of vector pBluescriptSK+ (Dr. D. Donald, pers. comm.). To utilise this cloning feature, an *EcoR* I site was introduced into the *Pst* I site of pSP900 II using a linker (see Section 5.1.3.4) in order that the *neo* gene could be introduced directly. Conversely, the *EcoR* I site in the polylinker of plasmid pSP900 II was deleted. To do this, 2 µg of pSP900 II DNA was cut with *EcoR* I,

The protruding DNA termini were then blunt-ended by incubation with dNTPs and the (DNA polymerase I fragment) Klenow enzyme.

A modified pSP900 II plasmid clone which was not linearized by the endonuclease *EcoR* I, and therefore lacked the polylinker *EcoR* I recognition sequence, was identified, and was named pSP900 III (see Figure 5.4).

Figure 5.4: Restriction Maps of pSP900, pSP900 II, pSP900 III and pSP900 IV

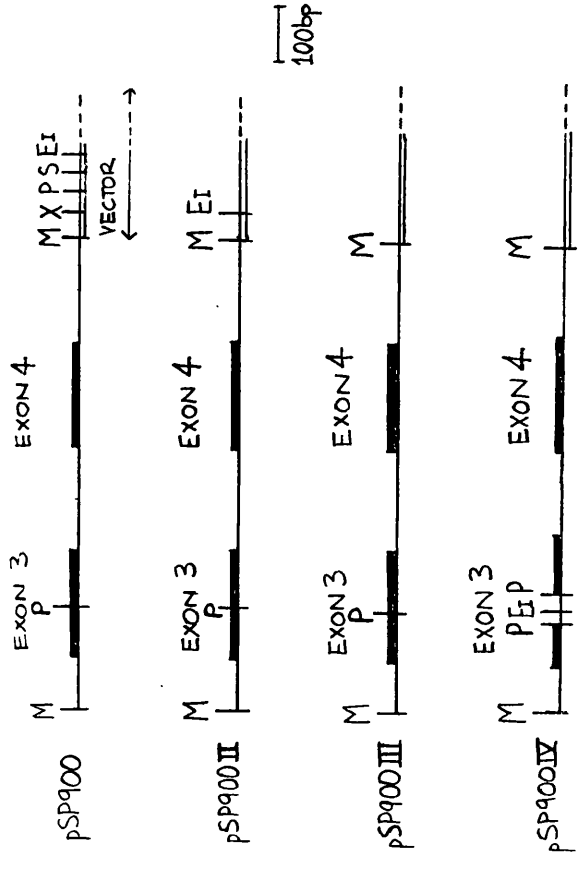


Figure 5.4: Linear representations of the PPT-A plasmid subclones pSP900, PSP900 II, pSP900 III and pSP900 IV (for details see Sections 5.1.3.1, 5.1.3.2, 5.1.3.3 and 5.1.3.4 respectively). The subclones are intermediates in the construction of the NK-1 targeting vector no. 1. X; *Xho* I, S; *Sal* I, M; *Mlu* I, P; *Pst* I, E1; *Eco* R I.

5.1.3.4 Insertion of an *EcoR* I Site into Exon 3: pSP900 IV

A small double stranded DNA linker sequence (shown below) was ligated into the *Pst* I site of subclone pSP900 III in order to introduce an *EcoR* I site at that location within exon 3. Two oligodeoxynucleotides were designed for this purpose. When annealed, the complimentary oligodeoxynucleotides generated a *Pst* I cohesive terminal at one end and a *EcoR* I cohesive terminal at the other end of the resultant double stranded DNA linker sequence.

Figure 5.5: Design and Sequence and Componeht parts of the *Pst* I- *EcoR* I linker

(a)

```
          5'GTATCCTAGAGG 3'
            ||||||||||||
3' ACGTCATAGGATCTCCTTAA 5'
```

(b)

```
                (1)      (2)        (3)
                  G    TATCCTAGAG    G
(Pst I End)       |    ||||||||||    |      (EcoR I End)
              ACGTC    ATAGGATCTC    CTTAA
```

(c)

EcoR I Site

```
                GTATCCTAGAGGAATTCCTCTAGGATACTGCA
(Pst I End)     ||||||||||||||||||||||||||||||      (Pst I End)
          ACGTCATAGGATCTCCTTAAGGAGATCCTATG
```

Figure 5.5: Sequence of the *Pst* I-*EcoR* I linker comprising two oligodeoxynucleotides which were synthesized, annealed to each other, phosphorylated, self-ligated, redigested with *Pst* I and cloned into the *Pst* I site of the plasmid pSP900 III to generate a modified plasmid plamid named pSP900 IV. Part (a) shows the oligodeoxynucleotides annealed, whereas part (b) shows the linker sub-divided to demonstrate the component sections; (1) a *Pst* I cohesive terminal, (2) 10 bp of random, complimentary sequence, and (3) an *EcoR* I cohesive terminal. Part (c) shows the arrangement of the linker after self-ligation and restriction with *Pst* I. In this form, an intact, central *EcoR* I recognition sequence was generated flanked by two *Pst* I cohesive termini. This was therefore ideal as a means introduce an *EcoR* I site into a chosen *Pst* I site. In part (c), the ligated linkers are differentiated from each other using bold text. The *EcoR* I recognition sequence is underlined.

Originally, the strategy for cloning the *neo* gene into the (*Pst* I site) of the third exon of the PPT-A gene involved ligating the above single linker (Figure 5.5, part b) onto the compatible *EcoR* I-restricted ends of a *neo* gene DNA fragment. This DNA fragment was generated by restriction of the *neo* gene plasmid subclone, pSK+*Neo*, with *EcoR* I. This cloning step proved difficult whereapon an alternate strategy was adopted.

The above oligodeoxynucleotides were synthesized by Cruachem Ltd. and 200 ng of each DNA was annealed to the other as described in Section 2.32.2. The annealed double stranded DNA linker was phosphorylated by polynucleotide kinase (Section 2.21.1) and ligated to itself under standard conditions. The ligation was demonstrated to have been successful by electrophoresis and visualisation of a sample on a 1.5 % agarose gel. Of the total 10 µl ligation reaction, 2 µl was electrophoresed alongside a sample of non-ligated linker. The smear which was visualised in the 'ligated linker' lane, but not in the other, indicated the existence of a range of multimers of this short DNA linker successfully ligated to itself. Obviously, the linkers had the capability to self-ligate at both the *Pst* I and *EcoR* I ends. However, only linker doublets constituting *Pst* I sticky ends flanking an intact *EcoR* I site were required for cloning purposes. Therefore, a sample of extended linker ligation reaction was restricted with *Pst* I to generate the extended linker shown above in figure 5.5, part (c). To remove any remaining endonuclease the reaction was phenol and chloroform extracted, ethanol precipitated overnight and resuspended in dH_2O to a concentration of 10 ng.µl. This cleaned-up extended linker DNA was then incubated under 'conventional' (see Section 2.21.1) ligation conditions with a sample of plasmid pSP900 III DNA which had been restricted with *Pst* I and dephosphorylated with SAP (Section 2.22.2). A number of ligation reactions, including controls, were performed including a range of insert to vector DNA ratios. A 5 µl sample of ligation reaction was transformed into competent *E. coli* TG1 cells. Ten colonies were picked and cultured overnight. DNA was extracted and then analysed, in the first instance, by digestion with *EcoR* I. 9 of the 10 clones analysed were indeed linearized by *EcoR* I. Clone number 1 was further analysed with the endonucleases *Mlu* I and *Pst* I (plus *EcoR* I

again and combinations of the three) and was shown to contain the correct extended linker insert. This clone was named pSP900 IV.

5.1.3.5 Insertion of the Neomycin Resistance Gene: pSP900 V and pSP900 VI

To summarise, subclone pSP900 (version 1) was modified several times (as detailed above) so as to create a plasmid which contained a unique *EcoR* I site within exon 3 and upstream of the NK-1-encoding sequence. This was achieved in the form of the plasmid pSP900 IV (version 4). The 1.8 kb 'neomycin resistance gene, as mentioned briefly in section 5.1.3.4, was released from its vector plasmid by *EcoR* I restriction and was cloned into the *EcoR* I site of pSP900 IV.

5 μg of plasmid pSP900IV was linearized with *EcoR* I to generate a 3.8 kb DNA fragment which was recovered from an agarose gel by the Spin-X method (Section 2.27.2) and dephosphorylated with SAP. The two DNA fragments were ligated under standard conditions and a 2 μl sample of the reaction was transformed into *E. coli* TG1 cells. 20 colonies were selected, cultured and analysed by restriction with *EcoR* I. Several positive recombinant clones bearing the 1.8 kb *neo* gene inserted into the *EcoR* I (formerly *Pst* I) site of plasmid pSP900 IV were identified. One clone was shown (using known restriction sites in the insert DNA) to contain the insert in the *opposite* 5'/3' orientation as the PPT-A gene into which it was inserted and was named pSP900 VI (version 6). In another clone, the 5'/3' orientation of the newly inserted neomycin resistance gene was the same as for the PPT-A gene into which it was inserted. This plasmid was named pSP900 V (version 5) and was used in the subsequent cloning step.

Figure 5.6: Restriction Maps of PPT-A Subclones pSP900 IV, pSP900 V & pSP900 VI

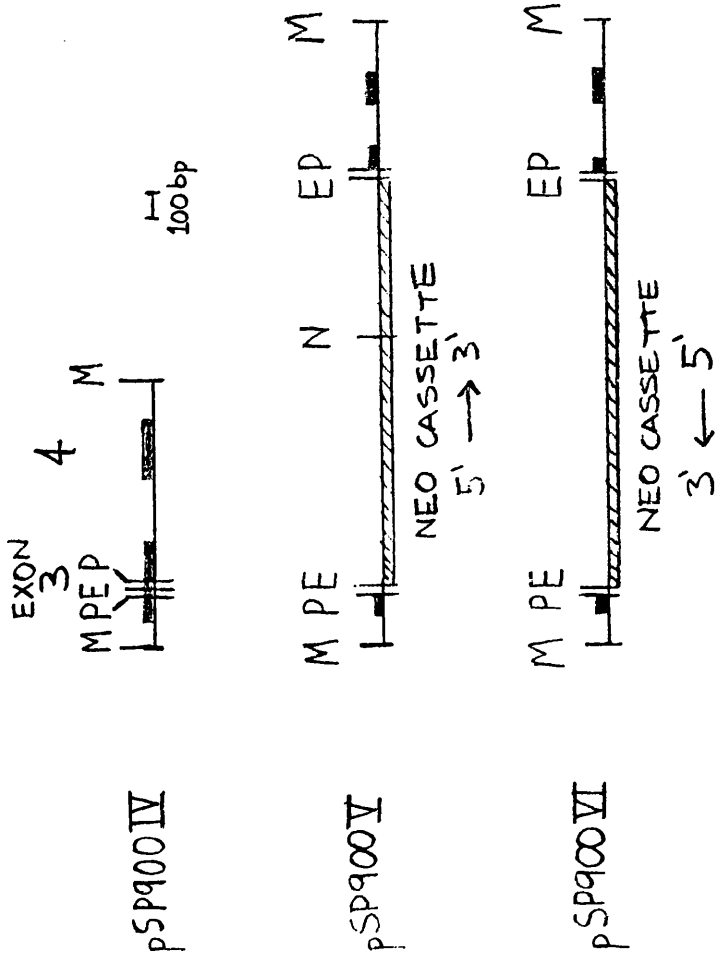


Figure 5.6: Linear maps of the PPT-A subclones pSP900 IV, pS900 V and pSP900VI (for details, see sections 5.1.3.4, 5.1.3.5 and 5.1.3.5 respectively). Plasmid pSP900 V contains the *neo* gene cloned in the correct 5' to 3' orientation for use in the next step of targeting vector construction. M; *Mlu* I, P; *Pst* I, E1; *Eco* R I.

5.1.4 Re-assembly of the PPT-A Gene: pSP7.2 III and pSP7.2IV

To recapitulate, originally, the 900 bp *Mlu* I-*Mlu* I fragment of plasmid pSP9, which contained PPT-A exons 3 and 4, was subcloned to generate the plasmid pSP900 (Section 5.1.3.1). This subcloned 900 bp fragment was modified so as to incorporate a *neo* cassette at a precise location in exon 3 (generating plasmid pSP900 V, section 5.1.3.5). This 900 bp fragment plus the *neo* insert amounted to the 2.7 kb total insert of plasmid pSP900 V. This 2.7 Kb was then cloned into the *Mlu* I site of plasmid 'pSP7.2 II (Section 5.1.2). Consequently, the original genomic organisation of the mouse PPT-A gene (between exons 1 and 6) was re-assembled except for the desired insertion of the *neo* gene into exon 3 (and the unavoidable introduction of the short DNA linker). This cloning step generated subclones pSP7.2 III and pSP7.2 IV which differed only in the orientation of the 2.7 Kb *Mlu* I/*Mlu* I insert (see Figure 5.3).

To do this, plasmid pSP900 V was digested with *Mlu* I to release the 2.7 Kb *neo*/NK-1 exons 3 and 4 insert which was purified as before. The plasmid pSP7.2 II (which differed from the original pSP7.2 in that it lacked the 900 bp *Mlu* I-*Mlu* I fragment, (Section 5.1.2) was linearized with *Mlu* I to generate a 9.3 kb fragment (which was dephosphorylated with SAP). The 2.7 and 9.3 kb *Mlu* I-*Mlu* I fragments were ligated under conventional ligation condition (Section 2.21.1) and transformed into *E. coli* nova blue competent cells. Transformants were picked, cultured and analysed by restriction with several endonucleases to confirm both the presence of the 2.7 kb *Mlu* I insert and its orientation. The insert was noted in two different clones but in the opposite orientation to each other. The modified pSP7.2 II plasmid which contained the 2.7 kb insert from plasmid pSP900 V in the *opposite* 5' to 3' orientation to the PPT-A gene into which it was inserted, was named pSP7.2 III. The recombinant plasmid with the 2.7 Kb insert in the *same* orientation as the PPT-A gene into which it was inserted was named pSP7.2 IV. This plasmid was of particular interest because it was further modified (see below).

5.1.5 Insertion of the Thymidine Kinase Gene

In order to select positive recombinant ES cell clones which had been generated as a result of homologous recombination with the completed PPT-A targeting vector, the thymidine kinase (TK) gene from *herpes simplex* virus 1 (HSV1) was introduced into the targeting vector pSP7.2 IV for the purpose of counter-selection. The TK gene was inserted into the polylinker of plasmid pSP7.2 IV so that it immediately flanked one of the two genomic DNA fragments in that plasmid. (which in turn flanked the recently inserted *neo* gene, section 5.1.3.5). The 2.9 kb TK gene was cloned from the 7239 bp plasmid pPNT (E. Gallagher, pers. comm.) into the *EcoR* I and *HinD* III sites of plasmid pGEM9 (Promega Corp.) upon which the plasmid was re-named pGEM9.TK.

In order to clone the TK gene from plasmid pGEM9.TK into the polylinker of plasmid pSP7.2 IV, the remaining unique restriction sites in the polylinker of pSP7.2 IV had to be used. To permit the accurate introduction of the TK gene, each of the chosen enzymes had to cleave only one site in the plasmid pSP7.2 IV (without excising the whole insert). The *Xho* I and *Not* I polylinker sites were found to be intact and suitable for this purpose.

The TK gene was re-cloned from the *EcoR* I and *HinD* III sites of plasmid pGEM9.TK into the same restriction sites in plasmid vector pBluescriptKS$^+$ to generate the plasmid pKS.TK. The pBluescript vector was chosen because the polylinker *EcoR* I and *Hind*III sites are flanked by *Not* I and *Xho* I sites respectively. The TK gene could therefore be re-released by a double restriction with *Not* I and *Xho* I and cloned into the corresponding sites in pSP7.2 IV. This cloning step was successfully performed and resulted in the completion of the construction of the first of the two mouse PPT-A gene targeting vectors. The resultant plasmid was named pSP7.2 V (version 5), a map of which is shown in Figure 5.7 (below). For use in cultured ES cell transfection experiments, the targeting vector was linearized by the endonuclease *Not* I.

5.2 CONSTRUCTION OF PPT-A GENE TARGETING VECTOR No. 2

In addition to the plasmid pSP7.2 V (Section 5.1.5), a second mouse PPT-A gene targeting vector was designed and constructed. This vector displayed a strong similarity to the first because it was derived from the same source plasmids. The plasmids pSP7.2 II, pSK.Neo and pKS.TK (Sections 5.1.2, 5.1.3.5 and 5.1.5 respectively) again supplied the PPT-A genomic DNA fragment, the *neo* gene and the thymidine kinase (TK) counter-selection gene respectively. Vector No. 1 was designed to allow the introduction of the *neo* gene at a precise (upstream) location within an exon as a means of disrupting, and ultimately preventing, the expresssion of the (downstream) neuropeptide neurokin-1, in a transgenic mouse. In doing this, the *neo* gene was *inserted* (Section 5.1) but no part of the genomic PPT-A DNA was *deleted.* Targeting vector number1 involved 12 cloning steps in the making and was therefore relatively difficult to assemble. However, as a consequence of pursu ing this vector design strategy, the foundation was put in place to allow the future construction of various other PPT-A gene targeting vectors (based on plasmids pSP7.2 and pSP900). Those future vectors could be used to generate more subtle targeted mutations. (see discussion chapter). The second PPT-A gene targeting vector, for example, was just such a variant vector based on the first vector. The second vector was more crude in design and involved the deletion of the third and fourth exons of the mouse PPT-A gene. The construction is detailed below. In contrast to the first vector, in which no genomic PPT-A DNA was deleted, the second vector was designed so that exons 3 and 4 were deleted and the *neo* gene inserted in their place. The net DNA insert was 900 bp (the 1.8 kb insertion minus the 900 bp deletion). Since exon 3 is responsible, when translated, for the production of the tachykinin neuropeptide neurokinin-1, successful homologous recombination using a vector which lacks exon 3 would introduce into the mouse germ line a form of the PPT-A gene which would be incapable of synthesizing NK-1.

5.2.1 Insertion of the Neomycin Resistance Gene: pPPT.Neo I and pPPT.Neo II

5 μg of plasmid pSK.Neo (Dr. D. Donald, pers. comm.) was digested with *EcoR* I to release the intact *neo* gene from its plasmid vector (see Section 5.1.3.5). The 1.8 kb *neo* gene was band isolated from a 1% agrose gel by phenol/chloroform extraction and was then ethanol precipitated. The isolated DNA was resuspended in 20 μl of dH_2O. The overhanging *EcoR* I cohesive termini of the recovered DNA fragments were then blunt-ended by incubation with dNTPs and Klenow enzyme. The polymerase enzyme was removed by extraction with phenol/chloroform.

Using a spectro-photometer, the DNA concentration was found to be 50 ng.μl.

5 μg of plasmid pSP7.2 II was linearized with the endonuclease *Mlu* I. Complete digestion was confirmed by electrophoresis of a sample on an agarose gel. The *Mlu* I enzyme was removed by phenol and chloroform extraction. The DNA fragments were precipitated, resuspended and blunt-ended as for the *neo* gene (paragraph above). The pSP7.2 II fragment was treated with SAP.

Once blunt-ended, both the 1.8 kb *EcoR* I-*EcoR* I *neo* gene DNA fragment and the linearized 9.3 kb *Mlu* I-*Mlu* I pSP7.2 II DNA fragment were compatible and were ligated under 'conventional' ligation conditions (Section 2.21.1). An aliquot of the ligation reaction was transformed into *E. coli* nova blue competent cells. 20 bacterial clones were picked and cultured overnight. DNA was prepared from each culture and analysed by restriction mapping. No single restriction endonuclease could be used to analyse the clones because successful ligation of the two blunt-ended DNA fragments did not cause a recognition sequence to be re-consituted. Therefore the single *Nco* I site known to be included in the sequence of the subcloned neomycin gene was utilised. This site was located in the new subclone (pPPT.Neo, below) relative to the restriction sites which were known to lie within the insert of the recipient plasmid DNA, pSP7.2 II. Using this approach, several positive recombinant clones were identified and shown to contain the *neo* resistance gene.

This gene was cloned in both orientations. A clone containing the neomycin gene insert in the same orientation as the PPT-A gene into which it was cloned was identified and named pPPT.Neo I (version 1). A clone with the insert in the opposite orientation to the PPT-A gene into which it was cloned was named pPPT.Neo II. It was pPPT.Neo I, however, which was further modified by the introduction of the TK gene.

5.2.2 Insertion of the Thymidine Kinase Gene: pPPT.Neo.TK

The same thymidine kinase (TK) gene which was used in PPT-A gene targeting vector No. 2 and was released from the plasmid pKS.TK by restriction with the endonucleases *Not* I and *Xho* I was cloned into the corresponding polylinker sites of plasmid pPPT.Neo I (Section 5.2.1) in the same way as described in section 5.1.5. The resultant plasmid clone constituted the completed second mouse PPT-A gene targeting vector and was named pPPT.Neo.TK. When linearized with *Not* I, this plasmid was ready for use in tissue cultured ES cell transfection experiments. The completed PPT-A gene targeting vector number 2 (plasmid pPPT.Neo.TK) is shown in figure 5.7 below.

Figure 5.7: Maps of both PPT-A Targeting Vectors (pSP7.2 V and pPPT.Neo.TK)

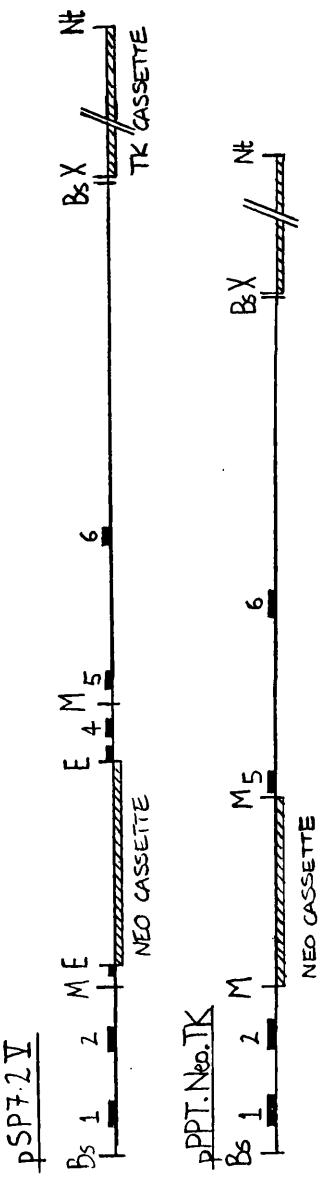


Figure 5.7: Linear map of both PPT-A gene targeting vectors (plasmids pSP7.2V and pPPT.Neo.TK respectively). For details, see sections 5.1.4 and 5.2.2, respectively. Bs; *Bst*E II, M; *Mlu* I, E; *Eco*R I, X; *Xba* I, Nt; *Not* I.

GENE TARGETING CONSTRUCT FOR THE MOUSE NK-1 RECEPTOR GENE

5.3 CONSTRUCTION OF THE MOUSE NK-1R GENE TARGETING VECTOR

The mouse NK-1R gene was known to comprise 5 exons (Sundelin *et al*, 1992 and Section 4.9). Exon (1), exon (2) and exons (3, 4 and 5) were subcloned from the bacteriophage clones λSPR8, λSPR15 and λSPR12 into the three plasmids pSPR8.5, pSPR6.8 and pSPR5.1 respectively. With respect to gene targeting vector design, the higher the degree of conservation of the genomic DNA in the targeting vector, the more effective the vector should theoretically be at promoting specific, targeted homologous DNA recombination events (see Section 5.1). Since exonic sequence is typically the most highly conserved, the higher the exonic content of the genomic DNA included in the vector, the more effective the vector should be, as indicated by the targeting efficiency. Since NK-1R exons 1, 2 and 3 were shown to be interspersed by introns of approximately 15 and 23 kb respectively (Sundelin *et al*, 1992), the smallest of the plasmid subclones, pSPR5.1, which contained a small 'cluster' of exons, namely exons 3, 4 and 5 (Section 4.9.3), was selected as the basis for the design of the first SPR gene targeting vector. In other words, subclone pSPR5.1 contained DNA of suitable 'quality' for use in a targeting vector. However, as discussed in section 5.1, the quantity of genomic DNA used in such a vector is also an important criterion to be considered; the length of the two genomic DNA fragments (which flank the neomycin resistance gene in the completed vector) when summed, should ideally be at least 6 kb in length. The 5.1 kb genomic DNA insert of subclone pSPR5.1 was considered to be too short for this purpose. A larger plasmid subclone with an insert which overlapped either exon 3, 4 or 5 of the mouse NK-1R gene was therefore sought.

5.3.1 A Larger, Overlapping NK-1R Subclone: pSPR8

On examination of the restriction map of plasmid pSPR5.1 (Section 4.9.3), two restriction sites of interest were noted; the *HinD* III site located between NK-1R exons 3 and 4, and the *EcoR* I site positioned between exons 4 and 5. The plasmid pSPR5.1 had been generated previously by subcloning a 5.1 kb *BamH* I-*BamH* I fragment of the bacteriophage λSPR12 (Section 4.9.3). Using bacteriophage λSPR12 DNA as template, a number of double restriction reactions were performed where, in each case, one of the endonucleases used was either *HinD* III or *EcoR* I. The reactions were electrophoresed capillary blotted membrane (Section 2.28). Because the *HinD* III site was located upstream of NK-1R exons 4 and 5, the blotted *HinD* III double restrictions were hybridised with the exon 5-specific oligodeoxynucleotide CS13 (Sections 4.8.2 and 2.29.2). This approach was used to locate restriction sites downstream of exon 5, the last NK-1R exon. The newly identified sites, when added to the known restriction sites in plasmid pSPR5.1, generated a more detailed restriction map of bacteriophage clone λSPR12. (N.B., the new sites were located relative to the *HinD* III site which was known to lie within subclone pSPR5.1). In the same way, the *EcoR* I site located in intron 4, between NK-1R exons 4 and 5, was used in concert with the exon 3-specific hybridisation probe, oligodeoxynucleotide CS11 (Section 4.8.2), to identify restriction sites upstream of exon 3; in intron 2. Like the probed *HinD* III double restrictions, these hybridisations proved useful in the identification of λSPR12 DNA fragments which both overlapped with subclone pSPR5.1 and were of a large enough size (greater than 6 kb) to be of use in the construction of a NK-1R gene targeting vector.

Figure 5.8a: Hybridisations of NK-1R Clone SPR12 DNA with Exon 3 Probe CS11

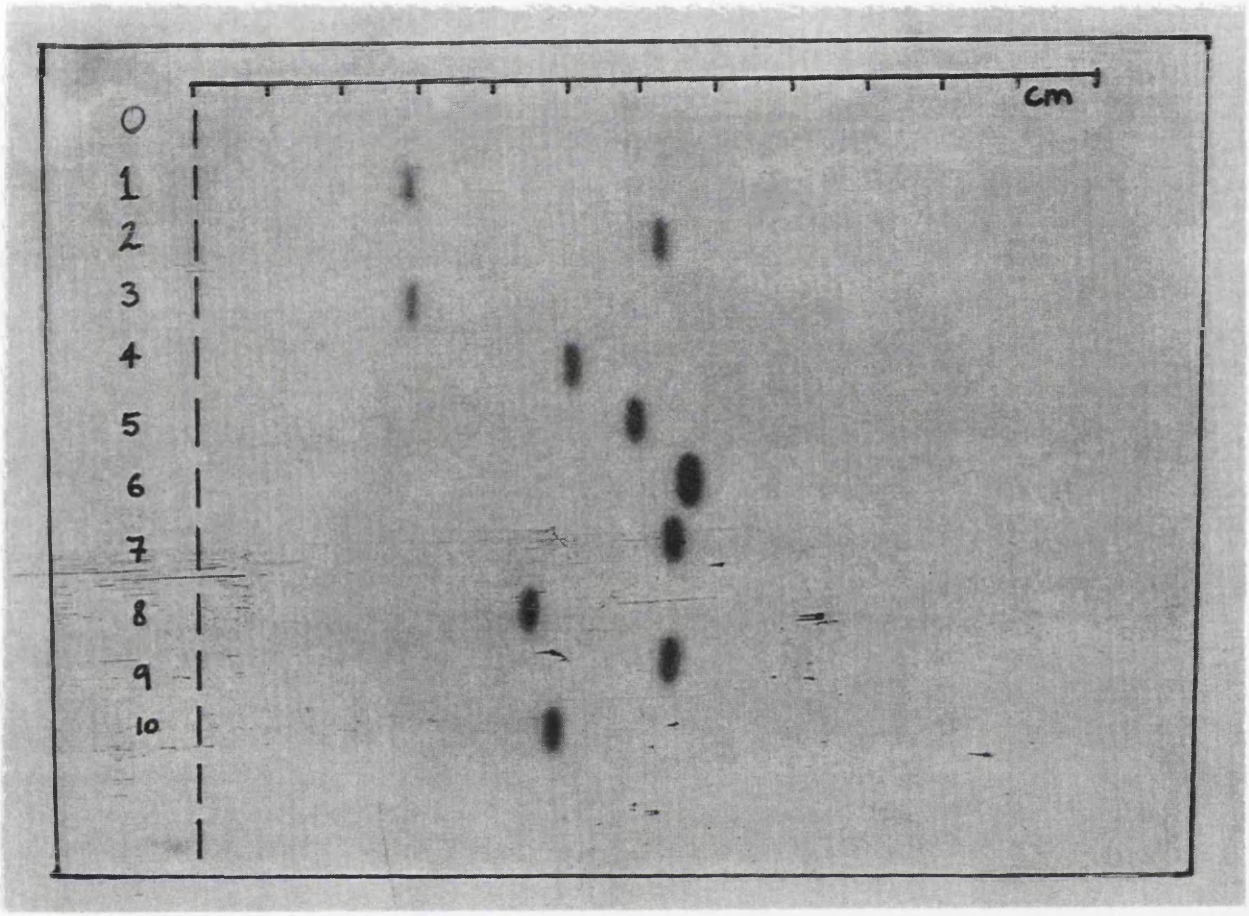


Figure 5.8a: A series of double digestions of NK-1R bacteriophage clone λSPR12 DNA resolved by electrophoresis, capillary blotted and hybridised with an NK-1R exon 3-specific oligodeoxynucleotide (CS11). 1 μg of phage DNA was digested in each lane. Lane 1; with *Eco*R I, lane 2; *Eco*R I/*Nco* I, lane 3; *Eco*R I/*Mlu* I, lane 4; *Eco*R I/*Bgl* II, lane 5; *Eco*R I/*Pst* I, lane 6; *Eco*R I/*Kpn* I, lane 7; *Eco*R I/*Xho* I, lane 8; *Eco*R I/*Xba* I, lane 9; *Eco*R I/ *Ava* I, lane10; *Eco* R I/ *BstE* II. 1Kb DNA ladder was used in lane 0 (not visualised here).

Figure 5.8b: Further NK-1R Clone SPR12 Hybridisations with Exon 3 Probe CS11

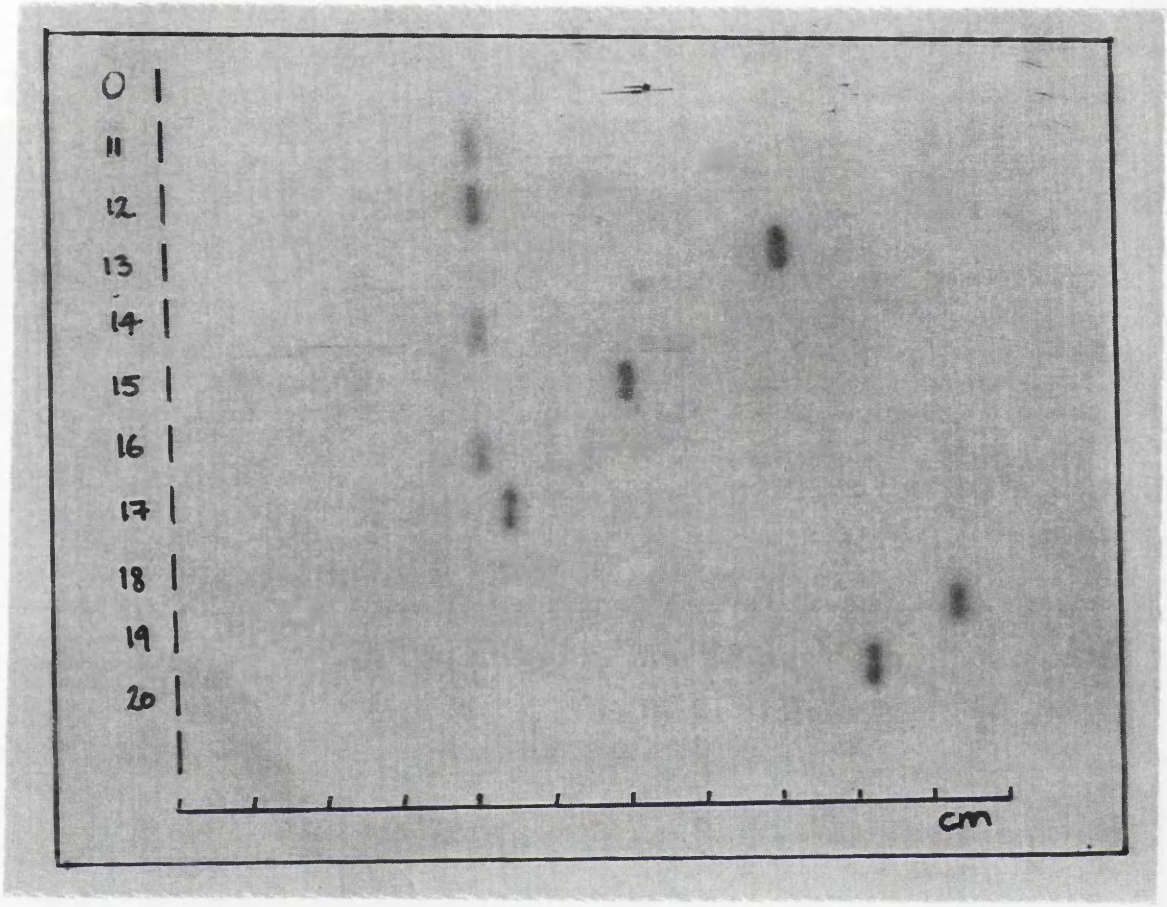


Figure 5.8b: A series of double digestions of NK-1R bacteriophage clone λSPR12 DNA resolved by electrophoresis, capillary blotted and hybridised with an NK-1R exon 3-specific oligodeoxynucleotide (CS11). 1 μg of phage DNA was digested in each lane. Lane 11; *Eco*R I/*Cla* I, lane 12; *Eco*R I/*Eco*R V, lane 13; *Eco*R I/*HinD* III, lane 14; *Eco*R I /*Not* I, lane 15; *Eco*R I/*Sst* I, lane 16; *Eco*R I/*Sal* I, lane 17; *Eco*R I/*Sma* I, lane 18; *Eco*R I/*Ssp* I, lane 19; *Eco*R I/*Stu* I, lane 20; *Eco*R I/*HinD* III. NB; the 8 Kb *Eco*R I/*Sst* I band in lane 15 which was subcloned to generate plasmid pSPR8 which was the basis for the NK-1R gene targeting vector described in this text. Scale is in centimetres. 1 Kb DNA ladder was used in lane 0 (not visualised here).

10 μg of λSPR12 DNA was double restricted with endonucleases *EcoR* I and *Sst* I in REact buffer 2 (B.R.L. Ltd.). The digestions were resolved on a gel and the 8 kb DNA fragment was band-isolated. The DNA was recovered from the agarose (see Qiagen Ltd. standard protocol) and resuspended. 5 μg of the plasmid vector pBluescriptKS+ was cut with the same endonucleases and recovered from an agarose gel slice in the same way. The linearized 3 kb plasmid vector and the 8 kb λSPR12 *EcoR* I-*Sst* I DNA fragments were ligated (see Section 2.21.1) and transformed into competent *E coli* cells (TG1 strain). The (white colony) transformants were analysed by endonuclease restriction. A clone containing the correct 8 kb insert was identified and named subclone pSPR8.

A restriction map for pSPR8 was compiled by single and double endonuclease digestion which allowed orientation of the insert restriction sites relative to the known pBluescriptKS+ polylinker sites. The agarose gels generated as a by-product of the mapping procedure, bearing digested plasmid pSPR8 DNA fragments were capillary blotted (Section 2.14), stored and at a later date positively hybridised with the oligodeoxynucleotide probe CS11. The restriction map and the relative locations of NK-1R exons 3 and 4 are shown below.

Figure 5.9 : Map of Targeting Vector pSPR.Neo.TK & Subclones pSPR8 and pSPR5.1

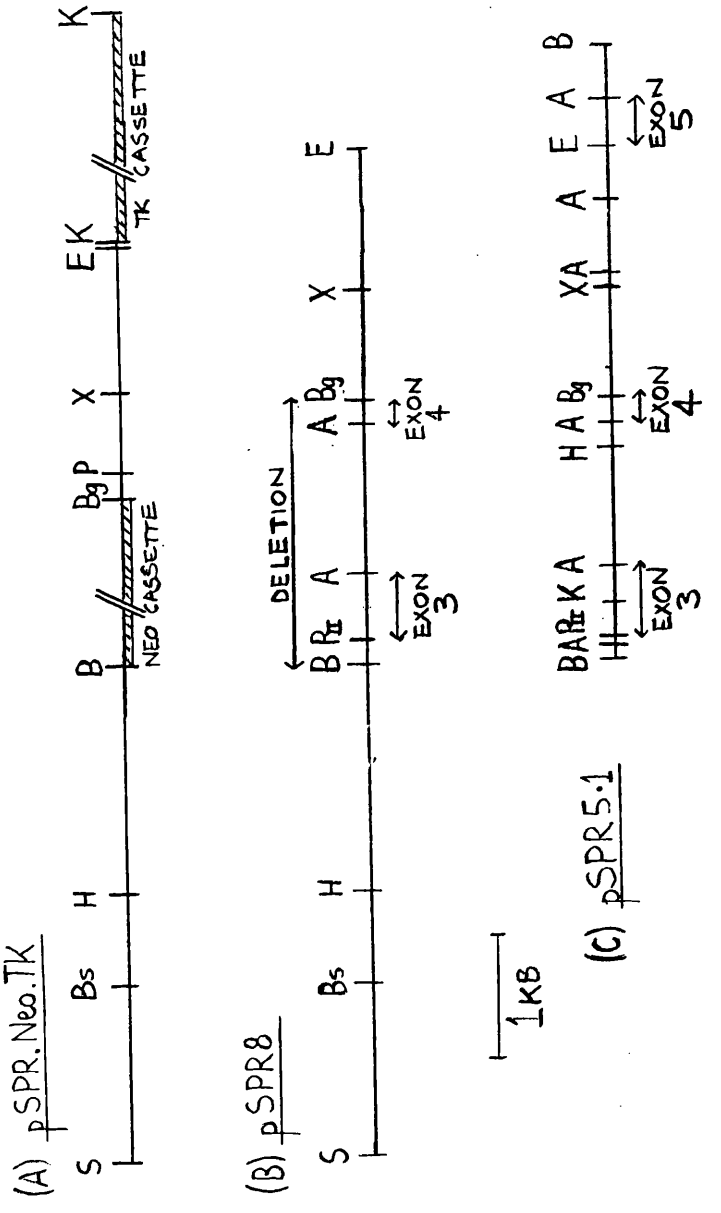


Figure 5.9: Restriction map of (a) the completed NK-1R gene targeting vector pSPR.Neo.TK (section 5.3), the NK-1R plasmid subclones pSPR8 (section 4.9.1) and pSPR5.1 (section 4.9.3). S; *Sst* I, Bs; *Bst*E II, H; *Hin*D III, Bg; *Bgl* II, P; *Pst* I, X; *Xho* I, E; *Eco*R I, K; *Kpn* I, PII; *Pvu* II, A; *Ava* II, B; *Bam*H I.

5.3.2 Insertion of the Neomycin Resistance Gene: pSPR.Neo

Based on the subclone pSPR8, the NK-1R gene targeting vector was designed to delete NK-1R exons 3 and 4 from the genomic DNA insert of plasmid pSPR8. The vector was constructed by cloning the *neo* gene and the thymidine kinase genes directly into plasmid pSPR8. To clone the *neo* gene, the *BamH* I restriction site located upstream of exon 3 and the *Bgl* II restriction site located downstream of exon 4 were utilised. Firstly, plasmid pSPR8 was cleaved with *Bgl* II. The protruding *Bgl* II cohesive ends were blunt-ended by incubation of the DNA fragments with dNTPs in the presence of Klenow enzyme. Residual Klenow enzyme was removed by extraction with phenol/chloroform. The linearized, blunt-ended pSPR8 DNA fragment was secondly restricted with *BamH* I and was electrophoresed on a 0.6% agarose gel. Of the two DNA fragments generated by this second restriction, the larger, 8.9 kb fragment was recovered from the agarose by phenol/chloroform extraction. The fragment which was not recovered, namely the 1.5 kb *Bgl* II-*BamH* I of pSPR8, constituted the section of genomic DNA which was deleted from the NK-1R gene within the gene targeting vector. This deleted genomic DNA fragment encompassd NK-1R exon 3 sequence.

The *neo* gene was cloned into pSPR8 in a way similar to the deletion of exon 3 as described above. Plasmid pSKNeo (Section 5.1.3.5) was restricted with *HinD* III and blunt-ended and purified in the same as described in the above paragraph. The DNA was then digested a second time with *BamH* I and recovered from an agarose gel as before. The double digestion released the *neo* gene as a 1.8 kb fragment with one blunt end and one *BamH* I cohesive terminal. The fragment could therefore be subcloned directly into the prepared pSPR8/*Bgl* II-*BamH* I fragment. The fragments were ligated under conventional conditions and transformed into *E. coli* nova blue cells. Bacterial clones were selected, cultured and analysed by restriction analysis. A 10.7 kb recombinant plasmid clone containing the *neo* gene at the correct location and which was linearised by *BamH* I was identified and named pSPRNeo.

5.3.3 Insertion of the Thymidine Kinase Gene: Subclone pSPRNeo.TK

The *Kpn* I site in the polylinker of plasmid pSPRNeo, which was one of only two remaining unique sites, was selected as the thymidine kinase (TK) gene insertion site. (The other unique site was used to linearise the final construct). The TK gene had previously been cloned into the vector pBluescriptKS+ to generate pKS.TK (Section 5.1.5). In order to clone the thymidine kinase gene from pKS.TK into the single *Kpn* I site of plasmid pSPRNeo, pKS.TK had to be modified so that a *Kpn* I site was located either side of the TK gene. One such flanking *Kpn* I site already existed in the polylinker of pKS.TK. The other *Kpn* I site was introduced at the opposite end of the TK gene in plasmid pKS.TK using a short double stranded DNA linker. The linker was generated by synthesizing two complementary oligodeoxynucleotides designed so that when they were annealed to each other, a *Kpn* I restriction site was generated and which was flanked by cohesive *EcoR* I termini as shown below.

Figure 5.10 : Components, Design and Sequence of the *EcoR* I- *EcoR* I Linker

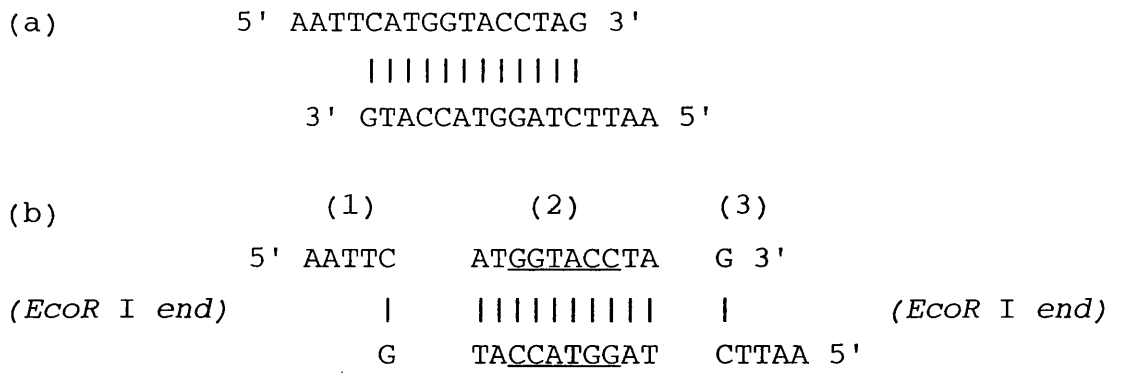


Figure 5.10: Sequence of the two oligodeoxynucleotides which were annealed, phosphorylated and cloned into the *EcoR* I site of plasmid pKS.TK to generate plasmid pKS.TK.Kpn which was in turn digested with the endonuclease *Kpn* I to release all 2.9 kb of the thymidine kinase counter selection gene. Part (a) shows the oligodeoxynucleotides annealed to each other, whereas part (b) shows the linker diagramatically sub-divided to demonstrate the 3 component sections; (1) a *EcoR* I cohesive terminal, (2) a *Kpn* I recognition sequence (which is underlined), and (3) another *EcoR* I cohesive terminal.

The two oligodeoxynucleotides were annealed, phosphorylated and prepared for ligation as previously described (Section 5.1.2). Plasmid pKS.TK was linearized with *EcoR* I and dephosphorylated with SAP in preparation for ligation. The ligation of the short DNA linker into the linearised 5.9 kb pKS.TK plasmid was conducted under conventional ligation conditions and with a 20 molar excess of linker relative to pKS.TK. A 2 µl sample of ligation mix was transformed into competent *nova blue* strain *E. coli* cells. Twenty transformants were picked, cultured and analysed by digestion with *Kpn* I. Unsuccessful clones were identified as single restriction fragments of 5.9 kb indicating the existence of only one *Kpn* I site. Successful clones were also identified by digestion with *Kpn* I where two DNA fragments were generated of 2.9 kb and 3 kb, corresponding to the TK gene and the pBluescript vector respectively. One such recombinant plasmid was identified and named pKS.TK.Kpn.

Plasmid pKS.TK.Kpn was restricted with *Kpn* I to release the TK gene as a single 2.9 Kb DNA fragment. The fragment was subcloned directly into the *Kpn* I site of plasmid pSPR.Neo (Section 5.3.2). The 2.9 Kb TK gene fragment was purified from an agarose gel by phenol/chloroform extraction and ethanol precipitation. The recipient DNA fragment for this ligation, namely the *Kpn* I-linearized pSPRNeo plasmid, was prepared for ligation by the same method but in addition was dephosphorylated with shrimp alkaline phosphatase (Section 2.22.2) The ligation was conducted under standard conditions and a 2 µl sample was transformed into *nova blue* cells as described above. Successful transformants were selected and cultured under antibiotic (ampicillin) selection in L-broth medium. Recombinant clones were identified and shown by restriction enzyme analysis to contain the subcloned thymidine kinase counter selection gene. A recombinant pSPRNeo plasmid clone which contained the TK gene was identified. The TK gene insert was shown by restriction enzyme digestion to be in the same 5'-3' orientation as the NK-1R gene into which it was cloned. The recombinant plasmid was named pSPRNeo.TK. This plasmid constituted the first completed mouse NK-1R gene targeting vector. Only the *Cla* I restriction site in the polylinker of the original plasmid vector (pBluescriptKS+) remained unique in

targeting vector pSPRNeo.TK. Prior to transfection into tissue cultured mouse ES cells, the targeting vector DNA was therefore linearized by digestion with the endonuclease *Cla* I. This resulted in the generation of a single 14.9 kb DNA fragment. A linear map of the completed NK-1R targeting vector pSPR.Neo.TK is shown in Figure 5.9).

5.4 FLOW DIAGRAMS DETAILING THE CONSTRUCTION OF THE PPT-A AND NK-1R GENE TARGETING VECTORS

This section comprises three flow diagrams which schematically explain the construction of the gene targeting vectors which are described in the above text. Figure 5.11 shows the the series of complex cloning steps required to assemble PPT-A vector number 1 (Section 5.1). The final vector was named plasmid pSP7.2 V. Figure 5.12 demonstrates the construction of PPT-A vector number 2 (Section 5.2), plasmid pPPT.Neo.TK. Figure 5.13 details how the targeting vector for the NK-1R gene was assembled (Section 5.3). The final plasmid vector was named pSPR.Neo.TK.

Figure 5.11 : Assembly of the PPT-A Gene Targeting Vector Number 1:pSP7.2 V

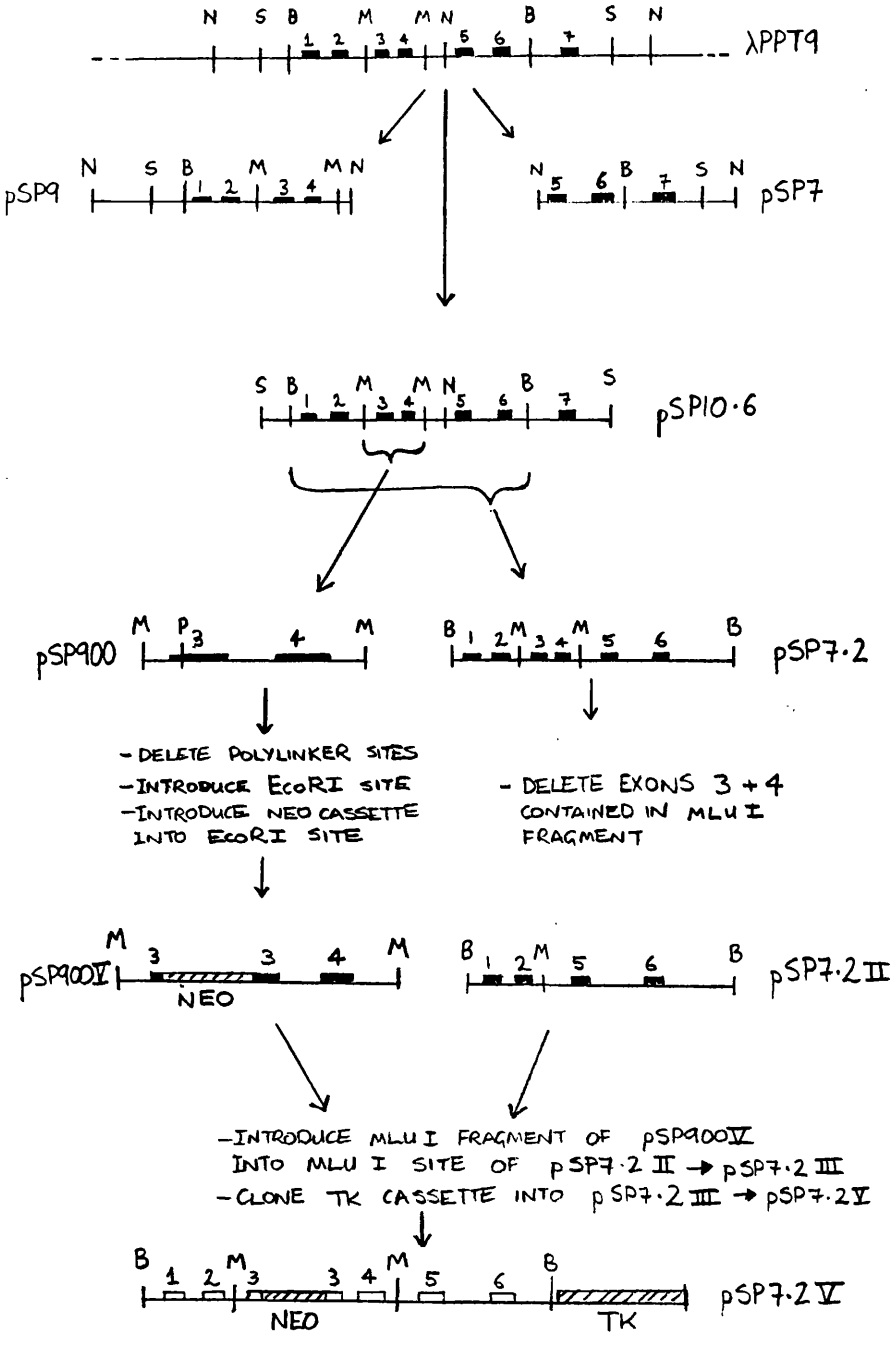


Figure 5.11: Schematic representation of the construction of the PPT-A gene targeting vector number 1, plasmid pSP7.2 V. Numbers refer to exons. N; *Nco* I, S; *Sst* I, B; *Bst*E II, M; *Mlu* I, P; *Pst* I.

Figure 5.12 : Assembly of the PPT-A Gene Targeting Vector Number 2:pPPT.Neo.TK

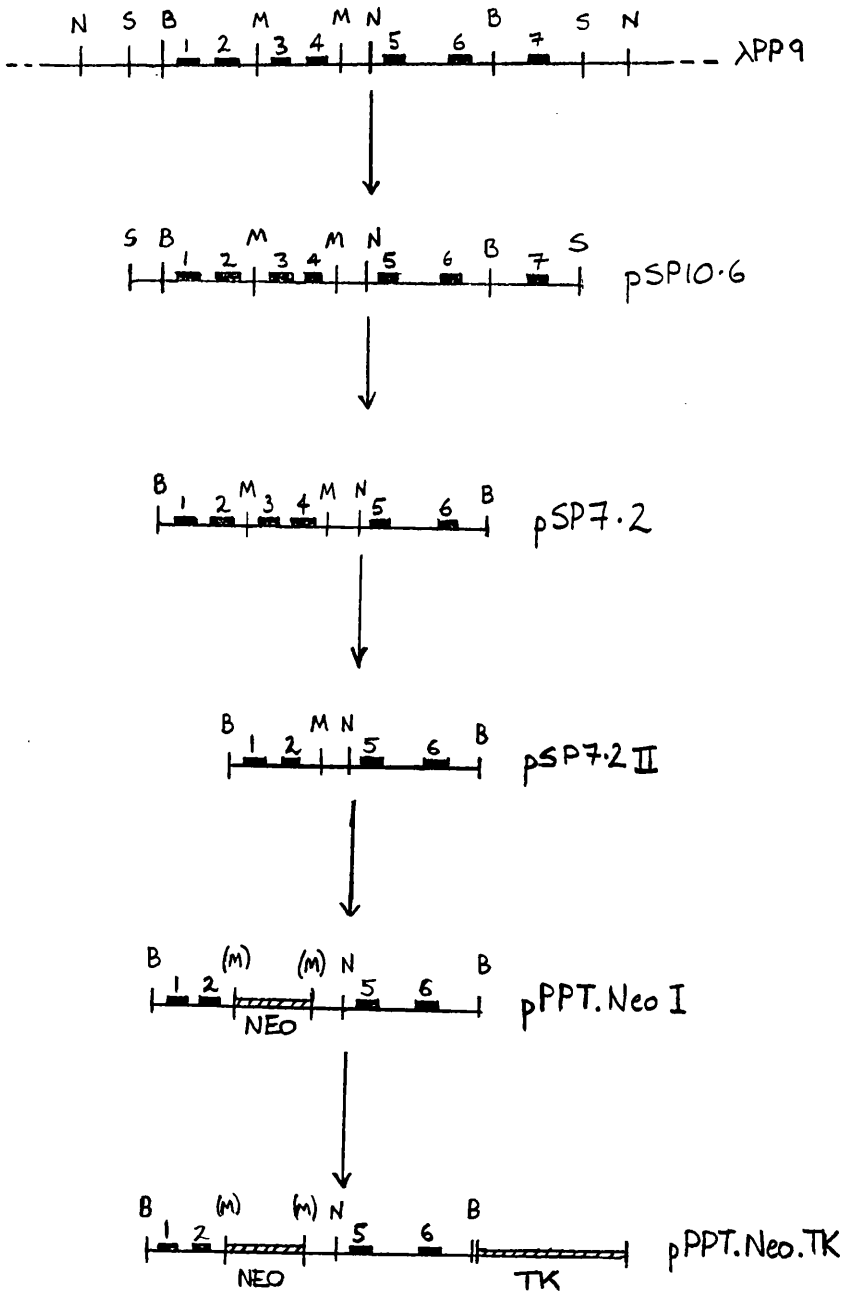


Figure 5.12: Schematic representation of the construction of the PPT-A gene targeting vector number 2, plasmid pPPT.Neo.TK. Numbers refer to exons. N; *Nco* I, S; *Sst* I, B; *Bst*E II, M; *Mlu* I, P; *Pst* I and "(M)" refers to former *Mlu* I recognition site destroyed by blunt-end cloning.

Figure 5.13 : Assembly of the NK-1R Gene Targeting Vector:pSPR.Neo.TK

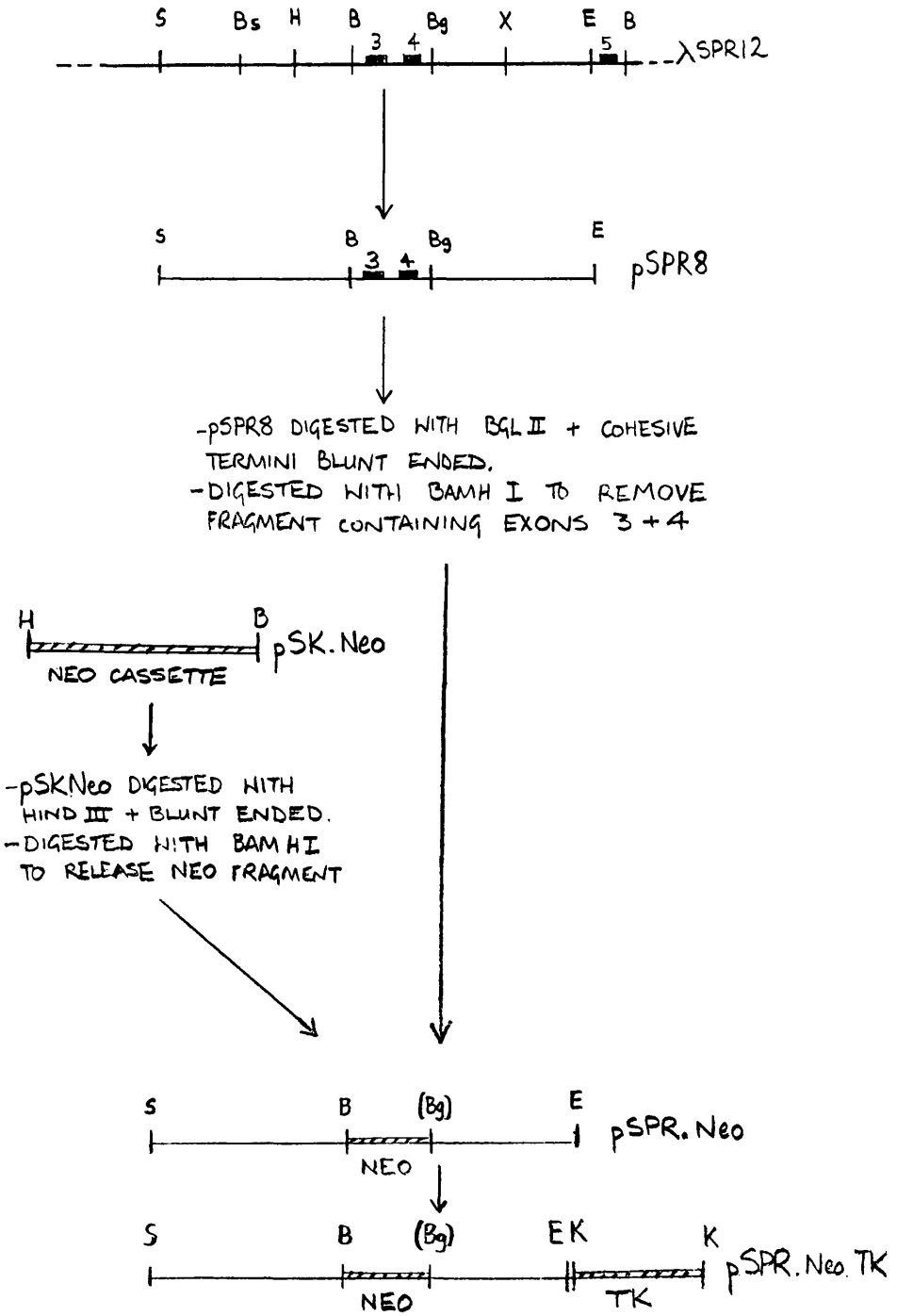


Figure 5.13: Schematic representation of the construction of the NK-1R gene targeting vector, plasmid pSPR.Neo.TK. Numbers refer to exons. S; *Sst* I, Bs; *Bst*E II,H; *Hin*D III, B; *Bam*H I, Bg; *Bgl* II, X; *Xho* I, E; *Eco*R I, K; *Kpn* I and "(Bg)" refers to former *Bgl* II recognition site destroyed by blunt-end cloning.

CHAPTER 6

DISCUSSION

SUMMARY AND DISCUSSION

In order to investigate the proposed involvement of the preprotachykinin-A precursor and NK-1R receptor genes in neuronal dysfunction, it is important to firstly understand, at the molecular level, the normal function and regulation of those genes. When this research project was designed, genomic or cDNA clones of neither the murine PPT-A and NK-1R genes had been isolated. In the interim, clones of both genes have been identified (Kako *et al*, 1992 and Sundelin *et al*, 1992, respectively). However, a full investigation into the organisation and the regulation of those genes remains outstanding. Similarly, an understanding of the alterations in phenotype, if any, which occur when those genes are rendered non-functional (by way of molecular gene targeting for example) also remains unclear. This study, by cloning those genes in the first instance, has made an initial investigation into the function of the PPT-A and NK-1R genes. This provides the bedrock for future, more comprehensive gene characterisation and gene targeting ('knock-out') analyses.

This chapter serves to firstly recapitulate, by way of a summary, the results obtained during the research project and to then relate those findings to the previously published research findings which are detailed in the Introduction. The technical shortcomings of the research presented here are discussed. Furthermore, future experiments which may help to further elucidate the function of the

tachykinin neuropeptide neurokinin-1 will also be discussed. It is worth bearing in mind that in the results chapters presented earlier in this text, the rationale behind the experiments undertaken is detailed, and there is some discussion and explanation as to why a particular experimental approach was adopted.

6.1 SUMMARY OF RESEARCH FINDINGS

Firstly, to clone the murine PPT-A gene from a genomic bacteriophage library, a 423 bp probe was generated by the polymerase chain reaction (PCR) using mouse genomic DNA as template. This fragment was subcloned and sequenced. Optimal hybridisation conditions were then determined by Southern analysis. The probe was used to isolate 10 lambda bacteriophage clones from a genomic library, two of which (λPPT7 and λPPT9) contained all seven PPT-A exons. Similarly, 186 positively hybridising clones of the NK-1R (receptor) gene were isolated from another bacteriophage library using an 865 bp rat NK-1R cDNA-derived probe. For the NK-1R screen, the same hybridisation conditions as for the PPT-A probe were successfully pre-tested by dot blot analysis and used.

Secondly, the PPT-A bacteriophage clones were analysed by endonuclease restriction mapping, PCR analysis and by hybridisation to exon-specific oligodeoxynucleotides. The seven exons of the murine PPT-A gene were then subcloned into plasmids named pSP9, pSP7 and pSP10.6. The subclones were restriction mapped and the exons were positioned relative to each other on those restriction maps. By using the oligodeoxynucleotide probes as sequencing primers, mouse PPT-A exon sequence was obtained and compared with other species to determine the extent of homology. High dgrees of homology (>95%) between the mouse and rat and bovine sequence at the 5' flanking, the start, the middle and the 3' end of the PPT-A gene were noted. A similar approach was used for the NK-1R bacteriophage clones. They were analysed by hybridisation, were subcloned, restriction mapped, and the distribution of the exons

within those subclones was determined (and found to be very similar to that of the rat PPT-A gene). Sequence analysis of the NK-1R gene is lacking. Other evidence presented here indicates that the murine NK-1R clones are authentic.

Thirdly, gene targeting vectors for use in homologous recombination experiments in mouse embryonic stem cells were then constructed for both genes (two for the PPT-A gene and one for the NK-1R gene). Multiple cloning and subcloning steps were required to complete the vectors. Several restriction sites were specifically deleted or introduced (by ligation of short DNA linker sequences). Principally, a *neo* resistance gene carrying its own promoter and poly(A) termination sequence was introduced into the vector (for example exon 3 of PPT-A vector number 1, plasmid pSP7.2 V). In PPT-A vector number 2 (plasmid pPPT.Neo.TK), the *neo* resistance gene was introduced in place of a deleted 2.4 kb genomic sequence comprising NK-1R exons 3 and 4. For the NK-1R vector (plasmid pSPR.Neo.TK), exons 3 and 4 of that gene were deleted and replaced by the *neo* resistance gene. In all three vectors, a thymidine kinase (TK) counter-selection gene from *herpes simplex* virus 1 (HSV1) was also inserted so as to flank the genomic DNA contained in the targeting vector.

Finally, published mouse PPT-A and NK-1R DNA sequence (Voorn *et al*, 1994, and Mick *et al*, 1994, respectively) was used in the design of 45-mer oligodeoxynucleotides for use in *in situ* hybridisations. 20 μm sagittal sections were prepared from two 10 day neonatal mouse brains. The brains were hybridised independently with one of the gene-specific oligodeoxynucleotide probes. The purpose of such experiments was to ascertain the distribution of mRNA transcripts encoded by the mouse PPT-A and NK-1R genes in the brains of normal, wild-type, mice. Furthermore, the results which were obtained in this way are of value with respect to future comparisons of *in situ* hybridisation results and demonstrate that there are specific but distinctly different expression patterns for both genes in the brain of the 10 day old (wild-type) mouse.

6.2 GENE CLONING DISCUSSION

Chapter 3 details the strategy adopted, and the ground work required, in order to isolate bacteriophage clones for the murine PPT-A and NK-1R genes.

For the PPT-A gene, the PCR-based approach was used to generate a library probe because no PPT-A clone DNA (genomic or cDNA) from any other species (and suitable for use as a probe) was available at that time. Such DNA, if available, would have been used as a heterologous probe for library hybridisations. Due to the design of the PCR primers, where the PCR product spanned (the neurokinin-1 encoding) third exon, in theory, any bacteriophage clones isolated using this probe would contain not only PPT-A sequence but specifically exon 3 of that gene. (Having analysed the bacteriophage isolated from both the Cambridge and Glaxo libraries, this has proved to be the case; see Section 3.3.3 and 3.3.4). This particular area of the gene, which was identified at an early stage, was considered to be of special interest because of its potential use in the design of future PPT-A gene targeting vectors. In that respect, it was considered an ideal insertion site for a *neo* resistance marker gene. Insertion of a *neo* cassette into the exon which encodes NK-1 offers the greatest probability of disrupting the PPT-A gene transcription in any subsequently targeted gene. Such an insertion should thereby alter the translation products and halt the synthesis of the normal form of the NK-1 peptide. Disruption of NK-1 synthesis is, after all, one of the long-term objectives of this targeting project. The inclusion (by design) of this site in the original probe sequence proved useful and this site was used at a later stage in the construction of one of the PPT-A gene targeting vectors. (This particular exon and its use in further targeting vectors is discussed in greater depth in section 7.4). If a suitable probe DNA had been readily available at the outset of this project, the generation of a PPT-A probe by PCR would not have been necessary and the time taken to isolate PPT-A clones would probably have been reduced. The PCR analysed did serve a positive purpose in that rat and mouse PCR products of same length were generated from

the respective genomic DNAs and this supports the authentic cloning of murine PPT-A gene as presented in this text. The PPT-A probe was sequenced prior to use to ensure its authenticity.

In the case of the NK-1R receptor gene, the rat NK-1R cDNA was available at the start of the project and a 965 bp fragment of this was used as the NK-1R library hybridisation probe. Subsequently, no PCR or subcloning step was required. The NK-1R probe sequence, like that of the PPT-A probe, was checked by sequence analysis before use. The use of a cDNA-derived probe to isolate clones of the murine NK-1R gene (which was believed to span more than 30 kb as it does in rat) was thought likely to result in the identification of clones which between them would contain all of the expected five NK-1R exons. This was the case simply because the rat cDNA probe sequence contained regions of high DNA sequence homology to each of the five mouse exons. Of the 186 positive clones identified at the primary stage, twenty were selected randomly for further analysis and were found to contain exons 1, 3, 4 and 5. Those exons were subcloned and the location of those exons relative to each other was detemined using exon specific oligodeoxynucleotides. Unfortunately, none of those 20 selected bacteriophage clones contained any exon 2 sequence (as concluded from a blank hybridisation to an oligodeoxynucleotide containing published mouse exon 2 sequence and derived from Sundelin *et al*, 1993, data not shown here). It was therefore necessary to isolate this exon. To do this, the bank of 186 primary clones was re-screened with the exon 2 specific oligodeoxynucleotide probe (rather than screening a whole library again). This appears to have resulted in the successful isolation of exon 2 sequence. This exon, which is flanked by 15 and 23 kb introns in rat, was shown to be included in an approximately 11 kb *Bam*H I/*Bam*H I fragment of the isolated clone λSPR14. In order to confirm that the isolated sequence from λSPR14 is indeed NK-1R exon 2, sequence analysis is necessary. Regrettably, no DNA from λSPR14 was either subcloned or sequenced. Therefore, the authenticity of the exon 2 DNA has not been unequivocally confirmed. Having used a 100% homologous probe (derived from published mouse sequence), which was found to hybridise specifically under stringent

conditions, it is, however, very likely that this bacteriophage clone contains authentic mouse NK-1R exon 2 sequence.

One possible future cloning experiment for both the PPT-A and NK-1R genes would be the isolation of cDNA clones from another library. Such clones would be useful in the future gene characterisation work as detailed in Section 7.3.

6.3 GENE CHARACTERISATION DISCUSSION

6.3.1 Future PPT-A Gene Characterisation Experiments

6.3.1.1 Sequence Analysis of the Murine PPT-A Gene

The oligodeoxynucleotides previously used as PCR primers and hybridisation probes were also used as sequencing primers in addition to standard sequencing primers such as T3 and T7 (Stratagene). This preliminary sequence analysis demonstrated very high sequence homologies between the mouse and the rat PPT-A sequences at both the 5' and 3' ends of the gene and also in central regions in the vicinity of exons 3 and 4 (Section 4.6). The sequence obtained for the murine PPT-A gene gas been presented as alignments with known, published database sequences. The exact location of the exons within those alignments remains to be detailed (by further computational analysis). This finding was deemed sufficient to confirm the cloned gene as the murine PPT-A precursor gene. However, the seven PPT-A exons were not sequenced fully or on both strands although oligodeoxynucleotides were shown to hybridise to each of the seven murine PPT-A exons in both subclones and bacteriophage clones (see Section 4.5). The relative positioning of those exons on a linear map (see Figure 4.6 after Section 4.4) showed the mouse exons to be dispersed in a similar way as in the rat PPT-A gene (Carter *et al*, 1990). Also, the introns appear to be of similar size in both the rat and mouse genes. This acted as further evidence supporting the authenticity of the cloned murine

PPT-A gene. The precise length of the PPT-A exons has not been determined by this study and remains outstanding. The seventh exon of the rat PPT-A gene is 2 kb in length and the mouse equvalent is likely to be similar. Ninety percent of that rat exon is untranslated. It would be of value to fully sequence the murine exon 7 and to determine the extent of the equivalent untranslated region which is likely to exist there. This could be used in the generation of a full genomic map of the murine PPT-A gene.

Similarly, the intron/exon boundaries of the murine PPT-A gene have not been accurately defined. At the same time as the identification of the PPT-A clones detailed here, the cDNA sequence was published by Kako *et al*, (1993). The published information could be of use in the identification of the murine PPT-A intron/exon boundaries. This information could be used to study the regulation of expression of the gene and to identify the mRNA splicing variants which are encoded, for example by using the enzyme S1 nuclease which specifically degrades unpaired single-stranded DNA sequences. S1 nuclease protection experiments could be undertaken to study the transcription patterns of the murine genes. For the PPT-A precursor gene, comparison with the known rat transcription pattern would be possible (for rat, see Figure 1.2). Furthermore, Northern blotting analysis could be used to study transcription. It is highly likely that several transcripts will be encoded by the mouse gene as for the rat PPT-A gene (Helke *et al*, 1990, see Section 1.3.4 and Figure 1.2). Those are likely to be translated into several neuropeptides. The sequence generated in that way would, at least, allow clarification of the physical organisation of the gene and may contribute to an understanding of its molecular activity.

6.3.1.2 PPT-A Promoter and Expression Studies

The rat PPT-A gene was isolated from a genomic library by Carter *et al* (1990) and from this, one major and two minor transcription initiation sites were identified upstream of the rat gene. Comparison of the rat and the bovine PPT-A sequences has demonstrated that there are highly conserved regions throughout the entire coding region and within the 5' flanking region. The mouse to

rat homologies for this region are also very high (see Section 4.6). This knowledge could help facilitate the further characterisation of the murine gene because clones λPPT7, λPPT8 and λPPT9 are known to contain genomic sequence 5' of the first exon. This could permit the identifcation of promoter (or enhancer) elements within those clones. In the human NK-1R gene, exons 1 to 5 are 977, 196, 151, 197 and approximately 3300 bp in size respectively and the exons are interspersed by introns of >22 kb, >26 kb, >7 kb and approximately 2.2 kb (Shigemoto *et al*, 1990). The human NK-1R, has a TATA box located 24-27 bp upstream of the putative transcription initiation site. (The TATA box is involved in the polymerase nucleotide selection which initiates transcription at a specific site). Furthermore, several promoter, enhancer and regulatory DNA elements have been identified in the 5'-flanking region, including a cyclic AMP responsive element (TGATGTCT) and an AP-1 sequence. In many species, the PPT-A precursor gene demonstrates a restricted expression pattern both during development and in the adult. The transcriptional regulation of the murine gene needs to be elucidated as does a more accurate understanding of its tissue-specific expression. Neurokinin-1, neurokinin A, neuropeptide K, and neuropeptide γ are each encoded from the PPT-A gene by tissue-specific, alternative post-transcriptional processing (alternative RNA splicing) and post-translational pathways (Carter *et al*, 1990). Alternative splicing of the primary PPT-A RNA transcripts generates three mRNAs which, in turn, generate several distinct precursor proteins. McDonald *et al* (1989) presented results which demonstrated that multiple tachykinins are produced and secreted as a consequence of post-translational processing of the three preprotachykinin-A precursor proteins, α-, β-, and γ- prepro-tachykinin. This form of processing in mouse needs to be clarified. The role of tachykinins such as neurokinin-1 in neurogenesis and in triggering the differentiation of mature neurons is also worth analysis. Further *in situ* hybridisation studies (see Section 7.5), which would compliment those presented in this text (see Chapter 6), may allow researchers do this by focussing on particular developmental stages during which processes such as neurogenesis are most or least active. The findings referred to in section 1.3.8.3 regarding the developmentally

transient expression of the NK-1R receptor in neuronal cells of the developing floor plate (Heath *et al*, 1995) indicate that research of this nature is already underway and contributing to the understanding the transient expression of neuronal gene products. This in turn will permit further understanding of the role of neuropeptides in the signalling systems of the mammalian nervous system. Using the PPT-A (and NK-1R) clones described here, the factors which mediate the regulation of neuropeptide expression in various tissues and at various developmental stages may be investigated.

The promoter for the rat PPT-A precursor gene has been cloned (see Mulderry *et al*, 1993, Section 1.3.5). The equivalent murine promoter has not. The murine promoter may well be located within the bacteriophage clones isolated during this research which are known to contain sequence 5' of PPT-A exon 1. The clones could therefore be sequenced in order to pinpoint regulatory motifs and regulatory factor binding sites within that promoter (using DNAase footprinting techniques and band-shift studies for example to do so). Mulderry *et al* (1993) used deletion analysis to study the rat PPT-A promoter and the regulation of expression. Sequences which are important to the regulation of this gene in rat dorsal root ganglia (DRG) neurons are located between 47 and 865 nucleotides upsteam of the start of the gene. There is also evidence of regulatory elements residing in the 5' untranslated regions of genes. For example, an iron responsive element is located in the 5' leader sequence of the ferritin gene (Goosen *et al*, 1990). Reporter assays could also be used to study a promoter. Furthermore, those promoter elements could be mutated so as to locate the critical sequences. Zajac-Kaye *et al* (1988) used the human oncogene c-myc as a model and utilised exonuclease enzymes in order to identify numerous potential binding sites for regulatory proteins which were mapped either upstream or within the human c-myc oncogene. A similar approach could be adopted in future studies of the PPT-A (or NK-1R) genes. Of the ten positively-hybridising PPT-A bacteriophage clones identified, λPPT9 contained all seven PPT-A exons and was used in the gene characterisation experiments described in Chapter 4. The other nine clones, which contain sequence both 5' and 3' of that clone, have not been fully

analysed (although the extent of overlap is shown in Figure 3.18). This flanking sequence could be examined for regulatory elements. Expression of a specifically mutated PPT-A (or NK-1R) gene promoter in tissue cultured neural-derived cells is another possible way to study promoter control and gene expresion. The rat PPT-A gene utilises alternative RNA splicing events to regulate expression. This is also likely to occur in the murine PPT-A gene. An analysis of this would also be of interest. There are other machanism of regulating gene expression that could be explored. For example, there may be several promoters for a given gene which could result in the the generation of several distinct mRNA isoforms. (As an example of this, the *Drosophila melanogaster* alcohol dehydrogenase gene, as reported by Benyajati *et al*, (1983), differs in its 5' end during different developmental stages). In addition to promoter studies, a similar approach to studying the regulation of the PPT-A gene (or the NK-1R gene) could be applied to the identification and activity of enhancer elements or repressor elements, if indeed there are any (see Quinn, 1987).

6.3.2 Future Murine NK-1R Gene Characterisation Experiments

The subcloned NK-1R exons 1, 3, 4, and 5, although subcloned, have not been fully sequenced on both strands and the exact length of the exons has not been determined. Similarly, the intron/exon boundaries have not yet been determined and this remains outstanding. Exon 2 also remains to be subcloned as well as fully sequenced. Completion of this would contribute to the further characterisation of this gene and by obtaining the sequence upstream of exon 1 where the NK-1R promoter and the transcriptional start site could possibly be identified. It would also be of use to look at the 166 NK-1R bacteriophage clones which were isolated from the Toronto library (Section 3.6), but were not studied further (186 in total, 20 already analysed). By defining the extent of overlap between them, it may be possible to determine whether the whole murine NK-1R gene, which is over 30 kb in rats, has been cloned. The first full genomic map of the NK-1R gene (which is not given in the

original murine NK-1R gene cloning publication from Sundelin *et al*, 1993), could be generated from this data.

The polymerase chain reaction technique could also be used as a further means of supporting the authenticity of the cloned murine NK-1R gene (in the absence of conclusive sequencing data). The isolated NK-1R bacteriophage (and the subsequently generated plasmid subclones) could be used as PCR templates to do this. Several oligodeoxynucleotides already exist that contain mouse NK-1R genomic exonic sequence (from Sundelin *et al*, 1993) and which appear to hybridise to the NK-1R exons of the isolated bacteriophage clones and could be used as PCR primers. Unfortunately, of the existing oligodeoxynucleotides specific to each of the five murine NK-1R exons, all are in the 5'-3' orientation. An anti-sense primer would therefore neeed to be designed and synthesised for use as the second primer in diagnostic PCRs.

There is currently very little known about the DNA sequences and trans-acting factors which control the expression of NK-1R in different tissues and during development. Identification, sequence analysis and mutational analysis of the promoter (for example using the exonuclease-based approach described in Section 7.3.2.2 for the PPT-A gene) would be equally useful in identifying potential regulatory elements of the NK-1R gene. Additionally, an investigation of the trancriptional control and the activity of this gene would be of interest. This activity is very likely to vary during development as it is known to do from one tissue to another.

Furthermore, there are a whole host of gene characterisation experiments that *could* be performed using the clones for the murine PPT-A and NK-1R genes bur which is too exhaustive to include here. The means to achieve a fully and comprehensively characterise those genes is however, presented here.

6.4 GENE TARGETING VECTOR DISCUSSION

Chapter 5 details the design and construction of gene targeting vectors for the murine PPT-A and NK-1R genes. The gene targeting vectors described are designed to have a major impact on gene activity. Taking the PPT-A targeting vectors as an example, the insertion of the neomycin resistance gene (with its own termination sequence) into the PPT-A gene should grossly disrupt the activity and terminate transcription prematurely in any subsequently generated transgenic mouse. Therefore the synthesis of all of the *several* neuropeptides encoded by that gene should be halted by this. A resultant transgenic mouse may therefore display phenotypes and behaviours which will correlate with the gross disruption of PPT-A as a whole but not with the specific deletion of a single neuropeptide gene product. A cause and effect relationship between the altered phenotype and the deletion of a particular gene product would not be established using this approach. A more subtle mutation would be required in order to disrupt the expression of a single neuropeptide specifically. The gene targeting vectors described in this text provide the capacity to permit the assembly of future gene targeting vectors that could successfully introduce specific, less gross, neuropeptide specific mutations.

On some occasions, a targeted mutation will be lethal to the host cell and in such instances, the host organism may not reach the point of developmental maturity. The function of a neuropeptide such as NK-1 in the adult organism cannot be studied in such instances. The adult stage is the point at which the mutant phenotype would typically be analysed (for example in mammalian model systems where normal function would be under investigation as opposed the fluctuations in gene activity that may occur during development). In order to study the function of transgenic PPT-A and NK-1R genes in the nervous system of a mature, adult mouse, a conditional gene targeting approach to introducing specific mutations may be required (as mentioned at the start of Section 7.4). This would permit the development of a "latently" transgenic animal to reach maturity before triggering the mutation of interest. The Cre-*lox*P system is one such way of generating such conditional gene 'knock

outs'. The Cre-*lox*P approach can be used to precisely halt the expression of a particular gene in a particular cell type (or at a particular developmental timepoint) as demonstrated by Gu *et al*, (1994). This report details the deletion of the promoter and exon 1 of DNA polymerase β gene, but only in murine T-cells. This technique utilises the Cre-*lox*P recombination system which is intrinsic to the bacteriophage P1. The Cre recombinase induces site specific recombination by aligning two *lox*P sites which are in the same (repeat) orientation. The enzyme then excises the intervening DNA and leaves a single *lox*P site. Therefore, if the expression of the recombinase enzyme can be harnessed, the timepoint at which the targeted mutation is induced can be controlled. The Cre-*lox*P system offers this capacity but requires two transgenic strains (of mice for example). One strain carries the target gene (such as PPT-A or NK-1R) flanked by a pair of *lox*P sites in the same 5'-3' orientation. The other mouse strain will carry a *cre* recombinase expressing transgene under the specific temporal and spatial control of a promoter (for example a promoter expressed solely in differentiated adult neurons). The complete deletion of the target gene, and therefore its expression in neuronal tissue, would be possible by adopting a mouse breeding program that crossed the *cre* expressing transgenic line with a transgenic line bearing the target gene flanked by *lox*P sites. A potential problem associated with this approach would be the identification of a suitable promoter for use in the construction of the *cre* transgene. If however, the *cre* transgene were targeted to the endogenous PPT-A (or NK-1R) gene locus, the *cre* gene would fall under the transcriptional control of the PPT-A (or NK-1R) promoter. Unfortunately, this approach would knockout PPT-A (or NK-1R) gene expression in *all* tissues. This approach would therefore not be of use if haploid insufficiency were responsible for preventing germline transmission of a mutation. The chromosomal insertion site of the *cre* transgene may contribute unpredictable expression patterns especially if the transgene is transiently expressed throughout development. Under such circumstances, the *cre* recombinase enzyme may have insufficient time to excise all DNA fragments flanked by *lox*P sites. This could explain why only an estimated 40% of T-cells were homozygous for the targeted deletion of the DNA polymerase-β gene as reported by

Gu *et al* (1994). Therefore, with respect to *cre* transgene expression, the choice of promoter is critical to the likelihood of a conditional mutation strategy working successfully for any given gene. This is equally relevent for the murine PPT-A and NK-1R genes, and as discussed previously (Section 7.3.1.2), identification of the promoters and regulatory elements for those genes would be of value in the design of future conditional targeting vectors.

There are other techniques currently used to generate tissue specific mutations. The yeast FLP enzyme can manipulate genes in a similar way to the Cre-*lox*P system, by inducing the deletion of DNA fragments flanked by frt sequences, (Fiering *et al*, 1993). This system was used recently to induce site-specific recombination in transgenic mice, (Dymecki *et al*, 1996). Another more subtle means of introducing a specific germ line mutation involves the insertion of a neomycin resistance gene lacking its intrinsic promoter and transcription termination sequence. This approach is detailed further in Section 7.4.1.

The existing plasmid subclones for the murine genes (pSP10.6 and pSPR5 for example) could be used to generate further gene targeting vectors for the PPT-A and NK-1R genes respectively, using the subtle targeting approaches detailed above. Similarly, the existing gene targeting vectors (pSP7.2 V and pSPR.Neo.TK) could be modified in order to focus the targeted mutations on specific neuropeptides translated by a precursor gene or on specific regions of a translated receptor peptide. The 186 isolated NK-1R bacteriophage clones and the other 9 PPT-A bacteriophage clones could be further analysed in order to locate DNA fragments for use in future targeting vectors.

Unfortunately, mice with transgenically altered PPT-A or NK-1R genes were not generated during this research project due to time restraints. Having cloned, characterised and generated targeting vectors for both murine genes, the transfections, the tissue culture of positive ES cell clones, the analysis of those clones (by Southern analysis) and, ultimately, the generation of knock-out mice lies ahead. (Indeed, some experience of tissue culture techniques has been gained). must be Therefore the generation of targeted ES cell

lines using the vectors described must be one immediate future aim in order to produce chimaeric mice and germline transmission of that mutation. The research presented in this text does, however, lay very strong foundations for successfully doing this in the near future. The vectors are indeed ready for transfection. Access to the targeting technology which can mutate a gene in a temporal and spatial manner also offers numerous possibilities for future investigation. A host of further experiments and comparative analyses will be possible when such transgenic animals are generated and a clearer picture of the role of neurokinin-1 in the mammalian nervous system should be obtained.

6.4.1 PPT-A Gene Targeting Vector and Future Vectors

As briefly mentioned in Section 7.2, sequence from the third PPT-A exon was included in the sequence of the PCR-derived PPT-A library probe (used in Section 3.3). As this exon encodes the NK-1 peptide in the rat, bovine and human PPT-A genes (section 1.4), it is very likely to also do so in the mouse (although this remains to be confirmed experimentally). This specific location was also of interest because it could be used as the insertion site for an alternative kind of marker gene in further targeting vectors. For example, a neomycin resistance gene lacking an intrinsic promoter and transcription termination sequence could be inserted. This approach has been used successfully by Vaulont *et al*, (1995) where the adenosine deaminase (ADA) gene was disrupted in ES cells using a promoter-trap gene trageting approach. A dicistronic targeting construct containing a promoter-less IRES (internal ribosomal entry site) B *geo* cassette was inserted into the ADA gene. This permits direct, cap-independent translation of the B *geo* reporter gene which encodes a protein with B-galactosidase and neomycin activities. This type of marker gene permits the target gene to function normally except that it disrupts the synthesis of the specific translation product encoded by the disrupted exon. (This approach is of particular relevence to neurokinin-1 which is one of several peptides encoded by the PPT-A gene). Further translation products encoded downstream of that marker gene insertion site would be unaffected by this type of mutation. This approach could be used to specifically

disrupt the production of *one* translation product encoded by the murine PPT-A gene, without affecting the normal translation and functional activity of the other peptides encoded by the gene. This type of subtle mutational modification of normal gene activity is a powerful tool in understanding the roles of the individual peptides. It is very likely that neurokinin A, neuropeptide K and neuropeptide gamma are encoded by murine PPT-A (as in the rat and bovine genes) and that they could be targeted individually, using this approach, in order to investigate the role of each peptide in the mammalian nervous system.

6.4.2 NK-1R Gene Targeting Vector and Future Vectors

The murine NK-1R targeting vector (pSPR.Neo.TK, Section 5.2) contains mouse genomic DNA sequence. 8 kb of genomic sequence which includes exons 3, 4 and 5 was taken as the starting DNA. 1.5 kb of that sequence, the *Bam*H I/*Bgl* II sequence encoding exons 3 and 4, was then deleted. The remaining 6.5 kb was cloned into the targeting vector pSPR.Neo.TK. In that vector the 6.5 kb was split into two shorter sequences by the insertion of a neomycin resistance gene (see Section 5.2.1). In rat, exon 3 encodes the fifth transmembrane domain and the third intracellular region of this G-protein coupled receptor. Also in the rat, exon 4 encodes the sixth transmembrane domain, the third intracellular region and the seventh transmembrane domain of the receptor. It is very likely that the corresponding mouse exons encode similar portions of the NK-1R receptor. In a successfully targeted transgenic *mouse*, because of the very high expected mouse to rat DNA sequence identity (and receptor amino acid sequence identity), those segments of the mouse receptor encoded by exons 3 and 4 would most likely be lacking. This targeting vector is crude by design, in that a relatively large segment of the receptor peptide sequence will (theoretically) be deleted in a resultant transgenic receptor. If successfully expressed, this deletion is extremely likely to render the receptor non-functional. This may even be a lethal mutation which would prevent development of the transgenic animal to the point of maturity (in which case the Cre-*lox*P approach detailed above may be appropriate). Only by performing the gene targeting experiment

itself will this predicted alteration in the transgenic receptor be either confirmed or rejected.

A more subtle approach to targeting a gene encoding a membrane receptor (as opposed to one encoding signal peptides) would be to perform single deoxynucleotide mutagenesis experiments using PCR oligodeoxynucleotide primers with known nucleotide changes included. A single nucleotide change could be used to alter a single codon and the single amino acid encoded at that corresponding point in a transgenic peptide. This, when analysed with the help of classical pharmacology (ligands, antagonists etc), in addition to molecular biology and biochemistry, would permit the identificiation of important areas of the three dimensional receptors' amino acid sequence. The specific segments of the receptor which are critical to receptor ligand binding could be identified in this way. Similarly, segments of the receptor important to three dimensional folding of the peptide may be identified. This form of mutational analysis is very precise and also presents numerous future targeting possibilities and the means to functionally analyse the murine NK-1R receptor.

6.5 GENERAL DISCUSSION POINTS AND FUTURE RESREARCH

There are numerous further experimental approaches which are available to the researcher wishing to explore the role of the murine neuropeptide neurokinin-1 and to comprehend the activities of the genes that encode that peptide and its corresponding receptor. (This could similarly be extended to the other tachykinins and the other peptides encoded by the PPT-A gene). Several of those approaches are detailed in the sections above.

The modern day molecular biologist has an increasing number of research tools. In addition to the future experiments which are proposed in the above sections, it may be of value to clone the cDNAs

for the NK-1R gene. cDNAs encoding the preprotachykinin-A and preprotachykinin-B protein precursors respectively, were isolated from a mouse brain cDNA library by Kako *et al* (1993), but the NK-1R (receptor) gene cDNAs have not been cloned. The genomic DNAs which have been cloned and are presented in this text could be used to facilitate this if they are used as homologous probes for cDNA library screens. Cloned cDNAs could then be transcribed *in vitro* using modern expression vectors and the resultant receptor studied at the biochemical, pharmacological and molecular levels. Similarly, such a receptor could be expressed in *Xenopus* oocytes. Many facets of gene expression could be studied using tissue cultured cell lines, for example, mapping promoter elements and other regulatory elements such as enhancers and repressors (see Section 7.3.1.2). This is typically conducted by determining the effects of induced deletions within such regulatory elements. Examination of the tissue specific expresssion of the PPT-A and NK-1R genes is also possible. For example, tissue-specific elements have been found in the rat insulin gene by Edlund *et al*, (1985), where tissue specific expression is mediated by elements in the 5' flanking region. Considering the range of tissues in which neurokinin-1 (as well as the other tachykinins) are expressed, research of this nature would be valuable in understanding NK-1 expression patterns. It may well be that elements that control PPT-A and NK-1R gene expression are also located in the 5' flanking region (also see Section 7.3.1.2). The immediate 5' flanking regions (upto 5 kb) of both genes are known to be contained in the plasmid clones detailed in Section 4.3 (Figure 4.6) for the PPT-A gene, and Section 4.8 (Figure 4.13) for the NK-1R gene.

A combination of classical pharmacological research (using agonists and antagonists for example) and modern molecular biology could be utilised to gain a fuller picture of neuropeptides, their functions and inherent properties. A clearer understanding of NK-1 synthesis and its role in nociception (see Section 1.2.2.1) could be attained in this way. Similarly, it would be of value to determine the specificity of the murine NK-1R receptor for its ligands (as detailed for rat in section 1.5.2) using classical pharmacology. To date, three tachykinin receptors have been identified; the neurokinin-1 (NK-1),

NK-2 and NK-3 receptors (described in Section 1.5). The three rat receptor gene cDNAs were cloned from mammalian tissue in 1987 (Masu *et al*, 1987, Nakanishi *et al*, 1987). These receptors are pharmacologically distinct (Watson *et al*, 1987), but when the three possible ligands are added at a high enough concentration, each of the endogenous mammalian tachykinin neuropeptides can elicit a response from each of the three receptors. This is most likely to be a consequence of the common carboxy terminal amino acid sequence characteristic of the tachykinins. As for the rat NK-1 receptor, the murine NK-1 receptor is likely to be activated by more than one such ligand. Will this activation at high ligand concentrations also occur with the eqivalent mouse receptors?. The cloning of the murine NK-1R gene is detailed in ths text, however, the molecular cloning of the murine NK-2R and NK-3R receptors could also be performed where DNA fragments of the NK-1R genomic clones described here could be used as heterologous library hybridisation probes at low stringency in order to isolate genomic or cDNA clones.

Neuropeptides are also believed to influence a range of other more complex, long-term behavioural changes such as learning, arousal, memory, emotional state and schizophrenia (Gilles *et al*, 1979; McKelvy *et al*, 1986). This area also merits further, more precise examination. On a more physical level, the ligand binding sites of the NK-1R receptor protein could be further examined, using X-ray crystallography for example, to clarify its tertiary structure. Furthermore, immunocytochemical and Western blotting methods could be used to detect and size the peptides encoded by the PPT-A and NK-1R genes and permit the quantification of those peptides in tissues. Phosphorylation, glycosylation proteolysis, and methylation studies of those peptides could also be included.

Chromosomal localisation of the PPT-A and NK-1R genes can also be undertaken using the latest fluorescent labelling technology. Cyanine based dyes can be used to label selected DNA probes (for example by nick translation) which can then be used to hybridise *in situ* to chromosomal metaphase spreads for example. The human NK-1R gene is located on chromosome 2 (Shigemoto *et al*, 1990). On which chromosome is the murine NK-1R (and PPT-A) gene located?.

The neuropeptides which mediate long-range communication between neurons are traditionally referred to as 'hormones', whereas those which mediate short-range communication and occur in membrane-bound vesicles are termed 'transmitters' (Burbach *et al*, 1992). The distribution of NK-1 both in and outwith the central nervous has been documented (see Section 1.4.2 and 1.6.1). The focus of the research presented here has been on NK-1 *within* the CNS. Therefore, there is obviously scope to research NK-1 and its activities in other non-neuronal tissues in the future.

As stated in Section 1.3.1, a common amino acid sequence (*H-Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH2*) has been identified and is characteristic of the tachykinin family of neuropeptides (where X is either a hydrophobic or aromatic amino acid residue [Erspamer *et al*, 1981]). Single base pair mutational analysis of the DNA sequence encoding those terminal amino acids (using the targeting technologies detailed in Sections 1.7.4.3 and 7.4) may also help to further our understanding of neurokinin-1.

The rat PPT-A gene has been shown to be regulated by a number of stimuli including growth factors (Kessler *et al*, 1981), conditioned media (Nawa *et al*, 1990), cocaine (Hurd *et al*, 1992), inflammation (Noguchi *et al*, 1992), and factors present in innervated target tissue (Barakat-Walter *et al*, 1991). Expression of the PPT-A gene and the level of tachykinin biosynthesis appears to be under homeostatic control in rat. The same approaches as those utilised by Kessler *et al* (1981) for example, could also be used to study the equivalent murine peptide. Tachykinin expression in the anterior pituitary gland has been shown to be under hormonal control in rats (Jonassen *et al*, 1987). Although likely, is this also the case in mouse?.

As documented in Section 1.3.8, mammalian tachykinins have been associated with sensory and motor function as well as immunologic, inflammatory, cardiovascular, respiratory and gastrointestinal function. Additionally, they may act as trophic and mitogenic factors and have a role in asthma (Ichinose *et al*, 1992). Furthermore, tachykinin peptides may have some involvement with

neurogenic inflammation (as indicated by Guard *et al*, 1991) and may subsequently be involved in the processes of arthritis. In addition, the locus coeruleus region of the mamalian brain plays an integral role in the control of sleep and alertness (Kolb *et al*, 1989). Tachykinins, including neurokinin-1, have been shown to be potent activators of locus coeruleus neurons and may therefore have some role to play in the regulation of sleep A host of different experimental approaches could be used to clarify those proposed roles of the tachykinins

Clearly, there is evidence to suggest that the tachykinin neuropeptide neurokinin-1 has some association, correlation or involvement in several, normal mammalian neuronal and non-neuronal cellular processes. However, the precise role of neurokinin-1 and its corresponding receptor in those processes remains unclear. Furthermore, its role in abnormal processes remains to be seen. The research presented in this text provides a very strong foundation to further investigate this neuropeptide and its multi-faceted nature.

CHAPTER 7

REFERENCES

A

- Anastasi, A, Erspamer, V., Cei, J. M. (1968). Isolation and amino acid sequence of Caerulein and active decapeptide of the skin of *hyla caerulea. Arch. Biochem. Biophys.* **125**:57-68.
- Anastasi, A., Erspamer, V., Endean, R., (1975). Structure of uperlein, a physalaemin-like endecapeptide occuring in the skin of *Uperleia rugosa* and *Uperleia marmorata. Experimentia.* **31**: 394-305.
- Anastasi, A., Falconieri Erspamer, G., (1970). Occurence of phyllomedusin, a physalaemin-like decapeptide in the skin of *Phyllomedusa bicolor. Experimentia.* **26**: 866-867.
- Antin, J., Gibbs, J., Holt, J., and Smith, G. P., (1975). Cholecystoskeletin elicits the complete behavioural sequence of satiety in rats. *J. Comp. Physiol. Psychol.* **89**: 784-799.
- Appell, K. C., Fragale, B. J., Loscig, K., Singh, S., Tomzuk, B. E., (1992). Antagonists that demonstrate species differences in neurokinin-1 receptors. *Mol. Pharmacol.* **41**: 772-778.
- Aratani, Y., Okazaki, R., Koyamu, H., (1992). End-extension repair of introduced targeting vectors mediated by homologous recombination in mammalian cells. *Nucleic Acids Res.* **20**: 4795-4801.
- Aronin, N., Coslovsky, R., Leeman, S. E., (1986). Substance P and neurotensin. *Ann. Rev. Physiol.* **48**:537-49.
- Attwood, T. K., Eliopoulos, E. E., Findlay, J. B. C., (1991). Multiple sequence alignment of protein families showing sequence homology: a methodological approach using database pattern-matching discriminators for G-protein-linked receptors. *Gene.* **98**:153-159.

B

- Bakarat-Walter, I., Affolter, H. U., Droz. B., (1991). Expression of substance P and preprotachykinin-A mRNA by primary sensory neurons in culture: regulation by factors present in peripheral and central target tissue. *Mol. Brain Res.* **10**: 107-114.
- Bannon, M. J., Freeman, A. S., Chiodo, L. A., Bunney, B. S., Roth, R. H., (1987). The electrophysiological and biochemical pharmacology of the mesolimbic and mesocortical dopaminergic neurons. *Handbook of Psychopharmacology.* Vol **19**. (Iversen, L. L., Iversen, S. D., and Snider, S. H., eds) p329-374. Plenum, New York.
- Bar-Shavit, Z. R., Goldman, Y., Stabinsky, E. A., (1980). Enhancement of phagocytosis:- a newly found activity of substance P residing in its N-terminal peptide sequence. *Biochem. Biophys. Res. Comm.* **94**: 1445-1449
- Bartho, E. R., Holzer, P., (1985). Search for a physiological role for substance P in gastrointestinal motility. *Neuroscience.* **6**: 1-32.
- Benuajati, C., Spoerei, N., Haymerle, H., Ashburner, M., (1983). The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5' end in different developmental stages. *Cell.* **33**: 125-133.
- Berger, H., Fencher, K., Albrecht, E., Neidrich, H., (1979). Substance P: In vitro inactivation by rat brain fractions and human plasma. *Biochem. Pharmacol.* **28**: 3173-3180.
- Bergstrom, L, Torrens, Y., Saffroy, M., Beaujouan, J. C., Lavielle, S., Chassaing, G., Morgat, J. L., Glowinski, J., Marquet, A., (1987). [^{3}H]neurokinin-B and ^{125}I-Bolton Hunter eledoisin label identical tachykinin binding sites in the rat brain. *J. Neurochem.* **49**: 125-133.
- Black, I. B., (1981). Nerve growth factor stimulates development of substance P in the embryonic spinal cord. *Brain Res.* **208**: 135-145.
- Blobel, B., Dobberstein, B., (1975). Transfer of proteins across membranes part 1: Presence of proteolytically processed and unprocessed nascent imunoglobulin light chains on membrane bound ribosomes of murine myeloma. *J. Cell Biol.* **67**:835-51.
- Blumberg, V., Teichberg, J.L., Charli, L. B., Hersh, B., McKelvy, J. F. (1980). Cleavage of substance P to an N-terminal tetrapeptide and

C-terminal heptapeptide by a post-proline cleaving enzyme from bovine brain. *Brain Res.* **192**:477-486.

- Bonner, T. I., Affolter, H. U., Young, A., Young, W. S., (1987). A cDNA encoding the precursor of the rat neuropeptide neurokinin-B. *Mol. Brain. Res.* **2**: 243-249.
- Bowen, D. M., (1988). Neurotransmitters in Alzheimer's disease. *Age.* **11**: 104-111.
- Brene, S., Lindefors, N., Herrera-Marschitz, M., Persson, H., (1993). Differential regulation of PPT-A mRNA expression in striatum by excitation of hippocamal neurons. *Eur. J. Neurosci.* **5**: 839-845.
- Brown, M. C. S., Biagi, M., Renzetti, A. R., Rovero, P., Mizrahi, J., (1992). [^{125}I]His-neurokinin-A binds selecyively to NK-2 receptors of the B-type in rat small intestine smooth muscle. *Eur. J. Pharmacol.* **227**: 163-171.
- Brownstein, M. J., Mroz, E. A., Kizer, J. S., Palkovits, M., Leeman, S. E., (1976). Regional distribution of substance P in the brain of rat. *Brain Res.* **116**: 299-305.
- Brownstein, M. J., Mroz, E. A., Tappaz, M. L., Leeman, S. E., (1976). Substance P and glutamic acid decarboxylase (GAD) in the a nigra. *Brain Res.* **135**: 315-323.
- Brun, A. (1983). in *Alzheimers Disease*, (Reisberg, B., Ed.) p37-47. The Free Press, New York.
- Buck, S. H., Helke, C. J., Burcher, E., Shults, C. W., O'Donohue, T. L., (1986). Pharmacological charcterisation and autoradiographic distribution of binding sites for iodinated tachykinins in the rat central nervous system. *Peptides.* **7**: 1109-1120.
- Buell, M. F., Schulz, S. J., Arkinstall, K., Maury, M., Missotten, N., Adami, F., Kawashima, E., (1992). Molecular characterisation, expression and localisation of human neurokinin-3 receptor. *FEBS Lett.* No. 1, **229**: 90-95.
- Bullock, W. O., Fernandez, J. M., , Short, J. M., (1987). XL1-Blue: a high effciency plasmid transforming *rec*A Escherichia coli strain with glycophosphatidylinositol-specific selection. *Biotechniques.* **5**: 376-378.
- Buongiorno-Nardelli, M., Amaldi, F., (1969). Autoradiographic detection of molecular hybrids between rRNA and DNA in tissue sections. *Nature.* **225**: 946-7.

- Burbach, J. P. H., and Meijer, O. C., (1992). The structure of neuropeptide receptors. *Eur. J. Pharmacol - Molecular Pharmacol. Section.* **227**:1-18.

C

- Carter, M. S., Krause, J. E., (1990). Structure, expression, and some regulatory mechanisms of the rat preprotachykinin gene encoding substance P, neurokinin-A, neuropeptide K, and neuropeptide gamma. *J. Neurosci.* **10** No. 7:2203-2214.
- Cascieri, M. A., Ber, E., Fong, T. M., Sadowski, S., Bansal, A., Burns, D., Strader, C. D., (1992). Characterisation of the binding of a potent, selective radioiodinated antagonist to the human neurokinin-1 receptor. *Mol. Pharmacol.* **42**: 458-463.
- Chalfie, M., Wolinsky, E., (1990). The identification and suppression of inherited neurodegeneration in *caenorhabditis elegans. Nature* **345**:410-416.
- Chambers, D. A., Cohen, R. L., Perlman, R. L., (1993). Neuroimmune modulation: signal transduction and catecholamines. *Neurochem. Int.* **22**: 95-110.
- Charlton,C. G., Helke, C. J., (1985). Autoradiographic localisation and characterisation of spinal cord substance P binding sites: high densities in sensory, autonomic, phrenic and Onuf's motor nuclei. *J. Neurosci.* **5**: 1653-1661.
- Chang, M.M., Leeman, S. E., Niall, H. D., (1971). Amino acid sequence of substance P. *Nature.* **232**: 86-87.
- Chartier-Harlin, M. C., Crawford, F., Houlden, H., Warren, A., Hughes, D., Fidani, L., Goate, A., Rossor, M., Roques, P., Hardy, J., Mullan, M., (1991). Early-onset Alzheimer's disease caused by mutation at codon 717 of the beta-amyloid precursor protein gene. *Nature.* **353**:844-845.
- Conlon, J. M., Deacon, C. F., O'Toole, L, Tim, L., (1986). Scyliorhinin I and II: two novel tachykinins from dogfish gut. *FEBS Lett.* **200**: 111-116.
- Corjay, M. H., Dobrzanski, D. J., Way, J. M., Viallet, J., Shapira, H., Sausville, E. A., Battey, J. F., (1991). Two distinct bombesin receptor subtypes are expressed and

functional in human lung carcinoma cells. *J. Biol. Chem.* **266**(28): 18771.

- Cuello, A. C., Del Fiacco, M., Paxinos, G., (1978). The central and peripheral ends of the substance P-containing neurones in the rat trigeminal system. *Brain Res.* **152**: 499-509.
- Czicki, U, Lennarz, W. J. (1977). Localisation of the enzyme system for glycosylation of proteins via the lipid-linked pathway in rough endoplasmic reticulum. *J. Biol. Chem.* **252**:7901-4.

D

- Dalsgaard, C. J., Jonnson, C. E., Hokfelt, K., Cuello, A. C., (1983). Localisation of substance P-immunoreactive fibres in the human skin. *Experientia.* **42**: 523-528.
- D'Amato, R. J., Zweig, R. M., Whitehouse, P. J., Wenk, J. L., Singer, H. S., Mayeux, R., Price, D. L., Snyder, D. H., (1987). Aminergic systems in Alzheimer's and Parkinson's disease. *Ann. Neurol.* **22**: 229-236.
- Davies, J. E., Dray, A., (1976). Substance P in the substantia nigra. *Brain Res.* **107**: 623-627.
- Davies, P., Katzman, R., Terry, R. D., (1980). Reduced somatostatin-like immunoreactivity in cerebral cortex from cases of Alzheimer's disease and Alzheimer's senile dementia. *Nature.* **288**: 279280.
- Davies, R. W., Gallagher, E. J., Savioz, A., (1994). Reverse genetics of the mouse central nervous system: Targeted genetic analysis of neuropeptide function and reverse genetic screens for genes involved in human neurodegenerative disease. *Prog. in Neurobiol.* **42**:319-31.
- De Felipe, C., Pinnock, R.D., Hunt, S.P., (1995). Modulation of chemotropism in the developing spinal cord by SP. *Science.* **267**:899-902.
- Dente, L., Gianni, C., Cortese, R., (1983). pEMBL: A new family of single stranded plasmids. *Nuc. Acids Res.* **11**:1645-1654.
- Dietl, M. M., Palacios, J. M., (1991). Phylogeny of taqchykinin receptor localisation in the vertebrate central nervous system: apparent absence of NK-2 and NK-3 binding sites in the human brain. *Brain Res.* **539**: 211-222.

- Dixon, R. A. F., Strader, C. D., Sigal, I. S., (1988). Structure and function of G-protein coupled receptors. *Annual Reports in Medicinal Chemistry*, ed. Allen, R. C., **23**: 221-233.
- Donehower, L.A., Harvey, M., Slagle, B.L., McArthur, M.J., Montgomery Jr, C.A., Butel, J.S., Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**: 215-221.
- Duffy, M. J., Mulhall, D., Powell, D., (1975). Subcellular localisation of substance P in bovine hypothalus and substantia nigra. *J. Neurochem.* **25**: 305-307.
- Dymecki, S.M. (1996). Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 6191-6196.

E

- Edlund, T., Walker, M. D., Barr, P. J., Rutter, W. J., (1983). Cell specific expression of the rat insulin gene: Evidence for the role of 2 distinct 5' flanking elements. *Science.* **230**: 912-916.
- Eipper, B. A., Mains, R. E. (1980). Structure and biosysnthesis of pro-adrenocorticotropin/endorphin and related peptides. *Endocrin. Rev.* **1**:1-27.
- Elling, C. E., Neilsen, S. M., Shwartz, T. W., (1995). Conversion of antagonist-binding site to metal ion site in the tachykinin NK-1 receptor. *Nature.* **374**: 74-77.
- England, P. J. (1980). Regulation of phosphorylation and dephosphorylation. *The Enzymology of Post-Translational Modification of Proteins.* Ed., R. B. Freedman and H. C. Hawkins. p1-423. London:Academic.
- Ericsson, A., Geenan, V., Robert, F., Legros, J-L., Vrindts-Gervert, Y., Franchimont, P., Brene, S., Persson, H., (1990). Expression of preprotachykinin-A and neuropeptide Y mRNA in the thymus. *Endocrinology.* **4**: 1211-1218.
- Erlich, H. A., ed., (1989). PCR Technology: Principles and applications for DNA amplification. Stockton Press.

- Erspamer, V., Anastasi, A., (1962). Structure and pharmacological actions of eledoisin, the active undecapeptide of the posterior salivary glands of *Eledone*. *Experimentia.* **18**: 58-59.
- Erspamer, V., Anastasi, A., Bertaccini, G., Cei, J. M., (1964). Structure and pharmacological actions of phylasaemin, the main active polypeptide of the skin of *Physalaemus fuscumaculatus. Experimentia.* **20**: 489-90.
- Erspamer, V., (1971). Biogenic amines and active polypeptides in the amphibian skin. *Ann. Rev. Pharmacol.* **11**: 327-50.
- Erspamer, V., The tachykinin peptide family (1981). *TINS.* **4**:267.
- Ellis, J., Bernstein, A., (1989). Gene targeting with retroviral vectors - recombination by gene conversion into regions of nonhomology. *Molr. Cellr. Biology.* **9**: 1621-1627.
- Euler, U. S., von, Gaddum, J. H., (1931). An unidentified suppresssor substance in certain tissue extracts. *J. Physiol.* **72**: 74-87
- Euler, U. S. v., Pernow, B., (1977). Substance P. *Nobel Symposium 37.* Raven Press, New York.

F

- Fahrbach, S. E., Morell, J. I., and Pfaff, D. W., (1986). *Physiol. Behav.* **37**:135-139.
- Fiering, S., Kim, C.G., Epner, E.M., Groudine, M. (1993). An 'in-out' strategy using gene targeting and FLP recombinase for the functional dissection of complex DNA regulatory elements: Analysis of the b-globin locus control region. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 8469-8473.
- Findlay, J., Eliopoulos, E., (1990). Three dimensional modelling of G-protein -linked receptors. *TIPS.* **11**: 492-497.
- Fitzsimmons, T. J. (1978). *Neurosci. Res. Prog. Bull.* **16**: 307-322.
- Fong, T. M., Andrson,S. A., Yu, H., Huang, R. R., Strader, C. D., (1992). Differential activation of intracellular effector by two isoforms of human NK-1 receptor. *Mol. Pharm.* **41**: 24-30.
- Fong, T. M., Huang, R. C., Stader, J. D., (1992). Localization of antagonist binding domains of the human neurokinin-1 receptor. *J. Biol. Chem.* **267**: 25664-2567.

- Fong, T. M., Cascieri, M. A., Yu, H., Strader, C. D., (1993). Amino-aromatic interaction between histidine 197 of neurokinin-1 receptor and CP96345. *Nature.* **362:** 350-353.
- Foreman, J., Jordan, C., (1983). Histamine release and vascular changes induced by neuropeptides. *Agent Actions.* **13**:105.
- Forman, M. M., and Moss, R. L., (1977). Effects of subcutaneous injection and intrahypothalamic infusion of releasing hormones upon lordotic response to repetitive coital stimulation. *Horm. and Behav.* **8**: 219-227.
- Forno, L. S., (1982). in *Movement Disorders, Neurology,* Vol. 2. (Marsden, C. D., Fahn, S., Eds.) p25-40, Butterworth Scientific, London.
- Furuta, H., Guo, D. F., Inagami, T., (1992). Molecular cloning and sequencing of the gene encoding human angiotensin II type 1 receptor. *Biochem. Biophys. Res Commun.* **183**(1): 8

G

- Gaddum, J. H., Schild, H., (1934). Depressor substances in extracts of intestine. *J. Physiol. London.* **83**:1-34.
- Gainer, H., Sarne, Y., Brownstein, M. J. (1977). Biosynthesis and axonal transport of rat neurohypophyseal proteins and peptides. *J. Cell. Biol.* **73**:366-81.
- Gale, J. S., Bird, E. D., Spokes, E. G., Iversen, L. L., Jessel, T., (1978). Human brain substance P: Distribution in controls and Huntingdons chorea. *J. Neurochem.* **30**: 633-634.
- Gamse, R., Petsche, U., Lembeck, F., Janso, G., (1982). Capsaicin applied to peripheral nerve inhibits axoplasmic transport of substance P and somatostatin. *Brain Res.* **239**: 447-462.
- Gamse, R., Holzer, P., Lembeck, F., (1983). Decrease in substance P in afferent neurons and impairment of plasma extravasion by capsaicin. *Br. J. Pharmacol.* **68**:207.
- Garrett, C., Carruette, A., Fardin, V., Moussanoui, S., Peyronil, J. F., Blanchard, J. C., (1991). Pharmacological properties of a potent and selective non-peptide substance P antagonist. *Proc. Natl. Acad. Sci.* **88**: 10,208-10,212.
- Gether, U., Yokota, Y., Edmionds-Alt, X., Breliere, J-C., Lowe, J. A., Snider, R. M., Nakanishi, S., Schwartz, T., (1993). Two non-peptide

tachykinin antagonists act through epitopes on corresponding segments of the NK-1 and NK-2 receptors. *Proc. Natl. Acad. Sci.* **90**: 6194-6198.

- Gether, U., Nillson, L., Lowe, J., Shwartz, T., (1994). Two specific residues at the top of transmembrane segment V and VI of the NK-1 receptor involved in binding of the non-peptide antagonist CP96345. *J. Biol. Chem.* **269**: 23959-23964.
- Gether, U., Johansen, T. E., Snider, R. M., Lowe, J. A., Nakanishi, S., Schwartz, T., (1993). Different binding epitopes on the NK-1 receptor for substance P and a non-peptide antagonist. *Nature*, **362**: 345-348.
- Gerard, N. P., Eddy, R. L., Shows, T. B., Gerard, C. (1990). The human neurokinin A (substance K) receptor. Molecular cloning of the gene, chromosomal localisation and isolation of a cDNA from tracheal and gastric tissues. *J. Biol. Chem.* **265**(33): 20455.
- Gerfin, C. R., (1991). Substance P (NK-1) receptor mRNA is selectively expressed in cholinergic neurons in the striatum and basal forebrain. *Brain Res.* **556**: 165-170.
- Gilles, G., Lowery, P., (1979). Corticotropin releasing factor may be modulated vasopressin. *Nature.* **278**: 463-465.
- Gilman, A. G., (1987). G Proteins: Transducers of receptor-generated signals. *Ann. Rev. Biochem.* **56**:615-49.
- Gluzman, Y., (1981). SV40-transformed simian cells support the replication of early SV40 mutants. *Cell.* **23**: 175-179.
- Goldman, J. E., (1983). The association of actin with Hirano bodies. *J. Neuropathol. Exp. Neurol.* **42**:146-152.
- Goosen,B., Caughman, S. H., Harford, J. B., Klausner, R. D., Hentze, M. W., (1990). Translational repression by a complex between the iron-responsive element of ferritin mRNA and its specific cytoplasmic binding protein is position-dependent *in vivo*. *EMBO.* **9**: 4127-4133.
- Gottfries, C. G., Gottfries, I, Roos, B. E., (1969). Homovanillic acid in the cerebrospinal fluid of patients with senile dementia, presenile dementia and parkinsinism. *J. Neurochem.* **16**: 1341-1345.
- Grabner, K., Lembeck, F., Neuhold, K., (1959). Substanz P im gehrin verscheidener species. *Naunyn-Schmeidebergs Arch. Exp. Pathol. Pharmakol.* **236**: 331-334.

- Gu, H. Marth, J.D., Orban, P.C., Mossmann, H., Rajewsky, K. (1994). Deletion of a DNA polymerase b gene segment in T cells using cell type-specific gene targeting. *Science.* **265**: 103-106.
- Guard, S., Watson, S. P, Maggio, J. E., Phon Too, H., Watling, K. J., (1990). Pharmacological analysis of [^{3}H]-senktide binding to NK-3 tachykinin receptors in guinea-pig ileum longitudinal muscle-myenteric plexus and cerebral cortex membranes. *Br. J. Pharmacol.* **99**: 767-773.
- Gubler, U., Seeburg, P., Hoffman, B. J., Gage, L. P. (1982). Molecular cloning establishes pro-enkephalin as precursor of enkephalin-containing peptides. *Nature.* **295**:206-8.

H

- Hanson, G. R., Alphs, L., Wolf, W., Levine, R., Lovenberg, W., (1981). Haloperidol-induced reduction of nigral substance P-like immunoreactivity: A probe for the interactions between dopamine and substance P neuronal systems. *J. Pharmacol. Exp. Ther.* **218**: 568-574.
- Harada, Y., Takahaski, T., Kuno, M., Nakayama, K., Masu, Y., Nakanishi, S., (1987). Expression of two different tachykinin receptors in *Xenopus* oocytes by exogenous mRNAs. *J. Neurosci.* **7**: 3265-3273.
- Hargrave, P. A., (1982). Rhodopsin. *Prog. Retinal Res.* **1**: 1-14.
- Harmar, A., Schofield, J. G., Keen, P., (1981). Substance P biosynthesis in dorsal root ganglia: An immunochemical study of (^{35}S)methionine and (^{3}H) Proline incorporation *in vitro. Neuroscience.* **6**: 1917-1922.
- Harmar, A. J., Keen, P. M., (1986). Methods for the identification of neuropeptide processing products: somatostatin and the tachykinins. *Meth. Enzymol.* **124**:335-33.
- Harmar, A. J., Armsrong, A., Pascall, J. C., Chapman, K., Rosie, R., Curtis, A., Going, J., Edwards, C. R. W., Fink, G., (1986). cDNA sequence of *B*-preprotachykinin, the common precursor to substance P and neurokinin-A. *FEBS Letters* **208** No.1:67-72.
- Harmar, A. J., Hyde, V., Chapman, K., (1990). Identification and cDNA sequence of δ-preprotachkinin, a fourth splicing variant of the rat subctance P precursor. *FEBS Letts.* **275**: 22-24.

- Hasty, P., Ramirez-Solis, R., Krumlauf, R., Bradley, A., (1991a). Introduction of subtle mutations into the Hox-2.6 locus in embryonic stem cells. *Nature.* **350**: 243-246.
- Hasty, P., Rivera-Perez, J. & Bradley, A. (1991b). The length of homology required for gene targeting in embryonic stem cells. *Molr. Cellr. Biol.* **11**: 5586-5591.
- Haverstick, D. M., Jeziorski, M., Bannon, M. J., (1989). Developmental profile of preprotachykinin mRNAs in rat striatum. *Soc. Neurosci. Abstr.* **15**: 1273.
- Henry, J. L., Couture, R., Cuello, A. C., Pelletier, G., Quirion, R., Regoli, D., (eds.) (1987). *Substance P and Neurokinins.* Springer-Verlag, New York.
- Henry, J. L., ed., (1989). *Substance P and Neurokinins.* Springer-Verlag, New York .
- Helke, C. J., Charlton, C. G., Wiley, R. G., (1986). Studies on the cellular localisation of spinal cord substance P receptors. *Neuroscience.* **19**: 523-533.
- Helke, S. J., Krause, J. E., Mantyh, P. W., Couture, R., Bannon, M. J., (1990). Diversity in mammalian tachykinin peptidergic neurons: multiple peptides, receptors, and regulatory mechanisms. *FASEB. J.* **4**:1606-1615.
- Helke, S. J., Krause, J. E., Mantyh, P. W., Couture, R., Bannon, M. J., (1990). Diversity in mammalian tachykinin peptidergic neurons: multiple peptides, receptors, and regulatory mechanisms. *FASEB. J.* **4**:1606-1615.
- Herkenham, M. (1987). Mismatches between neurotransmitter and receptor localisations in brain; observations and implications. *Neuroscience.* **23**: 1-38.
- Hershey, A. D., Krause, J. E., (1990). Molecular cloning of cDNAs for the rat substance P receptor gene. *Science.* **247**: 958-962.
- Hershey, A. D., Dykema, P. E.,Krause, J. E., (1991). Organisation, structure, and expression of the gene encoding the rat substance P receptor. *J. Biol. Chem.* **266** No. 7:4366-4374.
- Hokfelt, T., Johannson, A., Ljungdahl, A., Schultzberg, M., (1980). Peptidergic neurons. *Nature.* **285**:515.
- Hooper, M. L., (1993). Embryonic stem cells. In *Introducing changes into the animal germline.* Chapter 3. Harwood Academic Publishers: Switzerland.

- Hu, G., (1993). DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment. *DNA and Cell Biology.* **12**: 763-770.
- Hurd, Y. L., Brown, E. E., Finlay, J. M., Fibiger, H. C., Gerfen, C. R., (1992). Cocaine self-administration differentially alters mRNA expression of striatal peptides. *Mol. Brain. Res.* **13**: 165-170.

I

- Ichinose, M., Nakajima, N., Takahashi, T., Yamauchi, H., Inoue, H., Takishima, T., (1992). Protection against bradykinin-induced broncho-constriction in asthmatic patients by neurokinin receptor antagonist.*The Lancet.* **340**: 1248-1251.
- Innis, R. B., Andrade, R., Aghajanian, K., (1985). Substance K excites dopaminergic and non-dopaminergic neurons in rat substantia nigra. *Brain Res.* **334**:381-383.
- Iverson, L. L., Rossor, M. N., Reynolds, G. P., Hills, R., Rothh, M., Mountjoy, C. Q., Foote, S. L., Morrison, J. H., Bloom, F. E., (1983). Loss of pigmented dopamine B-hydroxylase positve cells from locus coeruleus in senile dementia of Alzheimer's type. *Neurosci. Lett.* **39**: 95-100.
- Iverson , L. L., Watling, K. J., McKnoght, A. T., Young, S. C., Williams, B. J., (1990). Pharmacology of the tachykinin system. In *Neuropeptides and their receptors.* ed Schwartz, T. W., Hilsted, L. M., Alfred Benzon Symposium No. 29: 363-372.
- Iwai, N., Yamano, Y., Chaki, S., Konishi, F., Bardham, S., (1991). Rat angiotensin II receptor: cDNA sequence and regulation of the gene expression. *Biochem. Biophys. Res. Commun.* **177**(1):299.

J

- Jasin, M., Berg., P. (1988). Homologous integration in mammalian cells without target gene selection. *Genes and Developt.* **2**: 1353-1363.
- Jeannotte, L., Rutz, C., Robertson, E. J., (1991). Low level Hox 1.3 gene expression does not preclude the use of promoterless

vectors to generate a targeted gene disruption. *Molec. Cell. Biol.* **11**: 5578-5585.

- Jonassen, J. A., Mullikin-Kilpatrick, D., McAdam, A., Leeman, S. E., (1987). Thyroid hormone status regulates PPT-A gene expression in male rat anterior pituitary. *Endocrinology.* **121**: 1555-1561.
- Ju, G., Liu, S., (1989) Relationships of substance P immunoreactive nerve fibres with somatotropes of the anterior pituitary in the monkey. *J, Neuroendocrinology.* **1**: 397-400.

K

- Kage, R., Leeman, S.E., Boyd, N. D., (1993). Biochemical characterisation of two different forms of the substance P receptor in the rat submaxillary gland. *J. Neurochem.* **60** No. 1:348-350.
- Kako, K., Munekata, E., Kosaka, M., Nakayama, K., (1993). Cloning and sequence analysis of mouse cDNAs encoding preprotachykinin A and B. *Biomed. Res.* **14** No.4: 253-259.
- Kandel, E. R., and Scwartz, J. H., (1981). *Principles of Neural Science.* New York: Elsevier North Holland.
- Kangawa, K., Minamino, N., Fukuda, A., Matsuo, H., (1983). Neuromedin K: a novel tachykinin identified in porcine spinal cord. *Biochem. and Biophys. Res. Comms.* **114**: No. 2, 533-540.
- Kawaguchi, Y., Hoshimaru, M., Nawa, H., Nakanishi, S., (1986). Sequence analysis of cloned cDNA for rat SP precursor: existence of a third SP precursor. *Biochem. Biophys. Res. Comms.* **139** No. 3:1040- 1046.
- Kelley, A. E., Iversen, S. D., (1978). Behavioural response to bilateral injections of substance P into the substantia nigra. *Brain Res.* **158**: 474-478.
- Kimura, S., Sugita, Y., Kanazawa, I., Munekata, E., (1983). Novel neuropeptides, neurokinin-β and γ, isolated from porcine spinal cord. *Acad. Ser. B. Phys. Biol. Sci.* **59**: 101-104.
- Kimura, S., Goto, K., Ogawa, T., Sugita, Y., Kanazawa, I., (1984). Pharmacological characterisation of novel mammalian tachykinins, neurokinin alpha and neurokinin beta. *Neurosci. Res.* **2**: 97-104.

- Kolb, B., Whishaw, I. Q., (1989) Fundamentals of neuropsychology, third edn. Freeman Press.
- Konishi, S., Otsuka, M., (1974). Excitatory action of hypothalamic substance P on spinal motorneurones of newborn rats. *Nature.* **252**:734-735.
- Kotani, H., Hoshimaru, M., Nawa, H., Nakanishi, S., (1986). Structure and gene organisation of bovine neuromedin K precursor. *Proc. Natl. Acad. Sci.* **83**:7074-78.
- Kovalic, D., Kwan, J.H., Weisblum, B., (1991). General method for direct cloning of DNA fragments generated by polymerase chain reaction. *Nuc. Acids Res.* **19**:1560-1562.
- Kowall, N. W., Beal, M. F., (1988). Cortical somatostatin,, neuropeptide Y, and NADPH diaphorase neurons: normal anatomy and alterations in Alzheimer's disease. *Ann. Neurol.* **23**: 105-114.
- Kowall, N. B., Beal, M. F., Busciglio, J., Duffy, K., Yankner, B. A., (1991). An in vitro model for the neurodegenerative effects of B-amyloid and protection by substance P. *Proc. Nat Acad. Sci.* **88**: 7247-7251.
- Krause, J. E., Cremins, J. D., Carter, M. S., Brown, E. R., McDonald, M. R., (1989). Solution hybridisation-nuclease protection assay for sensitive detection of differentially spliced substance P and neurokinin-A-encoding messenger ribonucleic acids. *Methods Enzymol.* **168**: 634-652.
- Krause, J. E., McDonald, M. R., Takeda, Y. (1989). The polyprotein nature of substance P precursors. *Bioessays.* **10**: 62-69.
- Kream, R. M., Schoenfield, T. A., Mancuso, R., Clancey, L. N., El-Bermani, L., (1985). Precursor forms of substance P in nervous tissue: detection with antisera to SP, SP-Gly, and SP-Gly-Lys. *Proc. Nat. Acad. Sci. USA.* **82**: 4832-36.

L

- Land, H., Schutz, G., Schmale, H., Richter, D., (1982). Nucleotide sequence of the cloned cDNA encoding bovine arganine vasopressin-neurophysin II precursor. *Nature.* **295**:299-303.
- Laneuville, O., Dorais, J., Couture, R., (1988). Characterisation of the effects produced by neurokinins and three agonists selective

for receptor subtypes in spinal nociceptive reflex in the rat. *Life Sci.* **42**: 1295-1305.

- Lee, C. M., Emerson, P. C., Iversen, L. L., (1980). The development and application of a novel N-terminal directed substance P antiserum. *Life Science.* **27**: 535-543.
- Lee, C. M., Iverson, L.L., Hanley, M. R., Sandberg, B. E. B., (1982). *Naunyn-Schmeidelbergs Arch. Pharmacol.* **318**: 281-287.
- Lee, C. M., (1982). Enzymatic inactivation of substance P in the CNS. In *Substance P in the nervous system* (Ciba Foundation sysmposium **91**) ed. by Porter, R., O'Connor, M., p165-185. Pitman Press, London.
- Leeman, S. E., Hammerschlag, R., (1967). Stimulation of salivary secretion by a factor from hypothalamic tissue. *Endocrinology.* **81**: 803-810.
- Lefkowitz, R. J., Kobilka, B. K., Caron, M. G., (1989). The new biology of drug receptors. *Biochem. Pharmacol.* **38**: 2941-2948.
- Li, X-Y., Wu, Y-N., North, R. A., Forte, M., (1992). Cloning, functional expression and developmental regulation of a neuropeptide Y receptor from Drosophila melanogaster. *J. Biol. Chem.* **267**(1): 9.
- Lin, F. L., Sperle, K., Sternberg, N., (1985). Recombination in mouse L cells between DNA introduced into cells and homologous sequences. *Proc. Natl. Acad. Sci.* **82**: 1391-1395.
- Lin, H. Y., Harris, T. L., Flannery, M. S., Aruffo, A., Kaji, E. H., Gorn, A., Lodish, H. F., Goldring, S. R., (1991). Expression and cloning of an adenylate cyclase-coupled calcitonin receptor. *Science.* **254**: 1022.
- Linderstrom-Lang, K. U., Ottesen, M., (1949). Formation of plakalbumin from ovalbumin. *C. R. Trav. Lab. Carlsborg.* **26**:403-412.
- Ljungdahl, A., Hokfelt, T., Nilsson, G., (1978). Distribution of substance P-like immunoreactivity in the central nervous system of the rat: cell bodies and nerve terminals. *Neuroscience* **3**: 861-943.
- Lundberg, J. M., Kokfelt, T., Anggard, A., Pernow, B., Emson, P., (1979). Immunohistochemical evidence for substance P immunoreactive nerves in the taste buds of the cat. *Acta. Physiol. Scand.* **107**: 389-391.

- Lundberg, J. M., Tatemoto, K., (1982). Pancreatic polypeptide family (APP, BPP, NPY, PYY) in relation to α-adrenoreceptor resistant sympathetic vasoconstrictor. *Acta. Physiol. Scand.* **116**: 393-402.
- Lundberg, J. M., Brodin, E., Saria, A., (1983). Origin and distribution of vagal capsaicin-sensitive substance P neurons with special reference to the lungs and trachea. *Acta. Phsiol. Scand.* **31**: 87092.

M

- Maggi, C. A., Giuliana, S., Santicioli, P., Regoli, D., Meli, A., (1987). Peripheral effects of neurokinins: functional evidence for the existence of multiple receptors. *J. Auton. Pharmac.* **7**:11-32.
- Maggi, C. A., Patachini, R., Guiliana, S., Rovero, P., Dion, S., Regoli, D., (1990). Competetive antagonists discriminate between NK-2 receptor subtypes. *Br. J. Pharmacol.* **100**: 588-592.
- Maggio, J. E., Sandberg, B. E. B., Bradley, C.V., Iversen, I. L., Santikarn, S., Williams, D. H., Hunter, J. C., Hanley, M. R., (1983). Substance K: a novel tachykinin in mammalian spinal cord. in *Substance P-Dublin*, 1983 (Skrabanek, P., Powell, D., Eds), Dublin, Ireland, p20-21.
- Maggio, J. E., Hunter, J. C., (1984). Regional distribution of kassinin-like immunoreactivity in rat central and peripheral tissues and the effect of capsaicin. *Brain Res.* **307**: 370-373.
- Maggio, J. E., (1988). Tachykinins. *Ann. Rev. Neurosci.* **11**:13-28.
- Mansour, S. L., Thomas, K. R., Capecchi, M. R., (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature.* **336**: 348-353.
- Mantyh, P. W., Pinnock, R. D., Downes, C. P., Goedert, M., Hunt, S. P., (1984). Substance P receptors: correlation with substance P induced inositol phospholipid hydrolysis in the rat central nervous system. *Nature.* **309**: 795-797.
- Mantyh, P. W., Gates, T., Mantyh, C. R., Maggio, J. E., (1989). Autoradiographic localisation and characterisation of tachkinin receptor binding sites in the rat brain and peripheral tissues. *J. Neurosci.* **9** No.1:258-279.

- Mantyh, P. W., Johnson, D. J., Boemer, C. J., Catton, M. D., Vinters, H. V., Maggio, J. E., Too, H. P., Vigna, S. R., (1989). Substance P receptor binding sites are expressed by glia *in vivo* after neural injury. *Proc. Natl. Acad. Sci.*. **86**: 5193-5197.
- Mantyh, P. W., Gates, T. S., Mantyh, C. R., Maggio, C. R., (1989). Autoradiographic localisation and characterisation of tachykinin receptor binding sites in the rat brain and peripheral tissues. *J. Neurosci.* **9**: 258-279.
- Marsden, C. D., (1982). Neurotransmitters and CNS disease: Basal ganglia disease. *The Lancet.* **11**:1141-1146.
- Marsden, C. D., (1986). Movement disorders and the basal ganglia. *TINS* Oct:512-515.
- Masu, Y., Nakayama, K., Tamaki, H., Harada, Y., Kuno, M., and Nakanishi, S., (1987). cDNA cloning of bovine substance K receptor through oocyte expression system. *Nature.* **329**:836-838.
- Matsumoto, S. G., (1994). Neuronal differentiation in cultures of murine neural crest: Neurotransmitter expression. *Devl. Brain Res.* **83**: 1-16.
- Matsuto, T., Yanagisawa, M., Otsuko, M., Kanazawa, I., Munekata, E., (1984).The excitatory action of the newly discovered mammalian tachykinins, neurokinin alpha, neurokinin beta, on neurons of the isolated spinal cord of the newborn rat. *Neurosci. Res.* **2**: 105-110.
- Matute, T., (1992). Distribution of neurons expressing SPR mRNA in immature anD adult cat visual cortex. *Exp. Brain Res.* **97**: 295-300.
- McDonald, M. R., Takeda, G., Rice, C. M., Krause, J. E., (1989). Multiple tachykinins are produced and secreted upon posttranslational processing of the three substance P precursor proteins, α-, β- and γ-preprotachykinin. *J. Biol. Chem.* **264**: 15578-15592.
- McEachern, A. E., Shelton, E. R., Bhakta, S., Obernolte, R., Bach, C., (1991). Expression cloning of a rat B2 bradykinin receptor. *Proc. Natl. Acad. Sci. USA.* **88**: 7724.
- McGregor, G. P., Hartel, R., Haberberger, R., Kummer, W., Voigt, V., (1995). Preprotachykinin-A gene expression occurs transiently in the developing rat endocrine pancreas and can be regulated in RINm5F cells. *Endocrinology.* **136**: No. 6:2538-2546.

- McKelvy, J. F., and Blumberg, S., (1986). Ann. Rev. Neurosci. **9**:277-305.
- McKnight, A. T., McGuire, J. J., Elliot, N. J., Fletcher, A. E., Foster, A. C., Tridgett, R., Williams, B. J., Longmore, J., Iversen, L. L., (1991). Pharmacological specificity of novel, sysnthetic cyclic peptides as antagonists at tachykini receptors. *Br. J. Pharmacol.* **104**: 355-360.
- McKnight, A. T., McGuire, J. J., Elliott, N. J., Williams, B. J., Iversen, L. L. (1991). Pharmacological specificity of novel, synthetic, cyclic peptides as antagonists at tachykinin receptors. *Regul. Peptides.* **22**: 355-360.
- McLean,S., Ganong, A. H., Seeger, T. F., Bryce, D. K., Pratt, K. G., Reynolds, L. S., Siok. C. J., Lowe, J. A., Heym, J., (1991). Activity and distribution of binding sites in brain of a non-peptide substance P (NK1) receptor agonist. *Science.* **251**:437-439.
- Mendelson, S. C., Quinn, J. P., (1995a). Characterisation of potential regulatory elements within the rat preprotachykinin-A promoter. *Neurosci. Letts.* **184**: 125-128.
- Mick, G., Maeno, H., Kiyama, H., Tohyama, M., (1994). Marginal topography of neurons expressing the substance P receptor in the rat suprachiasmatic nucleus. *Mol. Brain Res.* **21**:157-161.
- Milner, T. A., Pickel, V. M., Abate, C., Joh, T. H., Reis, D. J., (1988). Ultrastructural charatcterisation of substance P-like immunoreactive neurons in rostral ventrolateral medullas in relation to neurons containing catecholamine-synthesizing enzymes. *J. Comp. Neurol.* **270**: 427-445.
- Monnier, D., Colas, J., Rosay, P., Henm, R., Borelli, E., Maroteaux, L., (1992). NKD, a developmentally regulated tachykinin receptor in *Drosophila. J. Biol. Chem.* **262**(2): 1298.
- Morris, B. J., (1992). Neuropeptide Y gene expression. *Methods in Neuroscience.* **9**:151-165.
- Mulderry, P. K., Chapman, K. E., Lyons, V., Harmar, A. J., (1993). 5'-flanking sequences from the rat PPT-A gene direct high-level expression of a reporter gene in adult rat sensory neurons transfected in culture by microinjection. *Mol. and Cellr. Neurosci.* **4**: 164-172.

- Murphy, T. J., Alexander, R. W., Greindling, K. K., Bernstein, K. E., (1991). Isolation of a cDNA encoding the vascular type-1angiotensin II receptor. *Nature.* **351**: 233.
- Murre, C., McCaw, P. S., Vaessin, H., Jan, Y. N., Cabrera, C. V., Buskin, J. N., Haushka, S. D., Lassar, A. B., Weintraub, H., Baltimore, D., (1989). Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell.* **58**: 537-544.

N

- Nakamuta, M., Takayanagi, R., Sakai, S., Sakamoto, Y., Hagiwari, H., Mizuno, T., Saito, Y., Natawa, H., (1991). Cloning and sequence analysis of a cDNA encoding human non-selective type of endothelin receptor. *Biochem. Biophys. Res. Commun.* **177**(1): 34.
- Nakanishi, S., Inone, Y., Kita, K., Nakamura, M., Chang, A. C. Y., Cohen, S. N., Numa, SA., (1979). Nucleotide sequence for cloned cDNA for bovine corticotropin-B-lipotropin precursor. *Nature.* **278**:423-27.
- Nakanishi, S., (1986). Structure and regulation of the preprotachykinin gene. *TINS.* Jan.1986. **9**: 41-44.
- Nakanishi, S., (1987). Tachykinin neuropeptides. *Physiol. Rev.* **67**: 1117-1142.
- Nakanishi, S., (1993). The rising sun shines on substance P. *TIPS.* **14**: 81-82.
- Nakata, Y., Kusaka, Y., Segawa, T., Yajima, S., Kitagawa, K., (1978). Substance P: Regional distribution and specific binding to synaptic membranes in rabbit central nervous system. *Life Sci.* **22**: 259-268.
- Nakata, Y., Tanaka, H., Morishima, Y., Segawa, T., (1988). Solubilisation and characterisation of substance P binding protein from bovine brainstem. *J. Neurochem.* **50**: 522-527.
- Nathanson, N. M., (1987). Molecular properties of the muscarinic acetylcholine receptor. *Annu. Rev. Neurosci.* **10**: 195- 236.
- Nawa, H., Hirose, T., Takashima, H., Inayama, S., Nakanishi, S., (1983). Nucleotide sequence of cloned cDNAs for two bovine brain substance P precursors. *Nature.* **306**:32-36.

- Nawa, H., Kotani, H., Nakanishi, S., (1984). Tissue-specific generation of two preprotachykinin mRNAs from one gene by alternative RNA splicing. *Nature* **312**:729-735.
- Nawa, H., Sah, D. W. H., (1990). Different biological activities in conditioned media control the expression of a variety of neuopeptides in cultured sympathetic neurons. *Neuron.* **4**: 279-287.

O

- O'Dowd, B. F., Lefkowitz, R. J.,Caron, M. J., (1989). Structure of the adrenergic and related receptors. *Ann. Rev. Neurosci.* **12**:67-83.
- Ogata, N., Abe, H., (1981). Further support for the postsynaptic action of substance P and its blockade with baclofen in neurons of the guinea-pig hypothalamus *in vitro*. *Experientia (Basel).* **37**: 759-761.
- Ogawa, Y., Nakao, K., Arai, H., Nakagawa, H., Suga, S., Nakanishi, S., Imura, H., (1991). Molecular cloning of a non-iso-peptide selective human endothelin receptor. *Biochem. Biophys. Res. Commun.* **178**(1): 248

P

- Papir-Kricheli, D., Frey, J, Laufer, R., Gilon, C., Chorev, M., Selinger, Z, Devor, M., (1987). Behavioural effects of receptor-specific substance P agonists. *Pain.* **31**: 263-276.
- Paterson, J. M., Mendelson, S. C., Morrison, C. F., McAllister, J., Dobson, S. P., Grace, C., Quinn, J. P., (1995). Three immediate early gene response elements in the proximal preprotachykinin-A promoter in two functionally distinct domains. *Neuroscience.* **66**: No. 4: 921-931.
- Paterson, J.M., Morrison, C. F., Mendelson, S. C., McAllister, J., Quinn, J. P., (1995). An upstream stimulatory factor (USF) binding motif is critical for rat preprotachykinin-A promoter activity in PC12 cells. *Biochem. J.* **310**: 401-4406.
- Payan, D. G., Brewster, D.R., Goetzl, E. J., (1983). Specific stimulation of T-lymphocytes by substance P. *J. Immunol.* 131:1613.

- Payan, D. G., Levine, J. D., Goetzl, E. J., (1984). Modulation of immunity and hypersensetivity by sensory neuropeptides. *J. Immunology.* **123** No. 4:1601-04.
- Pernow, B., (1953). Studies on substance P. Purification, occurence and biological actions. *Acta Physiol. Scand.* **29**: 1-90.
- Pernow, B., (1983). Substance P. *Pharmacological Reviews.* **35** No. 2:85-141.
- Perry, E. K., Tomlinson, B. E., Blessed, G., Gibson, P. H., Perry, R. H., (1978). *Br. Med. J.* **2**:1457-1459.
- Phelps, C. F., (1980). in 'Glycosylation'. Published by Freedman and Hawkins. 1980 p105.
- Phyllis, J. W., Limacher, J. J., (1974). Substance P excitation of cerebral cortical Betz cells. Brain Res. **69**: 158-163.
- Picciolo, P., Ruberti, F., Biocca, S., Di Luzio, A., Werge, T., Bradbury, A., Cattaneo, A., (1991). Neuroantibodies: Molecular cloning of a monoclonal antibody against substance P for expression in the central nervous system. *Proc. Natl. Acad. Sci.* **88**: 5611-5615.
- Powers, J. M., Schlaefer, W. W., Willingham, M. C., Hall, B. J., (1981). An immunoperoxidase study of senile cerebral amyloidosis with pathogenetic considerations. *J. Neuropathol. Exp. Neurol.* **40**:592-612.
- Pradier, L., Menager, J., Le Guern, J., Bock, M. D., Heuillet, E., Fardin, V., Garret, C., Doble, A., Mayaux, J. F., (1993). Septide: an agonist for the NK 1 receptor acting at a site distinct from SP. *Mol. Pharmacol.* **45**:287-293.

Q

- Quinn, J.P., Holbrook, N., Levens, D., (1987). Binding of a cellular protein to the gibbon ape leukemia virus enhancer. *Mol. Cellular. Biol.* **21**: 2735-2744.

R

- Regoli, K., Dion, S., Rhalen, N. E., Rouissi, N., Tousignant, C., Jukic, D., D'Orleans-Juste, P., Drapeau, G., (1989). Selective antagonists for receptors for substance P and related tachykinins. *Biopolymers*, **28**: 81-90.

- Reynolds, S. E., and Truman, J., (1983). in *Endocrinology of Insects*, pp217-233, Alan R. Liss (Ed).
- Rimland, J., Xin, W., Sweetnam, P., Saijoh, K., Nestler, E. J., Dumn, R. S., (1991). Sequence and expression of a neuro-peptide Y receptor cDNA. *Mol. Pharmacol.* **8**:23.
- Rodbell, M., Birnbaumer, L., Pohl, S., Krans, H. M., (1971). The glucagon-sensitive adenyl cyclase system in plasma membranes of rat liver. *J. Biol. Chem.* **246**: 1877-1882.
- Roskilde, M. M., Cahir, M., Gether, U., Shwartz, T. W., (1994). Mutations along transmembrane segment II of the NK-1 receptor affect substance P competition with non-peptide antagonists but not substance P binding. *J. Biol. Chem.* **269**: 28160-28164.
- Russell, S., Bjorkroth, U., Xu, J. C., Folkers, K., (1983). The pharmacological profile of substance P antagonist. Evidence for the existence of subpopulations of substance P receptors. *Acta. Physiol. Scand.* **117**: 445-449.

S

- Saito, K., Konishi, S., Otsuka, M., (1975). Antagonism between lioresal (baclofen) and substance P in rat spinal cord. *Brain Res.* **97**: 177-180.
- Saito, Y., Mizuno, T., Itakura, M., Suzuki, Y., Ito, T., Hirose, S., (1991). Primary structure of bovine endothelin ETb receptor and identification of signal peptidase and metal proteinase cleavage sites. *J. Biol. Chem.* **266**(34): 23433.
- Sakamoto, A., Yangisawa, M., Sakurai, T., Masaki, T., (1991). Cloning and functional expression of human cDNA for the ETb endothelin receptor. *Biochem. Biophys. Res. Commun.* **178**(2):656.
- Sakurai, T., Yanagisawa, M., Takuwa, Y., Miyazaki, H., Masaki, T., (1990). Cloning of cDNA encoding a non-iso-peptide selective subtype of the endothelin receptor. *Nature.* **348**: 732.
- Sambrook, J., Fritsch, E. F., Maniatis, T., (1989). *Molecular Cloning: A Laboratory Manual.* 2nd Edn. Cold Spring Harbor Press.
- Sanger, F., Nicklen, S., Coulson, S.R., (1975). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**:5463-5467.

- Sasai, Y., Nakanishi, S., (1989). Molecular characteisation o the rat substance K receptor and its mRNA. *Biochem. Biophys. Res. Comm.* **165**, No. 2: 695-702.
- Scatton, B., Javoy-Agid, F., Rouquier, L., Dubois, B., Agid, Y., (1983). Reduction of cortical dopamine, noradrenaline, serotonin, and their metabolites in Parkinson's disease. *Brain Res.* **275**: 321-328.
- Selkoe, D. J., (1991). The molecular pathology of Alzheimers disease. *Neuron.* **6**:487-498.
- Sharkey, K. A., Williams, R. G., Dockray, G. J., (1984). Sensory substance P innervation of the stomach and pancreas: demonstration of capsacin-sensetive neurons in the rat by combined immunohistochemistry and retrograde tracing. *Gastroenterology.* **87**: 914-921.
- Shedl, A., Montilou, L., Kelsey, G., Shcutz, G., (1993). A yeast artificial chromosome covering the tyrosinase gene confers copy number-dependent expression in transgenic mice. *Nature.* **362**: 258-261.
- Shigemoto, R., Yokota, Y., Tsuchida, K., Nakanishi, S., (1990). Cloning and expression of a rat neurokinin K receptor cDNA. *J. Biol. Chem.* **265**: No. 3: 623-628.
- Short, J. M., Fernandez, J. M., Sorge, J. A., Huse, W. D., (1988). Lambda zap: a bacteriophage lambda expression vector with *in vivo* excision properties. *Nucl. Acids Res.* **16**: 7583-7600.
- Sidman, R. L., Cowen, J. S., (1979). Inherited nerve and muscle diseases in mice: a tabulation with commentary. *Ann. N. Y. Acad. Sci.* **317**: 497-505.
- Sivam, S. P., Krause, J. E., (1992). Tachykinin systems in the spinal cord and basal ganglia: influence of neonatal capsaicin treatment of dopaminergic intervention on levels of peptides, substance P-encoding mRNAs and SP receptor mRNA. *J. Neurochem.* **59**: 2278-2284.
- Siwicki, K. K., Beltz, B. S., Kravitz, E. A., (1987).Proctolin in identified serotonergic, dopaminergic and cholinergic neurons in the lobster, *Homarus americanus*. *J. Neuroscience.* **7**: 522-531.
- Skarnes, W. C., Auerbach, B. A., Joyner, A. L., (1992). A gene-trap approach in mouse embryonic stem cells; the *lacZ* reporter gene is

activated by splicing reflects endogenous gene expression and is mutagenic in mice. *Genes Dev.* **6**: 903-918.

- Smith, T., Cuzner, M. L., (1994). Neuroendocrine-immune interactions in homeostasis and autoimmunity. *Neuropath. and App. Neurobiol.* **20**: 413-422.
- Smithies, O., Gregg, R. G., Boggs, S.S., Koralewski, M. A., (1985). Insertion of DNA sequences into the human chromosomal B-globin locus by homologous recombination. *Nature.* **317**: 230-234.
- Snider, R. M., Constantine, C. W., Lowe, J. A., Longo, K. P., Lebel, W. S., Woody, H. A., Drozda, Z. E., Desai, M. C., Vinick, F. J., Spencer, R. W., (1991). A potent neuropetide antaonist of the substance P (NK-1) receptor.*Science.* **251**: 435-439.
- Stoessl, A. J., (1994). Localization of striatal and nigral tachkinin receptors in the rat. *Brain Res.* **646**:13-18.
- Straub, R. E., Frech, G. C., Joho, R. H., Gershengorn, M. C., (1990). Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-releasing hormone receptor. *Proc. Natl. Acad. Sci, USA.* **87**: 9514. receptor
- Strauss, M., Dausman, J., Beard, C., Johnson, C., Lawrence, J. B., Jaenisch, R., (1993). Germ line transmission of a yeast artificial chromosome spanning the murine α1 collagen locus. *Science.* **259**: 1904 1907.
- Struble R. G., Powers, R. E., Casanova, M. F., Kitt, C. A., Price, D. L., (1987). Neuropeptidergic systems in plaques of Alzheimer's disease. *J. Neuropathol. Exp. Neurol.* **46**: 567-584.
- Sundelin, L. B.,Provvedini, D. M., Wahlestedt, C. R., Laurell, H., Pohl , J. S., Peterson, P. A., (1993). Molecular cloning of the murine substance K and substance P receptor genes. *Eur. J. Biochem.* **203**:625-631.

T

- Takahashi, T., Otsuka, M., (1975). Regional distribution of substance P in spinal cordand nerve roots of the cat and the effect of dorsal root section. *Brain Res.* **87**: 1-11.
- Takahashi, K., Tanaka, A., Hara, M., Nakanishi, S., (1991). The primary structure and gene organisation of the human substance P and neuromedin K receptors. *Eur. J. Biochem.* **204**: 1025-1033.

- Tanaka, K., Masu, M., Nakanishi, S., (1990). Structure and functional expression of the cloned rat neurotensin receptor. *Neuron.* **4**: 847.
- Takeda, T., Chou, K. B., Takecda, J., Sachais, B. S., Krause, J. E., (1991). Molecular cloning, structural characterization and functional expression of the human substance P receptor. *Biochem. Bophys. Res. Comm.* **179**: 1232-1240.
- Thomas, K.R., Capecchi, M.R., (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell.* **51**: 503-512.
- Thomas, K. R., Deng, C., Capecchi, M., (1992). High-fidelity gene targetingin embryonic stem cells by using sequence replacement vectors. *Molecular and Cellular Biology.* **12**, No. 7: 2919-2923.
- Too, H. P., Hanley, M. R., (1988). Solubilisation and characterisation of substance P-binding sites from chick brain membranes. *Biochem. J.* **252**: 545-551.
- Tota, M. R., Strader, C. D., (1990). Characterisation of the binding domain of the b-adrenergic receptor with fluorescent antagonist carazolol: evidence for a buried ligand binding site. *J. Biol. Chem.* **265**: 16891.
- Tsuchida, K., Shigemoto, R., Yokoya, Y., Nakanishi, S., (1990). Tissue distribution and quantitiation of the mRNAs for three rat tachykinin receptors. *Eur. J. Biochem.* **193**: 751-757.
- Tublitz, N., Brink, D., Broadie, K. S., Loi, P. K., Sylwester, A. W., (1991). From behaviour to molecules: an updated approach to the study of neuropeptides. TINS. **14**, No. 6:254-259.

V

- Valancius, V., Smithies, O., (1991). Testing an "In-Out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. *Mol. Cellr. Biol.* **11** (No. 3): 1402-1408.
- Vaulont, S., Daines, S., Evans, M., (1995). Disruption of the adenosine deaminase (ADA) gene using a dicistronic promoterless construct: production of an ADA-deficient homozygous ES cell line. *Transgenic Research.* **4**: 247-255.
- Vigna, S. R., Bowden, J. J., McDonald, D. M., Fisher, J., Okamoto, A., McVey, D. C., Payan, D. G., Bunnett, N. W., (1994). Characterization

of antibodies to the rat substabce P (NK 1) receptor and to a chimeric SP receptor expressed in mammalian cells. *J. Neurosci.* **14** No. 2:834-845.

- Voorn, P., Docter, G. J., Jongen-Relo, A. L., Jonker, A. J., (1994). Rostocaudal subregional differences in the response of enkephalin, dynorphin and substance P synthesis in rat nucleus accumbens to dopamine depletion. *Eur. J. Neurosci.* **6**:486-496.

W

- Wada, E., Way, J., Shapira, H., Kusano, K., Coy, D., Jensen, R., Battey, J., (1991). cDNA cloning, characterisation, and brain region-specific expression of neuromedin-B-preferring bombesin receptor. *Neuron.* **6**:421.
- Walker, P., Ni, L., Riley, L., Jonakait, G., Hart, R., (1991). Serotonin innervation affects SP biosynthesis in rat neostriatum. *Ann. NY Acad, Sci.* **632**: 485-487.
- Ward, P., Ewan, G. B., Jordan, C.C., Ireland, S. J., Hagan, R. M., Brown, J. R., (1990). Potent and highly selective neurokinin antagonists. *J. Med. Chem.* **33**: 1848-1851.
- Warden, M. K., Young, W. S., (1984). Distribution of cells containing mRNAs encoding substance P and neurikinin B in the rat central nervous system. *J. Comp. Neurol.* **272**: 90-113.
- Warden, M. K., Young, W. S., (1988). Distribution of cells containing mRNAs encoding substance P and neurokinin-B in the rat central nervous system. *J. Comp. Neurol.* **272**: 90-113.
- Watson, S. P., (1987). Multiple receptors for substance P and other mammalian tachykinins. ISI Atlas of Science: *Pharmacology.* 1987 p82-85.
- Watson, S., Arkinstall, S., (1994). G-Protein linked receptors: Factbook. Academic Press. pp261-270.
- Webber, R. H., deFelice, R., (1970). Bone marrow response to stimulation of the sympathetic trunk in rats. *Acta. Anat.* **77**:92.
- Whitty, C.J., Walker, P. D., Goebel, D. J., Poosch, M. S., Bannon, M. J., (1995). Quantitation, cellular localization and regulation of neurokinin receptor gene expression within the rat substantia nigra. *Neuroscience.* **64**: No. 2: 419-425.

- Willis, W. D., Coggeshall, R.D., (1991). Sensory mechanisms of the spinal cord. Plenum Press, New York: 196-201.
- Wold, F., (1981). In vivo chemical modification of proteins. *Ann. Rev. Biochem.* **50**:783-814.
- Wood, J. N., Bevan, S. J., Coote, P. R., Dunn, P. M., Harmar, A., Hogan, P., Latchman, T. S., Morrison, C., Rougon, G., Thevaniau, M., Wheatley, S., (1990). Novel cell lines display properties of nociceptive sensory neurons. *Proc. R. Soc. Lond.* **241**: 187-194.

X

- Xiao-Jiang, L., Wolfgang, W., Yan-na, W., North, R. A., Forte, M., (1991). Cloning, heterologous expression and developmental regulation of a *Drosophila* receptor for tachykinin-like peptides. *EMBO. J.*, **10**: No. 11: 3221-3229.

Y

- Yamamoto, Y., Lagercrantz, H., Euler, C. V., (1981). Effects of substance P and TRH on ventilation and pattern of breathing in newborn rabbits. *Acta. Physiol. Scand.* **113**: 541-543.
- Yasuhara, T., Nakajima, T., Falconie Erspamer, G., (1981). New tachykinins Glu, Pro-kassinin and hylambatin in the skin of the African rhacophorid frog *Hylambates maculatus. Biomed. Res.* **22**: 613-617.
- Yokota, Y.,Sasai, Y., Tanaka, K., Fujiwara,T., Tsuchida, K., Shigemoto, R., Kakizuka, A., Ohkubo, H., Nakanishi, S., (1989). Molecular characterisation of a functional cDNA for rat substance P receptor. *J. Biol. Chem.* 264 No. **30**:17649-52.
- Young, W. S., Bonner, T. I., Brann, M. R., (1986). Mesencephalic dopamine neurons regulate the expression of neuropeptide mRNAs in the rat forebrain. *Proc. Natl. Acad. Sci. USA.* **83**: 9827-9831.

Z

- Zhang, L., Harlan, R. E., (1994). Ontogeny of the distribution of tachykinins in rat cerebral cortex: immunocytochemistry and *in situ* hybridisation histochemistry. *Dev. Brain Res.* **77**: 23-36.