Isolation, characterisation, and chromosomal localisation of the mouse and human vasoactive intestinal peptide receptor type 2 genes

by

Melanie E. H. Mackay

A thesis submitted for the degree of Doctor of Philosophy.

The University of Edinburgh,

1997.

Abstract

Vasoactive intestinal peptide (VIP) and pituitary adenylate cyclase activating polypeptide (PACAP) have been implicated in a wide range of functions in the central and peripheral nervous systems, including smooth muscle relaxation, the promotion of neuronal survival, modulation of the immune system and the control of embryonic growth. The VIP type 2 (VIP_2) receptor is a seven-transmembrane G-protein coupled receptor belonging to the secretin/glucagon receptor subfamily, and is one of two such receptors cloned to date that bind both VIP and PACAP with high affinity. Studies of the distribution of the VIP_2 receptor have provided some clues as to the possible functions of this receptor, and in particular have led to the proposal that this receptor may be largely responsible for mediating the neuroendocrine effects of VIP. However the absence of well characterised specific antagonists for the VIP_2 receptor has proved to be a major obstacle to the delineation of the physiological role of this receptor. The work presented in this thesis has focused on the isolation and characterisation of mouse and human genomic clones encoding the VIP_2 receptor, which should allow molecular genetic approaches to be used in future studies of the functions of this receptor.

As a result of hybridisation screening of a mouse genomic λ2001 library and a human genomic P1 library, 6 bacteriophage clones that together span the entire coding region of the mouse VIP_2 receptor gene (*Vipr2*), and a single P1 clone that was shown to contain at least part of human VIP_2 receptor gene (*VIPR2*), were isolated. Characterisation of the intron/exon structure of the mouse *Vipr2* gene, achieved mainly by direct sequencing of λ DNA and PCR amplification of introns, revealed that the gene contained at least 12 introns, with sizes that range from 66 bp to at least 12 kb, and spans a minimum of 50 kb in total. Following the isolation of a 6 kb subclone which included the first exon and 5' sequence from the *Vipr2* gene, 3 kb of the putative promoter region of the gene was sequenced. The proximal 5' region of the *Vipr2* gene was found to be GC rich and CpG rich. No TATA box sequences were apparent within the region immediately upstream of the proposed transcription start site, but a computer-based search for possible transcription factor binding sites identified potential Sp1 and AP2 sites. Other potential regulatory sites that were found within the 3 kb sequence included possible binding sites for islet-1-like factors, the pituitary factor Pit-1, cAMP responsive factors, serum response factor, interferon stimulated gene factor-2, STAT factors, and factors that bind to the conserved lymphokine element 0.

The chromosomal localisation of the VIP_2 receptor gene in human and mouse was determined through collaborative work, in which the *VIPR2* P1 clone and a 4 kb

subclone from the mouse gene were used as probes in fluorescence *in situ* hybridisation experiments. The mouse gene was mapped to the telomeric region of mouse chromosome 12 (12F2), and the human gene was mapped to 7q36.3. These results were interesting from two points of view: i) the genes localise to regions within the mouse and human genomes, that had not previously been recognised as being syntenic, and may therefore define a new region synteny, ii) the localisation of the human gene to 7q36.3 placed it within the previously defined minimal critical region for the brain developmental disorder holoprosencephaly type 3 (HPE3). However, subsequent, more detailed mapping of the position of the gene relative to the chromosomal deletions/rearrangements found in patients with HPE3, demonstrated that although one copy of the *VIPR2* gene is deleted in some patients who have HPE3 and may contribute to the phenotype observed in these cases, it is not the gene responsible for HPE3.

To Mum and Dad

with love.

Research is the process of going up alleys to see if they are blind.

—Marston Bates, 1967.

Acknowledgements

I would like to thank: Professor George Fink, for the opportunity to study at the Brain Metabolism Unit; the Medical Research Council, for funding these studies; and all of the members of the unit, at both George Square and the Royal Edinburgh Hospital, for making my time in the unit so memorable.

Special thanks are due to my supervisor, Professor Tony Harmar, for his excellent scientific advice, encouragement, incredible patience, and in particular for his apparently unfailing belief that one day this thesis would be finished.

I am also extremely grateful to, Dr Eve Lutz, who assisted in the supervision of much of the work presented here, provided constant encouragement, and helped me to keep a sense of humour during the more difficult stages of the work; and Dr Janice Patterson, who was involved in the supervision of the early stages of this work, and has been an invaluable source of advice, and a friend, throughout my studies.

A very big thank you goes to Katrine West, who worked with me on a large proportion of the studies that are reported here, taught me most of what I know about automated sequencing, and did her very best to accommodate what were at times rather strange working hours.

I would also like to thank my fellow students, past and present, for making life more interesting, if slightly less sober than it might otherwise have been. Its been fun. I am especially grateful to Pauline, my long-suffering friend and flatmate, who has provided me with continuous supplies of chocolate, brilliant technical advice on how to set mouse traps, and has been an intrinsic part of so many of the good times that I've had in Edinburgh.

Last, but definitely not least, my sincere thanks to Norma Brearley for her expert assistance in the final formatting of this thesis, and generous help with numerous faxes, forms, and parcels, over the years.

I declare that this thesis is my own composition and that the studies presented here are the results of my own work, with the exception of:

i) Screening and sequencing of clones from the λ2001 mouse genomic library (Chapters 3 & 6) which was carried out with assistance from Mrs. Katrine West.

ii) The chromosomal mapping of the *VIPR2* gene in mouse and human (Chapter 5) which was carried out in collaboration with research groups at the MRC Human Genetics Unit, Edinburgh, and the Hospital for Sick Children, Toronto, Canada. Chromosomal fluorescence *in situ* hybridisation studies of the human gene were carried out by Dr. Judy Fantes, and the chromosomal localisation of the mouse gene was carried out by Mrs. Muriel Gray. Physical mapping of the location of the human *VIPR2* gene relative to the 7q telomere was carried out in collaboration with Dr S. Scherer.

This work has not been, and is not concurrently being, submitted for any other degree or professional qualification.

Some of the results presented in this thesis have been published in:

Mackay, M., Fantes, J., Scherer, S., Boyle, S., West, K., Tsui, L.-C., Belloni, E., Lutz, E., Van Heyningen, V., Harmar, A. J. Chromosomal localisation in mouse and human of the vasoactive intestinal peptide receptor type 2 gene: a possible contributor to the holoprosencephaly 3 phenotype. Genomics 37: 345-353 (1996).

List of Abbreviations

A	adenine or adenosine
amp	ampicillin
AMV	avian myeloblastosis virus
ATP	adenosine triphosphate
bp	base pair(s)
BrdU	bromodeoxyuridine
BSA	bovine serum albumin
Bt_2cAMP	dibutyryl cyclic adenosine monophosphate
C	cytosine or cytidine
$CaCl_2$	calcium chloride
cAMP	cyclic adenosine monophosphate
CCD	charged-couple device
cDNA	complementary DNA
Ci	curie
CNS	central nervous system
CpG	cytidylyl (3'→5') guanosine
cpm	counts per minute
CsCl	caesium chloride
dCTP	deoxycytosine triphosphate
DAPI	4', 6 -diamidino-2-phenylindole (DAPI)
ddH_2O	deionised distilled water
DEAE	diethylaminoethyl
del	deletion
DEPC	diethyl pyrocarbonate
der	derivative
DMSO	dimethyl sulphoxide

DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
dup	duplication
dUTP	deoxyuridine triphosphate
E	embryonic day
ES-cell	embryonic stem cell
EDTA	ethylene diaminetetracetate
EtBr	ethidium bromide
EtOH	ethanol
FITC	fluorescein isothiocyanate
g	gravitational acceleration, 9.8 ms^{-2}
G	guanine or guanosine
G-banding	Giemsa-banding
GTP	guanosine triphosphate
GTPβS	guanosine 5'-o-(2 thiotriphosphate)
GTPγS	guanosine 5'-o-(3 thiotriphosphate)
HCl	hydrochloric acid
ICRF	Imperial Cancer Research Fund
IP_3	inositol triphosphate
IPTG	β-D-isopropyl-thiogalactopyranoside
$K_{act} = EC_{50}$	concentration required to produce half-maximal biological effect
kan	kanamycin
kb	kilobase(s), kilobase pairs
KCl	potassium chloride
lod	logarithm to the base 10 of the odds
L1	long interspersed nuclear element-1

M	moles/litre
Mb	megabase(s), megabase pairs
-mer	oligomer
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
MRC	Medical Research Council
mRNA	messenger RNA
N	gram equivalents/litre
NaCl	sodium chloride
NaOAc	sodium acetate
NaOH	sodium hydroxide
OD	optical density
ORF	open reading frame
P1	bacteriophage P1
^{32}P	radioisotope of phosphorus, emitting high energy β particles
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming units
PHA	phytohaemagglutinin
PIPES	piperazine-*N, N'*-bis(ethanesulfonic acid)
PMA	phorbol myristate acetate
PNS	peripheral nervous system
POU	pit-oct-unc
RACE	rapid amplification of cDNA ends
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute

RT	reverse transcriptase
SDS	sodium dodecyl sulphate
SSC	standard saline citrate
SSPE	salt-sodium phosphate-EDTA
T	thymine or thymidine
TAE	Tris-acetate-EDTA
TBE	Tris-borate -EDTA
TE	Tris-EDTA
ter	terminus
Tris	tris (hydroxymethyl) amino methane
tRNA	transfer RNA
UTR	untranslated region
UV	ultraviolet
-ve	negative
v/v	volume:volume ratio
X-gal	5-bromo-4-chloro-3-indolyl-β-galactopyranoside
YAC	yeast artificial chromosome

Nomenclature

The nomenclature used for the receptors of the secretin/glucagon family and the genes encoding these receptors, has been subject to several changes during the course of the studies reported in this thesis, and further changes are likely in the future. The situation with regard to receptors that have high affinity binding sites for vasoactive intestinal peptide and/or pituitary adenylate cyclase activating polypeptide is particularly confusing, and in the case of the VIP_2 receptor at least three different names for this receptor are currently in use in the published literature. However, as at present there is little or no consensus as to what the "correct" system for naming these receptors should be, the nomenclature that has been used in this thesis is as follows:

Receptors

i) The first high affinity VIP receptor to be isolated (Ishihara *et al.*, 1992) is referred to as the VIP type 1 receptor in this thesis, and is abbreviated to VIP_1 receptor. However, despite its name it should be noted that this receptor binds both VIP and PACAP with high and approximately equal affinities. This receptor is also sometimes referred to as the PACAP/VIP type 2 receptor, PVR2 (Rawlings and Hezareh, 1996).

ii) The second high affinity VIP receptor to be isolated (Lutz *et al.*, 1993) is referred to as the VIP type 2 receptor in this thesis, and is abbreviated to VIP_2 receptor. Like the VIP_1 receptor, this receptor binds both VIP and PACAP with high and approximately equal affinity. This receptor has also been called the PACAP/VIP type 3 receptor, PVR3 (Rawlings and Hezareh, 1996), and the "helodermin-preferring" VIP receptor (Robberecht *et al.*, 1988; Svoboda *et al.*, 1994). The mouse homologue of the receptor is sometimes referred to as PACAPR3 (Inagaki *et al.*, 1994)

iii) The only high affinity "PACAP-specific" receptor that has been isolated to date (reviewed by Arimura and Shioda, 1995) is referred to as the PACAP type I receptor in this thesis. This receptor, which is also called the PACAP/VIP type 1 receptor, PVR1 (Rawlings and Hezareh, 1996), does bind VIP, but with an affinity that is greatly reduced relative to that for PACAP.

Receptor Genes

i) The genes encoding the human VIP_1 receptor, VIP_2 receptor and PACAP type I receptor, are referred to as the *VIPR1* gene, *VIPR2* gene, and *PACAPR1* gene respectively. Lower case letters are used to designate the symbols for the corresponding mouse genes, which are: the *Vipr1* gene, the *Vipr2* gene and the *Pacapr1* gene.

ii) It is recognised that gene the word "gene" is not strictly necessary when the convention of italicising gene symbols has been followed, and that similarly if the distinction between human genes and those of other species had been made through the differential use of upper and lower case gene symbols, the word "human" is often superfluous when used immediately prior to the gene symbol. However, these additional words have frequently been used in this thesis, in an attempt to make its findings comprehensible to pharmacologists as well as geneticists and molecular biologists.

iii) Genes other than those for VIP, PACAP or their receptors are generally described using the full names of the protein products that they encode, and are not italicised. Exceptions to this include: cases where the gene name gives little clue to the nature of the gene product (e.g. the *goosecoid* gene); and situations where the same gene is mentioned several times.

cDNAs

Due to the current problems surrounding the receptor nomenclature, the gene symbols described above have also been used when describing the cDNAs, although in order to emphasise that the cDNA is not a direct product of the gene, the symbols are not italicised when used in this context.

Contents

Chapter 3: Isolation and characterisation of the mouse *Vipr2* gene

Chapter 5: Chromosomal localisation of the *VIPR2* gene in mouse and human

Chapter 6: Isolation and sequence analysis of the putative promoter region of the mouse *Vipr2* gene

Appendix 1

List of figures and tables

Chapter 1
General introduction

1.1 Vasoactive Intestinal Peptide (VIP)

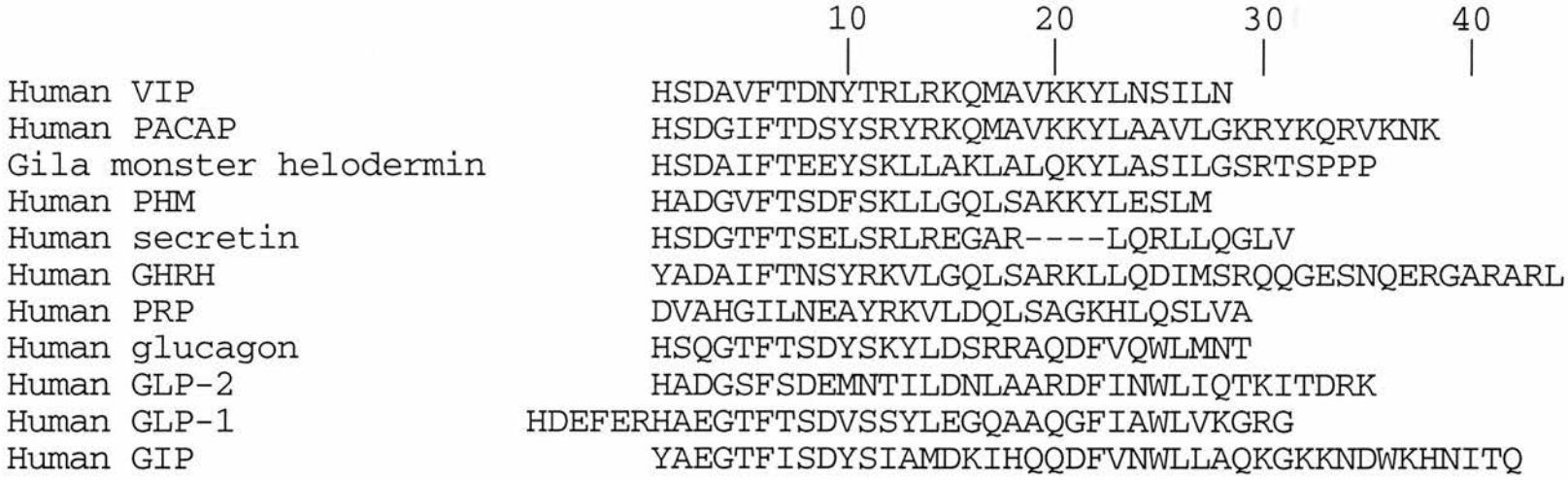
1.1.1 VIP: evolution, structure and distribution

VIP is a 28 amino acid carboxy-amidated peptide that was originally isolated from porcine gut (Said & Mutt, 1970), and has subsequently been shown to belong to a family of structurally related peptides (see Figure 1.1) that includes: secretin, peptide histidine isoleucine (PHI), peptide histidine methionine (PHM), helodermin, pituitary adenylate cyclase activating polypeptide (PACAP), gastric inhibitory peptide (GIP), glucagon, glucagon-like peptides (GLP) and growth hormone releasing hormone (GHRH). This family of peptides, which is often referred to as the secretin/glucagon family (Mutt, 1988), can be divided into two main subgroups, one including secretin, VIP, PHI, PACAP, and GHRH, and the other including glucagon, the glucagon like peptides, and probably also GIP (Campbell & Scanes, 1992). In evolutionary terms, it has been suggested that the secretin/glucagon peptide family has expanded through a cascade of gene duplication events (frequently coupled with exon duplications or losses) and mutations, which began with the duplication of a single ancestral gene (Campbell & Scanes, 1992). To date, VIP has been isolated from mammals (Carlquist *et al.*, 1982; Carlquist *et al.*, 1979; Dimaline *et al.*, 1984; Eng *et al.*, 1986), birds (Nilsson, 1975), reptiles (Wang & Conlon, 1993), amphibians (Chartrel *et al.*, 1995) and fish (Wang & Conlon, 1995). However, the detection of PACAP-like immunoreactivity in the *Drosophila* central nervous system (Zhong & Pena, 1995), and the finding that VIP-like and PHI-like proteins exist in nematodes (Foster & Lee, 1996), have strongly suggested that peptides of the secretin/glucagon family are also present in invertebrate species, and very recently two cDNAs (each encoding PACAP and another member of the peptide family) were isolated from a protochordate (McRory & Sherwood, 1997).

All peptides of the secretin/glucagon family are synthesised in the form of precursor proteins, which are subsequently cleaved to release the mature peptides. However, in vertebrates, the presence of more than one bioactive peptide within a precursor protein is a feature that is common to only three members of this family: the glucagon precursor, which when cleaved releases glucagon and several other peptides including the glucagon-like peptides (GLP-1 and GLP-2) (Bataille *et al.*, 1996; Lund *et al.*,

Figure 1.1 Alignment of the amino acid sequences of members of the secretin/glucagon family of peptides

Standard single letter abbreviations are used for amino acids (IUPAC-IUB Commission on Biochemical Nomenclature):
A=Ala, C=Cys, D=Asp, E=Glu, F=Phe, G=Gly, H=His, I=Ile, K=Lys, L=Leu, M=Met, N=Asn, P=Pro, Q=Gln, R=Arg, S=Ser, T=Thr, V=Val, W=Trp, Y=Tyr and (–) = gap/deletion.


10
20
30
40
Human VIP HSDAVFTDNYTRLRKQMAVKKYLNSILN
Human PACAP HSDGIFTDSYSRYRKQMAVKKYLAAVLGKRYKQRVKNK
Gila monster helodermin HSDAIFTEEYSKLLAKLALQKYLASILGSRTSPPP
Human PHM HADGVFTSDFSKLLGQLSAKKYLESLM
Human secretin HSDGTFTSELSRLREGAR----LQRLLQGLV
Human GHRH YADAIFTNSYRKVLGQLSARKLLQDIMSRQQGESNQERGARARL
Human PRP DVAHGILNEAYRKVLDQLSAGKHLQSLVA
Human glucagon HSQGTFTSDYSKYLDSRRAQDFVQWLMNT
Human GLP-2 HADGSFSDEMNTILDNLAARDFINWLIQTKITDRK
Human GLP-1 HDEFERHAEGTFTSDVSSYLEGQAAQGFIAWLVKGRG
Human GIP YAEGTFISDYSIAMDKIHQQDFVNWLLAQKGKKNDWKHNITQ

1982); the PACAP precursor, in which PACAP and PACAP-related polypeptide (PRP) are present (Ohkubo *et al.*, 1992); and the VIP precursor, which contains both VIP and PHI/PHM (Itoh *et al.*, 1983). In the case of the peptides released from the glucagon precursor protein, distinct tissue-specific roles for some of these peptides are starting to emerge (Bataille *et al.*, 1996), but at present the biological importance of PRP and PHI/PHM is not clear. No specific receptors for PHI/PHM have been identified to date, and the actions of these peptides (which appear to be mediated by VIP receptors) generally seem to mirror those of VIP, although they are often less potent (Huang & Rorstad, 1990; Jensen *et al.*, 1981; Robberecht *et al.*, 1982; Suzuki *et al.*, 1984).

The amino acid sequence of VIP has been strongly conserved throughout evolution, and is particularly well conserved in mammals where, with the exception of guinea pig (Eng *et al.*, 1986) and opossum (Eng *et al.*, 1992), its sequence is identical in all mammalian species from which it has been isolated. However, despite the rapid accumulation of information regarding the primary structure of VIP and other members of the secretin glucagon family, the characteristics of the 3-dimensional structures of these peptides (a critical factor in the understanding of ligand/receptor specificity), have emerged much more slowly and are still not entirely clear. In general though, it has been shown that members of the secretin/glucagon family adopt a largely α-helical conformation, particularly towards their carboxyl-terminal (C-terminal) regions (Gronenborn *et al.*, 1987; Sasaki *et al.*, 1975). NMR and computer-based analyses of VIP indicate that it adopts a bent N-terminal conformation (Filizola *et al.*, 1997), followed by a C-terminal region consisting of two α-helical stretches (spanning amino acids 7-15 and 19-27) that are linked by a more flexible region of undefined structure (Theriault *et al.*, 1991). The amino acids in the N-terminal region of VIP appear to be critical for the peptide's biological activity (Fournier *et al.*, 1984), whereas the C-terminal region in general and the second of the two helices (amino acids 19-27) in particular, have been strongly implicated in mediating peptide/receptor binding interactions and specificities (Gourlet *et al.*, 1996b; Wulff *et al.*, 1997), a process which may involve interactions between potential 'coiled coil' motifs found in the C-terminal regions of the secretin/glucagon peptide family and in the N-terminal domains of their receptors (Momany & Bowers, 1996).

Following its isolation from gut, VIP has been found to be widely distributed throughout mammalian central and peripheral nervous systems. Within the central nervous system (CNS) the highest concentrations of VIP are found in the cerebral cortex, the hypothalamus (especially in the suprachiasmatic nucleus), and the central amygdaloid nucleus (Lorén *et al.*, 1979), but VIP is also present at significant levels in

many other regions of the CNS, particularly in the brainstem (Eiden *et al.*, 1982), hippocampus (Roberts *et al.*, 1984; Samson *et al.*, 1979), and the dorsal horn and lumbosacral regions of the spinal cord (Anand *et al.*, 1983; Gibson *et al.*, 1984). Peripheral regions that are innervated by VIP-containing fibres include the smooth muscle and glands of the gastrointestinal tract (Polak *et al.*, 1974), the respiratory tract (Laitinen *et al.*, 1985; Lundberg *et al.*, 1984), and the genitourinary tract (Helm *et al.*, 1981; Larsen *et al.*, 1981). In the cardiovascular system VIP-immunoreactivity is found mainly in the nerves supplying arteries, especially the mesenteric and uterine arteries (Della *et al.*, 1983; Weihe *et al.*, 1984). Specific peripheral organs and tissues in which VIP has been detected include the intestine (Said & Mutt, 1970), lungs (Said, 1988), heart (Forsgren, 1989), kidney (Barajas *et al.*, 1983), eye (Lorén *et al.*, 1980), skin (Hartschuh *et al.*, 1984), and bone (Hohmann *et al.*, 1986). In the endocrine system VIP is found in the pancreas, (where heavy VIP innervation is seen around the islets; Bishop *et al.*, 1980), the adrenal cortex and medulla (Ehrhart-Bornstein *et al.*, 1991; Holzwarth, 1984), the ovaries (Gozes & Tsafriri, 1986), the vas deferens and prostate (Larsen *et al.*, 1981), the thymus (in which VIP mRNA has been detected; Gomariz *et al.*, 1993), the thyroid gland (Ahren *et al.*, 1980b), the pineal gland (Uddman *et al.*, 1980), the anterior pituitary (Lam, 1991), the median eminence (Ceccatelli *et al.*, 1988), and hypophysial portal blood (Shimatsu *et al.*, 1981). The presence of VIP-containing nerves within lymphoid tissues of the immune system is also well documented (Fink & Weihe, 1988; Weihe *et al.*, 1991). However, although it has been recognised for some time that certain non-neuronal cells seem to be able to synthesise VIP, e.g. Merkel cells in the skin (Hartschuh *et al.*, 1984), amacrine cells of the retina (Okamoto *et al.*, 1992), and eosinophils (Aliakbari *et al.*, 1987), it was only quite recently that it was discovered that VIP is also produced by B and T lymphocytes (Gomariz *et al.*, 1994; Leceta *et al.*, 1996).

At the transcriptional level, expression of the VIP gene can be induced by cyclic-AMP (cAMP) and by phorbol esters, which activate factors that bind to a cAMP/phorbol ester responsive element in the VIP gene's proximal promoter region (Yamagami *et al.*, 1988), and a more distal phorbol ester-specific responsive element (Hahm & Eiden, 1996). Leukemia inhibitory factor (LIF) and ciliary neurotrophic factor (CNTF) can stimulate VIP expression through a cytokine response element (CyRE) in the gene's promoter (Symes *et al.*, 1994), and a glucocorticoid receptor binding site which is involved in the pituitary modulation of VIP gene transcription has also been identified(Chew *et al.*, 1997). Other stimuli that have been shown to affect VIP levels in certain cells and tissues, include: steroids (e.g. oestrogen; Hammond *et al.*, 1997; Kasper *et al.*, 1992), prolactin (which exerts a negative regulatory effect in pituitary

Hammond *et al.*, 1997), retinoic acid (which increases VIP mRNA levels in neuroblastoma cells; Georg *et al.*, 1994), ascorbic acid (Brick *et al.*, 1985), and depolarising agents (Sun *et al.*, 1992). However, the modes by which these latter factors exert their effects on VIP levels are not entirely clear.

1.1.2 Actions of VIP

Although the role initially ascribed to VIP was that of a gastrointestinal hormone, it soon became apparent that this peptide also acts as a neurotransmitter and neuromodulator in both the central and peripheral nervous systems (Bryant *et al.*, 1976; Fahrenkrug, 1979; Larsson *et al.*, 1976). Early studies of the function of VIP led to the identification of a wide range of roles for this neuropeptide, including: neuronal excitation in regions such as the cerebral cortex, hippocampus, and spinal cord (Dodd *et al.*, 1979; Jeftinija *et al.*, 1982; Pawelzik *et al.*, 1992; Phillis *et al.*, 1978); smooth muscle relaxation, including vasodilatation and bronchodilatation within the cardiovascular and respiratory systems respectively (Altiere & Diamond, 1984; Bitar & Makhlouf, 1982; Morice *et al.*, 1983; Pompa *et al.*, 1990); the promotion of water and ion secretion (Krejs *et al.*, 1980; Nathanson *et al.*, 1983); and the regulation of metabolism (particularly glycogenolysis) in the CNS (Magistretti, 1990; Sorg & Magistretti, 1991) and liver (Feliu *et al.*, 1983).

VIP also has important actions on the endocrine system, where it is involved in the stimulation of steroidogenesis (Davoren & Hsueh, 1985; Kasson *et al.*, 1986; Tornell *et al.*, 1988), and the promotion of the release of several endocrine hormones. Within the pituitary VIP has been shown to increase oxytocin and vasopressin release from the neurohypophysis (Ottesen *et al.*, 1984), whereas its effects on the anterior pituitary include enhancement of interleukin-6 production (Spangelo *et al.*, 1990), the stimulation of prolactin secretion during lactation (Abe *et al.*, 1985; Escalada *et al.*, 1996; Kato *et al.*, 1978), and the potentiation of luteinising hormone release (Hammond *et al.*, 1993), responses which together with the high concentrations of VIP present in the hypophysial-portal blood, have led to the suggestion that VIP may act as a hypophysiotrophic hormone. Nevertheless, although VIP undoubtedly plays a critical role in the regulation of pituitary functions, the actions of VIP within the hypothalamic-pituitary axis by no means represent the limit of the effects of VIP on the endocrine system, as this peptide has also been found to be involved in many other endocrine pathways, for example, the promotion of catecholamine secretion from the adrenal medulla (Cheung & Holzwarth, 1986) and aldosterone and corticosterone secretion from the adrenal cortex (Cunningham & Holzwarth, 1988; Hinson *et al.*, 1996), stimulation of thyroid hormone secretion (Ahren *et al.*, 1980a), the enhancement of

serotonin N-acetyltransferase activity and melatonin release from the pineal gland (Simonneaux *et al.*, 1990; Yuwiler, 1983), and the stimulation of glucose-induced insulin secretion from pancreatic islets (Kato *et al.*, 1994; Schebalin *et al.*, 1977).

VIP is an important modulator of the immune system, and is involved in a diverse array of functions ranging from the regulation of cytokine levels (Ganea, 1996) to the modulation of adherence, chemotaxis, and phagocytosis (de la Fuente *et al.*, 1994; Litwin *et al.*, 1992). With regard to the regulation of cytokine expression, VIP has been shown to act at both the transcriptional and post-transcriptional levels to inhibit the expression of interleukin (IL)-2 in activated T-cells, and inhibits IL-4 production at the post-transcriptional level, at least partly as a result of IL-2 inhibition (Wang *et al.*, 1996). Inhibition of IL-10 by VIP in activated T-cells, has been shown to occur at the transcriptional level (Martinez *et al.*, 1996), and it also appears that VIP can act indirectly (through effects on macrophage-like cells) to suppress the proliferation of IL-7 responsive B-cells. In addition to these effects, VIP plays a role in the regulation of B-cell antibody production and isotype switching (Kimata & Fujimoto, 1995), and the ability of VIP to rescue thymocytes from glucorticoid-induced apoptosis suggests a possible function for this peptide in the regulation of thymocyte development (Ernström *et al.*, 1995).

The role of VIP in the regulation of cell growth and survival has been extensively studied, and represents an area in which the number of actions ascribed to VIP appears to be constantly expanding. Neverthless, it is in many ways one of the most poorly understood aspects of VIP's functions, as these effects of VIP appear to be highly dependent upon the system being examined, and it is clear that VIP acts as a modulator (Muller *et al.*, 1995), as opposed to a predefined promoter or inhibitor of growth, differentiation, and survival. The effects of VIP on growth and survival, can generally be divided into 3 main categories:

i) the neurotrophic and neuroprotective effects that are mediated by the actions of VIP on astroglial cells, a process in which VIP stimulation (at concentrations as low as 0.9 nM) leads to the release of several survival promoting factors including IL-1 (Brenneman *et al.*, 1996) and activity dependent neurotrophic factor (ADNF, a peptide that has neuroprotective effects even when present at femtomolar concentrations; Brenneman & Gozes, 1996)

ii) the actions of VIP on neuroblasts, neuroblastoma cells and other transformed cell lines, where it appears that VIP can either promote or inhibit cell proliferation depending on the particular cell line or type being studied (reviewed by Muller *et al.*,

1995), a characteristic that might conceivably reflect differences or alterations in the VIP effector systems in these cells (similar to the switch in the expression of PACAP and VIP receptors that has been reported to occur during phenotypic interconversion in human SH-IN neuroblastoma cells; Lelièvre *et al.*, 1996).

iii) the dramatic effects of VIP on embryonic growth and development, which were first reported by Gressens *et al.*, who showed that administration of VIP (10^{-7}-10^{-9}M) to whole cultured mouse embryos at embryonic day (E) 9.5, resulted in a dramatic increase in embryonic growth and (at 10^{-7} M) a 5-6-fold increase in cell mitosis (Gressens *et al.*, 1993).

Other growth-regulatory actions of VIP that do not fit into the above categories include: VIP-mediated suppression of apoptosis in the ovary (an effect that is partly antagonised by insulin-like growth factor-binding protein 3; Flaws *et al.*, 1995), and the recently reported involvement of VIP in stimulating the expression and secretion of human intestinal trefoil factor (hITF; Ogata & Podolsky, 1997), which has been found in the gut and in PC12 (pheochromocytoma) cells (Probst *et al.*, 1997), and is a member of a large family of growth factor-like peptides (De *et al.*, 1994).

Finally, although it would not be appropriate to try to list all the proposed functions of VIP in this space, the role of VIP in the regulation of circadian rhythms should be mentioned, as it is currently attracting a lot of attention in the scientific literature. VIP is present in both the pineal gland and in the suprachiasmatic nucleus (SCN), where in rodents, levels of the peptide display a diurnal variation (Kaku *et al.*, 1986; Okamoto *et al.*, 1991). Immunohistochemical and retrograde tracing studies support the idea that SCN derived VIP is involved in transmitting information to the pineal (Teclemariam-Mesbah *et al.*, 1997). The recent demonstration that antisense antagonism of VIP leads to the temporary abolition of circadian patterns of corticosterone secretion (but not stress related secretion) in rats (Scarbrough *et al.*, 1996), is particularly interesting, but it is not clear whether such findings can be extrapolated to humans as it has been reported that in the human SCN VIP levels do not display significant diurnal variation (Hofman *et al.*, 1996).

1.1.3 VIP receptors: evidence for VIP receptor heterogeneity

Many, but not all, of the effects of VIP result from the ability of this peptide to stimulate increases in intracellular cAMP levels through interactions with receptors on the cell surface, a process in which the involvement of G-proteins is indicated by sensitivity to GTP and GTP analogues (Christophe *et al.*, 1986; Quick *et al.*, 1978). Early studies of

VIP binding sites, carried out using a combination of ligand binding and biochemical/immunological approaches, clearly indicated that considerable heterogeneity existed amongst VIP receptors, both between tissues, and between different species. However, the nature/cause of this apparent heterogeneity was less clear, particularly in studies based solely on the determination of molecular weights of the binding sites by SDS-polyacrylamide gel electrophoresis following ligand-receptor crosslinking, since, as pointed out by Robberecht *et al.*, 1990, it is difficult to distinguish between variations due to differences in the receptor's intrinsic amino acid structure, and those that might be due to altered levels of glycosylation or proteolysis. In comparison, ligand-binding approaches produced results that were somewhat easier to interpret, and it soon became apparent that several tissues, including rat kidney (Magistretti *et al.*, 1988), human normal and cancerous gastric mucosa (Gespach *et al.*, 1988), and mammalian lung (Dickinson *et al.*, 1986; Schaffer *et al.*, 1987), contain two classes of receptor for VIP, comprising a high and a lower affinity binding site. Other lines of evidence which suggested that multiple types of VIP receptors exist, included the observation that in rat brain two distinct populations of VIP binding site can be distinguished on the basis of their sensitivity to GTP (which promotes the dissociation of ligands from G-protein-coupled receptors). The finding that in some regions of the brain, VIP binding is much more sensitive to GTP than in other regions, seems to indicate that two different types of VIP binding site exist, only one of which couples to a GTP-sensitive G-protein, are present in brain. Indeed, in some areas such as the olfactory bulb, the medial geniculate thalamic nucleus, the ventrolaterial thalamic nucleus, and the pituitary, the vast majority of VIP binding was GTP-insensitive. However at present, the basis of this differential sensitivity to GTP remains unclear. Finally, it should also be noted that several reports on VIP-receptor interations have described a VIP receptor, which binds the lizard-venom derived peptide helodermin with a higher affinity than it binds VIP. This receptor, which is found on the human SUP-T1 lymphoma cell line and some other neoplastic cell lines, is coupled to adenylate cyclase, but displays a distinctly different binding profile for VIP and related peptides, which results in the following order of potency: helodermin > PACAP-38 > PACAP-27 ≥ VIP (Christophe *et al.*, 1993; Robberecht *et al.*, 1996). In contrast to the situation with the GTP-insensitive binding sites, the identity of the helodermin-preferring VIP receptors is now clear and is discussed further in section 1.4.2.

1.2 Pituitary adenylate cyclase activating polypeptide (PACAP)

Despite the fact that PACAP was not identified until 19 years after the isolation of VIP, characterisation of this peptide and its actions has progressed extremely rapidly, and has

been greatly assisted by the fact that (in contrast to the situation with VIP, where there were gaps of about 10 years between isolation of the peptide, the cloning of the VIP gene and the cloning of VIP receptor cDNAs) the cloning of the PACAP gene (Ohkubo *et al.*, 1992), and the cloning of cDNAs encoding receptors for PACAP (Ishihara *et al.*, 1992; Lutz *et al.*, 1993; Pisegna & Wank, 1993), have followed the isolation of this peptide in fairly rapid succession .

1.2.1 PACAP, structure and distribution

PACAP exists in two post-translationally processed forms: a 38 amino acid amidated peptide (PACAP-38) from sheep hypothalamus, that was initially isolated during a search for novel peptides that could stimulate pituitary adenylate cyclase (Miyata *et al.*, 1989), and a C-terminally shortened form of this peptide, PACAP-27, which was subsequently isolated from side fractions obtained during PACAP-38 purification (Miyata *et al.*, 1990). PACAP-27 and the corresponding N-terminal region of PACAP-38 display 68% homology to the amino acid sequence of VIP, and were immediately recognised as members of the secretin/glucagon peptide family. Structural studies of PACAP indicate that it has conformation that is generally very similar to that of VIP and other members of the secretin/glucagon subfamily (Wray *et al.*, 1993; see 1.1.1). However it has been suggested that the distinct biological activities of VIP and PACAP-27 may be largely the consequence of a switch from Ala to Gly at amino acid 4 (Coy *et al.*, 1996; see also Figure 1.1), whereas in the case of PACAP-38, experiments involving the generation of chimeric peptides have indicated that the C-terminal extension 28-38 contains sequences that are important for binding affinity with regard to PACAP-specific binding sites (but not those shared with VIP), and may help to stabilise the peptide's structure (Gourlet *et al.*, 1996a).

In mammals, PACAP has a relatively wide distribution within the CNS, where the highest levels of this peptide are found in the hypothalamus (particularly the supraoptic nucleus and paraventricular nucleus), and the median eminence (Christophe, 1993), whereas lower levels can be detected in regions such as the hippocampus, frontal cortex and cerebellum (Vertongen *et al.*, 1992). Endocrine organs in which PACAP or PACAP transcripts have been detected include the posterior lobe of the pituitary gland and to a lesser extent the intermediated lobe (Kimura *et al.*, 1994); the pancreas (Arimura & Shioda, 1995); and the rat seminiferous tubules, in which PACAP mRNA is expressed in a stage specific manner (Kononen *et al.*, 1994). In the rat immune system PACAP is found to be present in most lymphoid tissues, in which PACAP immunoreactivity appears to be located mainly in lymphocyte-like cells (Gaytan *et al.*, 1994).

1.2.2 Actions of PACAP

The presence of PACAP in the hypothalamus and median eminence, together with its demonstrated actions in increasing intracellular cAMP in cultured pituitary cells, strongly suggest that this peptide might have hypophysiotrophic actions. PACAP has been shown to stimulate gonadotrophin release from perifused rat pituitary cells (Tsujii *et al.*, 1994), and has also been implicated in the promotion of pro-opiomelanocortin peptide secretion from the AtT-20 corticotroph cell line (Braas *et al.*, 1994), and stimulation of IL-6 secretion from cultured rat astrocytes (Gottschall *et al.*, 1994). Other reported functions of PACAP include the stimulation of catecholamine secretion from cells of the adrenal medulla (Watanabe *et al.*, 1995), the promotion of glucose-induced insulin secretion from the pancreatic islets (Bertrand *et al.*, 1996), and a possible involvement in the regulation of the early stages of spermatogenesis (Kononen *et al.*, 1994). Early studies of PACAP binding sites indicated that like VIP, the effects of PACAP were mediated by G-protein coupled receptors, and also suggested that in some tissues, particularly lung and liver, VIP and PACAP might share binding sites, a factor that was to explain at least to some extent, the partially overlapping actions of these peptides on some cells and tissues (Shivers *et al.*, 1991).

1.3 The secretin/glucagon receptor subfamily

The rat secretin receptor was first cloned by Ishihara *et al.*, in 1991, from an expression library containing cDNA transcripts derived from the NG108-15 (neuroblastoma/glioma hybrid) cell line, and from initial analysis of its cDNA and the deduced amino acid sequence, was predicted to have a seven-transmembrane structure that was very similar to that of the rhodopsin-like 7-TM G-protein coupled receptors. When expressed in COS cells the cloned receptor bound iodinated secretin, and interacted with the alpha subunit of G_S to activate adenylate cyclase. However, at the amino acid level, the secretin receptor was found to show little or no homology to the sequences of previously characterised members of the 7-TM G-protein coupled receptor superfamily (which had generally been shown to display at least ~20-30% homology to one another), and it was therefore suggested that the secretin receptor represented a new type of 7-TM G-protein coupled receptor (Ishihara *et al.*, 1991).

As is often the case when clones encoding novel types of proteins are isolated, the cloning of the rat secretin receptor sparked a search for related clones, which in this instance were thought likely to encode other peptide receptors. Two main approaches have been used in attempts to identify receptors that are related to the secretin receptor: i) expression-based cloning methods, which at least initially were based on the theory

(now substantiated) that peptides that are structurally related to secretin, i.e. the secretin/glucagon family, would be strong candidates as ligands for these receptors (Jelinek *et al.*, 1993; Thorens, 1992); and ii) homology-based cloning involving either low stringency hybridisation (often across species) of receptor cDNA clones to cDNA and genomic libraries (Gremlich *et al.*, 1995; Sreedharan *et al.*, 1993), or a PCR-based strategy using degenerate primers designed to bind to regions of the cDNA sequence that (following the isolation of other receptors related to the secretin receptor) have been found to be highly conserved in this G-protein coupled receptor subfamily (Lutz *et al.*, 1993; Mayo, 1992; Svoboda *et al.*, 1994). As predicted, a considerable number of clones encoding receptors with sequences related to that of the secretin receptor have now been isolated, including clones encoding receptors for the glucagon/secretin peptide family members: GHRH (Gaylinn *et al.*, 1993; Mayo, 1992), glucagon (Jelinek *et al.*, 1993; Lok *et al.*, 1994; MacNeil *et al.*, 1994), GLP-1 (Dillon *et al.*, 1993; Graziano *et al.*, 1993; Thorens, 1992), VIP (Ishihara *et al.*, 1992; Lutz *et al.*, 1993; Sreedharan *et al.*, 1993; Svoboda *et al.*, 1994), PACAP (Hashimoto *et al.*, 1993; Hosoya *et al.*, 1993; Morrow *et al.*, 1993; Ogi *et al.*, 1993; Pisegna & Wank, 1993; Spengler *et al.*, 1993; see also VIP receptor references above), and GIP (Gremlich *et al.*, 1995; Usdin *et al.*, 1993; Volz *et al.*, 1995; Yamada *et al.*, 1995b). However, somewhat surprisingly, very soon after the cloning of the rat secretin receptor, it emerged that the calcitonin receptor (Lin *et al.*, 1991) and the receptor for parathyroid/parathyroid hormone related peptide (PTH/PTHrP; Jüppner *et al.*, 1991) have sequences that are similar to that of the rat secretin receptor (although the amino acid sequences of these peptides do not display any significant homology to those of the secretin/glucagon peptide family), and it has subsequently been shown that in addition to the calcitonin and PTH/PTHrP receptors, the calcitonin gene related peptide (CGRP) receptor, and corticotrophin releasing factor (CRF) receptor, also belong to the secretin/glucagon 7-transmembrane G-protein coupled receptor subfamily (Aiyar *et al.*, 1996; Chen *et al.*, 1993; Kapas & Clark, 1995; Lovenberg *et al.*, 1995).

Other secretin-receptor-related clones that have been isolated include a rather unusual group of molecules that have a seven transmembrane region very similar to those found in members of the secretin/glucagon receptor subfamily, but also possess a greatly extended N-terminal domain region that contains numerous epidermal growth factor-like modules. So far, two clones encoding this type of molecule have been isolated from mammals: one of which encodes an EGF module-containing mucin-like hormone receptor 1 (EMR1; Baud *et al.*, 1995), whereas the other was found to encode an activation-induced antigen (CD97) from leukocytes (Hamann *et al.*, 1995). However, as other unusual properties of these molecules have gradually emerged, for example,

the requirement for intracellular processing of CD97 from its pro-protein form into a noncovalently associated two-subunit structure (Gray *et al.*, 1996), the perceived relationship between these molecules and the glucagon /secretin receptor subfamily has become more distant, and these molecules are now generally considered to form a new receptor subgroup that is closely related but distinct from the secretin/glucagon receptor subfamily.

All members of the secretin/glucagon receptor subfamily that have been cloned so far have been found to be capable of stimulating adenylate cyclase through coupling to G_S, and many of them appear to be able to activate additional pathways that result in increased intracellular calcium and inositol phosphate (reviewed by Segre & Goldring, 1993), although in most cases the details of the non-cAMP-mediated signalling pathways and the G-proteins involved in them, are not yet clear. Structural features that are shared by members of the secretin/glucagon receptor subfamily include the presence of a hydrophobic putative N-terminal signal sequence, a relatively large N-terminal extracellular domain containing at least 4 conserved cysteine residues (which are thought to be critical for the correct structure of the ligand binding domain), and several consensus sites for N-linked glycosylation (Laburthe *et al.*, 1996). Particularly high conservation of the amino acid sequences of the 3rd and 7th transmembrane domains has also been noted (Lutz *et al.*, 1993), and it has recently been reported that in comparison with the rhodopsin family, the predicted structure of secretin/glucagon subfamily receptors differs with respect to the relative positions of transmembrane helices 2, 3, and 7 (Donnelly, 1997).

1.4 Cloned receptors for VIP and/or PACAP

To date, three types of receptors that bind VIP and/or PACAP have been cloned, two of which bind both VIP and PACAP with approximately equal affinities, whereas the third type of receptor displays a distinct preference for PACAP and has a much lower affinity for VIP. Reports of the cloning of these three types of receptors were first published within an ~18 month period from April 1992- November 1993 (Ishihara *et al.*, 1992; Lutz *et al.*, 1993; Pisegna & Wank, 1993), since which time no further types of receptors for these peptides have been isolated, although pharmacological data relating to the apparent GTP-insensitivity of some VIP receptors (Hill *et al.*, 1992) has been used to argue that additional receptors for VIP exist (Gressens *et al.*, 1994; see Chapter 5 section 5.4.6 for further discussion of this issue).

1.4.1 The VIP_1 receptor

The VIP_1 receptor (the first of the VIP/PACAP receptors to be cloned) was initially isolated from a rat lung cDNA library by low stringency hybridisation using the rat secretin receptor cDNA as a probe (Ishihara *et al.*, 1992). The isolated clone which was found to encode a 459 amino acid protein (with a putative 30 amino acid N-terminal signal sequence), shares the predicted 7-transmembrane topology of the secretin receptor, and following expression of the clone in mouse COP cells the receptor was found to bind both VIP and PACAP with high affinity. In COSGs1 cells, this receptor, latter named the VIP_1 receptor, was found to stimulate increases in cAMP levels in response to PACAP-38, PACAP-27, and VIP (EC_{50} of VIP within this system = 0.57 nM), and also responded to helodermin, PHM , and secretin, although these peptides displayed significantly lower potencies than VIP or PACAP (Ishihara *et al.*, 1992). Studies of the VIP_1 receptor's distribution, carried out in parallel with the initial characterisation of the rat cDNA clone, indicated that transcripts encoding the VIP_1 receptor are found predominantly in the lung, liver, intestine, and are present at lower levels in the thymus. Relatively high levels of *Vipr1* transcripts were also detected in certain regions of the brain, especially the hippocampus, the mitral layer of the olfactory bulb, and the neurons of the cerebral cortex (Ishihara *et al.*, 1992). Later investigations (including those on the human homologue of this receptor) have largely confirmed this distribution pattern, whilst indicating that transcripts encoding the VIP_1 receptor are also present in other cells and tissues, such as the heart, vas deferens, some areas of the kidney, the walls of pancreatic blood vessels, the uterus, the adrenal medulla, Raji B-lymphoblasts and both stimulated and unstimulated T-cells (Delgado *et al.*, 1996b; Sreedharan *et al.*, 1993; Usdin *et al.*, 1994; Wei & Mojsov, 1996).

A cDNA encoding the human homologue of the rat VIP_1 receptor, was isolated by hybridisation screening of a human colonic adenocarcinoma HT29 cDNA library (again using a secretin receptor cDNA derived probe), and has been found to encode a protein that displayed 84% homology to the rat VIP_1 receptor at the amino acid level (Sreedharan *et al.*, 1993). When stably-transfected into chinese hamster ovary (CHO) cells or human embryonic kidney (HEK) cells the VIP_1 receptor was found to be able to couple to two distinct signalling pathways, one leading to the increase in intracellular cAMP that is characteristic of receptors of the secretin/glucagon subfamily, and the other leading to increases in intracellular calcium (Sreedharan *et al.*, 1994).

Actions of VIP and PACAP in which the VIP_1 receptor has been implicated include: the inhibition by VIP and PACAP of glucocorticoid-induced apoptosis of rat $CD4^+$ $CD8^+$

thymocytes (Delgado *et al.*, 1996a); inhibitory effects of VIP on the chemotaxis of the VIP_1 receptor expressing human T-cell line HuT 78, a process which in contrast is stimulated in cells expressing the VIP_2 receptor (Xia *et al.*, 1996a); and possibly, the modulation of the responsiveness of dorsal horn neurons to sensory stimuli (Dickinson *et al.*, 1997).

1.4.2 The VIP_2 receptor

The cDNA encoding the rat VIP_2 receptor was first isolated from rat pituitary and olfactory bulb by Lutz *et al.*, and contains an open reading frame (ORF) of 437 amino acids which includes a 22 amino acid hydrophobic signal sequence at the N-terminal end (Lutz *et al.*, 1993). Comparison of the predicted amino acid sequence of this receptor with those of other members of the secretin/glucagon receptor subfamily reveals that it is most closely related to the VIP_1 receptor and the PACAP type 1 receptor with which it shares ~50% homology at the amino acid level (Lutz *et al.*, 1993). However, the N- and C- terminal ends of these receptors display a high level of divergence, and although, like the VIP_1 receptor, the VIP_2 receptor binds both VIP and PACAP with high affinity, the order of potency of VIP, PACAP and related peptides for cAMP production in COS cells transfected with the rat Vipr2 cDNA is: VIP ~ PACAP-27 ~ PACAP-38 ~ helodermin > PHI >> rat GHRH (Lutz *et al.*, 1993). Thus, in contrast to the VIP_1 receptor where helodermin is less potent than VIP in stimulating the accumulation of cAMP (Ishihara *et al.*, 1992), VIP and helodermin appear to be approximately equipotent at the VIP_2 receptor. Other notable differences between the potencies of ligands at the VIP_1 and VIP_2 receptors, include: the complete insensitivity of the VIP_2 receptor to secretin (Usdin *et al.*, 1994), a peptide which at the VIP_1 receptor results in increases in cAMP albeit with a potency that is approximately 30-100 fold less than that of VIP (Ishihara *et al.*, 1992; Usdin *et al.*, 1994); and the observation that higher concentrations of VIP, PHI, and PACAP-38 are necessary for VIP_2 receptor activation than for activation of the VIP_1 receptor (Usdin *et al.*, 1994).

Studies of the tissue distribution of mRNA transcripts encoding the rat VIP_2 receptor (encoded by the *Vipr2* gene) have demonstrated that mRNA for this receptor is present at relatively high levels in the suprachiasmatic nucleus, thalamus, hippocampus, the internal granular layer of the olfactory bulb, and several hypothalamic nuclei (Lutz *et al.*, 1993; Sheward *et al.*, 1995; Usdin *et al.*, 1994). *Vipr2* transcripts are also found at moderate levels in many other areas of the CNS, particularly the midbrain and brainstem, and are present in several endocrine tissues, including: the pituitary (especially the anterior lobe, where the density of *Vipr2* transcripts rises dramatically during pregnancy), pancreatic islets, ducts and spermatocytes within the testes,

granulosa cells of the ovary, the pineal gland, the adrenal cortex, and the medulla and cortex of the thymus, a distribution that has led to the proposal that the VIP_2 receptor may mediate the neuroendocrine effects of VIP (Usdin *et al.*, 1994). Indeed, in many tissues, and particularly in the CNS (Usdin *et al.*, 1994), the distribution of the *Vipr2* transcripts appears to be complementary to that of the *Vipr1* transcripts , and it is interesting to note that in addition to its prevalence in neuroendocrine tissues, the VIP_2 receptor appears to be the only one of the cloned VIP or PACAP receptors (i. e. the VIP_1, VIP_2, and PACAP type I receptors) for which transcripts have been detected at significant levels in skeletal muscle (Wei & Mojsov, 1996), and is also the only VIP receptor that is present at significant levels in the stomach and duodenum (Usdin *et al.*, 1994).

Clones encoding the mouse and human homologues of the rat VIP_2 receptor have also been isolated (Inagaki *et al.*, 1994; Svoboda *et al.*, 1994), and in each instance the cellular origin of the cDNA library from which the clone was isolated has pointed to previously characterised functions of VIP in which the VIP_2 receptor has later been implicated.

The cDNA encoding the mouse VIP2 receptor (also known as PACAPR3) was initially isolated (using PCR-based methods) from a cDNA library that was derived from the mouse insulin-secreting β-cell line MIN6, and (in agreement with the pancreatic islet expression of rat *Vipr2* mRNA later described by Usdin *et al.*, 1994, see above) has been shown to be present in mouse pancreatic islets and several other insulin secreting cell lines (Inagaki *et al.*, 1994). Like the rat VIP_2 receptor, the mouse VIP_2 receptor is positively coupled to cAMP signalling pathways (EC_{50} of VIP = 3.2 nM in COSGs1 cells). However, following the expression of the Vipr2 clone in *Xenopus* oocytes, PACAP-27, PACAP-28, and VIP, at 0.1 nM, have been found to be capable of stimulating calcium-activated chloride currents, an observation which led to the still somewhat controversial suggestion that the mouse VIP_2 receptor can be coupled to phospholipase C (PLC) signalling pathways in some systems (Inagaki *et al.*, 1994). Nevertheless, although there have been no other reports of recombinant VIP_2 receptors coupling to PLC in cell lines, PACAP-38 has been shown to stimulate insulin secretion in the MIN6 cell line from which the mouse VIP_2 receptor was cloned (Inagaki *et al.*, 1994), and there is now substantial evidence to indicate that the VIP_2 receptor mediates the effects of VIP and PACAP on glucose-induced insulin secretion (as indicated by the approximately equipotent effects of VIP and PACAP, but the failure of secretin to stimulate insulin secretion; Bertrand *et al.*, 1996), which in addition to increases in

cAMP and intracellular calcium, involves a novel signalling pathway that is sensitive to the phosphatidylinositol 3-kinase inhibitor wortmannin (Straub & Sharp, 1996).

While the isolation of the mouse VIP_2 receptor prompted new studies on the actions of VIP and PACAP in the stimulation of insulin secretion, the isolation and characterisation of the human homologue of this receptor (Svoboda *et al.*, 1994) has indicated that other aspects of VIP's actions in which the VIP_2 receptor might play an important role, are the modulation of the immune system and the regulation of cell growth. Indeed, it seems that as suggested by Svoboda *et al.*, the human VIP_2 receptor, which was cloned (using a method based on low stringency PCR) from cDNA derived from the human $CD4^+$ T-cell lymphoma cell line SUP-T1, is in fact the receptor that was previously described as the 'helodermin-preferring VIP receptor (see section 1.2.3; Christophe *et al.*, 1993; Robberecht *et al.*, 1988), and thus by inference the human VIP_2 receptor is likely to be present on several other cancer cell lines including the human THP-1 monocyte/macrophage cell line, the human lung carcinoma small cell line NCJ-H345, and the B16 mouse melanoma cell line (reviewed by Christophe *et al.*, 1993; Robberecht *et al.*, 1990). Moreover, as it has been suggested that in some cases receptors with helodermin-preferring binding characteristics are revealed in malignant cell lines, but are virtually undetectable in the corresponding normal tissues (Robberecht *et al.*, 1990), it is interesting to note that in $CD4^+$ and $CD8^+$ T-lymphocytes, expression of transcripts encoding the VIP_2 receptor is only induced following stimulation through the T-cell receptor-associated CD3 molecule, whereas in cell lines representing immature thymocytes and T-helper 2 cells, transcripts encoding the VIP_2 receptor have been shown to be present at significant levels, and are the only VIP receptor transcripts detected (Delgado *et al.*, 1996b).

Other actions of VIP and PACAP, in which the VIP_2 receptor has been implicated to date include: i) the inhibitory effects of VIP (and to a lesser extent PACAP) on interleukin-2 (IL-2) production by $CD4^+$ T-lymphocytes (Tang *et al.*, 1996), in which a role for the VIP_2 receptor has been indicated by the ability of the recently characterised selective VIP_2 receptor agonist Ro 25-1553 (a synthetic VIP analogue; Gourlet *et al.*, 1997) to inhibit IL-2 production (Tang *et al.*, 1995); ii) the effects of PACAP in the regulation of testicular germ cell protein synthesis (West *et al.*, 1995) and its (proposed) role in early spermatogenesis (Kononen *et al.*, 1994), which are thought to be mediated solely by VIP_2 receptors (since transcripts encoding the VIP_1 and PACAP type I receptors cannot be detected by RT-PCR in the germ cells; Krempels *et al.*, 1995); and iii) the recently reported antiproliferative effect of VIP and PACAP on the T98G human

glioblastoma cell line, a cell line which appears to express *Vipr2* transcripts but not those encoding the VIP_1 or PACAP type I receptors (Vertongen *et al.*, 1996).

As the principal cytokine responsible for T-cell proliferation, IL-2 is an important target for the action of immunosuppressive drugs (reviewed by Cavalieri, 1996). Therefore elucidation of the pathways by which VIP causes inhibition of IL-2, and in particular the possible effects of the VIP/PACAP system on calcineurin (a calcium/calmodulin regulated protein phosphate implicated in pathways leading to IL-2 gene expression; Batiuk *et al.*, 1997) will be important, as it may suggest new ways in which IL-2 levels can be modulated. Similarly, the characterisation of the modes by which the VIP_2 receptor acts to regulate insulin secretion and the regulation of cell growth, could also potentially open new therapeutic possibilities for diabetes and cancer respectively. Nevertheless, although of great interest, it seems almost certain that the functions of the VIP_2 receptor that have been characterised to date, represent only a small fraction of the number of processes in which this receptor will eventually be shown to act. The generally low levels of *Vipr2* gene transcripts even in tissues where the transcripts are most abundant (Usdin *et al.*, 1994) is likely to have hindered, at least to some extent, attempts to correlate the distribution of this receptor with its possible functions, but as it is now becoming increasingly apparent that in some cells and tissues *Vipr2* expression is developmentally regulated (e.g. in the testes) and/or can be dramatically induced under certain conditions (e.g. upon T-cell activation or during lactation), critical clues to the identity of the stimuli responsible for increasing *Vipr2* expression levels might be gained from examination of the *Vipr2* gene's promoter region.

1.4.3 The PACAP type I receptor

The cDNA encoding the rat PACAP type I receptor was first isolated by Pisegna and Wank (1993), who used a rat Vipr1 cDNA probe to screen a cDNA library constructed from the rat pancreatic acinar carcinoma cell line AR4-2J. Characterisation of the isolated cDNA clone revealed that it contained an open reading frame that is predicted to encode a protein of 495 amino acids (including a 19 amino acid N-terminal signal sequence), and displays 50% identity to the rat VIP_1 receptor at the amino acid level (Pisegna & Wank, 1993). Following transfection of the Pacapr1 cDNA into COS-7 cells, comparative studies of possible ligands for this receptor demonstrated that PACAP-38 and PACAP-27 acted approximately equipotently in the stimulation of the accumulation of intracellular cAMP through this receptor (EC_{50} 0.6-1.6 nM), whereas VIP and secretin had little effect. Thus, in contrast to the characteristics of the VIP_1 and VIP_2 receptors, where VIP and PACAP have similar potencies to one another, the AR4-2J derived receptor displayed a very definite preference for PACAP, and in

recognition of this was named the PACAP type I receptor. Within weeks of the publication of the rat Pacapr1 cDNA sequence, four other laboratories also published the sequences of the PACAP type I receptor (Hashimoto *et al.*, 1993; Hosoya *et al.*, 1993; Morrow *et al.*, 1993; Spengler *et al.*, 1993), which in one case (Spengler *et al.*, 1993) included a report of the isolation of 5 splice variants of the receptor (with sequence alterations in the region encoding the receptor's third intracellular loop). In contrast to the VIP_1 and VIP_2 receptors, the PACAP type I receptor has been shown to be capable of stimulating both intracellular cAMP and inositol triphosphate (IP_3) accumulation when the receptor is stably transfected into CHO cells (Delporte *et al.*, 1995). However as splice variation in this receptor has been shown to be tissue-specifically and developmentally regulated (D'Agata *et al.*, 1996), the discovery that the expression of different variants of the PACAP type I receptor leads to changes in both ligand recognition and G-protein coupling (Spengler *et al.*, 1993), indicates that splice variation is likely to represent an additional level at which the end result(s) of this receptor's action can be modulated.

Studies of the distribution of the the PACAP type I receptor and its transcripts indicate that in the CNS this receptor is abundant in many regions, including: the gray matter of the spinal cord, several hypothalamic nuclei (particularly the supraoptic and paraventricular nuclei), thalamus, olfactory bulb, brainstem, cerebral cortex, cerebellum, hippocampus, and basal ganglia, whereas in the endocrine system transcripts encoding this receptor are prevalent in the anterior pituitary, the pineal gland, the adrenal medulla, pancreas, and testicular spermatids (see Arimura & Shioda, 1995; Christophe, 1993; Spengler *et al.*, 1993). Thus, it appears that each of the receptors for PACAP and/or VIP (i.e. the VIP_1, VIP_2 and PACAP type I receptors) has a different distribution (and probably also different functions) within the nervous system, although a considerable degree of overlap in these patterns exists in some cells and tissues.

Functions of VIP and PACAP in which the PACAP type I receptor has been implicated to date, include: the regulation of sympathetic neuron catecholamine secretion (Braas & May, 1996), and mediation in the hypothalamic tract of the phase-shifting effects that dark pulses during the daytime have on the circadian clock (Hannibal *et al.*, 1997).

1.5 Aims

The main aim of the work presented in this thesis was to isolate and characterise mouse genomic clones encoding the VIP_2 receptor, as in the absence of suitable antagonists that are specific for the VIP_2 receptor, the production of transgenic *Vipr2*-null mice is currently the most promising approach to the determination of the functions of this

receptor. However, as the process of isolating clones suitable for use in transgenic constructs, was likely to include the isolation of clones containing the 5' flanking sequences of the *Vipr2* gene, it was also hoped that by studying the promoter region, it would be possible to identify some of the factors that might be responsible for regulating the tissue-and cell-restricted expression patterns of this gene.

In view of the role of VIP in the regulation of cell growth and survival, the isolation of a human clone encoding the VIP_2 receptor, and the subsequent determination of the localisation of the gene in mouse and human, was also a major aim of this study, as this would greatly assist the identification of any previously mapped mouse or human genetic disorders in which this gene might be involved.

Chapter 2
General methodology

2.1 Isolation and purification of plasmid DNA

2.1.1 Plasmid minipreps

Standard SDS-alkaline lysis method

A single bacterial colony was used to inoculate 3 ml of L-broth (containing the appropriate antibiotic), and the culture was grown overnight in a 37°C incubator with gentle shaking. The bacterial cells were collected in an Eppendorf tube by centrifugation at 10,000*g* for 5 minutes in a microcentrifuge. The supernatant was removed, and the cells were then resuspended in 100 µl of alkaline lysis solution 1 (see 2.5.2). Alkaline lysis solution 2 (200 µl; see 2.5.2) was then added to the cell suspension, and the tube contents were mixed by rapidly inverting the tube. After the tubes had been incubated on ice for 5 minutes, to allow cell lysis to occur, 150 µl of ice cold alkaline lysis solution 3 (see 2.5.2) was added, and the reactions were mixed by flicking the tubes. A further 5 minute incubation on ice was followed by centrifugation at 10,000*g* for 7 minutes, and the supernatant (~400 µl) was then transferred to a fresh tube where it was mixed with an equal volume of phenol:chloroform, and centrifuged for 2 minutes at 10,000*g* in a microcentrifuge. The supernatant was carefully removed and transferred to another Eppendorf tube. Two volumes of ethanol were added to the supernatant to precipitate the DNA, and after a 2 minute incubation at room temperature, the precipitated DNA was collected by centrifugation at 10,000*g* for 5 minutes. The DNA pellet was then washed in 70% ethanol, dried at 37°C for 5-10 minutes, and was finally resuspended in 50 µl of TE (pH 8.0) containing 20 µg/ml RNase A (50 units/mg; Boehringer Mannheim).

SDS-alkaline lysis/PEG precipitation method

This method was used to isolate plasmids for automated fluorescent sequencing, prior to the availability of Wizard™ 373 DNA purification kits (Promega). The SDS-alkaline lysis section of the protocol is almost identical to that described above, with the exception of the addition of an RNase A step (final concentration 20 µg/ml, 20 minute incubation at 37°C). However, following ethanol precipitation of the plasmid DNA , the dried DNA pellet was resuspended in 33.6 µl of ddH_2O, and then precipitated using 6.4 µl of 5M NaCl and 40 µl of 13% PEG. The precipitation reaction was mixed

briefly and then left on ice for 20 minutes, after which the DNA was pelleted by centrifugation at ~12,000 rpm (10,000*g*) in a microcentrifuge for 15 minutes at room temperature. The supernatant was carefully removed using a glass Pasteur pipette, and the pellet was washed with 70% ethanol, air dried at room temperature, then dissolved in 20 µl of ddH_2O. To estimate the volume of miniprep DNA that would be needed for a sequencing reaction, a 2 µl aliquot of the plasmid DNA was run on an agarose minigel, alongside plasmid DNA of known concentration

Wizard™ 373 plus DNA purification system (Promega)

This commercial kit is specifically designed to be used to isolate plasmid miniprep DNA for use in automated sequencing reactions for the Applied Biosystems (ABI) 373 sequencer. Although the kit was not available at the start of this project, it replaced the modified SDS-alkaline lysis/PEG precipitation method during the later stages of this work, and was used in the preparation of template DNA from most of the plasmid subclones of the mouse *Vipr2* gene.

Plasmid DNA was prepared according to the manufacturer's instructions for the isolation of high copy number plasmids (although a vacuum manifold was not available), and the plasmid DNA was eluted from the Wizard column by the addition of 100 µl of ddH_2O (preheated to 65°C). The volume of plasmid DNA solution that was recovered following purification using a Microcon™-100 microconcentrator (Amicon) ranged from 4 -50 µl , but in all cases the final volume was made up to at least 20 µl with ddH_2O, and (after 2 µl had been checked on an agarose minigel) 0.5-5 µl of DNA was generally used in each sequencing reaction.

2.1.2 Plasmid maxipreps

Alkaline lysis followed by equilibrium ultracentrifugation in a caesium chloride-ethidium bromide gradient

A 2 ml overnight bacterial culture (grown from a single bacterial colony grown in L-broth + 50 µg/ml ampicillin) was used to inoculate 500 ml of media (L-broth + 50 µg/ml ampicillin), which was in turn grown overnight at 37°C in an orbital shaker. Plasmid DNA was then isolated according to Treisman's alkali lysis method for the large scale preparation of plasmid DNA (published in Sambrook *et al.*, 1989), and the DNA was subsequently purified by equilibrium centrifugation in a CsCl-ethidium bromide gradient.

The bacterial cells from the 500 ml culture were collected by centrifugation at 6000 rpm for 10 minutes in a Sorvall GSA rotor at 4°C, and the cell pellet was then resuspended in 18 ml of alkaline lysis solution 1. Following the addition of 40 ml of alkaline lysis solution 2, the contents of the centrifuge bottle were mixed by inversion and left at room temperature for 5-10 minutes. Ice cold alkaline lysis solution 3 (20 ml) was then added, and the bottle was shaken vigorously, then left on ice for 10 minutes. After centrifugation at 6000 rpm for 15 minutes in a Sorvall GSA rotor at 4°C, the rotor was stopped without braking, and the supernatant was filtered through 4 layers of cheesecloth, into a clean centrifuge bottle. Isopropanol (0.6 of a volume) was mixed with supernatant to precipitate the DNA, and following a 10 minute incubation at room temperature, the DNA was pelleted by centrifugation at 8000 rpm for 15 minutes at room temperature (Sorvall GSA rotor). The DNA pellet was washed with cold 70% ethanol, dried, and then resuspended in 3 ml of TE.

To prepare the DNA for equilibrium centrifugation in a CsCl-ethidium bromide gradient, the DNA solution was transferred to a 10 ml total volume of the DNA sample was made up to exactly 4 ml with TE. Ethidium bromide (400 µl of a 10 mg/ml stock solution) and CsCl (4.4 g) were then dissolved in the DNA solution, and the tube was left on ice for 5-10 minutes to allow protein to precipitate. After centrifugation at 2000 rpm for 5 minutes in an MSE Chilspin centrifuge (Fisons), a Pasteur pipette was used to carefully transfer the supernatant to a 6 ml polyallomer quick seal ultracentrifuge tube (Sorvall® Instruments). The tube was then topped up with balance solution (400 µl of ethidium bromide + 4.4 g of CsCl + 4 ml TE), sealed, and centrifuged (together with a balance tube if necessary) overnight at 45,000 rpm in a TV-1665 rotor (Sorvall ultracentrifuge) at 20°C. Plasmid DNA (the lower of the two bands) was removed using a 19-gauge needle and a 2 ml disposable syringe, and the ethidium bromide was removed from the solution by repeated extraction with an equal volume of isopropanol/H_2O/CsCl (1:1 isopropanol:H_2O, saturated with CsCl). The lower (aqueous) layer was then transferred to a 15 ml disposable plastic tube, where 0.3 ml of sodium acetate (pH 5.2) was added, and the total volume of the solution was made up to 3 ml with ddH_2O. Ethanol (6 ml) was then added, and the tube was left at -20°C for 1-2 hours to allow the DNA to precipitate. The DNA was collected by centrifugation at 2000 rpm for 10-15 minutes in a Fisons MSE Chilspin centrifuge, dried, and then dissolved in TE. The DNA was then reprecipitated, washed in 70% ethanol, and was finally resuspended in 0.5-1 ml of TE.

Qiagen plasmid maxi kit (Qiagen Ltd)

Qiagen plasmid maxi kits (Qiagen Ltd.) were generally used in situations where DNA from only 1 or 2 different plasmids was required. Plasmid DNA was prepared from a 250 ml overnight culture of bacteria, according to the manufacturer's instructions for the isolation of high copy number plasmids, and was found to be comparable in quality to that obtained following CsCl-ethidium bromide gradient ultracentrifugation of plasmid DNA.

2.2 Manipulation of purified DNA

2.2.1 Restriction enzyme digestion of DNA

Restriction enzymes used in these studies were purchased from Promega Ltd. and Boerhinger Mannheim Ltd.

Unless otherwise stated, restriction enzyme digests of plasmid DNA or PCR products, were generally set up in a total reaction volume of 20-30 µl containing: 0.5-1.5 µg of DNA, 5-10 units of restriction enzyme, and ddH_2O. The reaction components were mixed by vortexing, spun briefly in a microcentrifuge, then incubated at 37°C for 3-12 hours. If removal of the restriction enzyme(s) was necessary for subsequent manipulations, a Magic™/Wizard™ DNA clean-up kit (Promega) was used (according to the manufacturer's instructions) to purify the restriction fragments. Digested DNA that was not to be used immediately, was stored at -20°C until required.

2.2.2 Agarose gel electrophoresis of DNA

Agarose gels (0.8-1.5% w/v) were prepared using Ultrapure agarose (Gibco-BRL), and contained 1× Tris-borate EDTA (TBE), which was also used as the running buffer. However, in cases where a Geneclean II® kit (BIO 101 Inc.; see 2.2.3) was to be used to isolate DNA from the gel, Tris-acetate EDTA (TAE) was used in the gel and as the running buffer (final concentration 1×), since this avoided the need for the use of the Geneclean TBE modifier.

The agarose was dissolved in the electrophoresis buffer, by heating (4 minutes at medium power in a microwave oven), and once the solution had cooled to about 60°C, ethidium bromide (10 mg/ml stock solution) was added to a final concentration of 1 µg/ml. Agarose minigels (30 ml), were set up and run using Horizon® 58 minigel kits

(Gibco/BRL). Standard agarose gels (100 ml) were poured in 11×13.9 cm gel trays with 14 well combs, and were run in horizontal tanks for submerged gel electrophoresis (Gibco/BRL).

DNA samples were mixed with loading dye (see 2.5.1) prior to loading, and were run alongside DNA molecular weight markers. The DNA markers used were either *Hind* III digested lambda DNA restriction fragments, or a 1 kb DNA ladder (Gibco-BRL). Electrophoresis was carried out at room temperature using a voltage gradient of 1-4 volts/cm. Gels were then examined on a UV transilluminator, and photographed as described in 2.3.5.

2.2.3 Isolation of gel slices

Following the separation of DNA fragments (restriction fragments or PCR products) by electrophoresis on agarose gels, a sterile razor blade was used to excise gel slice(s) containing any DNA fragments that were required for further experiments. Excision of the gel slice(s) was carried out on a UV transilluminator (in the minimum time possible), and after any excess agarose had been cut away, each agarose embedded DNA fragment was transferred to a separate 1.5 ml Eppendorf tube. If at all possible, the DNA was extracted from the gel slices immediately (see 2.2.4), but when this was not feasible, the Eppendorf tubes containing the slices were wrapped in tinfoil and stored at -20°C.

2.2.4 Extraction of DNA from agarose gel slices

Extraction of DNA from agarose gel slices was carried out using either a Geneclean II® kit (BIO 101 Inc.) or a QIAEX gel extraction kit (Qiagen Ltd.). The QIAEX kit was used in the vast majority of experiments, and was generally preferred to the Geneclean II® kit, as it allowed the use of TBE buffer in the gel without the necessity for additional steps during DNA extraction. The kits were used according to the manufacturers' instructions, and the DNA was eluted from the silica gel/matrix by the addition of 20-30 µl of TE or ddH_2O.

2.2.5 Preparation of plasmid vector DNA for subcloning

Plasmid vector DNA (5 µg) was linearised by digestion with the appropriate restriction enzyme(s) in a total reaction volume of 50 µl (containing: 10× restriction enzyme reaction buffer, 30-50 units of restriction enzyme, and ddH_2O) at 37°C for 2-3 hours.

In cases where digestion of the vector DNA had produced DNA ends that were compatible with one another (and could potentially recircularise during ligation reactions), the 5' ends of the linearised vector DNA were dephosphorylated to prevent intramolecular ligation. Twenty eight microlitres of linearised vector DNA was added to 16 µl of ddH_2O, 5 µl of calf intestinal alkaline phosphatase (CIAP) buffer (Boehringer Mannheim), and 1 µl of CIAP (molecular biology quality, ca. 20×10^3 units/ml; Boehringer Mannheim). The dephosphorylation reaction was incubated for 30 minutes at 37°C, after which the reaction was terminated by the addition of EDTA to a final concentration of 5 mM.

The linearised DNA was separated from any residual supercoiled plasmid DNA, by electrophoresis through a 0.8% agarose 1× TAE gel. A gel slice containing the linear vector DNA was cut from the agarose gel, and the DNA was isolated from the gel slice using a QIAEX gel extraction kit (Qiagen Ltd.). A few microlitres of the QIAEX isolated DNA was then checked on an agarose minigel (alongside DNA markers) to estimate the final concentration of the vector DNA.

2.2.6 Ligation reactions/subcloning

Insert DNA for use in subcloning was usually prepared by electrophoretic separation of the DNA of interest on an agarose gel, followed by QIAEX extraction of DNA from a gel slice, or when electrophoretic separation of the DNA product(s) was not required, purification of the DNA using a Magic™/Wizard™ DNA clean-up kit (Promega).

Ligation reactions were generally set up in a total reaction volume of 10 µl containing: 100-200 ng of linearised vector DNA, insert DNA (to give a 2:1 molar ratio of insert:vector DNA), 1 µl of T4 DNA ligase buffer (Promega), 1 µl (1 unit) of T4 DNA ligase (Promega), and ddH_2O to 10 µl.

PCR products were ligated into pGEM®-T vector (Promega; see section 2.6 for vector map), using pGEM-T vector system I (Promega), according to the manufacturer's instructions.

Depending on the facilities available, ligation reactions were either incubated in a water bath at 12°C in the cold room overnight, or were started in a beaker of water (10°C) in an ice bucket and then left to revert to room temperature overnight. The reactions were then spun briefly in a microcentrifuge, and if necessary, stored at 4°C until required. Half of the ligation reaction (5 µl) was used to transform 200 µl of competent cells (see 2.4.3).

2.3 Analysis of purified DNA

2.3.1 Southern blotting

Two methods of Southern blotting were used:

DNA transfer under high salt conditions (neutral pH)

This method was used for Southern blotting of the λEMBL3 G4 clone, the subclones of this lambda clone, and the bacteriophage P1 *VIPR2* clone (Chapters 3 and 4; see section 2.6 for vector maps). Restriction enzyme digests of DNA from the clones were run on an agarose gel (0.8-1.0% agarose, 1× TBE pH 8.3, 1 μg/ml ethidium bromide), which was photographed prior to blotting. The agarose gel was then placed in a small plastic box containing depurination solution, and was incubated in this solution for 15 minutes at room temperature with gentle shaking. After removal of the depurination solution, the gel was rinsed briefly with ddH_2O, and was then soaked in denaturing solution for 30 minutes at room temperature (again with gentle shaking). Another brief rinse with ddH_2O was then followed by the addition of neutralisation solution, and the gel was then left on the shaker at room temperature for at least 30 minutes (during which time the original batch of neutralisation solution was replaced with fresh solution after about 15 minutes). Details of the depurination, denaturation, and neutralising solutions are given in section 2.5.

A piece of Genescreen Plus nylon membrane (DuPont/NEN) was cut to the size of the agarose gel, and left to soak in 10× SSC for 15 minutes. DNA from the neutralised gel was then transferred to the nylon membrane by capillary blotting, using 10× SSC as the transfer solution. Standard transfer apparatus and procedures (Sambrook *et al.*, 1989) were employed, and the assembled blot was left for at least 12 hours to allow transfer to take place. Once the transfer apparatus had been disassembled, the nylon membrane was washed in 0.4 N NaOH for 30-60 seconds, after which it was transferred to a solution consisting of 0.2 M Tris and 2× SSC, and left there for 2-3 minutes. The membrane was then left to air dry at room temperature, before being placed between 2 pieces of 3MM paper (Whatman) and baked at 80°C for 1.5 hours to immobilise the transferred DNA.

DNA transfer under alkaline conditions

For all Southern blots of the λ2001 *Vipr2* clones and subclones (Chapters 3 and 6; see section 2.6 for vector map), alkaline transfer conditions were used with Appligene

Positive™ nylon membrane (Appligene Oncor). Restriction enzyme digests of the clones were run on agarose gels, and photographed prior to blotting, as described in the previous method for DNA transfer under high salt conditions. The gel was then denatured in a solution of 0.4 N NaOH, for 30 minutes with gentle shaking at room temperature.

Standard capillary transfer Southern blot apparatus (Sambrook *et al.*, 1989) was used to transfer the DNA onto the membrane, and 0.4 N NaOH was used as the transfer solution. Once the blot had been set up, it was left for 8-12 hours to allow transfer to occur, after which the blot was disassembled, and the membrane was rinsed in 2× SSC. The membrane was then either transferred immediately into prehybridisation solution, or was placed in a small zip-seal plastic bag with a few millilitres of 2× SSC and stored at 4°C until required.

Prehybridisation and hybridisation of Southern blots

Prehybridisation and hybridisation of Southern blots was carried out in sealed plastic bags which were placed in a box of water on a shaking table in a large incubator. The incubator, water, prehybridisation solution and wash solutions were preheated to the appropriate prehybridisation or wash temperature(s), and were then maintained at that temperature until required.

Blots were prehybridised and hybridised in 20 ml of a solution containing : 50% formamide (deionised), 5× Denhardt's reagent, 0.5% SDS, and 100 µg/ml herring testis DNA (sheared and denatured). After the blot had been prehybridised for at least 1 hour, the radiolabelled probe (see 2.3.2) was added to the prehybridisation solution, and hybridisation of the probe to the membrane was then allowed to occur overnight. Prehybridisation and hybridisation were carried out at 45°C unless otherwise stated, and after removal of the hybridisation solution, the membranes were usually washed at 55°C, first in a solution of 2× SSC, 0.1% SDS (2 washes, 20-30 minutes each), then in a solution of 0.2× SSC, 0.1% SDS (2 washes, 20 minutes each). However, variations were made in the duration and temperature of the wash steps in some cases where the radioactive signal from the membrane was particularly weak or strong. The washed membranes were left to air dry at room temperature, then wrapped in clingfilm and exposed to Fuji RX film (Fuji) in an autoradiography cassette with two intensifying screens. The exposure times (which ranged from a few seconds to 24 hours) and temperatures employed (-70°C or room temperature) depended upon the strength of the radioactive signal (as judged using a Geiger counter).

2.3.2 Preparation, radiolabelling, and purification of probes

Probes for blots and library screens were prepared from agarose gel slices (1× TAE gel) containing the appropriate DNA restriction fragment or PCR product (which in most cases was derived from the rat Vipr2 cDNA). The DNA was isolated from the gel slices using either a QIAEX gel extraction kit (Qiagen Ltd.) or in a few cases a Geneclean II® kit (BIO 101 Inc.), and was eluted in 20-40 µl of ddH_2O. A small volume (2-4 µl) of the purified DNA was then run alongside DNA markers (1 kb DNA ladder, Promega) on an agarose mini gel, so that the concentration of the DNA could be estimated, and 50-100 ng of DNA was then radiolabelled with α-^{32}P dCTP (3000 Ci/mmol; from either Amersham or DuPont/NEN).

During the early work that was carried out on the λEMBL3 G4 *Vipr2* clone (Chapter 3), probes for blots and plaque screens, were labelled by primer extension (using hexanucleotide primers of random sequence) according to the method of (Feinberg & Vogelstein, 1983). However, all probes used during studies of the λ2001 *Vipr2* clones, subclones (Chapters 3 and 6), and the bacteriophage P1 clone (Chapter 4), were labelled using a Ready-To-Go™ DNA labelling kit (Pharmacia) according to the manufacturers instructions. In both cases the probes were purified (to remove unincorporated nucleotides) by centrifugation (at 2000 rpm for 4 minutes) through a Sephadex G50 spun-column (Nu-Clean D50 column, Kodak).

Probes were boiled in a water bath for ~5 minutes (immediately prior to use) and then cooled on ice before they were added to the hybridisation solution.

2.3.3 Automated fluorescent sequencing of DNA

Sequencing reactions

Sequencing reactions were carried out using either an ABI PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit or its successor the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA Polymerase, FS (Perkin Elmer). Plasmid DNA for use in sequencing reactions was prepared using either the SDS-alkaline lysis/PEG precipitation method or a Wizard™ 373 plus DNA purification system (Promega), both of which are described in section 2.1.1. PCR products that were to be sequenced directly (i.e. without being subcloned) were purified using Microcon™ -100 concentrators (Amicon), as described in Chapter 3 , where details of the methods used in the preparation and sequencing of lambda DNA clones are also given. All sequencing reactions were set up according to the manufacturer's

instructions unless otherwise stated. Sequencing products were resuspended in 4 µl of sequencing gel DNA loading buffer (see 2.5.1), and were heated to 90°C for 3-5 minutes then cooled on ice, immediately before being loaded onto the gel.

Preparation of sequencing gel

Sequencing reactions were run on 24-well, 24 cm well-to-read, denaturing polyacrylamide gels (4.75% acrylamide). To prepare the gel mix, 40 g of urea (ultra PURE®, Gibco-BRL) was dissolved in 32.5 ml of ddH_2O (with stirring and gentle heating), and after the solution had been allowed to cool to room temperature, 9.5 ml of a 40% 19:1 acrylamide:bisacrylamide solution (Bio-Rad) was added. The gel solution was deionised by mixing the solution with Amberlite MB-1 resin (Sigma) for 5-10 minutes, and both the deionised gel solution and 8 ml of 10× TBE (pH 8.8) were then filtered through a 0.45 µm cellulose nitrate membrane filter (Whatman), using a vacuum filtration apparatus, to remove the resin and de-gas the gel mix. Polymerisation reagents [400 µl of a 10% (w/v) ammonium persulphate solution and 45 µl of *N,N,N',N'* -tetramethyl-1,2-diaminoethane (TEMED; Sigma)] were mixed with the gel solution just before use, and the gel was then poured immediately (using a 50 ml syringe). Gels were left to polymerise (at room temperature) for at least 2 hours before use.

Electrophoresis conditions

Sequencing reactions were run using an ABI 373 DNA Sequencer (Applied Biosystems), operating on the full scan sequencing option. Sequencing gels were pre-run for at least 10 minutes before the sequencing reactions were loaded onto the gel, and the samples were then electrophoresed at 30 watts for 12-16 hours. The running buffer used for both the pre-run and main electrophoresis run, was 1× TBE (pH 8.8).

Sequence analysis

Collection and preliminary analysis of the sequence data was carried out by the ABI 373 data collection and analysis programs (Applied Biosystems).

The processed data was then transferred to the DNA sequence analysis program Genejockey II (Biosoft), which was used to carry out the initial steps in the analysis of the sequence data e.g. removal of vector sequences, sequence comparisons, and alignment of sequence contigs. Sequence database homology searches were generally carried out using the NCBI BLASTN program (National Centre for Biotechnology Information, Bethesda, USA).

2.3.4 PCR amplification of DNA

With the exception of some of the early PCR amplifications from *VIPR2* P1 clone template DNA (Chapter 4), all PCR reactions were set up in a flow-hood that was not used for any other nucleic acid work. Standard precautions taken to prevent contamination of the PCR reactions, included the use of separate pipettes, and aerosol resistant pipette tips. All reagents were stored in small aliquots, and master mixes were used wherever possible (to reduce the number of pipetting steps involved).

Thermocycling was carried out using a Perkin Elmer DNA thermal cycler. The specific cycling conditions used for particular reactions, and details of the reagents and primers used, can be found in the methods section(s) of the relevant chapter(s).

Primers were obtained from Cruachem Ltd.

2.3.5 Photography and autoradiography

Photographs of UV transilluminated agarose gels were originally taken using a Polaroid® CU-5 land camera (5" lens) and Polaroid® 667 film.

Autoradiography was carried out using Fuji RX medical X-ray film (Fuji) and autoradiography cassettes with two intensifying screens (Fuji). Filters were wrapped in clingfilm before being placed in the cassettes. Films were exposed at -70°C, or at room temperature if the radioactive signal from the filter was very strong, and were then developed using Kodak LX24 X-ray developer and Kodak Industrex manual fixer.

2.4 Microbiological methods

2.4.1 General growth and maintenance of bacterial cultures

Standard aseptic technique procedures were used when working with bacterial cultures, and all glassware used had been baked at 200°C for 2-3 hours (in a Gallenkamp hotbox oven) to sterilise it. Bacteria were grown at 37°C unless otherwise stated. Agar plates on which bacteria had been plated were inverted and left in a 37°C incubator overnight. Liquid cultures of bacteria were placed in an orbital incubator (at 37°C, 200 rev/min) overnight. Liquid cultures of ≤ 50 ml were set up from single isolated colonies of plated bacteria, whereas larger volume cultures were generally set up from 2 ml overnight liquid cultures.

Ampicillin (50 μg/ml) was added to culture media that was to be used for selection and growth of bacterial clones containing pBluescript or pGEM plasmid vectors (see section 2.6 for vector maps), whereas kanamycin (25 μg/ml) was added to culture media that was to be used for the growth of bacterial clones containing the bacteriophage P1 vector p*Ad*10-*sacB*II (see section 2.6 for vector map).

Plates of bacterial colonies were sealed with parafilm and stored in the cold room for up to 3 weeks. Glycerol stocks were used for the long term storage of bacteria at -70°C (see 2.4.5)

2.4.2 Preparation of competent bacterial cells

A single isolated colony of XL1-Blue or JM109 bacterial cells (grown on M9+thiamine-HCl plates, see 2.5.5), was used to inoculate 2 ml of L-broth, and the culture was grown overnight at 37°C. The 2 ml culture, was then in turn used to inoculate a flask containing 40 ml of L-broth, and growth was allowed to continue until the OD_{600} of the bacterial culture was between 0.3 and 0.5. The flask of cells was then placed on ice, and from this point onwards all steps were carried out on ice, and centrifugations were carried out at 4°C. Once the culture had cooled, it was transferred to a sterile pre-chilled Oakridge tube, and the cells were collected by centrifugation at 6500 rpm for 5 minutes in a Sorvall SS-34 rotor. The supernatant was removed, and the pellet was resuspended in 10 ml of ice-cold 0.1 M $MgCl_2$. The cell suspension was then centrifuged again (6500 rpm for 5 minutes in a Sorvall SS-34 rotor) to pellet the cells, which this time were resuspended in 2 ml of ice-cold 0.1 M $CaCl_2$, and then left on ice for 45 minutes. Following a final centrifugation (6500 rpm for 5 minutes in a Sorvall SS-34 rotor), the cells were resuspended in 2 ml of ice cold MOPS/$CaCl_2$/glycerol solution (see below), and then incubated on ice for a further 20 minutes. The competent cells were then aliquoted into ice-cold Eppendorf tubes (200 μl/tube), snap frozen (in a dry ice/ethanol bath), and stored at -70°C until required.

The MOPS/$CaCl_2$/glycerol solution contained:

0.1 M 3-(*N*-morpholino)propanesulfonic acid (MOPS) pH 6.5

50 mM $CaCl_2$

20% glycerol

2.4.3 Transformation of bacterial cells

For each transformation reaction, a 200 µl aliquot of competent cells was taken from storage at -70°C, and allowed to defrost on ice. Once the cell suspension had thawed, 3 µl of AnalaR® DMSO (BDH Chemicals Ltd.) was added to the tube, and was mixed with the cell suspension by very gently flicking the tube. Approximately 5 µl of ligation mix was then added, and after the contents of the tube had been gently mixed again, the transformation reactions were left on ice for at least an hour. The cells were then heat shocked in a 42°C water bath for 2-3 minutes, and returned to ice where they were left for a further 15 minutes. Transformation reactions that involved the retransformation of a previously isolated plasmid, were plated directly onto selection plates (see 2.4.4) at this stage. However, in cases where ligation reactions were being used to transform bacterial cells (and the expected number of transformants was low), the cells were transferred to a 5 ml sterile disposable plastic tube, and following the addition of 2 ml of L-broth, the cells were grown at 37°C for 1 hour with gentle shaking. The cells were then pelleted by centrifugation, resuspended in 100 µl of L-broth, and spread onto selection plates (see 2.4.4).

2.4.4 Selection of transformed cells

Blue/white colour screening (based on the α-complementation system) was used to select transformed JM109 or XL1-blue cells. After transformation with plasmid DNA (see 2.4.3) the bacteria were spread onto AIX selection plates (see 2.5.5) that had been dried and pre-warmed to 37°C. The plates were then inverted, and left in a 37°C incubator overnight to allow the bacteria to grow. If colonies were visible on the plates the next morning, the plates were sealed and left in the cold room for 1-2 hours to allow the colour to develop fully. Although pure white colonies were usually indicative of transformants, colonies which had small blue centres were sometimes selected (in addition to any white colonies) in cases where the plasmid inserts of interest were small (≤ 200 bp), as these fragments do not always completely inactivate the amino terminal encoding region of the *lacZ* gene.

2.4.5 Preparation of glycerol stocks

Glycerol stocks containing a final concentration of ~20% glycerol were prepared by adding 700 µl of sterile 80% glycerol solution to a test tube containing a 2 ml overnight bacterial culture. Each glycerol stock was divided between 3 pre-cooled Eppendorf tubes (~900 µl/tube). The glycerol stocks were then snap frozen (in a dry ice/ethanol bath), and stored at -70°C.

2.5 Composition of buffers, solutions, and media

2.5.1 Buffers

TE (pH 8.0)

10 mM Tris.Cl (pH 8.0)

1 mM EDTA (pH 8.0)

TAE

stock solution (50×)	working concentration (1×)
242 g of Tris base	0.04 M Tris-acetate
57.1 ml of glacial acetic acid	0.001 M EDTA
100 ml of 0.5 M EDTA (pH 8.0)	
deionised H_2O to 1 litre	

TBE

stock solution (10×)	working concentration (1×)
108 g of Tris base	0.09 M Tris-borate
55 g of boric acid	0.002 M EDTA
40 ml of 0.5 M EDTA (pH 8.0)	
deionised H_2O to 1 litre	

Batches of TBE stock solution that were to be used for agarose gel electrophoresis were adjusted to pH 8.3 with boric acid, whereas those that were going to be used in DNA sequencing experiments were adjusted to pH 8.8.

SM buffer

500 ml of SM buffer contained:

2.9 g of NaCl

1.0 g of $MgSO_4.7H_2O$

25 ml of 1 M Tris.Cl (pH 7.5)

2.5 ml of 2% gelatine solution

H_2O to 500 ml

The buffer was sterilised by autoclaving at 121°C, 15 lb in^{-2} for 20 minutes, and 100 ml aliquots were then transferred to sterile glass bottles.

Sequencing gel DNA loading buffer

The buffer used for loading sequencing reactions onto ABI 373 sequencing gels consisted of a 5 to 1 ratio of deionised formamide: 25 mM EDTA (pH 8.0) containing 50 mg/ml blue dextran.

Agarose gel DNA loading buffer (6×)

0.25% bromophenol blue

0.25% xylene cyanol FF

30% glycerol in water

2.5.2 Alkaline lysis solutions for plasmid isolation

Alkaline lysis solution 1

50 mM glucose

25 mM Tris.Cl (pH 8.0)

10 mM EDTA (pH 8.0)

This solution was sterilised by autoclaving at 121°C, 15 lb in^{-2} for 20 minutes, and was then stored at 4°C.

Alkaline lysis solution 2

0.2 N NaOH (diluted from 5 N stock)

1% SDS

Alkaline lysis solution 3

100 ml of alkaline lysis solution 3 contained:

60.0 ml of 5 M potassium acetate

11.5 ml of glacial acetic acid

28.5 ml of ddH_2O

2.5.3 Solutions for Southern blotting

Solutions used for blotting of DNA onto Genescreen Plus nylon membrane (NEN/DuPont)

Depurination :	0.25 N HCl
Denaturation :	0.4 N NaOH, 0.6 M NaCl
Neutralisation :	1.5 M NaCl, 0.5 M Tris.Cl (pH 7.0)
Transfer:	10× SSC (see 2.5.4)

Solution used for alkaline blotting of DNA onto positive-charged nylon membrane (Appligene Oncor)

Denaturation and transfer: 0.4 N NaOH

2.5.4 Hybridisation solution components

20× SSC (standard saline citrate)

1 litre of 20× SSC contained: 175.3 g of NaCl and 88.2 g of sodium citrate, which were initially dissolved in 800 ml of deionised water. The pH of this solution was adjusted to 7.0 using 10 N NaOH, and deionised water was then added to give a final volume of 1 litre. The solution was then divided into 3 batches of ~330 ml, which were sterilised by autoclaving at 121°C, 15 lb in^{-2} for 20 minutes.

20× SSPE (salt-sodium phosphate-EDTA)

1 litre of 20× SSPE contained: 175.3 g of NaCl, 27.6 g of $NaH_2PO_4.H_2O$, and 7.4 g of EDTA, which were initially dissolved in 800 ml of deionised water. The pH of this solution was then adjusted to pH 7.4 using 10 N NaOH, and the final volume of the solution was subsequently made up to 1 litre by the addition of deionised water. The solution was then divided into 3 batches of ~330 ml, which were sterilised by autoclaving at 121°C, 15 lb in^{-2} for 20 minutes.

Denhardt's reagent

Denhardt's reagent was prepared as a 50× stock solution containing:

5 g of Ficoll type 400

5 g of polyvinylpyrrolidone

5 g of bovine serum albumin (BSA) fraction V

and deionised H_2O to 500 ml

The solution was filtered, aliquoted into 50 ml batches, and then stored at -20°C.

Deionised formamide

Deionised formamide was prepared using Amberlite MB-1A resin (Sigma). The resin was added to the formamide (~5 g per 100 ml of formamide) and stirred (with a magnetic stirrer) at room temperature for at least one hour. After being filtered through Whatman No. 1 paper (to remove the resin), the deionised formamide was aliquoted into 50 ml batches and stored at -20°C

Non-homologous DNA

A 10 mg/ml solution of herring testis DNA in deionised water was prepared, and passed through a 17-gauge syringe needle several times in order to shear the DNA. The sheared DNA was boiled for 10-15 minutes in a water bath (to denature the DNA), and then cooled on ice. In cases where the DNA was not used immediately, it was aliquoted, and stored at 4°C. These aliquots were boiled again (for 5-10 minutes) and cooled on ice before use.

2.5.5 Bacterial media and antibiotics

B-agarose

To prepare 1 litre of B-agarose, 10 g of Bacto®-tryptone (Difco) and 8 g of NaCl were dissolved in deionised water to give a final volume of 1 litre. This solution was transferred into glass bottles, where 1.5 g of agarose was added per 100 ml, and the bottles of media were then autoclaved at 121°C, 15 lb in^{-2} for 20 minutes.

B soft top-agarose

Prepared as described for B-agarose, but 0.7 g of agarose was added per 100 ml of solution.

L-broth

To prepare 1 litre of L-broth, 10 g of Bacto®-tryptone (Difco), 5 g of Bacto®-yeast extract (Difco), and 5 g of NaCl were dissolved in ~900 ml of deionised water. The pH of the solution was adjusted to pH 7.0-7.5 using 5 N NaOH, and deionised water was added to give a final volume of 1 litre. The L-broth was then transferred into glass bottles and autoclaved at 121°C, 15 lb in^{-2} for 20 minutes

L-agar

Prepared as described for L-broth, but after the solution had been transferred into glass bottles 1 g of Bacto®-agar was added per 100 ml of solution.

Minimal agar +1 mM thiamine-HCl (M9 media for maintenance of JM109 cells)

1 litre of M9 agar + thiamine-HCl was prepared from:

6 g of disodium hydrogen phosphate (Na_2HPO_4)

3g of potassium dihydrogen phosphate (KH_2PO_4)

1g of ammonium chloride (NH_4Cl)

0.5g of NaCl

15g of Bacto®-agar

to which ddH_2O was added to 1 litre. The pH was then adjusted to 7.4 by addition of 10 N NaOH, and the solution was autoclaved at 121°C, 15 lb in^{-2} for 20 minutes. The

solution was then allowed to cool below 50°C before the following (autoclaved or filter sterilised) components were added:

10.0 ml of 20% glucose

2.0 ml of 1M $MgSO_4$

0.1 ml of 1M $CaCl_2$

1.0 ml of 1M thiamine-HCl

2× YT-broth

To prepare 1 litre of 2× YT-broth, 16 g of Bacto®-tryptone, 10 g of Bacto®-yeast extract, and 5 g of NaCl, were dissolved in 900 ml of deionised water. The pH of the solution was adjusted to 7.0 with 5 N NaOH (if necessary), and after the final volume of the solution had been adjusted to 1 litre (with deionised water), the 2× YT-broth was transferred into glass bottles, and autoclaved at 121°C, 15 lb in^{-2} for 20 minutes.

Antibiotics

Ampicillin and kanamycin solutions were prepared in ddH_2O, and were filter sterilised through a 0.2 μm filter.

Ampicillin

stock solution: 50 mg/ml

working concentration: 50 μg/ml (in both liquid cultures and agar plates)

Kanamycin

stock solution: kanamycin monosulfate 25 mg/ml

working concentration: 25 μg/ml (in both liquid cultures and agar plates)

Melted agar solutions were allowed to cool to 50°C before antibiotics were added to them.

Isopropylthio-β-D-galactoside (IPTG)

0.1 M stock solution contained:

0.24 g of IPTG

ddH2O to 10 ml

1 M stock solution contained:

2.38g of IPTG

ddH2O to 10 ml

IPTG solutions were filter sterilised using a 0.2 µm filter, and then stored at -20°C

5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal)

Stock solutions of X-gal (50 mg/ml) in N, N'-dimethylformamide were purchased from Promega, and were stored in light-tight containers at -20°C

AIX plates

AIX plates (for blue/white colour screening of potential transformants) contained 50 µg/ml ampicillin, 0.1 mM IPTG, and 50 µg/ml X-gal, and were prepared by adding the appropriate stock solutions (see above) to molten L-agar that had been allowed to cool to 50°C. The supplemented L-agar was poured into small petri dishes (90 mm diameter), and allowed to set. The AIX plates were then dried at 37°C for ~1 hour, and either used immediately or stored at 4°C until required (3 weeks maximum).

2.6 Cloning vectors

Plasmid vectors used during these studies included: pBluescript II SK (Stratagene; see Figure 2.1 for vector map); pGEM®-7Zf (Promega; see Figure 2.2 for vector map); and pGEM®-T vector (Promega; see Figure 2.3 for vector map).

Mouse genomic libraries that were screened had been constructed in the bacteriophage lambda based vectors λEMBL3 (Frischauf *et al.*, 1983; see Figure 2.4 for vector map) or λ2001 (Karn *et al.*, 1984; see Figure 2.5 for vector map), and the human genomic library had been constructed in bacteriophage P1 based vector p*Ad*10-*sacB*II (Sternberg, 1990; Pierce and Sternberg, 1992a; Pierce and Sternberg, 1992b; see Figure 2.6 for vector map).

Figure 2.1: Map of pBluescript II SK (Stratagene)

The pBluescript® II SK (+/-) phagemid is a 2961-bp phagemid derived from pUC19. The SK designation indicates the polylinker is oriented such that *lacZ* transcription proceeds from *Sac* I to *Kpn* I.

f1 (+) origin: (3-459 bp) f1 filamentous phage origin of replication allowing recovery of the sense strand of the *lacZ* gene when a host strain containing the pBluescript II phagemid is co-infected with helper phage.

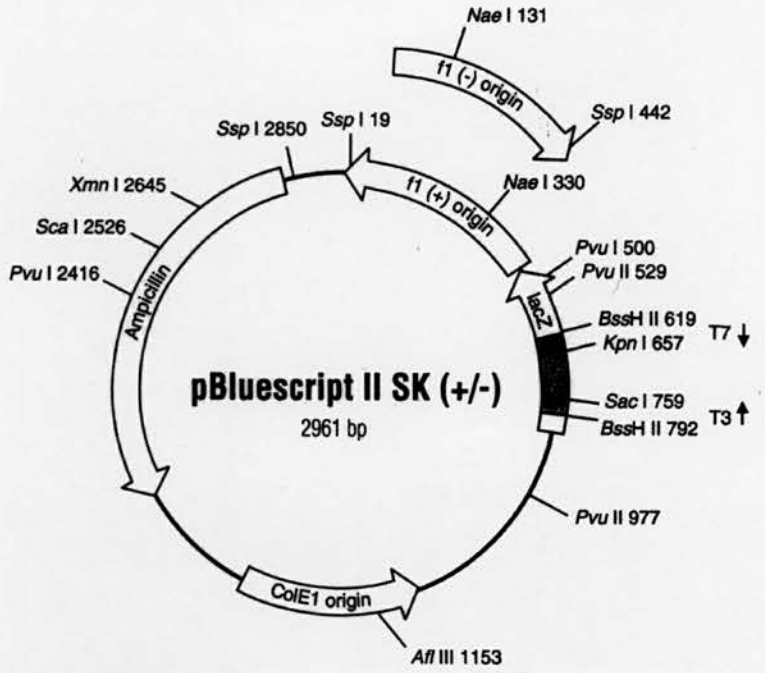
f1 (-) origin: (3-459 bp) f1 filamentous phage origin of replication allowing recovery of the antisense strand of the *lacZ* gene when a host strain containing the pBluescript II phagemid is co-infected with helper phage.

ColE1 origin: (1032-1972 bp) Plasmid origin of replication used in the absence of helper phage.

***lacZ* gene:** (*lac* promoter: 816-938 bp) This portion of the *lacZ* gene provides α-complementation for blue/white color selection of recombinant phagemids. An inducible *lac* promoter upstream from the *lacZ* gene permits fusion protein expression with the β-galactosidase gene product.

MCS: (657-759 bp) Multiple cloning site flanked by T3 and T7 RNA promoters (please see the polylinker sequence below).

Ampicillin: (1975-2832 bp) Ampicillin-resistance gene (Amp^r) for antibiotic selection of the phagemid vector.



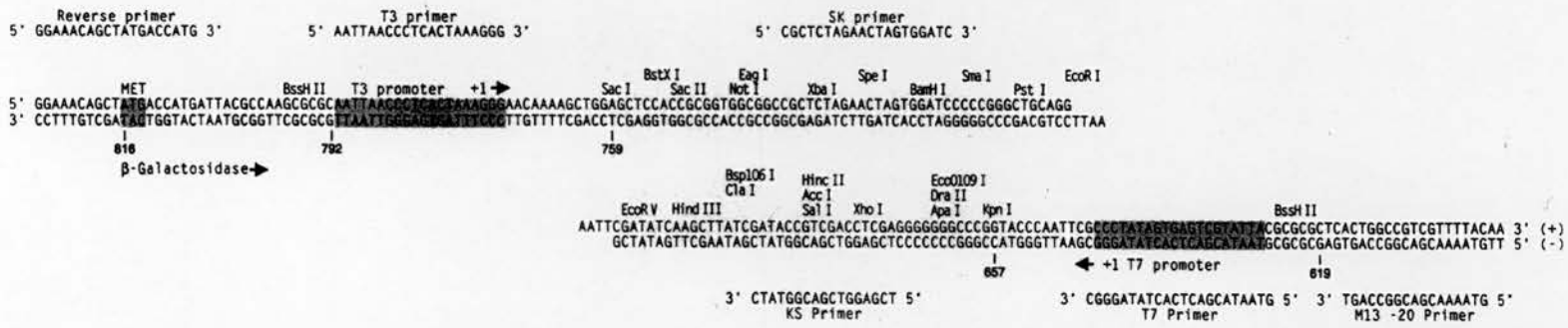
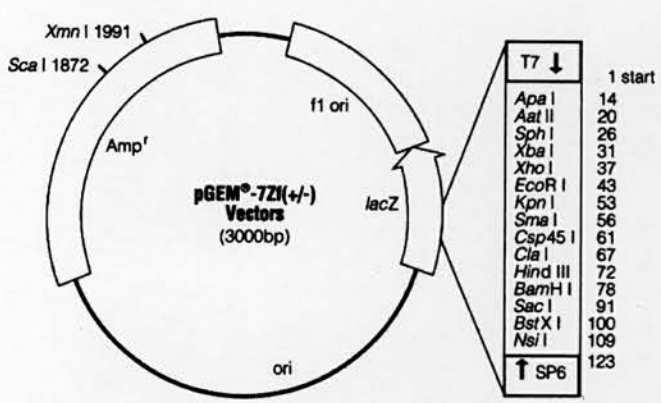


Figure 2.2: Map of pGEM®-7Zf (Promega)



Vector Map Notes:

1. Sequence reference points:

a. T7 RNA polymerase transcription initiation site	1
b. SP6 RNA polymerase transcription initiation site	123
c. T7 RNA polymerase promoter	2984-6
d. SP6 RNA polymerase promoter	118-140
e. multiple cloning sites	10-110
f. *lacZ* start codon	162
g. *lac* operon sequences	2821-2981; 148-377
h. *lac* operator	182-198
i. β-lactamase (Amp^r) coding region	1319-2179
j. phage f1 region	2365-2820
k. binding site of pUC/M13 Forward Sequencing Primer	2941-2957
l. binding site of pUC/M13 Reverse Sequencing Primer	158-174

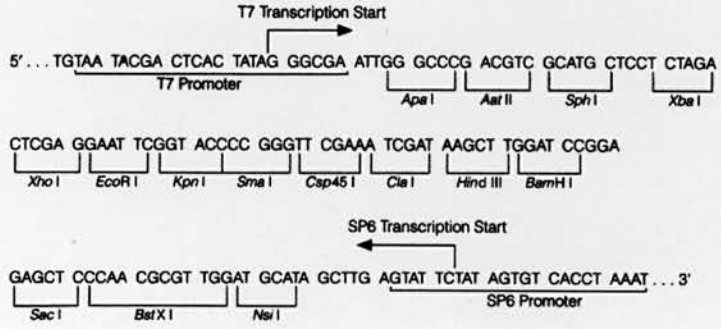
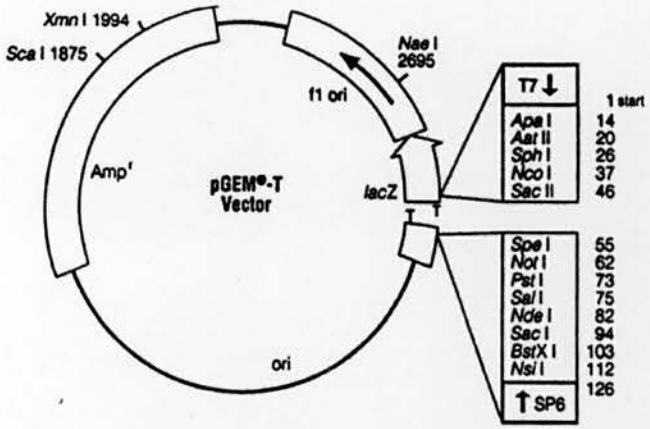


Figure 2.3: Map of pGEM® -T vector (Promega)



Sequence reference points:

a.	T7 RNA Polymerase transcription initiation site	1
b.	SP6 RNA Polymerase transcription initiation site	126
c.	T7 RNA Polymerase promoter	2987-3003
d.	SP6 RNA Polymerase promoter	127-143
e.	multiple cloning site	10-113
f.	*lacZ* start codon	165
g.	*lac* operon sequences	2824-2984, 151-180
h.	*lac* operator	151-201
i.	β-lactamase coding region	1322-2182
j.	phage f1 region	2368-2823
k.	binding site of pUC/M13 Forward Sequencing Primer	2944-2960
l.	binding site of pUC/M13 Reverse Sequencing Primer	161-177

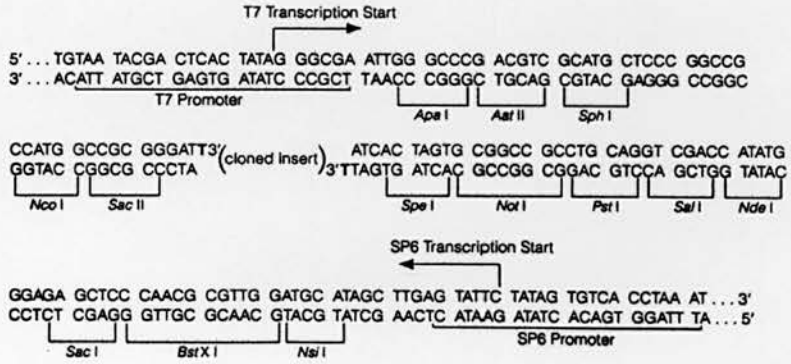
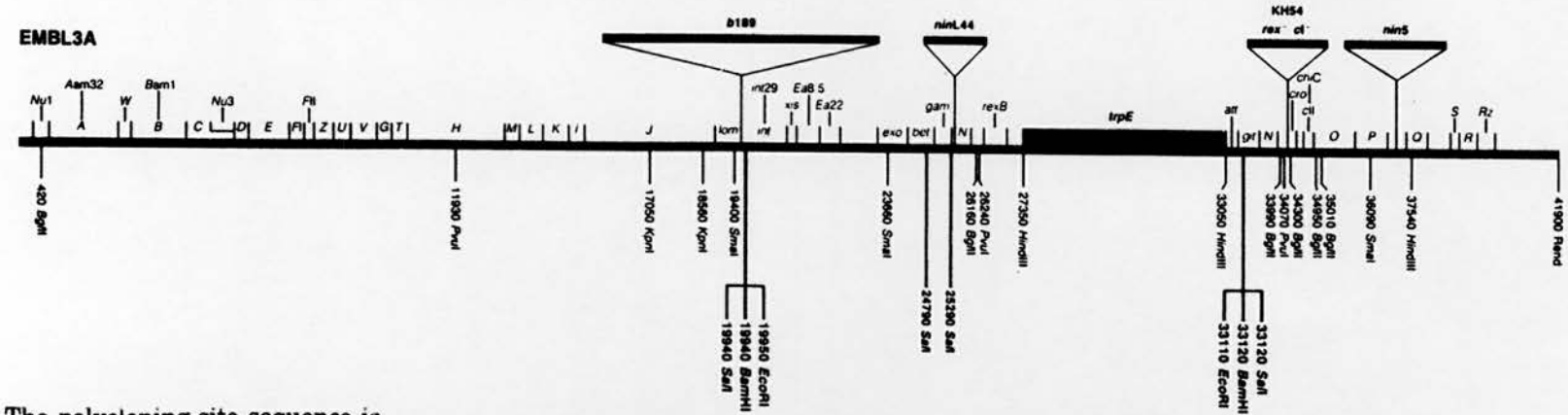


Figure 2.4: Map of λEMBL3A



The polycloning-site sequence is

GGATCTGGGTCGACGGATCCGGGGAATTCCCAGATCC

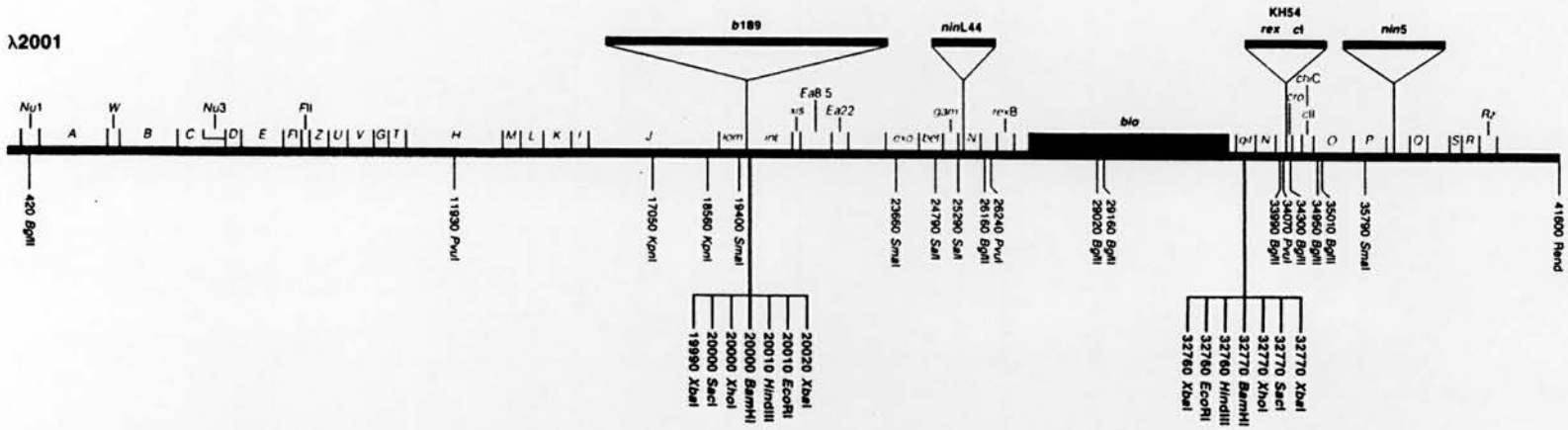
SalI BamHI EcoRI

left arm = 19.9 kb
right arm = 8.8 kb
insert size = 7-20 kb

N.B. the amber mutations *A*am32 and *B*am1 shown in this map of λEMBL3A, are not found in λEMBL3

References: Frischauf *et al.*, 1983; Sambrook *et al.*, 1989.

Figure 2.5: Map of λ2001



The polycloning-site sequence is

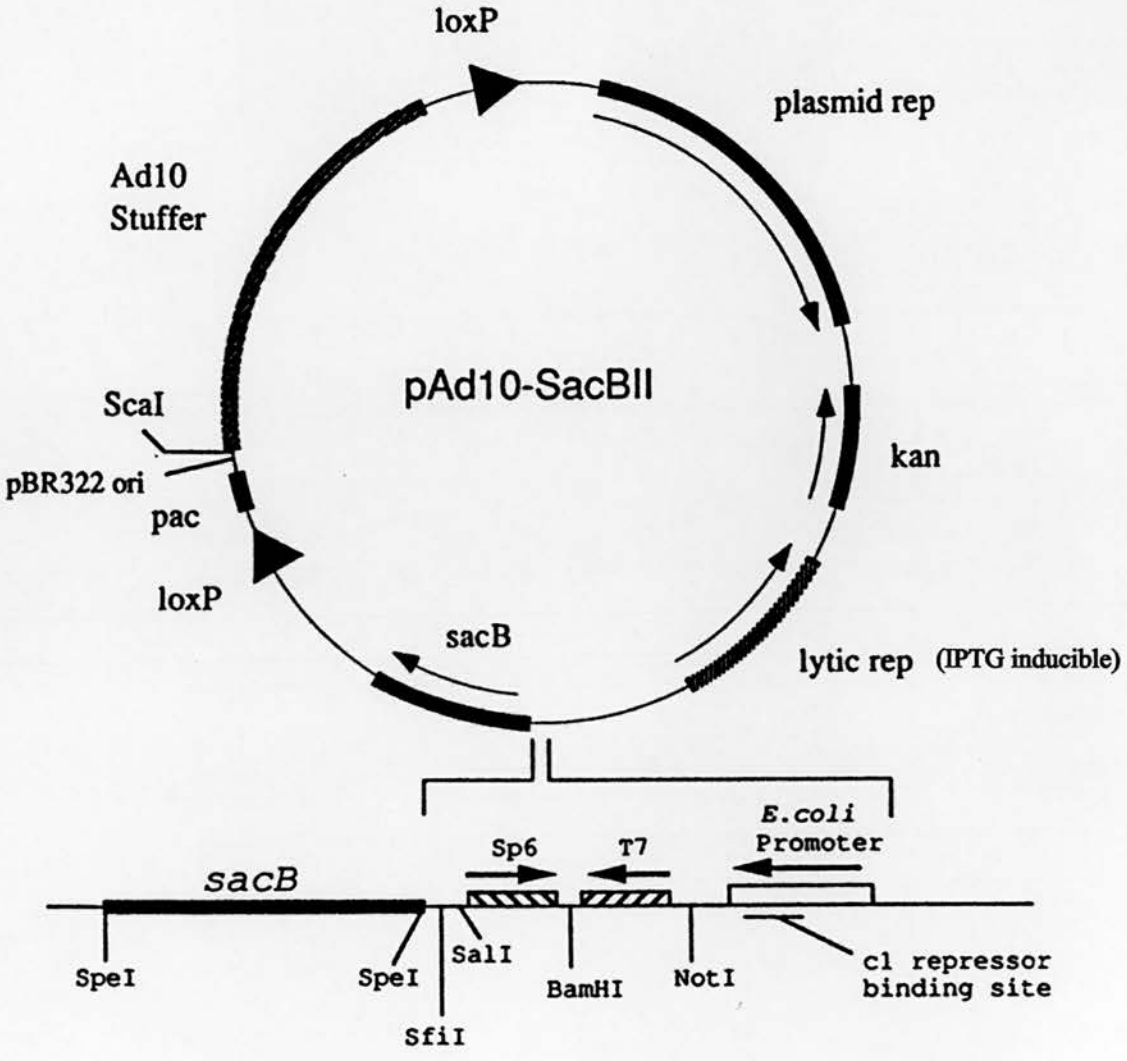
TCTAGAGCTCGAGGATCCAAGCTTGAATTCTAGA

XbaI SacI XhoI BamHI HindIII EcoRI XbaI

left arm = 20 kb
right arm = 9 kb
insert size = 10-22 kb

References: Karn *et al.*, 1984; Sambrook *et al.*, 1989.

Figure 2.6: Map of the bacteriophage P1 cloning vector p*Ad*10-*sacB*II



loxP = 34 bp viral recombination sites on which the P1 cyclisation recombinase protein (Cre) can act.
kan = kanamycin resistance gene
sacB = gene encoding the enzyme levansucrase
Ad10 stuffer = 10 kb adenovirus derived stuffer fragment
pac = 162 bp site that is recognised and cleaved by phage-encoded pacase proteins (orients the DNA for packaging)

Reference: Pierce and Sternberg, 1992b.

Chapter 3

Isolation and characterisation of the mouse vasoactive intestinal peptide type 2 receptor gene (*Vipr2*)

3.1 Introduction

The work that led up to the studies described in this chapter, began with the PCR screening of rat pituitary cDNA, using degenerate PCR primers that had been designed to bind to the highly conserved sequences that are found within the regions encoding the predicted 3rd and 7th transmembrane domains of members of the secretin/glucagon receptor subfamily (Lutz *et al.*, 1993). Of the three different classes of PCR product that were identified, two appeared to represent novel receptor encoding sequences, whereas the third class of PCR product corresponded to part of the published calcitonin receptor cDNA sequence (Gorn *et al.*, 1992; Lin *et al.*, 1991). Subsequent isolation (from a rat olfactory bulb cDNA library) and characterisation of full length cDNA clones corresponding to the two novel sequences, revealed that one of the clones encoded a splice-variant of the PACAP type 1 receptor (Morrow *et al.*, 1993), and the other encoded a second receptor for VIP and PACAP, the vasoactive intestinal peptide type 2 (VIP_2) receptor (Lutz *et al.*, 1993), which became the focus of the work reported in this thesis (see Chapter 1).

While the isolation and characterisation of full length clones of these novel cDNAs was in progress, the 500 bp PCR product (clone 4 /Ac(500)4) representing the region of cDNA encoding transmembrane domains 3-7 of the VIP_2 receptor, was used as a probe to screen a λEMBL3 mouse genomic library. One positive clone was isolated, and this clone, which was named λG4 (to indicate that it was the genomic clone that had been isolated by the Ac(500)4 Vipr2 partial cDNA probe), provided the starting point for the characterisation of the mouse *Vipr2* gene.

Although cDNAs encoding at least 6 members of the secretin/glucagon receptor subfamily had been isolated by this time (Ishihara *et al.*, 1991; Ishihara *et al.*, 1992; Jüppner *et al.*, 1991; Lin *et al.*, 1991; Mayo, 1992; Thorens, 1992), the only published information relating to the structure of the genes that encoded these receptors was an abstract (Kronenberg, 1992) which stated that the rat PTH/PTHrP receptor gene contains "a series of introns that separate individual putative transmembrane domains and discrete portions of the amino-terminal extracellular domain", and a paper on the

cloning of a cDNA encoding the rat GHRH receptor, which reported the existence of a possible site of splice variation within the region encoding the 3rd intracellular loop of the receptor, and described the rat GHRH receptor gene as "quite large" and "interrupted by numerous introns within the coding region" (Mayo, 1992). Together, these reports indicated that the presence of introns was probably a common feature of the genes that encoded members of this receptor subfamily. However, they provided few clues to suggest what the total size of the mouse *Vipr2* gene might be, or give any clear indication of how many introns the gene was likely to contain. Therefore, although there were several possible approaches that could be taken to characterise the structure of the mouse *Vipr2* gene (including PCR amplification of introns, restriction mapping, or direct sequencing of the clones), the characterisation of the initial mouse genomic clone (λG4) was begun without any specific plan as to how the structure of the gene would finally be determined, as the choice of method(s) would depend to a large extent on the size of the introns (if any) within the *Vipr2* gene.

Preliminary characterisation of the λG4 clone indicated that, like the rat GHRH receptor gene (Mayo, 1992) and PTH/PTHrP receptor gene (Kronenberg, 1992), the mouse *Vipr2* gene contained several introns. However, some problems were encountered with this genomic clone, and by the time that a second genomic library (which had the additional advantage of being ES-cell derived and therefore preferable for the long term transgenic aims of this work) had been screened for further *Vipr2* gene clones, the gene structures of several other members of the receptor gene subfamily had been published (Kong *et al.*, 1994; Lin *et al.*, 1993; Maget *et al.*, 1994; Zolnierowicz *et al.*, 1994). These included: the gene structure of the rat, mouse and human PTH/PTHrP receptor genes (reported to contain at least 15 exons; Kong *et al.*, 1994), the pig calcitonin gene (at least 14 exons; Zolnierowicz *et al.*, 1994), the majority of the rat glucagon receptor gene structure (at least 12 exons; Maget *et al.*, 1994), and the positions of the introns in the mouse GHRH receptor gene (at least 13 exons; Lin *et al.*, 1993). Furthermore, a common pattern had begun to emerge in the intron/exon structure of these receptor genes, which allowed the positions of the mouse *Vipr2* gene's intron/exon boundaries to be predicted by comparison of its cDNA sequence with the exon sequences of other members of the receptor gene subfamily.

Following the isolation of a series of λ clones that together spanned the mouse *Vipr2* gene (with the exception of part of intron 4), the predicted structure of the *Vipr2* gene was used as a template to design primers that could be used to sequence (outwards from the exon sequences) across the predicted exon/intron junctions. These primers were used to establish the sequences and positions of the *Vipr2* gene's exon/intron

junctions, and the sizes of all but one of the identified introns were subsequently determined using a combination of PCR and restriction mapping.

3.2 Methods

3.2.1 Preparation of phage lysates

Plate lysates

A single bacterial colony of the *E. coli* host strain (LE392 in the case of the λEMBL3 library, and Q358 in the case of the λ2001 library), was used to inoculate 10 ml of L-broth containing 10 mM $MgSO_4$ + 0.2% maltose, and the culture was left to grow overnight at 37°C, with gentle shaking. The bacteria were then pelleted in 30 ml Oakridge tubes, by centrifugation at 4000 rpm for 10 minutes at 4°C in a Sorvall SS-34 rotor. After removal of the supernatant, the pellet was resuspended in a small volume (3-5 ml) of 10 mM $MgSO_4$, and then diluted further (with 10 mM $MgSO_4$) until the bacterial suspension reached an OD_{600} of ~0.5. Phage stocks were diluted in SM buffer (see Chapter 2) to give a final concentration of $\sim 3.6 \times 10^5$ plaque forming units (pfu) per 100 µl.

Large (140 mm diameter) B-agar plates were prepared (see Chapter 2), dried, and then warmed in an incubator at 37°C. Twenty plates were prepared for each λ clone in cases where the lysate was to be used for a DNA maxiprep, and 5 plates were prepared when the lysate was going to be used for DNA midipreps or kept for use as a phage stock. While the plates were being warmed, an appropriate volume of soft top-agarose (~10 ml per plate) was melted (in a microwave oven), allowed to cool to 46°C, and then placed in a water bath (pre-heated to 46°C) where it was maintained at that temperature until required. Adsorption of the phage to the bacteria, was carried out in 10 ml test tubes that had been pre-warmed to 37°C in a heating block. For each plate that had been prepared, 0.5 ml of the bacterial cell suspension and 70 µl of the phage dilution were added to one of the 10 ml test tubes and gently mixed before being left at 37°C for ~20 minutes to allow adsorption to take place. Magnesium sulphate was added to the soft top-agarose (1 ml of 0.1 M $MgSO_4$ stock/100 ml of soft top-agarose) just prior to use, and 7 ml of soft top-agarose was then added to the first test tube. After addition of the agarose, the contents of the tube were mixed by inverting the tube, and then poured immediately onto one of the pre-warmed B-agar plates. This process was repeated with each of the remaining test tubes. The plates were left at room temperature, until the top-

agarose had set, and were then incubated (upright) in an oven at 37°C for 6-8 hours or until the plaque formation (resulting from lysis of the bacteria) was confluent. Once lysis was complete, 10 ml of SM buffer was added to each plate, and the plates were left on a slowly rotating shaker in the cold room overnight. The following morning, the SM buffer was carefully removed from the plates using disposable Pasteur pipettes, and transferred to a centrifuge bucket, where chloroform was added (1 µl per ml of lysate). The chloroform and lysate were mixed gently, to kill any remaining bacteria, and centrifuged for 10 minutes at 7000 rpm in a Sorvall GSA rotor at 4°C, to pellet any debris. The supernatant (phage lysate) was then removed, and stored at 4°C (over a few drops of chloroform) until required.

Liquid lysates

A single bacterial colony of the *E. coli* host strain (LE392 in the case of the λEMBL3 library, and Q358 in the case of the λ2001 library), was used to inoculate 10 ml of L-broth containing 10 mM $MgSO_4$ + 0.2% maltose, and the culture was left to grow overnight at 37°C, with gentle shaking. This 10 ml culture was then used to inoculate 250 ml of L-broth containing 10 mM $MgSO_4$ and 0.2% maltose, which in turn was left to grow at 37°C in an orbital shaker until the OD_{600} of the culture reached about 0.5. At this point, 6×10^{10} pfu of phage stock was added to the culture, and the culture was then shaken quite vigorously (~200 cycles/minute) at 37°C, until lysis of the bacterial cells by the phage caused a very obvious clearing of the previously opaque bacterial culture and string-like clumps of bacterial debris became visible at the bottom of the flask. Chloroform (0.5 ml) was then added to the culture and incubation at 37°C in the orbital shaker was continued for a further 10 minutes, after which the contents of the flask were transferred into a centrifuge bottle and centrifuged for 10 minutes at ~8000 rpm in a Sorvall GSA rotor at 4°C. The supernatant was carefully decanted into a clean 250 ml bottle, and was stored at 4°C (over chloroform) until required.

3.2.2 Titration of phage stocks/libraries

For each titration 7 small (90 mm diameter) B-agar plates (see Chapter 2) were prepared, dried, and then warmed in an incubator at 37°C just prior to use. *E. coli* host cells for the phage infections were grown as described in section 3.2.1 (plate lysate method), and resuspended at an OD_{600} of ~0.5 in 10mM $MgSO_4$.

Serial 10-fold dilutions of the phage suspension were made in SM buffer, to generate a series of dilutions that spanned a concentration range from 10^{-5} to 10^{-10} of the concentration of the original phage suspension, and each of the dilutions within this

range was then used to infect the *E. coli* host cells. Seven 10 ml test tubes were pre-heated to 37°C in a heating block (one test tube for each phage dilution, and one tube for a negative control reaction). Each of the tubes was labelled to indicate which one of the six phage dilutions (or -ve control reaction) would be added to it, and the seven B-agar plates were also labelled accordingly.

A 100 µl aliquot of the appropriate phage dilution was added to each of the labelled tubes, and 100 µl of SM buffer was added to the -ve control tube. This was followed by the addition of 200 µl of the bacterial cell suspension, to each of the seven tubes, and after gently mixing, the reactions were left at 37°C for 20 minutes, so that the phage could adsorb to the bacteria. Soft top-agarose containing 10 mM $MgSO_4$ was prepared as described in section 3.2.1 (plate lysate method), and 3 ml of the soft top-agarose was added to the first test tube. After addition of the agarose, the contents of the tube were quickly mixed by inversion, and then immediately poured onto the appropriately labelled B-agar plate. The process was repeated with the other six test tubes, and the plates were then left at room temperature to set, before being inverted and incubated at 37°C overnight to allow plaques to form. Plates on which plaque growth was confluent, or the plaques had begun to merge with one another were discarded, as were any plates on which there were no plaques. The number of plaques on one of the remaining plates (preferably one with 10-300 plaques) was then counted, and used to calculate the titre of the original phage stock.

The phage dilutions used in this assay should be suitable for the titration of phage stocks whose titres lie between 3×10^8 pfu/ml and 1×10^{12} pfu/ml.

3.2.3 Isolation of λ DNA

Qiagen column kits

Qiagen lambda midi- and maxi-kits (Qiagen Ltd.) were used as the standard method for the isolation of λ DNA throughout this study. Phage lysates were prepared using either the plate lysate or liquid lysate methods described in section 3.2.1, the choice of method depending on the rate of growth of the clone in question. For midipreps of λ DNA ~40 ml of phage lysate was used as the starting material, and for maxipreps of λ DNA ~150 ml of phage lysate was used. Isolation of λ DNA, using the Qiagen lambda kits, was carried out according to the manufacturer's instructions.

Phage purification on DEAE-cellulose columns, and subsequent isolation of λ DNA by proteinase K-SDS treatment

During attempts to obtain satisfactory preparations of λ DNA from the λEMBL3 *Vipr2* clone (λG4), a second method of λ DNA isolation was also tried. This method which was published by Raikhinstein and Hanukoglu (1992), combined: ultracentrifugation of the phage lysate, purification of the phage particles on a DEAE-cellulose minicolumn, and a proteinase K-SDS lysis-based protocol for the recovery of λ DNA.

Six phage lysate plates were prepared from phage stocks as described in section 3.2.1 (plate lysate method), and were incubated overnight at 37°C to allow lysis to occur. The following morning, the plates were transferred to the cold room, where they were cooled for at least an hour before being overlaid with 10 ml of cold SM buffer and left at 4°C with gentle shaking for a further 2 hours. The lysate was then carefully removed using Pasteur pipettes, and centrifuged: i) for 10 minutes at 12,000*g* in a Sorvall SS-34 rotor at 4°C (to pellet debris); ii) for 30 minutes at 100,000*g* in a TH-641 swingout ultracentrifuge rotor (Sorvall) at 4°C (to precipitate the phage particles). After resuspension in 3 ml of 10 mM Tris (pH 8), the phage suspension was applied to the top of a DEAE-cellulose column (2 ml of preswollen Whatman DE-52 in a syringe barrel) that had been equilibrated with 10 mM Tris (pH 8). The column was then washed sequentially with: i) 3 ml of wash buffer (5 mM magnesium acetate, 30 mM sodium acetate, 10 mM Tris pH 8), ii) 1 ml of elution buffer (50 mM magnesium acetate, 10 mM Tris pH 8), and iii) 0.6 ml of elution buffer. The phage enriched fraction should be eluted following the addition of the second aliquot (iii) of elution buffer. Therefore, the runthrough resulting from the addition of i and ii was discarded whereas the liquid eluted following the addition of iii was carefully collected in an Eppendorf tube. The phage enriched fraction was treated with ribonuclease A (2 µl of a 10 mg/ml stock), and incubated at 37°C for 30 minutes. Proteinase K (15 µl of a 1 mg/ml stock) and 25 µl of 10% SDS were then added to the reaction, and incubation at 37°C was continued for another 15 minutes. Addition of 100 µl of 3 M potassium acetate, was followed by an incubation at 88°C for 15 minutes, after which the reaction was cooled on ice, then centrifuged at 12,000*g* for 15 minutes in a microcentrifuge in the cold room. The supernatant was carefully transferred to a new tube, where 700 µl of isopropanol was added to precipitate the DNA, and the reaction was left at -20°C for 1-2 hours. Finally, the precipitated DNA was pelleted by centrifugation at 12,000*g* for 15 minutes at 4°C, washed with 70% ethanol, dried, and then resuspended in 80 µl of 10 mM Tris pH 8.

3.2.4 Restriction enzyme digestion and Southern blotting of λ clone DNA

In general, ~1 μg of DNA from the λEMBL3 G4 clone, or ~2-3 μg of DNA from a λ2001 clone (for vector maps see Chapter 2 Figures 2.4 and 2.5), was digested in a total reaction volume of 20 μl containing the appropriate 10× reaction buffer and ~20 units of restriction enzyme (10 units/μl stock). The reactions were incubated for at least 4 hours at 37°C.

Southern blotting of DNA from the λ clones was carried out as described in Chapter 2. GeneScreen Plus membrane (NEN/Dupont) was used with neutral transfer conditions in the case of the λEMBL3 G4 clone, whereas Appligene Positive™ nylon membrane (Appligene Oncor) was used with alkaline transfer conditions for the Southern blotting of clones that had been isolated from the λ2001 mouse genomic library.

3.2.5 Probes

Rat Vipr2 cDNA derived probes that were used to isolate mouse genomic clones included: a 5' 260 bp probe which spans the region from bp 1to bp 260 of the rat cDNA; a 5' 1 kb probe which spans the region from bp 1 to bp 999 of the rat cDNA; and a full length rat cDNA probe which spans the entire published rat Vipr2 cDNA sequence (Lutz *et al.*, 1993). The 500 bp PCR product Ac(500)4 spans the region corresponding to bp 711 to bp 1192 of the rat cDNA sequence, and the PCR amplified cDNA derived 'gap' probe used to screen for intron 4 containing clones, spans the region corresponding to bp 222 to bp 553 of the rat cDNA sequence.

Q12, the mouse Vipr2 cDNA 5' RACE product, extends from 54 bp upstream of the published mouse Vipr2 cDNA sequence (Genebank accession number D28132; Inagaki *et al.*, 1994) to bp 192 of the published sequence.

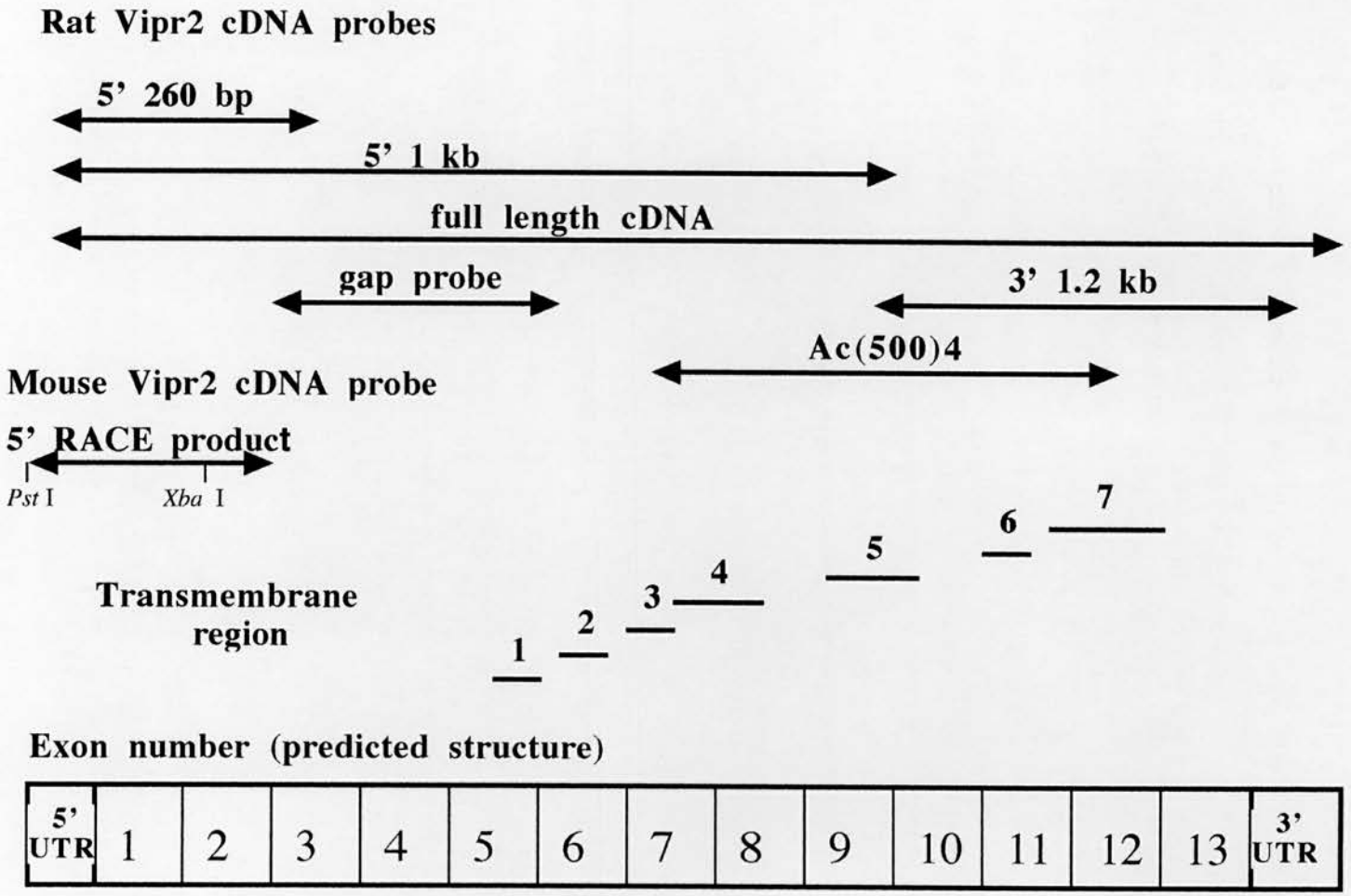
A schematic diagram of the relationship between these cDNA derived probes and the predicted exon structure of the mouse *Vipr2* gene is shown in Figure 3.1.

3.2.6 Rapid freezing method for transformation of bacteria

This one step method, which simply consists of a 1 minute incubation of the bacterial cell/plasmid mixture in liquid nitrogen, was published by Takahashi *et al.*, (1992), who developed it as a means of transforming non-competent cells with plasmid DNA. However, the authors also found that the use of this method could produce a tenfold increase in the number of transformants obtained from competent DH5$_{\alpha}$ cells, relative

Figure 3.1

Schematic diagram illustrating the relationship between the predicted exons of the mouse *Vipr2* gene and the probes used in the isolation and characterisation of clones containing the gene


Rat Vipr2 cDNA probes
5' 260 bp
5' 1 kb
full length cDNA
gap probe
3' 1.2 kb
Ac(500)4
Mouse Vipr2 cDNA probe
5' RACE product
Pst I
Xba I
Transmembrane region
1
2
3
4
5
6
7
Exon number (predicted structure)
5' UTR
1
2
3
4
5
6
7
8
9
10
11
12
13
3' UTR

to the number obtained following heat shock treatment. In this study the method was tried on competent XL-1 blue cells.

3.2.7 Plaque purification

The aim of this process was to re-isolate a single positive plaque from a phage stock that had up to this point been presumed to be homogeneous. In order to obtain phage plates from which single plaques could be easily isolated, the λG4 phage stock (titre 1.5×10^6 pfu/μl) was diluted by a factor of 10^6 using SM buffer, and adsorption reactions were then set up containing 250 μl, 150 μl, 100 μl, 50 μl, or 10 μl of the phage dilution + 500 μl of bacterial cell suspension (see 3.2.1, plate lysate method). Plating and growth of the phage was carried out as described in section 3.2.1 (plate lysate method), and plaque lifts were then taken from two of the phage plates (in this case those generated from 250 μl of phage dilution and 50 μl of phage dilution) as described in section 3.2.10. Following hybridisation of the filters (see 3.2.10) with the Ac(500)4 probe (see 3.2.5), two positive plaques were picked and each was placed in a separate Eppendorf tube containing 1 ml of SM buffer and a drop of chloroform. The tubes were left at room temperature for ~2 hours to allow the phage to diffuse out of the agarose. The phage suspensions were then used in the generation of two phage lysate stocks (from plate lysates, see 3.2.1), one of which was subsequently used to prepare a liquid lysate (see 3.2.1) from which λG4 DNA was isolated using a Qiagen lambda maxi-kit (Qiagen Ltd.) as described in 3.2.3.

3.2.8 Cycle sequencing of λ DNA using ^{32}P end-labelled primers

Cycle sequencing of the λG4 clone, and a control plasmid containing the 500 bp PCR product Ac(500)4 from the rat Vipr2 cDNA (see 3.2.5), was carried out using a *fmol*® DNA cycle sequencing system (Promega). The primer used in the reactions was 992W 5' GGGCAAAGTTGACTAC 3', which had been designed to bind to bases 253-238 of Ac(500)4. Sequencing of the plasmid and λ DNA was carried out according to the manufacturer's recommendations.

3.2.9 Dot blots of λ clone DNA

Dot blots of DNA from the λG4 clone were prepared on GeneScreen Plus nylon membrane (NEN/DuPont). A 6×6 grid of 1 cm^2 squares was lightly pencilled onto the membrane, and the membrane was then wet in 6× SSC for a few minutes before use. Two DNA solutions were prepared, one containing ~200 ng/μl of λG4 DNA in 6× SSC, and the other (a vector-only control) containing ~200 ng/μl of λEMBL3 vector arms in 6× SSC. Both DNA solutions were boiled for 10 minutes in a water

bath, to denature the DNA, immediately cooled on ice, and then centrifuged briefly. After residual 6× SSC had been allowed to drain from the nylon filter, the filter was placed on a lid from a small plastic box, so that only the edges of filter touched the lid. Eleven 2 µl samples of the λG4 DNA were then applied to the filter (two samples on each row of the grid, except for the last row), followed by four 2 µl samples of the λEMBL3 vector DNA control (one sample on each of the first four rows of the grid). The filter was then left for 10 minutes on a pad of 3 MM paper (Whatman) that had been soaked in denaturing solution (1.5 M NaCl, 0.5 M NaOH), before being moved to a pad of 3 MM paper (Whatman) that had been soaked in neutralising solution (1 M NaCl, 0.5 M Tris-HCl pH 7.0), where the filter was left for a further 10 minutes. After removal from the neutralising solution, the filter was left to dry at room temperature, and then baked at 80°C for 2 hours. Prehybridisation and hybridisation of the filter were carried out as described in Chapter 2 (Southern blot method).

3.2.10 Screening of a mouse ES-cell-derived genomic library

The mouse ES-cell-derived genomic library used in this study was constructed in the bacteriophage λ vector λ2001 (Karn *et al.*, 1984; see Figure 2.5 for vector map), and was provided by Dr. H. Cooke (MRC Human Genetics Unit).

The preparation of the host bacteria (*E. coli* strain Q358) that were used in the plating of the phage library, timing and temperature conditions for the phage adsorption steps, preparation of B-agar plates and soft top-agarose, and subsequent plating of the infection reactions, were carried out as described in section 3.2.1 (plate lysate method) unless otherwise stated below.

Plating of phage for primary screens

On receipt of the library its titre was determined as described in 3.2.2 and was found to be ~10^9 pfu/ml. For primary screens of the library a 1/2500 dilution of the library stock was made with SM buffer (to give a phage suspension with ~4×10^4 pfu/100 µl), and a series of 20 adsorption reactions were set up. Each adsorption reaction contained 100 µl of the phage dilution + 600 µl of bacterial host cell suspension, and was thus expected to lead to the generation of ~4×10^4 plaques. Following adsorption, the phage/bacteria infection reactions were plated onto 20 large (140 mm diameter) B-agar plates, and once the agarose had set, the plates were inverted and incubated at 37°C until the plaques were almost confluent (with only a thin web of bacterial growth still visible). The plates were then removed from the incubator, sealed with parafilm, and stored in the cold room overnight.

Plaque lifts

Plaque lifts were carried out using BioBlot-NC nitrocellulose blotting membrane 132 mm discs (Costar®). Two duplicate filters were prepared from each of the plates. The filters were labelled in pencil on the 'wrong side', to indicate which plate they represented, and whether they were the first or second plaque lift from that plate. The first filter was then placed on the surface of the appropriate agar plate, and was left in contact with the plaques for 1 minute, while orientation marks were applied to both the plate and the filter (in positions corresponding to 12, 6, and 8 o'clock) using an 18 gauge needle. The filter was then transferred sequentially to trays containing: i) denaturing solution (1.5 M NaCl, 0.5 M NaOH) in which the filter was left to float for 30 seconds before being immersed and left for a further 1 minute; ii) neutralising solution (1.5 M NaCl, 0.5 M Tris pH 8.0) in which the filter was immersed for 5 minutes; iii) a wash solution consisting of 3× SSC, in which the filter was left for a few minutes before being transferred to 3MM paper (Whatman) where it was left to dry. A second filter was then placed on the same phage plate, and was left in contact with the plaques for 3 minutes. Orientation marks were made on the second filter at identical positions to those made on the plate and the first filter, and the filter was then passed through the denaturation, neutralisation, and wash steps, described in i-iii. This process was repeated for each of the phage plates, and the filters were then placed between two sheets of 3MM paper (Whatman) and baked in an oven for 1.5 hours at 80°C.

Hybridisation

Prehybridisation and hybridisation of the filters were carried out in a solution containing: 50% deionised formamide, 6× SSPE, 5× Denhardt's solution, 0.5% SDS, and 100 µg/ml of single stranded herring testis DNA (added just prior to prehybridisation). The filters were divided into two sets, one set comprising the first lift that was taken from each of the plates, and the other set comprising the second lift from each plate. Each set of filters was placed in a plastic bag to which ~100 ml of prehybridisation solution was added, and the bags were then sealed. For prehybridisation, the bags were placed in a large plastic box containing water that had been preheated to 42°C, and the box was then placed on a gently rotating shaker in an incubator that had also been preheated to 42°C. Probe labelling was carried out as described in Chapter 2, and after prehybridisation had been allowed to proceed for at least an hour, the probe was boiled in a water bath for 5 minutes, cooled immediately on ice, and then added to the filter bags (which were subsequently resealed). Following an overnight hybridisation at 42°C, the filters were removed from the bags,

carefully separated, and then washed at 50°C with gentle shaking, first in a solution containing 2× SSC, 0.1% SDS, and then in a solution containing 0.5× SSC, 0.1% SDS. Both solutions had been preheated to 50°C before use, and two separate 20 minute washes were carried out in each solution. After these washes, the filters were left to dry at room temperature for a few minutes before being taped onto sheets of 3MM paper (Whatman) that had been cut to fit the inside of autoradiography cassettes. The orientation marks on the filters were highlighted with fluorescent paint, and each sheet of paper and its attached filters were then covered with clingfilm, placed in an autoradiography cassette with two intensifying screens and Fuji RX film (Fuji), and stored at -70°C overnight.

Picking positive plaques

Once the autoradiographs had been developed, autoradiographs from duplicate filters were aligned over a light box, so that the positions of positive clones could be identified. If positive clones were seen, the appropriate phage plate was placed on top of the autoradiographs on the light box, and the orientation marks on the plate were aligned with those on the autoradiographs of the filters from that plate. Following primary screens, positive plaques were picked using the wide end of a yellow (200 µl) pipette tip, and following secondary or tertiary screens, positive plaques were picked using the fine end of a glass Pasteur pipette. Each agarose plug that was selected was placed in a separate 1.5 ml sterile Eppendorf tube that contained 250 µl of SM buffer and 10 µl of chloroform. The tubes were left at 4°C overnight to allow the phage to elute from the agarose plugs, and were then stored at this temperature until required.

Plating of phage for secondary screens

For each positive clone that had been picked in the primary screen of the library, 1 in 10^6 and 1 in 3×10^6 dilutions were prepared from the phage that had eluted from the agarose plug. Two adsorption reactions containing 100 µl of phage dilution + 600 µl of bacterial host cell suspension were set up for each positive clone (one reaction for each of the phage dilutions). The reaction mixtures were plated onto B-agar plates, and once the agarose had set, the plates were inverted and incubated at 37°C until plaques were clearly visible. One of the two plates that had been produced from each primary screen positive was then selected for use in the secondary screen. The aim was to try to maximise the chances of obtaining positive clones on the plate used for the secondary screen while trying to ensure that if positives were identified they could be isolated as single plaques (ie. the density of plaques on the plated should be high but not so high that the plaques are converging), and in most instances the plates from the 1 in 3×10^6

dilutions of the phage were selected. Plaque lifts, hybridisation, and the selection of positive plaques from the secondary screens, were carried out as described for the primary screens of the library. Tertiary screens were not usually necesary.

Preparation of phage lysates, and subsequent isolation of DNA from positive clones, were carried out as described in sections 3.2.1 and 3.2.3.

3.2.11 Direct automated fluorescent sequencing of λ DNA

Direct sequencing of λ clone DNA was carried out using an ABI 373 automated fluorescent sequencer and an 'ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq® DNA polymerase, FS (Perkin Elmer). Lambda DNA for use in sequencing was isolated using Qiagen lambda maxi- or midi- kits (see 3.2.3), resuspended in ddH_2O (~100 µl in the case of midipreps and 300 µl in the case of maxipreps), and 5 µl (~1-2 µg) of the DNA was then used as a template in each sequencing reaction.

Sequencing reactions were carried out according to the manufacturer's instructions, with the following exceptions: i) The amount of template DNA used was greater than the recommended amount (0.25-0.5 µg) for double stranded templates, but the increase was found to give better sequencing results from λ DNA, presumably due to the increase in the concentration of target binding sites, as this would be substantially lower in a given weight of a λ DNA clone than in the same weight of smaller molecular weight plasmids or PCR products; ii) all sequencing reactions were heated to 90°C for 5 minutes (as opposed to the recommended 2 minutes) as this was also found to lead to better results.

The sequencing primers used in these reactions were designed using the 'Primer' software package (HGMP Resource Centre, Cambridge, UK). To design primers that would bind to the mouse *Vipr2* gene, the mouse Vipr2 cDNA sequence was used as a template for the primer design program, and the entire sequence was searched for the presence of possible 'forward or 'reverse' primer binding sites for primers that were at least 18 bp long and had a melting temperature that was between 57°C and 63°C. Using a prediction of the mouse *Vipr2* gene's structure (that had been derived from an alignment of the mouse Vipr2 cDNA sequence with the exon sequences of members of the receptor gene sub-family whose structures had already been determined by other groups), a few primers were initially selected that would bind to exons that lay towards the 5' and 3' ends of the *Vipr2* gene. However, as the analysis of the *Vipr2* gene progressed additional primers were selected and new primers were designed until

eventually there were enough primers to sequence all the known intron/exon boundaries of the *Vipr2* gene, and sequence across some of its introns. The PCR primers used in direct sequencing of the mouse *Vipr2* gene clones (or in the subsequent PCR amplification of *Vipr2* gene introns), are listed in Table 3.1. Primers 8484 and 7363 (also included in Table 3.1) are primers that were designed to bind to sequences in the left and right arm respectively of the λ2001 vector (see Figure 2.5 for vector map).

3.2.12 PCR amplification of mouse *Vipr2* gene introns

From the possible primers that were identified in the initial 'Primer' search of the mouse Vipr2 cDNA sequence, and other sequencing primers that had been subsequently designed to sequence the *Vipr2* gene (see 3.2.11), pairs of primers were selected, in which the individual primers were expected to bind to opposite strands of adjacent exons of the *Vipr2* gene, and lead to the amplification of the intervening intron when used in PCR reactions containing the *Vipr2* gene as template. Twelve pairs of primers were required initially (one pair for each intron), and although in some cases it proved necessary to design new primers for use in PCR (either because of potential primer dimer formation or because of large differences in the melting temperatures of the primers) in the vast majority of cases the primers used were those that had already been successfully used in sequencing of the *Vipr2* gene λ clones. The PCR primers used in the direct sequencing of the mouse *Vipr2* gene clones, or in the PCR amplification of *Vipr2* gene introns, are listed in Table 3.1.

For each PCR reaction 1 µl of midiprep DNA from an appropriate λ clone was used as a template in a reaction mix containing: 2 µl of 10 mM dNTP mix (Boehringer Mannheim), 10 µl of magnesium-free 10× PCR reaction buffer (Promega), 4 µl of 50 mM $MgCl_2$, 2 µl of each of the two primers (15 µM stock solutions), 78 µl of tissue culture grade water (Sigma), and 1 µl of *Taq* DNA polymerase (Promega). The reaction components were mixed by vortexing, then spun briefly in a microcentrifuge. The reactions were then carefully overlaid with two drops of light white mineral oil (Sigma), and the tubes were placed in the PCR machine. After an initial denaturation at 94°C for 5 minutes, the reactions underwent 35 cycles consisting of: denaturation at 94°C for 1 minute, annealing at 57°C for 1 minute, and extension at 72°C for 1.5 minutes + 4 seconds/cycle. Following a final extension period at 72°C for 10 minutes, the reactions were cooled to 4°C and maintained at this temperature until they could be collected. Analysis of the PCR products was carried out by electophoresis of 20 µl of each reaction through a 0.8% agarose TBE gel.

Table 3.1

PCR and sequencing primers used to characterise the mouse *Vipr2* genomic clones were designed using the published mouse cDNA sequence (Inagaki *et al.*, 1994) as a template where possible (i.e. for exon primers). The number given in the far left hand column 'Primer' is the number that has been used to refer to that primer in the main text. The primer sequences are presented in a 5'→ 3' direction. Where applicable the position of the primer within the published mouse cDNA sequence is given in the far right hand column, and in these cases the direction in which the primer faces when bound to its target sequence within the *VIPR2* gene can be deduced from order of the numbers given in this column, i.e. if the site in the mouse cDNA is given as 1-20 then the primer will bind to the top strand of the mouse cDNA sequence (shown in Figure 3.10) and sequence derived from the primer will extend downstream towards the 3' end of the mouse cDNA, whereas if the sequence is given as 20-1, the primer will bind to the bottom strand of the mouse cDNA sequence and will extend upstream towards the 5' end of the mouse cDNA sequence.

Primers marked with an asterisk (*) are designed to bind to the λ2001 vector arms, and the superscript [a] has been used to denote that primer 4336 was originally designed to bind to human as opposed to mouse Vipr2 cDNA sequences, and contains a two bp mismatch in comparison to the mouse target sequences.

Primers used in the characterisation of the mouse *Vipr2* gene

Primer	Number of bases	Primer sequence (5'→ 3')	*Vipr2* gene exon or intron number	'Site' in mouse cDNA
3755	22	TGGGACAACATCACATGCTGGC	exon 3	215-236
4340	20	CAGATGTTGGTGGCAATGAC	exon 10	987-1006
4341	22	CAAGTCCACACTGCTGCTAATC	exon 11	1030-1051
4342	23	GTATAGAACTGCTACCACCAGGC	exon 12	1174-1152
4343	22	GATTAGCAGCAGTGTGGACTTG	exon 11	1051-1030
4344	20	GCCTCGCCATCTTCTTTTCA	exon 13	1225-1206
4336	20	AGTGGCGTCTGGGACAACAT	exon 3	206-225[a]
5990	20	GGCCTGGTGGTAGCAGTTCT	exon 12	1151-1170
5991	22	GGCCATTTATACCTTGGGCTAC	exon 5	430-451
5992	20	GCAGTAGCAGGTCAGCACCA	exon 1	85-66
5993	21	TCTCCTTCATGCTGAGAGCCA	exon 6	549-569
7122	20	TCTGATTCTCCGTTTGGCTG	exon 2	197-178
7363	22	AGAAGTCCAACCCAGATAACGA	N/A*	N/A*
7853	20	CTAAGCAGCCAAACGGAGAA	exon 2	173-192
7854	22	GCCCAAGGTATAAATGGCCTTC	exon 5	448-427
7890	20	TCTGGATGGATGCTGCTCAC	exon 2	120-101
7891	20	CTGAATACTTTGGGGCAGGG	exon 3	288-269
7892	20	AGAGAGGCGAGTTGCTGTCC	exon 8	844-825
7893	20	GTGAGCAGCATCCATCCAGA	exon 2	101-120
7894	20	TACTGGCTTCTGGTGGAGGG	exon 7	701-720
7895	20	CACAAACGACCACAGCATCC	exon 9	868-887
7896	20	ACAATGCTGATGAAGAGGGC	exon 10	957-938
7897	20	ATGCTGTGGTCGTTTGTGTC	exon 9	885-866
7898	21	ATGTCTCTGACCATCCATCGC	exon 4	353-333
7899	22	CACTAGCGATGGATGGTCAGAG	exon 4	328-349
7900	23	GCAGTAGACCTGAGCTGGAGTAG	exon 6	620-598
7901	23	GACAGCAACTCGCCTCTCTTTAG	exon 8	826-848
7992	20	AGTGCTGCTGGGTTTGACAT	3' UTR	1391-1410

Primer	Number of bases	Primer sequence (5'→ 3')	*Vipr2* gene exon or intron number	'Site' in mouse cDNA
8482	20	CACTCCCTGCCTCTGTCATC	N/A*	N/A*
8483	18	CAAGGACCGAGGCAGCAC	exon 1	23-40
8484	20	GCACCAGCAACCAGCAGTAG	exon 1	98-79
9200	19	TGGGAGGAAGGATGGCTAC	exon 7	761-743
9628	23	AGGAGAGGAAGAGGTTTAGGTGG	exon 6	554-532
9629	22	AGGTGGATGTAGTTCCTTGTGC	exon 6	537-516
9630	21	GGACAGCGTGCTCTACTCCAG	exon 6	586-606
14367	25	CTGGCCTTTTGTGTCTAAACCTGCT	intron 10	N/A
14368	25	GATAATTCTGGTGGGACTGGAAAGG	intron 10	N/A
14400	25	CTACTTCCCTCTGAAGTGCAGGACC	5'	N/A
14891	25	AAGATTATGCTGAAAGGACCCTGAT	5'	N/A
14892	25	CCTCTCCTGTAAAATGGAATGGATG	intron 10	N/A
15026	25	ATCACTAAGCACACAGCCTATCTCC	intron 10	N/A
15075	23	GTAGCCCCTAATGGGTCGAGTGT	5'	N/A
15132	24	GGTATGGGGGGACTTTTGGTATAG	5'	N/A
15133	25	GTGAAGAAGGTCTTTGGATGGATAC	intron 10	N/A
15585	23	TGCTTTGGTCATGGTGTCTGTTC	intron 10	N/A
15586	23	CCAACCCTGATAGAACCCAACAT	intron 10	N/A
15587	23	ATTGCTGACCCAAAACTTCACTG	5'	N/A
15588	23	ATCTGCCCATTATCTTTCTCCCC	5'	N/A
17982	22	TTGCTTGTCTTAGCTCTTTGGC	intron 4	N/A
17983	20	GGTAGGACAAGGCCACAGGA	intron 2	N/A
17984	21	GAGCCTCCTTCTTCCCTGATG	intron 2	N/A
18439	22	GAGACATGGAAGACTCGAAAGC	intron 5	N/A
18440	24	ACAGAAAGTAAATCCTCAGTGGTC	intron 2	N/A
18441	23	CAGGGTAAGGACATTGAGTCCAG	intron 2	N/A
19458	26	TAGTAACCTACATCTGACCAAAGTCC	intron 5	N/A
19459	20	GCAGGATGTCTCATGCCAGT	intron 2	N/A
19460	24	CAGTGTTGCTACCATAAGTAATCC	intron 2	N/A

Primer	Number of bases	Primer sequence (5'→ 3')	*Vipr2* gene exon or intron number	'Site' in mouse cDNA
19461	20	CAGTAGCCTGGGTATGCTGG	intron 6	N/A
19773	23	TACAGGGAATTACTGCTGGATGG	intron 2	N/A
19774	23	ATATTACCCAGCTCTGCCCAGTC	intron 2	N/A
19775	23	GATTCCTACCTGTCAATGGAAGA	intron 4	N/A
20847	23	CCCAGAAGAAAGCAAGAGATTCC	intron 5	N/A
20848	22	GGGGGAAACAAACACAGTTGAC	intron 5	N/A
22567	23	GTGTATGCTCATCTACGAAAGGA	intron 6	N/A
22568	23	GCTAGACACCAGTCAGCCTTCTG	intron 6	N/A
22569	22	AATGTCTGTCCAAGTCAGCACG	intron 6	N/A
22570	21	TTATCCAAACAGATCGCTGCC	intron 6	N/A

3.2.13 Preparation of PCR products for direct automated fluorescent sequencing

PCR products that were to be sequenced directly (i.e. those that were not subcloned into plasmids before being sequenced), were purified using Microcon™-100 microconcentrators (Amicon), according to the manufacturer's instructions for the purification of PCR products.

3.2.14 Long PCR

Attempts at PCR amplification of introns 2 and 4 were carried out using the Expand™ long template PCR system (Boehringer Mannheim). The primers used in attempts to amplify intron 2 using λ clone ES9 as template DNA were 13896 5' CCATCCAGAATGTCGCTTTCATCTAGAAATAC 3' and 13897 5' GTAGAAATTGCTGAATACTTTGGGGCAGG 3', and those used in attempts to amplify intron 4 from mouse total genomic DNA (Promega) were 13898 5' GGAAACATAAGCAAAAACTGCACTAGCGAT 3' and 13899 5' GTAAGAGACATCAGAGAAACACTGTAGCCCAAG 3'.

PCR reactions were carried out in 0.5 ml thin-walled GeneAmp reaction tubes (Perkin Elmer). PCR buffer 1 (10× concentration, with 17.5 mM $MgCl_2$) which was used in the reactions, is the recommended buffer for the amplification of human genomic DNA fragments of 0.5-12 kb or λ DNA fragments of 0.5-25 kb, and all PCR reactions were set up according to the manufacturer's instructions for the amplification of fragments within these size ranges (PCR system 1; Boehringer Mannheim).

An initial denaturation at 94°C for 2 minutes, was followed by ten cycles of: denaturation at 94°C for 10 seconds, annealing at 65°C for 30 seconds, and elongation at 68°C for 15 minutes. This in turn was followed by fifteen cycles of: denaturation at 94°C for 10 seconds, annealing at 65°C for 30 seconds, and elongation at 68°C for 15 minutes +20 seconds/cycle. After a final extension phase at 68°C for 7 minutes, the reactions were cooled to 4°C and were maintained at this temperature until they could be collected. Analysis of the PCR products was carried out by electrophoresis through a 0.8 % agarose TBE gel.

3.2.15 'Shotgun' subcloning of λ clone fragments from 'long introns'

Generally, this method was used in cases where several different restriction fragments had to be subcloned from a single λ clone. However, it was also occasionally used

when attempts to clone a single restriction fragment using other methods (e.g. those involving the isolation of DNA from an agarose gel slice) had failed.

In most cases, two restriction digest reactions were set up, in which approximately 5 µg of DNA from the λ clone of interest, was digested in a total reaction volume of 40 µl, containing the appropriate restriction enzyme(s) and 10× reaction buffer. After incubation at 37°C for 6-8 hours, the two digests were pooled, and the restriction fragments were purified using a Wizard™ DNA clean-up system (Promega) according to the manufacturer's instructions. The λ DNA fragments were eluted from the Wizard™ column by the addition of 40 µl of ddH_2O (preheated to 65°C). Plasmid vector DNA (usually pGEM®-7Zf(+), Promega; see Figure 2.2) was prepared as described in Chapter 2. Various ratios of plasmid: λ DNA were tested in a series of ligation reactions, which in addition to λ DNA and vector, contained 1 µl of T4 DNA ligase (Promega), 1 µl of 10× ligase buffer (Promega), and ddH_2O to 10 µl. The ligation reactions were incubated at 4°C overnight, and each ligation was then used to transform competent bacterial cells (as described in Chapter 2).

Following the use of this 'shotgun' subcloning method, it was often necessary to select more transformants than normal for plasmid minipreps, particularly if the target fragments were relatively large in size, as subclones containing the smaller products from the λ DNA digest tended to predominate.

3.2.16 PCR amplification of plasmid insert fragments

PCR amplification of *Vipr2* gene intron 5 and intron 6 fragments, from the ES11 derived plasmid subclone Xho (2), was achieved using two separate pairs of PCR primers that each consisted of: one plasmid vector primer, and one internal *Vipr2* gene primer (that bound to sequence from exon 6 of the *Vipr2* gene).

An M13/pUC reverse primer (GIBCO) was used in conjunction with the primer 9629 5' AGGTGGATGTAGTTCCTTGTGC (exon 6) to amplify the intron 5 derived sequence from the Xho (2) subclone, and an M13/pUC forward primer (GIBCO) was used together with primer 9630 5' GGACAGCGTGCTCTACTCCAG 3' (exon 6) to amplify the intron 6 derived portion of the plasmid's insert. Xho (2) plasmid DNA was prepared using a Wizard™ 373 plus DNA purification kit (Promega), and for each PCR reaction 1 µl (~0.5 µg) of the plasmid miniprep was used as a template in a reaction mix containing: 2 µl of 10 mM dNTP mix (Boehringer Mannheim), 10 µl of magnesium-free 10× PCR reaction buffer (Promega), 4 µl of 50 mM $MgCl_2$, 2 µl of the appropriate M13/pUC primer (10 µM stock), 6.5 µl of the appropriate exon 6 primer

(3.2 μM stock), 1 μl of *Taq* DNA polymerase (Promega) and tissue culture grade water (Sigma) to 100 μl. The reactions were vortexed, spun briefly in a microcentrifuge, then overlaid with light white mineral oil (Sigma). Thermocycling and subsequent analysis of the PCR products were carried out as described in section 3.2.13.

3.3 Results

3.3.1 Restriction enzyme digestion and Southern blot analysis of the putative *Vipr2* mouse genomic clone λG4

Initial restriction enzyme digests carried out on λG4 DNA that had been prepared using a plate lysate derived phage suspension and a Qiagen lambda maxi-kit (see 3.2.1 and 3.2.3), gave surprisingly poor results. Despite the relatively high concentration of restriction enzyme used in the digestion reactions and the fairly long incubation time of these reactions (see 3.2.4), it seemed that only partial digestion of the DNA was occurring (as indicated by the disproportionate weakness of some of the restriction fragment bands relative to their molecular weight when the digest products were analysed on an agarose gel; see Figure 3.2 for example). Possible explanations for this included: the presence of impurities in the DNA preparation, too short an incubation time, the presence of insufficient restriction enzyme in the digestion reactions, or the presence of too much restriction enzyme in the digestion reactions -which can lead to star activity with some enzymes. However, although various different sets of reaction conditions were tried, using both the original Qiagen column isolated λG4 DNA and a new preparation of λG4 DNA that had been isolated from DEAE-cellulose column purified phage by proteinase K-SDS treatment (see 3.2.3), satisfactory restriction enzyme digests of the λG4 DNA were not obtained.

Southern blot analysis of a *Bam*H I digest of the λG4 clone, clearly identified a single hybridising band of ~4 kb following hybridisation with the rat Ac(500)4 probe (Figures 3.3a and 3.3b). Unfortunately, the strong hybridisation signal on the autoradiograph originated from a band of DNA that seemed to be significantly underrepresented in *Bam*H I digests of the λG4 clone, and in some digests of the clone this 4 kb fragment and a smaller *Bam*H I restriction fragment of < 2 kb were only just visible on the 1% agarose gel (see Figures 3.3a and 3.3b). Nevertheless, since it appeared that the isolation of this 4 kb *Bam*H I fragment would be likely to lead to confirmation that the λG4 clone contained *Vipr2* gene sequence, and provide a starting point for the analysis of the mouse *Vipr2* gene, attempts to subclone this restriction fragment were pursued.

Figure 3.2

Gel showing restriction enzyme digests of clone λG 4

Under-represented bands are intermingled with stronger bands

λ *Hind* III markers
λG4 *Eco*R I
λG4 *Bam*H I
λG4 *Eco*R I/*Xho* I
λG4 *Eco*R I/*Cla* I
λG4 *Eco*R I/*Pst* I
λG4 *Eco*R I/*Xba* I
λG4 *Eco*R I/*Not* I
λG4 *Eco*R I/*Eco*R V
23.1 kb
9.42 kb
6.56 kb
4.36 kb
2.32 kb
2.03 kb

Figure 3.3a

_Bam_H I restriction digest of λG4 DNA separated on an agarose gel in preparation for Southern blotting

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane)

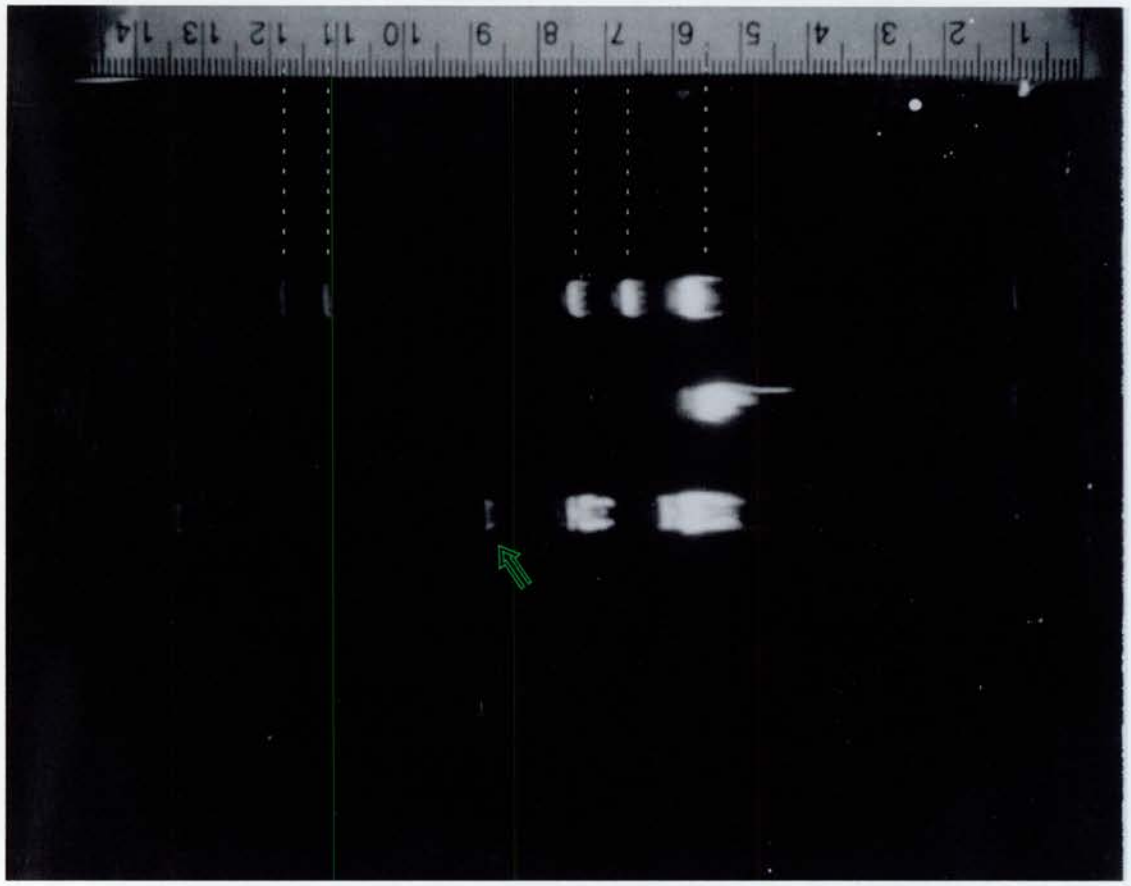
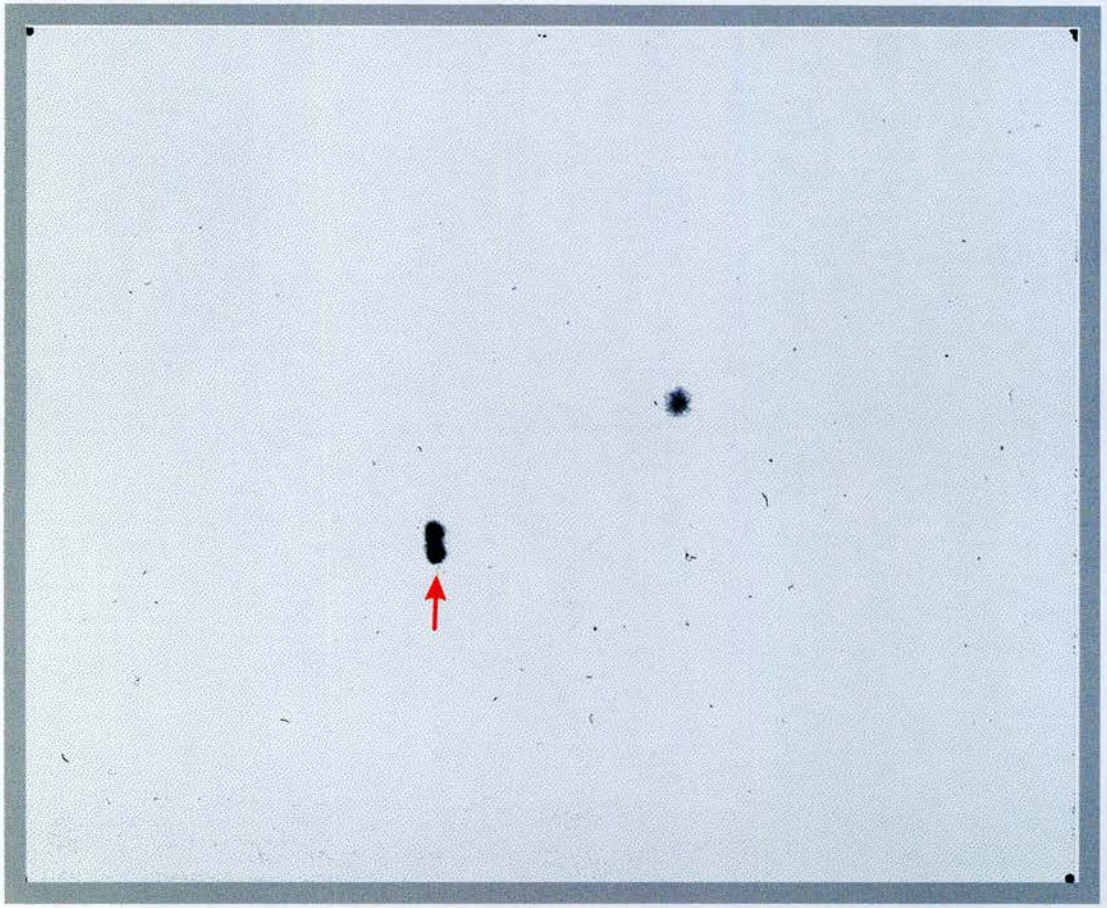

23.1 kb
9.42 kb
6.56 kb
2.32 kb
2.03 kb
λ *Hin*d III markers
λG4 uncut
λG4 *Bam*H I

Figure 3.3b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested λG4 DNA shown in Figure 3.3a is incubated with the radiolabelled Vipr2 rat Ac(500)4 cDNA probe

The faint 4 kb *Bam*H I restriction fragment arrowed in 3.3a hybridises strongly to the probe

(The autoradiograph is shown at actual size)



λ *Hin*d III markers

λG4 uncut

λG4 *Bam*H I

3.3.2 Attempts to subclone the 4 kb *Bam*H I restriction fragment from the λG4 clone

As anticipated, the subcloning of the λG4 4 kb *Bam*H I fragment proved to be difficult. The very limited amount of this DNA fragment that could be isolated in an agarose gel slice following electrophoresis of λG4 *Bam*H I digests, was the main problem, and even when gel slices from several lanes were pooled prior to purification of the DNA from the agarose, very little if any DNA could be visualised when 4 µl out of a total of 20 µl of the GeneClean® purified restriction fragments (see Chapter 2) was run on a 1% agarose mini gel. In cases where the 4 kb restriction fragment was visible on the gel, ligation reactions with Bluescript SK plasmid (Stratagene) were set up (see Chapter 2), and after initial subcloning attempts in which the standard heat shock protocol was used (see Chapter 2) had failed to generate colonies containing the 4 kb subclone of interest, a recently published 'freeze shock' transformation method was tried (Takahashi *et al.*, 1992; see 3.2.6). However, although the latter method seemed to result in an approximately 2-fold increase in the number of transformants obtained relative to the numbers seen following use of the 'heat shock' based method, transformants containing plasmids with the 4 kb *Bam*H I insert fragment were not obtained.

3.3.3 Further Southern blot analysis of the λG4 clone

The failure of attempts to subclone the 4 kb *Bam*H I restriction fragment from the λG4 clone, led to efforts to identify other restriction fragments that hybridised to the Ac(500)4 probe, and in particular, hybridising fragments that were more abundant within λG4 DNA digests than the previously identified 4 kb *Bam*H I fragment. A Southern blot was therefore prepared from a gel on which *Cla* I, *Eco*R V, and *Pst* I digests of λG4 DNA had been separated (Figure 3.4a), and the blot was then hybridised with the Ac(500)4 probe. The results of this hybridisation clearly showed that in the lanes containing *Cla* I or *Eco*R V digested λG4 DNA, only undigested or large molecular weight fragments near the top of the gel, hybridised to the probe, whereas in the lane containing *Pst* I digested λG4 DNA, the probe had hybridised only to a single band of just under 4 kb (Figures 3.4a and 3.4b). The fact that hybridisation had occurred only to this single band within the *Pst* I digest, and was quite intense, ruled out the possibility of partial digestion as the cause of any underrepresentation of this band within the digest, as if partial digestion had occurred, hybridisation to other fragments in the digest or to uncut DNA would have been expected. Nevertheless, when the band from which the hybridisation signal had originated was located on the

Figure 3.4a

Restriction digests of λG4 DNA separated on an agarose gel in preparation for Southern blotting

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane)

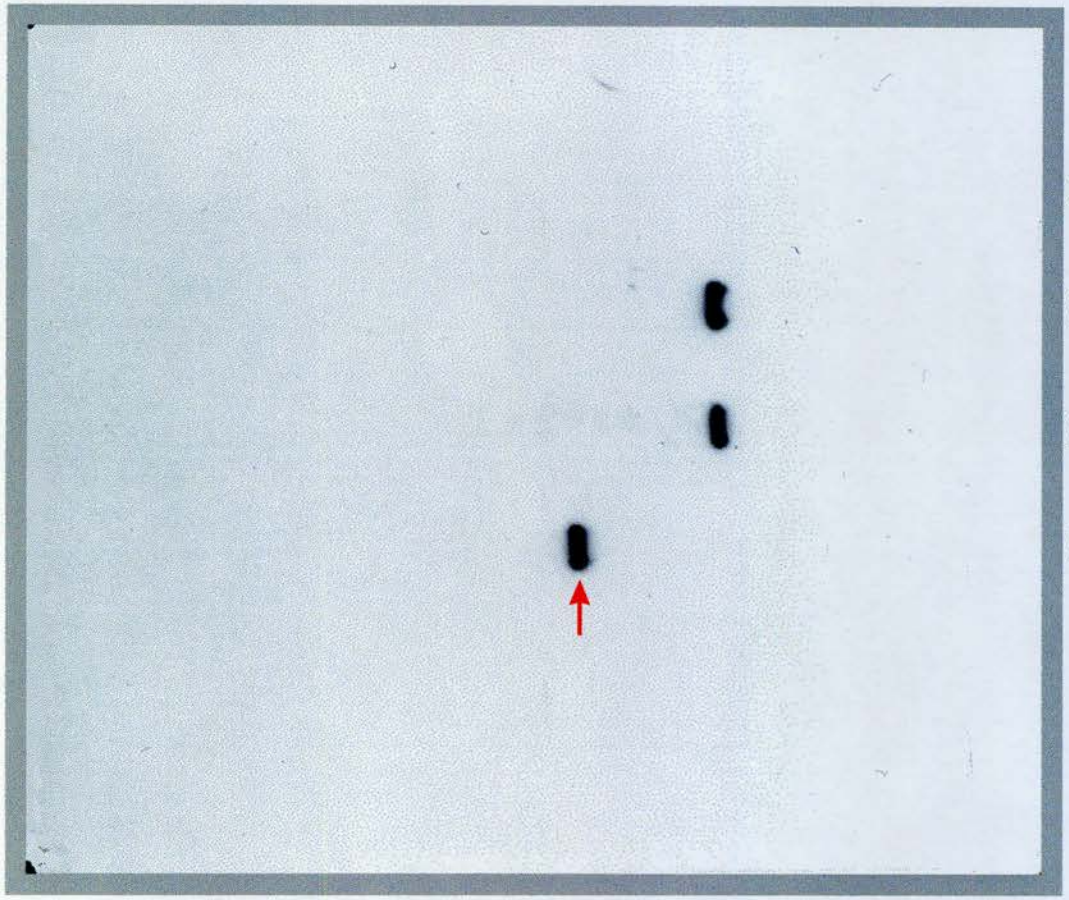
λ Hind III markers
λG4 Cla I
λG4 EcoR V
λG4 Pst I
23.1 kb
6.56 kb
4.36 kb
2.32 kb
2.03 kb

Figure 3.4b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested λG4 DNA shown in Figure 3.4a is incubated with the radiolabelled Vipr2 rat Ac(500)4 cDNA probe

The under-represented band of just under 4 kb (arrowed in Figure 3.4a) hybridises strongly to the probe.

(The autoradiograph is shown at actual size)


λ *Hind* III markers
λG4 *Cla* I
λG4 *Eco*R V
λG4 *Pst* I

photograph of the agarose gel that was taken prior to Southern blotting (Figure 3.4a), it was clear that this *Pst* I restriction fragment was underrepresented to an even greater degree than the previously identified 4 kb *Bam*H I fragment, and consequently there seemed little point in trying to subclone it.

3.3.4 Plaque purification of the λG4 clone

The identification of two λG4-derived restriction fragments that hybridised to the Ac(500)4 probe, but were both significantly underrepresented in restriction digests, was a cause for serious concern. However, as it now appeared that partial digestion of the λG4 DNA was not likely to be the cause of these problems, other potential causes had to be considered, including the possibility that a contaminating phage might have become predominant in the phage stocks. After the SM buffer and bacterial stocks had been checked to ensure that they were not the source of any contaminating phage or plasmid DNA, the decision to carry out a plaque purification of the λG4 phage stock, was finally made. The results subsequently obtained from the hybridisation screening of the λG4 phage stock during the plaque purification process (see 3.2.7 for methods), were strange, and difficult to interpret. Of the two phage plates that were used to take plaque lifts for use in the hybridisation screening of the phage stock, one phage plate had been prepared using one fifth of the number of phage that had been used in the preparation of the second plate. Following hybridisation of the plaque lift filters with the Ac(500)4 probe, the hybridising plaques from the plate that had the greater density of plaques (Figure 3.5), were found to represent ~77% of the total number of plaques on that plate, a value which, allowing for inefficiencies in the plaque lift and hybridisation procedures, might be consistent with that expected from stocks of plaque purified phage clones. In contrast, the hybridising plaques from the plate with the lower density of plaques, were found to represent only ~14% of the total number of plaques on the plate. However, the rather odd background pattern seen on the filters from the lower density plaques (Figure 3.5), suggested that factors affecting the plaque lift process from this plate may have been responsible for the low levels of hybridisation in this case, and consequently the plaque purification process was continued (see 3.2.7) following selection of two positive plaques from the plate that had the higher plaque density.

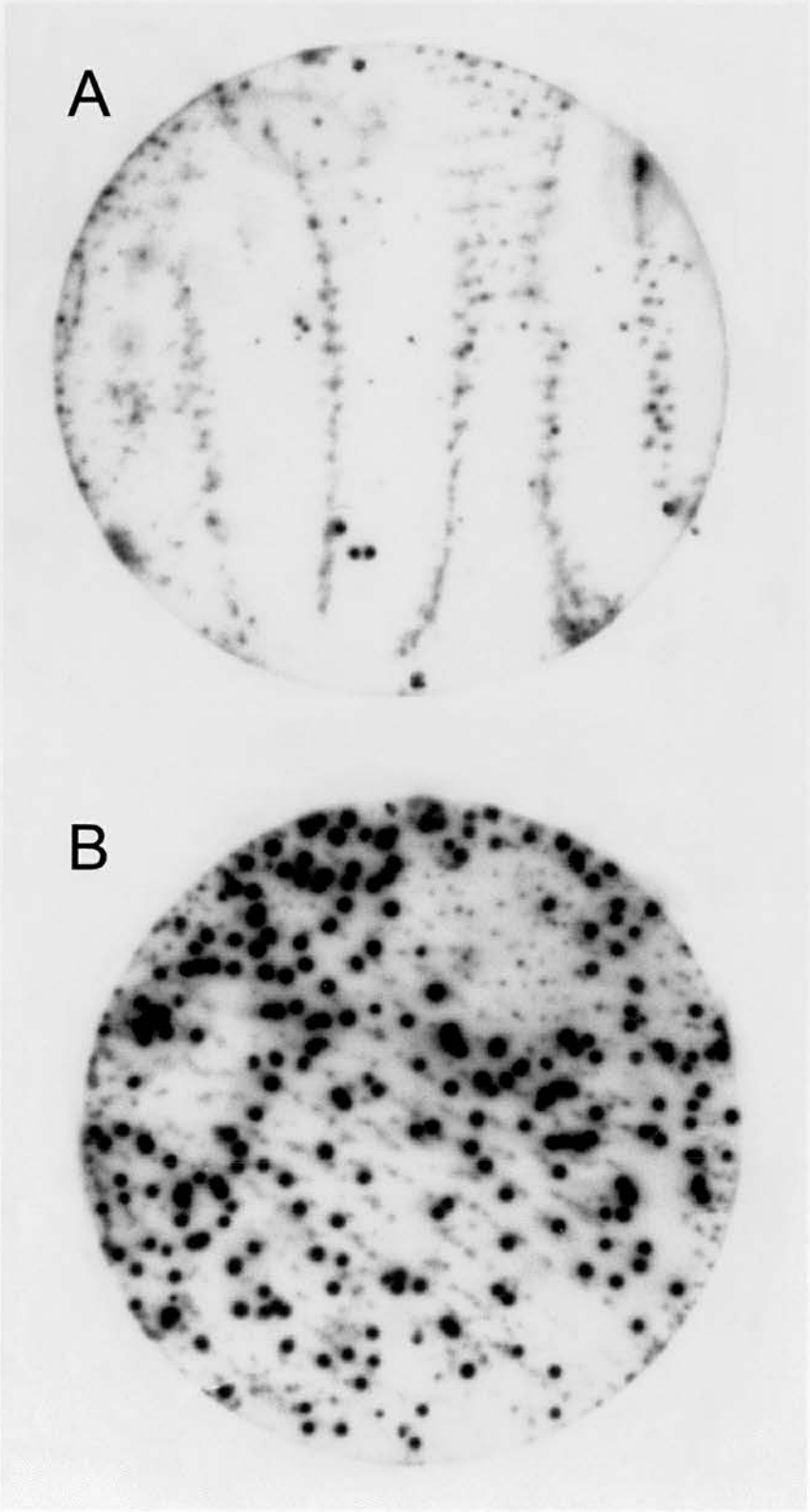
3.3.5 Attempt at direct sequencing of λ DNA using a ^{32}P end-labelled primer

Direct sequencing of λG4 DNA using the *fmol*® DNA cycle sequencing system (Promega; see 3.2.8), was tried as an alternative approach to confirming that the λG4

Figure 3.5

Autoradiographs of filters from the plaque purification plates, following hybridisation with the radiolabelled Ac(500)4 cDNA probe

(A=low plaque density filter, B=higher plaque density filter)


A
B

clone contained mouse *Vipr2* gene sequence, following the failure of subcloning attempts. The primer (992W) that was selected for use in these sequencing reactions was chosen because it had been shown to bind to the central region of the rat Vipr2 cDNA probe Ac(500)4 that was originally used to isolate the λG4 clone. Unfortunately, this approach proved to be unsuccessful, as no sequence data was produced from λG4 template DNA in sequencing reactions using this primer, although the generation of a very small amount of sequence data from the Ac(500)4 insert of the control plasmid demonstrated that the primer was priming correctly in this sequencing reaction.

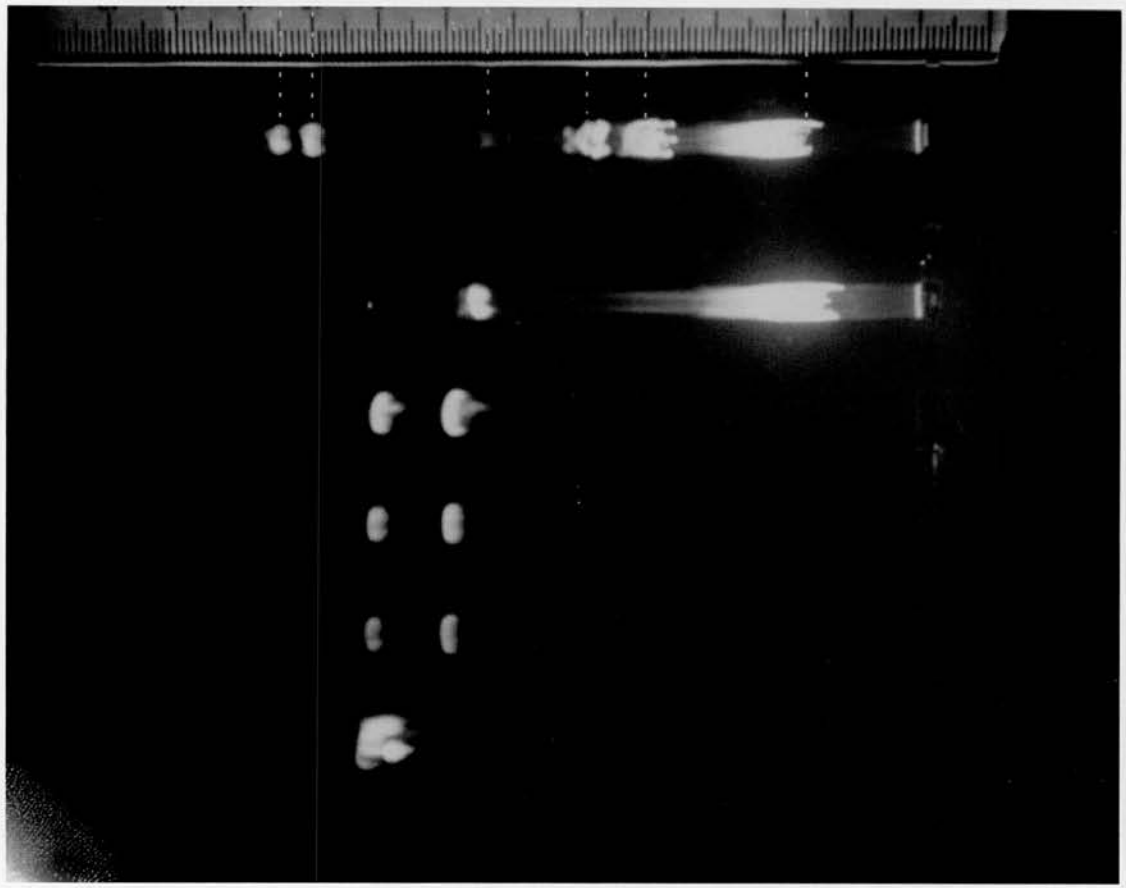
3.3.6 Preparation, restriction digestion, and subcloning of λG4 DNA prepared from the new λG4 phage lysate stocks

The titre of the phage stocks that were produced following plaque purification of the λG4 phage was significantly higher than that of the original λG4 stock, 10^{13} pfu/ml as opposed to 1.5 10^{9} pfu/ml, and therefore, although a liquid lysate method of phage preparation had been tried unsuccessfully with the previous λG4 phage stock, this method (see 3.2.1) was tried again, and this time was successful. Following isolation of DNA from the phage lysate, using a Qiagen lambda maxi-kit (Qiagen Ltd; see 3.2.3), a *Bam*H I restriction digest of the newly isolated DNA was set up in order to see whether underrepresented bands were still generated. The difference between the products of this DNA digest and previous digests of λG4 DNA with the same enzyme, when the restriction products were analysed on an agarose gel, was amazing. Some quite strong bands that had been seen previously were now completely absent, and most importantly, the 4 kb *Bam*H I restriction fragment that had previously proved impossible to subclone, now appeared as a strong easily visible band. The 4 kb *Bam*H I restriction fragment from the new batch of λG4 DNA, was then subcloned into pBluescript SK (Stratagene; for vector map see Chapter 2 Figure 2.1) without any problems, and a Southern blot was prepared from a gel on which *Bam*H I digests of both the new batch of λG4 DNA and the λG4 4 kb *Bam*H I subclone (known as BSc4) had been run (Figure 3.6a). Hybridisation of the Southern blot with the Ac(500)4 probe, resulted in strong hybridisation of the probe to the 4 kb fragment from the λG4 DNA and also to the 4 kb insert fragment from the BSc4 plasmid (Figures 3.6a and 3.6b), thus confirming that the hybridising fragment had been successfully subcloned. Preliminary sequencing of the BSc4 subclone (using forward and reverse M13/pUC vector primers), showed that the sequence from one end of the 4 kb insert corresponded to mouse intronic repeat sequences, whereas sequence from the other end of the insert showed a high degree of similarity to sequence from the 3' untranslated

Figure 3.6a

Restriction digests of new (plaque purified) λG4 DNA and subcloned 4 kb *Bam*H I restriction fragments, separated on an agarose gel in preparation for Southern blotting.

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane)



λ *Hin*d III markers

λG4 *Bam*H I

BSc4 *Bam*H I

BSc4 *Bam*H I

BSc4 *Bam*H I

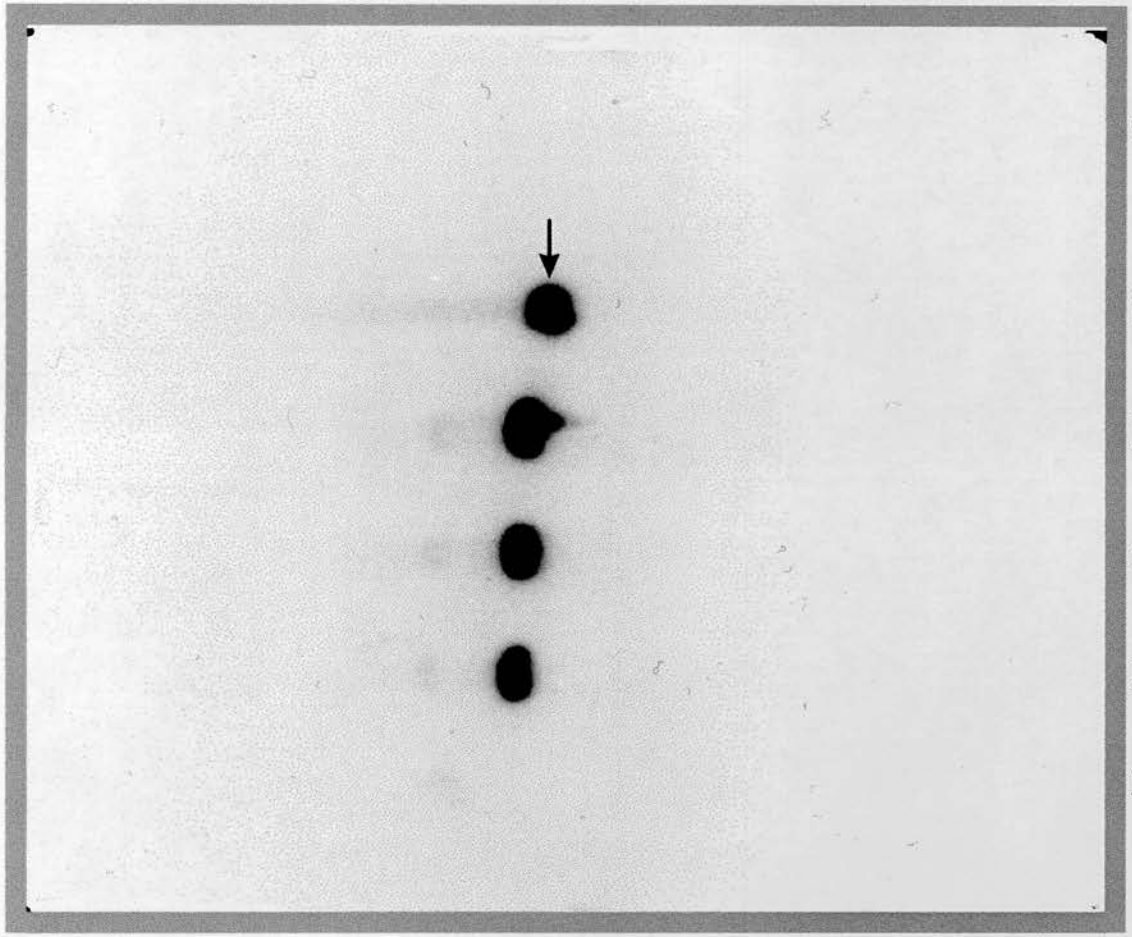
pBluescript *Bam*H I

Figure 3.6b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested DNA shown in Figure 3.6a is incubated with the radiolabelled Vipr2 rat Ac(500)4 cDNA probe

The 4 kb fragments from the BSc4 subclones and the 4 kb *Bam*H I fragment from the λG4 digest, all hybridise strongly to the probe.

(The autoradiograph is shown at actual size)



λ *Hin*d III markers

λG4 *Bam*H I

BSc4 *Bam*H I

BSc4 *Bam*H I

BSc4 *Bam*H I

pBluescript *Bam*H I

region of the rat Vipr2 cDNA. This sequence data, together with the clear hybridisation of the BSc4 insert to the Ac(500)4 probe, strongly suggested that the BSc4 subclone contained mouse *Vipr2* gene sequence that began upstream of sequence corresponding to the 3' end of the Ac(500)4 probe (bp 1192 of the rat Vipr2 cDNA) and continued into the 3' untranslated region of the gene.

3.3.7 Analysis of dot blots of λG4 DNA, using probes from the 5' region of the rat Vipr2 cDNA

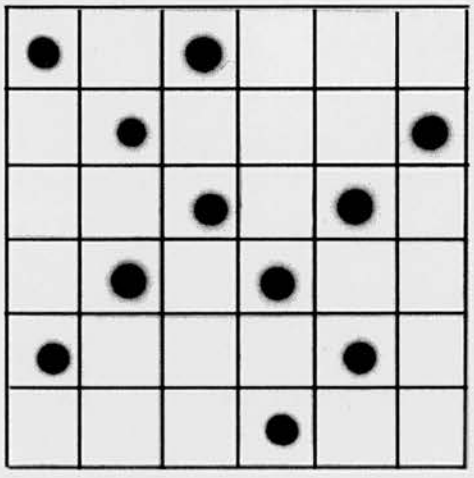
In order to try to establish whether the λG4 clone contained the 5' end of the mouse *Vipr2* gene, two dot blots of λG4 DNA were set up (as described in section 3.2.9) for use in hybridisation reactions with probes derived from the 5' region of the rat Vipr2 cDNA. The first dot blot was hybridised with the rat Vipr2 5' 1 kb cDNA probe, and following autoradiography a clear hybridisation signal was detected in regions containing λG4 DNA (Figure 3.7). However hybridisation of the second blot with the rat Vipr2 5' 260 bp cDNA probe, did not produce any hybridisation signal following autoradiography. Since the rat Vipr2 5' 1 kb cDNA fragment overlaps with the 5' half of Ac(500)4 and ends at bp 999 of the rat cDNA sequence, the positive hybridisation result with the 5' 1 kb probe suggested that the λG4 clone contains sequence 5' of the mouse genomic equivalent of rat Vipr2 cDNA bp 999, but the failure of the λG4 DNA to hybridise to the 5' 260 bp probe, indicated that the 5' end of the *Vipr2* gene is not present in this λ clone.

3.3.8 Subcloning of restriction fragments from BSc4

A preliminary restriction map of the BSc4 subclone was constructed on the basis of the results obtained from *Bam*H I, *Hind* III, and *Pst* I, single and double restriction enzyme digests of the plasmid (Figure 3.8). An ~1.25 kb *Bam*H I/*Hind* III restriction fragment (BH1.25) from the 5' end of the BSc4 insert, the directly adjacent ~1.8 kb *Hind* III/*Pst* I restriction fragment (HP1.8) which spanned the central region of the insert, and an ~700 bp *Pst* I/*Bam*H I restriction fragment containing the remaining sequence from the 3' end of the insert, were then subcloned into pBluescript SK (Stratagene; see Figure 2.1). Initially there was some uncertainty as to whether the 3' *Pst* I site at the 3' end of the BSc4 insert, was within the insert sequence or was part of the plasmid polylinker sequence, and consequently the 3' ~700 bp restriction fragment from the insert was cloned in the form of both a *Pst* I fragment (P700) and a *Pst* I/*Bam*H I fragment (PB700). However, P700 and PB700 are essentially the same (see Figure 3.8).

Figure 3.7

Dot blot of λG4 DNA following hybridisation to the radiolabelled rat Vipr2 5' 1 kb cDNA probe

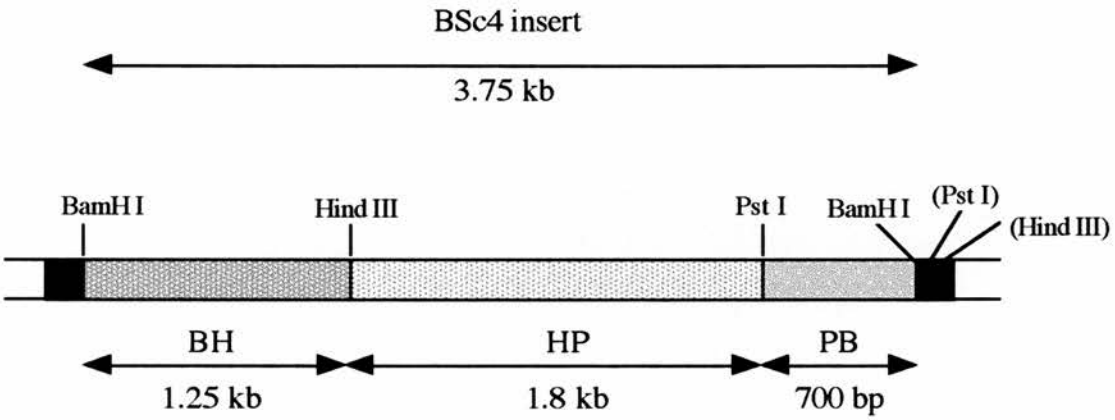


G4		G4		λ	
	G4		λ		G4
λ		G4		G4	
	G4		G4		λ
G4				G4	
			G4		

Figure 3.8

Preliminary restriction map of the λG4 subclone BSc4

λG4 subclone BSc4



3.3.9 Southern blot analysis and sequencing of subcloned fragments from BSc4

As a first step towards identifying the regions of the *Vipr2* gene that were present in BSc4, a Southern blot was prepared from a gel on which the insert DNA from the BSc4 subclones BH1.25, HP1.8, PB700, and P700, had been separated from the vector sequence, and this blot was used in hybridisation reactions with probes that had previously been shown to hybridise to λG4 DNA.

Following an attempted hybridisation of the blot with the rat Vipr2 1 kb 5' cDNA probe no hybridisation to the insert fragments of any of the subclones was observed. However, after hybridisation of the blot with the Ac(500)4 probe, bands representing hybridisation to the insert fragment of the HP1.8 subclone and the insert fragment of the BSc4 control plasmid, could be clearly seen (Figures 3.9a and 3.9b). This suggested that the HP1.8 subclone contained mouse *Vipr2* exon sequence which corresponded to at least part of the sequence present within the Ac(500)4 probe i.e. bp 711 to bp 1192 of the rat Vipr2 cDNA sequence. Since BSc4 was thought to end within the 3' untranslated region (UTR) of the mouse *Vipr2* gene, the failure of P700 and PB700 (which were derived from the 3' end of BSc4) to hybridise to the Ac(500)4 probe was not surprising, as the Ac(500)4 probe ends 187 bp upstream of the 3' UTR in the rat *Vipr2* cDNA. However, the absence of any detectable hybridisation between BH1.25 and either the 5' 1 kb probe or the Ac(500)4 probe, seemed to indicate that either this subclone was composed largely of intron sequence, or if *Vipr2* gene exon sequences were present in BH1.25 then they were significantly different from those found in rat.

Fortunately, at around this stage of the study, the mouse Vipr2 cDNA sequence was published (Genebank accession number D28132; Inagaki *et al.*, 1994), and by aligning this sequence with the published exon sequences of other members of the receptor gene sub-family (whose gene structures had already been published), a prediction of the structure of the mouse *Vipr2* gene was made (see Figure 3.10).

Subsequent sequencing of subclones BH1.25, HP1.8, P700 and PB700, generally confirmed the conclusions that had been suggested by the Southern blot analysis, and through comparison with the mouse cDNA sequence, produced the first definitive information relating to the mouse *Vipr2* gene's structure.

The data obtained from the use of forward and reverse vector primers demonstrated that subclones P700 and PB700 were identical with the exception of a few additional base

Figure 3.9a

Restriction enzyme digests of DNA from BSc4 and BSc4-derived subclones (BH1.25, HP1.8, P700, and PB700), separated on an agarose gel in preparation for Southern blotting.

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane)

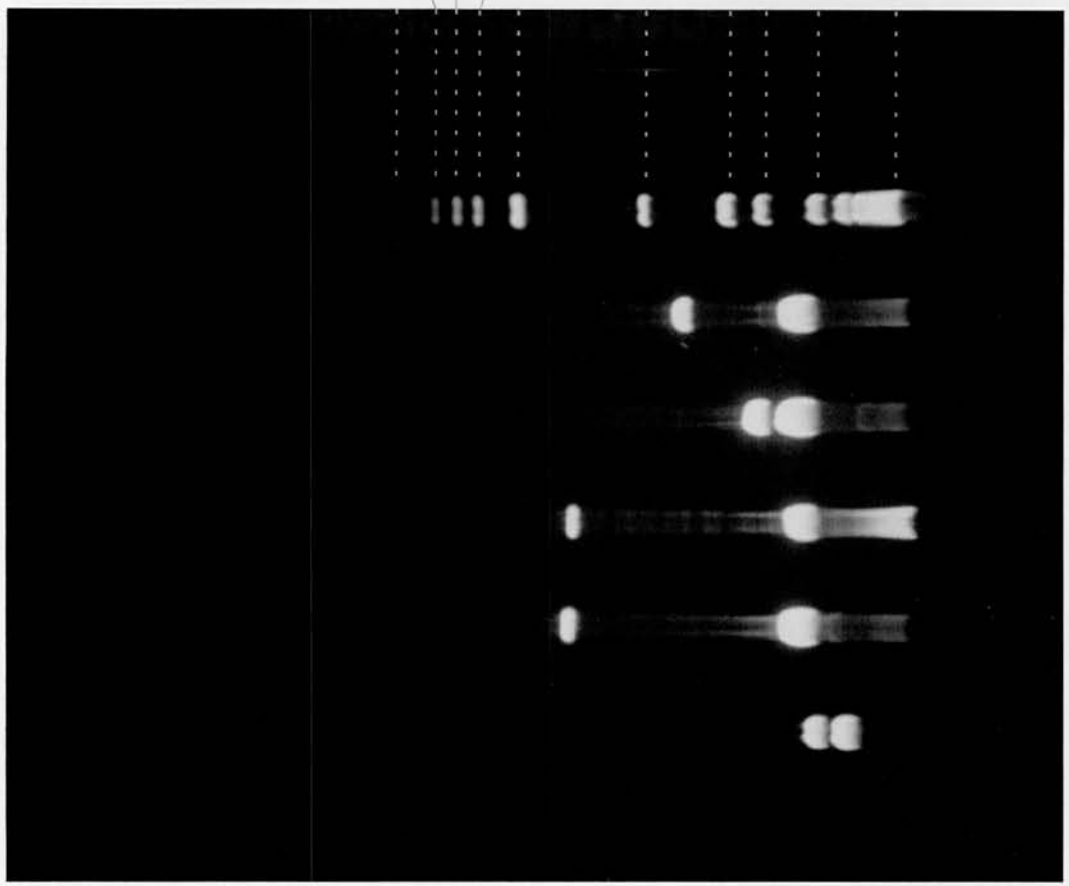
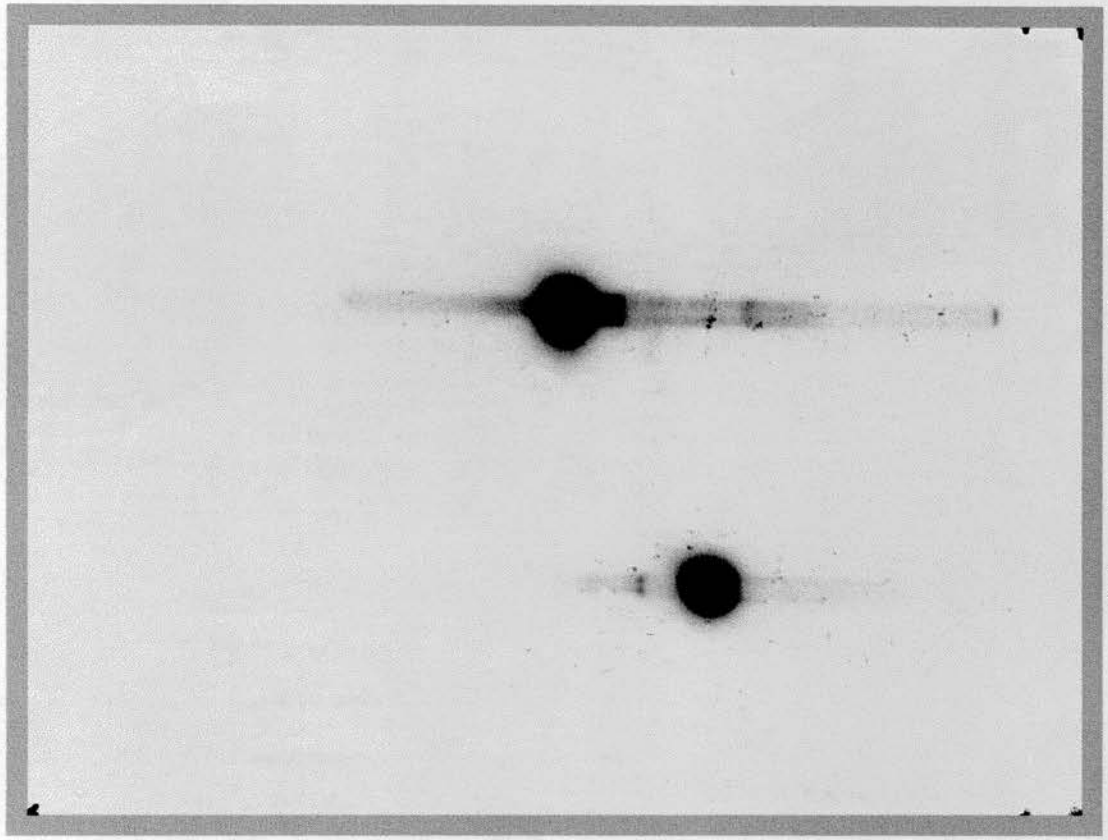

12.2 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
~517 bp
396 bp
344 bp
298 bp
~200 bp
1 kb ladder
BH 1.25 *Bam*H I/*Hin*d III
HP1.8 *Hin*d III/*Pst* I
P700 *Pst* I
PB700 *Pst* I/*Bam*H I
Bsc4 *Bam*H I

Figure 3.9b

Hybridisation pattern observed when a Southern blot taken from the gel shown in Figure 3.9a is incubated with the radiolabelled Vipr2 rat Ac(500)4 cDNA probe

The probe can be seen to have hybridised to the plasmid inserts of HP1.8 and BSc4

(The autoradiograph is shown at actual size)



1 kb ladder

BH 1.25 *Bam*H I/*Hin*d III

HP1.8 *Hin*d III/*Pst* I

P700 *Pst* I

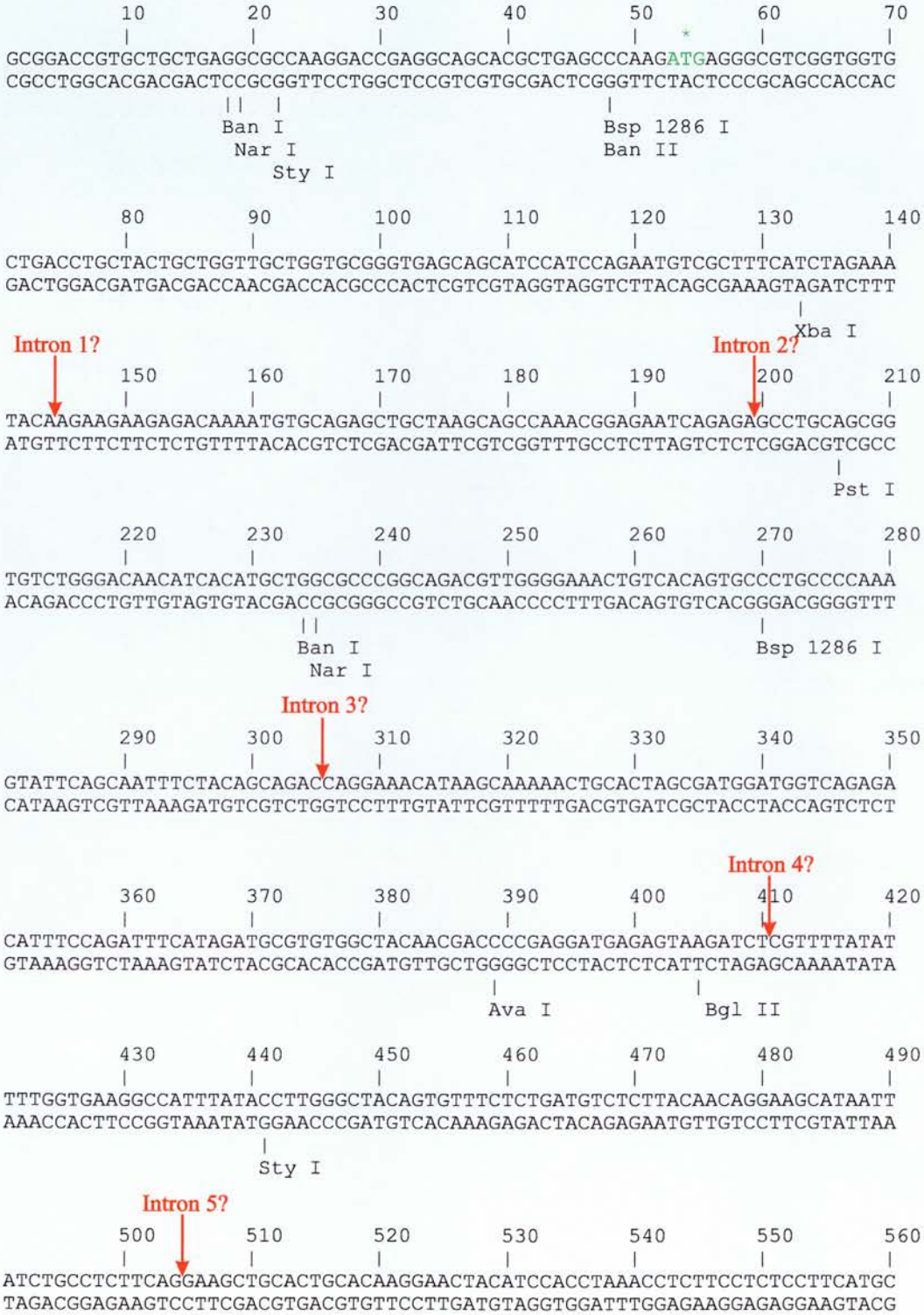
PB700 *Pst* I/*Bam*H I

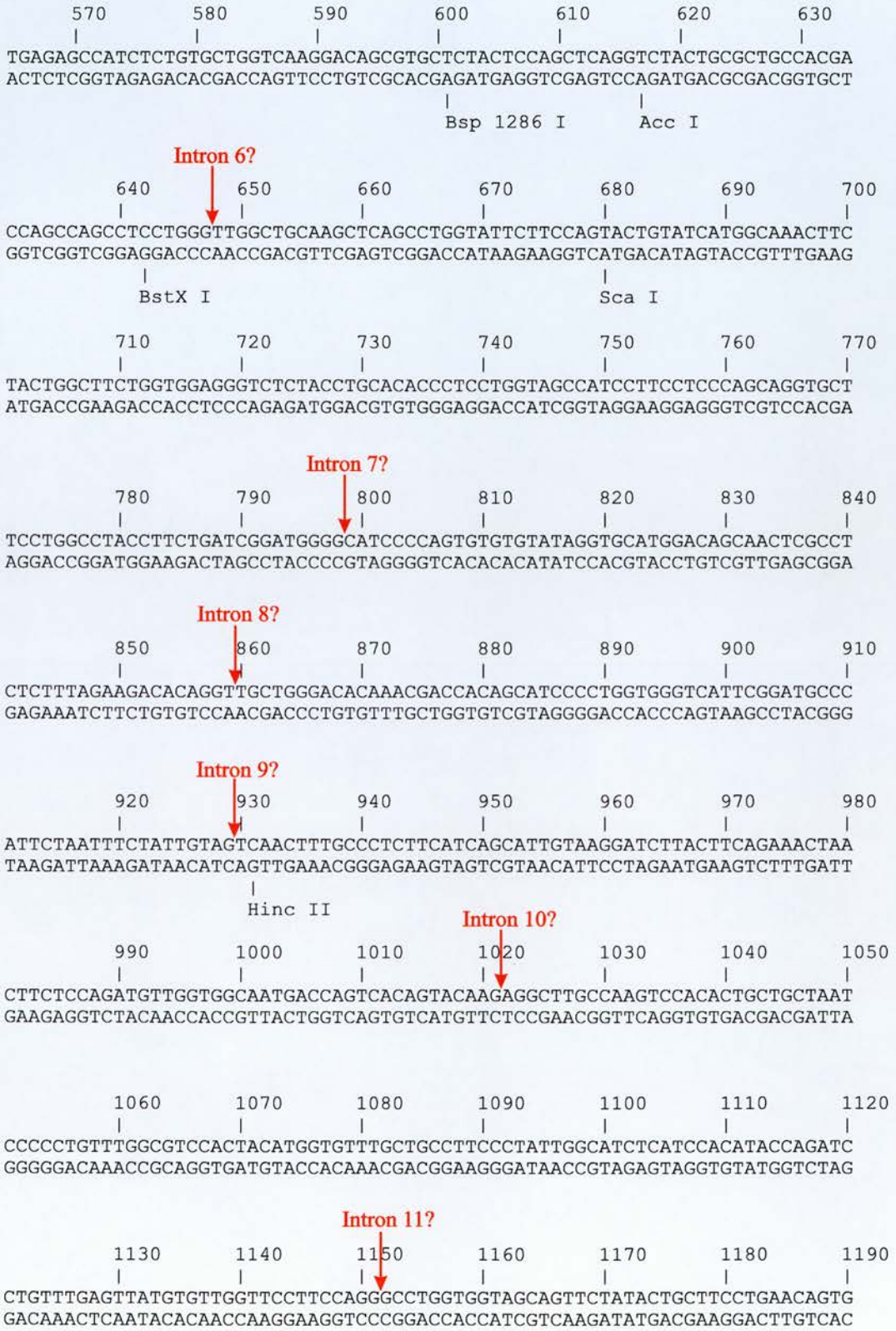
Bsc4 *Bam*H I

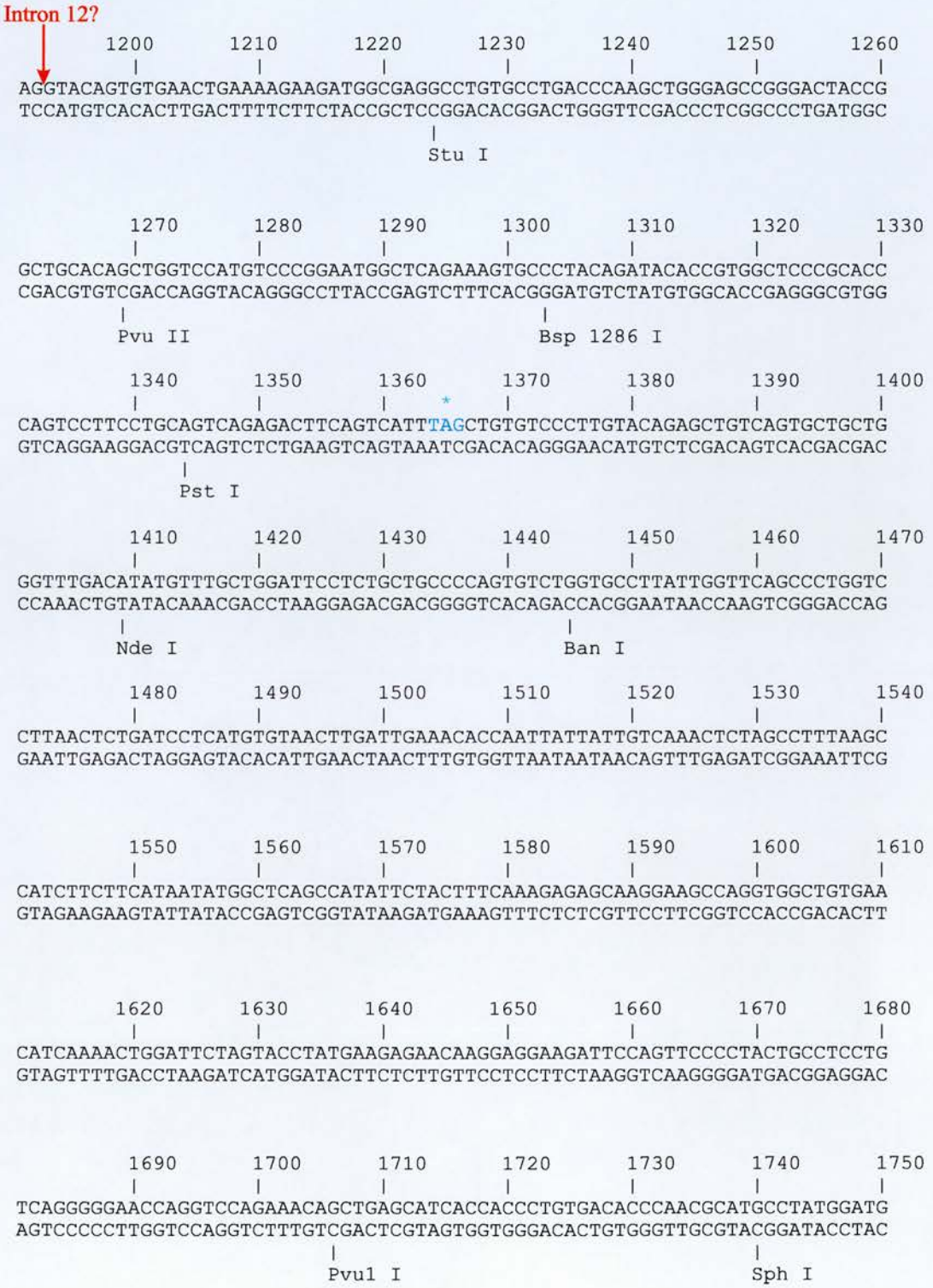
Figure 3.10

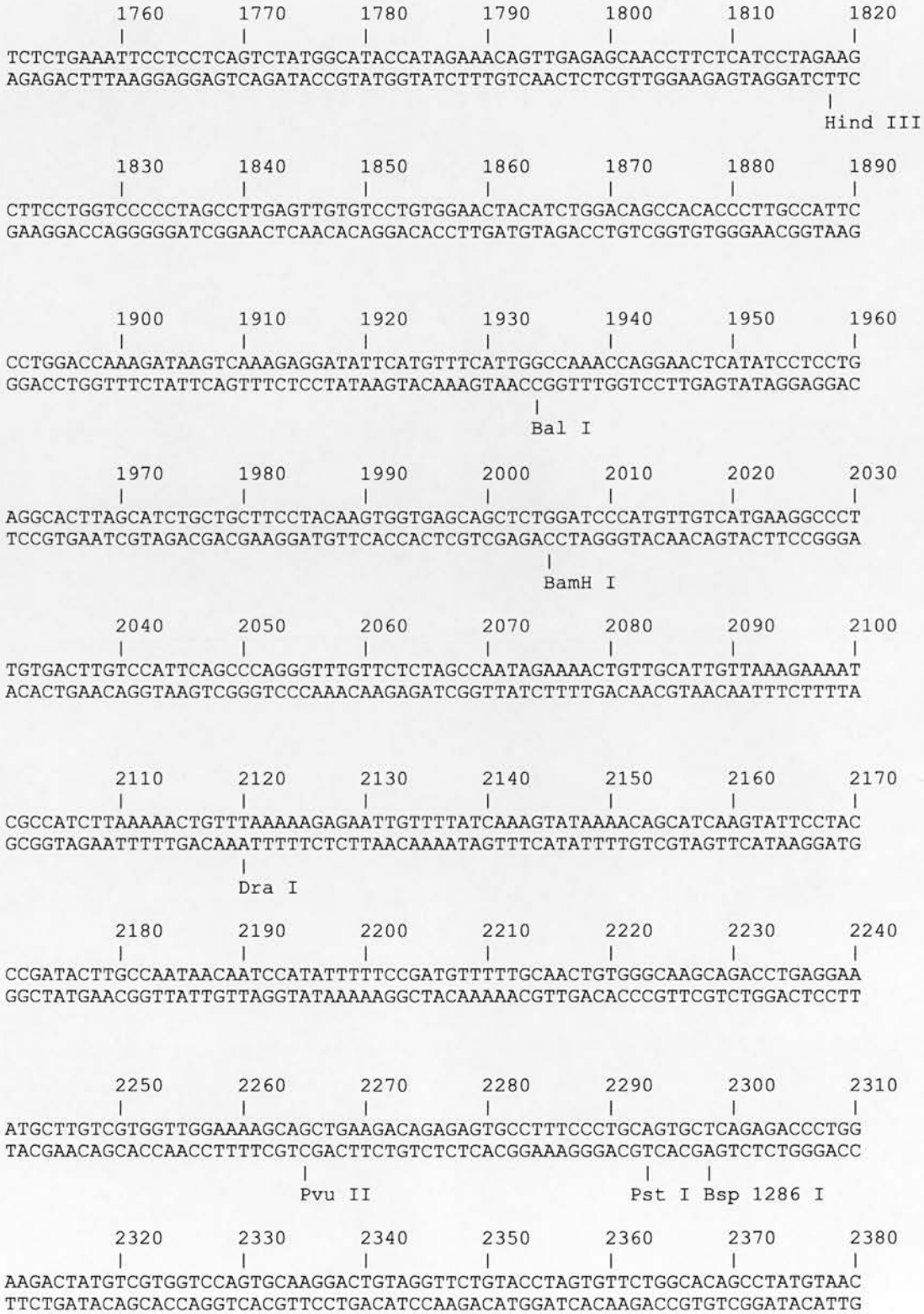
Diagram showing the predicted positions of introns within the mouse *Vipr2* gene (as determined from sequence alignment with other members of the secretin/glucagon receptor gene subfamily).

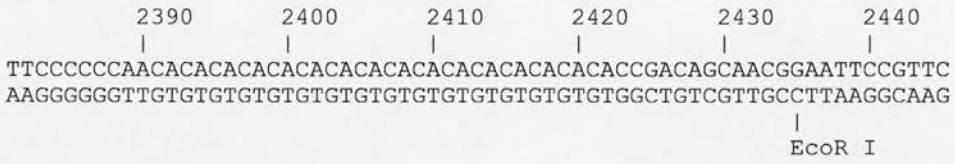
Within the mouse Vipr2 cDNA sequence, positions that are predicted to correspond to intron sites within the *Vipr2* gene are indicated with a red arrow, and are labelled according to intron number.


10 20 30 40 50 60 70
GCGGACCGTGCTGCTGAGGCGCCAAGGACCGAGGCAGCACGCTGAGCCCAAGATGAGGGCGTCGGTGGTG
CGCCTGGCACGACGACTCCGCGGTTCCTGGCTCCGTCGTGCGACTCGGGTTCTACTCCCGCAGCCACCAC
Ban I
Nar I
Sty I
Bsp 1286 I
Ban II
80 90 100 110 120 130 140
CTGACCTGCTACTGCTGGTTGCTGGTGCGGGTGAGCAGCATCCATCCAGAATGTCGCTTTCATCTAGAAA
GACTGGACGATGACGACCAACGACCACGCCCACTCGTCGTAGGTAGGTCTTACAGCGAAAGTAGATCTTT
Xba I
Intron 1?
Intron 2?
150 160 170 180 190 200 210
TACAAGAAGAAGAGACAAAATGTGCAGAGCTGCTAAGCAGCCAAACGGAGAATCAGAGAGCCTGCAGCGG
ATGTTCTTCTTCTCTGTTTTACACGTCTCGACGATTCGTCGGTTTGCCTCTTAGTCTCTCGGACGTCGCC
Pst I
220 230 240 250 260 270 280
TGTCTGGGACAACATCACATGCTGGCGCCCGGCAGACGTTGGGGAAACTGTCACAGTGCCCTGCCCCAAA
ACAGACCCTGTTGTAGTGTACGACCGCGGGCCGTCTGCAACCCCTTTGACAGTGTCACGGGACGGGGTTT
Ban I
Nar I
Bsp 1286 I
Intron 3?
290 300 310 320 330 340 350
GTATTCAGCAATTTCTACAGCAGACCAGGAAACATAAGCAAAAACTGCACTAGCGATGGATGGTCAGAGA
CATAAGTCGTTAAAGATGTCGTCTGGTCCTTTGTATTCGTTTTTGACGTGATCGCTACCTACCAGTCTCT
Intron 4?
360 370 380 390 400 410 420
CATTTCCAGATTTCATAGATGCGTGTGGCTACAACGACCCCGAGGATGAGAGTAAGATCTCGTTTTATAT
GTAAAGGTCTAAAGTATCTACGCACACCGATGTTGCTGGGGCTCCTACTCTCATTCTAGAGCAAAATATA
Ava I
Bgl II
430 440 450 460 470 480 490
TTTGGTGAAGGCCATTTATACCTTGGGCTACAGTGTTTCTCTGATGTCTCTTACAACAGGAAGCATAATT
AAACCACTTCCGGTAAATATGGAACCCGATGTCACAAAGAGACTACAGAGAATGTTGTCCTTCGTATTAA
Sty I
Intron 5?
500 510 520 530 540 550 560
ATCTGCCTCTTCAGGAAGCTGCACTGCACAAGGAACTACATCCACCTAAACCTCTTCCTCTCCTTCATGC
TAGACGGAGAAGTCCTTCGACGTGACGTGTTCCTTGATGTAGGTGGATTTGGAGAAGGAGAGGAAGTACG


570 580 590 600 610 620 630
TGAGAGCCATCTCTGTGCTGGTCAAGGACAGCGTGCTCTACTCCAGCTCAGGTCTACTGCGCTGCCACGA
ACTCTCGGTAGAGACACGACCAGTTCCTGTCGCACGAGATGAGGTCGAGTCCAGATGACGCGACGGTGCT
Bsp 1286 I
Acc I
Intron 6?
640 650 660 670 680 690 700
CCAGCCAGCCTCCTGGGTTGGCTGCAAGCTCAGCCTGGTATTCTTCCAGTACTGTATCATGGCAAACTTC
GGTCGGTCGGAGGACCCAACCGACGTTCGAGTCGGACCATAAGAAGGTCATGACATAGTACCGTTTGAAG
BstX I
Sca I
710 720 730 740 750 760 770
TACTGGCTTCTGGTGGAGGGTCTCTACCTGCACACCCTCCTGGTAGCCATCCTTCCTCCCAGCAGGTGCT
ATGACCGAAGACCACCTCCCAGAGATGGACGTGTGGGAGGACCATCGGTAGGAAGGAGGGTCGTCCACGA
Intron 7?
780 790 800 810 820 830 840
TCCTGGCCTACCTTCTGATCGGATGGGGCATCCCCAGTGTGTGTATAGGTGCATGGACAGCAACTCGCCT
AGGACCGGATGGAAGACTAGCCTACCCCGTAGGGGTCACACACATATCCACGTACCTGTCGTTGAGCGGA
Intron 8?
850 860 870 880 890 900 910
CTCTTTAGAAGACACAGGTTGCTGGGACACAAACGACCACAGCATCCCCTGGTGGGTCATTCGGATGCCC
GAGAAATCTTCTGTGTCCAACGACCCTGTGTTTGCTGGTGTCGTAGGGGACCACCCAGTAAGCCTACGGG
Intron 9?
920 930 940 950 960 970 980
ATTCTAATTTCTATTGTAGTCAACTTTGCCCTCTTCATCAGCATTGTAAGGATCTTACTTCAGAAACTAA
TAAGATTAAAGATAACATCAGTTGAAACGGGAGAAGTAGTCGTAACATTCCTAGAATGAAGTCTTTGATT
Hinc II
Intron 10?
990 1000 1010 1020 1030 1040 1050
CTTCTCCAGATGTTGGTGGCAATGACCAGTCACAGTACAAGAGGCTTGCCAAGTCCACACTGCTGCTAAT
GAAGAGGTCTACAACCACCGTTACTGGTCAGTGTCATGTTCTCCGAACGGTTCAGGTGTGACGACGATTA
1060 1070 1080 1090 1100 1110 1120
CCCCCTGTTTGGCGTCCACTACATGGTGTTTGCTGCCTTCCCTATTGGCATCTCATCCACATACCAGATC
GGGGGACAAACCGCAGGTGATGTACCACAAACGACGGAAGGGATAACCGTAGAGTAGGTGTATGGTCTAG
Intron 11?
1130 1140 1150 1160 1170 1180 1190
CTGTTTGAGTTATGTGTTGGTTCCTTCCAGGGCCTGGTGGTAGCAGTTCTATACTGCTTCCTGAACAGTG
GACAAACTCAATACACAACCAAGGAAGGTCCCGGACCACCATCGTCAAGATATGACGAAGGACTTGTCAC


Intron 12?
1200 1210 1220 1230 1240 1250 1260
AGGTACAGTGTGAACTGAAAAGAAGATGGCGAGGCCTGTGCCTGACCCAAGCTGGGAGCCGGGACTACCG
TCCATGTCACACTTGACTTTTCTTCTACCGCTCCGGACACGGACTGGGTTCGACCCTCGGCCCTGATGGC
Stu I
1270 1280 1290 1300 1310 1320 1330
GCTGCACAGCTGGTCCATGTCCCGGAATGGCTCAGAAAGTGCCCTACAGATACACCGTGGCTCCCGCACC
CGACGTGTCGACCAGGTACAGGGCCTTACCGAGTCTTTCACGGGATGTCTATGTGGCACCGAGGGCGTGG
Pvu II
Bsp 1286 I
1340 1350 1360 1370 1380 1390 1400
CAGTCCTTCCTGCAGTCAGAGACTTCAGTCATTTAGCTGTGTCCCTTGTACAGAGCTGTCAGTGCTGCTG
GTCAGGAAGGACGTCAGTCTCTGAAGTCAGTAAATCGACACAGGGAACATGTCTCGACAGTCACGACGAC
Pst I
1410 1420 1430 1440 1450 1460 1470
GGTTTGACATATGTTTGCTGGATTCCTCTGCTGCCCCAGTGTCTGGTGCCTTATTGGTTCAGCCCTGGTC
CCAAACTGTATACAAACGACCTAAGGAGACGACGGGGTCACAGACCACGGAATAACCAAGTCGGGACCAG
Nde I
Ban I
1480 1490 1500 1510 1520 1530 1540
CTTAACTCTGATCCTCATGTGTAACTTGATTGAAACACCAATTATTATTGTCAAACTCTAGCCTTTAAGC
GAATTGAGACTAGGAGTACACATTGAACTAACTTTGTGGTTAATAATAACAGTTTGAGATCGGAAATTCG
1550 1560 1570 1580 1590 1600 1610
CATCTTCTTCATAATATGGCTCAGCCATATTCTACTTTCAAAGAGAGCAAGGAAGCCAGGTGGCTGTGAA
GTAGAAGAAGTATTATACCGAGTCGGTATAAGATGAAAGTTTCTCTCGTTCCTTCGGTCCACCGACACTT
1620 1630 1640 1650 1660 1670 1680
CATCAAAACTGGATTCTAGTACCTATGAAGAGAACAAGGAGGAAGATTCCAGTTCCCCTACTGCCTCCTG
GTAGTTTTGACCTAAGATCATGGATACTTCTCTTGTTCCTCCTTCTAAGGTCAAGGGGATGACGGAGGAC
1690 1700 1710 1720 1730 1740 1750
TCAGGGGGAACCAGGTCCAGAAACAGCTGAGCATCACCACCCTGTGACACCCAACGCATGCCTATGGATG
AGTCCCCCTTGGTCCAGGTCTTTGTCGACTCGTAGTGGTGGGACACTGTGGGTTGCGTACGGATACCTAC
Pvu1 I
Sph I


1760 1770 1780 1790 1800 1810 1820
TCTCTGAAATTCCTCCTCAGTCTATGGCATACCATAGAAACAGTTGAGAGCAACCTTCTCATCCTAGAAG
AGAGACTTTAAGGAGGAGTCAGATACCGTATGGTATCTTTGTCAACTCTCGTTGGAAGAGTAGGATCTTC
Hind III
1830 1840 1850 1860 1870 1880 1890
CTTCCTGGTCCCCCTAGCCTTGAGTTGTGTCCTGTGGAACTACATCTGGACAGCCACACCCTTGCCATTC
GAAGGACCAGGGGGATCGGAACTCAACACAGGACACCTTGATGTAGACCTGTCGGTGTGGGAACGGTAAG
1900 1910 1920 1930 1940 1950 1960
CCTGGACCAAAGATAAGTCAAAGAGGATATTCATGTTTCATTGGCCAAACCAGGAACTCATATCCTCCTG
GGACCTGGTTTCTATTCAGTTTCTCCTATAAGTACAAAGTAACCGGTTTGGTCCTTGAGTATAGGAGGAC
Bal I
1970 1980 1990 2000 2010 2020 2030
AGGCACTTAGCATCTGCTGCTTCCTACAAGTGGTGAGCAGCTCTGGATCCCATGTTGTCATGAAGGCCCT
TCCGTGAATCGTAGACGACGAAGGATGTTCACCACTCGTCGAGACCTAGGGTACAACAGTACTTCCGGGA
BamH I
2040 2050 2060 2070 2080 2090 2100
TGTGACTTGTCCATTCAGCCCAGGGTTTGTTCTCTAGCCAATAGAAAACTGTTGCATTGTTAAAGAAAAT
ACACTGAACAGGTAAGTCGGGTCCCAAACAAGAGATCGGTTATCTTTTGACAACGTAACAATTTCTTTTA
2110 2120 2130 2140 2150 2160 2170
CGCCATCTTAAAAACTGTTTAAAAAGAGAATTGTTTTATCAAAGTATAAAACAGCATCAAGTATTCCTAC
GCGGTAGAATTTTTGACAAATTTTTCTCTTAACAAAATAGTTTCATATTTTGTCGTAGTTCATAAGGATG
Dra I
2180 2190 2200 2210 2220 2230 2240
CCGATACTTGCCAATAACAATCCATATTTTTCCGATGTTTTTGCAACTGTGGGCAAGCAGACCTGAGGAA
GGCTATGAACGGTTATTGTTAGGTATAAAAAGGCTACAAAAACGTTGACACCCGTTCGTCTGGACTCCTT
2250 2260 2270 2280 2290 2300 2310
ATGCTTGTCGTGGTTGGAAAAGCAGCTGAAGACAGAGAGTGCCTTTCCCTGCAGTGCTCAGAGACCCTGG
TACGAACAGCACCAACCTTTTCGTCGACTTCTGTCTCTCACGGAAAGGGACGTCACGAGTCTCTGGGACC
Pvu II
Pst I Bsp 1286 I
2320 2330 2340 2350 2360 2370 2380
AAGACTATGTCGTGGTCCAGTGCAAGGACTGTAGGTTCTGTACCTAGTGTTCTGGCACAGCCTATGTAAC
TTCTGATACAGCACCAGGTCACGTTCCTGACATCCAAGACATGGATCACAAGACCGTGTCGGATACATTG


2390
2400
2410
2420
2430
2440
TTCCCCCCAACACACACACACACACACACACACACACACACACCGACAGCAACGGAATTCCGTTC
AAGGGGGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGGCTGTCGTTGCCTTAAGGCAAG
EcoR I

pairs of vector polylinker sequence that was present at the end of P700, and both subclones spanned genomic sequence corresponding to region from bp 1345 (~20 bp upstream of the TAG translation 'stop' codon) to bp 2005 (in the 3' UTR) of the mouse Vipr2 cDNA sequence.

Initial sequencing of subclone HP1.8, using vector primers, showed that this subclone contained genomic sequence corresponding to bp 1021-1344 of the mouse Vipr2 cDNA sequence, and also contained ~465 bp of intron sequence 5' to the start of the exon sequence (cDNA bp 1021). Since this region of the mouse cDNA sequence was thought to be encoded by predicted exons 11, 12, and 13 of the mouse gene, primers (4341, 4342, 4343, 4344 and 5990) that were designed to bind to these exons, were also used to sequence this subclone. The data obtained from these sequencing reactions, confirmed that the intron/exon junctions in this region were present at the positions predicted in Figure 3.10. An intron of 66 bp was found at the site of predicted intron 11, and an intron of unknown size (>1 kb) was found at the site of predicted intron 12. These results are consistent with the Southern blot results which indicated that only sequences 3' of those corresponding to the rat 5' 1 kb Vipr2 cDNA probe (i.e. the start of predicted exon 10) were present in the HP1.8 subclone.

In contrast to the other subclones, sequencing of subclone BH1.25 with vector primers, did not reveal any sequences that displayed homology to the mouse Vipr2 cDNA sequence (although ~500 bp of sequence was obtained from each side of the insert). However, as this subclone was derived from the region that lies directly upstream of subclone HP1.8 (which appeared to begin in intron 10 of the mouse *Vipr2* gene), did not hybridise to rat Vipr2 cDNA probes on Southern blots, and could not be sequenced using the exon 10 primer 4340, it seemed likely that BH1.25 consisted entirely of sequence from intron 10.

3.3.10 Initial screen of the λ2001 ES-cell-derived mouse genomic library, for clones containing the *Vipr2* gene

Although the BSc 4 subclone had been successfully isolated from the new λG4 phage stocks, and was shown to contain sequences from the 3' end of the mouse *Vipr2* gene, attempts to isolate further *Vipr2* gene derived restriction fragments from the λG4 clone were discontinued when it was found that later preparations of DNA from the phage stock were beginning to display the same type of anomalous restriction patterns that had been observed with the original λG4 DNA preparation prior to plaque purification. The reason for the apparent instability of the λG4 clone was not clear. However, as it seemed that repeated plaque purifications would be required if the study of this clone

was to be pursued any further, it was decided that it would be better in the long run, to isolate new clones from another library.

Nevertheless, there was a positive side to this unexpected change of plan, as it provided the opportunity to screen a mouse genomic library that was derived from ES-cells, a factor which was not important with respect to the characterisation of the mouse *Vipr2* gene's structure, but could be critical if the clones were to be used in gene targeting studies at a later stage.

Plaque lifts were prepared from a mouse ES-cell derived genomic library (constructed in the vector λ2001; see Figure 2.5 for vector map) as described in section 3.2.10, and following hybridisation of these filters with the full length rat Vipr2 cDNA probe, 31 positive clones were detected. Sixteen of the positive clones hybridised very strongly to the Vipr2 cDNA probe, whereas the other fifteen hybridised more weakly. All 31 positive plaques were 'picked' (see 3.2.10), but initially only the ten plaques (clones A1-A10) which were thought to show the strongest hybridisation to the probe, were used in the initial secondary screen.

In the secondary screen, as in the primary screen, two filters were prepared from each of the phage plates (see 3.2.10). However as one of the aims of the work on the *Vipr2* gene was to isolate and characterise clones that contained the promoter region of the *Vipr2* gene, there was a particular interest in isolating clones containing the 5' end of the *Vipr2* gene's coding sequence, and consequently in the secondary screen, one filter from each of the plates was hybridised with the full length rat Vipr2 cDNA probe, while the other filter was hybridised with the 5' 260 bp probe. The results of these hybridisations demonstrated that whereas all but one (A2) of the original 10 positive clones hybridised to the full length rat Vipr2 cDNA probe in the secondary screen, none of the 10 positive clones hybridised to the 5' 260 bp probe. Therefore, using the autoradiographs from the filters that had been hybridised with the full length cDNA probe, single plaques were isolated from the 9 positive plates, and were stored in SM buffer until required (see 3.2.10). The filters that had been hybridised with the 5' 260 bp probe were then rehybridised with the rat Vipr2 5' 1 kb cDNA probe in an attempt to determine whether any of these clones contained *Vipr2* gene sequences that lay 5' of those isolated from the λG4 subclone BSc4, and with this probe, 5 positive clones (A1, A3, A8, A9, and A10) were identified.

3.3.11 Sequencing of selected clones from the initial screen of the library

DNA for use in sequencing was prepared from 5 of the clones that had been isolated in the secondary screen of the library. These clones included 4 clones that had hybridised to both the 5' 1 kb probe and the full length Vipr2 cDNA probe (A1, A3, A8, and A9), and one clone (A7) that had hybridised only to the full length cDNA probe. Preliminary sequence analysis of these clones was carried out by testing the ability of two sequencing primers 3755 (which binds to sequence within predicted exon 3 of the mouse gene) and 7992 (which binds to sequence within the 3' UTR of the mouse gene) to sequence the clones. As suggested by the failure of these clones to hybridise to the 5' 260 bp rat Vipr2 cDNA probe (which ends at a point which corresponds to predicted exon 3 of the mouse gene), none of the 5 clones was successfully sequenced by primer 3775, although all 5 clones were sequenced correctly by the primer 7992. These results suggested that all 5 clones originated from the 3' end of the mouse *Vipr2* gene and did not contain exon sequence that extended further 5' than exon 4 of the gene.

3.3.12 Screening of the λ2001 mouse genomic library, for clones containing the 5' region of the mouse *Vipr2* gene

Since the 5 strongly hybridising clones that had been sequenced, all contained the 3' end of the mouse *Vipr2* gene, it began to appear that the clones that had initially shown weaker hybridisation to the rat full length Vipr2 cDNA probe, might contain the 5' end of the *Vipr2* gene. To investigate this possibility, another secondary screen of the λ2001 library was carried out, this time using 11 plaques (clones B1-B11) that had hybridised fairly weakly in the primary screen of the library. In this secondary screen one set of filters was hybridised with the rat 5' 1 kb Vipr2 cDNA probe, while the replica filters were hybridised with the rat 5' 260 bp probe. Out of the 11 clones, 8 clones (B2, B3, B4, B5, B7, B8, B9, and B11) hybridised to the rat 5' 1 kb Vipr2 cDNA probe, and 2 of these clones (B7 and B9) also hybridised to the 5' 260 bp probe. A single plaque representing each of the hybridising clones was picked, and stored in SM buffer until required.

3.3.13 Restriction digests, Southern blot analysis, and sequencing of *Vipr2* clones isolated from the λ2001 mouse genomic library

Further analysis of DNA from clone A8 (one of the 4 clones which had hybridised to both the full length cDNA and the 1 kb 5' probe), and initial analysis of DNA from

clones B7 and B9 (which had hybridised to the 5' 260 bp probe), was carried out using a selection of sequencing primers (including 4336, 4340, 4343, 5991, 5992, 5993, 7122, 7898 and 9200) that were designed to bind to different exons of the mouse *Vipr2* gene (see Table 3.1). The results of these preliminary sequencing reactions demonstrated that DNA from B7 and B9 could be sequenced with the predicted exon 3 and exon 4 primers (4336 and 7898 respectively), and B9 DNA could also be sequenced with the predicted exon 2 primer 7122. DNA from A8 was successfully sequenced with predicted exon 7, 10, and 11 primers (9200, 4340, and 4343), and as it had been previously been established that A8 DNA could be sequenced with primer 7992 from the 3' UTR of the mouse *Vipr2* gene, it appeared that clone A8 began in the intron 6/exon 7 region and extended into the 3' UTR of the mouse *Vipr2* gene. With the exception of intron 1 which did not appear to be present in either of the 5' clones, the only *Vipr2* gene exons that were not spanned by A8 or B9, were exons 5 and 6. However, by testing exon 5 and exon 6 primers (5991 and 5993) on template DNA from some of the other positive clones from secondary screen B (3.3.12), two clones were identified: B3 (which contains exon 6), and B11 (which contains exons 5 and 6).

Subsequent studies then focused on the characterisation of clones B3, B7, A8, B9, and B11, which from this point onwards are referred to as ES3, ES7, ES8, ES9, and ES11 respectively (regardless of whether they were isolated in secondary screen A or B). Restriction digestion and Southern blotting of these clones was carried out first, with the aim of identifying restriction fragments that contained *Vipr2* gene exon sequences, and obtaining some indication of the extent of overlap between the 5 clones. Sequencing results had already demonstrated that clones ES7 and ES9 both contained exons 3 and 4 of the mouse *Vipr2* gene. However, when Southern blots of ES7 and ES9 DNA were hybridised with the 5' 260 bp Vipr2 cDNA probe (which contained sequences corresponding to regions within exons 1-3 of the mouse gene), some distinct differences were found in the hybridisation patterns produced by these two clones (Figures 3.11a, 3.11b, 3.12a, 3.12b), indicating that the homology between might not extend very far beyond exons 3 and 4 (a conclusion which was in agreement with the failure of the exon 2 primer 7122 to sequence ES7 DNA). ES3 and ES11, which had both been shown to sequence with the exon 6 primer, showed some similarities in their hybridisation patterns, when blots of these clones were hybridised with the 5' 1 kb Vipr2 cDNA probe (originally used to isolate these clones). Hybridising restriction fragments that were of similar sizes from these two clones included: an ~4 kb *Xho* I fragment, an ~ 4 kb *Ava* I fragment, and an ~5 kb *Bam*H I/*Sca* I fragment (Figures 3.13a, 3.13b, 3.14a, 3.14b). Thus, it appeared that at least within the region spanned by the 5' 1 kb cDNA probe, there was probably a substantial degree of overlap between

Figure 3.11a

Restriction enzyme digests of DNA from clone ES7, separated on an agarose gel in preparation for Southern blotting.

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane).

1 kb ladder
ES7 EcoR I
ES7 Hind III
ES7 Nde I
ES7 Pvu I
ES7 Sac I
ES7 Sca I
ES7 Stu I
ES7 Xba I
ES7 Xho I
ES7 Not I
ES7 Sal I
ES7 Spe I
1 kb ladder
12.2 kb
7.13 kb
6.11 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb

Figure 3.11b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested ES7 DNA shown in Figure 3.11a is incubated with the radiolabelled Vipr2 rat 5' 260 bp cDNA probe

(The autoradiograph is shown at actual size)

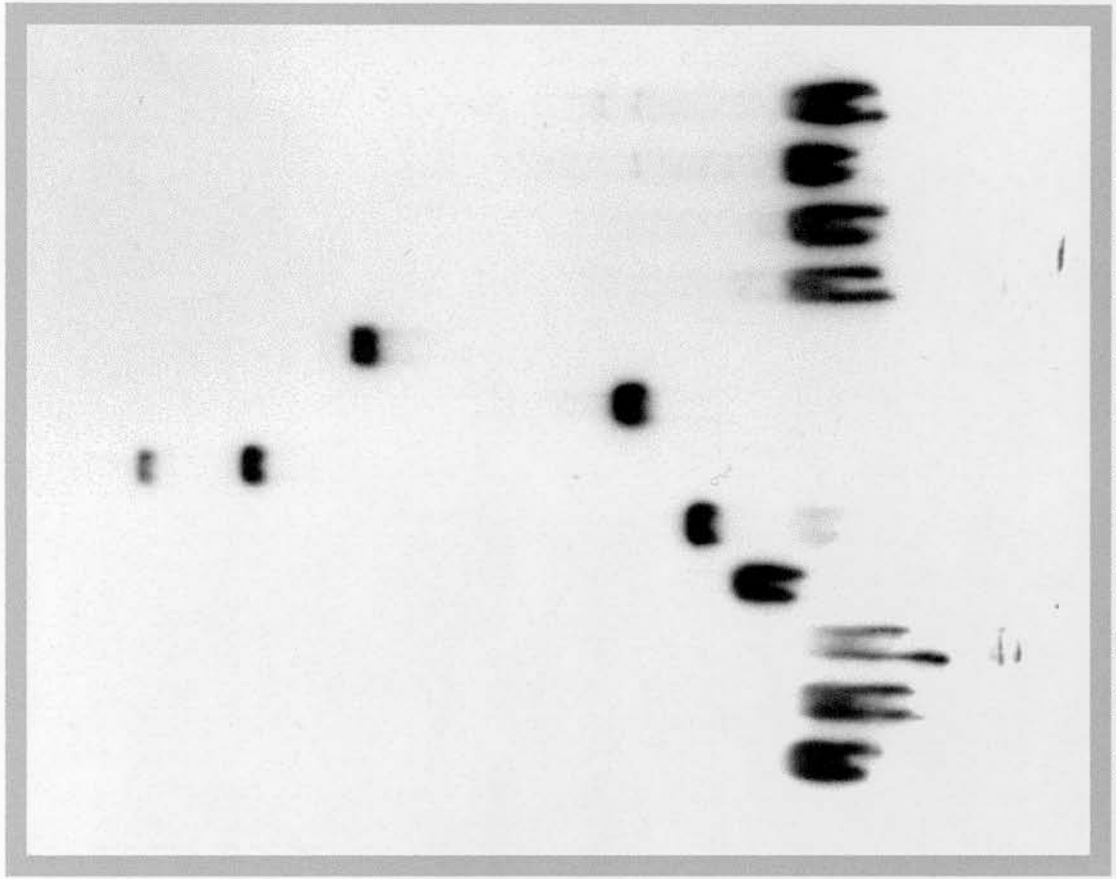

1 kb ladder
ES7 *Eco*R I
ES7 *Hin*d III
ES7 *Nde* I
ES7 *Pvu* I
ES7 *Sac* I
ES7 *Sca* I
ES7 *Stu* I
ES7 *Xba* I
ES7 *Xho* I
ES7 *Not* I
ES7 *Sal* I
ES7 *Spe* I
1 kb ladder

Figure 3.12a

Restriction enzyme digests of DNA from clone ES9, separated on an agarose gel in preparation for Southern blotting.

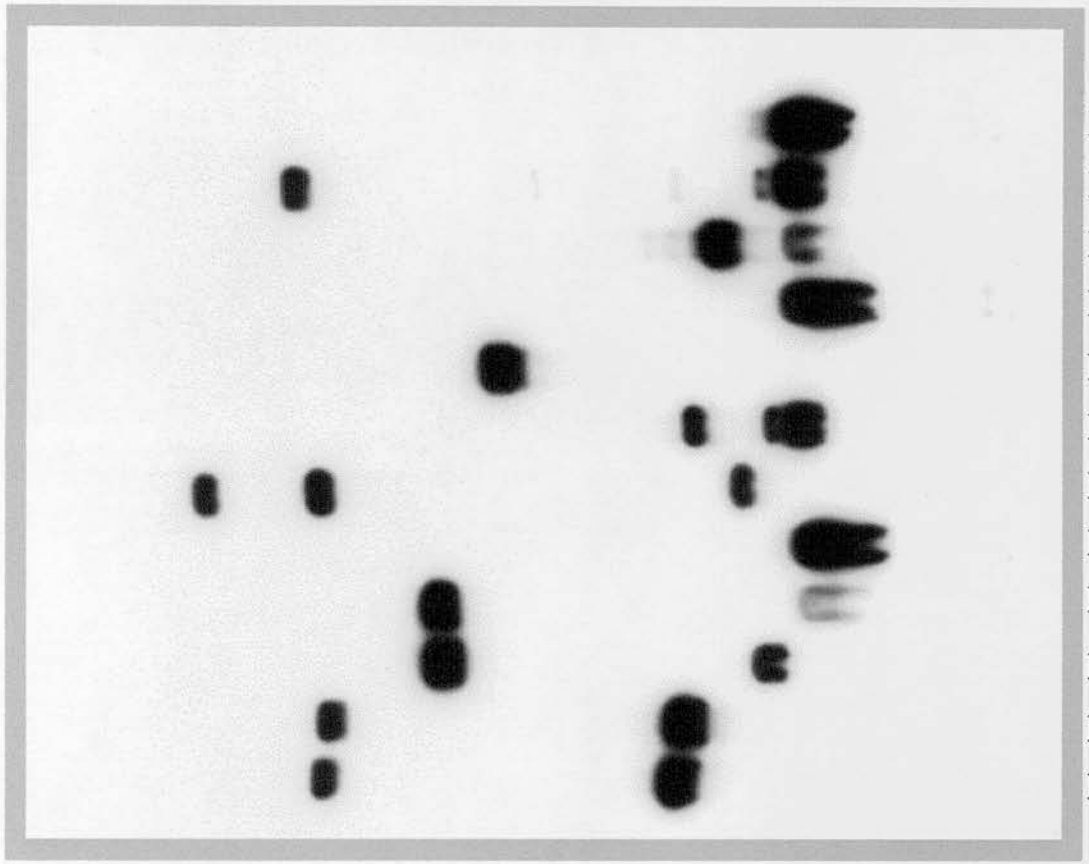
(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane).

1 kb ladder
ES9 EcoR I
ES9 Hind III
ES9 Nde I
ES9 Pvu I
ES9 Sac I
ES9 Sca I
ES9 Stu I
ES9 Sal I
ES9 Spe I
ES9 Xho I/Spe I
ES9 Xho I/EcoR I
ES9 Xho I/Hind III
1 kb ladder
12.2 kb
6.11 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb

Figure 3.12b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested ES9 DNA shown in Figure 3.12a is incubated with the radiolabelled Vipr2 rat 5' 260 bp cDNA probe

(The autoradiograph is shown at actual size)



1 kb ladder
ES9 *Eco*R I
ES9 *Hin*d III
ES9 *Nde* I
ES9 *Pvu* I
ES9 *Sac* I
ES9 *Sca* I
ES9 *Stu* I
ES9 *Sal* I
ES9 *Spe* I
ES9 *Xho* I/*Spe* I
ES9 *Xho* I/*Eco*R I
ES9 *Xho* I/*Hin*d III

Figure 3.13a

Restriction enzyme digests of DNA from clone ES3, separated on an agarose gel in preparation for Southern blotting.

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane).

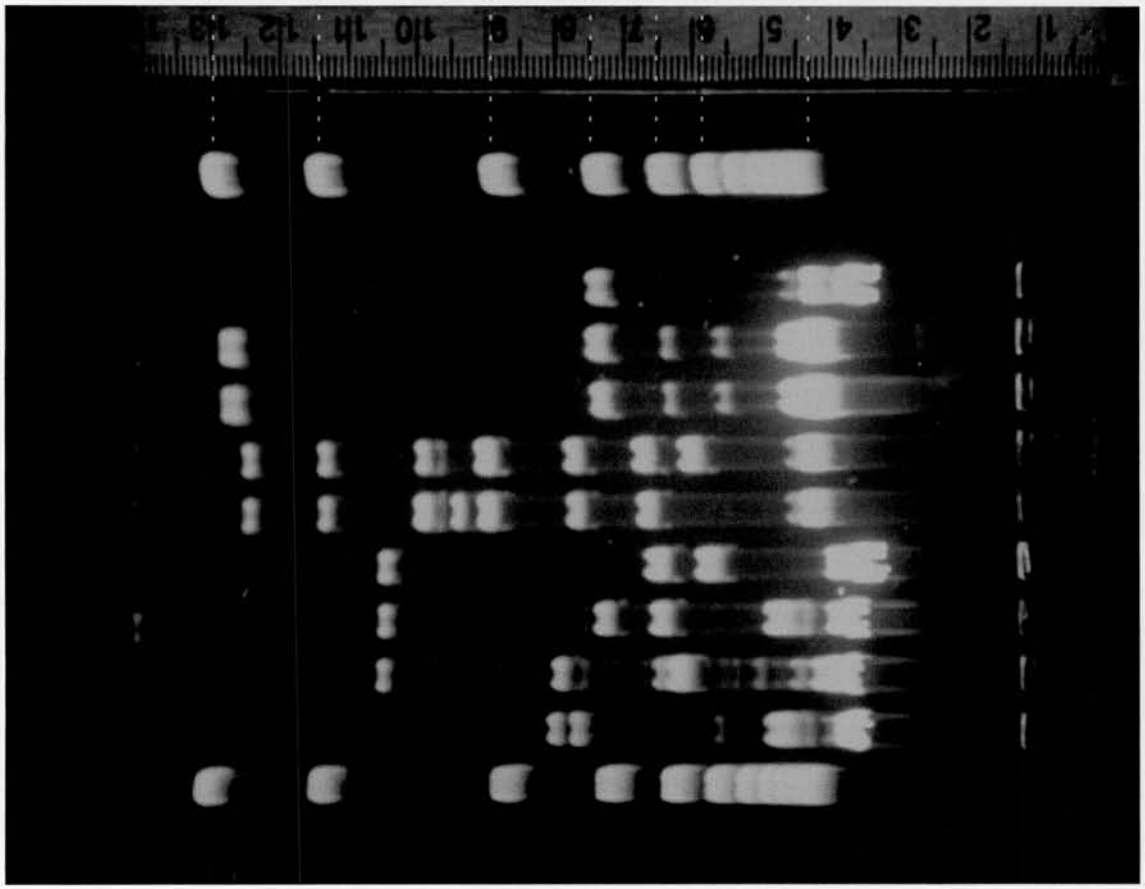

1.64 kb
2.04 kb
3.05 kb
4.07 kb
5.09 kb
6.11 kb
12.2 kb
1 kb ladder
ES3 *Xho* I
ES3 *Ava* I
ES3 *Xho* I/*Ava* I
ES3 *Pst* I
ES3 *Pst* I/*Xho* I
ES3 *Sca* I
ES3 *Xho* I/*Sca* I
ES3 *Sca* I/*Bam*H I
ES3 *Bam*H I/*Xho* I
1 kb ladder

Figure 3.13b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested ES3 DNA shown in Figure 3.13a is incubated with the radiolabelled Vipr2 rat 5' 1 kb cDNA probe

An ~4 kb hybridising *Xho* I fragment from ES3, which also appears to be present in ES11 is arrowed in green, whereas an ~5 kb *Bam*H I/*Sca* I fragment that hybridised to the probe and was also detected in Southern blots of ES11, is arrowed in blue.

(The autoradiograph is shown at actual size)

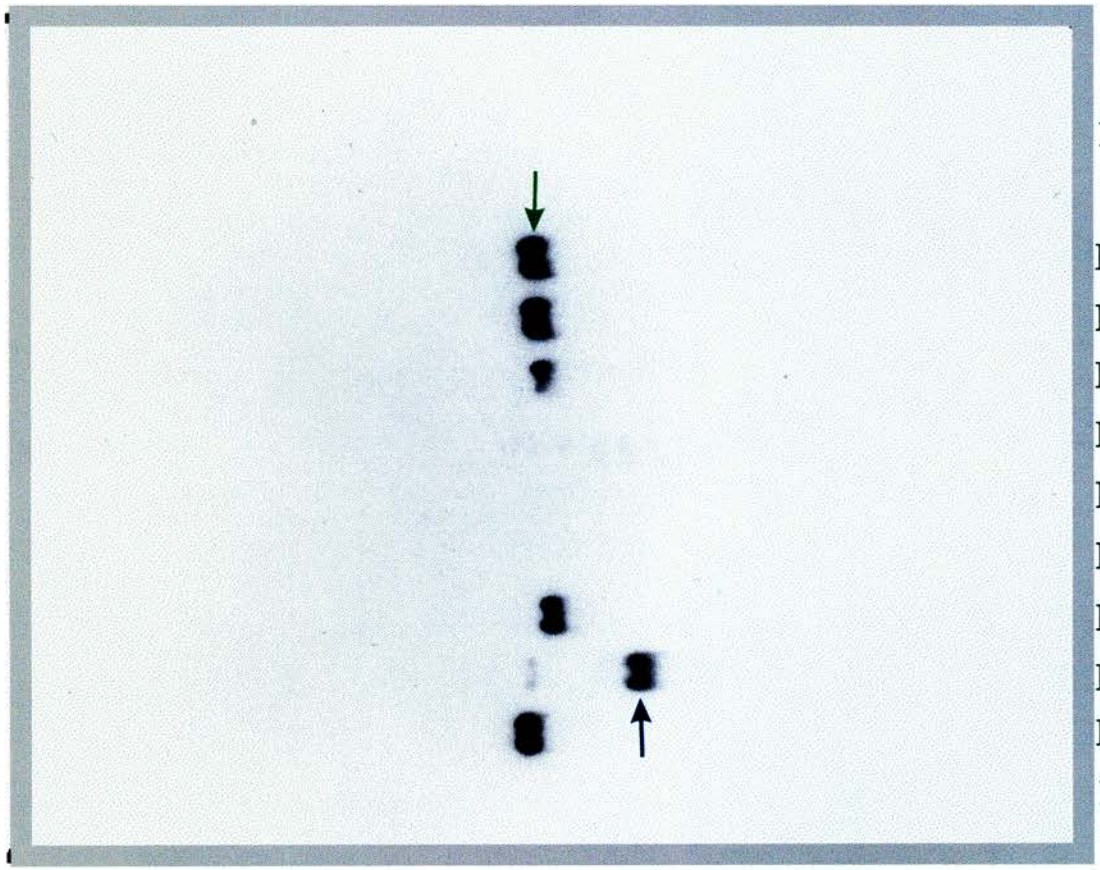

1 kb ladder
ES3 *Xho* I
ES3 *Ava* I
ES3 *Xho* I/*Ava* I
ES3 *Pst* I
ES3 *Pst* I/*Xho* I
ES3 *Sca* I
ES3 *Xho* I/*Sca* I
ES3 *Sca* I/*Bam*H I
ES3 *Bam*H I/*Xho* I
1 kb ladder

Figure 3.14a

Restriction enzyme digests of DNA from clone ES11, separated on an agarose gel in preparation for Southern blotting.

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane).

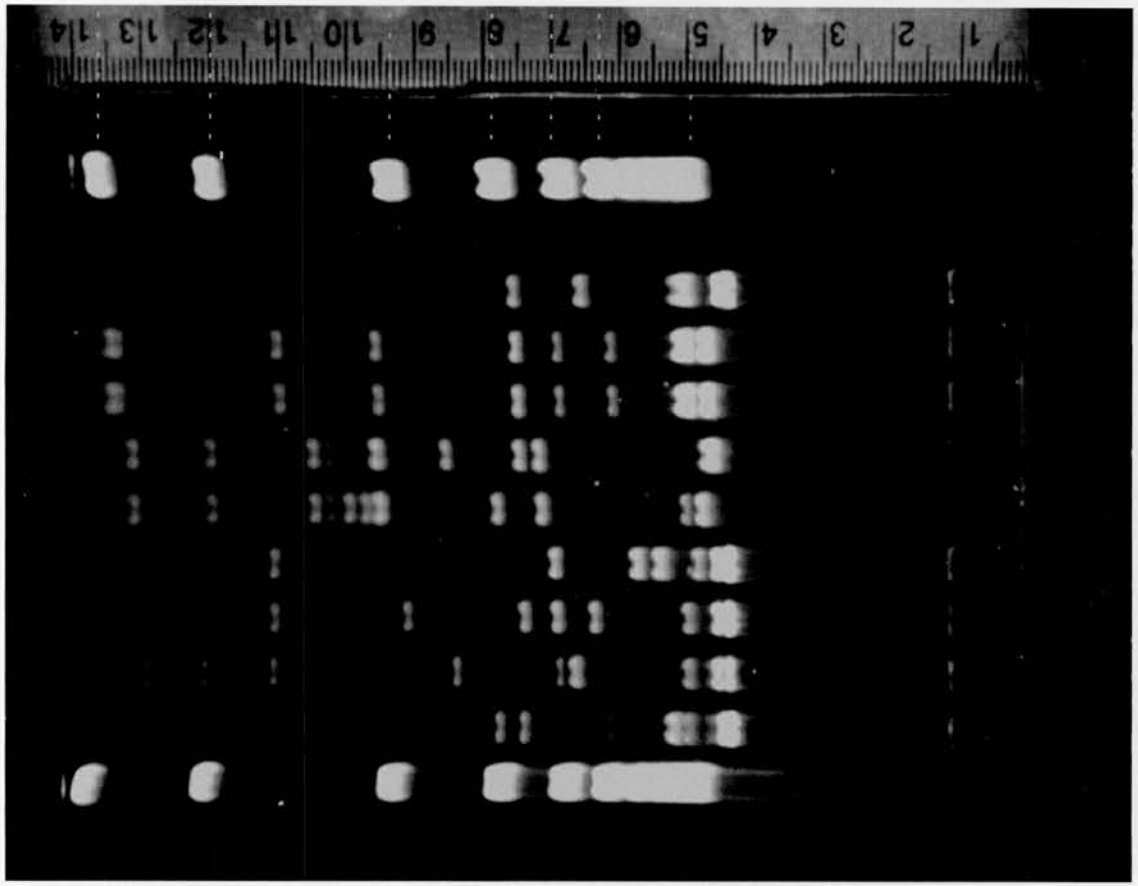
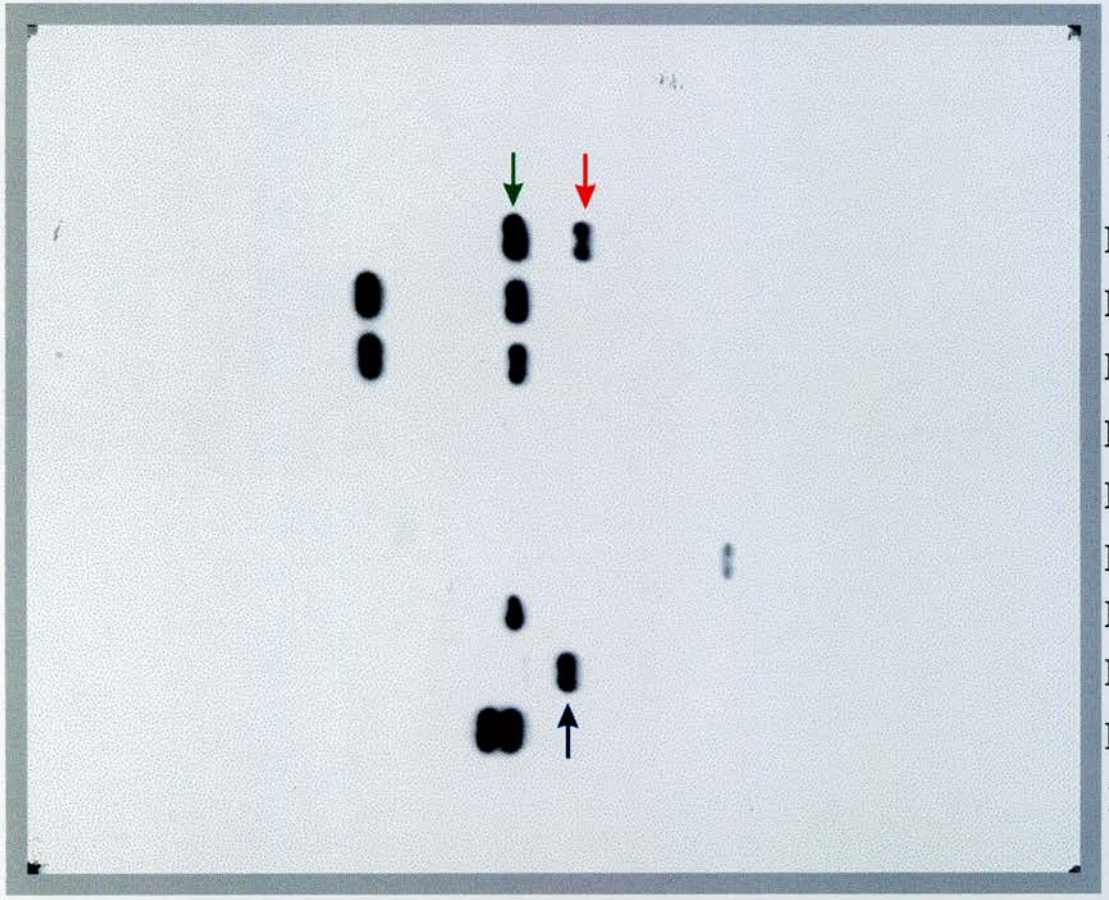

1.64 kb
2.04 kb
3.05 kb
4.07 kb
5.09 kb
6.11 kb
12.2 kb
1 kb ladder
ES11 *Xho* I
ES11 *Ava* I
ES11 *Xho* I/*Ava* I
ES11 *Pst* I
ES11 *Pst* I/*Xho* I
ES11 *Sca* I
ES11 *Xho* I/*Sca* I
ES11 *Sca* I/*Bam*H I
ES11 *Bam*H I/*Xho* I
1 kb ladder

Figure 3.14b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested ES11 DNA shown in Figure 3.14a is incubated with the radiolabelled Vipr2 rat 5' 1 kb cDNA probe

A hybridising *Bam*H I/*Sca* I fragment of ~5 kb that also appears to be present in ES3 and ES8, is arrowed in dark blue; and an ~4 kb hybridising *Xho* I fragment that is found in both ES11 and ES3 is arrowed in green. The ~5.5 kb *Xho* I fragment arrowed in red is unique to ES11 (i.e. is not present in ES3) and is thought to contain the region of ES11 that lies 5' to the start of ES3.

(The autoradiograph is shown at actual size)



1 kb ladder

ES11 *Xho* I

ES11 *Ava* I

ES11 *Xho* I/*Ava* I

ES11 *Pst* I

ES11 *Pst* I/*Xho* I

ES11 *Sca* I

ES11 *Xho* I/*Sca* I

ES11 *Sca* I/*Bam*H I

ES11 *Bam*H I/*Xho* I

1 kb ladder

ES3 and ES11. Indications that ES3 and ES11 may also overlap with the 5' end of clone ES8, were provided by the identification of an ~5 kb *Bam*H I/*Sca* I fragment from ES8, which (as in the case of the similar sized *Bam*H I/*Sca* I fragments from ES3 and ES11) hybridised to the 5' 1 kb cDNA probe (Figures 3.15a and 3.15b), and in agreement with sequencing data, which had shown that ES8 contained sequences corresponding to predicted exons 7-13 of the mouse *Vipr2* gene, the number of hybridising bands increased substantially when the ES8 blot was subsequently hybridised with the full length Vipr2 cDNA probe (Figure 3.16).

Having gained some idea of which regions of the *Vipr2* gene were contained within each of the λ clones, the λ vector primers 7363 and 8482 were used to obtain sequence data from the extreme 5' and 3' ends of the clones, and together with further sequencing studies in which the exact exon content of each clone was determined (using the exon specific primers listed in Table 3.1) this allowed the construction of a map of the positions of these clones relative to the exons and introns of the mouse *Vipr2* gene. The results obtained from these studies (and those from subsequent experiments in which the mouse *Vipr2* gene's intron sizes were determined; see sections 3.3.14 - 3.3.22), are summarised in Figure 3.17.

Primers that were designed to sequence outwards from the predicted exon sequences, across exon/intron boundaries, were then used (with λ DNA from the appropriate clone as a sequencing template) to determine the sequence of each of the exon/intron boundaries within the mouse *Vipr2* gene (with the exception of the 5' boundary of intron 1 for which an appropriate λ template had not yet been isolated). The sequences of all of the known exon/intron boundaries within the mouse *Vipr2* gene, and the positions of any of the mouse VIP_2 receptor's amino acids that are interrupted by these boundaries, are given in Table 3.2 The positions of the sequencing primers used are shown in Figure 3.18 and their sequences are given in Table 3.1

3.3.14 PCR amplification of introns from the mouse *Vipr2* gene

Once the intron/exon boundaries present within the insert sequences of the *Vipr2* gene clones had been established, attempts were made to amplify the intron sequences present within these clones, so that the sizes of the introns within the mouse *Vipr2* gene could be determined. Twelve pairs of primers were designed initially, one pair to span each intron of the mouse *Vipr2* gene, and following PCR reactions in which these primers were used with the appropriate λ clone as template DNA, a single strong band of PCR product was obtained from 6 of the PCR reactions (those designed to amplify introns 3, 7, 8, 9, 11, and 12; Figures 3.19 and 3.20).

Figure 3.15a

Restriction enzyme digests of DNA from clone ES8, separated on an agarose gel in preparation for Southern blotting.

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane).

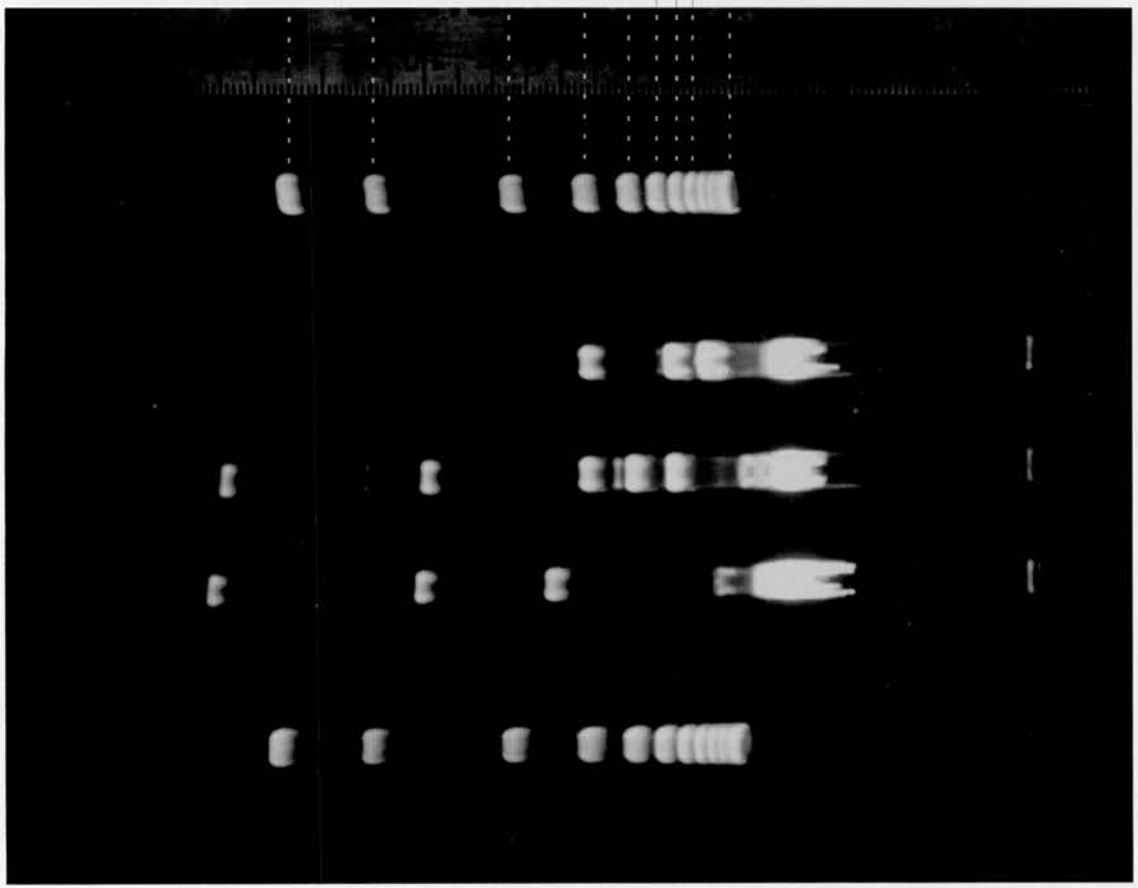

12.2 kb
8.14 kb
7.13 kb
6.11 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb
1 kb ladder
ES8 *Xho* I/*Bam*H I
ES8 *Bam*H I/*Sca* I
ES8 *Sca* I/*Xho* I
1 kb ladder

Figure 3.15b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested ES8 DNA shown in Figure 3.15a is incubated with the radiolabelled Vipr2 rat 5' 1 kb cDNA probe

An ~5 kb ES8 *Bam*H I/*Sca* I fragment, that hybridised to the 5' 1 kb probe, and also appears to be present in ES3 and ES11, is arrowed in dark blue.

(The autoradiograph is shown at actual size)

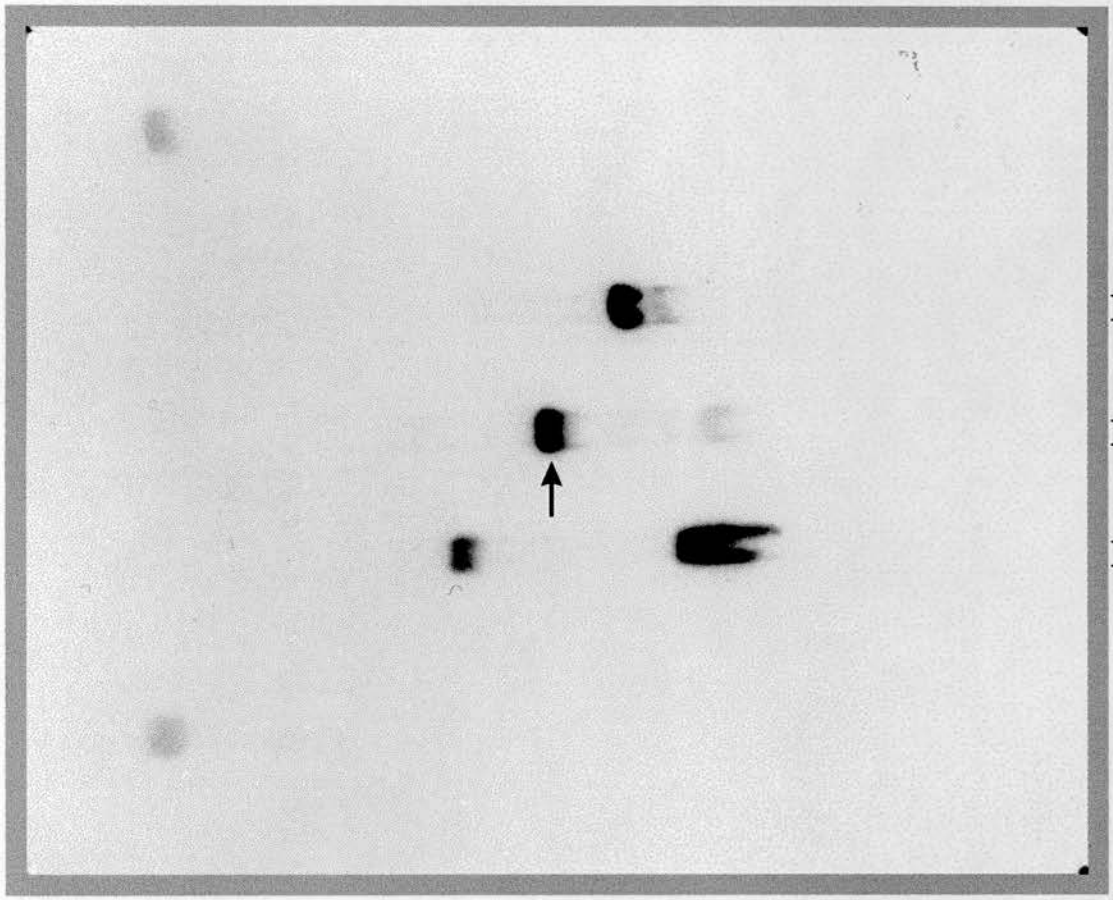
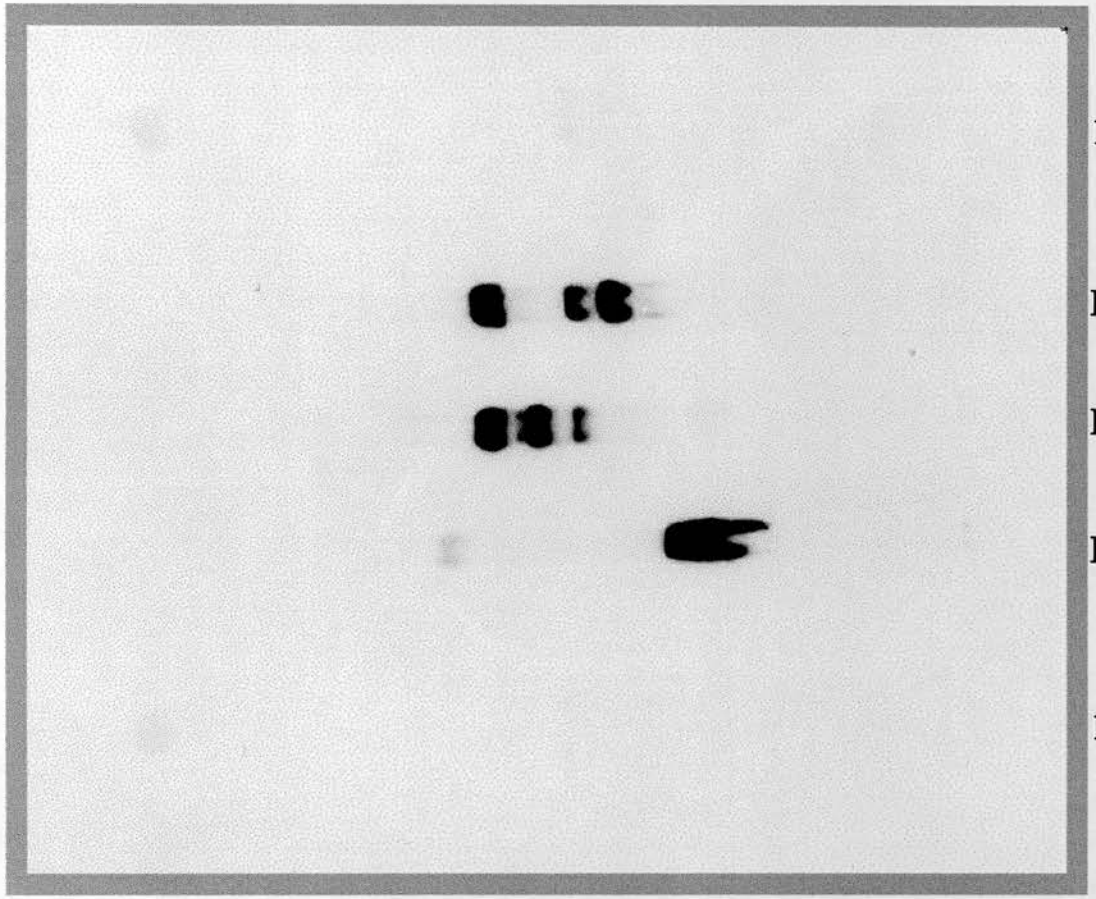

1 kb ladder
ES8 *Xho* I/*Bam*H I
ES8 *Bam*H I/*Sca* I
ES8 *Sca* I/*Xho* I
1 kb ladder

Figure 3.16

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested ES8 DNA shown in Figure 3.15a is incubated with the radiolabelled full-length rat Vipr2 cDNA probe

Notably more bands are seen to hybridise to the full-length probe than to the 5' 1 kb Vipr2 cDNA probe (see Figure 3.15b)

(The autoradiograph is shown at actual size)



1 kb ladder

ES8 *Xho* I/*Bam*H I

ES8 *Bam*H I/*Sca* I

ES8 *Sca* I/*Xho* I

1 kb ladder

Figure 3.17

Map of the mouse *Vipr2* gene (including a partial map of *Eco*R I/*Xho* I restriction sites)

*Eco*R I and *Xho* I digests of lambda clones encoding the mouse *Vipr2* gene, have so far only allowed the construction of a partial map of the *Vipr2* gene's *Eco*R I and *Xho* I restriction sites. The restriction map presented above is thought to be complete for the regions encompassed by clones ES9, ES11, ES3, ES8, and for the region of clone D1 that is shown (i.e. 3.2 kb of the putative promoter region, and the area from exon 1to the start of intron 2). However, within the intron 4 regions marked by blue arrows * additional unmapped *Eco*R I restriction sites are known to exist (at least 2 sites in ES7, 4 sites in ES10B), and in both ES7 and ES8 there are some regions (marked in grey as opposed to black) where the exact distances spanned by the clones are uncertain.

Key

E = *Eco*R I restriction site

X = *Xho* I restriction site

= 5' flanking sequence

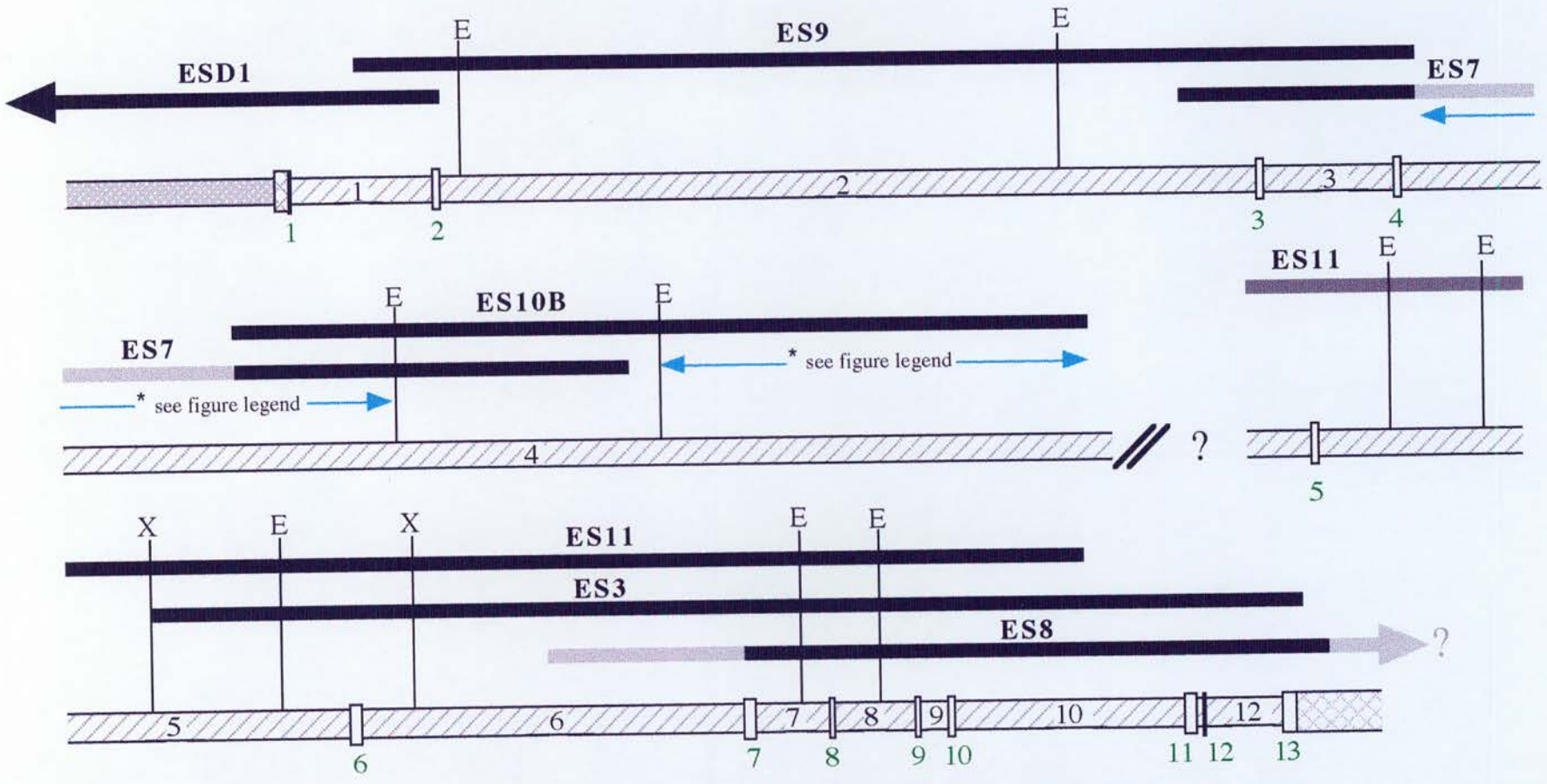
= intron sequences (1-12)

= 5' or 3' untranslated regions

= regions spanned by λ2001 *Vipr2* clones

= exon sequences (1-13)

// ? represents the site of a `gap' within the region of the λ contig that corresponds to intron 4 of the *Vipr2* gene.


E
ES9
E
ESD1
ES7
1
2
3
1
2
3
4
ES11
E
E
E
ES10B
E
ES7
* see figure legend
* see figure legend
4
?
5
X
E
X
ES11
E
E
ES3
ES8
?
5
6
6
7
8
9
10
12
7
8
9
10
11
12
13

Scale

1 kb

Table 3.2

Exon/Intron boundaries of the mouse *Vipr2* gene

The exon/intron boundaries surrounding each intron of the mouse *Vipr2* gene, are paired in this table. Exon sequences are shown in black and intron sequences are shown in green. In cases where the exon/intron junction interrupts an amino acid codon, the amino acid and its position are given in the right hand column. However, when the junction lies between codons, the amino acids situated on either side of the junction are shown in brackets.

Boundary Sequences	Amino Acid Interrupted
exon 1/ intron1 junction GCTGGTGCGGGTGAGTGAACCC intron1/exon 2 junction CTTCTATTTCTTTCACTTGAACAGGTGAGCAGCAT	None (Arg-16-Val-17)
exon 2/intron 2 junction GAATCAGAGAGGTAGGTTTGGGGA intron 2/exon 3 junction GCCTCTCCTTCTAGCCTGCAGCGGTGT	Ala-50
exon 3/ intron 3 junction TCTACAGCAGACCAGGTATTTGCATTGATT intron 3/ exon 4 junction AAATTTACTTTCCAGGAAACATAAGCAA	Gly-86
exon 4/ intron 4 junction GGATGAGAGTAAGGTAGGTGCCCGACCCG intron 4/ exon 5 junction TTTTTTCATTTCTTTGTAGATCTCGTTTTATA	None (Lys-118-Ile-119)
exon 5/intron 5 junction CTGCCTCTTCAGGTAAGGAACAAAA intron 5/ exon 6 junction CTCTCCCACTTCAGGAAGCTGCACTGCA	Arg-151
exon 6/ intron 6 junction CAGCCAGCCTCCTGGGTAAGACTATCCTG intron 6/ exon 7 junction TTCTCTCACTGTCTCTGGTTCCAGGTTGGCTGCAAGCTCAG	None (Trp-198-Val-199)
exon 7/intron 7 junction CTGATCGGATGGGGTAAGGGCTTCATATTC intron 7/ exon 8 junction ATGGGTCTCTCCACAGCATCCCCAGTGT	Gly-249

Boundary Sequences	Amino Acid Interrupted
exon 8/ intron 8 junction TAGAAGACACAGGGTAAGTGCTATATGTGC intron 8/ exon 9 junction TTTCTTTCTCTCTAACAGTTGCTGGGACACAAACG	Gly-269
exon 9/ intron 9 junction AATTTCTATTGTAGTAAGTAGCTGCC intron 9/ exon 10 junction TCCTTTCTTGTCTCTGTGCAGGTCAACTTTGCCCTCTT	None (Val-292-Val-293)
exon 10/ intron 10 junction AGTCACAGTACAAGTGAGTATCTGCTTG intron 10/ exon 11 junction TCCTTCTCCTTTTACTAGGAGGCTTGCCAAGTC	Lys-323
exon 11/ intron 11 junction TTGGTTCCTTCCAGGTAAGTATGGGAGGTG intron 11/ exon 12 junction CTCTGTCTCTTTGCAGGGCCTGGTGGTAGCAG	None (Gln-366-Gly-367)
exon 12/ intron 12 junction CTTCCTGAACAGTGAGGTAAGTCCTTTGTCAGC intron 12/ exon 13 junction CACGCTCATCCCCAGGTACAGTGTGAACTGA	None (Glu-380-Val-381)

Figure 3.18

Diagram showing the structure of the mouse *Vipr2* gene, and the approximate locations of the sequencing primers that were used in the determination of the gene's exon/intron boundaries.

The gene structure has been drawn to a scale of 1 cm =1 kb. Regions of the gene that have been sequenced are represented by a green line (see Appendix 1 for sequence data); unsequenced regions are shown in black. The approximate positions of primer binding sites are indicated by black arrows below the exons (or introns in the case of intron 10). Primer sequences can be found (listed according to primer number) in Table 3.1. For the positions of the binding sites for sequencing primers that were used in the characterisation of the 5' flanking region of the gene (i.e. the region upstream of exon 1), see Figure 6.4 (Chapter 6).

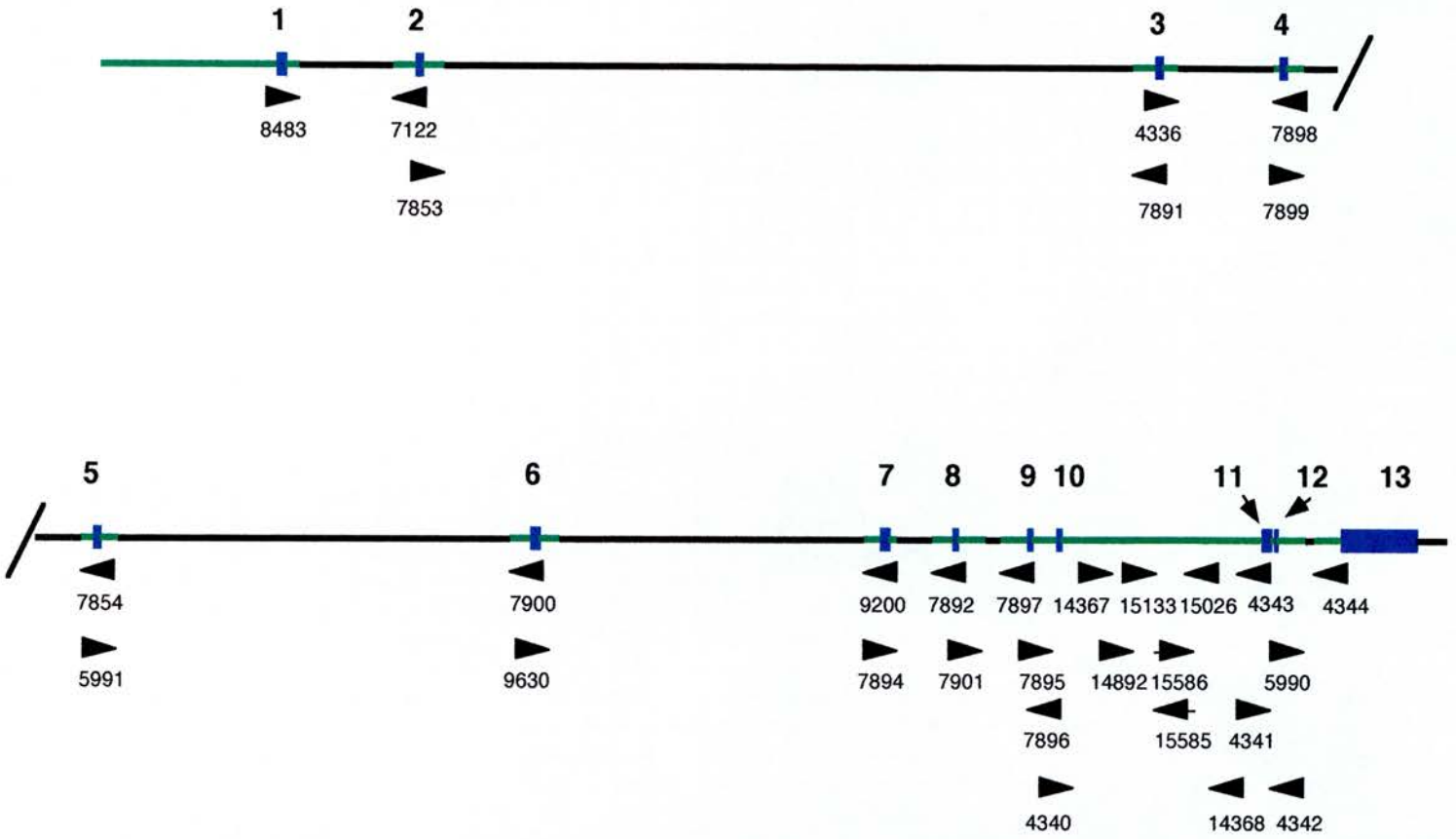

1
2
3
4
8483
7122
7853
4336
7898
7891
7899
5
6
7
8
9
10
11
12
13
7854
5991
7900
9630
9200
7894
7892
7901
7897
7895
7896
4340
14367
14892
15586
15585
4341
14368
15133
15026
4343
5990
4342
4344

Figure 3.19

PCR products obtained from the amplification of intron 3 of the mouse *Vipr2* gene

Reactions in which the exon 3 primer 4336 and the exon 4 primer 7898 were used on ES9 template DNA, resulted in the production of an intron 3 product of ~2.0 kb.

(Also shown on this gel is the ~2.5 kb product produced by primer 4336 and the λ2001 vector primer 7363 with ES9 template DNA.)

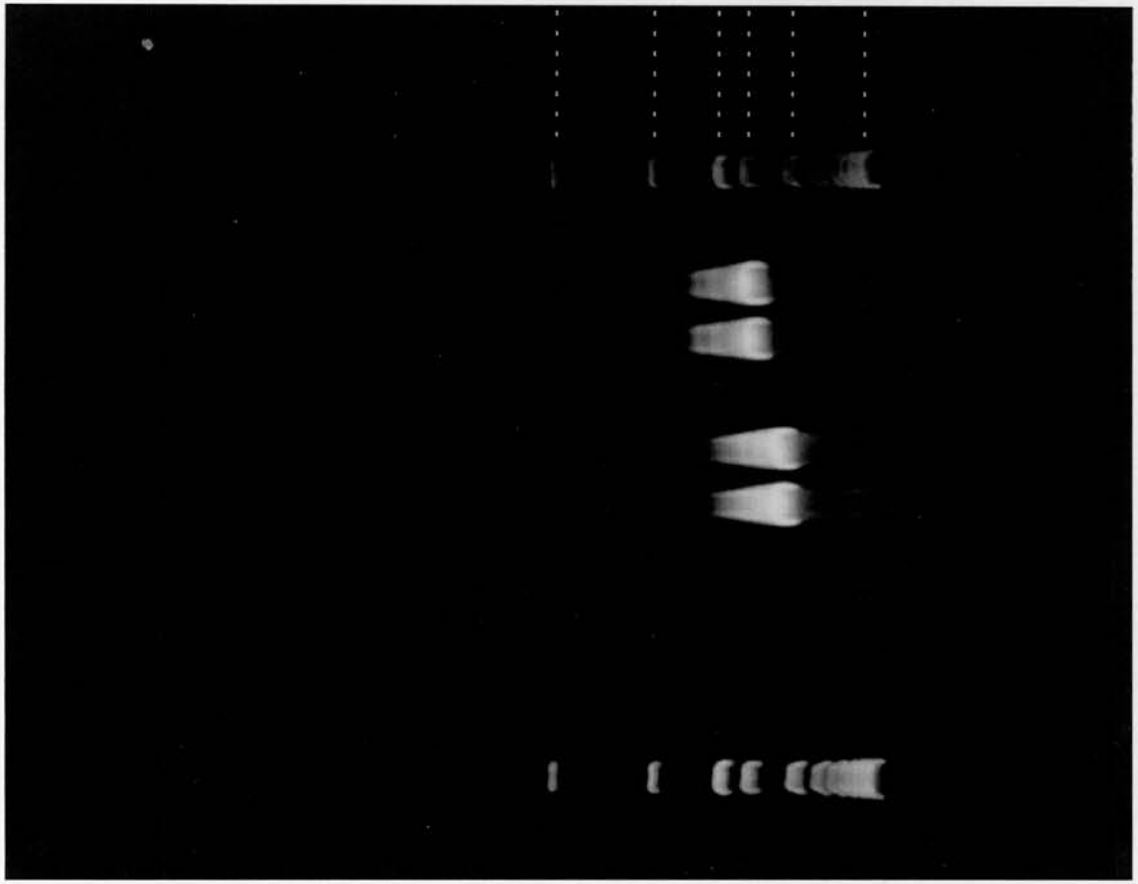
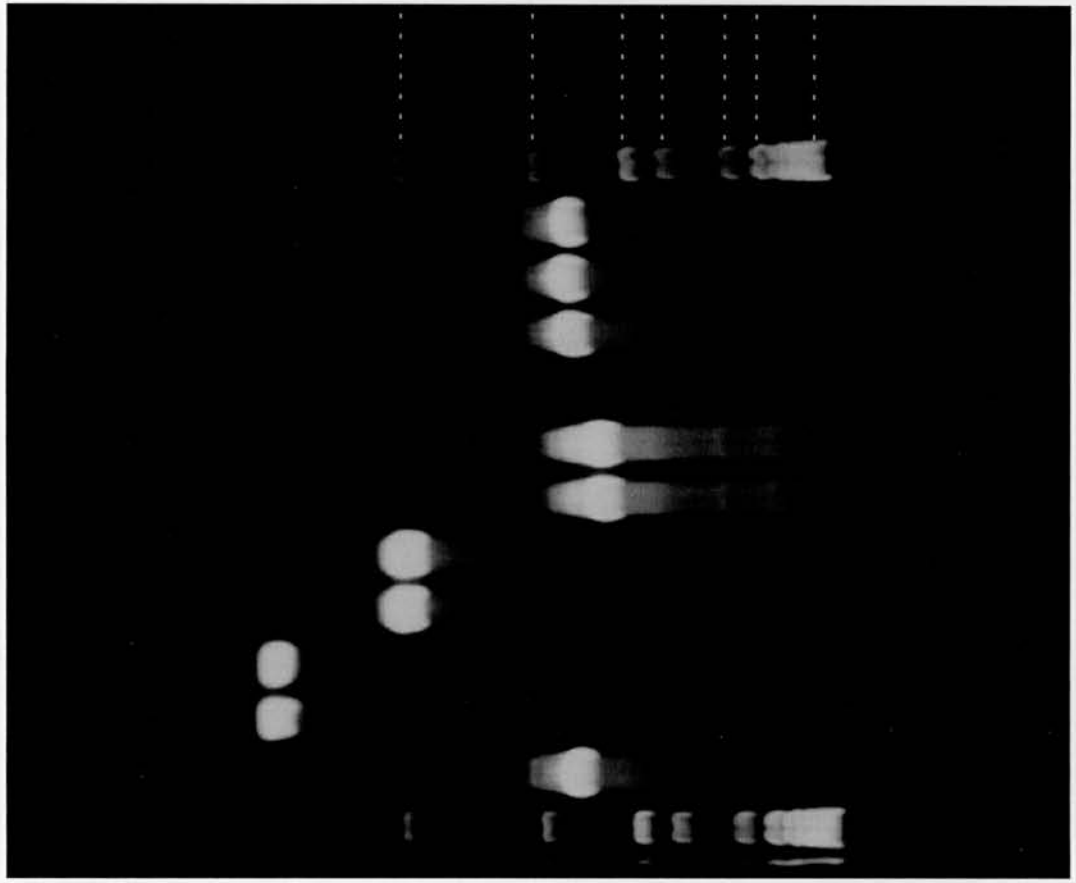

12.2 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
~517 bp
1 kb ladder
ES9 4336/7898 (int.3)
ES9 4336/7898 (int.3)
ES9 4336/7363
ES9 4336/7363
-ve control 4336/7898
-ve control 4336/7363
1 kb ladder

Figure 3.20

PCR products obtained from the amplification of introns, 7, 8, 9, 11, and 12 of the mouse *Vipr2* gene

ES8 DNA was used as a template in all cases, and a strong single band of DNA was produced from each of the amplification reactions.
Exon 7 primer 7894 and exon 8 primer 7892 produced an intron 7 product of ~1.1 kb
Exon 8 primer 7901 and exon 9 primer 7897 produced an intron 8 product of ~1.2 kb
Exon 9 primer 7895 and exon 10 primer 7896 produced an intron 9 product of 405 bp (previously determined by sequencing)
Exon 11 primer 4341 and exon 12 primer 4342 produced an intron 11 product of 66 bp (also previously determined by sequencing)
Exon 12 primer 5990 and exon 13 primer 4344 produced an intron 12 product of ~1.1 kb


12.2 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
~517 bp
1 kb ladder
ES8 5990/4344 (int.12)
ES8 7894/7892(int.7)
ES8 7894/7892(int.7)
ES8 7901/7897(int.8)
ES8 7901/7897(int.8)
ES8 7895/7896(int.9)
ES8 7895/7896(int.9)
ES8 4341/4342(int.11)
ES8 4341/4342(int.11)
ES8 5990/4344 (int.12)
1 kb ladder

The PCR primers (4340 and 4343) that had been used in the initial attempt to amplify intron 10, resulted in the generation of two bands of PCR products, one of ~ 1.2 kb and one of ~3.5 kb (Figure 3.21). However, PCR reactions that were carried out using the same exon 10 primer (4340) with the exon 12 primer (4342), also produced a single band of PCR product that was in the region of 3.5 kb (Figure 3.22), and as sequencing of the λ clones had previously established that exon 11 and intron 11 were 130 bp and 66 bp long respectively, it appeared that intron 10 probably spans about 3.5 kb.

Similar problems were encountered in attempts to amplify intron 5 (using primers 5991 and 7900) and intron 6 (using primers 5993 and 9200), although in these cases the problem was more extreme, as the primers produced multiple PCR products over a wide range of sizes. Nevertheless, in both cases the replacement of the original exon 6 primer (7900 for intron 5 and 5993 for intron 6) with an alternative primer (9628 for intron 5 and 9630 for intron 6), resulted in the generation of a single strong band of DNA which was the predominant PCR product. In the case of the intron 5 amplification reactions this product was ~600 bp long, whereas the predominant product from the intron 6 amplification reactions was ~1.2 kb long (Figure 3.23).

The failure of clone ES9 to sequence with primers that had been designed to bind to predicted exon 1 of the mouse *Vipr2* gene, suggested that this clone did not extend 5' of intron 1 of the gene, a theory which was supported by the finding that sequencing of this clone with the vector primer 8482 (which sequenced the 5' end of the ES9 clone) did not produce recognisable *Vipr2* exon sequence. However, as it remained possible that the failure of the clone to sequence with exon 1 primers was due to the high GC content of the first exon of the *Vipr2* gene, PCR reactions were carried out in which the exon 2 primer 7122 and the vector primer 8482 were used to amplify those sequences within clone ES9 that lay upstream of exon 2. A PCR product of ~1.2-1.3 kb was produced in these reactions (Figure 3.24), and was subcloned into pGEM T-vector (Promega; see Figure 2.3 for vector map) in preparation for sequencing.

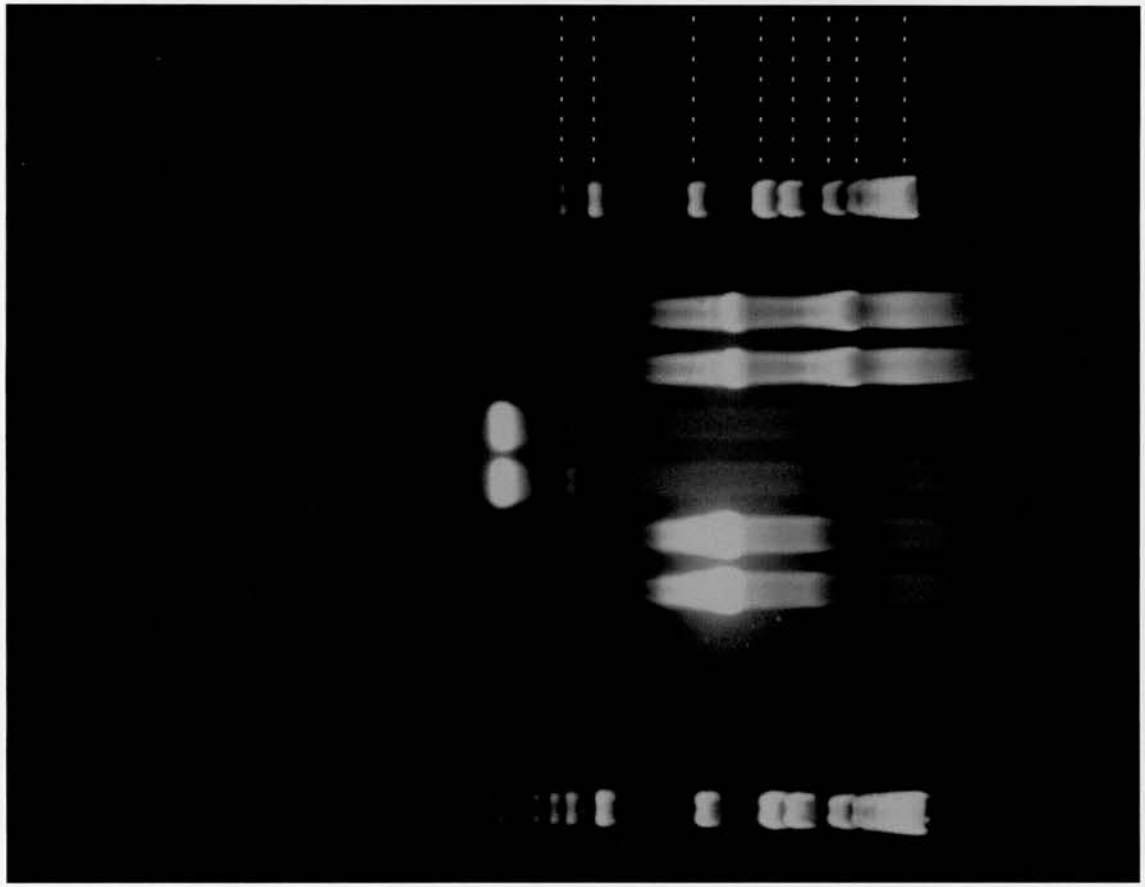
Attempts to amplify intron 2 from ES9 template DNA (using primers 7853 and 7891, or 7893 and 7891) were unsuccessful, and as indicated in section 3.3.13, none of the clones that had been characterised spanned intron 4 of the *Vipr2* gene.

The primer pairs used in the amplification of the mouse *Vipr2* gene introns, and the results obtained, are summarised in Table 3.3

Figure 3.21

PCR products obtained from the amplification of introns 10, 11, and 12 of the mouse *Vipr2* gene

In attempts to amplify intron 10 (using the exon 10 primer 4340 and the intron 11 primer 4343 on ES8 template DNA) PCR products of ~3.5 kb and ~1.2 kb were produced. The intron 11 and intron 12 amplification reaction products shown on this gel have previously been shown in Figure 3.20.



1kb ladder

ES8 4340/4343 (int.10)

ES8 4340/4343 (int.10)

ES8 4341/4342 (int. 11)

ES8 4341/4342 (int. 11)

ES8 5990/4344 (int.12)

ES8 5990/4344 (int.12)

-ve control 4340/4343

-ve control 4341/4342

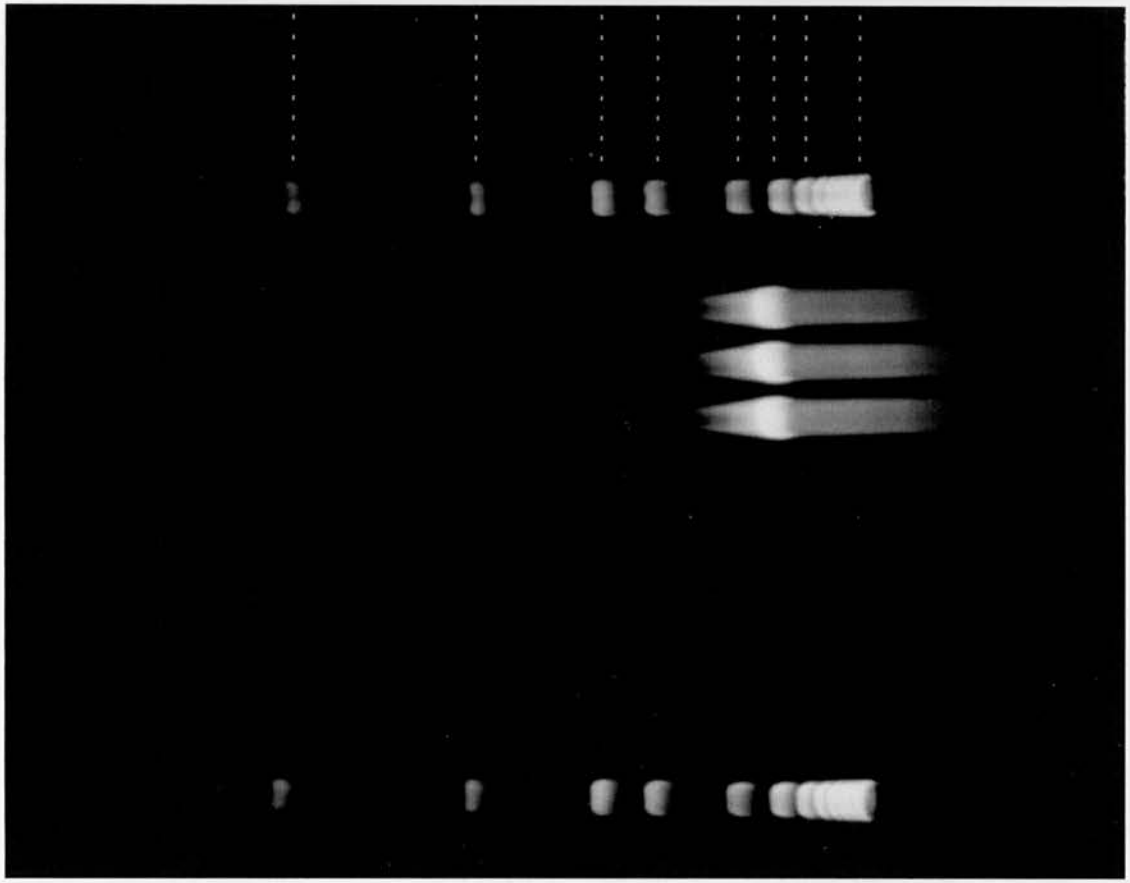
-ve control 5990/4344

1 kb ladder

Figure 3.22

PCR products obtained from the amplification of a fragment that spans intron 10 and intron 11 of the mouse *Vipr2* gene

The exon 10 primer 4340 and exon 12 primer 4342 were used with ES8 template DNA to amplify a fragment of *Vipr2* DNA that spans intron 10 and intron 11. The reaction was set up in triplicate, and in each case a single PCR product of ~3.5 kb can be seen.



1 kb ladder

ES8 4340/4342 (int.10+11)

ES8 4340/4342 (int.10+11)

ES8 4340/4342 (int.10+11)

1 kb ladder

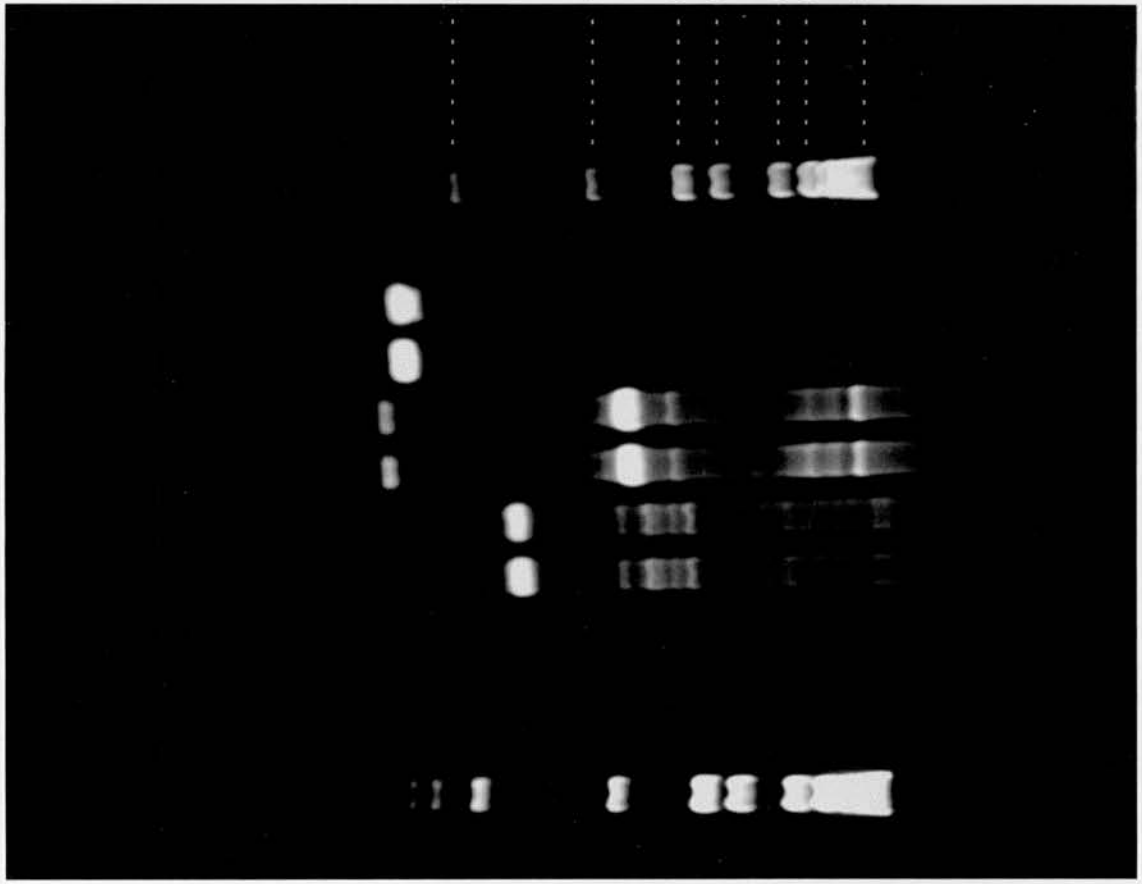
Figure 3.23

PCR products obtained from attempts to amplify intron 5 (using primers 5991 and 9628) and intron 6 (using primers 9630 and 9200 or 8482 and 9200)

In intron 6 amplification reactions using the exon 6 primer 9630 and the exon 7 primer 9200 on ES11 template DNA, a predominant band of 1.2 kb can be seen.

In intron 5 amplification reactions using the exon 5 primer 5991 and the exon 6 primer 9628 on ES11 template DNA, a predominant band of 600 bp can be seen.

[Intron 6 amplification reactions using the λ2001 vector primer 8482 and exon 7 primer 9200 on ES3 template DNA were unsuccessful (only low molecular weight primer dimers were produced)]



1 kb ladder

ES3 9200/8482

ES3 9200/8482

ES11 9200/9630 (int.6)

ES11 9200/9630 (int.6)

ES11 9628/5991 (int.5)

ES11 9628/5991 (int.5)

-ve control 9200/8482

-ve control 9200/9630

-ve control 9628/5991

1 kb ladder

Figure 3.24

PCR products obtained from the amplification of the region of intron 1 of the *Vipr2* gene that is present at the 5' end of clone ES9

Two primer pairs were tried in these reactions:
i) the λ2001 vector primer 8482 and the exon 1 primer 8484 (ES9 template DNA)
ii) the λ2001 vector primer 8482 and the exon 2 primer 7122 (ES9 template DNA)
However, as can be seen from the gel opposite, the 8482/8484 reaction appeared unsuccessful, whereas the 8482/7122 reactions produced an extremely strong band of PCR product which had a size of approximately 1.2-1.3 kb.

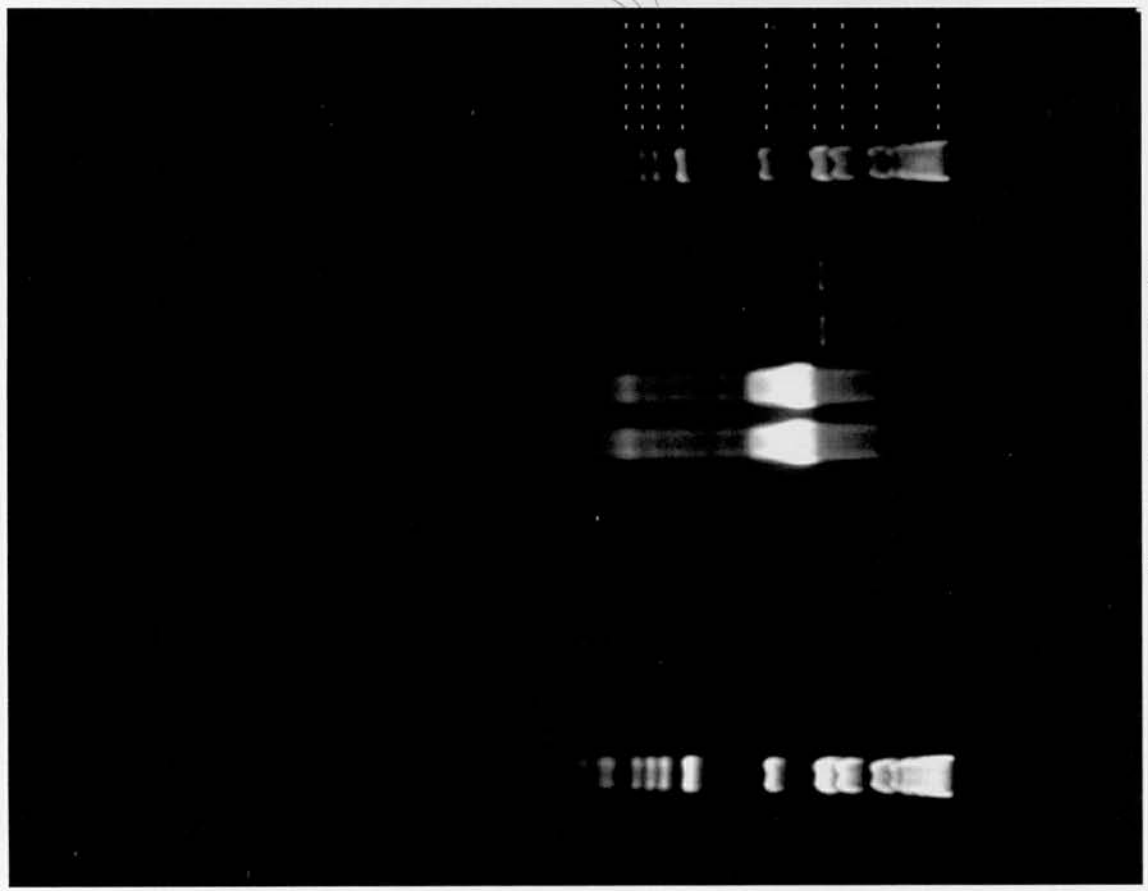

12.2 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
~517 bp
396 bp
344 bp
298 bp
1 kb ladder
ES9 8482/8484
ES9 8482/8484
ES9 8482/7122
ES9 8482/7122
-ve control 8482/8484
-ve control 8482/7122
1 kb ladder

Table 3.3

Summary of initial attempts to amplify the introns of the mouse *Vipr2* gene by PCR

(-) indicates that a reaction was not attempted or was unsuccessful

(√) indicates that the sequencing reaction was successful and produced the expected data

(×) indicates that the sequencing reaction was successful, but did not produce the expected data (i. e. the PCR reaction had not amplified the desired intron)

* indicates that the size of the intron had previously been determined from sequencing data

INTRON	UPSTREAM PRIMER	DOWNSTREAM PRIMER	TEMPLATE DNA	SIZE OF PREDOMINANT PCR PRODUCT	SEQUENCING RESULTS AS EXPECTED FOR INTRON
1	-	-	-	-	-
2	7853	7891	ES9	-	-
	7893	7891	ES9	-	-
3	4336	7898	ES9	~2.0 kb	√
4	-	-	-	-	-
5	5991	7900	ES11	-	-
5	5991	9628	ES11	~0.6 kb	×
6	5993	9200	ES11	-	-
6	9630	9200	ES11	~1.2 kb	×
7	7894	7892	ES8	~1.1 kb	√
8	7901	7897	ES8	~1.2 kb	√
9	7895	7896	ES8	0.4 kb*	√*
10	4340	4343	ES8	-	-
10+11	4340	4342	ES8	~3.5 kb	√
11	4341	4342	ES8	0.066 kb*	√*
12	5990	4344	ES8	~1.2 kb	√

3.3.15 Direct sequencing of the PCR products obtained from the amplification of the mouse *Vipr2* gene's introns

Sequencing of the intron PCR products with primers that had been used in the original PCR reactions, demonstrated that the PCR products from introns 3, 7, 9, 10 + 11, and 12, contained sequence corresponding to that found at the appropriate *Vipr2* gene intron/exon junctions in the λ clone inserts. The intron 8 PCR product was also found to contain the expected sequences, although in this case the proximity of the PCR primers to the ends of the exons, meant that the PCR product had to be cloned into a plasmid vector and sequenced using vector primers in order to establish the sequence of the intron/exon boundaries. In contrast, the PCR products from introns 5 and 6, were each found to contain the expected sequences at one end of the PCR product but not at the other, and it appeared that in both cases the primer that had been designed to bind to exon 6 (primer 9628 in the intron 5 PCR reaction, and 9630 in the intron 6 PCR reaction) had mis-primed.

Subcloning and Sequencing of intron 10 of the mouse Vipr2 gene

As intron 10 of the mouse *Vipr2* gene was positioned within the region predicted to encode the third intracellular loop of the VIP_2 receptor, a site at which splice variation had been detected (Mayo, 1992; Spengler *et al.*, 1993) in transcripts encoding two other members of the secretin/glucagon receptor subfamily, the complete sequence of intron 10 was determined in an attempt to establish whether there was the potential for splice variation to occur within this region of the mouse *Vipr2* gene. Following the subcloning of the intron 10+11 PCR product (see 3.3.14 for details of the PCR reaction), a total of 9 primers were used to successively sequence inwards from either end of intron 10 (Figure 3.18). However, although analysis of the sequence data obtained from intron 10 (see Appendix 1) revealed that a sequence matching the consensus sequence for splice acceptor sites (Mount, 1982) was present at bp 1315-1342 of intron 10, subsequent analysis of potential open reading frames (ORFs) and a search for possible splice donor sites did not lead to the identification of any potential patterns of splice variation that seemed likely to lead to the generation of receptor variants.

3.3.16 Further attempts at the PCR amplification of introns 2, 4, 5 and 6 of the mouse *Vipr2* gene

Of the introns whose sizes had not yet been determined, there was reason to suspect that at least two of these introns, introns 2 and 4, might be too large to amplify by

conventional PCR methods. In the case of intron 4, these suspicions were based purely on the fact that the Vipr2 cDNA probe had failed to detect any clones that spanned intron 4 of the *Vipr2* gene. However, for intron 2, strong evidence in support of this idea was provided by the finding that intron 3 spanned only ~2 kb (see 3.3.14), as it had previously been established that clone ES9 (which contains introns 2 and 3) began ~1.2 kb before the end of intron 1 (see 3.3.14) and ended a short distance (365 bp) into intron 4 (Figure 3.17). Consequently, during further attempts to amplify intron 2 from ES9 template DNA, and attempts to isolate intron 4 from total mouse genomic DNA, long PCR reagents and thermocycling conditions were used (which are reported to be capable of amplifying fragments of up to 27 kb from λ DNA and up to 40 kb from total genomic DNA, see 3.2.14). Nevertheless, despite these efforts to extend the size range of the fragments that could potentially be amplified during these reactions, no PCR products with sizes that were compatible with the predicted minimum sizes of introns 2 or 4 were obtained.

As there was no reason other than the failure of initial PCR attempts, to suggest that introns 5 and 6 of the mouse *Vipr2* gene were particularly large, further attempts to amplify these introns were made using standard PCR conditions. In both cases it appeared that a lack of specificity on the part of the exon 6 primers was largely to blame for the failure of the initial PCR reactions, but this was puzzling, as all of these exon 6 primers generated the expected sequence data from ES11 template DNA in sequencing reactions. Since the main difference between the conditions used for sequencing reactions and those used in the PCR reactions was the primer/template annealing temperature, which was 57°C in the PCR reactions as opposed to 50°C in sequencing, the effects of lowering the annealing temperature in the PCR reactions were investigated. However, in contrast to the clear results obtained from sequencing reactions, the use of a 50°C annealing temperature in PCR reactions aimed at amplifying either intron 5 or intron 6, resulted in the production of multiple PCR products which ranged in size from <500 bp to >5 kb (Figure 3.25).

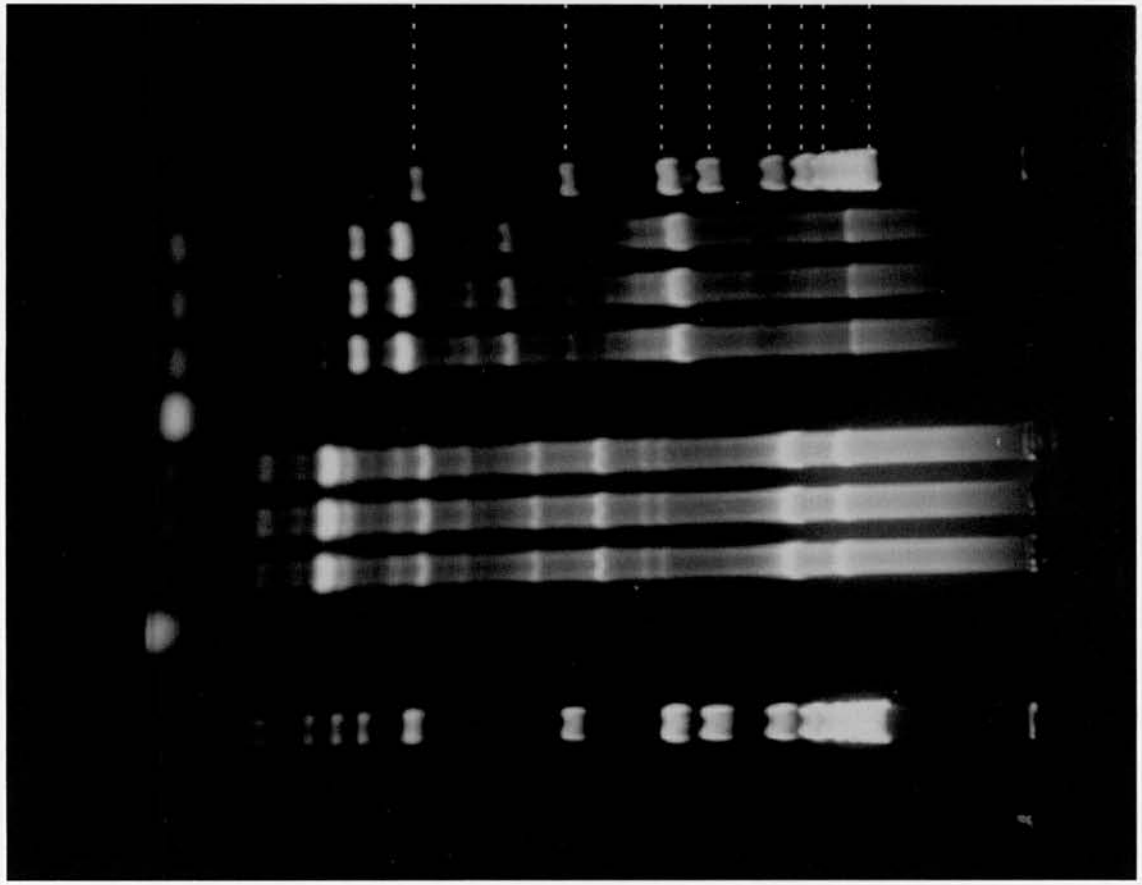
3.3.17 Rescreen of the λ2001 mouse genomic library, for clones that span intron 4 of the mouse *Vipr2* gene

In order to try to isolate a genomic clone that spanned intron 4 of the mouse *Vipr2* gene, a rat Vipr2 cDNA probe that spanned the coding region corresponding to exon 3-exon 6 of the mouse gene, was amplified from the full length rat Vipr2 cDNA sequence using primers 4336 and 9629 (Table 3.1), radiolabelled, and then used to rescreen the λ2001 mouse genomic library. Six positive clones (GC4, GD2, GD3, GD6, GD7, and GD9) were identified in the primary screen, and following secondary and tertiary

Figure 3.25

PCR products obtained in attempts to isolate intron 5 and intron 6 of the mouse *Vipr2* gene using $MgCl_2$ concentrations and primer annealing temperatures similar to those used for automated fluorescent cycle sequencing.

Reactions using primer pair 5991/9629 (designed to amplify intron 5) and those using primer pair 9630/9200 (designed to amplify intron 6), were carried out with ES11 template DNA. However in both cases a large number of bands were produced including several that were of fairly high intensity, a result that made it very difficult to identify the 'correct' product.



1 kb ladder

ES11 5991/9629 (int.5)

ES11 5991/9629 (int.5)

ES11 5991/9629 (int.5)

ES11 9630/9200 (int.6)

ES11 9630/9200 (int.6)

ES11 9630/9200 (int.6)

1 kb ladder

screens, 5 positive clones (GC4, GD3, GD6, GD7, and GD9) were isolated. However, sequencing of DNA prepared from each of these clones, showed that 4 of the clones (GC4, GD3, GD6, and GD7) could be successfully sequenced with the exon 4 primer 7899, but not the with the exon 5 primer 7854 or the exon 6 primer 7900, whereas the other clone (GD9) sequenced with the exon 5 and 6 primers but not with the exon 4 primer. Therefore, although sequencing reactions with the λ vector primers 7363 and 8482 indicated that the 5 clones were all different from one another, it appeared that none of these clones spanned intron 4 of the mouse *Vipr2* gene.

3.3.18 Rescreen of the λ2001 mouse genomic library, for clones that contain the 5' end of the mouse *Vipr2* gene

Although the mouse and rat Vipr2 cDNA sequences showed very little divergence within the region of coding sequence that lay directly downstream of the translation start site, no clones containing sequence from predicted exon 1 of the mouse *Vipr2* gene had been isolated in the initial screen of the mouse genomic library. However, although there appeared to be very little sequence divergence between the 5' UTRs of the mouse and rat Vipr2 cDNA sequences, it was possible that the relatively short region (~90 bp) of homology between the rat and mouse sequences prior to intron 1 had not been sufficient for the detection of clones containing this region of the gene. Therefore, when a 5' RACE PCR product (Q12) from the mouse Vipr2 cDNA was isolated in our laboratory, and was shown to contain an additional 34 bp of sequence extending 5' of the start of the rat Vipr2 cDNA, a decision was made to rescreen the library using an ~ 190 bp *Pst* I/*Xba* I restriction fragment from the 5' end of the 5' RACE product as a probe.

Hybridisation of the radiolabelled 5' RACE product restriction fragment (Q12 *Pst* I/*Xba* I) to the plaque lift filters (previously used for the 'intron 4 rescreen' see 3.3.17) resulted in the identification of 3 positive clones (ESC4, ESC5, and ESD1). However the signal from clone ESC5 was weaker than that from the other two clones, and in secondary screens only the ESC4 and ESD1 plaque picks produced positive hybridisation results. Initial sequence analysis of ESC4 and ESD1 DNA (prepared from liquid lysates, using Qiagen columns) was carried out using the mouse *Vipr2* gene predicted exon 1 primers 8483 and 8484 . DNA from clone ESC4 could not be successfully sequenced with these primers, but the sequence data obtained from ESD1 DNA clearly demonstrated that this clone contained sequence corresponding to that found at the extreme 5' end of the published mouse Vipr2 cDNA sequence (GeneBank accession number D28132; Inagaki *et al.*, 1994), and that found at the 5' end of the mouse Vipr2 5' RACE product Q12.

In addition to verifying that clone ESD1 contained the 5' end of the coding region of the mouse *Vipr2* gene, the initial sequence data obtained from ESD1 also provided: i) the sequence of the exon 1/intron 1 junction (primer 8483) -the only known intron/exon junction sequence that remained to be determined within the *Vipr2* gene, and ii) the first sequence data (~300 bp) from the putative promoter region of the mouse *Vipr2* gene (primer 8484).

Subsequent sequencing reactions with the λ vector primers 7363 and 8482, showed that the 3' end of clone ESD1 was situated a short distance (~36 bp) into intron 2 of the mouse *Vipr2* gene, whereas the 5' end of the clone was within a region of unrecognised DNA sequence. Nevertheless, as later characterisation of the ESD1 clone showed that intron 1 of the mouse *Vipr2* gene spanned ~2.2 kb (see 3.3.19), it appeared that clone ESD1 must contain several kilobases of the putative promoter region of the mouse *Vipr2* gene. This clone therefore provided the starting point for studies aimed at the identification and subcloning of fragments from the putative promoter region, the results of which are described in Chapter 6.

3.3.19 PCR amplification of intron 1 of the mouse *Vipr2* gene

As mentioned in the previous section, the isolation of clone ESD1, and the subsequent demonstration (from sequencing data) that this clone contained intron 1 of the mouse *Vipr2* gene, finally provided a λ DNA template from which this intron could be amplified by PCR. Three combinations of PCR primer pairs and 3 slightly different DNA templates were used in attempts to isolate this intron: i) the exon 1 primer 8483 and the exon 2 primer 7122 were used with the ESD1 *Xba* I subclone D1X2* as template DNA; ii) the exon 1 primer 8483 and the vector primer 7363 (from the 3' end of ESD1) were used with ESD1 template DNA; and iii) the exon 1 primer 8483 and the plasmid vector primer T7 were used with the ESD1 *Xba* I subclone D1X8* as template DNA. Reactions i and ii resulted in the generation of strong single bands of PCR product that were ~2.2kb and ~2.3 kb respectively (Figure 3.26). In reaction iii, a predominant PCR product of ~1.6 kb was generated, but the presence of at least two other weaker PCR products in this reaction, hinted at the possibility of mispriming, and this was subsequently confirmed by sequencing reactions which demonstrated that primer 8483 had not primed correctly in PCR reaction iii. Nevertheless, sequencing of the PCR products from reactions i and ii indicated that in each of these cases, the PCR primers had primed correctly, and as the sequence obtained from the ends of these two

Figure 3.26

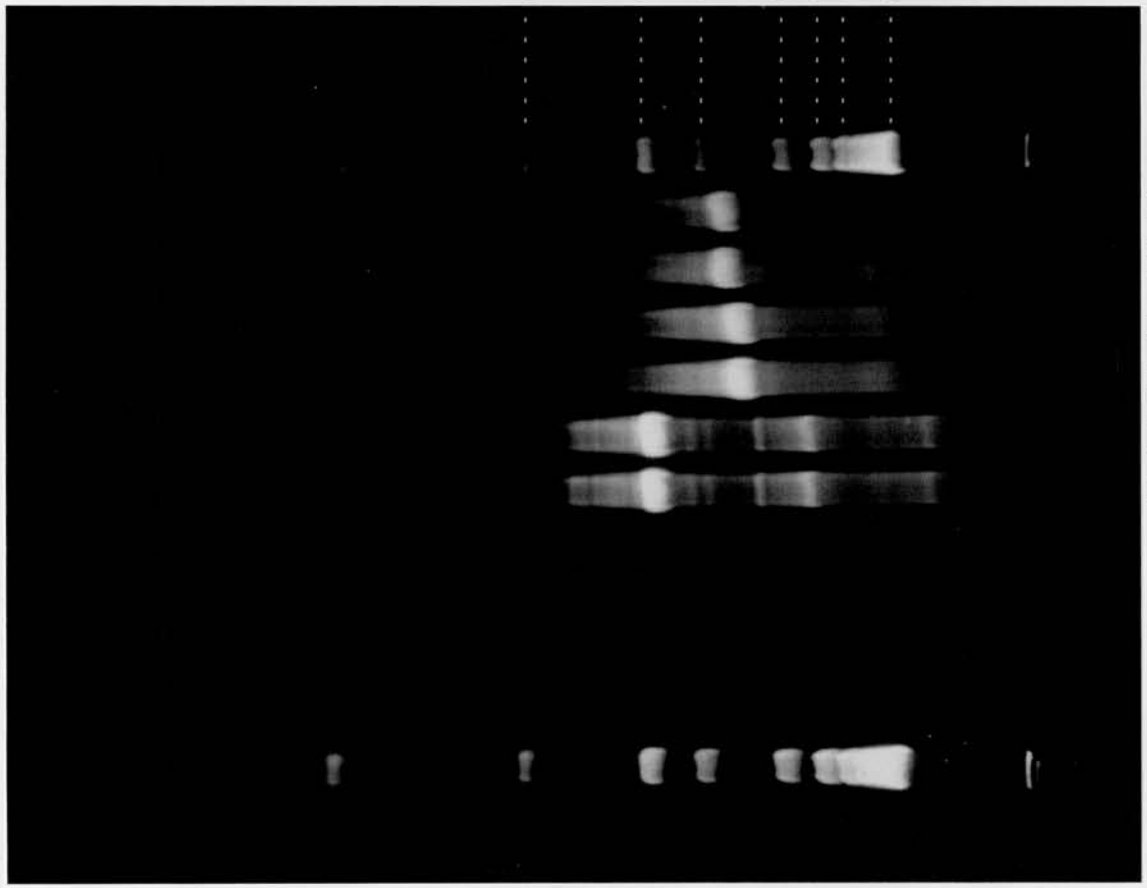
PCR products obtained from the amplification of intron 1 of the mouse *Vipr2* gene

Three different sets of primer pairs were tested for their ability to amplify intron 1 of the *Vipr2* gene, and produced the following results:

i) in the reactions where the exon 1 primer 8483 and the exon 2 primer 7122, were used with *Xba* I subclone D1X2 as template DNA, a product of ~2.2 kb can be seen.

ii) in the reactions where the exon 1 primer 8483 and the vector primer 7363 (from the 3' end of ESD1), were used with ESD1 template DNA, a product of ~2.3 kb can be seen

iii) in the reactions where the exon 1 primer 8483 and the plasmid vector primer T7, were used with the ESD1 *Xba* I subclone D1X8 as template DNA, a predominant product of 1.6 kb can be seen but some fairly strong secondary bands are also present suggesting that this reaction might be less specific.


12.2 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
1 kb ladder
D1X2 8483/7122 (int.1)
D1X2 8483/7122 (int.1)
ESD1 8483/7363
ESD1 8483/7363
D1X8 8483/T7
D1X8 8483/T7
1 kb ladder

PCR products included both the 5' and 3' boundaries of intron 1, this demonstrated that intron 1 of the mouse *Vipr2* gene spans ~2.2 kb.

*Further details of these subclones (which both contain intron 1 of the mouse *Vipr2* gene, are given in Chapter 6.

3.3.20 Subcloning and sequencing of DNA fragments spanning introns 2, 5, and 6 of the mouse *Vipr2* gene

The sizes of introns 2, 5, and 6 of the mouse *Vipr2* gene were eventually determined through the subcloning and sequencing of restriction fragments from the λ clones that were known to span these introns (ES9 for intron 2, and ES11 for introns 5 and 6). The restriction enzymes that were selected for use in the 'shotgun' subcloning (see 3.2.15) of restriction fragments from these introns were enzymes that did not cut within the λ vector arms other than at the polylinker site. *Xba* I was used in the case of intron 2, as it was known to release the insert fragment from clones and was also predicted to cut at a site within exon 2 of the *Vipr2* gene (thus providing a reference point for determining the positions of the other *Xba* I restriction fragments), whereas for introns 5 and 6 a combination of single and double digests was carried out using *Xho* I (which releases the insert fragment) and *Sca* I (which was predicted to cut in exon 7 of the mouse *Vipr2* gene).

Following digestion of ES9 DNA using *Xba* I, 8 different restriction fragments could be seen when the products were run on an agarose gel (Figure 3.27), and a series of plasmid subclones containing these restriction fragments was isolated using the method described in 3.2.15. Sequencing of the ES9 subclones using vector primers (T7 and SP6), allowed the identification of the 2 subclones that were derived from the 5' and 3' ends of ES9, and the order and positions of the intervening fragments were then determined by sequencing the junction regions between the subclones (using ES9 DNA as a template, and primers that were designed to sequence outwards from the appropriate end of subclones). The final order of the subclones, their sizes, and the positions of the primers that were used, are shown in Figure 3.28. All but one of the subclones contained at least some intron 2 sequence, and once a small gap in the middle of the subclone contig had been closed by direct sequencing of ES9 DNA (see Figure 3.28), the size of intron 2 of the mouse *Vipr2* gene was estimated to be ~12.6 +/-1 kb.

The first step towards determining the sizes of introns 5 and 6 was the subcloning of two *Xho* I restriction fragments (of ~4 kb and ~5.5 kb) from clone ES11 (which begins towards 3' end intron 4). The presence of the 4 kb *Xho* I fragment, but not the 5.5 kb fragment, in *Xho* I digests of ES3 (which begins in intron 5) had indicated that

Figure 3.27

Gel electrophoresis of ES9 DNA following restriction digestion

In lane 3 (ES9 *Xba* I) the 8 *Xba* I restriction fragments (that were later subcloned into pGEM®-7Zf) can be seen to range from ~600 bp (just visible, appears as two dots) to ~4 kb

1 kb ladder
ES9 *Xho* I
ES9 *Xba* I
ES9 *Xho* I/*Xba* I
ES9 *Pst* I
ES9 *Xho* I/*Pst* I
ES9 *Xho* I/*Ava* I
ES9 Ava I
ES9 Bgl II
ES9 Xho I/Bgl II
ES9 Not I
1 kb ladder
12.2 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
~517 bp

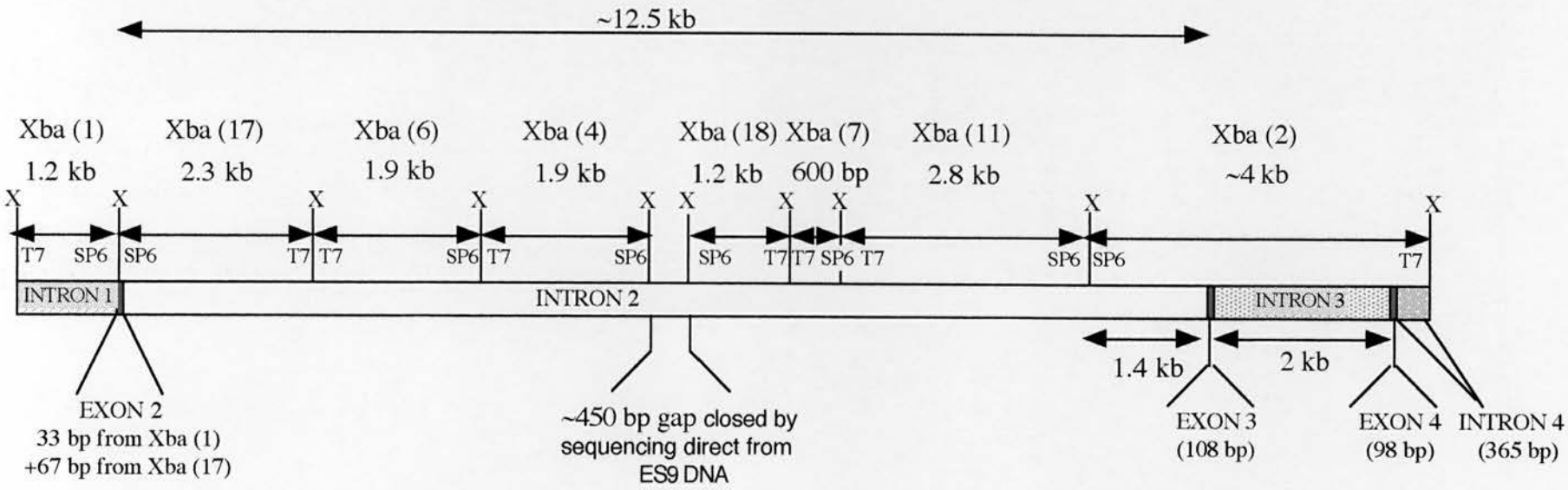
Figure 3.28

Schematic diagram showing the relative positions of the *Xba* I restriction fragments that were isolated (and subcloned) from the intron 2 region of the mouse *Vipr2* gene.

X= *Xba* I restriction site within the *Vipr2* gene (with the exception of the extreme left and right *Xba* I sites, which are thought to originate from the λ2001 polylinker sequence). Double headed arrows spanning the regions between adjacent *Xba* I restriction sites, represent subcloned *Xba* I restriction fragments the sizes of which are given directly above each arrow. Exons are represented by pink/red vertical bars.

(1.3 cm = ~1 kb)

MAP OF ES9 Xba I RESTRICTION FRAGMENTS FROM THE INTRON 2 REGION



the 5.5 kb fragment originated from the region of ES11 that lay 5' to the start of ES3, and this was confirmed by the sequence data obtained from the clones, which showed that the 5.5 kb *Xho* I subclone (ES11 Xh5) could be successfully sequenced with the exon 5 primer 5991 but not with exon 6 primers 9629 or 9630, whereas the 4 kb subclone (ES11 Xh4) could be sequenced with the exon 6 primers but not the exon 5 primers. To determine whether the two *Xho* I fragments were situated directly adjacent to one another within ES11, the region corresponding to presumed junction between the two subclones was sequenced (using primer 19458 and ES11 template DNA). The results obtained clearly indicated that the two fragments were directly adjacent, and therefore together spanned intron 5 of the mouse *Vipr2* gene. However, as ES11 Xh4 contained regions of both intron 5 and intron 6, it was not known what proportion of the ES11 Xh4 insert belonged to intron 5. Therefore, PCR reactions were carried out in which the regions of the insert between the plasmid vector primers (pUC/M13 forward and reverse primers) and the exon 5 primers 9629 (which extends towards intron 5) or 9630 (which extends towards intron 6) were amplified. A single band of PCR product of ~ 3 kb was produced using primer 9629 in conjunction with the pUC/M13 reverse primer, and a single product of ~1.1 kb was generated using primer 9630 in conjunction with the pUC/M13 forward primer (Figure 3.29). This demonstrated that ~3 kb of ES11 Xh4 was composed of intron 4 sequence, and ~1 kb of ES11 Xh4 was composed of intron 5 sequence. The proportion of intron 5 sequence in the subclone ES11 Xh5, and the distance from the 3' end of ES11 Xh4 to the start of exon 7, were subsequently determined through the mapping of the positions of 3 subcloned *Sca* I restriction fragments and 3 subcloned *Sca* I/*Xho* I restriction fragments, from ES11. The results of these mapping studies are summarised in Figure 3.30, from which it can be seen that intron 5 of the mouse *Vipr2* gene spans ~ 7.3 kb and intron 6 spans ~5.8 kb.

3.3.21 Retrospective PCR analysis of introns 5 and 6 of the mouse *Vipr2* gene

The estimated sizes of introns 5 and 6 (~7.3 kb and ~5.8 kb respectively; see 3.3.20) suggested that these introns were close to the upper size limit of fragments that could be amplified by conventional PCR. Nevertheless, products of approximately these sizes could be seen amongst the multiple PCR products that had been obtained from the intron 5 and 6 amplification reactions in which the primer/template annealing temperature had been lowered to 50°C (see section 3.3.16 and Figure 3.25).

Attempts to increase the specificity of these reactions, by reducing the number of PCR cycles from 35 to 30, had a particularly significant effect on the intron 5 amplification

Figure 3.29

PCR products obtained from the amplification of the regions of subclone ES11 Xh4 that are situated between:

i) the vector pUC/M13 reverse primer (RP) site and that for the exon 6 upstream-facing primer 9629,
ii) the pUC/M13 forward primer (FP) site and that for the exon 6 downstream-facing primer 9630.

The product seen in the two RP/9629 reactions clearly shows that the 9629 primer site is ~3 kb from one side of the plasmid vector, whereas the two FP/9630 reactions show that the 9630 primer site is ~1.1 kb from the other side of the plasmid vector.

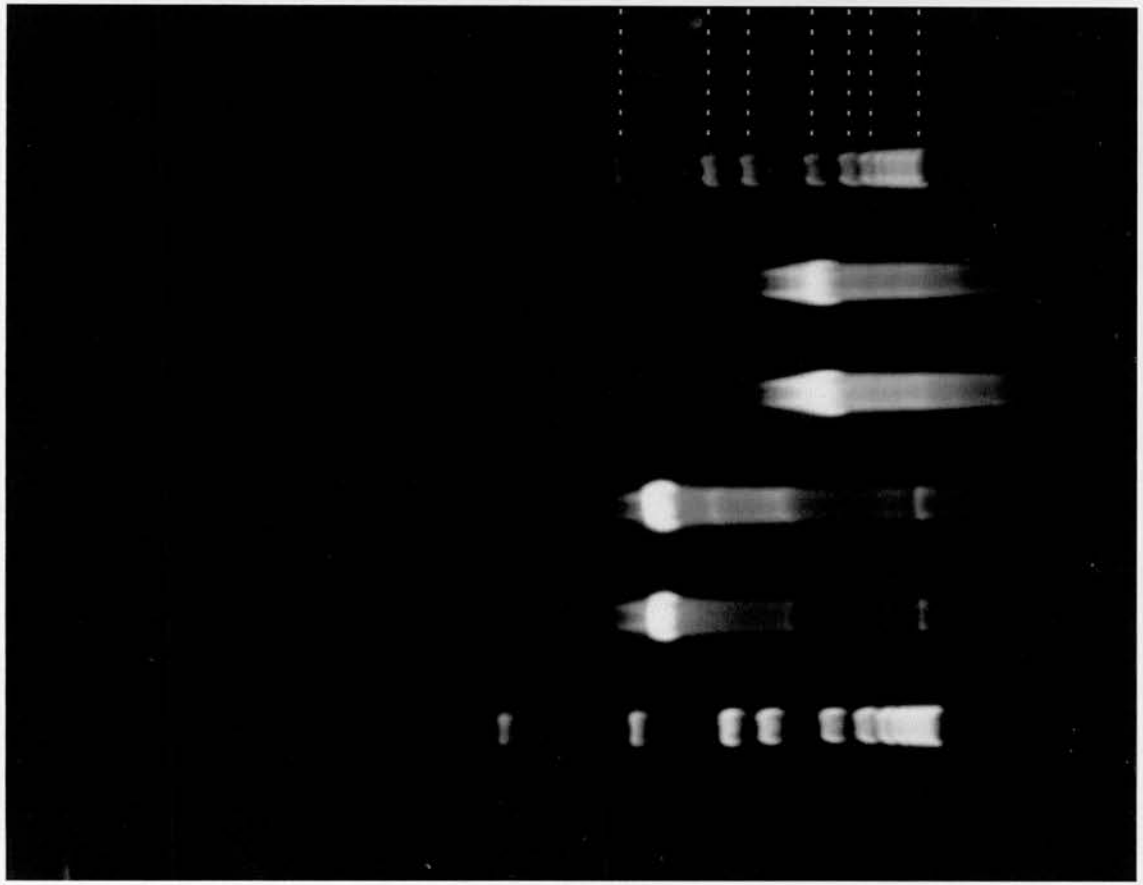

12.2 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
1 kb ladder
ES11 Xho (2) RP/9629
ES11 Xho (2) RP/9629
ES11 Xho (2) FP/9630
ES11 Xho (2) FP/9630
1 kb ladder

Figure 3.30

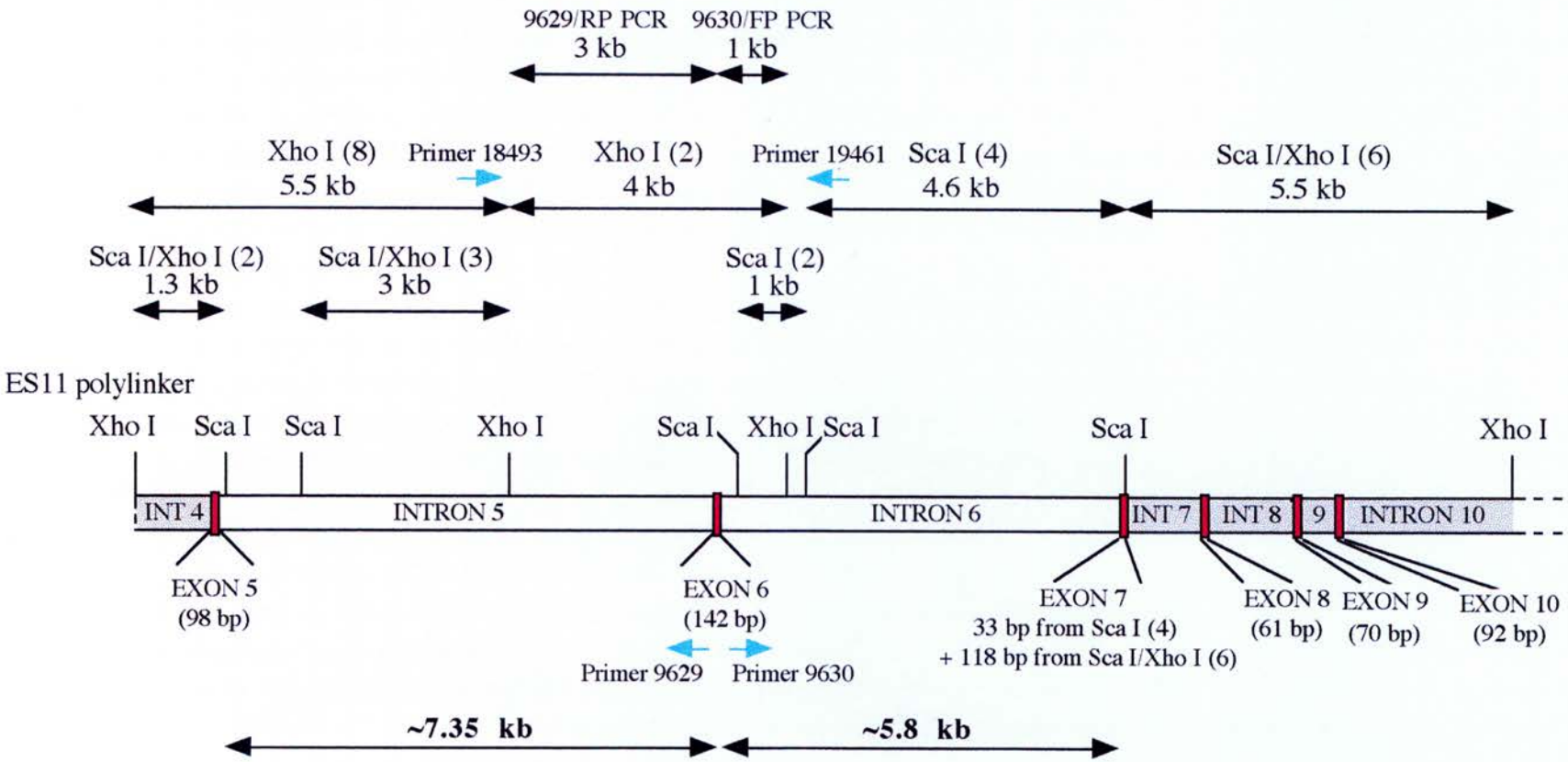
Schematic diagram of the intron 5/intron 6 region of the mouse *Vipr2* gene

The regions spanned by the ES11 subclones (derived from *Xho* I, *Sca* I, or *Xho* I/*Sca* I restriction fragments) and related PCR products, are represented by the double headed arrows towards the top of the diagram and both the size of the restriction fragment is and the subclone number (where applicable) is given directly above each of these arrows. The *Xho* I restriction sites shown at the extreme left and right of this diagram, are derived from the λ2001 polylinker sequence, but all other restriction sites shown are thought to represent sites within the *Vipr2* gene.

Blue arrows indicate the approximate positions of the binding sites of PCR/sequencing primers that were used in the characterisation of the subclones from this region of the gene (see Table 3.1 for primer sequences). Exons are shown in pink/red, and all introns except for introns 5 and 6 are shaded.

(1 cm = ~ 1 kb)

MAP OF SUBCLONES FROM THE INTRON 5 / INTRON 6 REGION



reactions, and when the reaction products were separated on an agarose gel, a single band of ~7 kb and a low molecular weight band of 'primer dimers' were seen (Figure 3.31). Unfortunately, the effect that this modification had on the intron 6 amplification reactions was less drastic (Figures 3.31 and 3.32), and although a distinct band of ~5.5-6 kb could be seen on an agarose gel, this was not the only PCR product produced in the intron 6 amplification reactions.

Following Microcon™-100 purification of the 7 kb intron 5 PCR product (see 3.2.13), and gel extraction of the ~5.5-6 kb intron 6 PCR product (using a QIAEX gel extraction kit; see Chapter 2), the isolated PCR products were sequenced using primers that were designed to prime to internal regions of the appropriate intron and generate sequencing products that extended outwards to the ends of the introns. This worked well with the intron 5 PCR product, which, when sequenced with primers 20847 and 20848, generated sequence data that corresponded to known intron 5 sequence and terminated at the exon target sites of the primers that had been used in the intron 5 PCR amplification reactions. However, the results of the sequence analysis of the intron 6 PCR product, were less clear, as when this PCR product was sequenced with the intron 6 primers 22568 and 22569, a dramatic decrease in the strength of the sequencing signal, as opposed to termination of the signal, was seen at the site of the original exon 7 primer, and a more gradual decrease in the signal strength was observed around the site of the original exon 6 primer (Figures 3.33 and 3.34), suggesting that non-specific secondary products from the intron 6 PCR reaction had been present in the gel slice.

3.3.22 Second rescreen of the λ2001 mouse genomic library, for clones that would close the intron 4 'gap'

Since intron 4 was now the only *Vipr2* gene intron for which a size had not been determined, a final effort was made to isolate this intron .

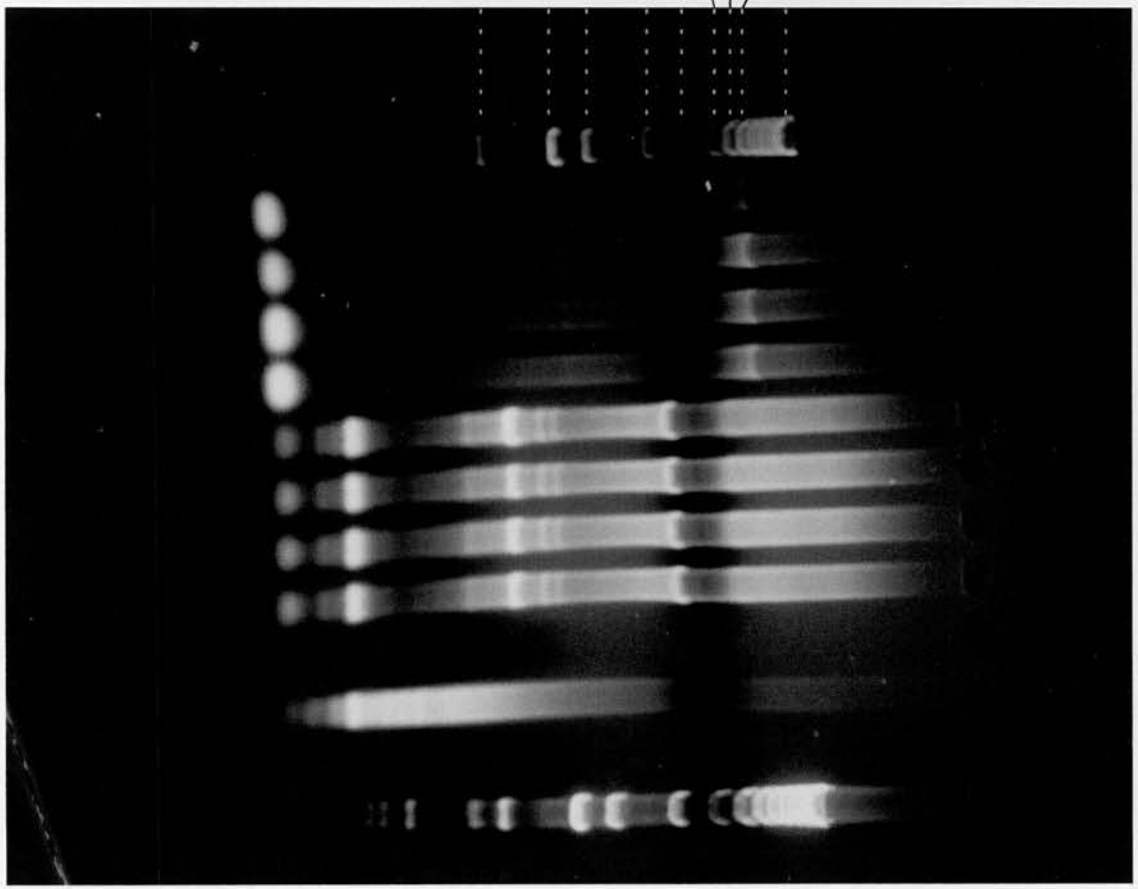
During the sequencing of *Xba* I restriction fragments from the intron 2 region of clone ES9 (3.3.20), it had been noticed that the sequence obtained from the 5' end of the ES9 *Xba* I subclone 2 (see Figure 3.28), aligned (from ~bp 316) with sequence data that had been obtained from reactions in which clone ES7 (also known as B7; see 3.3.12) was sequenced using the λ vector primer 7363. This placed the 5' end of clone ES7 ~ 1.1 kb from the intron 2/exon 3 junction, and since the total length of exon 3, intron 3 and exon 4 was ~2.2 kb, it appeared that all but ~3.3 kb of the ≥ 14 kb ES7 insert must be from intron 4 of the mouse *Vipr2* gene. Further sequencing of clone ES7, confirmed that this clone contained exons 3 and 4 of the mouse *Vipr2* gene, but the

Figure 3.31

PCR products obtained from reactions in which primers designed to amplify intron 5 (5991 and 9629) and intron 6 (9630 and 9200) of the mouse *Vipr2* gene were used on ES11 template DNA.

In each of the four intron 5 amplification reactions an isolated DNA band of approximately 7 kb (which is very similar to the estimated size of intron 5 that was predicted from restriction mapping studies ; see Figure 3.30) is the only clearly visible product apart from the 'primer dimers' seen towards the bottom of the gel. In the intron 6 amplification reactions multiple products can be seen, but the largest of these is close in size (~5-6 kb) to that estimated from restriction mapping studies (5.8 kb; see Figure 3.30) and the reaction was therefore re-run on another gel to obtain a more accurate estimate of this product's size.

See Figure 3.32.



1 kb ladder
ES11 5991/9629 (int.5)
ES11 5991/9629 (int.5)
ES11 5991/9629 (int.5)
ES11 5991/9629 (int.5)
ES11 9630/9200 (int.6)
ES11 9630/9200 (int.6)
ES11 9630/9200 (int.6)
ES11 9630/9200 (int.6)

1 kb ladder

Figure 3.32

PCR products obtained from a PCR reaction in which primers (9630 and 9200) designed to amplify intron 6 of the mouse *Vipr2* gene were used on ES11 template DNA.

Six identical reactions were set up, and although in each case multiple PCR products were produced, a DNA band of just under 6 kb (a size very close to that expected from restriction mapping studies of the intron 6 region ; see Figure 3.30) is clearly visible in each lane.

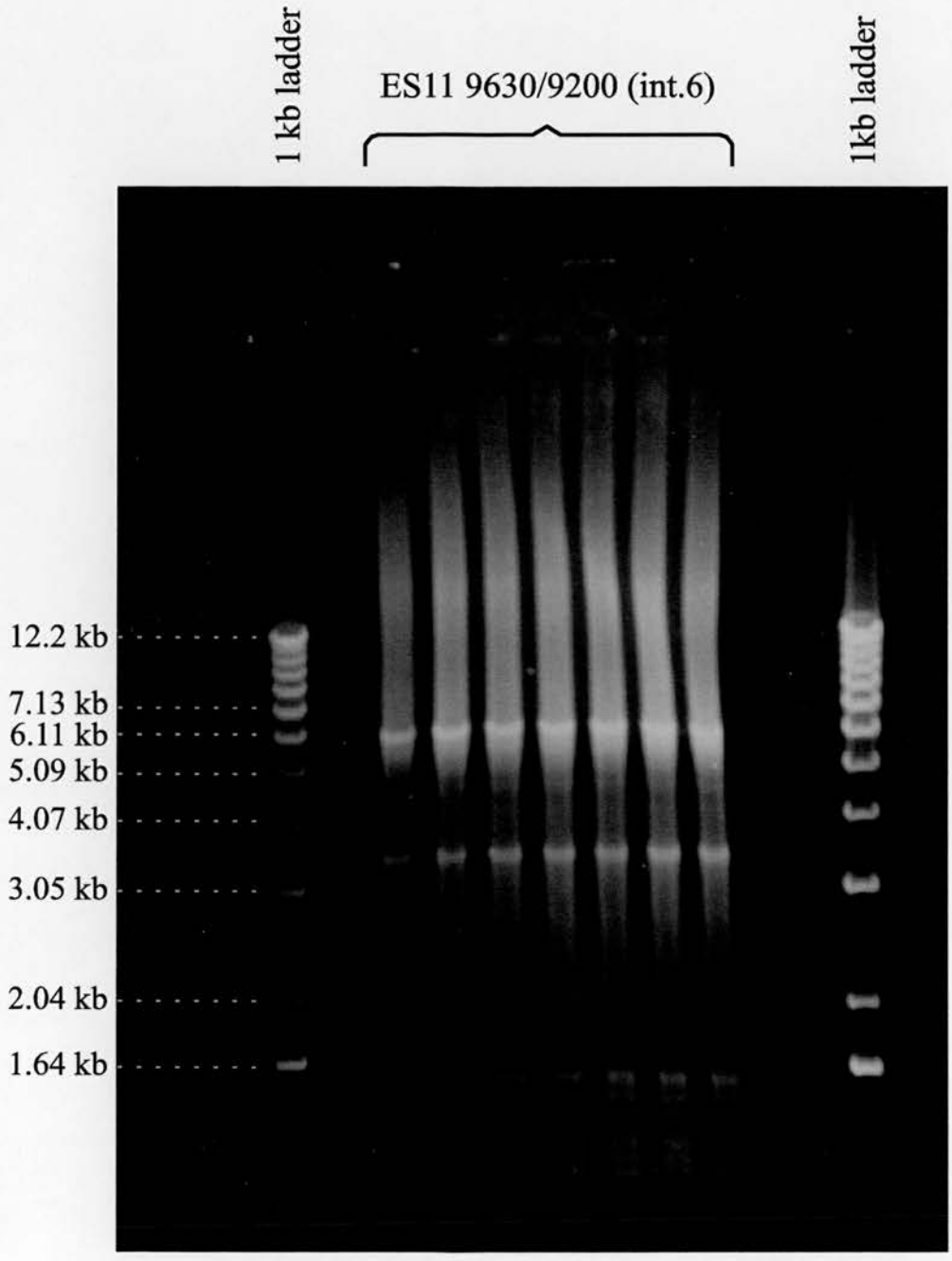

1 kb ladder
ES11 9630/9200 (int.6)
1kb ladder
12.2 kb
7.13 kb
6.11 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb

Figure 3.33

Sequence data obtained from the intron 6/exon 7 boundary of the mouse *Vipr2* intron 6 PCR product

The binding site of the *Vipr2* exon 7 PCR primer, i.e. the expected end of the intron 6 PCR product's sequence, is underlined in red, and a sharp drop in the strength of the sequencing signal can be seen at this point.

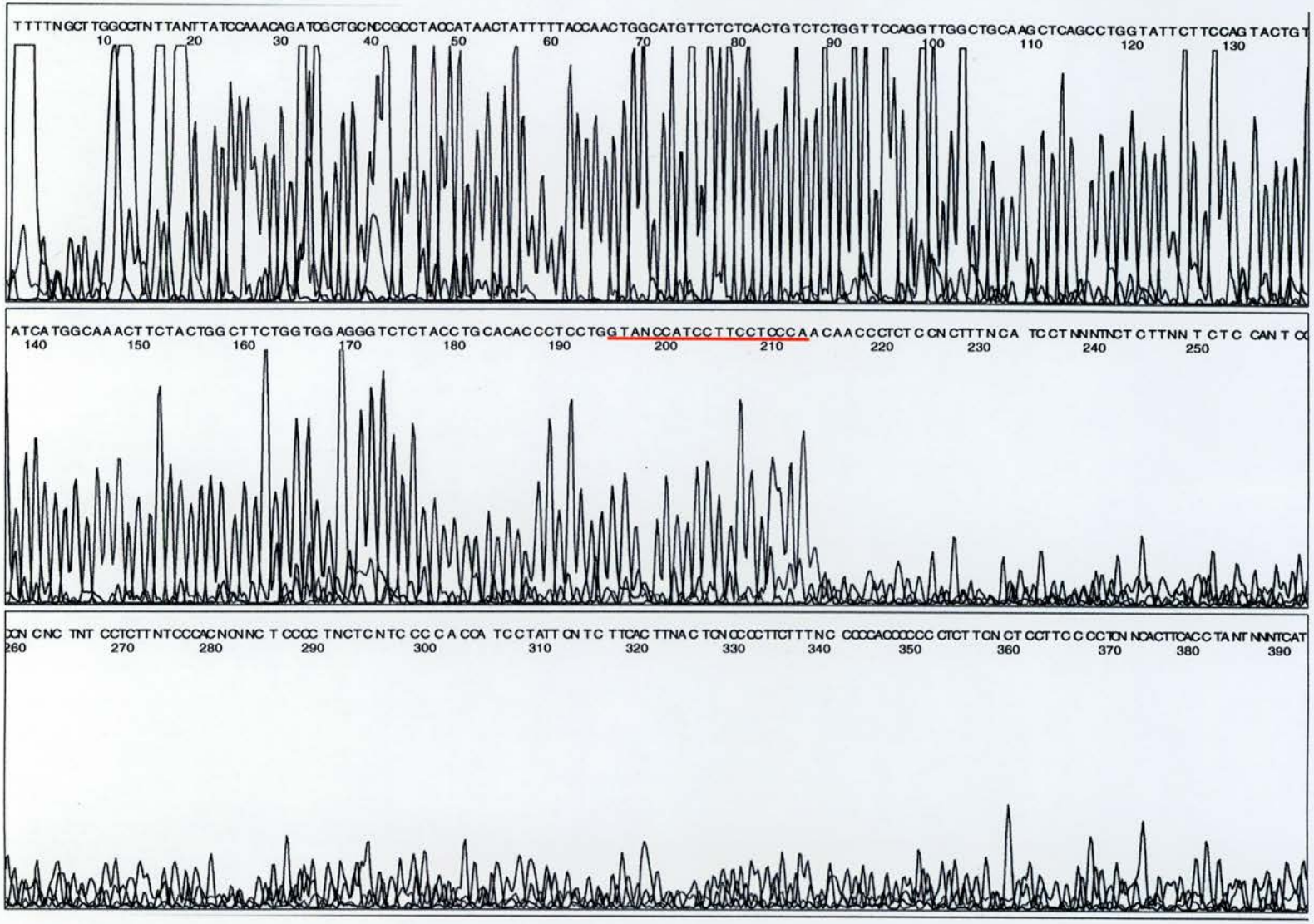
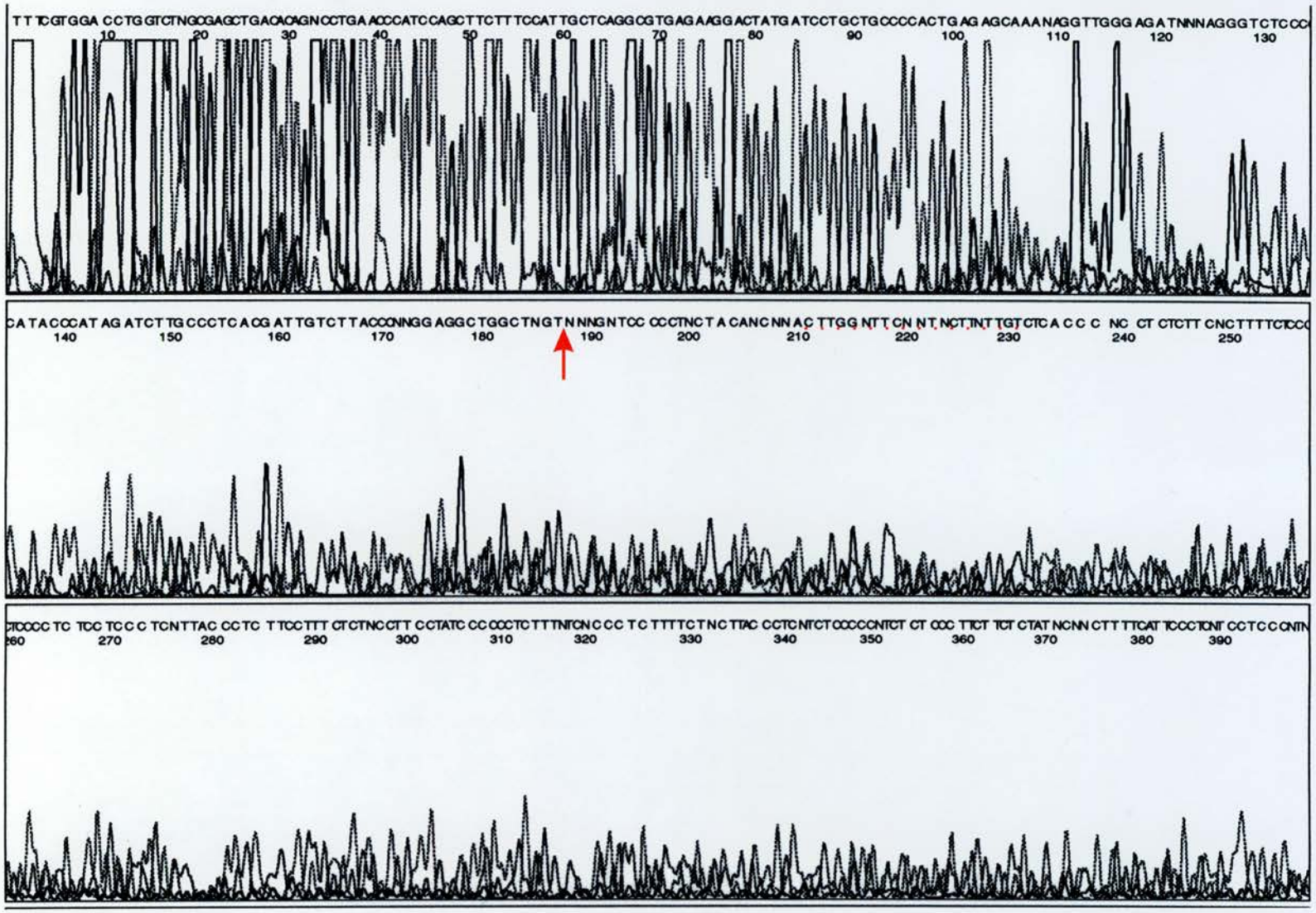

TTTTNGCT TGGCCTNTTANTTATCCAAACAGATCGCTGCNCCGCCTACCATAACTATTTTTACCAACTGGCATGTTCTCTCACTGTCTCTGGTTCCAGGTTGGCTGCAAGCTCAGCCTGGTATTCTTCCAGTACTGT
10 20 30 40 50 60 70 80 90 100 110 120 130
ATCATGGCAAACTTCTACTGGCTTCTGGTGGAGGGTCTCTACCTGCACACCCTCCTGGTANCCATCCTTCCTCCCAACAACCCTCTCCNCTTTNCA TCCTNNNTNCTCTTNN T CTC CAN T CC
140 150 160 170 180 190 200 210 220 230 240 250
CN CNC TNT CCTCTTNTCCCACNCNNC T CCCC TNCTCNTC CC C A CCA TCCTATT CN TC TTCAC TTNA C TCN CC CCTTCTTTNC CCCCACCCCCC CTCT TCN CT CCTTC C CCTCN NCACTTCACC TA NT NNNTCAT
260 270 280 290 300 310 320 330 340 350 360 370 380 390

Figure 3.34

Sequence data obtained from the intron 6/exon 6 boundary of the mouse *Vipr2* intron 6 PCR product

The position of the binding site of the *Vipr2* exon 6 PCR primer, i.e. the expected end of the intron 6 PCR product's sequence is indicated by a red arrow. In contrast to the dramatic decrease in signal at the exon 7 primer site (see Figure 3.33), a gradual decrease in the strength of the sequencing signal can be seen around the binding site of the exon 6 primer.


TTTCGTGGA CCTGGTCTNGCGAGCTGACACAGNCCTGA ACCCATCCAGCTTCTTTCCATTGCTCAGGCGTGAGAAGGACTATGATCCTGCTGCCCCACTGAGAGCAAANAGGTTGGGAGATNNNAGGGTCTCCC
10 20 30 40 50 60 70 80 90 100 110 120 130
CATACCCATAGATCTTGCCCTCACGATTGTCTTACCONNGGAGGCTGGCTNGTNNNGNTCC CCCTNCTACANCNNACTTGGNTTCNNTNCTTNTTGTCTCACCC NC CT CTCTT CNCTTTTCTCCC
140 150 160 170 180 190 200 210 220 230 240 250
CTCCCC TC TCC TCC C TCNTTAC CC TC TTCCTTT CTCTNCCTT CCTATC CC CCCTCTTTNTCNCCC TC TTTTCTNCTTAC CCTCNTCTCCCCCNTCT CT CCC TTCT TCT CTAT NCNN CTT TTCAT TCCCTCNT CCTCCCNTN
260 270 280 290 300 310 320 330 340 350 360 370 380 390

failure of attempts to sequence this clone using the exon 5 primer 7854, indicated that ES7 did not contain the whole of intron 4. To establish whether any of the clones that had been isolated previously, overlapped with the 3' end of ES7, a primer (17982) that was designed to bind to a site near the 3' end of the ES7 insert, was tested in sequencing reactions containing template DNA from clones ES9, GC4, GD3, GD6, GD7, GD9, ES11, and ES7 (positive control). However, apart from ES7, none of the clones were successfully sequenced by this primer 17982, a result which suggested that ES7 extended further into intron 4 (from the exon 4/intron 4 junction) than clones ES9, GC4, GD3, GD6, or GD7, but did not extend far enough to overlap with the 5' ends of clones GD9 or ES11 (clones that contain the intron 4/exon 5 junction).

As the minimum size of intron 4 was at this stage estimated to be ~12 kb (at least 11 kb from ES7 + ~1 kb from ES11), two genomic *Vipr2* probes (as opposed to the rat cDNA probe used earlier) were used in the final attempt to close the intron 4 'gap'. Both probes were subcloned genomic restriction fragments from intron 4 of the mouse *Vipr2* gene, one was a 3.5 kb *Eco*R I/*Xba* I restriction fragment that contained the 3' end of clone ES7, and the other was a 1.3 kb *Sca* I/*Xho* I restriction fragment from the 5' end of ES11 (ES11 S/X 2; see Figure 3.30). Twenty pairs of plaque lift filters were screened in total, and in order to try to isolate a clone that spanned the gap between the two probes, one set of filters was incubated with each of the radiolabelled probes, in the hope of identifying a clone that hybridised to both probes. However, in the primary screen no plaques were seen to hybridise to both probes, although several plaques were found to hybridise to one or other of the probes. Six plaques were selected from the primary screen (3 plaques that hybridised to the ES7 probe, and 3 that hybridised to the ES11 probe), but following secondary screens only 2 clones (ES9B and ES10B) were found to produce duplicate positives with the ES7 derived probe, and 1 clone (ES8A) produced duplicate positives with the ES11 derived probe.

Sequencing of λ DNA from ES9B, ES10B, and ES8A, revealed that only ES10B could be successfully sequenced with primer 17982 (which was designed to bind to the 3' end of ES7). However, in agreement with the failure of this clone to hybridise to the ES11 derived probe from the 3' end of intron 4, ES10B could not be sequenced with the exon 5 primer 7854. Therefore, although it seemed likely that ES10B extended at least as far into intron 4 (from the 5' end of the intron) as ES7 and possibly further, it was unlikely that ES10B on its own could close the gap in the intron 4 contig. Unfortunately, sequencing of clone ES8A (the only clone isolated using the ES11 derived probe) using the λ vector primer 8482, showed that the 5' end of ES8A was

Figure 3.35

Schematic diagram of the intron 4 region of the mouse *Vipr2* gene

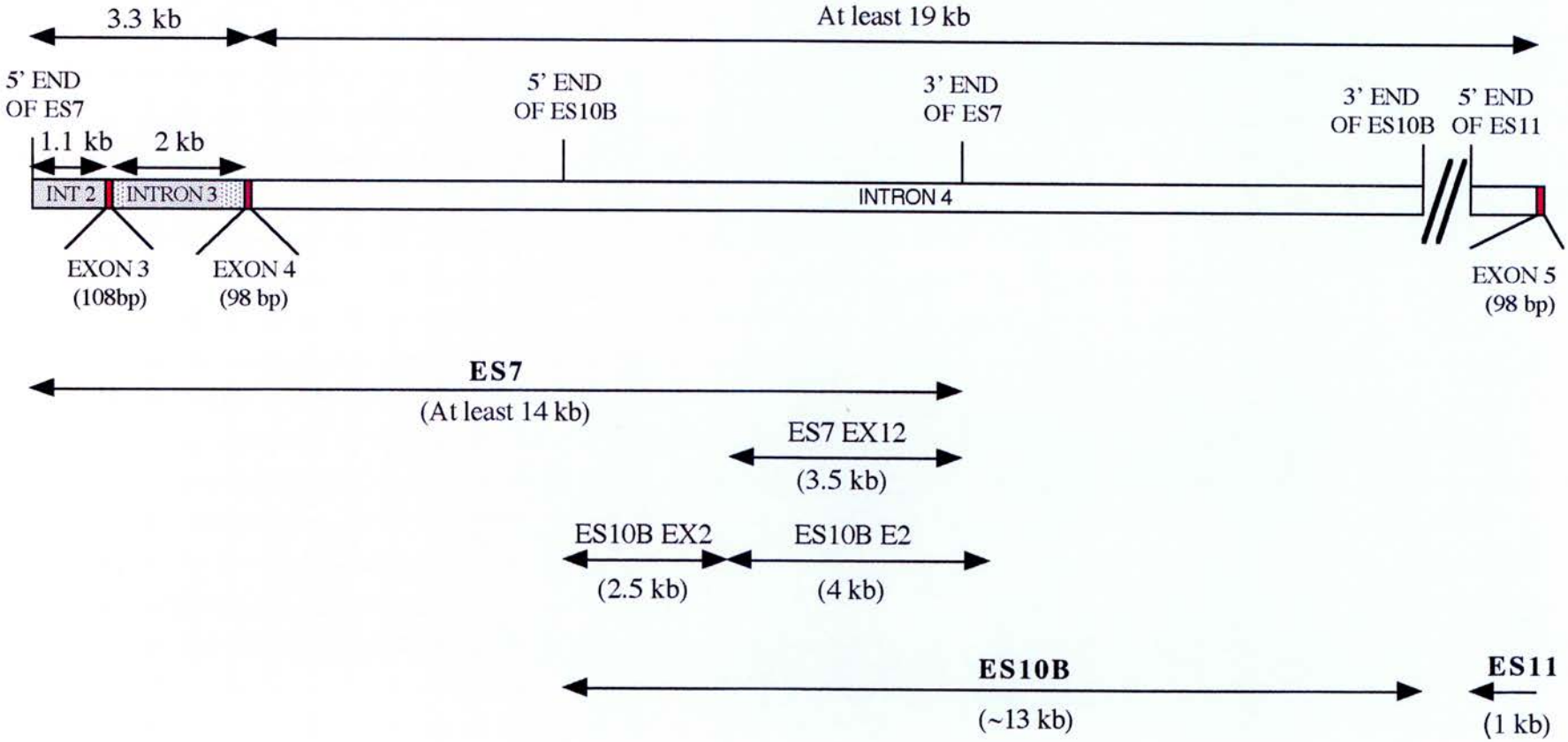
The regions spanned by: lambda clones ES7, ES10B, ES11, and the subclones ES7 EX12, ES10B EX2, and ES10B E2, are indicated by the double headed arrows under the names of these clones. Exons are shown in pink/red, and all introns other than intron 4 are shaded. In the contig of lambda clones spanning the *Vipr2* gene, a gap which is indicated by a double slash (//), exists between the 3' end of ES10B and the 5' end of ES11.

ES7 EX12: is a 3.5 kb fragment generated from an *Eco*R I/*Xho* I double digest of ES7 DNA
ES10B EX2: is a 2.5 kb fragment generated from an *Eco*R I/*Xho* I double digest of ES10B DNA
ES10B E2: is a 4 kb fragment generated from an *Eco*R I digest of ES10B DNA

Both ES7 and ES10B contain additional, unmapped, *Eco*R I restriction sites.

(1 cm = ~ 1 kb)

MAP OF INTRON 4 REGION



located only ~100 bp upstream of the start of ES11, and no overlap between the 5' end of ES8A and the 3' end of ES10B was found in the sequence data.

Nevertheless, to determine how far ES10B extended beyond the end of ES7, an *Eco*R I/*Xho* I and several *Eco*R I restriction fragments were subcloned from ES10B (using the 'shotgun' cloning method described in 3.2.15). Preliminary sequence analysis of these subclones (using T7 and SP6 primers) resulted in the identification of two subclones that contained previously determined intron 4 sequences: i) an ~2.5 kb *Eco*R I/*Xho* I restriction fragment (ES10B E/X) which began at the 5' end of the ES10B insert sequence, and ii) an ~4 kb *Eco*R I restriction fragment (ES10B E2) that began at the same point in intron 4 as the ~3.5 kb *Eco*R I/*Xho* I subclone from the 3' end of ES7, and was shown (by sequencing of the junction region using ES10B λ DNA as a template) to lie directly adjacent to the 3' end of subclone ES10B E/X (see Figure 3.35). This placed the 3' end of clone ES7 ~3.5 kb from the 5' end of the 4 kb subclone ES10B E2, and since it had been shown (through the sequencing of ES10B E/X) that the 5' end of ES10B lay ~2.5 kb upstream of ES10B E2, the overlap between the 5' end of the ~13 kb insert sequence from ES10B and the 3' end of the ES7 insert, was estimated to be ~6 kb. Therefore, as ES10B extended ~7 kb further into intron 4 than ES7, but did not overlap with ES11 sequence from the 3' end of the intron, it appears that intron 4 of the mouse *Vipr2* gene spans >19 kb.

3.4 Discussion

The results presented in this chapter indicate that the introns and exons within the protein encoding region of the mouse *Vipr2* gene are likely to span over 59 kb in total (see Figures 3.17, 3.18, and Table 3.4). The exons that were identified range in size from 42 bp to 171 bp (excluding the 5' and 3' UTRs), and the introns range in size from 66 bp to ≥19 kb. With the possible exception of the intron 7/exon 8 junction (where the sequence data obtained is suspected to contain an error), all of the exon/intron junctions conform to the consensus GT-AG rule for splice junctions (Breathnach *et al.*, 1978). Twelve introns and thirteen exons were found in the mouse *Vipr2* gene, and the positions of the intron sequences within the gene were very similar to those of introns in other members of the secretin/glucagon receptor gene family. Indeed, apart from intron 1, all of the *Vipr2* gene introns were found very close to the locations that had been predicted from the alignment of the mouse Vipr2 cDNA sequence with the exon sequences of other members of the receptor gene subfamily.

Table 3.4

All the sizes given are approximate with the exception of those for intron 9 and intron 11 (which were calculated from data produced during sequencing of the mouse *Vipr2* clones). The size of intron 4 is unknown but is estimated to be at over 19 kb (see main text for details). *As the position of the polyadenylation site of the mouse *Vipr2* transcripts had not been established at the time this work was carried out, and the total length of the cDNA was therefore unknown. The transcript size (~ 3.7 kb) used in the calculation of the minimum size of the mouse *Vipr2* gene was taken from the results of northern blots (probed for the presence of *Vipr2* transcripts), published by Inagaki *et al., 1994*..

Summary of mouse *Vipr2* gene intron sizes

Intron number	Intron size (kb)
1	~2.2
2	~12.5 (+/- 1)
3	~2.0
4	unknown (>19 kb?)
5	~7.35
6	~5.8
7	~1.1
8	~1.2
9	0.405
10	~3.5
11	0.066
12	~1.1

Total length of introns isolated so far = ~56.2 kb
Length of mRNA sequence* = ~3.7 kb

59.9 kb

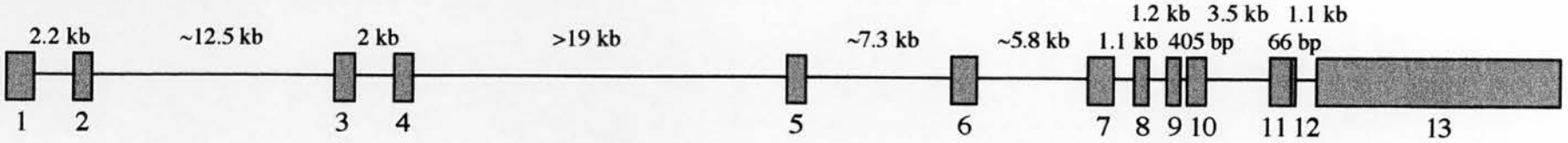
As further information about the gene structures of other members of the secretin/glucagon receptor family has accumulated (Aino *et al.*, 1995; Albrandt *et al.*, 1995; Burcelin *et al.*, 1995; Liaw *et al.*, 1996; Lok *et al.*, 1994; Sreedharan *et al.*, 1995; Tsai-Morris *et al.*, 1996; Yamada *et al.*, 1995b), it has become increasingly apparent that the positions of the exon/intron boundaries, particularly those in the regions encoding the transmembrane domains, are very highly conserved among members of this receptor gene subfamily. The positions of the introns that were identified in the mouse *Vipr2* gene indicate that: the predicted transmembrane region (TM) 1 of the receptor is encoded by exon 5; TM2 is encoded by exon 6; TM3 and first half of TM4 (TM4a) is encoded by exon 7; TM4b is encoded by exon 8; TM5a is encoded by exon 9; TM5b is encoded by exon 10; TM6 and TM7a are encoded by exon 11; TM7b is encoded by exon 12; and the predicted cytoplasmic tail and 3' UTR are encoded by exon 13 (see Figure 3.36). This basic intron/exon structure is shared with the majority of other receptor genes from this subfamily that have been characterised to date, including: the PTH/PTHrP receptor gene (Kong *et al.*, 1994), the human GIP receptor gene (Yamada *et al.*, 1995b), the rat CRF type I receptor gene (Tsai-Morris *et al.*, 1996), the human VIP type I receptor gene (*VIPR1*; Sreedharan *et al.*, 1995), the mouse PACAP type I receptor gene (*Pacapr1*; Aino *et al.*, 1995), the mouse GHRH receptor gene (Lin *et al.*, 1993), and the human and mouse glucagon receptor genes (Burcelin *et al.*, 1995; Lok *et al.*, 1994). However, two notable exceptions exist: i) the pig and human calcitonin receptor genes, in which a single exon spans the region encoding TM5b, TM6, and TM7a (Albrandt *et al.*, 1995; Zolnierowicz *et al.*, 1994), i.e. there is no intron at the position that corresponds to that of intron 10 of the mouse *Vipr2* gene; and ii) the rat glucagon receptor gene, in which a single exon spans the region encoding TM6 and TM7 (Maget *et al.*, 1994), i.e. there is no intron at the position that corresponds to that of intron 11 of the mouse *Vipr2* gene. Interestingly, the human gene encoding the secretin/glucagon receptor family related molecule CD97 (see Chapter 1), which is thought to have evolved from an ancestral secretin/glucagon receptor family gene, is similar to the rat glucagon receptor gene in that TM6 and TM7 are encoded by a single exon (Hamann *et al.*, 1996). However, as it has been proposed that the rat glucagon receptor may represent an intermediate form in the evolution of the more commonly observed two exon division of this region (TM6+TM7a, and TM7b) (Tsai-Morris *et al.*, 1996), it seems likely the CD97 gene diverged from the secretin/glucagon receptor gene family before the current structures of genes like those encoding the CRF and VIP receptors had evolved.

The regions of the secretin/glucagon receptor family genes that encode the predicted N-terminal extracellular domains of the receptors, display greater variability in their

Figure 3.36

Diagram showing the relationship between the exon structure of the mouse *Vipr2* gene and the protein structure of the VIP_2 receptor

The diagram shown above illustrates the intron/exon structure of the mouse *Vipr2* gene and its relationship to the predicted protein structure of the VIP_2 receptor. In this diagram the exons of the gene are represented by shaded rectangles (and are numbered 1-13). Introns are represented by the black vertical lines between adjacent exons, and are drawn to a scale of 3 cm = 10 kb. Exons are drawn to a scale of 3 cm = 1 kb. Amino acids within the VIP_2 receptor are shown as small circles and are labelled using standard single letter abbreviations are used for amino acids (IUPAC-IUB Commission on Biochemical Nomenclature):
A=Ala, C=Cys, D=Asp, E=Glu, F=Phe, G=Gly, H=His, I=Ile, K=Lys, L=Leu, M=Met, N=Asn, P=Pro, Q=Gln, R=Arg, S=Ser, T=Thr, V=Val, W=Trp, Y=Tyr. Alternate regions of light and dark grey shading have been used throughout the VIP_2 receptor's amino acid sequence, to distinguish between regions of the receptor that are encoded by different exons of the *Vipr2* gene. Therefore, all regions of the protein that are encoded by odd numbered exons (1, 3 ,5 etc.) are shown in light grey, and the regions encoded by even numbered exons (2, 4, 6 etc.) are shown in dark grey. Sites corresponding to positions where splice variation has been detected in some of the other secretin/glucagon receptor family members, are indicated by arrows.


2.2 kb
~12.5 kb
2 kb
>19 kb
~7.3 kb
~5.8 kb
1.2 kb
3.5 kb
1.1 kb
1.1 kb
405 bp
66 bp
1
2
3
4
5
6
7
8
9
10
11
12
13


EXTRACELLULAR
INTRACELLULAR

organisation than those encoding the transmembrane-spanning regions and C-terminal tails. As with the amino acid sequences of the N-terminal regions of these receptors, the exon/intron structures of the N-terminal encoding regions of the receptor genes are generally well conserved across species for a given receptor, but often vary significantly between different members of the receptor gene family. Nevertheless, despite the increased variability within this region, a common intron/exon pattern in which the N-terminal region is encoded by 4 exons, is found in several members of the receptor gene family including the genes encoding the GHRH receptor (Lin *et al.*, 1993), calcitonin receptor (Albrandt *et al.*, 1995; Zolnierowicz *et al.*, 1994), glucagon receptor (Burcelin *et al.*, 1995; Lok *et al.*, 1994; Maget *et al.*, 1994), CRF_1 receptor (Tsai-Morris *et al.*, 1996), and the VIP_1 receptor (Sreedharan *et al.*, 1995). The data obtained in this study shows that the mouse *Vipr2* gene also has 4 exons encoding the predicted N-terminal extracellular domain of the receptor (Figure 3.36), and therefore differs from: i) the mouse PACAP type I receptor gene (the protein product of which shares ~50% sequence homology with the mouse VIP_2 receptor; Inagaki *et al.*, 1994) and the PTH/PTHrP receptor gene, which both contain additional exons within this region (2 additional exons in the case of the PACAP type I receptor gene, and 1 additional exon in the case of the PTH/PTHrP receptor gene: Aino *et al.*, 1995; Kong *et al.*, 1994); ii) the human CRF type 2 receptor gene, in which 3 exons encode the predicted N-terminal extracellular domain of the receptor (Liaw *et al.*, 1996).

The PTH/PTHrP receptor gene, pig calcitonin receptor gene, and the GIP receptor gene, have each been shown to contain 5' untranslated exons (3, 2, and 1 respectively; McCuaig *et al.*, 1995; Yamada *et al.*, 1995b; Zolnierowicz *et al.*, 1994), and there is evidence to suggest that several other genes from the secretin/glucagon receptor subfamily also contain 5' untranslated exons. PCR-based studies of rat PACAP type I receptor cDNAs indicate that the *Pacapr1* gene is likely to contain at least 3 non-coding exons within its 5' UTR (Chatterjee *et al.*, 1996a), and it appears that the rat glucagon receptor also contains at least 2 untranslated exons in its 5' UTR (Maget *et al.*, 1996b). However, in many cases these 5' untranslated exons had not been detected when the structures of the genes were initially published, and therefore although no 5' untranslated exons have been identified within the *VIPR1* gene as yet, the possibility that such exons exist cannot be excluded. Similarly, despite the fact that the *Vipr2* gene exons identified during this study span the entire 5' untranslated and coding regions of the published mouse cDNA sequence, it is entirely possible, and perhaps even probable, that unidentified 5' untranslated exons exist within the mouse *Vipr2* gene, particularly if they are transcribed in a tissue specific fashion (as occurs in the PTH/PTHrP receptor gene (Amizuka *et al.*, 1997; McCuaig *et al.*, 1995; and probably

also in the rat glucagon receptor gene; Maget *et al.*, 1996a; Maget *et al.*, 1996b). Attempts to define the transcriptional start site(s) within the *Vipr2* gene have been made (see Chapter 6), and suggest that at least in the mouse AtT20 cell line, the 5' untranslated region of the *Vipr2* transcripts is encoded by the same exon as the ATG translational start site. To date, these studies have not been extended to other cell lines or tissues. Nevertheless, if candidate 5' untranslated exons are eventually identified from studies of Vipr2 cDNA sequences, there is a reasonable chance that any corresponding regions of mouse genomic DNA would be present within clone ESD1, which appears to have a total insert size of ≥ 20 kb (close to the upper size limit of ~23 kb for λ2001 inserts; Karn *et al.*, 1984) and is predicted to contain at least 17.5 kb of *Vipr2* gene 5' flanking sequence.

The genes encoding members of the secretin/glucagon receptor family, range considerably in size. At the lower end of the size range, are the rat and mouse glucagon receptor genes, which have coding regions that span ~4 kb, and were characterised almost entirely by subcloning and sequencing of λ clones that were isolated from genomic libraries (Burcelin *et al.*, 1995; Maget *et al.*, 1994). In contrast the pig calcitonin receptor gene, which is the largest member of the secretin/glucagon receptor gene family to be isolated to date, spans approximately 70 kb in total (including its two 5' untranslated exons), and was characterised using a combination of partial digestion restriction mapping, subcloning, and PCR amplification of λ clones (Zolnierowicz *et al.*, 1994). With a coding region that is estimated to span at least 59 kb of genomic DNA, the mouse *Vipr2* gene is one of the largest of the members of this receptor gene family that have been cloned to date, and at present appears to be second in size only to the pig calcitonin receptor gene [although if the human calcitonin receptor gene (Albrandt *et al.*, 1995) has a 5' UTR exon structure similar to that found in pig (Zolnierowicz *et al.*, 1994), it may also be larger than the mouse *Vipr2* gene]. Another large gene from this family is the mouse *Pacapr1* gene, which is reported to span over 50 kb (Aino *et al.*, 1995), a size which includes only one of the three 5' untranslated exons that have been shown to be present in the rat gene (Chatterjee *et al.*, 1996a). However, the human *VIPR1* gene, (the product of which, shows ~50% homology to the rodent VIP_2 receptor at the amino acid level; Sreedharan *et al.*, 1995) is much smaller than either the mouse *Vipr2* gene or the *Pacapr1* gene, and with a protein coding region that spans ~22 kb of genomic DNA (Sreedharan *et al.*, 1995), is similar in size to the coding region of the PTH/PTHrP receptor gene (Kong *et al.*, 1994).

The two largest introns identified within the mouse *Vipr2* gene, introns 2 and 4 (which are ~12.5 and ≥19kb respectively), are located in the region that encodes the predicted

N-terminal extracellular domain of the receptor (see Figure 3.36). Problems were encountered in determining the sizes of both of these introns, and although the approximate size of intron 2 was eventually estimated by the addition of restriction fragments sizes, this was not an ideal way to determine the size of this intron, and there could easily be an error of +/- 1 kb in the final value. The reason for the failure of attempts to amplify intron 2 using long PCR, is not fully understood, but it may be due at least in part, to the quality of the ES9 template DNA that was used, as although no problems were encountered with the use of this DNA in other procedures (including automated fluorescent sequencing, which is generally very sensitive to template quality), the success of the Expand™ long template PCR system (Boehringer Mannheim) is reported to be highly dependent on the purity of the template DNA that is used. Similar problems may have affected attempts to amplify intron 4 using this method, as although human genomic DNA specifically produced by Boehringer Mannheim for use with the Expand™ long template system, has been successfully used in our laboratory to amplify fragments from the human *VIPR2* gene, mouse genomic DNA is not available from this company. However, it is equally possible that the conditions used in attempts to amplify intron 4 were not optimal (as this becomes increasingly critical when amplifying longer fragments, particularly from total genomic DNA templates), or that intron 4 of the mouse *Vipr2* gene is larger than the maximum size (~27 kb) that can be efficiently amplified from genomic DNA by this system.

Although the isolation of mouse genomic clones that span intron 4 of the *Vipr2* gene, is not essential for the determination of this intron's size (as if necessary the intron size could be estimated by long range restriction mapping and blotting) it would greatly assist the process. Since it has now been shown that this intron is likely to span over 19 kb, the inability of the original cDNA probe to isolate clones that spanned intron 4, is no longer surprising. However, although the switch from a cDNA probe to the use of genomic probes (see 3.3.22) helped in the isolation of clones that extended further into intron 4 from its 5' end, the failure of any of the library screens to identify a clone that extended more than ~1 kb into intron 4 from the 3' end of this intron, seemed strange when the library had provided more than adequate coverage of other regions of the mouse *Vipr2* gene, and could be an indication that this region is somewhat resistant to cloning.

Interestingly, introns 2 and 4 are two of only three introns whose sizes have not yet been determined in the human *VIPR2* gene, (although intron 2 is known to span >5 kb; E. Lutz, unpublished data), but the problems that have been encountered in attempts to

amplify these introns by PCR from human genomic DNA, suggest that they may also be amongst the largest introns in the human gene.

As indicated by the large variation in the sizes of the genes that encode members of the secretin/glucagon receptor subfamily, the sizes of the introns present in these genes vary widely from one receptor gene to another, and there is no overall pattern in the spacing of the exons in genes that encode different members of the receptor subfamily. On the other hand, the size of a given receptor gene does appear to be conserved to some extent across species boundaries. For example, the coding region of the glucagon receptor gene spans ~4 kb in rat and mouse (Burcelin *et al.*, 1995; Maget *et al.*, 1994), and ~5.5 kb in human (Lok *et al.*, 1994), whereas the rat, mouse, and human PTH/PTHrP receptor genes each span ~22 kb (including one of the 5' untranslated exons; Kong *et al.*, 1994). Attempts were made to examine (from the published literature) whether the sizes (or relative sizes) of individual introns within a given receptor gene were generally conserved across species boundaries, but the low number of secretin/glucagon receptor subfamily genes that have been cloned from more than one species, and the small scale diagrammatic form in which the gene structures are often published, made it difficult to obtain any conclusive answers to this question. Nevertheless, from the sizes of the mouse *Vipr2* gene introns that were determined in this study, it appears that several of the mouse *Vipr2* gene's introns (1, 7, 8, 11, and 12) are similar or identical in size to those found in the human gene, despite distinct species differences in the sizes of introns 3, 5, 6, 9, and 10. The similarities in the sizes of some of these introns could of course be completely coincidental, but the pairwise distribution of the conserved (7&8, 11&12) and nonconserved (5&6, 9&10) intron sizes is curious, and although its significance, if any, is not entirely clear, it may in some way reflect the rearrangements that occurred during the evolution of the VIP receptor type 2 gene.

Splice variation has been found to occur in transcripts of several members of the secretin glucagon receptor gene family (see also Chapter 1). The positions within the receptor transcripts at which alternative splicing has been detected, vary between different members of the receptor subfamily, and include inserts in the regions encoding: the C-terminal section of the 3rd intracellular loop [as in the the PACAP type I receptor (Spengler *et al.*, 1993) and the GHRH receptor (Mayo, 1992)]; the first intracellular loop [as in the CRF type I receptor receptor (Chen *et al.*, 1993) and pig calcitonin receptor (Zolnierowicz *et al.*, 1994)]; and the first extracellular loop [as in the rodent calcitonin receptor (Sexton *et al.*, 1993)]. More recently there have also been reports of: variation (a 29 amino acid insertion) within the C-terminal tail of the human

GIP receptor (Gremlich *et al.*, 1995); two additional rat CRF type I receptor gene transcript variants in which the sequences corresponding to exon 3 or exons 7, 11, and 12 of the gene are deleted (Tsai-Morris *et al.*, 1996); splice variation within rat CRF type 2 receptor gene transcripts that results in the generation of receptor variants that have different N-terminal domains (Lovenberg *et al.*, 1995); human calcitonin receptor gene transcripts that lack both the 48 bp insert in the region encoding the 1st intracellular loop and the region encoding the first 47 amino acids of the receptor's N-terminal extracellular domain (Albrandt *et al.*, 1995); a mouse and human PACAP type I receptor variant, which (in contrast to the previously identified 3rd intracellular loop variants of this receptor), results in the deletion of a 21 amino acid region (encoded by exons 5 and 6) from the N-terminal extracellular domain of the receptor (Pantaloni *et al.*, 1996); the identification of four patterns of splice variation in transcripts of the rat PTH/PTHrP receptor gene, including one variant which affects the protein coding region and leads to the deletion of the exon (exon S) encoding the receptor's putative signal peptide (Joun *et al.*, 1997); and the isolation of two variants of human PTH/PTHrP receptor gene mRNAs, both of which lack the sequence corresponding to the first exon (E1) of the N-terminal-encoding region, although in one of these transcripts this is replaced by sequence from the intron directly upstream of exon E2 (Jobert *et al.*, 1996).

So far, the only evidence to suggest that splice variation might exist in transcripts of the *VIPR1* gene, comes from an isolated report, which describes the characterisation of two human VIPR1 cDNA clones that display complete divergence within the region encoding the 67 amino acid N-terminal domain (Couvineau *et al.*, 1994). However, although the human *VIPR1* gene has now been cloned (Sreedharan *et al.*, 1995), the point of amino acid sequence divergence between the predicted sequences encoded by the two cDNA variants does not correspond to an identified exon/intron junction within the *VIPR1* gene, and the mechanism underlying the generation of these two forms of *VIPR1* transcripts has not been established.

To date, no variants of the Vipr2 cDNA sequence have been published, but the existence of such variants cannot be discounted, especially in light of the recent cascade of publications relating to splice variation in the transcripts of other members of the secretin/glucagon receptor family. Analysis of genomic sequence data is not the usual way to begin to search for splice variation, as normally these variants can be detected by reverse-transcriptase (RT)-PCR of RNA from appropriate tissues. However, although RT-PCR experiments in our laboratory had failed to detect any transcripts encoding VIP_2 receptor variants, it was possible that the generally low levels of *Vipr2*

transcripts (Usdin *et al.*, 1994), perhaps combined with even lower levels of some splice variant transcripts [as has been found for several other splice variants in this receptor subfamily (Albrandt *et al.*, 1995; Jobert *et al.*, 1996; Joun *et al.*, 1997; Zolnierowicz *et al.*, 1994)], could make the detection of any splice variants very difficult.

Intron 10 of the mouse *Vipr2* gene was selected for sequence analysis primarily because: i) it corresponded (in terms of position within the receptor's gene structure) to the region in the rat *Pacapr1* gene that contains a "cassette" of two alternatively spliced 84 bp sequences (leading to the generation of 5 splice variants; Spengler *et al.*, 1993), and at that time was the only identified source of splice variation within the gene encoding the PACAP type I receptor (the receptor which together with the VIP_1 receptor is the most closely related receptor to the VIP_2 receptor at the amino acid level); ii) the rat PACAP type I receptor was the only member of this receptor subfamily for which splice variation had clearly been shown to result in alterations in receptor signal transduction , a factor that was of particular interest in light of the controversy over apparent discrepancies between the signalling properties of the cloned VIP/PACAP receptors and those of previously characterised endogenous VIP receptors within the brain (Gressens *et al.*, 1994; Hill *et al.*, 1994; Hill *et al.*, 1992; Oláh *et al.*, 1994).

Analysis of the sequence data obtained from intron 10 of the mouse *Vipr2* gene did not lead to the identification of any potential patterns of splice variation, and it seems unlikely that splice variation exists within this region of the *Vipr2* gene. Nevertheless, it should be noted that most of the intron 10 derived sequence data was obtained from the sequencing of only one strand of the receptor gene, and only one clone containing the PCR amplified intron 10 PCR product was sequenced (whereas ideally both strands would have been sequenced in each of several independent clones containing the PCR product). Therefore, the error rate within the intron 10 sequence data may be slightly higher than normal, and although improbable, it is conceivably that this might have hindered the recognition of splice site consensus sequences.

Other *Vipr2* gene introns that we would have liked to have sequenced include introns 5 and 6, which are situated in regions of the gene that encode the first intracellular loop and first extracellular loop respectively (see Figure 3.36), sites at which splice variation has been detected in other members of the receptor subfamily. However, these introns proved to be extremely difficult to amplify by PCR. In retrospect it is now clear that this was due at least in part to the length of the introns, which in the case of intron 5 (~7.3 kb), was approaching the size limit for amplification using standard PCR

conditions, but as this was not established until late in these studies, there was not enough time to sequence these introns.

Another difficulty that was puzzling at the time, but can now be explained, is the failure of the early attempts to sequence the λEMBL3 G4 *Vipr2* clone using the sequencing primer 992W and the Promega *fmol*® DNA cycle sequencing system (3.3.5), as following the determination of the *Vipr2* gene's exon/intron boundaries it became apparent that this primer [which would bind to bp 941-926 of the published mouse Vipr2 cDNA sequence; GeneBank accession number D28132; (Inagaki *et al.*, 1994)] crosses the exon 10 intron 9 junction (resulting in bp mismatches between the genomic target sequence and the primer's three most 3' nucleotides).

The isolation of clone ESD1, which contains the first protein coding exon of the mouse *Vipr2* gene, was particularly important with respect to the long term aim of generating *Vipr2*-null mice (because inclusion of the gene's ATG translation start site and first coding exon in a gene replacement construct should help to ensure that the truncated gene products, which could be partially functional, are not produced). In contrast to the other *Vipr2* clones that were isolated from the λ2001 library, a probe derived from a mouse Vipr2 5' RACE product had to be used in order to identify a clone (ESD1) containing the 5' end of the mouse *Vipr2* gene's coding sequence. As this probe was both mouse specific and contained an extra 34 bp of 5' UTR sequence (relative to the rat 260 bp Vipr2 cDNA probe), there were several possible explanations for its success, including the possibility that the rat and mouse sequences diverged towards the start of their 5' UTRs, a feature that has been noticed in the rat and human glucagon receptor genes (Lok *et al.*, 1994). However, since it has subsequently been shown that the rat Vipr2 cDNA contains 5' untranslated sequence that corresponds very closely to that of the mouse Vipr2 5' RACE product, the most likely explanation for the failure of the rat 260 bp 5' probe to detect clones containing exon 1, is that the length of homology between the probe and its target DNA, prior to intron 1 of the *Vipr2* gene, was not enough to allow efficient hybridisation.

During the characterisation of clones encoding the mouse VIP_2 receptor, a large amount of sequence data was obtained from the exons and introns of the *Vipr2* gene, and although most of this data consists of sequence derived from one strand of the receptor gene, each region of the contigs that were generated from this data was generally derived from the results of at least 3 separate sequencing reactions. These contigs, which include the sequence surrounding all the known exon/intron boundaries of the mouse *Vipr2* gene, and a substantial amount of intron sequence (including the complete sequences of introns 9, 10, and 11), are shown in Appendix 1.

In summary, the results presented in this chapter have demonstrated that the exon/intron organisation of the mouse *Vipr2* gene is very similar to that of other members of the secretin/glucagon receptor gene subfamily, and at a size of ≥59 kb this gene lies towards the upper end of the size range for those genes from this subfamily that have been characterised to date. Further work, including the use of pulse field gel electrophoresis and long range restriction mapping of genomic DNA will probably be necessary in order to establish the size of intron 4, which at present is estimated to span at least 19 kb. However, these studies, which also resulted in the isolation of a clone containing the first coding exon and a substantial amount of the gene's 5' flanking sequence (including the putative promoter region), represent the first characterisation (in any species) of the structure of the gene encoding the VIP_2 receptor, and have subsequently enabled transgenic approaches to the characterisation of the VIP_2 receptor's functions to be initiated in our laboratory.

Chapter 4

Isolation of a P1 clone encoding the human *VIPR2* gene

4.1 Introduction

Although much of this project focused on the isolation and characterisation of the mouse *Vipr2* gene, we were also interested in obtaining clones that encoded the human homologue of this gene, as it was hoped that this would allow: i) determination of the chromosomal localisation of the human gene and the investigation of any possible links with human genetic diseases; ii) comparison of the sequence and structure of the mouse and human genes; and iii) examination of the promoter and other regulatory sequences of the human gene.

Work on the isolation of the human *VIPR2* gene was initiated shortly after the isolation of the 4 kb subclone from the 3' region of the mouse *Vipr2* gene (see Chapter 3). At this stage the human VIPR2 cDNA sequence had not been published, and consequently rat Vipr2 cDNA sequences had to be used as probes in the isolation of the human gene. Nevertheless, those receptor cDNAs from this subfamily, that had been cloned from both rat and human, had generally been shown to display a high level of homology between the two species (Lok *et al.*, 1994; Mayo, 1992; Ogi *et al.*, 1993; Perrin *et al.*, 1993; Schipani *et al.*, 1993; Sreedharan *et al.*, 1993), and although there were only a few cases where the rodent and human gene structures from members of this subfamily had been published, there were indications that the intron/exon structure of these receptor genes might also be highly conserved between species (Kong *et al.*, 1994; Lok *et al.*, 1994; Maget *et al.*, 1994).

Previous studies of the 4 kb subclone from the 3' end of the mouse *Vipr2* gene had shown that this fragment of the gene contained only a small proportion of the *Vipr2* gene's total coding sequence (Chapter 3), and thus suggested that the total length of the mouse gene might be quite large. The reported sizes of other members of the receptor gene subfamily that had been characterised at this time, ranged from ~4 kb for the coding regions of the rat and mouse glucagon receptor genes (Burcelin *et al.*, 1995; Maget *et al.*, 1994) to ~70 kb for the porcine calcitonin receptor gene (Zolnierowicz *et al.*, 1994). This raised the possibility that the full length sequence of the human *VIPR2* gene might be too large to be found within any single clone from a λ-based genomic library, a factor which could cause difficulties if it was later necessary to reconstruct the

gene from overlapping clones in order to express it in cultured cells and/or transgenic mice . The possibility of screening a genomic library constructed in a high capacity vector was therefore examined, as the use of this type of library would significantly increase the chance of isolating the entire human *VIPR2* gene within one clone. This option also offered the added advantage that clones in high capacity vectors such as YACs and cosmids usually produce much stronger signals than λ clones when used as probes for the chromosomal localisation of genes by *in situ* hybridisation.

The main obstacle to the use of any of the high capacity vectors that were available, was that we did not have pulse field gel electrophoresis equipment (which is required for the resolution of DNA fragments of > ~20 kb; Birren & Lai, 1994). This meant that it would have been unpractical to work with YAC clones (insert size range ~100 kb to 1 mb; Burke *et al.*, 1987), and left the option of either cosmid libraries (maximum insert size ~45 kb; Collins & Hohn, 1978), or the more recently developed bacteriophage P1 libraries (maximum insert size ~100 kb; Sternberg, 1990). A human genomic library constructed in the bacteriophage P1 vector was chosen in preference to a cosmid library on the basis of the greater insert size of the P1 clones (which at an average of ≥75 kb was larger than the porcine calcitonin receptor gene), and the availability of a pre-gridded human genomic P1 library which would reduce 'hands-on' time and allow the rapid identification of any possible positive clones. Other advantages of P1 clones over YAC and cosmid clones are thought to include lower levels of cloning artifacts, and greater stability of the clones during propagation (Lengauer *et al.*, 1994; Weitzel & Patel, 1994), the latter factor being particularly important considering the problems that had been encountered with the mouse λEMBL3 *Vipr2* clone (Chapter 3).

The copy number of the P1 plasmid in bacterial cells is controlled by the *lac* operon promoter, which permits the copy number of the otherwise single-copy P1 plasmid to increase by up to 20-fold in the presence of the *lac* inducer isopropyl-β-D-thiogalactopyranoside (IPTG), and should therefore allow the relatively easy isolation of large quantities of DNA from these clones. However, although a putative human *VIPR2* gene clone was identified from the library, severe problems were encountered in trying to obtain sufficient quantities of P1 DNA to analyse this clone further, and it became increasingly apparent that extensive subcloning of this P1 clone in order to isolate the human *VIPR2* gene was not a realistic aim within the time available. Nevertheless, despite these difficulties, enough P1 DNA was eventually isolated to allow the human chromosomal localisation of this clone to be determined by fluorescence *in situ* hybridisation (see Chapter 5), and confirmation that the P1 clone contained *VIPR2* gene sequence was subsequently obtained by PCR.

4.2 Methods

4.2.1 Hybridisation screening of the P1 human genomic library

The human genomic P1 library that was used in these studies was obtained while still in preliminary form, from Dr. F. Francis at the ICRF in London, but has now been completed, characterised and incorporated into the ICRF Reference Library Database System. The library (which is derived from a human lymphoblastoid cell line) is constructed in the bacteriophage P1 vector p*Ad*10-*sacB*II (see Chapter 2 Figure 2.6 for vector map) and is presented in the form of duplicate pairs of filters which carry gridded arrays of P1 clones. At the time that we obtained the library it consisted of two pairs of duplicate filters each of which carried a gridded array of 144 × 144 clones (the 2 filters together providing an ~1× genome coverage).

The probes used to screen the P1 library were derived from the rat Vipr2 cDNA sequence. The regions of the rat cDNA sequence that are spanned by the Ac(500)4 and 5' 1 kb probes have been described in Chapter 3. The rat Vipr2 cDNA 3' ~1.2 kb probe spans the region from bp 910 to bp 2125 of the published rat Vipr2 cDNA sequence (Lutz *et al.*, 1993). Probe preparation and labelling reactions were carried out as described in Chapter 2.

Hybridisations were carried out according to the recommendations that were provided with the P1 library. Thus, in contrast to the ES library screens that were described in Chapter 3, the P1 library filters were prehybridised and hybridised at 65°C in a solution containing 7% SDS, 0.5 M sodium phosphate buffer (pH 7.2), 1% BSA, and 1 mM EDTA. The filters were prehybridised for a minimum of 2 hours, after which the probe was added and hybridisation was left to proceed overnight. The following morning, the hybridisation solution was removed and the filters were washed in a solution containing 40 mM sodium phosphate buffer (pH 7.2) and 0.1% SDS. Two 30 minute washes were carried out, at 65°C, in 500 ml of wash solution which had been pre-heated to that temperature. After these washes, excess wash solution was allowed to drain from the filters, which were then wrapped in cling film and exposed to Fuji RX film (Fuji) for 8-12 hours at -70°C in cassettes with two intensifying screens. When a potential positive clone was identified on the autoradiographs, the filters were then re-exposed for a longer time period (~24 hours) so that the 'background signal' produced by the gridded array of P1 clones was clearly visible and the relative position of the potential positive clone within the grid could be easily identified.

4.2.2 Colony Blots of the P1 clone

The colony blot method used was based on the method that was described in Sambrook *et al.*, (1989). The P1 stab culture was streaked onto an L-Agar plate containing 25 μg/ml kanamycin (kan), inverted, and incubated overnight at 37°C. Single colonies were then selected from the plate and patched in a grid pattern, onto each of two BioBlot-NC nitrocellulose blotting filters (Costar®) lying on L-agar +kan (25 μg/ml) plates, and then directly onto a master plate which also consisted of L-agar +kan (25 μg/ml) but did not have a nitrocellulose filter on it. These three plates were inverted and incubated overnight at 37°C to allow the bacterial colonies to regenerate. The master plate was then sealed with parafilm and stored at 4°C until required. After removal from the agar plates, the filters were placed colony side up on a pad of 3MM paper (Whatman) that was saturated in 10% SDS and was left there for 3 minutes to remove debris from the filter. The filters were then left for 5 minutes on each of a series of pads of 3MM paper (Whatman) that were saturated in: i) denaturing solution (1.5 M NaCl, 0.5 M NaOH), ii) neutralising solution (1M NaCl, 0.5 M Tris-HCl pH 7.0), and iii) 2× SSC. After the filters had air dried (~30 minutes), they were placed in an oven and baked at 80°C for 1.5 hours. Before prehybridisation, the filters were soaked briefly in 2× SSC, and then pre-washed for 30 minutes, at 50°C, with gentle shaking, in a solution containing: 5× SSC, 0.5% SDS, and 1 mM EDTA. Any loose debris that remained was carefully removed from the filters using tissues that had been soaked in prewash. Prehybridisation, hybridisation, and autoradiography, were carried out as described in section 4.2.1.

4.2.3 Isolation of P1 clone DNA

Method 1

(based on the method provided by the genome analysis group at the ICRF)

A single P1 colony was used to inoculate 50 ml of 2× YT + kan (25 μg/ml). After growth overnight at 37°C, the culture was divided in two, and 25 ml of culture was transferred into each of two conical flasks containing 800 ml of 2× YT broth + kan (25 μg/ml). The 800 ml cultures were grown at 37°C for 1 hour, at which point, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Growth of the cultures was then allowed to continue at 37°C for another 4 hours.

The cells were collected by centrifugation at 4000 rpm for 20 minutes at 4°C in a Sorvall GSA rotor. After removal of the supernatant, the bacterial pellets from each

800 ml culture were resuspended in 10 ml of alkaline lysis solution 1, (see Chapter 2) transferred to two 250 ml pre-cooled centrifuge bottles, and incubated on ice for 15 minutes. Alkaline lysis solution 2 (30 ml; see Chapter 2) was then added to each centrifuge bottle while gently swirling the bottle, and the bottles were left on ice for another 5 minutes. Once the cells had lysed, 22.5 ml of alkaline lysis solution 3 was added, and the bottles were shaken sharply, then left on ice for 30 minutes. The precipitate (formed on addition of alkaline lysis solution 3) was pelleted by centrifugation for 30 minutes at 13 000 rpm, and the supernatant was transferred to fresh tubes. Isopropanol (45 ml) was added to precipitate the DNA, and the bottles were mixed gently and left at room temperature for 5 minutes, then centrifuged at 9000 rpm for 15 minutes (Sorvall GSA rotor) to pellet the DNA. After removal of the supernatant, the DNA pellets were air dried for ~20 minutes, and each pellet was resuspended in 10 ml of TE. The DNA solution was transferred to a 30 ml Oakridge tube, where 19.66 g of CsCl was added and mixed until dissolved. EtBr (1.5 ml of a 10 mg/ml stock) was then added, and the DNA/CsCl/EtBr solution was mixed gently. Following centrifugation at 10,000 rpm for 10 minutes in a Sorvall SS-34 rotor, the supernatant was transferred to 6 ml polyallomer quick seal ultracentrifuge tubes (Sorvall® Instruments). These tubes were then sealed, and centrifuged (together with suitable balance tubes) in a Sorvall ultracentrifuge (TV-1665 rotor) at 45,000 rpm for 16 hours at 20°C. Isolation and purification of the P1 DNA following CsCl banding was carried out as described in Chapter 2.

* The original method recommended the use of 25 × 89 mm polyallomer ultracentrifuge tubes and a VTi 50 rotor, but we did not have access to this type of rotor.

Method 2

(based on the method of Pierce and Sternberg, 1992a)

Although this method was originally published as miniprep method for P1 DNA isolation (which yielded 1-2 µg of DNA from a 10 ml culture), the authors reported that the method could be scaled up to a 250 ml culture with the expected increase in DNA recovery (Pierce & Sternberg, 1992a).

During this work, several minor modifications were made to the original method when it was scaled up to a 250 ml culture: i) a 2 ml overnight P1 culture was used to inoculate 250 ml of L-broth +kan (25 µg/ml), instead of the single P1 colony that was used to inoculate 10 ml of culture medium in the miniprep method; ii) the cultures were grown for the maximum recommended time of 12 hours after the addition of IPTG; iii)

the cells were pelleted in a 250 ml centrifuge bottle by centrifugation in a Sorvall GSA rotor at ~6000 rpm for 7 minutes at 4°C; iv) following potassium acetate treatment, centrifugation was carried out at 6500 rpm for 25 minutes at 4°C; v) phenol extractions, chloroform extractions, and ethanol precipitation of the DNA, were carried out in 30 ml corex tubes.

Method 3

(modification of the Qiagen cosmid maxiprep DNA isolation protocol suggested by technical staff at Qiagen Ltd)

Isolation of P1 DNA was carried out using a Qiagen plasmid maxi kit (Qiagen Ltd). The isolation procedure used was based on the manufacturer's recommendations for the large scale isolation of cosmid DNA. However, several modifications which were published in the Qiagen Winter 92/93 handbook and/or were suggested by members of their technical assistance group, were also incorporated. These alterations involved: i) growth of the 250 ml bacterial cultures as described in method 2 (see above); ii) resuspension of the bacterial pellet in a volume of P1 solution that was calculated on the basis of 10 ml of P1 solution for every 500 ml of the original bacterial culture volume; iii) increases in the volumes of solutions P2 and P3 used, so that they matched the volume of P1 buffer used; and iv) preheating of the elution buffer (buffer QF) to 37°C prior to use.

N.B. Solutions P1, P2 and P3 are plasmid alkaline lysis solutions that are provided with the Qiagen plasmid maxi kit.

Method 4

(a modified version of the alkaline lysis/CsCl-ethidium bromide gradient protocol for plasmid maxipreps published by Sambrook et al., 1989)

Four test tubes containing 3 ml of L-broth +kan (25 µg/ml) were each inoculated with a single P1 colony, and left to grow at 37°C overnight. Each of these 3 ml cultures was then used to inoculate one of four 2 litre flasks that contained 500 ml of L-broth + kan (25 µg/ml). The 500 ml cultures were grown for at least 20 hours at 37°C, after which the bacterial cells were pelleted into two 250 ml centrifuge bottles, by centrifugation at 6000 rpm for 7 minutes at 4°C in a Sorvall GSA rotor. Each bacterial pellet was resuspended in 18 ml of alkaline lysis solution 1 (see Chapter 2). Forty millilitres of alkaline lysis solution 2 (see Chapter 2) was added to the bacterial suspensions, which were mixed by gently inverting the centrifuge bottles several times, and then left at

room temperature for 5-10 minutes. Once the cells had lysed, ice-cold alkaline solution 3 (20 ml; see Chapter 2) was added, and the suspension was mixed thoroughly by gently shaking the bottles. The bottles were then stored on ice for 20 minutes before pelleting the precipitate by centrifuging the lysate at 6500 rpm (Sorvall GSA rotor) for 25 minutes at 4°C. After centrifugation the supernatant was filtered through 4 layers of cheesecloth (to remove any residual precipitate) and transferred to new centrifuge bottles. Isopropanol (0.6 volume) was then added to the supernatant, mixed well, and the bottles were left at room temperature for at least 30 minutes, to allow the DNA to precipitate. Finally, the DNA was pelleted by centrifugation at 6500 rpm for 20 minutes (Sorvall GSA rotor) at room temperature, the pellets were washed in 70% ethanol, dried, and resuspended in 2 ml of TE (pH 8.0). Two batches of this DNA were then pooled to give a total of 4 ml of DNA solution which was further purified by CsCl-EtBr gradient ultracentrifugation as described in Chapter 2.

4.2.4 Restriction analysis, gel electrophoresis, and hybridisation of Southern blots, of P1 clone and human total genomic DNA

Restriction analysis and Southern blotting of P1 clone DNA (see Figure 2.6 for vector map) and human genomic DNA (Promega) were carried out using methods that differ slightly from those described in Chapter 2.

Restriction enzyme digestion of P1 clone DNA and human genomic DNA was generally carried out in a total reaction volume of 50 µl, containing either 5-8 µl of P1 clone maxiprep DNA or 10 µg of human genomic DNA, 10-20 units of restriction enzyme, 5 µl of an appropriate 10× reaction buffer, and ddH_2O. Digests of P1 clone DNA were incubated at 37°C for a minimum of 5 hours but for no longer than 8 hours as it was found that overnight incubation of these reactions led to signs of degradation of the DNA samples (presumably due to the presence of nucleases in the DNA prep.). Human genomic DNA digests were incubated at 37°C overnight.

Agarose gel electrophoresis of P1 clone and human genomic DNA was carried out at 40-50 volts over a period of approximately 8-12 hours. Southern blotting of the gels (onto GeneScreen Plus membrane; NEN/DuPont) was carried out as described in Chapter 2. The blots were prehybridised and hybridised in 50 ml of solution containing: 25 ml of 50% formamide (deionised), 4.75 ml of 10% SDS, 12.5 ml of 20× SSPE, 5 ml of 50× Denhardt's solution (see Chapter 2), 2.5 ml of 1 M sodium phosphate buffer (pH 7.2), and 250 µl of denatured herring testis DNA (10 mg/ml stock) which was added just prior to prehybridisation. After prehybridisation at 42°C for at least 1 hour, the probe was added and the blots were left to hybridise overnight at

42°C. Finally, the blots were washed at 60°C for 2× 15 minutes, in a solution containing 2× SSC and 0.1% SDS. The blots were then left to air-dry, placed in a cassette with Fuji RX film (Fuji) and two intensifying screens, and stored at -70°C, for periods of 24-96 hours.

4.2.5 Amplimer-mediated PCR of P1 DNA restriction fragments

This method was provided by Dr. A. Wright (MRC Human Genetics Unit, Edinburgh). The method was originally intended for use on YAC DNA and consequently some of the initial DNA restriction digestion and clean-up steps have been modified for use on P1 clone DNA.

Five microlitres of P1 maxiprep DNA was digested using *Bam*H I and *Bgl* II in a total reaction volume of 50 µl . The restriction reaction was incubated at 37°C for 8 hours, after which the digested DNA was purified using the Wizard™ DNA Clean-up System (Promega), and eluted in 40 µl of ddH_2O.

The synthetic amplimers 5' P-GATCGTCGACGGTACCGAATTCT 3' (H3729) and 5' GTCAAGAATTCGGTACCGTCGAC 3' (H3730) were annealed to one another by heating equimolar concentrations of the amplimers to 98°C for 15 minutes and then leaving them to cool to room temperature. Annealing of these amplimers results in the production of a double stranded amplimer which has a *Sau*3A I cohesive end 5' GATC, and is thus compatible with the DNA ends created by the restriction enzymes *Sau*3A I, *Mbo* I, *Bam*H I, *Bgl* II, *Xho* II, and *Bcl* I. Ligation of the double stranded amplimers to the *Bam*H I / *Bgl* II digested P1 clone DNA was carried out at 16°C overnight in a total volume of 20 µl containing either 2 µl, 5 µl, or 10 µl of *Bam*H I/*Bgl* II digested P1 DNA, 1 unit of T4 ligase (Promega), 2 µl of 10× T4 ligation buffer (Promega), 3 µl of a 25 µM double stranded amplimer stock solution, and tissue culture grade water (Sigma). After ligation overnight, the reactions were spun briefly in a microcentrifuge, diluted to 100 µl with ddH_2O, and then heated to 65°C for 10 minutes in order to inactivate the ligase.

The PCR reactions were carried out in triplicate. In each case, 1 µl of the diluted amplimer-ligated P1 DNA was used as a template in a reaction mix containing: 5 µl of magnesium-free 10× PCR reaction buffer (Promega), 2 µl of 50 mM $MgCl_2$, 1.25 µl of 10 mM dNTP mix (Boehringer Mannheim), 1.6 µl of master primer (H3729) 0.8 µM final concentration, 1 µl of Taq Extender (Stratagene), 3 µl of TaqStart™ anti-*Taq* antibody (Clontech), 34.15 µl of tissue culture grade water (Sigma), and 1 µl of *Taq* DNA polymerase (Promega). The reactions were spun in a microcentrifuge for ~20

seconds, overlaid with 2 drops of light white mineral oil (Sigma), and then spun again briefly before being transferred to the PCR machine. An initial denaturation at 94°C for 3.5 minutes was then followed by 35 cycles of: 1.5 minutes at 94°C, 1 minute at 64°C, and 6 minutes at 72°C. After a final extension at 72°C for 15 minutes the reactions were cooled to 4°C and maintained at that temperature until they could be collected.

The PCR reaction products were separated by electrophoresis through a 1% Seaplaque® GTG agarose (FMC Bioproducts) TAE gel. Gel slices were taken from regions of the gel that contained strong bands of amplified DNA, and extraction of the DNA from these slices was carried out using a Geneclean II® kit (Bio 101 Inc.). The purified DNA fragments were then restricted with *Eco*R I (which cuts within the amplimer sequence at the ends of the PCR products), and after heat inactivation of the restriction enzyme (65°C for 20 minutes), a Ready-To-Go™ ligation kit (pUC18/*Eco*R 1/BAP; Pharmacia) was used to subclone the fragments into pUC18. Ligations were carried out according to the manufacturer's instructions.

4.2.6 PCR assay for the detection of a 3' region of the human *VIPR2* gene

Primers 4334 5' CCAGGTATGGGGTTTAGTGGAC 3' and 4335 5' GCCGTCCTCTACTGTTTCCTGA 3' were designed to bind to bases 1640-1619 and 1276-1297 of the human VIPR2 cDNA sequence (Svoboda *et al.*, 1994) respectively. These primer binding sites corresponded to sequences within predicted exons 12 and 13 of the gene, and the primers were expected to amplify the final intron and exon (including part of the 3' UTR) of the human *VIPR2* gene.

Each PCR reaction was carried out in a total reaction volume of 100 µl, containing: 0.5 µl, 1.0 µl, or 1.5 µl of P1 DNA, 2 µl of 10 mM dNTP mix (Boehringer Mannheim), 10 µl of magnesium-free 10× PCR reaction buffer (Promega), 4 µl of 50 mM $MgCl_2$, 2 µl of primer 4334 (15 µM stock), 2 µl of primer 4335 (15 µM stock), 1 µl of *Taq* DNA polymerase (Promega), and tissue culture grade water (Sigma) to 100 µl*. Negative control reactions without template DNA were also set up. The reactions were spun briefly in a microcentrifuge and then overlaid with 2-3 drops of light white mineral oil (Sigma). An initial denaturation of the DNA at 94°C for 5 minutes, was then followed by 30 cycles of: denaturation at 94°C for 1 minute, annealing at 57°C for 1 minute, and extension at 72°C for 1.5 minutes + 4 seconds/cycle. After a final extension at 72°C for 10 minutes, the reactions were cooled to 4°C. Preliminary analysis of the products was carried out by running 8 µl of each PCR reaction on a 1.5% agarose TBE gel.

*In early experiments using this assay TaqStart™ anti-*Taq* antibody (Clontech) was used (2 µl per reaction) in order to obtain a 'hot start' as contamination problems had made it necessary to set up the PCR reactions in another building. In these cases the volume of water added was therefore adjusted accordingly.

4.2.7 PCR assay for the detection of a 5' region of the human *VIPR2* gene

Two different pairs of primers were used in separate experiments. The first primer pair consisted of primer 4336 5' AGTGGCGTCTGGGACAACAT 3' and 4337 5' TCAGGATGAAGGACAGGAACAG 3' which were designed to bind to bases 319-338 and bases 676-655 of the human VIPR2 cDNA sequence (Svoboda *et al.*, 1994) respectively, and to predicted exons 3 and 6 of the *VIPR2* gene. The second primer pair also included primer 4336, but this time it was used in conjunction with primer 9796 5' CGTCTCTGACCATCCGTCACT 3' which was designed to bind to bases 465-445 of the human VIPR2 cDNA sequence (Svoboda *et al.*, 1994), and predicted exon 4 of the *VIPR2* gene. The PCR reaction conditions used in these experiments were essentially the same as those for the 3' *VIPR2* gene assay described in section 4.1.6.

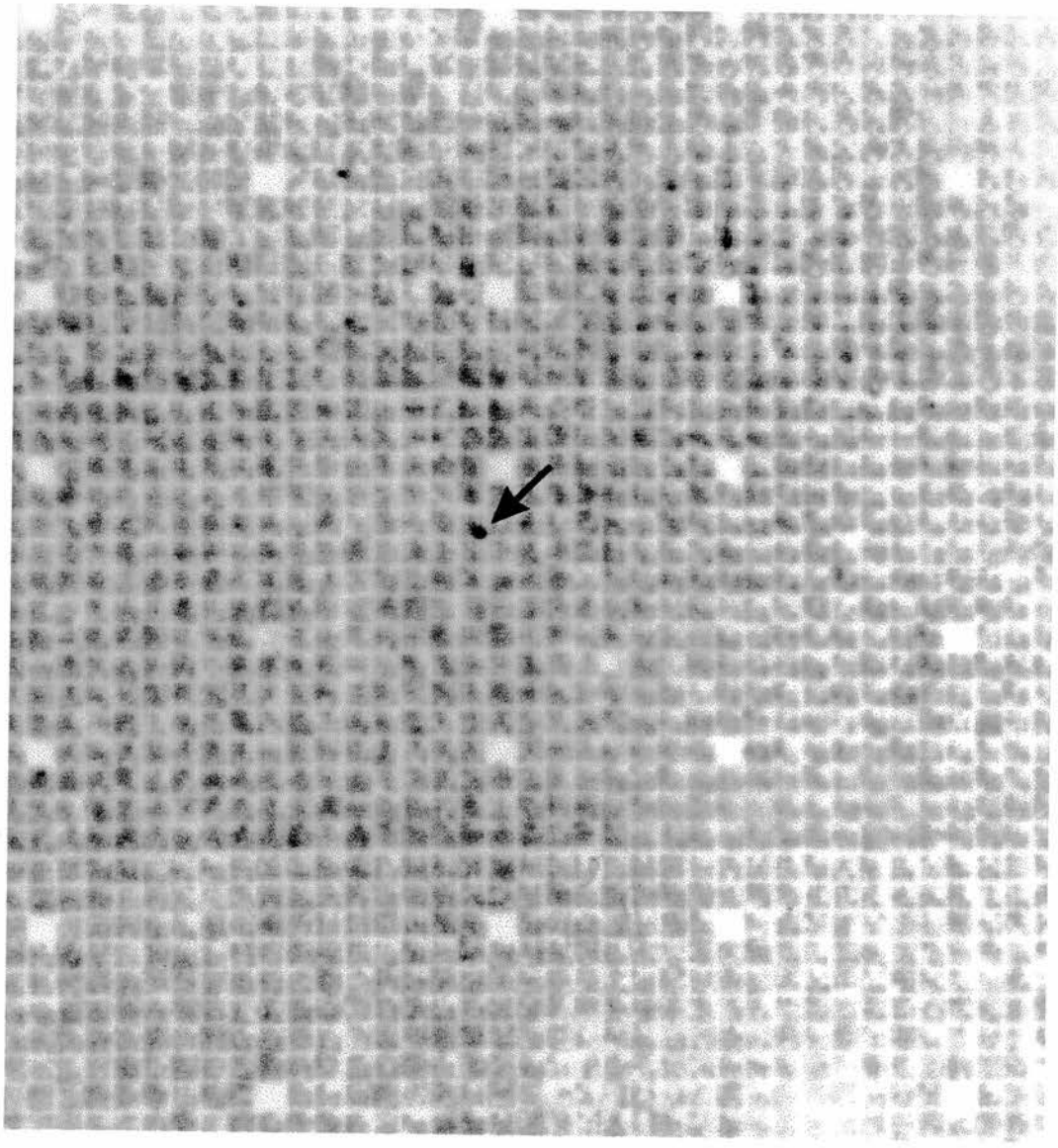
4.3 Results

4.3.1 Screening of the bacteriophage P1 human genomic library

The first screen of the P1 library, which was carried out using the rat Ac(500) 4 probe, did not result in the identification of any putative positive clones on the filters, a result that was somewhat surprising as the probe had previously been successfully used in the isolation of the full length rat Vipr2 cDNA sequence from an olfactory bulb cDNA library (Lutz *et al.*, 1993), and in the isolation of the mouse *Vipr2* genomic clone λG4 from the λ2001 library (Chapter 3). Further attempts to identify a human *VIPR2* clone by screening the P1 library with a probe from the 3' half of the rat Vipr2 cDNA (see 4.2.1 and Figure 3.1) were also unsuccessful. However, a final screen of the library, this time using the 1 kb probe from the 5' end of the rat Vipr2 cDNA, resulted in the identification of a single putative positive clone (to which hybridisation of the probe was observed on duplicate filters; Figure 4.1). This clone was then ordered from the ICRF.

Figure 4.1: Human P1 genomic library filters, hybridised with the radiolabelled 1 kb 5' rat Vipr2 cDNA probe

(hybridisation signal from identified positive clone is arrowed)



4.3.2 Colony blot of the P1 clone

On receipt of the P1 clone stab culture from the ICRF, a colony blot was carried out using the rat 1 kb 5' Vipr2 cDNA probe. This was done in order to: i) verify the results of our P1 library screen and confirm that the correct P1 clone had been selected, and ii) to ensure that the bacterial colonies that were to be grown for the generation of glycerol stocks of the P1 clone, had not lost the P1 plasmid and/or insert DNA. Out of a total of 34 single colonies that were randomly selected for use in the colony blot (filters were prepared by K. West), 22 colonies displayed strong hybridisation to the Vipr2 1 kb cDNA probe, 4 colonies showed weaker hybridisation, and 8 colonies showed little or no hybridisation to the probe (Figure 4.2). Only colonies that had shown strong hybridisation to the probe were used in the generation of P1 clone stocks.

4.3.3 Isolation of DNA from the P1 clone

When we first identified the putative *VIPR2* P1 clone, there were very few published methods for the isolation of P1 clone DNA. Initial attempts to isolate P1 DNA using the alkaline lysis/CsCl-EtBr gradient ultracentrifugation method that was provided by the ICRF (see 4.2.3, method 1) were unsuccessful, and despite repeated efforts, did not produce visible amounts of *VIPR2* P1 clone DNA following CsCl-EtBr gradient banding. The reason for the failure of this method in our hands was not clear. However, during a search of other methods for isolating P1 DNA, it became apparent that the times that were suggested by different investigators, for growth of the bacterial cultures before beginning DNA isolation, varied significantly. Generally, investigators seemed to agree that the initial phase of growth prior to the addition of IPTG should be relatively short, for example 1 hour (ICRF; 4.2.3 method 1), or until the bacteria reach an OD_{590} of 0.1- 0.2 (Pierce & Sternberg, 1992a; Pierce & Sternberg, 1992b), but for the length of time of bacterial growth after the addition of IPTG, recommendations included: 4 hours (ICRF; 4.2.3 method 1), 7 hours (L. Mullins, personal communication), 5-12 hours (Pierce & Sternberg, 1992a) and 5-14 hours (Pierce & Sternberg, 1992b). The method published by Pierce and Sternberg (1992a) was of particular interest as it differed from that suggested by the ICRF group in that it was based on an alkaline lysis P1 DNA miniprep method (that could be scaled up to a 250 ml culture) which included phenol:chloroform extractions, but no CsCl-EtBr gradient ultracentrifugation. Furthermore, the authors of this paper also reported that they had successfully purified P1 DNA using commercially available column chromatography protocols such as those available from Qiagen Ltd. However, although a small amount of DNA was isolated from our P1 clone when a scaled up version of the alkaline lysis

Figure 4.2: Colony blot of the P1 clone, hybridised with the radiolabelled 1 kb 5' rat Vipr2 cDNA probe

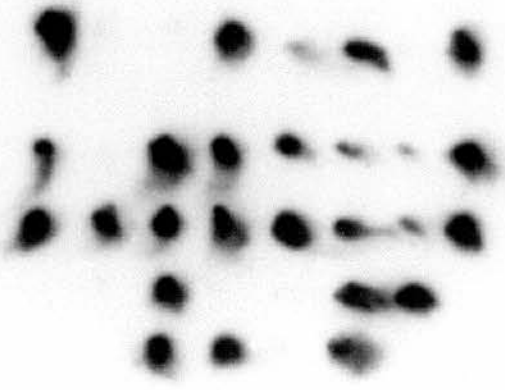
1 2 3 4 5 6 7 8

9 10 11 12 13 14 15 16

17 18 19 20 21 22 23 24

25 26 27 28 29 30

31 32 33 34



based miniprep method was used, this DNA was not cut well by restriction enzymes and partial digestion products were often seen. Nevertheless, the fact that this method had allowed us to isolate DNA from our P1 clone for the first time, seemed to indicate that the extended period of growth after addition of IPTG, may have been beneficial. Efforts to isolate P1 DNA using Qiagen columns, after overnight (12 hour) growth of the cultures in the presence of IPTG, were also successful in obtaining small quantities of P1 DNA, but again the DNA did not cut well with restriction enzymes and there was some evidence of high salt concentration as indicated by narrowing of the lanes of DNA on an agarose gel. As CsCl-EtBr gradient banding of DNA (used in method 1) is generally accepted as being the best method for obtaining high quality DNA, and extending the length of time of bacterial culture (as in methods 2 and 3) seemed to be necessary in order to obtain even small amounts of DNA from our P1 clone, various methods were then tried that included both a long period of bacterial growth and a CsCl-EtBr gradient ultracentrifugation step. The method that was finally devised for use in preparing DNA from the P1 clone, was a modified version of the method published by Sambrook *et al.*, (1989), for the large scale preparation of plasmid DNA by alkaline lysis and CsCl-EtBr gradient banding (see 4.2.3 method 4). In many ways this method does not differ greatly from the method provided by the ICRF (method 1), but the most significant changes are probably those relating the length of time for which the bacteria were grown before isolation of the P1 DNA. In the final method, a 2 ml overnight culture was used to inoculate a flask containing 500 ml of L-broth +kan, and the culture was then allowed to grow for at least 20 hours prior to isolation of the P1 DNA, a time which sounds excessively long, but did result in increased DNA yields. IPTG was not added to these cultures, as surprisingly it was found to make little difference to the final DNA yield obtained with this clone. The DNA obtained using this method appeared to be much purer than that obtained using the other methods and was cut relatively easily by most restriction enzymes. Nevertheless, DNA yields from the 500 ml cultures were still extremely low, to the extent that it was necessary to pool the DNA obtained from two 500 ml cultures, prior to the addition of CsCl and EtBr, in order to be able to visualise the P1 DNA band (which was generally no more than 1 mm deep) following ultracentrifugation. Extremely high RNA to DNA ratios, which had been observed with all of the methods tried, also proved to be a problem when using this method, even after CsCl-EtBr gradient ultracentrifugation. RNase treatment of the DNA solution either prior to or following CsCl-EtBr gradient banding, alleviated this problem to some extent , but considerable amounts of RNA remained. Consequently quantitation of the DNA by UV spectrophotometry was not practical, and estimates of the DNA concentration had to be made by running a small volume of the DNA alongside DNA markers on an agarose gel .

4.3.4 Southern blots of P1 clone DNA

As only very small amounts of P1 DNA could be isolated from each maxiprep, the number of different restriction digests of P1 DNA that could be examined was severely limited. After preliminary Southern blot analysis of the P1 DNA had been carried out using the rat 1 kb 5' Vipr2 cDNA probe on a variety of single restriction enzyme P1 digests, the enzymes *Pst* I and *Bam*H I were selected for use in subsequent Southern blots. Unlike the majority of enzymes tested, *Pst* I and *Bam*H I each generated a series of P1 clone restriction fragments whose pattern of hybridisation to the probe included hybridising bands of relatively low molecular weight, an important factor as we did not have pulse field gel electrophoresis equipment.

The main aim at this stage was to confirm that the P1 clone contained at least part of the *VIPR2* gene, as this would allow us to use the clone as a probe to determine the human chromosomal localisation of the gene. However, because isolation of P1 DNA from this clone was difficult, it was important to be certain that the hybridisation of the cDNA probe to the P1 DNA was specific and represented hybridisation of the probe to the human genomic DNA in the clone's insert, before any attempts at subcloning were made. A Southern blot was therefore carried out on an agarose gel on which *Pst* I and *Bam*H I digests of P1 clone DNA and total human genomic DNA digested with the same enzymes, had been run in parallel (Figure 4.3a). Hybridisation of the blot using the rat 1 kb 5' Vipr2 cDNA probe, resulted in the identification of strongly hybridising bands of ~1.7 kb and ~3.0 kb in the *Bam*H I digested P1 DNA, and of ~2.2 kb and ~2.8 kb in the *Pst* I digested P1 DNA. Hybridisation of the probe to *Bam*H I digested total human genomic DNA, appeared to consist almost entirely of hybridisation to high molecular weight or uncut DNA, and may indicate that digestion of this DNA was incomplete. However, the hybridisation pattern of the probe on *Pst* I digested total human genomic DNA closely mirrored the pattern produced on the *Pst* I digested P1 clone DNA (with hybridising bands present at both ~2.2 kb and ~2.8 kb; Figure 4.3b); a result which strongly suggested that the P1 clone contained at least part of the *VIPR2* gene. Indeed, although it was recognised that this result did not represent conclusive confirmation of the identity of the P1 clone, we were sufficiently confident at this stage to proceed with the establishment of collaboration with the MRC Human Genetics Unit (Edinburgh, UK) in order to determine the chromosomal localisation of the clone (see Chapter 5).

Finally, it may be worth noting that evidence which might indicate that the 5' end of the *VIPR2* gene is present in the P1 clone, was provided by a Southern blot of a *Pst* I digest of the P1 clone and corresponding digests of total human genomic DNA.

Figure 4.3a: Restriction digests of *VIPR2* P1 clone DNA and total human genomic DNA

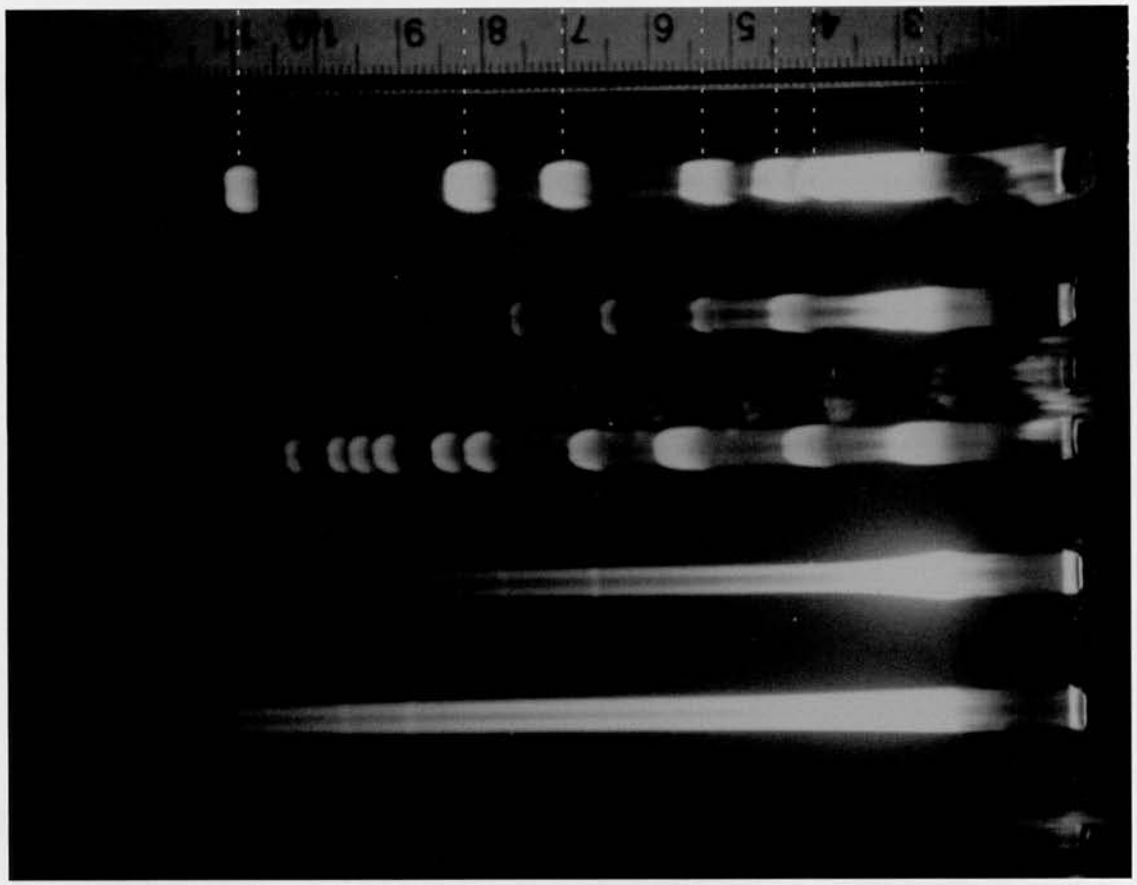
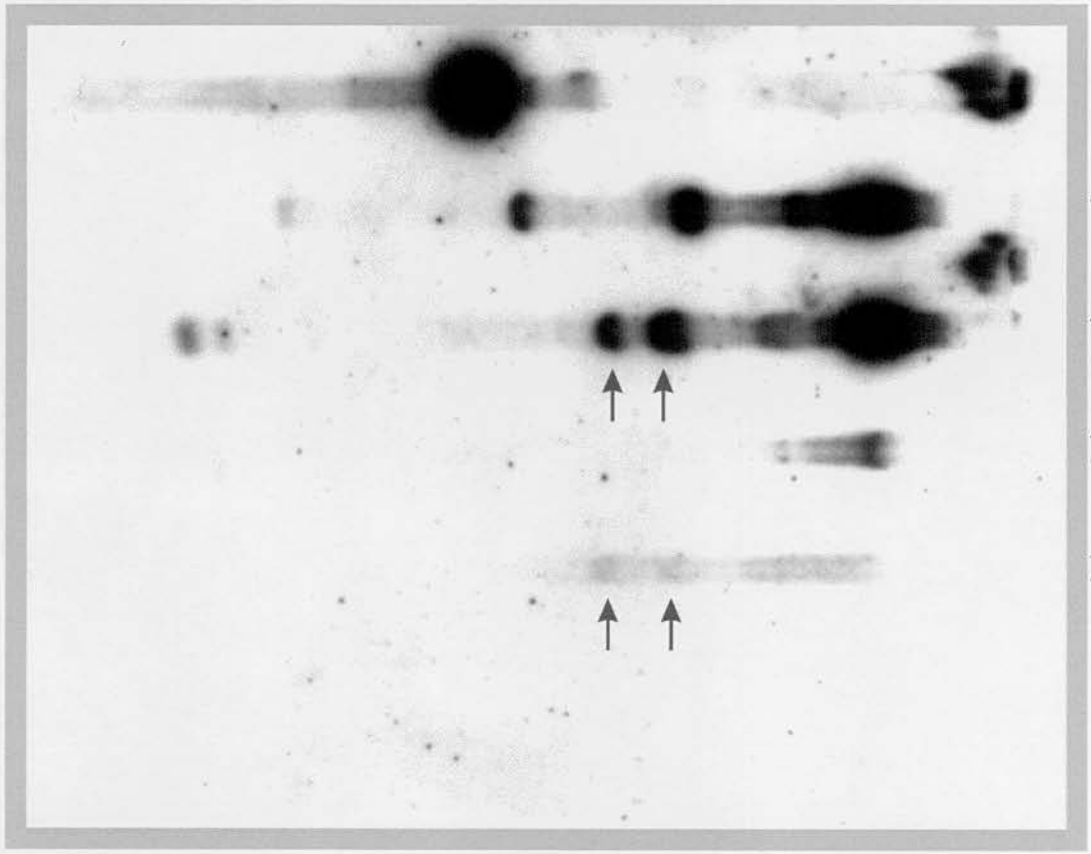

1.02 kb
1.64 kb
2.04 kb
3.05 kb
4.07 kb
5.09 kb
12.2 kb
1 kb ladder
VIPR2 P1 clone *Bam*H I
VIPR2 P1 clone *Pst* I
human genomic DNA *Bam*H I
human genomic DNA *Pst* I

Figure 4.3b: Southern blot of gel shown in 4.3a hybridised with the radiolabelled 1 kb 5' rat Vipr2 cDNA probe

Parallel hybridising bands (of approximately 2.2 kb and 2.8 kb) that are present in *Pst* I digests of both the *VIPR2* P1 clone and total human genomic DNA, are arrowed in red.

(Autoradiograph is shown at actual size)



1 kb ladder

VIPR2 P1 clone *Bam*H I

VIPR2 P1 clone *Pst* I

human genomic DNA *Bam*H I

human genomic DNA *Pst* I

However in contrast to the pattern of hybridisation that was observed when the 1 kb 5' probe was used, the hybridisation seen following the use of the 260 bp 5' probe, consisted almost exclusively of hybridisation to high molecular weight or uncut P1 DNA at the top of the gel. No clear bands of hybridisation to the digests of total human genomic DNA were seen, and although two faint bands of hybridisation were observed, hybridisation to these bands was not much stronger than the non-specific hybridisation to some of the DNA marker bands. This made it very difficult to reach any firm conclusion as to the whether or not the region of DNA spanned by this probe was present in the P1 clone. The presence of some plasmid vector contamination of the probe DNA, indicated by the strong hybridisation of the probe to some bands of the DNA markers, further complicated interpretation of these results as it brought into question the specificity of the hybridisation to the P1 DNA. Although additional Southern blots using different restriction enzymes would probably have resolved this issue, it was not pursued any further at this stage, as P1 clone DNA was scarce, and the results were unlikely to make an immediate contribution to our main aim which was still the verification of the clone's identity.

4.3.5 Attempts at direct subcloning of *VIPR2* gene fragments from the P1 clone

Having isolated a P1 clone that was thought to contain at least part of the human *VIPR2* gene, and identified two *Pst* I restriction fragments that hybridised to the rat 5' 1 kb Vipr2 cDNA probe, the most obvious route to confirming that the P1 clone did contain the *VIPR2* gene, was to subclone and sequence the hybridising bands (which would be expected to contain *VIPR2* gene exon sequences). Isolation and subcloning of DNA from gel slices (see Chapter 2) that were thought to contain the 2.2 kb and 2.8 kb *Pst* I fragments that had hybridised to the Vipr2 cDNA probe (4.3.4.), resulted in the identification of six subclones that contained inserts of ~2.2 kb and two subclones that contained inserts of ~2.8 kb. As the original bands at these positions on the gel had appeared to contain multiple restriction fragments, it would have been preferable to have obtained more transformants and thus increase the chance of isolating the hybridising fragments, but once again, P1 DNA supplies were limiting. Sequencing of the subclones using T7 and SP6 primers did not produce any sequence data that contained regions of similarity to the rodent Vipr2 cDNA sequences. However the sequence obtained from one of the 2.8 kb subclones clearly matched P1 vector sequence, thus confirming that multiple restriction products were present in the original gel slice. Southern blots carried out on the subclone inserts did not show any

hybridisation to the rat Vipr2 full length cDNA probe (probe 6.3), suggesting that these subclones did not contain *VIPR2* gene exon sequences.

4.3.6 PCR amplification and subcloning of P1 *Bam*H I/*Bgl* II amplimer-ligated restriction fragments

An alternative approach to the isolation of fragments containing *VIPR2* sequence from the P1 clone, was suggested by Dr. A. Wright (MRC Human Genetics Unit, Edinburgh). The proposed method (see 4.2.5) which was PCR based (and thus to a large extent overcame the problems that we had previously encountered in obtaining enough DNA for subcloning), involved the ligation of double stranded amplimers (annealed primer pairs with restriction fragment-compatible ends) to restriction fragments generated by *Bam*H I/*Bgl* II double digestion of the P1 DNA, and the subsequent amplification of these fragments using primers that bind to the amplimer sequence.

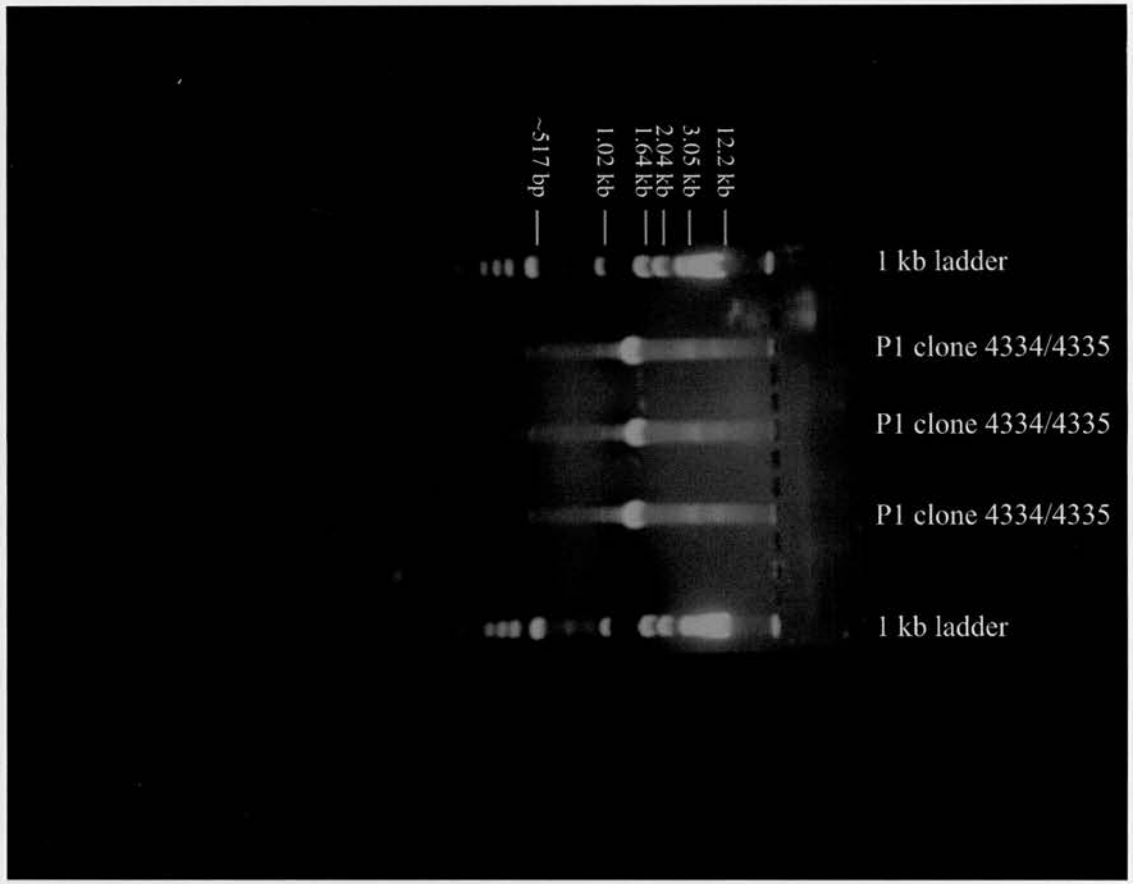
PCR amplification of the amplimer-ligated P1 restriction fragments generated strong bands of PCR product at ~500 bp, 1 kb, 1.7 kb, and 2.3 kb, which were cut out of the gel and purified. Following subcloning of these fragments into pUC18 vector, plasmids that contained inserts of approximately 1 kb (2 different clones), 1.7 kb (1 clone), and 2.3 kb (2 different clones), were isolated. However, sequencing of this rather limited number of subclones, using forward and reverse primers to the vector sequence, did not result in the identification of any sequences that appeared to be *VIPR2* gene coding sequences.

4.3.7 Direct PCR from P1 clone DNA

The publication of the human VIPR2 cDNA sequence by Svoboda *et al.*, (1994), finally allowed specific PCR primers to be designed for the amplification of regions of the human *VIPR2* gene. Two pairs of primers were designed initially, one pair (4334 and 4335) to amplify the predicted last intron and exon of the human *VIPR2* gene, and one pair (4336 and 4337) to amplify a region from the 5' end of the gene (see 4.2.6).

PCR reactions in which the primers 4334 and 4335 were used on P1 template DNA, resulted in the generation of a single PCR product of ~1.3 kb (Figure 4.4). Restriction analysis of this product was carried out using the enzyme *Pst* I. Within the human VIPR2 cDNA sequence (Svoboda *et al.*, 1994), a *Pst* I restriction site is found at bp 1421, which lies in the predicted final exon of the gene, ~200 bp upstream of the 4334 primer site. Thus the release of a fragment of ~200 bp from the 1.3 kb 4334/4335 PCR

Figure 4.4: PCR amplification of a 1.3 kb fragment from the *VIPR2* P1 clone, using primers designed to bind to the 3' region of the VIPR2 cDNA sequence


12.2 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
~517 bp
1 kb ladder
P1 clone 4334/4335
P1 clone 4334/4335
P1 clone 4334/4335
1 kb ladder

product following digestion with *Pst* I (Figure 4.5), strongly indicated that this PCR product was derived from the 3' region of the human *VIPR2* gene.

Reactions carried out with the aim of isolating a region from a more 5' area of the human *VIPR2* gene, using either primers 4336 and 4337, or primers 4336 and 9796, on P1 template DNA (see 4.2.7) were unsuccessful.

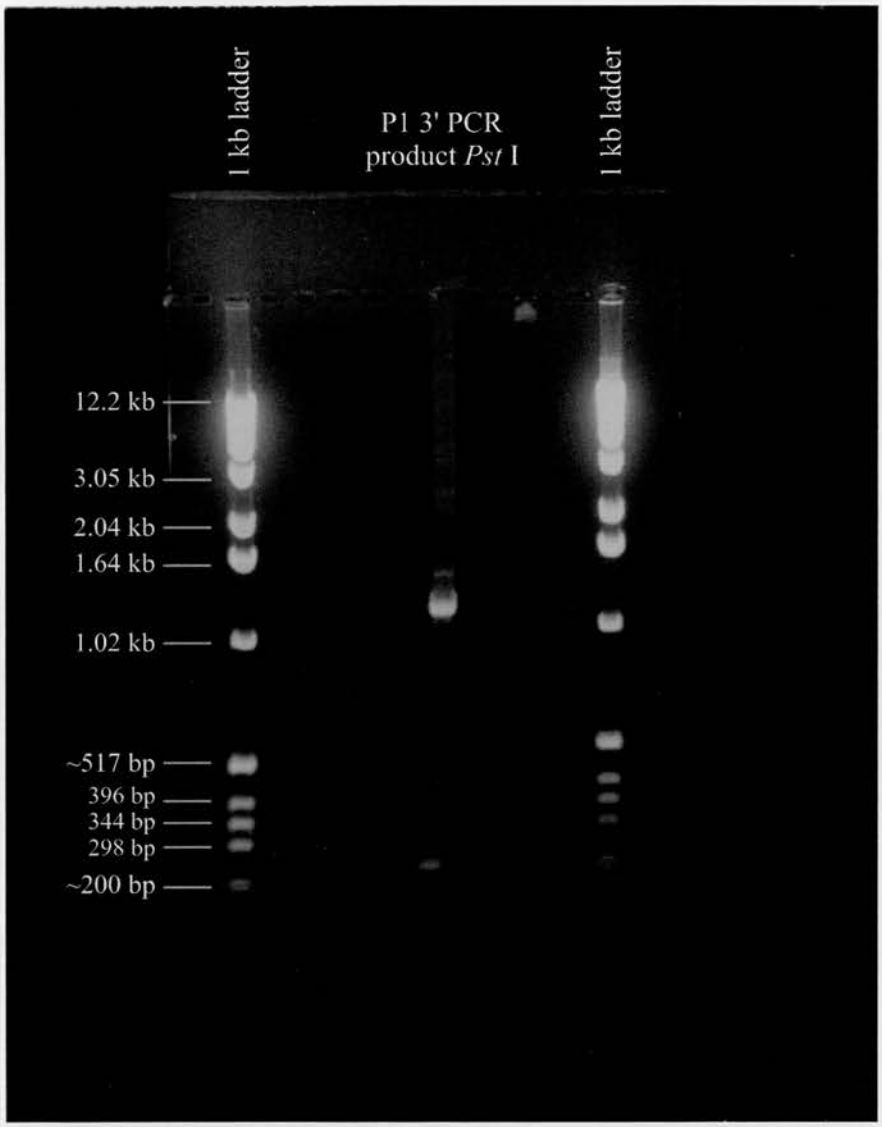
4.3.8 Subcloning and sequencing of the human 3' *VIPR2* gene PCR product

Following subcloning of the putative 3' *VIPR2* gene PCR product (see 4.3.7) into pGEM®-T vector (Promega; see Chapter 2), plasmid DNA was isolated from 16 randomly selected transformants. Restriction analysis of these plasmids indicated that they all contained DNA inserts of ~1.3 kb, and two of these subclones were subsequently sequenced using T7 and SP6 primers. As expected, the sequence data clearly showed that the 1.3 kb product contained sequence that began at the 4335 primer site (just 5' to an exon/intron junction) and followed the human VIPR2 cDNA sequence (Svoboda *et al.*, 1994) for an additional 8 bp before entering intron sequence. In contrast, the sequence extending from the 4334 primer site at the 3' end of the PCR product, spanned ~300 bp of exon sequence that corresponded to the 3' end of the published VIPR2 cDNA sequence (Svoboda *et al.*, 1994) before entering intron sequence. This data conclusively demonstrated that the final intron and exon of the human *VIPR2* gene were present in our P1 clone, and indicated that the final intron of the human gene was ~1 kb.

4.4 Discussion

Unexpected problems, relating to the selection of a suitable probe for the library screen and the subsequent isolation of DNA from the putative *VIPR2* gene P1 clone, were encountered during this work. These problems were largely the result of: i) trying to isolate a genomic clone using cDNA fragments as probes in the absence any knowledge of the gene's intron sizes, and ii) working with a cloning system (P1) that was still relatively new and was not yet in wide use. If this work had been started 9-12 months later, it is likely that some of the problems could have been avoided, as by that stage the structure of the mouse *Vipr2* gene had been much more extensively characterised, and other groups of investigators had reported difficulties in obtaining sufficient amounts of good quality DNA from P1 clones and had begun to look at ways of overcoming this

Figure 4.5: *Pst* I restriction digest of the 3' *VIPR2* PCR product


1 kb ladder
P1 3' PCR
product *Pst* I
1 kb ladder
12.2 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
~517 bp
396 bp
344 bp
298 bp
~200 bp

problem. However, as it was known that other groups were also working on the isolation of this gene, it would not have been wise to delay the work in this way.

The detection of the human *VIPR2* P1 clone by the rat 1 kb 5' Vipr2 cDNA probe was surprising considering the failure of the rat Ac(500)4 Vipr2 cDNA probe and the rat Vipr2 3' ~1.2 kb cDNA probe to detect this clone, especially as the Ac(500)4 probe had been successfully used in the previous isolation of the mouse genomic *Vipr2* clone from the λEMBL3 library. Furthermore, PCR reactions carried out using the P1 clone as a template clearly showed that the predicted final intron and exon (intron 12 and exon 13) of the human *VIPR2* gene were present in the P1 clone, so the absence of the 3' end of the *VIPR2* gene from the P1 clone was not the reason for the failure of the rat 3' Vipr2 cDNA probe to detect this clone. The presence of large introns within the regions that were spanned by these cDNA probes, would be the most obvious explanation for the problems. However, subsequent work on the mouse *Vipr2* gene structure has shown that the introns found towards the 3' end of the mouse gene are generally smaller than those in the 5' half of the gene (Chapter 3), and preliminary results from our lab indicate that this is also the case in the human *VIPR2* gene.

The decision to screen a bacteriophage P1-based genomic library was in some ways the right decision, even in light of the problems that were met in trying to isolate DNA from the P1 clone, as it is now known that the mouse *Vipr2* gene spans over 59 kb in total (Chapter 3), and studies of the human gene (carried out in our laboratory) suggest that the human gene may be even larger. Thus, as had been suspected, it is extremely unlikely that the entire human *VIPR2* gene could have been isolated from a single cosmid clone. It also appears that earlier concerns regarding the stability of YAC and cosmid clones, may have been justifiable. The isolation of an unstable mouse *Vipr2* gene clone from the EMBL3 genomic library did not in itself suggest that the insert DNA was responsible. However, later experiments carried out by S. Scherer and colleagues with the aim of determining the position of the human *VIPR2* gene within a YAC contig (Chapter 5), showed that although the presence of the extreme 3' end of the gene within the YAC contig could be demonstrated using the 3' *VIPR2* gene PCR assay described in this chapter, the result could not be repeated using hybridisation-based methods. This led them to conclude that at least part of the gene had been deleted from the YAC contig during culturing, leaving enough DNA to assay by PCR but not enough to detect by hybridisation (S. Scherer, personal communication), and together with the problems encountered with the mouse clone, this begins to raise the possibility that regions of the *VIPR2* gene may be particularly prone to deletion/rearrangement in clones.

Following the screening of early P1 libraries by several groups it has become increasingly apparent that in many cases the DNA yield from these clones is significantly lower than had been anticipated: an observation that prompted Sternberg and co-workers to report that " the preparation of P1 plasmid DNA with cloned inserts has proven to be a bit more problematical than expected" (Sternberg *et al.*, 1994). The low levels of P1 DNA that were obtained from the *VIPR2* gene P1 clone in this study were probably the result of several factors. First, it seems likely that the *VIPR2* gene P1 clone was a very slow growing clone in comparison with the majority of clones studied by other workers. Similar 'slow growing' P1 clones were noted by Pierce and Sternberg, (1992b), who pointed out that such clones would be very difficult to detect if grown in pooled populations as opposed to an arrayed format. Since the *VIPR2* gene P1 clone was found to require a culture time of over 20 hours as opposed to the 5-14 hours recommended for the growth of P1 clones, it is likely that this clone would have been significantly under-represented in a pooled population, and the decision to screen an arrayed library, although taken purely on the basis of convenience, may have been fortunate.

The apparent failure of IPTG to lead to a rise in the P1 plasmid's copy number and hence an increase in the yield of P1 DNA from the *VIPR2* gene P1 clone, is puzzling. However, as IPTG seemed to have little if any effect on the DNA yields obtained from this clone, there seemed to be little point in continuing to add it to the culture medium. The fact that the P1 DNA was isolated from cells in which it may have been present only as a single copy plasmid, could have had a very significant effect on the DNA yield obtained from this clone, as in general the addition of IPTG to culture medium during the propagation of P1 clone is thought to lead to increases of up to 20-fold in the copy number of the P1 plasmid (Pierce & Sternberg, 1992b). However, the actual copy number of the *VIPR2* P1 plasmid in the absence of IPTG or in its presence is not known, and the reason for the apparent failure of the copy number elevation mechanism in the *VIPR2* P1 clone is not understood, although it may be worth noting that in contrast to the other bacterial strains that can be used for the propagation of P1 DNA, the *lacI^q* gene in the host strain containing the *VIPR2* P1 plasmid, is present on an F' plasmid (Francis *et al.*, 1994).

Another contributing factor to the low yield of DNA obtained from this P1 clone, is almost certainly the bacterial cell type in which the library was constructed and the isolated *VIPR2* clone was subsequently propagated, as there is now considerable evidence to suggest that growth of P1 clones in bacterial host strains that express the Cre recombinase gene (*Cre*), results in reductions in the yield and quality of P1 DNA,

relative to that obtained from cells that do express the *Cre* gene (Kimmerly *et al.*, 1994; Sternberg *et al.*, 1994). Unfortunately, the ICRF P1 human genomic library (Francis *et al.*, 1994) that the *VIPR2* P1 clone was isolated from, was constructed in the Cre^+ host strain NS3145, and at the stage that attempts to isolate DNA from the *VIPR2* P1 clone were made, the link between the presence of the *Cre* gene and problems with P1 DNA isolation was not yet apparent. However, subsequent reports have shown that P1 DNA yields from the Cre^+ host NS3592 are up to 10-fold lower than those from the otherwise identical Cre^- host NS3516 (Sternberg *et al.*, 1994), and that the quality of P1 DNA as determined by its ability to successfully act as a template for fluorescent automated sequencing is significantly higher when the DNA is prepared from Cre^- $endA^-$ DH10B host cells, as opposed to the Cre^+ $endA^+$ strains NS3145 or NS3529 (Kimmerly *et al.*, 1994). Thus, although in the DH10B host cells it is likely that the *endA* mutation contributes at least partly to the improved quality of the P1 DNA isolated, there seem to be distinct advantages in transferring the P1 plasmid DNA from a Cre^+ bacterial host into a Cre^- strain before attempting to isolate P1 DNA.

After the *VIPR2* P1 clone had been isolated a lot of time was spent trying to establish the best method of preparing P1 DNA from this clone, and it was finally decided that when paired with an extended growth time for the bacterial cells, a slightly modified version of the standard alkaline lysis/CsCl-EtBr gradient method for large scale plasmid preparation, produced the best results for our clone. Since then, several reports in which methods for preparing P1 DNA are compared have been published, but the conclusions reached by the authors are quite varied. In most of these reports the primary factor considered is the quality of the P1 DNA obtained using different isolation methods, as the isolation of P1 DNA that is of sufficiently high quality for procedures such as automated sequencing, appears to be particularly problematic. The methods that are generally used for the isolation of P1 DNA do not differ greatly from those used for the isolation of smaller plasmids, and again, alkaline lysis followed by, phenol:chloroform extractions, CsCl-EtBr gradient ultracentrifugation, or purification through commercially available matrix columns e.g. Qiagen columns, seem to be the most commonly used protocols. However, in this study, although some success was achieved in isolating small quantities of DNA from the *VIPR2* gene P1 clone, using each of the methods listed above, the results presented in this chapter suggest that the quality of the P1 DNA obtained using either Qiagen columns or alkaline lysis+phenol:chloroform extractions, was far below that obtained when alkaline lysis was combined with CsCl-EtBr gradient banding of the P1 DNA. These results are in agreement with those of Kimmerly *et al.*, (1994), who compared the success rates of automated fluorescent cycle sequencing reactions carried out on P1 DNA that had been

isolated using: alkaline lysis, triton lysozyme lysis, or Qiagen columns. They found that although the amounts of DNA isolated by these three methods were roughly the same, the highest percentage of successful sequencing reactions was obtained when using DNA that had been isolated using alkaline lysis in conjunction with CsCl-EtBr gradient ultracentrifugation, whereas P1 DNA isolated using Qiagen columns produced the least successful results when sequenced (only ~1/3 of the success rate achieved from alkaline lysis + CsCl-EtBr gradient banding; Kimmerly *et al.*, 1994). In order to avoid the time consuming process of CsCl-EtBr gradient ultracentrifugation of the DNA, this group then went on to develop a modified alkaline lysis method that includes several additional purification steps (phenol:chloroform extractions, PEG precipitation, and spot dialysis) in place of purification by ultracentrifugation. However although this modified alkaline lysis protocol was shown to produce results that were at least equivalent to and possibly slightly better than those obtained using the alkaline lysis + CsCl-EtBr gradient ultracentrifugation method, it is interesting to note that when using this improved method of P1 DNA isolation (which led to a sequencing success rate of 74% when the DNA was prepared from DH10B cells and 26% when the DNA was prepared from NS3592 cells), they were still unable to successfully sequence any P1 DNA that had been isolated from NS3145 cells (Kimmerly *et al.*, 1994). Therefore it appears that even in comparison with P1 DNA isolated from the Cre^+ strain NS3529, P1 DNA isolated from the NS3145 strain (in which the *VIPR2* P1 clone used in this study was grown) is often of particularly poor quality.

The suitability of alkaline lysis methods, for the isolation of P1 DNA that is to be used in *Taq* polymerase catalysed fluorescent end-labelled primer cycle sequencing, was also commented on by Pan *et al.*, (1994), who additionally devised a modified diatomaceous earth (Prep-A-Gene, Biorad) -based method which they found preferable for the isolation of P1 DNA which was to be sequenced by Sequenase catalysed fluorescent dye-terminator sequencing (Pan *et al.*, 1994). Nevertheless, the alkaline lysis method used by Pan *et al.*, (1994) differs from that used by Kimmerly *et al.*, (1994), in that the former group used a very simple alkaline lysis method which did not involve either phenol:chloroform extractions or CsCl-EtBr gradient ultracentrifugation, as they were not deemed necessary. In contrast, Kimmerly *et al.*, (1994) emphasised the importance of CsCl-EtBr gradient ultracentrifugation for the isolation of high quality P1 DNA, and although this step was omitted in their modified alkaline lysis protocol it was replaced by a battery of other purification steps including three phenol:chloroform extractions, PEG precipitation, and dialysis of the DNA. Extractions with phenol and chloroform also featured in the method suggested by Pierce and Sternberg, (1992a), whereas McCormick *et al.*, (1994) reported that the use of Qiagen columns resulted in

the isolation of better quality P1 DNA than that isolated by phenol:chloroform extraction and ethanol precipitation of the lysate. Of the methods of P1 DNA isolation that were tried during the studies of the *VIPR2* P1 clone described in this chapter, only method 2 contained phenol:chloroform extraction steps (see 4.2.3). However as neither this method nor the Qiagen column method that was tried, resulted in the isolation of DNA that could be satisfactorily cut by restriction enzymes, the differences in the results produced by these two isolation methods seemed to be marginal in the case of the *VIPR2* gene P1 clone.

Many of the problems that were encountered in trying to subclone fragments of the *VIPR2* gene from the P1 clone, were directly related to the fact that only very limited amounts of P1 DNA were available. However, in retrospect it seems likely that if there had been more time in which to pursue some of these approaches, they might eventually have been successful. This is especially true for the PCR amplification and subcloning of amplimer-ligated DNA restriction fragments from the P1 clone, as small amounts of P1 DNA were quite sufficient as starting material for this process. It is also possible that a random 'shotgun' approach to the cloning of *VIPR2* gene fragments from the P1 clone, similar to that described for the subcloning of fragments from large introns of the mouse *Vipr2* gene (Chapter 3), might have been more productive than trying to isolate specific bands from agarose gel slices. However, in both cases it would probably be more efficient to generate a 'mini-library' of subclones that could then be screened for the presence of *VIPR2* gene exon sequence (by hybridisation with a rat Vipr2 cDNA probe), than to screen the clones by sequencing as was attempted during this study. This is because although the human *VIPR2* gene is predicted to be quite large (at least as large as the mouse gene and probably considerably larger) the regions encoding exons are likely to occupy a very small proportion of the total DNA in the clone (P1 vector + insert DNA), and consequently the chance of identifying *VIPR2* gene exon sequences within the small number of *Bam*H I/*Bgl* II fragment subclones that were sequenced in this study was very low. If time had been available, the 'mini-library' approach to subcloning *VIPR2* gene fragments would certainly have been tried, but by this stage, the main aim of this work was the confirmation of the P1 clone's identity, and consequently, upon publication of the human VIPR2 cDNA sequence (Svoboda *et al.*, 1994) the subcloning approaches described above were superseded by attempts to amplify regions of the human gene by PCR, using primers that were specific for human *VIPR2* exon sequences.

The successful PCR amplification of an ~1.3 kb region of DNA which spans the last intron and exon of the human *VIPR2* gene (including part of the 3' UTR), not only

supplied the final proof that the P1 clone contained at least part of the human *VIPR2* gene, but also provided us with a PCR-based assay that could be used to detect the presence of the *VIPR2* gene in DNA samples from other sources. This was to prove particularly useful when, following the determination of the *VIPR2* gene P1 clone's chromosomal localisation, we were able to use this PCR assay to determine the position of the 3' end of the human *VIPR2* gene within a YAC contig that was derived from the same chromosomal region, and hence provide a starting point for further mapping studies (see Chapter 5). Thus, although due to problems encountered in isolating DNA from the *VIPR2* gene P1 clone, characterisation of this clone was not pursued beyond the point required for the determination of the *VIPR2* gene's chromosomal localisation, these studies eventually led indirectly to the identification of YAC and cosmid clones encoding the *VIPR2* gene, and characterisation of these clones is currently underway in our laboratory.

Since the *VIPR2* P1 clone was initially identified using a rat cDNA probe that spanned the region from predicted exon 1 to predicted exon 8 of the gene, and this probe was also successfully used in Southern blots of restriction digests of the *VIPR2* P1 plasmid, it appears that sequences located 5' of the end of predicted exon 8 are present in this clone. However, as inconclusive results were obtained from the one Southern blot that was hybridised with the 260 bp probe from the extreme 5' end of the rat cDNA, the question of whether or not the 5' end of the human *VIPR2* gene's coding sequence is present in the P1 clone remains unanswered.

Subsequent characterisation of YAC and cosmid clones encoding the human *VIPR2* gene, carried out in our laboratory, has demonstrated that the introns 3 and 5 of the human gene are 7 kb and >12 kb respectively (E. Lutz, unpublished data), thus explaining the failure of attempts to use PCR to amplify regions of the *VIPR2* gene that spanned introns 3, 4 and 5, or even intron 3 alone. Nevertheless, as intron 1 of the human *VIPR2* gene has now been shown to be ~2.1 kb in size (E. Lutz, unpublished data), this intron and adjacent exon sequences would be a suitable target for future PCR assays to detect the 5' end of the human gene.

Chapter 5

Chromosomal localisation of the *VIPR2* gene in mouse and human

5.1 Introduction

The initial aim of investigators involved in the human and mouse genome mapping projects, is to generate high density physical and genetic maps of these genomes (Dietrich *et al.*, 1996; Guyer & Collins, 1995; Hunter *et al.*, 1996; Spurr, 1996). However, the chromosomal localisation of a gene, in addition to providing a new marker on a genomic map, has the potential to lead to a wide variety of information about the gene and its product(s), including: determination of the gene's copy number; evidence for evolutionary relationships between closely related genes; and in some instances, indications of possible linkage to mapped genetic diseases.

The mouse is often the first choice as a model in which to study genetic diseases that have been initially characterised in humans, due to the strong similarities between mouse and human developmental pathways, and the relative ease with which the mouse genome can be manipulated (Wynshaw-Boris, 1996). Comparative mapping of genes in the mouse and human genomes has led to the identification of many regions of synteny (groups of genes that map to the same chromosome) that are conserved between the two species (O'Brien *et al.*, 1993). To date over 900 loci have been mapped to the chromosomes of both species, and the syntenic regions defined by these loci often allow predictions to be made regarding the likely localisation of the human homologue of a gene that has been mapped in mouse and vice versa (Searle *et al.*, 1994).

The diverse nature of the physiological functions in which VIP and PACAP have been implicated, together with the relatively widespread distribution of *Vipr2* mRNA within the body, albeit often at very low levels (see main introduction), suggested many possibilities for the involvement of the *VIPR2* gene in genetic disease, but did not immediately point towards a specific disease locus. Nevertheless, in light of the reported neurotrophic properties of both VIP and PACAP, it is interesting to note that the human *PACAP* gene maps to the human chromosomal region 18p11 (Hosoya *et al.*, 1992) which, as the authors stated, is associated with holoprosencephaly, a developmental defect of the brain (Muenke, 1994). The human *VIP* gene maps to the human chromosomal region 6q24-q27 (Gotoh *et al.*, 1988; Gozes *et al.*, 1987).

The genes encoding some very closely related G-protein coupled receptors have been shown to be quite tightly linked within the human genome, for example the α2/β1 and α1/β2 adrenergic receptor gene pairs (which are found on chromosomes 10 and 5 respectively; Yang-Feng *et al.*, 1990), or the olfactory receptor gene cluster on chromosome 17 (Ben-Arie *et al.*, 1994). However, the recent localisation of the human *PACAPR1* and *VIPR1* genes to 7p14 (Stoffel *et al.*, 1994) and 3p22 (Sreedharan *et al.*, 1995) respectively, indicates that this is unlikely to be the case for VIP and PACAP receptors. Instead, these receptors seem to follow the same pattern as other members of the secretin/glucagon receptor family (Table 5.1) and the G-protein coupled receptor superfamily as a whole, in that the genes of family members are widely scattered throughout the human genome (Wilkie *et al.*, 1993), and their chromosomal localisation is generally unrelated to that of any peptide ligand(s) (Table 5.2).

In collaboration with colleagues at the MRC Human Genetics Unit in Edinburgh, we have determined the chromosomal localisation of the mouse and human *VIPR2* genes by fluorescence *in situ* hybridisation (FISH). Several other methods such as PCR analysis of somatic cell hybrids, or analysis of the European collaborative interspecific backcross (EUCIB) panel (for the mouse gene), could potentially have been used in our own laboratory to provide similar information, but the relatively high resolution attainable with FISH on both mouse and human chromosomes (Heiskanen *et al.*, 1996) made this our preferred method. Our initial results from the chromosomal localisation of the human *VIPR2* gene suggested the possible involvement of this gene in at least one previously mapped human genetic disorder, and we therefore proceeded to map the position of the gene more precisely (with assistance from Dr. S. Scherer and colleagues at the Hospital for Sick Children, Toronto, Canada), using previously localised YAC and cosmid contigs from the chromosomal region surrounding the *VIPR2* gene.

5.2 Methods

5.2.1 Chromosomal mapping by FISH

Metaphase preparations were made from short-term peripheral blood cultures, lymphoblastoid or fibroblast cultures and mouse ES cells by standard methods.

Table 5.1: Chromosomal localisations of genes encoding endogenous ligands of members of the secretin/glucagon receptor subfamily

Peptide	Chromosomal localisation of the gene in human	Chromosomal localisation of the gene in mouse	Reference(s)
Calcitonin / calcitonin gene-related peptide	11p15.2-p15.1	chromosome 7	(Hoovers *et al.*, 1993; Lalley *et al.*, 1987)
Corticotropin-releasing hormone	8q13	chromosome 3 (~8 cM from centromere)	(Arbiser *et al.*, 1988; Knapp *et al.*, 1993)
Gastric inhibitory polypeptide	17q21.3-q22		(Fukushima *et al.*, 1989; Inagaki *et al.*, 1989)
Glucagon	2q36-q37	chromosome 2 (~37 cM from centromere)	(Lalley *et al.*, 1987; Schroeder *et al.*, 1984)
Growth hormone-releasing hormone	20q11.2	chromosome 2 (~89 cM from centromere)	(Pezzolo *et al.*, 1994)
Parathyroid hormone	11p15.3-p15.1	chromosome 7 band F (~53 cM from centromere)	(Lalley *et al.*, 1987; Tonoki *et al.*, 1991)
Parathyroid hormone-related protein	12p12.1-p11.2	chromosome 6 band F-G (~75 cM from centromere)	(Mangin *et al.*, 1988; Seldin *et al.*, 1992)
Pituitary adenylate cyclase activating polypeptide	18p11.32		(Chang *et al.*, 1993)
Vasoactive intestinal peptide	6q24-q27		(Gotoh *et al.*, 1988; Gozes *et al.*, 1987)

Table 5.2: Chromosomal localisations of genes encoding members of the secretin/glucagon receptor subfamily

Receptor	Chromosomal localisation of the gene in human	Chromosomal localisation of the gene in mouse	Reference(s)
Calcitonin receptor	7q21.3	proximal region of chromosome 6	(Perez Jurado *et al.*, 1995; Yamin *et al.*, 1994)
Corticotropin-releasing hormone receptor	17q12-q22	chromosome 11 (~62 cM from centromere)	(Burrows *et al.*, 1995; Polymeropoulos *et al.*, 1995)
Gastric inhibitory polypeptide receptor	19q13.3		(Gremlich *et al.*, 1995; Stoffel *et al.*, 1995)
Glucagon receptor	17q25		(Lok *et al.*, 1994; Menzel *et al.*, 1994)
Glucagon-like peptide-1 receptor	6p21	proximal region of chromosome 17	(Kershaw *et al.*, 1995; Stoffel *et al.*, 1993)
Growth hormone-releasing hormone receptor	7p15-p14	chromosome 6 (~26 cM from centromere)	(Gaylinn *et al.*, 1994; Godfrey *et al.*, 1993)
Parathyroid hormone / parathyroid hormone-related peptide receptor	3p21.1-p22	chromosome 9 band F (~58 cM from centromere)	(Gelbert *et al.*, 1994; Pausova *et al.*, 1994)
Parathyroid hormone type 2 receptor	2q33		(Usdin *et al.*, 1996)
Pituitary adenylate cyclase activating polypeptide type 1 receptor	7p14		(Stoffel *et al.*, 1994)
Secretin receptor	2q14.1		(Mark & Chow, 1995)
Vasoactive intestinal peptide type 1 receptor	3p22		(Sreedharan *et al.*, 1995)
CD97 antigen	19p13.2-p13.12		(Hamann *et al.*, 1995)
Egf module-containing, mucin-like hormone receptor	19p13.3		(Baud *et al.*, 1995)

Probes

VIPR2 gene probes

The probe used in the chromosomal localisation of the mouse the *Vipr2* gene was a 4 kb *Bam*H1 restriction fragment containing the final three introns and exons of the mouse gene, which had been isolated and subcloned into pBluescript (Stratagene). The isolation and characterisation of this mouse genomic fragment are described in detail in Chapter 3. Plasmid DNA for use in FISH was prepared using a Qiagen plasmid maxi kit (Qiagen Ltd.) according to the manufacturer's instructions.

For the chromosomal localisation of the human *VIPR2* gene, a human *VIPR2* gene P1 clone, was used as the probe. The isolation and characterisation of this clone are described in Chapter 4. P1 clone DNA for use in FISH was prepared by alkaline lysis followed by CsCl banding and RNase treatment. A human VIPR2 cDNA (isolated by E. Lutz in our laboratory) was also used as a FISH probe, to confirm the initial result obtained with the P1 clone.

YACs and cosmids that map to distal 7q

YACs HSC7E207 and HSC7E526 both map to 7q36, YAC HSC7E207 contains the DNA markers *D7S392* and *D7S550*, and YAC HSC7E526 contains *D7S68* (Kunz *et al.*, 1994). Cosmid 29f5 contains marker *D7S168* . Cosmid 30 has been described previously (Riethman *et al.*, 1993) and contains marker *D7S590*. YAC HTY146 also known as HTY3006 (provided by Anna Jäuch, University of Heidelberg, Institute of Human Genetics, Heidelberg; and Harold Riethman, Wistar Institute, Philadelphia) is a 240 kb chromosome 7 telomeric YAC (insert size 230 kb) originally isolated by H. Riethman (Riethman *et al.*, 1993).

Cell lines

The cell lines used in this study were obtained from the National Institute of General Medical Sciences (NIGMS) Human Mutant Cell Repository (Camden, USA). Details of the cell lines are given in Table 5.3.

Probe labelling, hybridisation, and detection

This method was provided by J. Fantes (MRC Human Genetics Unit , Edinburgh) who also carried out the FISH studies shown in this chapter.

Table 5.3: Details of cell lines used in FISH studies

(MC=microcephaly, MDF=mildly dismorphic facial features, MR=mental retardation, CP=cleft palate, HPE=holoprosencephaly)

Cell line	Phenotype	Reported Cytogenetic abnormalities	Comments and references	FISH results with VIPR2
GM03240	MC, MDF	46, XY, del (7) (q34>qter) deletion of 7q34-7qter	Gurrieri *et al.*, 1993	One copy of VIPR2 on normal 7q
GM00657	MC, MR, CP, small chin	46, XY, t (7;18) (7pter>7q34::18q12.2>18qter; 18pter >18q12.2::7q36>7qter) mat deletion of 7q34-7q36	Punnett *et al.*, 1979. Gurrieri *et al.*, 1993.	One copy of VIPR2 on normal 7q
GM07216	HPE	46, XX, -7, +der (7) t (3;7) (7pter>7q36::3p21.3>3pter) deletion of 7q36; duplication of 3p21.3>3pter	Kurtzman *et al.*,1987. Gurrieri *et al.*, 1993.	One copy of VIPR2 on normal 7q
GM10064	HPE	46, XX, -7, +der (7) t (7;13) (7pter>7q34::13q12.3>13qter) pat deletion of 7q34-qter; duplication of 13q12.3>13qter	Gurrieri *et al.*, 1993.	One copy of VIPR2 on normal 7q One copy of VIPR2 on normal 13p
GM10067	normal	46, XY, t (7;13) (7pter>7q34::13q12.3>13qter; 13pter>13q12.3::7q34>7qter) balanced translocation between 7 and 13	Father of GM10064	One copy of VIPR2 on normal 7q One copy of VIPR2 on normal 13p One copy of VIPR2 on 13q der

The P1 clone, cDNA, cosmids, YACs, and plasmid, were labelled by nick translation with biotin-16 or digoxigenin-16 dUTP. Chromosome 7 paint was produced by Alu-PCR of DNA from somatic cell hybrids containing a single human chromosome 7 (Kofman-Alfaro *et al.*, 1994). Amplified DNA was labelled with biotin-16 or digoxigenin-16 dUTP by nick translation. Chromosome 13 paint labelled with biotin was obtained from Cambio.

Hybridisation was carried out as described in Fantes *et al.*, 1992 (Fantes *et al.*, 1992).

After single probe hybridisation, biotin was detected by successive layers of avidin FITC, biotinylated anti-avidin and avidin-FITC. When two probes, labelled with either biotin or digoxigenin, were hybridised simultaneously, the signals were detected using successive layers of FITC conjugated sheep anti-digoxigenin, FITC-conjugated anti-sheep antibody plus avidin-Texas Red, biotinylated anti-avidin and avidin-Texas Red. The *VIPR2* gene was localised on prometaphase chromosomes produced by early S-phase synchronisation with methotrexate of PHA stimulated peripheral blood cultures, followed by release with bromodeoxyuridine (BrdU). In this case hybridisation signals were detected by successive layers of avidin-Texas Red, biotinylated anti-avidin and avidin-Texas Red. A final incubation with FITC conjugated anti-BrdU gave a replication G-banding pattern (Craig & Bickmore, 1994). The chromosomes were counterstained with DAPI. Texas Red hybridisation signals were visualised simultaneously with FITC signals through a Chroma Pinkel #1 filter block fitted to a Zeiss Axioplan microscope, with a Photometrics CCD camera and Digital Scientific software.

5.2.2 Physical mapping of the position of the human *VIPR2* gene, relative to the 7q telomere

PCR analysis of YAC contigs

Unless otherwise stated, ~1 ng of YAC DNA template was used in each PCR reaction. PCR reaction conditions and cycling parameters were the same as those described for the characterisation of the human *VIPR2* gene P1 clone (see Chapter 4, section 4.2.6). The human genomic DNA (used as a control), was obtained from Promega. DNA from somatic cell hybrids containing a single human chromosome-7, was provided by J. Fantes (MRC Human Genetics Unit, Edinburgh).

Use of YAC and cosmid clones to determine the position and orientation of the human *VIPR2* gene relative to the 7q telomere

This method was provided by Dr. S. Scherer (Hospital for Sick Children, Toronto) who also carried out the studies relating to the fine mapping of the position and orientation of the human *VIPR2* gene that are shown in this chapter.

Previously isolated YAC and cosmid clones

YAC clones HSC7E526 and HSC7E1357 were isolated from a chromosome 7-specific library prepared from human-hamster somatic cell hybrid lines (Scherer *et al.*, 1992), E145A7 was isolated from the St. Louis total human genomic YAC library (Brownstein *et al.*, 1989), and HTY146 (locus *D7S427*) also known as HTY3006 was isolated from a telomeric YAC library that was described previously (Riethman *et al.*, 1993). The cosmid clones 2, 3 (*D7S591*), 6 (*D7S592*), 14, 19, 30 (*D7S590*), 48 (*D7S593*), and 49 were derived from a library made from the YAC clone HTY146 and have been described previously (Riethman *et al.*, 1993).

Hybridisation screening of cosmids.

Hybridisation screening of the Lawrence Livermore National Laboratory chromosome 7-specific library, and of the HYT146-derived cosmid clones described above, was carried out using a human VIPR2 cDNA (isolated by Dr. E. Lutz in our laboratory) corresponding to the entire coding region (13 exons) of the published sequence (Svoboda *et al.*, 1994) as a probe .

5.3 Results

The FISH studies shown in this chapter were carried out by J. Fantes at the MRC Human Genetics Unit, Edinburgh, with assistance from M. Gray and S. Boyle. Screening of the chromosome 7 cosmid library, and related work on the fine mapping of the position and orientation of the human *VIPR2* gene, was carried out by Dr. S. Scherer and colleagues at the Hospital for Sick Children, Toronto, Canada.

5.3.1 Chromosomal localisation of the mouse *Vipr2* gene

To determine the chromosomal localisation of the mouse *Vipr2* gene, FISH analysis was carried out on metaphase preparations from mouse embryonic stem cells, using a

subcloned 4 kb *Bam*H1 restriction fragment of the mouse gene as a probe (see section 5.2 and chapter 3). Initial experiments showed a strong clear signal on the telomeric region of mouse chromosome 12. However the use of FISH for the chromosomal localisation of mouse genes is still relatively rare, and requires a high degree of skill, because the acrocentric nature of mouse chromosomes makes them quite difficult to type. Therefore, to verify the identity of the chromosome to which mouse *Vipr2* probe hybridised, a co-localisation experiment was carried out using the mouse *Vipr2* gene probe and a cosmid clone containing the *Spi-2* gene. The *Spi-2* gene has previously been mapped to mouse chromosome 12 using the interspecific backcross method (Hill *et al.*, 1985). The results of the co-localisation experiment (Figure 5.1) clearly showed that the mouse *Vipr2* gene is located on mouse chromosome 12 band F2.

5.3.2 Chromosomal localisation of the human *VIPR2* gene

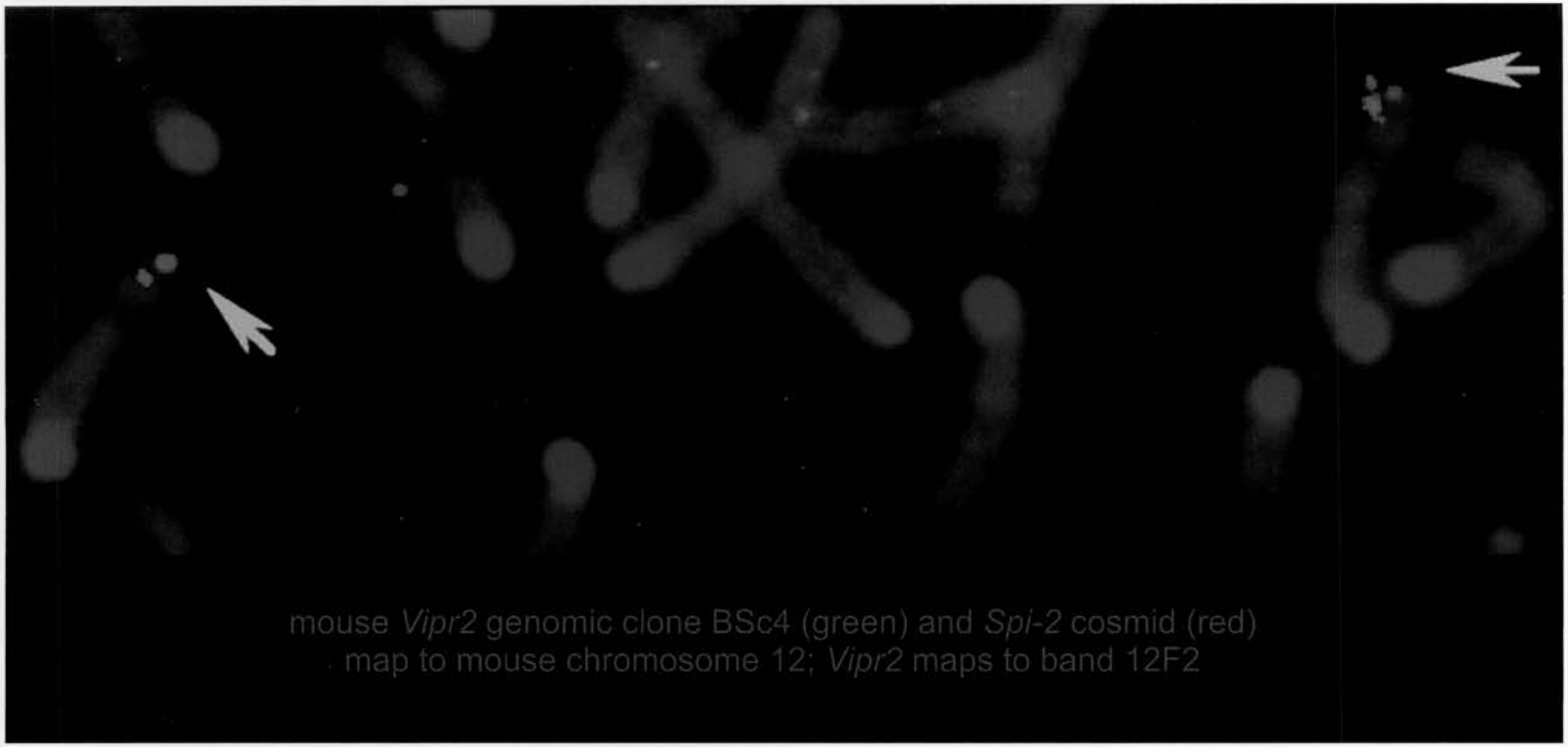
Localisation of the human *VIPR2* gene was carried out by FISH on human prometaphase chromosomes, using a human *VIPR2* gene P1 clone (see section 5.2 and Chapter 4) as a probe. Two signals, corresponding to hybridisation of the probe to both sister chromatids, were seen on chromosomal region 7q36.3 (Figure 5.2). No reproducible hybridisation to any other chromosomal region was observed. This result was very interesting as several human genetic diseases map to 7q36, and one of these, holoprosencephaly (HPE), is also associated with the chromosomal region to which the *PACAP* gene has been mapped (Hosoya *et al.*, 1992).

The localisation of the human *VIPR2* gene was confirmed by FISH of a human VIPR2 cDNA probe (isolated by E. Lutz in our laboratory), and we then proceeded to examine cell lines derived from patients who had HPE associated with 7q deletions.

5.3.3 Localisation of the *VIPR2* gene in holoprosencephaly cases associated with 7q deletions.

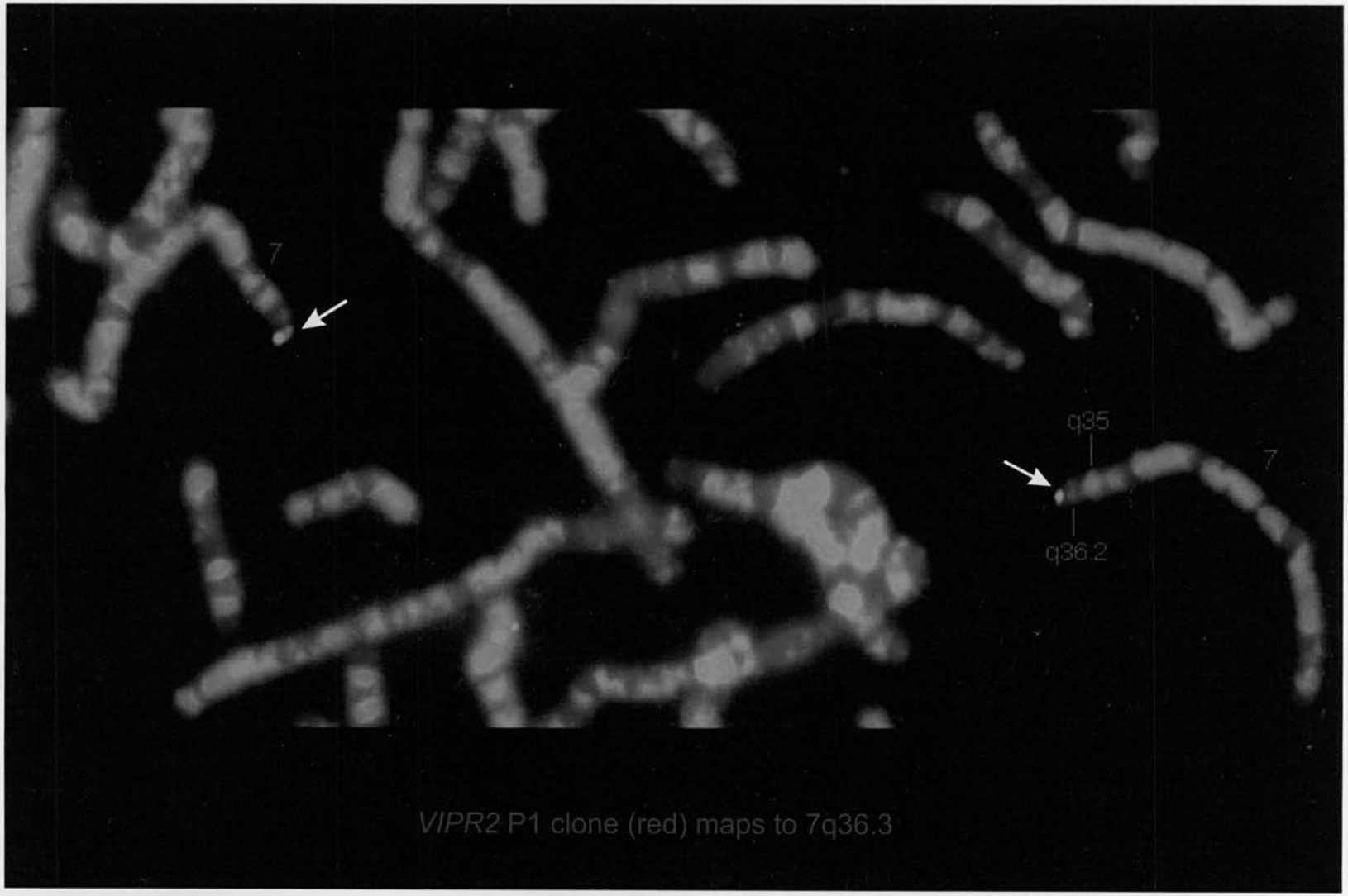
Previous studies by Gurrieri *et al.*, (1993), on cell lines from HPE 7q deletion cases, have led to the definition of a minimal critical region for 7q associated HPE (HPE3). The distal extent of the deletion in cell line GM00657 defines the distal boundary of the *HPE3* minimal critical region, and the proximal extent of the deletion in cell line GM03240 (and 2 other cell lines that have deletions very similar to those of GM03240) defines the proximal boundary of the *HPE3* minimal critical region (Gurrieri *et al.*, 1993)

Figure 5.1: Chromosomal localisation of the mouse *Vipr2* gene



mouse *Vipr2* genomic clone BSc4 (green) and *Spi-2* cosmid (red)
map to mouse chromosome 12; *Vipr2* maps to band 12F2

Figure 5.2: Chromosomal localisation of the human *VIPR2* gene



VIPR2 P1 clone (red) maps to 7q36.3

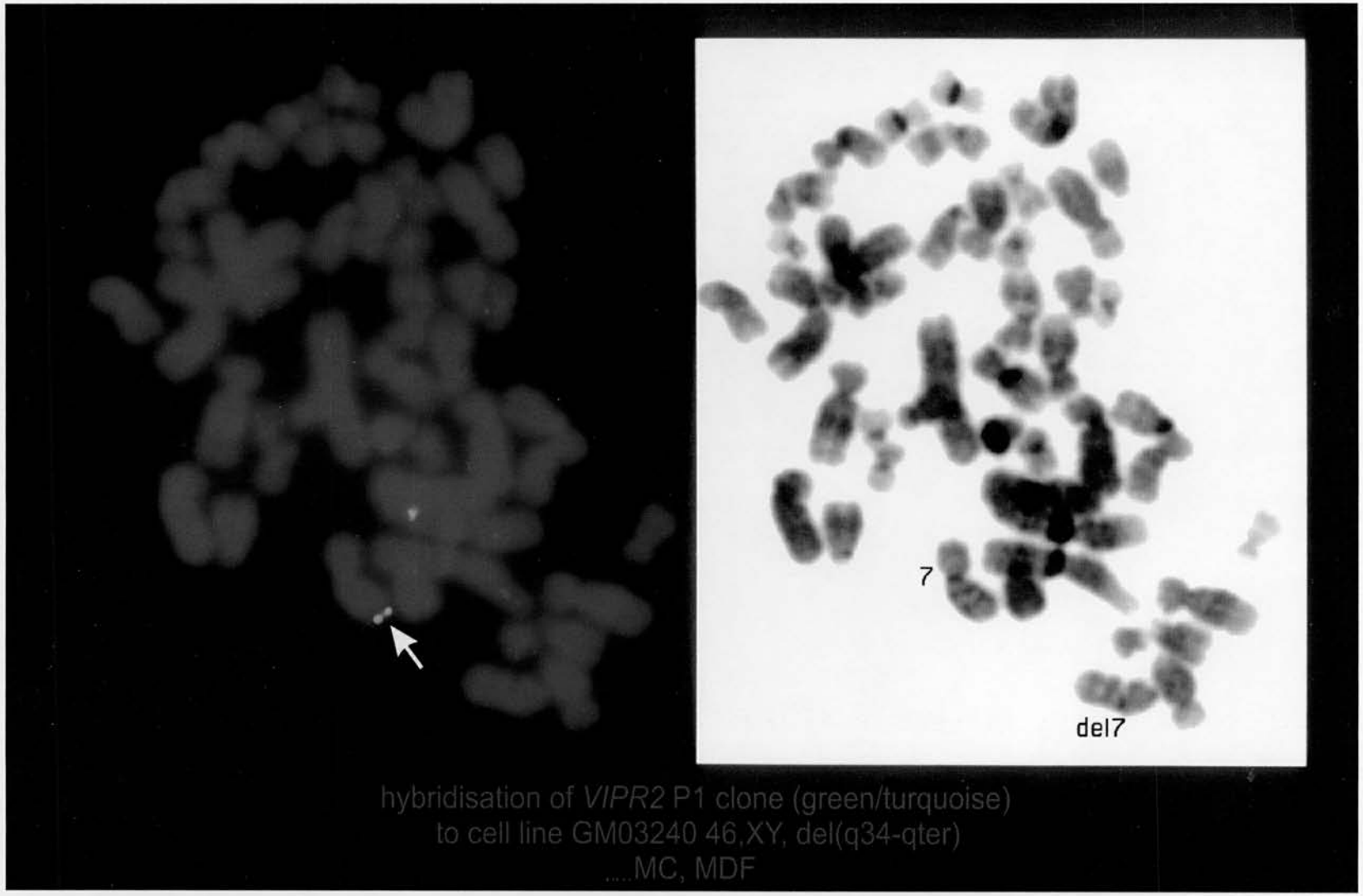
FISH was used to localise the *VIPR2* gene in cell lines derived from 4 of the patients described by Gurrieri *et al.,* (1993) and one unaffected relative. Details of the cell lines used (and a summary of the results obtained) are given in Table 5.3.

Analysis of metaphase preparations from cell lines GM03240 and GM00657 (the cell lines used by Gurrieri *et al.* to define the *HPE3* minimal critical region), using the *VIPR2* gene P1 clone as a probe, showed that only one copy of the *VIPR2* gene was present (on the normal chromosome 7) in each of these cell lines (GM03240 is shown in Figure 5.3; GM00657 is shown in Figure 5.4). These results showed that the *VIPR2* gene is located within the *HPE3* minimal critical region, as defined by Gurrieri et al., (1993). Analysis of cell line GM07216, which has a deletion that extends to the telomere of 7q from a point proximal to the breakpoint in GM03240 (Gurrieri *et al.*, 1993), also showed that only a single copy of the *VIPR2* gene was present (on the normal chromosome 7; Figure 5.5). This result is in agreement with those obtained from GM03240 and GM00657.

Cell line GM10064 is thought to contain an interstitial deletion of 7q that extends from a point centromeric to the proximal boundary of the *HPE3* minimal critical region, to a point that is slightly telomeric to the distal boundary of the *HPE3* minimal critical region (Gurrieri *et al.*, 1993). Since we had already determined that the *VIPR2* gene lies within the *HPE3* minimal critical region, it was expected that one copy of the *VIPR2* gene would be present, on the normal chromosome 7 in this cell line. We were therefore surprised to find that two copies of the *VIPR2* gene were present in GM10064. However, while one copy of the *VIPR2* gene was present on the normal chromosome 7 as expected, the other copy of the *VIPR2* gene was present on the otherwise apparently normal (paternal) 13p. Examination of cell line GM10067 which was derived from the father of GM10064, and is thought to contain a balanced translocation between chromosomes 7 and 13, verified that a copy of the *VIPR2* gene was indeed present at the distal end of the short arm of an otherwise normal chromosome 13, and showed that a total of three copies of the *VIPR2* gene were present in this cell line, one on the otherwise normal 13p, one on the normal chromosome 7, and one on the 13q derivative. The phenotype of this patient is normal (Gurrieri *et al.*, 1993). These results were confirmed by the use of chromosome 7 and chromosome 13 paints, which also revealed that although GM10064 contains two copies of the *VIPR2* gene, the majority of chromosome 13 is present in three copies in this cell line (GM10064 is shown in Figure 5.6; GM10067 is shown in Figure 5.7). Trisomy 13 is strongly associated with the occurrence of HPE, and it has been estimated that HPE occurs in 70% of trisomy 13 cases (Taylor, 1968). In contrast,

Figure 5.3: Chromosomal localisation of the *VIPR2* gene on metaphase chromosomes from cell line GM03240

(MC=microcephaly, MDF= mildly dismorphic facial features)



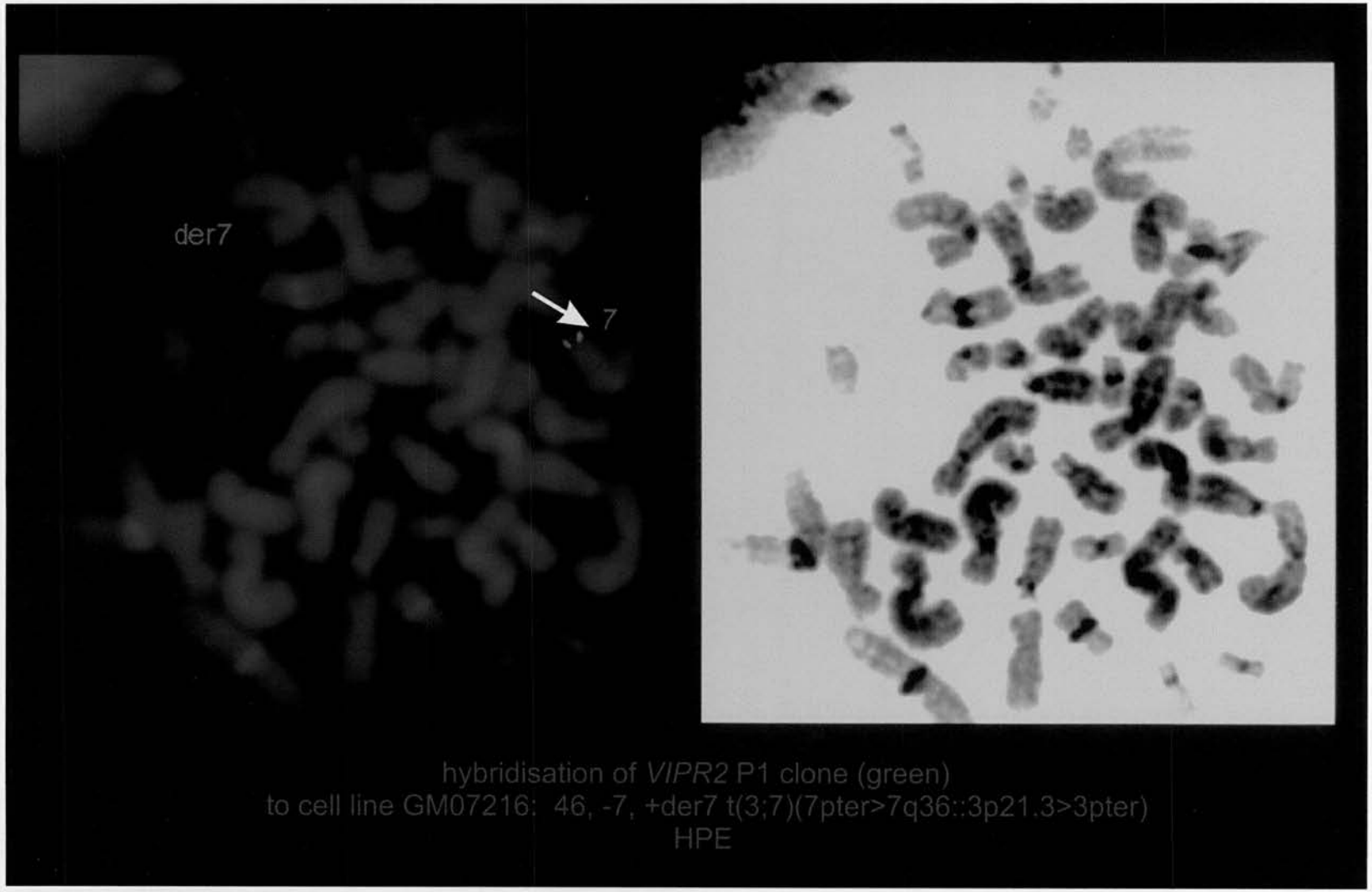
hybridisation of *VIPR2* P1 clone (green/turquoise)
to cell line GM03240 46,XY, del(q34-qter)
....MC, MDF

Figure 5.4: Chromosomal localisation of the *VIPR2* gene on metaphase chromosomes from cell line GM00657

(MC=microcephaly, MR=mental retardation, CP=cleft palate)

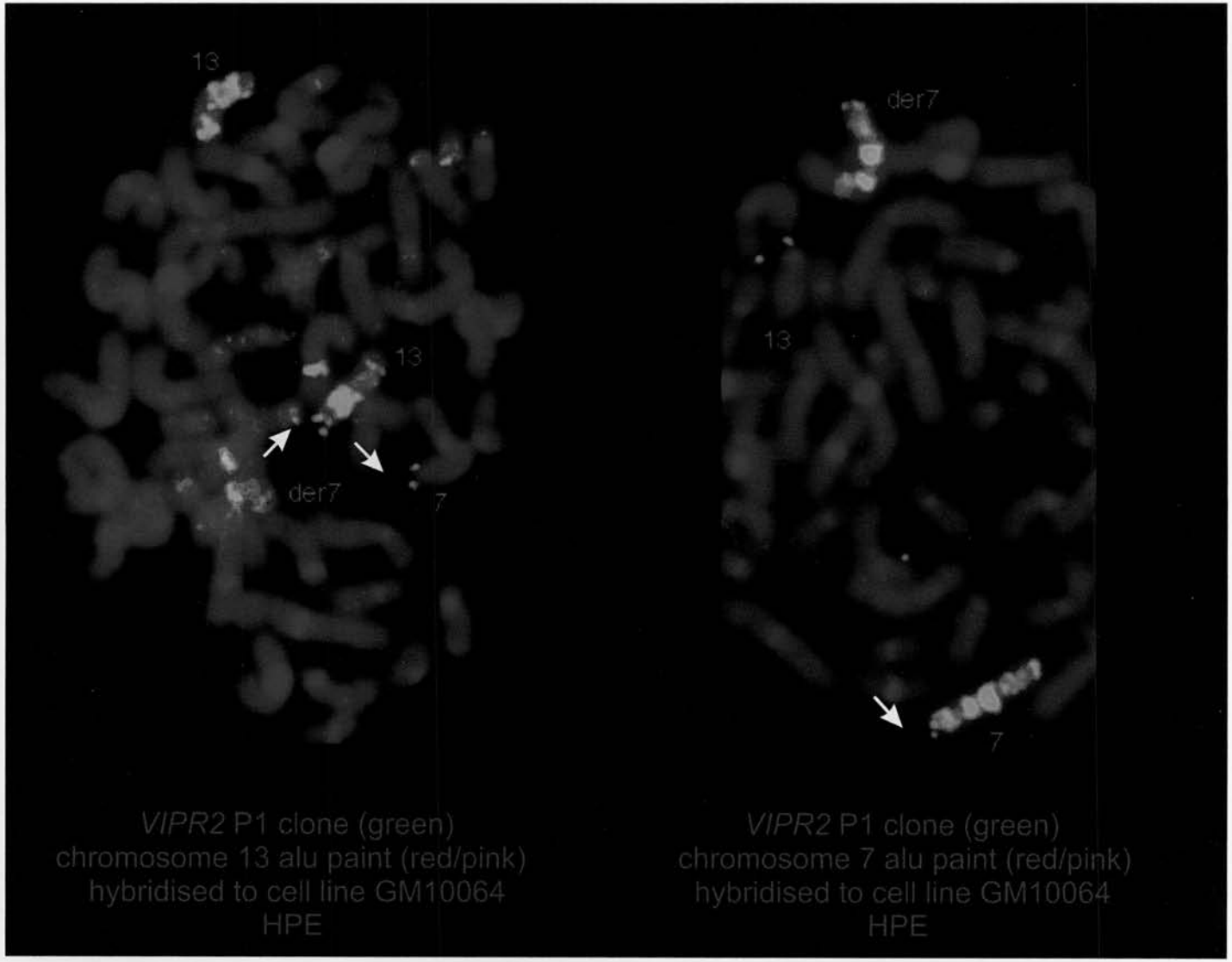
hybridisation of *VIPR2* P1 clone (green/turquoise)
to cell line GM00657: 46,XY, t(7;18)
(7pter>7q34::18q12.2>18qter; 18pter>18q12.2::7q36>7qter)
MC, MR, CP, small chin

Figure 5.5: Chromosomal localisation of the *VIPR2* gene on metaphase chromosomes from cell line GM07216



hybridisation of *VIPR2* P1 clone (green)
to cell line GM07216: 46, -7, +der7 t(3;7)(7pter>7q36::3p21.3>3pter)
HPE

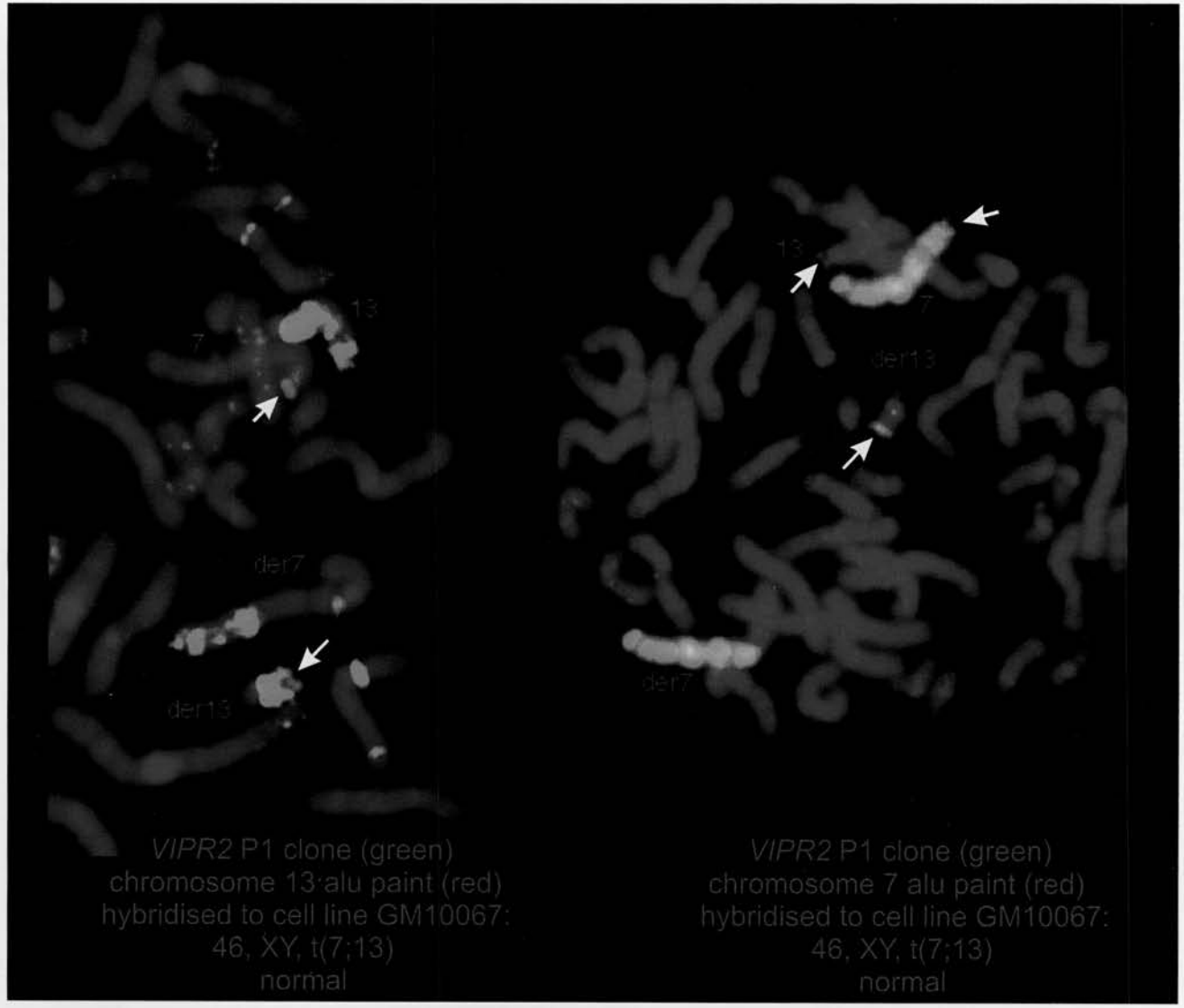
Figure 5.6: Chromosomal localisation of the *VIPR2* gene on metaphase chromosomes from cell line GM10064



VIPR2 P1 clone (green)
chromosome 13 alu paint (red/pink)
hybridised to cell line GM10064
HPE

VIPR2 P1 clone (green)
chromosome 7 alu paint (red/pink)
hybridised to cell line GM10064
HPE

Figure 5.7: Chromosomal localisation of the *VIPR2* gene on metaphase chromosomes from cell line GM10067



VIPR2 P1 clone (green)
chromosome 13 alu paint (red)
hybridised to cell line GM10067:
46, XY, t(7;13)
normal

VIPR2 P1 clone (green)
chromosome 7 alu paint (red)
hybridised to cell line GM10067:
46, XY, t(7;13)
normal

estimates of the occurrence of HPE in patients with deletions of 7q36 vary from ~30% (Grass *et al.*, 1995) to ~50% (M. Muenke, unpublished observation, cited in Overhauser *et al.*, 1995). Consequently, although either trisomy 13 or del (7q)(36) could have resulted in HPE in the patient from which cell line GM10064 was derived, the balance of probability suggests that it is more likely to have been caused by trisomy 13, and GM10064 was therefore excluded from our study of chromosome-7-linked HPE.

5.3.4 PCR analysis of a YAC contig spanning the *HPE3* minimal critical region

Having established that the *VIPR2* gene was located in the *HPE3* minimal critical region, we were interested in defining more precisely the position of the *VIPR2* gene within this region. In order to achieve this, a collaboration was established with Dr S. Scherer and colleagues at the Hospital for Sick Children in Toronto, Canada, who have been extensively involved in the construction of a genetic map of the long arm of human chromosome 7, and in the mapping of the *HPE3* locus. This group provided us with DNA samples from a contig of 15 YACs which together spanned the 2 Mb *HPE3* minimal critical region (Belloni *et al.*, 1995). However, PCR analysis of the YACs from this contig, using primers 4334 and 4335 which flank the last intron of the human *VIPR2* gene (see 5.2 and chapter 4), failed to identify any YACs that contained this region of the *VIPR2* gene. Nevertheless, the positive control reaction, to which 100ng of human genomic DNA had been added as a template, did contain a product that was the expected size (~1.3 kb). These results indicated that the *VIPR2* gene was not present within the YAC contig that spanned the *HPE3* minimal critical region, a result that appeared to conflict with those obtained from the FISH studies.

5.3.5 FISH localisation of YACs on interphase nuclei, and subsequent PCR analysis of a YAC contig from the telomeric region of 7q

The clue that was to eventually lead to the resolution of the apparent discrepancies between the FISH studies and our physical mapping data, came from a FISH study on interphase nuclei, in which the position of the *VIPR2* P1 probe was compared with those of YACs HSC7E207 and HSC7E526. Both of these YACs have previously been localised to 7q36 by Kunz *et al.* (1994), who also showed that HSC7E207 contains markers that are found centromeric to the proximal boundary of the *HPE3* minimal critical region, whereas HSC7E526 contains markers that place this YAC telomeric to the distal boundary of the *HPE3* minimal critical region as defined by Gurrieri *et al.*, (1993). Our results from FISH of these probes on interphase nuclei, suggested that a

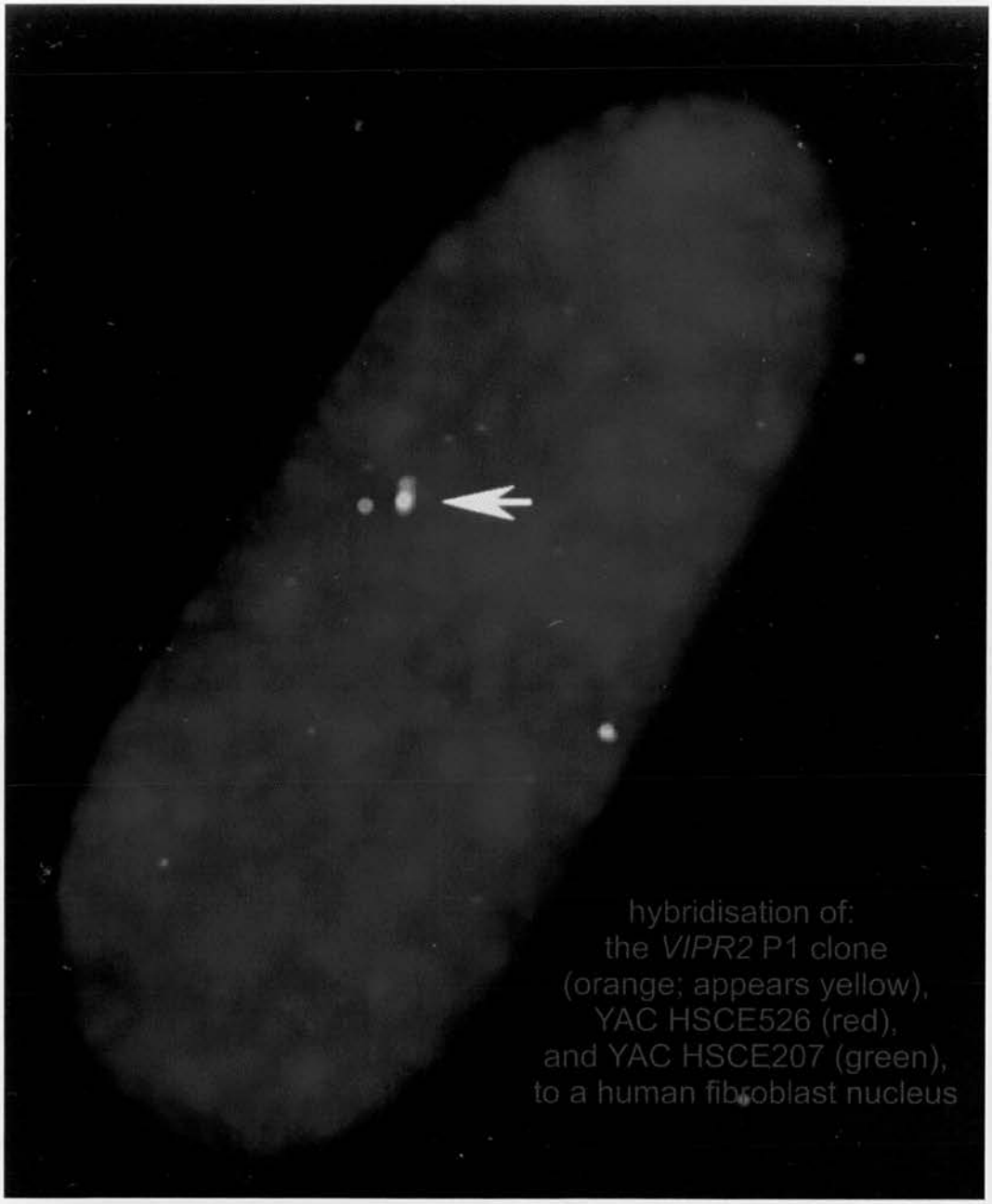
region of overlap was present between HSC7E526 and the *VIPR2* P1 probe, as indicated by the overlapping red coloured signal from HSC7E526 and the orange/yellow coloured signal from the *VIPR2* P1 probe (Figure 5.8). This result was confirmed by using PCR (with primers 4334 and 4335) to demonstrate that at least part of the human *VIPR2* gene was present within YAC HSC7E526. A product of the expected size (~1.3 kb) was produced from ~1 ng of HSC7E526 template DNA, and also from positive control reactions containing total human genomic DNA (100 ng and 300 ng) or DNA from a human chromosome 7 somatic hybrid cell line (Figure 5.9). The failure of the PCR reactions containing higher amounts of HSC7E526 template DNA (~2 ng and ~3 ng) was slightly surprising (Figure 5.9), however it was suspected that this may be due to the presence of contaminants in the mini-prep DNA that was used in these reactions. Further experiments using a different batch of YAC HSC7E526 DNA (provided by Dr S. Scherer) produced the expected DNA product at all concentrations tested.

HSC7E526 was originally isolated from a library of YACs whose average insert size is around 475 kb (Kunz *et al.*, 1994; Scherer *et al.*, 1992). Therefore, to localise the *VIPR2* gene within this region, a contig of 10 YACs (provided by Dr. S. Scherer) from the region to which HSC7E526 maps, was screened using the *VIPR2* gene PCR assay described above. A positive result (a product of ~1.3 kb) was obtained from one of the YACs in this contig, HSC7E1355 (Figures 5.10 and 5.11). YACs HSC7E526 and HSC7E1355 both contain the chromosome 7 marker *D7S593* which is located about 200 kb from the 7q telomere (S. Scherer, personal communication; Hing *et al.*, 1993). These results therefore suggest that the *VIPR2* gene lies close to *D7S593* and hence the 7q telomere.

5.3.6 Determination of the position and orientation of *VIPR2* relative to the 7q telomere.

Fine mapping of the position and orientation of the *VIPR2* gene was carried out by Dr S. Scherer and colleagues at the Hospital for Sick Children, Toronto, Canada. This group used our PCR assay for the 3' region of the *VIPR2* gene (see above), to demonstrate that in addition to HSC7E526 and HSC7E1355, two other YACs HSC7E1357 and E145A7 also contained the 3' region of the *VIPR2* gene. All of these YACs contain the marker *D7S593* (S. Scherer, personal communication) which was originally derived from the 7q telomeric YAC HTY146 (Hing *et al.*, 1993; Riethman *et al.*, 1993). Therefore it appeared that YACs HSC7E526, HSC7E1357 and E145A7 each overlap with YAC HTY146, thus generating a contig of YACs that extends inwards from the 7q telomere (see Figure 5.11).

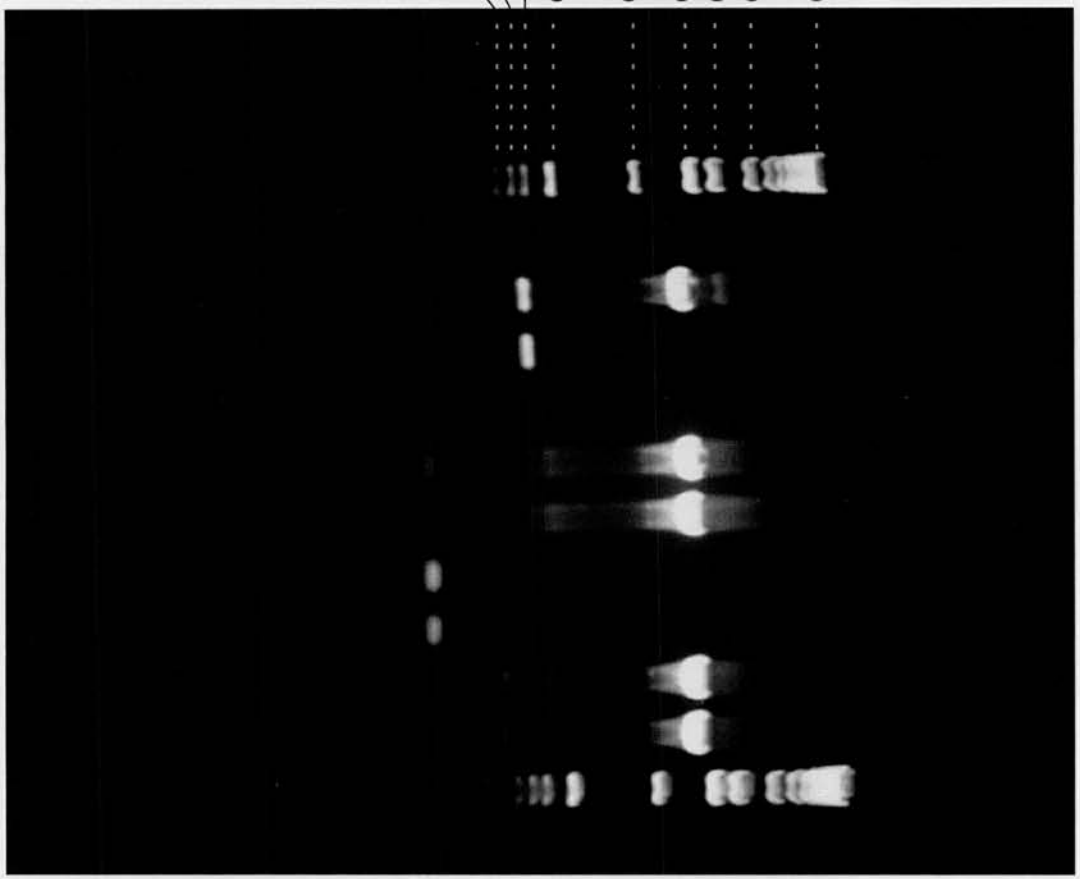
Figure 5.8: Localisation of the *VIPR2* gene relative to the positions of YACs HSC7E526 and HSC7E207 on a human fibroblast interphase nucleus



hybridisation of:
the *VIPR2* P1 clone
(orange; appears yellow),
YAC HSCE526 (red),
and YAC HSCE207 (green),
to a human fibroblast nucleus

Figure 5.9: PCR amplification of a 3' fragment of the *VIPR2* gene from YAC HSC7E526 and a human chromosome 7 somatic hybrid cell line

(Expected size of PCR product = 1.3 kb)



1 kb ladder

YAC E526 4334/4335

YAC E526 4334/4335

YAC E526 4334/4335

Chrom.7 somatic cell hybrid 4334/4335

Chrom.7 somatic cell hybrid 4334/4335

-ve control 4334/4335

-ve control 4334/4335

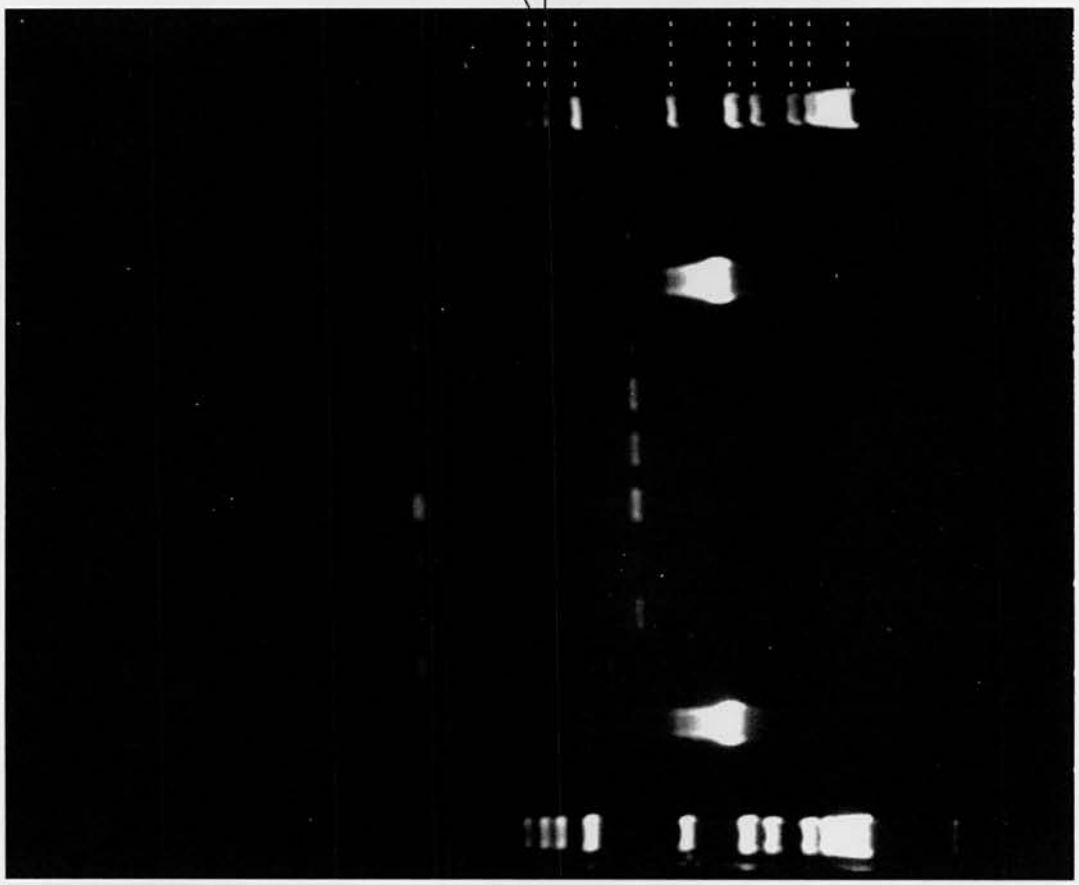
human genomic DNA (100 ng) 4334/4335

human genomic DNA (300 ng) 4334/4335

1 kb ladder

Figure 5.10: PCR amplification of a 3' fragment of the *VIPR2* gene from a contig of human chromosome 7-derived YACs surrounding HSC7E526

(YACs HSC7E1355 and HSC7E526 produced a product of the expected size, 1.3 kb)



1 kb ladder

YAC 1353 4334/4335

YAC 1355 4334/4335

YAC 1356 4334/4335

YAC 1369 4334/4335

YAC 1370 4334/4335

YAC 1372 4334/4335

YAC 1373 4334/4335

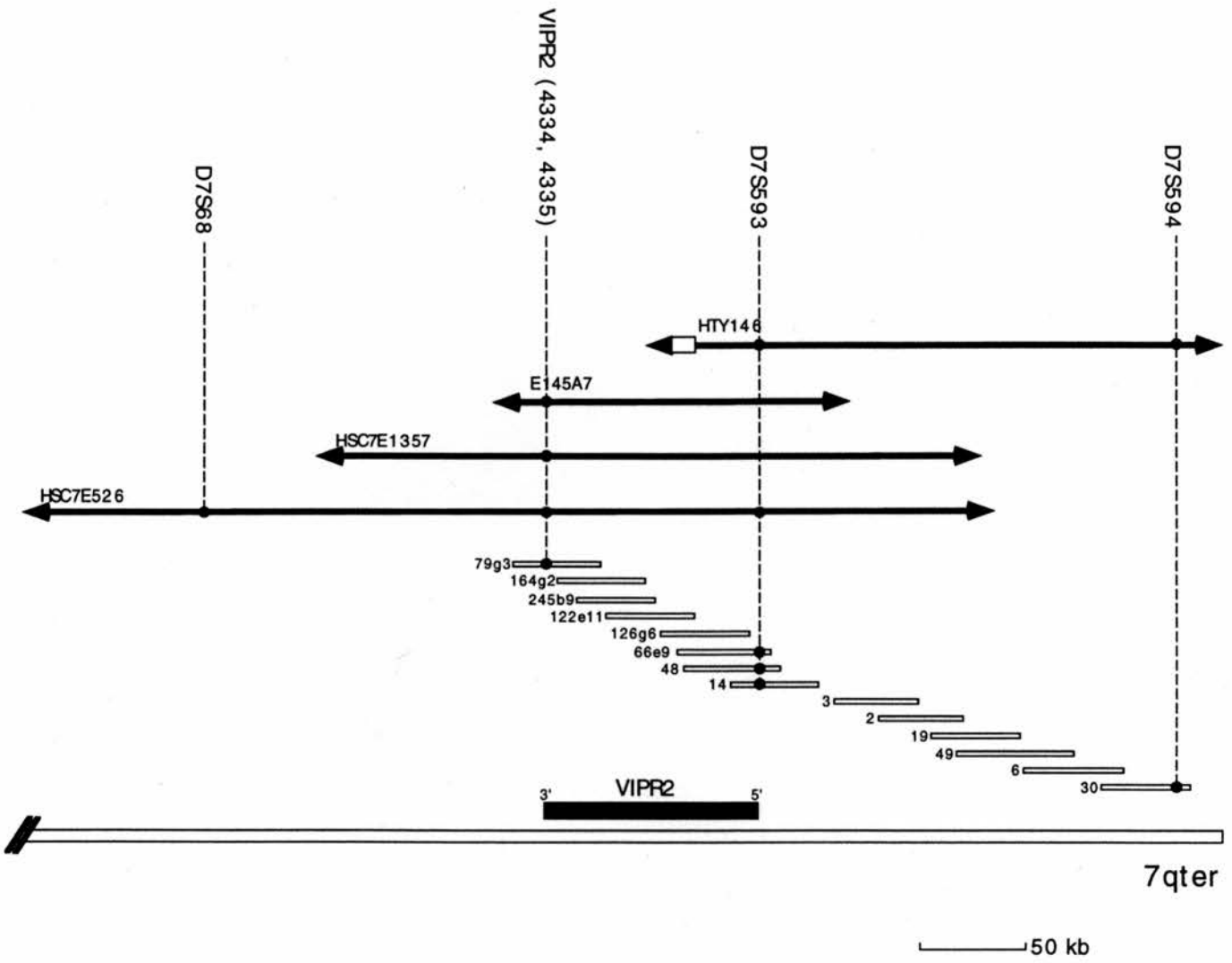
YAC 1374 4334/4335

YAC 1377 4334/4335

YAC E526 4334/4335

1 kb ladder

Figure 5.11: Schematic diagram showing the positions of 4 YAC clones and 14 cosmids relative to the position of the *VIPR2* gene and the telomere of human chromosome 7


D7S68
VIPR2 (4334, 4335)
D7S593
D7S594
HTY146
E145A7
HSC7E1357
HSC7E526
79g3
164g2
245b9
122e11
126g6
66e9
48
14
3
2
19
49
6
30
3'
VIPR2
5'
7qter
50 kb

YAC HTY146 has previously been subcloned into cosmids, and a contig of cosmid clones spanning the length of this YAC has been established (Riethman *et al.*, 1993). Eight of these cosmids, which together span the entire length of HTY146, were screened by hybridisation, for the presence of the *VIPR2* gene, using a human VIPR2 cDNA clone as the probe (see 5.2). Two positive clones were identified, cosmid 14, and cosmid 48 (*D7S593*), which based on the map generated by Riethman *et al.*, (1993), places part of the *VIPR2* gene ~200 kb from the 7q telomere. Hybridisation screening of a chromosome 7-specific cosmid library, again using the human VIPR2 cDNA as a probe, resulted in the isolation of a further 22 positive cosmid clones (8d1, 14e10, 66e9, 77e7, 79b6, 79g2, 79g3, 81e9, 107g9, 116h3, 122e11, 124g2, 126g6, 134a8, 135c5, 137g2, 164g2, 215a5, 222f1, 230e6, 245b9, and 250h9), and PCR analysis (using primers 4334 and 4335) indicated that four of these clones (8d1, 77e7, 79g3, and 137g2) contained the 3' end of the *VIPR2* gene. The relative positions of all 24 clones, and the degree of overlap between them, was determined by a combination of hybridisation analysis and restriction enzyme mapping. Figure 5.11 shows the positions, relative to the 7q telomere, of 8 representative cosmids that together span the entire known coding region of the *VIPR2* gene. The proximal locations of the cosmids that have been shown to contain the 3' end of the gene, in comparison with those of cosmids containing other regions of the gene, revealed that the orientation of the *VIPR2* gene along the chromosome is: *7cen-D7S68-(3'-VIPR2-5')-D7S593-D7S591-D7S592-D7S590-D7S594-7qtel.* Together, these results show that the known coding region of the *VIPR2* gene is encoded by ~100 kb of genomic DNA, and place the entire *VIPR2* gene within approximately 300-350 kb of the 7q telomere.

5.3.7 Reconciliation of the cytogenetic and physical mapping data: redefinition of the distal extent of the *HPE3* minimal critical region

Although our FISH results clearly showed that the cell lines used by Gurrieri *et al.*, to define the *HPE3* minimal critical region, were monosomic for the *VIPR2* gene, a major discrepancy existed between these results and those obtained from the physical mapping study. The FISH localisation of the *VIPR2* gene to the *HPE3* minimal critical region as defined by Gurrieri *et al.*, (1993) would be expected to place the *VIPR2* gene somewhere in the region between the chromosome 7 markers *D7S392* (which is just centromeric to the proximal boundary of the minimal critical region) and *D7S292* (which is located at the distal boundary of the minimal critical region). However, our physical mapping data suggest that the *VIPR2* gene lies close to the marker *D7S593*, which is at least 1.5 Mb telomeric to the proposed distal boundary of the *HPE3* minimal

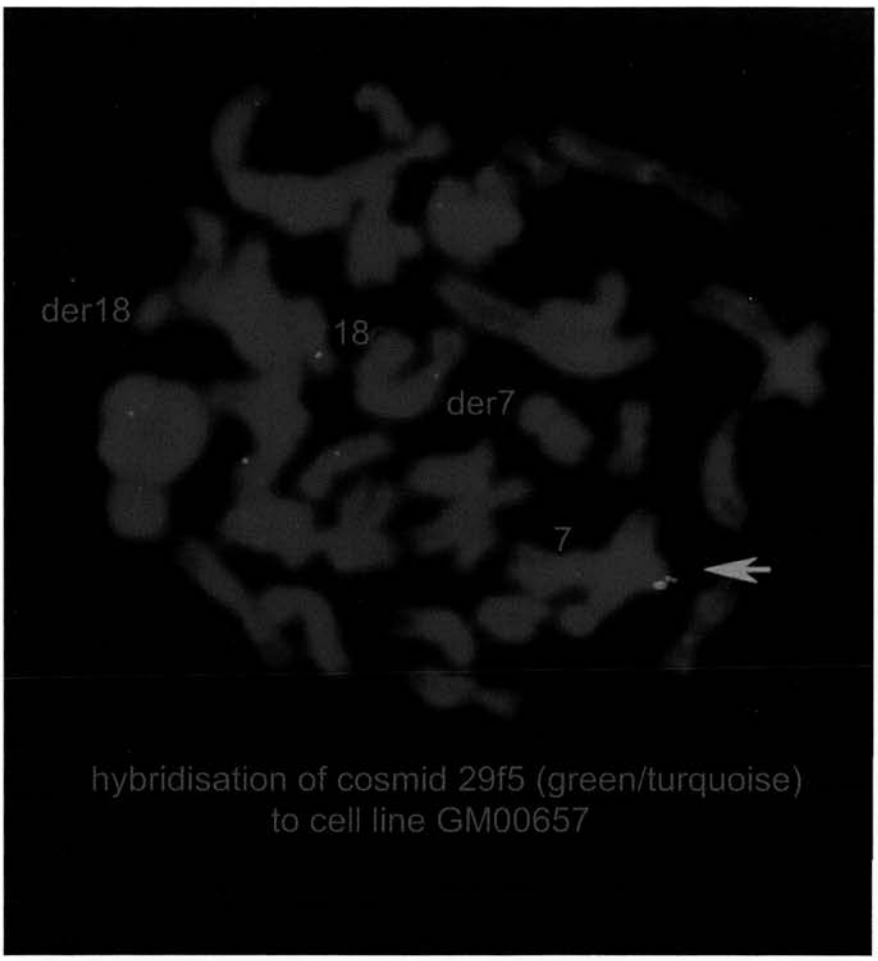
Figure 5.12: Map of the telomeric region of human chromosome 7

(Large dark arrows at the foot of the diagram represent the extent of the chromosomal deletion present in each of the cell lines that were studied, and the vertically positioned dotted line represents the position of a translocation breakpoint in a patient with a mild HPE phenotype)

The proximal deletion breakpoint in GM03240 defines the proximal boundary of the HPE3 critical region (Gurrieri *et al.*, 1993). The distal boundary of the HPE3 critical region is now defined as the 7q telomere.

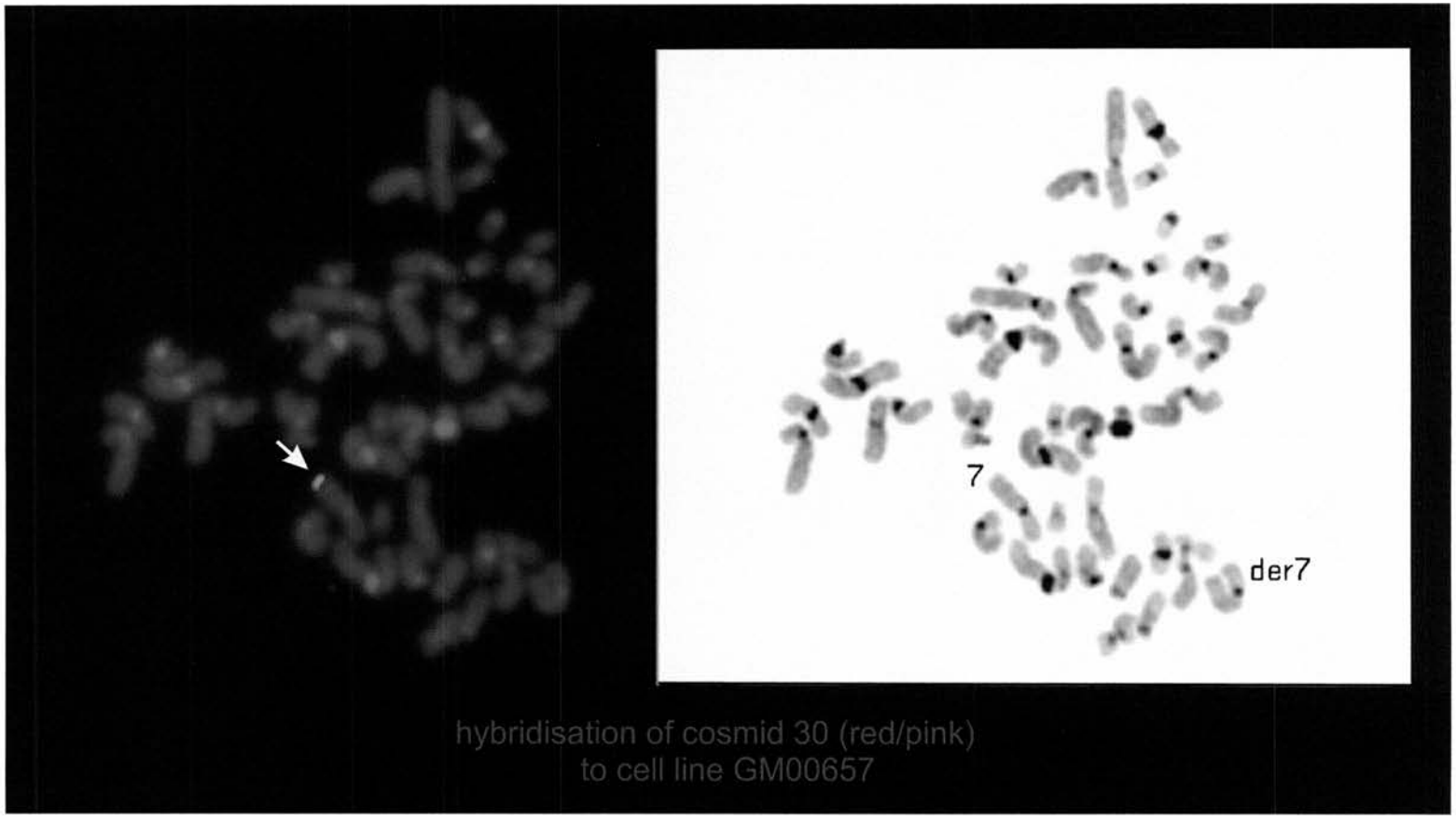
HPE deletions
GM03240
GM00657
GM07216
GM10064
YACs
cosmids
DNA Markers
C665a11
HSC7E207
HSC7E5
HSC7E70
HSC7E445
cos29f5
Scale (kb)
500
400
300
200
100
0
GAP
HSC7E1376
HSC7E526
HSC7E1357
E145A7
D7S427
cos30
cos6
cos3
cos48
cos126g6
cos122e11
cos79g3
D7S550
D7S392
D7S468
D7S104
D7S168
D7S292
D7S559
D7S22
D7S98
D7S68
VIPR2-3'
VIPR2-5'
D7S593
D7S591
D7S592
D7S590
7qter

Figure 5.13: Localisation of cosmid 29f5 on metaphase chromosomes from cell line GM00657



hybridisation of cosmid 29f5 (green/turquoise)
to cell line GM00657

Figure 5.14: Localisation of cosmid 30 on metaphase chromosomes from cell line GM00657


7
der7
hybridisation of cosmid 30 (red/pink)
to cell line GM00657

critical region (S. Scherer, unpublished data; Gurrieri *et al.*, 1993; see Figure 5.12). The most likely explanation for this discrepancy was that either the deletion in cell line GM100657 (used by Gurrieri *et al.*, (1993) to define the distal boundary of the *HPE3* minimal critical region) extended much further towards the 7q telomere than was previously thought, or that a second deletion existed in this cell line, distal to the deletion that had originally been characterised. Therefore, to examine more closely the extent of the deletion(s) present in GM100657, FISH studies on this cell line were carried out, using cosmids containing *D7S168* (cos29f5) which lies just proximal to the distal boundary of the *HPE3* minimal critical region, *VIPR2* (cos79g3 and cos126g6), and the telomeric marker *D7S594* (cos30) as probes (see Figure 5.12 for locations). Only one copy of each of these probes, was found to be present (on the normal chromosome 7) in cell line GM100657 (the results obtained with cos29f5 and cos30 are shown in Figures 5.13 and 5.14 respectively). The telomeric nature of the deletion in GM10064, indicated by the presence of only one copy of *D7S594* (cos30), was confirmed by using the 240 kb 7q telomeric YAC HTY146 (Riethman *et al.*, 1993) as a FISH probe on this cell line. As expected, only one copy of this YAC was present (on the normal chromosome 7) in GM100657. Thus, although it remains possible that a region of 7q36 between markers D7S168 and the *VIPR2* gene is present in cell line GM100657, our results clearly show that substantial regions telomeric to the proposed distal boundary of the *HPE3* minimal critical region (*D7S292*), including the *VIPR2* gene and the 7q telomere, are deleted in this cell line. These results redefine the *HPE3* minimal critical region, which now extends from *D7S392* to the telomere of 7q, and consequently identify the *VIPR2* gene as a positional candidate for the *HPE3* locus.

5.4 Discussion

5.4.1 Chromosomal location of the human *VIPR2* gene

The results shown in this chapter placed the 5' end of the human *VIPR2* gene within 200 kb of the human chromosome 7q telomere, and the whole gene within about 300-350 kb of the telomere. Currently, the localisation of the 5' end of the gene is based on the position of the 5' end of a human VIPR2 cDNA within a cosmid contig from the telomeric region, and it is not yet clear how much further towards the telomere the regulatory elements (e.g. promoter and enhancer elements) of the *VIPR2* gene lie. Genes that have been localised to 7q36 include: the 5-hydroxytryptamine receptor 5A gene (7q36.1) (Schanen *et al.*, 1996); the *GBX1* gene which encodes a homeobox-containing protein (7q36.1) (Matsui *et al.*, 1993); the endothelial nitric oxide synthase

gene (7q36) (Robinson *et al.*, 1994); the engrailed 2 gene (Poole *et al.*, 1989); the sonic hedgehog gene (7q36) (Marigo *et al.*, 1995); the aquaporin 3 gene (7q36.2-7q36.3) (Ishibashi *et al.*, 1995); the cyclin-dependent kinase 5 gene (7q36) (Demetrick *et al.*, 1994); and the *RHEB2* gene which encodes a Ras-related GTP-binding protein (7q36) (Mizuki *et al.*, 1996). However, of the genes localised to this region so far, the *VIPR2* gene is the closest to the telomere.

The 7q telomeric YAC HTY146 (also known as HTY3006) has been characterised and subcloned into cosmids by Riethman *et al.*, (1993) who have shown that the ~230 kb insert of this YAC contains all but the last 12 kb of the terminal 240 kb of 7q (Riethman *et al.*, 1993). Analysis of these clones indicates that in contrast to many human chromosome telomeres, the telomeric region of 7q contains single copy DNA sequences that are positioned very close to the telomere terminal repeat sequences, and an unmethylated CpG island has been identified which lies at a distance of about 195 kb from the end of 7q (Riethman *et al.*, 1993). This is almost exactly the position to which we have mapped the 5' end of the *VIPR2* gene (using the cDNA), and it seems very likely indeed that this unmethylated CpG island forms part of the 5' end of the human *VIPR2* gene.

Other interesting aspects of the region immediately surrounding the 5' end of the human *VIPR2* gene's transcribed sequence include: the marker *D7S593* (Hing *et al.*, 1993) for which at least 4 allele systems can be detected following restriction digestion with *Hind*III, three of which show a combined heterozygosity of 70% (Hing *et al.*, 1993); and the presence of an *L1*-type repetitive DNA element (Riethman *et al.*, 1993). An *L1* repetitive sequence was also found in the putative promoter region of the mouse *Vipr2* gene (see chapter 6), suggesting that sequence conservation between the mouse and human *VIPR2* gene may extend beyond the coding regions of the genes.

In addition to holoprosencephaly type 3 (Gurrieri *et al.*, 1993), the following human genetic diseases have been mapped to 7q35-36 or 7q36: long QT syndrome (cardiac arrhythmia, 7q35-36) (Jiang *et al.*, 1994) which is now known to be caused by mutations of the HERG gene (Curran *et al.*, 1995); various forms of polydactyly and polysyndactyly (7q36) (Heutink *et al.*, 1994; Hing *et al.*, 1995; Tsukurov *et al.*, 1994); and an autosomal dominant form of sacral agenesis (7q36) (Lynch *et al.*, 1995).

Deletions involving 7q have been found in many malignant tumours (Mertens *et al.*, 1997), particularly testicular, prostate, and ovarian tumours (Atkin & Baker, 1993), and in several tumour types (especially breast and lung carcinomas) an increase in deletion frequency has been observed towards the 7q telomere (Mertens *et al.*, 1997),

leading to the suggestion that an unknown tumour-suppressor gene may be located on distal 7q (Atkin & Baker, 1993; Mertens *et al.*, 1997). In contrast, isochromosome 7q is thought to be associated with hepatosplenic γ/δ T cell lymphoma (Jonveaux *et al.*, 1996), gains of the distal region of 7q were among the chromosomal imbalances noted in a study of primitive neuroectodermal tumours (Schütz *et al.*, 1996), and duplication of 7q has been found in about 50% of low-grade astrocytomas occurring in adults (Schrock *et al.*, 1996). This latter observation is of considerable interest with respect to the reported role of VIP and PACAP in the stimulation of astrocyte proliferation (Brenneman *et al.*, 1990; Tatsuno *et al.*, 1996). Nevertheless, despite the frequently reported stimulatory effects of VIP on neuronal growth and survival, it has become increasingly apparent that VIP acts as a modulator of cell growth and differentiation (Muller *et al.*, 1995). VIP has been shown to inhibit the growth of some human neuroblastoma cell lines, and to promote differentiation of these cells (Hoshino *et al.*, 1993; Pence & Shorter, 1993). VIP also inhibits mitogen- and antigen-induced T-cell proliferation (Metwali *et al.*, 1993), and both basal and histamine-stimulated proliferation of human airway smooth muscle cells (Maruno *et al.*, 1995). However, although the role of VIP receptors in the transduction of the growth inhibiting and differentiation promoting effects of VIP would not be inconsistent with the function(s) of a tumour suppressor gene, a strong candidate gene for the tumour supressor locus on distal 7q is the neural precursor cell-expressed developmentally down-regulated gene-2 (*NEDD2*) which is involved in the regulation of programmed cell death, and maps to 7q34-35 (Kumar *et al.*, 1995).

5.4.2 *VIPR2*: a possible contributor to the HPE3 phenotype

Analysis of the cell lines used by Gurrieri *et al.*., (1993) to define the *HPE3* minimal critical region, clearly showed that the human *VIPR2* gene is located within this minimal critical region. However, inconsistencies between the results that we obtained and those reported by Gurrieri *et al.*, (1993) have led us to redefine the *HPE3* minimal critical region, which we have shown to extend from the original proximal boundary to the 7q telomere.

Holoprosencephaly (HPE) is a developmental abnormality that results from incomplete cleavage of the forebrain at the stage when the prosencephalon would normally divide to form the cerebral hemispheres, diencephalon and telencephalon, and the olfactory and optic bulbs (reviewed in Siebert *et al.*, 1990). The HPE phenotype is characterised by midline structural defects of the brain and/or face, the extent of the underlying brain abnormalities usually, but not always, being reflected by the degree of facial malformation (DeMyer *et al.*, 1964). Alobar HPE, the most severe form of HPE, is

found when a complete failure of forebrain cleavage has occured (DeMyer & Zeman, 1963). In alobar HPE a single brain ventricle (or 'holosphere') is present with little or no evidence of interhemispheric fissure formation, the olfactory bulbs and tracts are frequently absent and the pituitary is often absent or hypoplastic. Facial abnormalities found in association with alobar HPE may include cyclopia or extreme orbital hypotelorism, proboscis formation and/or arhinia, absent nasal bones, or a single nostril, and median cleft lip. This form of HPE is not usually compatible with postnatal survival. Nevertheless, the severity of a given HPE phenotype is generally inversely related to its frequency of occurence, such that alobar HPE with cyclopia is a relatively rare finding. At the other end of the HPE phenotypic spectrum, are the HPE-like phenotypes which often include mild facial dysmorphism, and/or mental retardation in conjunction with a recognised HPE microform eg. microcephaly, hypotelorism, or a single frontal maxillary incisor. However, although they occur more frequently, these HPE-like phenotyes are often much more difficult to diagnose unambiguously, because many of these features are also associated with disorders other than HPE.

HPE has been estimated to occur at a frequency of 1/16000 live births and 1/250 embryos (Cohen, 1989). HPE can occur sporadically or in the form of familial HPE, and approximately 50% of cases are thought to be associated with chromosomal abnormalities (Muenke, 1994). However, although there are strong indications that HPE in humans can be caused by either genetic or environmental factors, the identification of environmental teratogens that may be responsible for HPE has proved to be difficult. To date, the only clear case for a link between an environmental factor and HPE in humans is the increase in the risk of HPE that is associated with maternal diabetes (Barr *et al.*, 1983), but HPE may also be linked to fetal alcohol syndrome (Coulter *et al.*, 1993), and maternal use of topical retinoids (Rosa *et al.*, 1994). Cytogenetic studies have identified several human chromosomal regions in which deletions or duplications appear to be non-randomly associated with HPE: del(2)(p21), dup(3)(p24-pter), del(7)(q36), del(18p), del(21)(q22.3), duplications or deletions involving 13q, trisomy 13, and trisomy 18 (reviewed by Muenke, 1994). As a result of these findings, four genetic loci at which genes responsible for HPE are thought to reside have been designated, *HPE1* at 21q22.3, *HPE2* at 2p21, *HPE3* at 7q36, *HPE4* in 18p, and minimal critical regions have now been defined for each of these four loci (Gurrieri *et al.*, 1993; Muenke *et al.*, 1995; Overhauser *et al.*, 1995; Schell *et al.*, 1996). However, the variability of the HPE phenotype, cannot be explained by the idea that different phenotypes may be linked with each of the *HPE* loci, as HPE cases resulting from deletions at any one of these four loci, display the same wide spectrum of phenotypes that is associated with HPE in general. This high degree of phenotypic

variability is exemplified by the familial autosomal dominant (AD) form of HPE, where even within a given AD HPE family, the phenotype displayed by affected family members can range from alobar HPE through to the apparently normal phenotype of some carriers of this form of the disease (Collins *et al.*, 1993; Muenke, 1994). Linkage studies have shown that in the majority of families with AD HPE, strong linkage can be demonstrated between this form of HPE and markers from the *HPE3* region on 7q36, suggesting that the *HPE3* gene may also be responsible for a large proportion of cases of AD HPE . The estimated risk for the occurence of HPE in the offspring of an obligate carrier of AD HPE is ~30-35%, indicating that this form of HPE displays incomplete penetrance (reviewed in Siebert *et al.*, 1990). Indeed, in contrast to the phenotypic spectrum of HPE which is wide, but does not seem to vary between loci, the degree of phenotypic penetrance that is observed for HPE, appears to be a locus specific phenomenon. For example, 30-50% of patients with deletions that encompass the *HPE3* minimal critical region, have some symptoms of HPE (Grass *et al.*, 1995; M. Muenke, unpublished observation cited in Overhauser *et al.*, 1995), a figure that is generally consistent with that for AD HPE (which is also thought to be linked to the *HPE3* locus). However, in contrast to this, it has been reported that all patients with deletions that span the *HPE2* minimal critical region on 2p21, have some form of HPE (Muenke, 1994), and only 10% of patients with deletions encompassing the *HPE4* minimal critical region on 18p show signs of HPE (Schinzel, 1984). Moreover, even when HPE does occur as a result of deletions in 18p or 7q36, the severity of the phenotype does not appear to correspond to the extent of the chromosomal deletion, and consequently it could be very difficult to distinguish between very mild forms of HPE and phenotypic abnormalities that might be associated with the deletion of other genes in the region.

The presence of two copies of the *VIPR2* gene (S. Scherer personal communication), in two unrelated patients with mild craniofacial HPE symptoms and balanced chromosomal translocations involving 7q36 (Belloni *et al.*, 1995), indicates that the *VIPR2* gene cannot be the sole gene responsible for the HPE3 phenotype. Nevertheless, one of these patients is the mother of four children, and it is intriguing to note that the two children who have a deletion of distal 7q resulting from the unbalanced segregation of the mother's chromosomal translocation (7;9) (q36;q34), have a more severe HPE phenotype than the two children who carry the same balanced translocation as the mother (Belloni *et al.*, 1995; Gurrieri *et al.*, 1993; Hatziioannou *et al.*, 1991). The association between a more severe phenotype and deletion of the previously translocated region of 7q in these cases could of course be pure chance. However, it does raise the possibility that one or more gene(s) lying distal to the

translocation breakpoint on 7q may be responsible for the increased severity of the phenotype observed when the region of 7q that lies distal to the breakpoint is lost.

Although research into the function, distribution, and signalling characteristics of the VIP_2 receptor is still in its early stages, several findings reported to date appear to support the idea that the *VIPR2* gene may contribute to the phenotype observed in some cases of HPE3. One of the first, and perhaps the most compelling piece of evidence, came from a study carried out by Gressens *et al.*, (1993) in which the effects of VIP on whole cultured mouse embryos were investigated. This study showed that treatment of explanted E9.5 mouse embryos with VIP results in a dramatic concentration dependent increase in embryonic growth, and at a concentration of 10^{-7} M VIP produces a 5-6 fold increase in cell mitosis in comparison with control embryos (Gressens *et al.*, 1993). The inverse experiment, in which pregnant mice were treated with a specific VIP antagonist (a neurotensin-VIP hybrid; Gozes *et al.*, 1991) from E9.3 to E11.8, resulted in embryonic growth retardation, and in particular, severe microcephaly (Gressens *et al.*, 1994). Microcephaly is thought to be one of the most reliable minor symptoms for the identification of possible carriers of autosomal dominant HPE (Collins *et al.*, 1993).

Within the developing CNS, the growth-promoting effects of VIP are thought to be mediated by a population of GTP-insensitive VIP receptors first described by Hill *et al.*, (1992). Comparison of the binding of ^{125}I-VIP to rat brain sections in the presence and absence of the GTP analogue guanylyl-imidodiphosphate (Gpp(NH)p), identifies two distinct populations of VIP receptors in the rat CNS, one GTP-insensitive and the other GTP-sensitive (Hill *et al.*, 1992). In mouse embryos at E9.5, the vast majority of VIP binding sites are confined to neural tissues and GTP-insensitive VIP binding sites are estimated to represent 73% of total VIP binding (Gressens *et al.*, 1993). The suggestion that the GTP-insensitive VIP receptors mediate the growth effects of VIP in the CNS of mouse embryos (Gressens *et al.*, 1993), stems from studies using the VIP antagonist neurotensin-$VIP_{(7-28)}$. This antagonist which partially blocks VIP-induced embryonic growth at E9.5 (Gressens *et al.*, 1993) especially that of the brain (E9.3-E11.8) (Gressens *et al.*, 1994), and completely blocks VIP stimulation of mitosis in the CNS at E9.5 (Gressens *et al.*, 1993), appears to induce a selective increase in the density of GTP-insensitive VIP receptor sites in mouse embryos but has no effect on the density of GTP-sensitive VIP binding sites (Gressens *et al.*, 1993).

Vipr1 and *Vipr2* mRNAs can be detected in the anterior (head) segment of mouse embryos at E9.5 (Sheward *et al.*, 1996b), and are thus present at the developmental

stage where VIP is thought to exert its greatest effects on growth in the embryonic CNS. To date, no reports have been published on the detailed distribution of *Vipr1* or *Vipr2* mRNA within the embryonic brain, or on the early embryonic distribution of the GTP-insensitive VIP receptors. Nevertheless, the adult distribution of the GTP-insensitive VIP receptors in the rat CNS has been studied (Hill *et al.*, 1994; Hill *et al.*, 1992) and appears to be similar but not identical to the pattern that has been reported for *Vipr2* mRNA (Lutz *et al.*, 1993; Sheward *et al.*, 1995; Usdin *et al.*, 1994). Both are found at relatively high levels in the olfactory bulb and pituitary (structures that can be severely affected by HPE), the dorsomedial nucleus of the hypothalamus, several thalamic nuclei (especially the medial geniculate nucleus), and the superior colliculi. Furthermore, from E14-E16 (the earliest stages examined) in rat embryos GTP-insensitive VIP binding is reported to be dense and restricted to a few midline structures of the brainstem and spinal cord, including the roof and floor plates (Hill *et al.*, 1994). The relationship between the GTP-sensitive/GTP-insensitive VIP receptor populations and the VIP_1receptor and VIP_2 receptor is unclear, and will be discussed in more detail later. However, considering the intrinsic association between midline structure dysgenesis and HPE, it will be important to establish the timepoint at which VIP receptors first appear in these structures.

PACAP and PACAP binding sites can be detected in the brain of rat embryos from E14 onwards by radioimmunoassay and radioreceptor assay respectively (Tatsuno *et al.*, 1994), and *Pacapr1* mRNA has been shown to be present in mouse embryos from E9.5 onwards (Sheward *et al.*, 1996a). Unlike the VIP_1 receptor and VIP_2 receptor, which bind both VIP and PACAP with high affinity, the PACAP type I receptor binds PACAP with much higher affinity than VIP (see chapter 1). The extremely high level of *Pacapr1* mRNA found in mouse embryos, relative to those encoding the VIP_1 receptor and VIP_2 receptor (Sheward *et al.*, 1996a), suggest that a high proportion of the PACAP binding observed during studies on rodent embryos is likely to represent binding to PACAP type 1 receptors, and a significant amount of low affinity VIP binding may also comprise binding to PACAP type 1 receptors. The neurotrophic properties of PACAP (Arimura *et al.*, 1994), together with the observation that differential expression of alternately spliced forms *Pacapr1* mRNAs occurs during development of the rat brain (D'Agata *et al.*, 1996), suggest that PACAP may play an important role in the development of the CNS. As noted earlier, the human *PACAP* gene maps to the chromosomal region 18p11 (Hosoya *et al.*, 1992), and its localisation has been refined to 18p11.32 by Chang et al., (1993)(Chang *et al.*, 1993). Deletions of 18p have been associated with HPE, and a minimal critical region for *HPE4* on 18p11.3 has recently been defined (Overhauser *et al.*, 1995). The *PACAP* gene maps

within this minimal critical region. However, although the authors note that the *PACAP* gene and three other genes lie within the minimal critical region they state that "the known expression pattern and function of each of these genes does not implicate them as likely candidates for *HPE4*" (Overhauser *et al.*, 1995). We believe that the evidence presented above provides grounds for the consideration of the *PACAP* and *VIPR2* genes as possible contributors to the HPE phenotype. Although the role of PACAP in neuronal development is currently less well defined than that of VIP, PACAP is also indirectly implicated in the effects of VIP by its ability to compete with VIP for high affinity binding sites on both the VIP_1 receptor and VIP_2 receptor. Indeed, the observations that: 10% of 18p deletion cases, 50% of 7q36 deletion cases, and all cases of 2p21 deletion, have symptoms of HPE (M. Muenke unpublished observation reported in Overhauser *et al.*, 1995) could be interpreted as supporting this type of interaction between the genes at *HPE* loci. Unfortunately, the abundance of the PACAP type 1 receptor makes it unlikely that the effects of PACAP binding to VIP_1 receptors and VIP_2 receptors during embryogenesis will be clarified until suitable specific antagonists are found, or the results from transgenic/gene knockout studies of the receptor genes are available.

Although VIP is readily detectable in rodent embryos from at least as early as E11 (Hill *et al.*, 1996; Waschek *et al.*, 1996), controversy exists over its source. The basis of the problem, is the apparent absence of *VIP* mRNA in the rat embryo during early embryogenesis, despite the presence of both VIP and VIP receptors at this time. It has been reported that *VIP* mRNA cannot be detected in the rat CNS prior to birth, and cannot be found in any part of the rat embryo before E16 (Hill *et al.*, 1994). This has led to the suggestion that embryonic VIP may originate either from the placenta or the mother (Hill *et al.*, 1994). Subsequent studies have shown that a peak in the VIP content of the rat embryo at E11 (with VIP levels four times higher than those at E17) coincides with a peak in the VIP levels of maternal serum, and the transfer of radiolabelled VIP from maternal blood to the embryo has also been demonstrated (Hill *et al.*, 1996). *VIP* mRNA could not be detected in the placentas of rat embryos by in situ hybridisation, but was detected in E15-E18 placentas by RT-PCR (Hill *et al.*, 1996). However, another group has recently reported that *VIP* mRNA is present in the CNS of mouse embryos at E11 (Waschek *et al.*, 1996). As noted by Waschek et al, possible reasons for the difference between their results and those of Hill et al., include species-specific variations in the levels of *VIP* mRNA in embryos, or differences in the sensitivity of the assays used by the two groups (Waschek *et al.*, 1996). The possibility of developmental differences between rat and mouse makes it difficult to draw any firm conclusions from these experiments at present. The finding that *VIP*

mRNA is present in mouse embryos, is very exciting, as it may lead to the resolution of a longstanding problem in this area of research, especially if the methodology is successfully applied to rat embryos. Nevertheless, the coincident peak of VIP levels in rat embryos and maternal serum at E11 is also interesting, and may represent a potential mechanism by which maternal VIP could augment embryonic VIP levels during critical stages in embryonic development.

The effects of maternal factors on the developing embryo are of particular interest in the study of specific congenital malformations that are thought to occur at a higher than normal frequency in the offspring of mothers who have insulin-dependent diabetes (IDD). HPE has been reported to display a 200 fold increase in frequency in infants of mothers who have IDD and thus belongs to this group of malformations (Barr *et al.*, 1983), which also includes neural tube defects, sacral agenesis (also known as caudal dysplasia), and congenital heart abnormalities (Cousins, 1983; Mills *et al.*, 1979). However, although the link between maternal IDD and specific congenital malformations in infants seems to be widely acknowledged, it is not universally accepted, and it has been suggested that the retrospective nature of many of the studies has led to a severe bias in the data (Kalter, 1993). While debate over the existence of a specific diabetic embryopathy continues, a more general link between maternal IDD and developmetal abnormalities in infants is not in question. Experiments in which E10.5 rat embryos are grown *in vitro* using serum obtained from women who have type I diabetes result in a 71% incidence of embryonic abnormalities in comparison with a 9% incidence of abnormalities in embryos cultured in serum that was obtained from non-diabetic women (Ornoy *et al.*, 1995). Serum from diabetic women is also found to decrease the size of embryos and to reduce the number of somites present relative to those of control embryos (Ornoy *et al.*, 1995), an effect that is the opposite of that observed following administration of VIP to E9.5 mouse embryos grown *in vitro* (Gressens *et al.*, 1993).

The hyperglycemic environment that results from streptozotocin (STZ) induced diabetes in mice, has been shown to result in a two-fold increase in the frequency of mutation of a *lacI* transgene in mouse embryos, suggesting that a general increase in the level of DNA mutations may occur in embryos of diabetic mothers (Lee *et al.*, 1995a). However, there is also evidence to support the idea that alterations in the expression levels of specific genes may also be involved in diabetic embryopathy. A recent study on the effects of STZ induced maternal diabetes on the embryonic expression of mRNAs for insulin-like growth factors (IGFs) and insulin-like growth factor binding proteins (IGFBPs) in rat, showed that there is a specific increase in the level of

IGFBP-1 mRNA in E14 embryos from diabetic mothers, and has led to the suggestion that elevated levels of IGFBP-1 may be involved in the growth retardation that occurs in the embryos of diabetic mothers (Streck *et al.*, 1995).

The *VIPR2* gene is expressed at high levels in the insulin-secreting pancreatic β-cell line MIN6 from which the mouse Vipr2 cDNA was cloned, and is also expressed in the insulin-secreting cell lines HIT-T15 and RINm5F (Inagaki *et al.*, 1994). Both PACAP and VIP have been shown to amplify glucose-stimulated insulin-secretion *in vitro* (Bertrand *et al.*, 1996; Straub & Sharp, 1996), and the pharmacological characteristics of this response (ineffectiveness of secretin but approximately equipotent effects of VIP and PACAP at 10^{-11}-10^{-8} M) strongly suggest that it is mediated by the actions of VIP and PACAP on the VIP_2 receptor (Bertrand *et al.*, 1996). The involvement of VIP, PACAP, and the VIP_2 receptor, in the control of insulin secretion, indicates that the expression levels of the genes encoding these products may be modified in the diabetic state.

5.4.3 Sacral agenesis

Although sacral agenesis and HPE are midline developmental defects that affect opposite ends of the developing neural tube, the frequency at which these abnormalities are found to coexist suggests that their association is not random (Martínez-Frías *et al.*, 1994), and has led to the proposal that they may result from similar deficiencies in the early development of the neural tube. Sacral agenesis and HPE display similar increases in incidence (~200 fold) in embryos of diabetic mothers (Barr *et al.*, 1983), and both developmental defects have been associated with cytogenetic abnormalities involving trisomy 13, trisomy 18, or deletion of the terminal region of 7q (Martínez-Frías *et al.*, 1994; Morichon-Delvallez *et al.*, 1993). The basis of at least one of these associations has been partially explained by the recent finding that a gene responsible for the autosomal dominant form of sacral agenesis, maps to the holoprosencephaly region on 7q36 (maximum lod scores are obtained with *D7S559* and *D7S594*) (Lynch *et al.*, 1995). However it is not yet clear whether or not the same gene is responsible for both defects.

5.4.4 Chromosomal location of the mouse *Vipr2* gene

The mouse *Vipr2* gene was localised to mouse chromosome 12 band F2, a result that was unexpected with regard to previously defined regions of synteny between the mouse and human genomes.

For several years the only indication of synteny between human chromosome 7 and mouse chromosome 12 has been the laminin B1 gene. The human laminin B1 gene (*LAMB1*) has been mapped to human chromosome 7q22-q31 (Modi *et al.*, 1987; Pikkarainen *et al.*, 1987), and its mouse homologue *lamb1-1*, has been localised to the proximal region of mouse chromosome 12 (O'Brien *et al.*, 1993). In the absence of any other indications of synteny between these two chromosomes, doubts have been expressed over the validity of this single syntenic marker, particularly with regard to the functional status of the mouse *lamb1-1* gene (Khan et al., 1994). Nevertheless, a recent report that the dihydrolipoamide dehydrogenase gene (*DLD*) which is located on human chromosome 7q31-32 (Scherer *et al.*, 1991) has a homologue on mouse chromosome 12 (Tsui *et al.*, 1995), together with the finding that the mouse homeobox gene *Meox2* which maps to the proximal region of mouse chromosome 12 (Candia *et al.*, 1992) has a human homologue on 7p22.1-p21.3 (Grigoriou *et al.*, 1995), strongly suggests that some conservation of synteny does exist between human chromosome 7 and mouse chromosome 12.

Human chromosomal region 7q36 is generally considered to be syntenic with the proximal region of mouse chromosome 5 (Tsui *et al.*, 1995), and to date at least four human genes from 7q36 have been shown to have homologues on mouse chromosome 5 (Marigo *et al.*, 1995; Martin *et al.*, 1990; Matthes *et al.*, 1993; Ohshima *et al.*, 1995). However, as the *VIPR2* gene appears to be the most distally located gene that has been mapped to 7q36 so far, the possibility exists that a chromosomal region showing synteny with the telomeric region of mouse chromosome 12 may be located distal to the recognised region of synteny with mouse chromosome 5.

At present, a large section of the distal region of mouse chromosome 12 bordered proximally by the paired box homeotic gene 9 (*Pax9*) locus and distally by the immunoglobulin heavy chain-C (*Igh-C*) locus, is believed to be homologous in both its gene content and gene order, to the long arm of human chromosome 14 (D'Eustachio & Riblet, 1996). However, since both *Igh-C* (D'Eustachio & Riblet, 1996) and *Vipr2* map to 12F2, we do not know whether the mouse *Vipr2* gene lies within or distal to the defined region of synteny with human chromosome 14q, and consequently there are three possible explanations for our findings: i) the mouse *Vipr2* gene lies distal to the defined region of synteny with human chromosome 14q, and the subtelomeric extremes of human chromosomal region 7q36 and mouse chromosome 12 could form a previously unrecognised region of synteny, ii) the mouse *Vipr2* gene lies within the previously defined region of synteny with human chromosome 14q, and other markers from the distal region of 7q36 also have mouse homologues within the same region of

mouse chromosome 12, thus dividing the region that is syntenic with human 14q, iii) the mouse *Vipr2* gene may be the only gene from 7q36 that maps to the distal region of mouse chromosome 12, and for some reason, its subtelomeric location has been preserved at the expense of conservation of regions of synteny between the two species.

Although further work will be required in order to clarify the relationship between the subtelomeric region of human chromosome 7q36 and the distal region of mouse chromosome 12, the localisation of the mouse *Vipr2* gene to chromosome 12F2 is likely to be of considerable interest with regard to mouse/human comparative genomic mapping, and may define a new region of synteny between the mouse and human genomes.

Since we suspect that monosomy at the human *VIPR2* locus may contribute to HPE3 phenotype, it was possible that a similar phenotype might be associated with the locus of the mouse *Vipr2* gene. However, a search of the mouse genome databases did not reveal the existence of any mapped mouse genetic mutations that result in an HPE-like phenotype. This is not as suprising as it might seem, and may be a reflection of developmental differences between mice and humans. A review by Searle *et al.*, indicates that a corresponding pathological variant in the mouse, is found for only about 20% of the genes that have been shown to be responsible for genetic disease in humans (Searle *et al.*, 1994). Moreover, in the case of genes responsible for human craniofacial abnormalities, the percentage of conditions for which a spontaneously occurring mouse homologue of the defect occurs, appears to be even lower (~10%), although it has been suggested that this may be due to difficulties in recognising some of these conditions in mice (Winter, 1996).

HPE can be experimentally induced in mouse embryos by ochratoxin A (Wei & Sulik, 1993), alcohol (Sulik & Johnston, 1982), and all-trans retinoic acid (Sulik *et al.*, 1995), each of which can cause a wide spectrum of developmental abnormalities, the type and severity being dependent to a large extent on the dose and the gestational stage at which the teratogen is administered (Sulik *et al.*, 1995; Sulik & Johnston, 1982; Wei & Sulik, 1993). A study carried out by Wei and Sulik (1993) on the effects of ochratoxin A is particularly interesting, as in addition to showing a correlation between the gestational day of toxin administration (GD7 or GD8) and variations in the HPE-like phenotype of the embryos, parallels have also been drawn between the patterns of ochratoxin A-induced cell death in the developing forebrain, and the resultant phenotype (Wei & Sulik, 1993). However, the effects of these teratogens at the level of gene expression are not fully understood, and the extent to which the patterns

observed in these models reflects the pathogenesis of chromosomal HPE in humans is unknown.

The possibility of examining the effects of monosomy at the mouse *Vipr2* locus, through the use of mice that carry translocations of the distal region of chromosome 12, was considered, as several such mouse lines are available. Similar experiments have been carried out previously by Beechey *et al.*, (1980), who used mice carrying the reciprocal translocation T(5;12)31H which has breakpoints at bands 5B and 12F1 (just distal to the immunoglobulin heavy chain-1 locus, *Igh-1*), to generate mice that are monosomic (5, 12, 12^{5}) for the distal region of chromosome 12, and mice that are trisomic for this region (5, 5, 12, 12, 5^{12}). Interestingly, both the trisomic and the monosomic mice show growth retardation, and some of the monosomic mice have abnormal skull and vertebral fusions (Beechey *et al.*, 1980). However, by definition, these mice are also either monosomic or trisomic for regions of chromosome 5, thus making it impossible to reach any clear conclusion about the cause of the observed phenotype. Moreover, even from the relatively few genes that have been mapped to the telomeric region of mouse chromosome 12 to date, there are at least two genes that are thought to have roles in craniofacial development. The *goosecoid* gene (*Gsc*), which is essential for normal craniofacial development (Rivera-Pérez *et al.*, 1995; Yamada *et al.*, 1995a), and maps slightly proximal to *Igh-1* (D'Eustachio & Riblet, 1996; Guénet *et al.*, 1995), is probably not involved in the T(5;12)31H translocation; but the *legless* (*lgl*) insertional mutation, which is situated near to the *situs inversus viscerum* (*iv*) locus (Singh *et al.*, 1991), distal to *Igh-1* (D'Eustachio & Riblet, 1996), almost certainly lies within the T(5;12)31H translocated region. The *legless* mutation (which is thought to disrupt more than one gene) is particularly interesting, as in addition to situs inversus and limb abnormalities, the phenotypes observed in homozygous mutants include: absent olfactory bulbs, malformed cerebral hemispheres, hydrocephaly, and in ~50% of cases craniofacial clefts (Singh *et al.*, 1991), characteristics that correspond to some of those seen in deletions of human chromosomal region 7q36. Mice that are heterozygous for either the *legless* insertional mutation or the *Gsc*-null mutation, are phenotypically normal, indicating that monosomy at these loci is unlikely to be responsible for the effects seen in the mice characterised by Beechey et.al., (1980). However, even in the absence of the potential complications caused by these two known genes, the difficulties associated with interpreting the information gained from the study of mice derived from the T(5;12)31H translocation line (Beechey *et al.*, 1980), suggested that this type of approach would be of limited use for the study of VIP_2 receptor function, at present.

5.4.5 Although initially excluded as a candidate, the*Sonic Hedgehog* gene has now been shown to be the *HPE3* gene.

As discussed previously, the presence of the *VIPR2* gene in two unrelated patients who carry balanced translocations of 7q, and have mild HPE phenotypes, suggests that the *VIPR2* gene cannot be the sole gene responsible for HPE3, although it may contribute to the phenotype observed in some cases.

The combination of the high gene density that appears to be associated with the subtelomeric regions of most human chromosomes (Craig & Bickmore, 1994), and the large number of genes that are already known to be involved in brain and craniofacial development, indicated that the *HPE3* minimal critical region on 7q36 was likely to contain several genes that would be strong candidates for the *HPE3* locus. During our studies of the *VIPR2* gene as a candidate gene for the *HPE3* locus, two other candidate gene's were identified by other laboratories (Marigo *et al.*, 1995; Mizuki *et al.*, 1996):

i) The Sonic hedgehog gene (*SHH*), was mapped to 7q36, by Margio *et al.*, soon after we began our cytogenetic studies of the *VIPR2* gene. However, although the authors initially examined the *SHH* gene as a potential candidate for the polysyndactyly locus that is situated in this region (Marigo *et al.*, 1995), a possible role for this gene in HPE was not suggested at this stage.

ii) The Ras homologue enriched in brain 2 gene (*RHEB-2*) was mapped to 7q36 by Mizuki *et al.*, shortly after the completion of our studies of the *VIPR2* gene, and was suggested as a candidate gene for holoprosencephaly type 3 and/or sacral agenesis (Mizuki *et al.*, 1996), although no cytogenetic mapping studies of the gene in cell-lines derived from individuals with HPE had been carried out.

Of these two positional candidates, the *SHH* gene was the most promising. *SHH* encodes a secreted signalling molecule and is a member of a family of vertebrate genes that are related to the *Drosophila hedgehog* (*hh*) gene (Echelard *et al.*, 1993). So far, at least seven vertebrate *hedgehog* genes have been identified (including 3 in mammals), and, like their *Drosophila* counterpart, these genes appear to play an important role in patterning activities during embryonic development (reviewed by Hammerschmidt *et al.*, 1997). The expression pattern of the *Shh* gene is generally consistent with a role for this gene in HPE, as *Shh* is expressed in midline structures including the notochord and floorplate during embryogenesis, and has been implicated in the control of patterning during embryonic limb, brain, and eye development (Chang *et al.*, 1994; Ekker *et al.*, 1995; Macdonald *et al.*, 1995). Ectopic expression of *Shh* in zebrafish

embryos results in eye defects and absence of the midbrain- hindbrain-boundary constriction, whereas in homozygous *cyclops* mutant (*cyc*$^{-/-}$) embryos the ventral neural structures that usually express *Shh* are absent (Ekker *et al.*, 1995). Linkage studies of the *SHH* gene suggested that the gene is located distal to *D7S550* (see Figure 5.12 for location of this marker) and close to the *EN2* gene (Marigo *et al.*, 1995), a result which would probably place the *SHH* gene within or slightly centromeric to the proximal boundary of the *HPE3* minimal critical region (Gurrieri *et al.*, 1993). Taken together, these linkage studies and expression data suggested that *SHH* would be a strong candidate for *HPE3*, and we were therefore very surprised to read that the *SHH* gene had been excluded as a candidate for the *HPE3* locus (M. Muenke, personal communication cited in Winter, 1996).

Indeed, the exclusion of *SHH* as a candidate for *HPE3* became even more puzzling, when it was reported that the phenotype observed in *Shh*$^{-/-}$ mice, includes midbrain and forebrain defects, cyclopia, a proboscis-like extension from the rostral midline, severe growth retardation, and severe craniofacial malformations (Chiang *et al.*, 1996). These characteristics are very similar to those found in patients who have HPE, and consequently, it was not unexpected when, in November 1996, evidence indicating that *SHH* is the primary gene responsible for HPE, was finally published (Belloni *et al.*, 1996; Roessler *et al.*, 1996). Following an extensive search for candidate genes in the region spanning the translocation breakpoints of patients who have balanced translocations involving 7q36, and have mild HPE phenotypes, the *SHH* gene (which was found to lie 15 kb from the closest HPE-associated translocation breakpoint and 250 kb from another HPE-associated translocation breakpoint) was found to be one of only two candidate genes that could be identified within this region, and was the only candidate gene for which a role in HPE could be established (Belloni *et al.*, 1996). Subsequent analysis of 30 families who had autosomal dominant HPE, revealed point mutations within the coding region of the *SHH* gene in 5 of these families, thus confirming that *SHH* is likely to be the *HPE3* gene (Roessler *et al.*, 1996).

These findings suggest that distinct differences exist between mice and humans with regard to their sensitivity to changes in *SHH* expression (and possibly the level of functional redundancy or dosage requirements for the *Sonic hedgehog* gene product), as mice that are heterozygous for *Shh* gene disruption are phenotypically normal (Chiang *et al.*, 1996), whereas it appears that in humans, mutations or deletions that affect only one copy of the *SHH* gene can cause HPE3 (Belloni *et al.*, 1996; Roessler *et al.*, 1996; Roessler *et al.*, 1997). This sensitivity to changes in *SHH* expression is thought to be particularly relevant to the ability of chromosomal rearrangements outside

the immediate vicinity of the human gene, to cause HPE, presumably as a result of reducing or altering *SHH* expression, through position effects (which have been reported to be capable of acting over distances of up to at least 400 kb; Bedell *et al.*, 1996) or the disruption of distal regulatory elements (Belloni *et al.*, 1996). Nevertheless, although variations in the precise nature of the *SHH* mutations or deletions that lead to HPE3, may to some extent explain the wide range of phenotypes that can be associated with this condition, some questions relating to the genotype-phenotype correlations observed in this disorder remain difficult to answer, and it has been suggested that additional genetic and/or environmental factors may affect the severity of the observed phenotype (Belloni *et al.*, 1996; Dean, 1996). These genotype-phenotype inconsistencies are exemplified by the findings that: i) the phenotypes associated with the deletion of one copy of the *SHH* gene, range from microcephaly (a phenotype that is considered to be an HPE microform), to severe HPE (Roessler *et al.*, 1997), and ii) in some cases, deletions of regions over 250 kb away from the *SHH* gene, can result in a phenotype more severe than that found in some instances where there are deletions that appear to encompass the entire *SHH* coding region. Interestingly, the other human *hedgehog* genes that have been identified to date do not map to HPE-associated loci (Dean, 1996), although proteins that are involved in SHH signalling pathways might prove to be strong candidates. As pointed out by Roessler *et al.*, (1997), further studies on the functional consequences of the *SHH* mutations identified in HPE patients, will ultimately be required to confirm that haploinsufficiency for SHH is responsible for 7q36-associated HPE. However, it is now generally accepted that *SHH* is the *HPE3* gene, and the emphasis of future work on HPE3 is likely to centre on determining the molecular basis for the phenotypic variability of this disorder.

Finally, it may be worth mentioning that an improbable, but nevertheless possible twist in the SHH/HPE3 story, might conceivably exist, as it has recently been shown that the effects of Indian hedgehog (Ihh; another member of the hedgehog family) on chondrocyte differentiation are mediated by PTHrP through the formation of a negative feedback loop (Vortkamp *et al.*, 1996). These effects of PTHrP are in turn mediated by the PTH/PTHrP receptor, and chondrocytes from mice that are homozygous for a targeted deletion of the PTH/PTHrP receptor gene are resistant to the effects of both PTHrP and Sonic Hedgehog (Lanske *et al.*, 1996). This raises the intriguing possibility that other members of the secretin/glucagon receptor family including the VIP_2 receptor, may also be involved in mediating some of the effects of Hedgehog proteins during embryonic development, although at present there is no further evidence to support this suggestion.

5.4.6 Is the VIP_2 receptor involved in mediating the dramatic effects of VIP on embryonic growth? Current arguments from a pharmacological/signal transduction perspective

Much of the evidence supporting the existence of a role for VIP_2 receptors in embryonic growth, hinges on the similarity between the adult distribution of the VIP_2 receptor (Lutz *et al.*, 1993; Sheward *et al.*, 1995; Usdin *et al.*, 1994) and the GTP-insensitive population of high affinity VIP receptors (Hill *et al.*, 1994; Hill *et al.*, 1992) in the rat CNS; a relationship that was first noted by Harmar and Lutz (1994) (Harmar & Lutz, 1994). However, although mRNAs encoding the VIP_1 and VIP_2 receptors are present during early embryogenesis (Sheward *et al.*, 1996b), the known ligand binding and signal transduction properties of these receptors do not appear to correspond with those of the GTP-insensitive population of VIP receptors (discussed in Gressens *et al.*, 1994). These discrepancies between the properties of the pharmacologically characterised and the cloned VIP receptors, have led several investigators to infer that a third high affinity VIP-specific receptor may exist. Nevertheless, despite extensive efforts, no one has yet cloned a receptor that fits this description.

The current arguments against the involvement of the VIP_2 receptor in mediating the growth-promoting effects of VIP in embryos, centre on, the failure of PACAP to stimulate embryonic growth (J. M. Hill, unpublished data, cited in Hill et al., 1996), the failure of PACAP to prevent the inhibition of embryonic growth produced by the VIP antagonist neurotensin-$VIP_{(7-28)}$, and the inability of PACAP antagonist $PACAP_{6-27}$ to duplicate the growth inhibitory effects of the neurotensin-VIP_{7-28} during embryogenesis (Gressens *et al.*, 1994). These observations provide reasonable grounds for the authors' conclusion that VIP receptors which bind both VIP and PACAP are unlikely to be involved in mediating the effects of VIP on embryonic growth. However, it may be worth noting that in reaching this conclusion several assumptions seem to be made.

The first of these assumptions is that both PACAP and VIP bind to exactly the same site on VIP/PACAP receptors, and that once bound they have the same effect on receptor signalling. This may prove to be correct, but it cannot be regarded as a certainty as there is already evidence from other G-protein coupled receptor systems, including the PACAP type 1 receptor, which indicates that this might not always be the case. A large proportion of this evidence centres on the theory of agonist-receptor trafficking, a concept that is based on the idea that different agonists acting at a given receptor may be capable of selectively stabilising the formation of different ternary complexes, which in turn can couple to different G-proteins [(Kenakin, 1995) reviewed

by (Gudermann *et al.*, 1996)]. One of the examples that is frequently cited in support of this theory, is the reversal of relative agonist potency that occurs between $PACAP_{1-27}$ and $PACAP_{1-38}$ for the stimulation of cAMP and inositol phosphate accumulation mediated by the short form of the PACAP type 1 receptor in transfected LLC PK1 cells (Spengler *et al.*, 1993). A similar effect is also found at the *Drosophila* octopamine-tyramine receptor in transfected CHO cells, where it has been shown that a reversal of agonist potencies occurs between octopamine and tyramine, depending upon whether cAMP inhibition, or the elevation of intracellular calcium is studied (Robb *et al.*, 1994). In the case of the PACAP type 1 receptor, studies of a recently isolated mRNA splice variant that results in the deletion of a 21 amino acid domain from the N-terminal extracellular region of the receptor, have demonstrated that this 21 amino acid domain of the PACAP type I receptor is responsible for the differential potencies observed for $PACAP_{1-27}$ and $PACAP_{1-38}$ in stimulating the accumulation of inositol phosphate, although the same domain appears to have little or no effect on the relative potencies of these peptides in the stimulation of cAMP (Pantaloni *et al.*, 1996). Inextricably linked with the concept of agonist-receptor trafficking, is the idea that different agonists binding to a given receptor could be differentially sensitive to the effects of antagonists acting at that receptor (Kenakin, 1995). Thus, if agonist-receptor trafficking was occurring at the VIP_2 receptor, it would explain to some extent many of the effects that to date have been interpreted as arguing against the involvement of this receptor in mediating the effects of VIP on embryonic growth.

The second assumption that appears to have been made, is that either PACAP type 1 receptors are not present in the developing embryo, or that the effects of PACAP binding at these receptor sites are the same as those of PACAP binding to VIP/PACAP receptors. As mentioned previously, it is now known that in E9.5 mouse embryos, mRNA encoding the PACAP type I receptor is much more abundant than mRNAs encoding either the VIP_1 receptor or VIP_2 receptor (Sheward *et al.*, 1996a), and it seems unlikely that these receptors have the same function. Therefore, the results of the PACAP antagonist study, in which it was shown that $PACAP_{6-27}$ could not reproduce the growth inhibition seen after administration of the neurotensin-VIP hybrid antagonist, will almost certainly represent the net effect of blocking the actions of PACAP at both the PACAP type 1 receptor and the VIP/PACAP receptors, and may not necessarily provide a true reflection of the actions of PACAP at VIP/PACAP receptors. Indeed, in the experiments carried out by Gressens et al., (1994), it was found that the addition of neurotensin-VIP hybrid antagonist + PACAP, resulted in a dramatic decrease in embryonic body growth that was not observed after the addition of the antagonist alone. Interestingly, a recent report on the expression of PACAP type 1

and VIP/PACAP receptors during phenotypic interconversion in the human neuroblastoma cell line SH-IN, has shown that a switch from large substrate-adherent slow-growing cells to focal aggregates of small rapidly growing cells, can be correlated with a change in the expression of receptors for VIP and PACAP from 95% PACAP type 1 receptors and 5% VIP/PACAP receptors in the large slow-growing cells to only 13% PACAP type 1 receptors and 87% VIP/PACAP receptors in the small fast-growing cells (Lelièvre *et al.*, 1996). The finding that VIP and PACAP are approximately equipotent in the stimulation of cell growth in both the large and small cell-types led the authors to suggest that the stimulation of the growth of these cells is mediated by VIP/PACAP receptors. However, whereas the addition of either 10 nM or 20 nM VIP or PACAP significantly stimulated the growth rate of both cell-types, the addition of 10 nM VIP + 10 nM PACAP had very little effect on proliferation (Lelièvre *et al.*, 1996). As noted by the authors, this finding indicates that under certain circumstances VIP and $PACAP_{27}$ may have mutually antagonistic effects (Lelièvre *et al.*, 1996), a factor which could be crucial in the interpretation of the experiments on mouse embryonic growth carried out by Gressens *et al.* (1994).

Do the growth-promoting VIP receptors couple to G-proteins in the embryonic CNS?

The apparent GTP-insensitivity of the growth-mediating VIP receptors in the embryonic rat CNS (Gressens *et al.*, 1994; Hill *et al.*, 1994; Hill *et al.*, 1992), is puzzling, because it seems to imply that no heterotrimeric G-proteins, or at least none of the known heterotrimeric G-proteins, are involved in the signal transduction mechanism(s) of this receptor in embryos. This phenomenon is not unique to VIP receptors, as in certain cell lines, agonist binding to other 7-transmembrane G-protein coupled receptors, including the 5-hydroxytryptamine type 2 receptor ($5HT_2$), and the angiotensin II type 2 receptor (AT_2), is also reported to be insensitive to GTP analogues (Buisson *et al.*, 1995; Szele & Pritchett, 1993). However, although the addition of 1mM GTPγS has no effect on the levels of angiotensin II binding to the AT_2 receptor in NG108-15 cells, the inhibition of T-type calcium currents that is mediated by the AT_2 receptor in this cell line, is clearly dependent on a GTPase activity (as indicated by the ability of GTPγS to replicate the effects of the agonist, and the ability of GDPβS to reverse these effects) (Buisson *et al.*, 1995). These studies, and those carried out by Szele and Pritchett (1993) on $5HT_2$ receptors in transfected HEK 293 cells, indicate that unknown G-proteins may exist for which the presence of GTP-analogues does not affect the binding of agonists by receptors (Buisson *et al.*, 1995; Szele & Pritchett, 1993). Interestingly, the signalling pathways in which these 'GTP-insensitive G-proteins' have been implicated, involve calcium regulation (AT_2

receptor), and inositol phosphate hydrolysis ($5HT_2$ receptor), pathways that are similar to those through which the GTP-insensitive VIP receptors are thought to signal (Hill *et al.*, 1994; Hill *et al.*, 1992). The status of the cloned VIP_1 receptor and VIP_2 receptor with respect to non-cAMP mediated signalling is unclear at present, but there is increasing evidence to suggest that the VIP_2 receptor can stimulate increases in intracellular calcium (Inagaki *et al.*, 1994; Straub & Sharp, 1996; Xia *et al.*, 1996b), and may also signal via a novel wortmannin-sensitive pathway (Straub & Sharp, 1996).

It also appears that the effects of Gpp(NH)p (the nonhydrolysable GTP analogue used to identify the GTP-sensitive receptors) on VIP binding to the VIP_2 receptor may differ from those expected, as it has been shown that in SUP-T1 cells, partial agonists (including VIP) binding to the helodermin preferring (VIP type 2) receptor in the presence of 10 μM Gpp(NH)p, show a decrease (3.3 fold for VIP) in their K_{act} for adenylate cyclase stimulation, relative to the K_{act} of the same partial antagonist in the presence of 10 μM GTP (Christophe *et al.*, 1990). The same effect is also seen in cholera toxin treated SUP-T1 cells, and binding studies confirm that the observed decrease in the K_{act}, reflects an increase in receptor affinity for partial agonists, and a decrease in receptor selectivity (Christophe *et al.*, 1990; Robberecht *et al.*, 1989). However, it is not known whether or not this effect of Gpp(NH)p is specific to VIP_2 receptors.

Other explanations for the GTP-insensitivity of the embryonic growth mediating VIP receptors, include the possibility that these receptors signal through a heterotrimeric G-protein independent pathway, similar to the Jak/STAT signalling pathway that has recently been proposed for the angiotensin II type 1 receptor (Marrero *et al.*, 1995). This would in many ways be consistent with the transduction of mitogenic signals by these receptors, as this type of signalling pathway has traditionally been associated with cytokine receptor signalling.

Could cAMP be involved after all?

While there is substantial evidence to suggest that high affinity VIP receptors that signal through non-cAMP pathways exist (Oláh *et al.*, 1994), it is less clear that these VIP receptors are solely responsible for mediating the growth effects of VIP in the embryonic CNS, as the neurotensin-VIP hybrid antagonist that implicated the GTP-insensitive receptors in this process appears to have some unusual properties. The original suggestion that the GTP-insensitive population of VIP receptors is responsible for mediating these growth effects, stemmed from the observation that neurotensin-VIP

hybrid antagonist produces an increase in the density of GTP-insensitive VIP receptors, but has no effect on the density of GTP-sensitive VIP receptors (Gressens *et al.*, 1993). However, although this does seem to indicate that the antagonist is acting selectively on the GTP-insensitive receptors, previous reports have clearly demonstrated that the neurotensin-VIP hybrid antagonist significantly inhibits VIP stimulated cAMP production in CNS cell cultures, at antagonist concentrations equivalent to those used in the Gressens *et al.*, (1993) study (Gozes *et al.*, 1991). The reason for the antagonist's failure to affect the density of GTP-sensitive VIP receptors (which are thought to signal via cAMP) is unclear, but it seems unwise to interpret these results as ruling out the involvement of cAMP or GTP-sensitive VIP receptors in the effects of VIP on embryonic growth.

Summary

Although there is considerable evidence in support of the existence of a third VIP-specific VIP receptor, this evidence is not indisputable, as many of the VIP-induced signalling pathways that have been attributed to this receptor, could equally be found to be mediated by the VIP_1 receptor or VIP_2 receptor under certain circumstances. The signalling mechanisms mediated by 7-transmembrane G-protein coupled receptors, appear to be highly dependent on the cell type in which the receptor is expressed, and are likely to reflect, both the complement of intracellular signalling components present in the cells (eg. G-protein and adenylate cyclase subtypes) and the types and levels of receptors expressed by these cells. However, the majority of studies on VIP_1receptor and VIP_2 receptor to date, have been carried out in a limited number of transfected cell lines, eg.COS, CHO, and HEK 293 cells, and the properties of these receptors in other cell types remain largely unknown. Finally, it is also possible that splice variants of the VIP_1 or VIP_2 receptors could be involved in regulating embryonic growth, as although no splice variants of the VIP_1 or VIP_2 receptors have been identified to date, their existence has not been excluded. Further studies of these receptors should help to clarify whether or not the VIP_2 receptor could be involved in the control of embryonic growth, but unless a VIP-specific VIP receptor is cloned, this issue seems likely to remain unresolved for some time.

Chapter 6

Isolation and sequence analysis of the putative promoter region of the mouse *Vipr2* gene

6.1 Introduction

The transcription factors that bind to the promoter region of a gene, hold the key to many of the mechanisms that are responsible for the control of the spatial and temporal patterns of the gene's expression. Ubiquitous, tissue specific, and inducible transcription factors that can activate or repress transcription, have all been shown to play important roles in the control of gene expression.

The tissue- and cell-specific expression pattern of the *Vipr2* gene, together with the proposed role of the VIP_2 receptor as the neuroendocrine VIP receptor (Usdin *et al.*, 1994), strongly suggest that the transcriptional regulation of this gene is likely to involve tissue-specific and hormonally-regulated transcription factors.

However, in order to unambiguously identify the promoter region of a gene, it is first necessary to determine the position of the gene's transcriptional start site(s). The published sequence of the mouse Vipr2 cDNA, contains 52 bp of 5' untranslated sequence (GeneBank accession number D28132, (Inagaki *et al.*, 1994)), but to date there have been no reports of the use of RACE-PCR, primer extension, or RNase protection to determine the position(s) of the transcriptional start site(s) of the *Vipr2* gene. Preliminary studies carried out by E. Lutz in our laboratory, using 5' RACE-PCR on RNA isolated from the mouse AtT20 cell line, resulted in the isolation of several PCR products the largest of which extended 54 bp 5' of the published Vipr2 cDNA sequence. This clearly demonstrates that transcribed sequences of the *Vipr2* gene are present upstream of the published cDNA sequence, and although the isolation of multiple RACE products indicates that premature termination of the extension reaction may be occurring, the repeated identification of this extension product in several separate experiments seemed to suggest that the 5' end of the *Vipr2* transcript in AtT20 cells lies at or close to this point.

The 5' regulatory regions of several members of the secretin/glucagon receptor gene family have been isolated and sequenced, including that of the *VIPR1* gene (Aino *et al.*, 1995; Burcelin *et al.*, 1995; McCuaig *et al.*, 1995; Sreedharan *et al.*, 1995; Tsai-Morris *et al.*, 1996). In general, the proximal promoter regions of these genes appear to be GC-rich, and contain multiple binding sites for the transcription factor Sp1 instead of a

TATA box (Aino *et al.*, 1995; Burcelin *et al.*, 1995; McCuaig *et al.*, 1995; Sreedharan *et al.*, 1995; Tsai-Morris *et al.*, 1996). In some cases a CCAAT box is also present upstream of the Sp1 sites (Aino *et al.*, 1995; Sreedharan *et al.*, 1995).

In order to isolate the 5' region of the mouse *Vipr2* gene, a restriction fragment from the 5' RACE product described above, was used as a probe on Southern blots of lambda clones ESC4A and ESD1, the latter of which had already been identified (see chapter 3) as containing sequence corresponding to the 5' end of the published mouse Vipr2 receptor cDNA sequence (GeneBank accession number D28132 (Inagaki *et al.*, 1994)). Sequence analysis of a fragment isolated from ESD1 demonstrated that the characteristics of the region immediately upstream of the RACE product were consistent with those of promoter regions from this family of receptor genes, and also raised the possibility that additional promoter or enhancer regions may exist further upstream. In an attempt to confirm the position of the *Vipr2* gene's transcription start site(s), primer extension experiments were carried out using template RNA isolated from mouse olfactory bulbs and AtT20 cells. However, although control experiments were successful, no extension product from *Vipr2* mRNA was observed. Possible reasons for this, and the results that have subsequently been obtained in our laboratory through the use of alternative approaches, are discussed.

6.2 Methods

6.2.1 Southern blotting, subcloning, and sequencing, of clones and subclones from the 5' region of the mouse *Vipr2* gene

Southern blotting of clones ESC4A and ESD1 was carried out using Appligene Positive™ nylon membrane (Appligene Oncor) and alkaline transfer conditions (see Chapter 2). The *Xba* I restriction fragments from ESD1 were subcloned using the 'shotgun' subcloning method described in Chapter 3, and sequencing of these subclones was then carried out as described in Chapter 2.

6.2.2 RNA preparation

All glassware and plasticware that was to be used in the preparation of RNA, or in the subsequent primer extension experiments (6.2.2), was treated with diethyl pyrocarbonate (DEPC) solution (0.1% in water) overnight at room temperature, then autoclaved for 20 minutes at 15 lb/sq. in. to remove residual DEPC. Solutions used in

these experiments were made with DEPC-treated water where possible, and were also autoclaved before use.

Olfactory bulb RNA

Total RNA was prepared from 50 mouse olfactory bulbs, using the guanidinium thiocyanate/caesium chloride centrifugation method (Sambrook *et al.*, 1989).

AtT20 cell poly A$^+$ mRNA

The poly A$^+$ mRNA used in this work was isolated by Dr. E Lutz.

Catriomox-14 surfactant (V. H. Bio) was used to extract total RNA from confluent AtT20 cells. Two millilitres of surfactant was added to each 75 cm^2 flask of cells. The cells were gently scraped into the Catriomox solution, then pelleted by centrifugation at ~1000*g* for 10 minutes, and washed several times in 2 M lithium chloride to remove DNA. After the washed pellet had been resuspended in DEPC-treated H_2O and reprecipitated using sodium acetate (pH 4.0) and ethanol, poly A$^+$ mRNA was isolated using an Oligotex™ mRNA isolation kit (Qiagen Ltd) according to the manufacturer's instructions.

6.2.3 Primer extension assay

Primer labelling

A 30-mer DNA oligonucleotide (primer 19043) dCAGCAACCAGCAGTAGCAGGTCAGCACCAC which is complementary to bases 94-65 of the published mouse Vipr2 cDNA sequence (GeneBank accession number D28132, (Inagaki *et al.*, 1994)), was used as a probe for this assay. For each experiment 100 ng of primer was end-labelled with [γ-^{32}P]ATP, using a Ready-To-Go T4 Polynucleotide Kinase kit (Pharmacia Biotech). Labelling reactions were carried out according to the manufacturers instructions.

Primer extension method 1

For each extension reaction 500,000 cpm of 5' end-labelled primer and 50 µg of total RNA from mouse olfactory bulbs were co-precipitated in a 1.5 ml Eppendorf, using sodium acetate (pH 4.0) and ethanol, and the pellet was resuspended in 20 µl of a solution containing 80% formamide, 40 mM PIPES pH 6.4, 400 mM NaOH. This RNA/primer mix was overlaid with 30 µl of mineral oil, and heated to approximately

90°C for 5 minutes, then placed in a PCR machine which had been pre-heated to 50°C, and left there overnight to allow the primer to anneal.

The following day, the RNA/primer mix was taken from the PCR machine, and chilled on ice. The mix was then spun briefly in a microfuge, after which the reaction mixture was removed from beneath the mineral oil. Following precipitation using 70 µl of 0.3 M NaOAc and 2.5 volumes of EtOH, the RNA/primer pellet was resuspended in 20 µl of reverse transcriptase buffer (50 mM Tris pH 7.5, 60 mM KCl, 10 mM $MgCl_2$, 5 mM dNTP, 1 mM DTT, 50 µg/ml actinomycin D, and 5% v/v DMSO). One µl of RNasin (Promega) per 100 µl of buffer and 2 µl of Superscript II reverse transcriptase (Gibco BRL) per 100 µl of buffer, were added to the buffer just before the buffer was added to the RNA. The extension reaction was incubated at 42°C for 1 hour, after which 80 µl of RNase solution (10 mM Tris pH 7.5, 10 mM EDTA, 10 µg/ml RNase A) was added and the reaction was incubated at 37°C for a further 30 minutes. Finally, 1/10th of a volume of 3M NaOAc was added to the extension reaction, which was then phenol/chloroform extracted, and EtOH precipitated using 10 µg of tRNA as a carrier. The reaction was resuspended in 80% formamide/dye mix, heated at 90°C, and then loaded onto a sequencing gel beside dideoxy sequencing reactions from plasmid clone D1X8 (a genomic DNA subclone which contains the primer target site) that had been sequenced with the same 30-mer primer (19043). Sequencing reactions were carried out using a Sequenase 2.0 kit (United States Biochemicals) according to the manufacturer's instructions.

Primer extension method 2

Primer labelling, DNA marker labelling, and primer extension reactions were carried out using the Promega AMV Reverse Transcriptase Primer Extension System (Promega). Reactions were carried out according to the manufacturer's instructions. Sequencing reactions were carried out as described in method 1 (see above).

6.3 Results

6.3.1 Identification and isolation of a ~5.5 kb fragment which includes the putative promoter region of the mouse *Vipr2* gene

Southern blot analysis of the λ2001 ES clones C4A and D1 (see chapter 3) was carried out using an ~190 bp *Pst* I/*Xba* I restriction fragment from the 5' end of the mouse Vipr2 5' RACE product Q12 as a probe. Q12 (isolated by E. Lutz, and previously

described in Chapter 3) is the largest of a series of 5' RACE products generated from RNA from the mouse pituitary AtT20 cell line, and extends from base pair 192 of the published mouse cDNA sequence (GeneBank accession number D28132, (Inagaki *et al.*, 1994)) to a point 54 bp 5' of the published sequence (246 bp in total). Southern blots of gels containing *Pst* I, *Xba* I, and *Pst* I/*Xba* I, restriction digests of ESC4A and ESD1 DNA, demonstrated that only clone ESD1 contained DNA fragments that hybridised to the Q12 derived probe. The ESD1 restriction fragments that showed strong hybridisation to the probe were a 1.7 kb fragment from the *Pst* I digest, a 1.7 kb fragment from the *Pst* I/*Xba* I digest, and a 5-6 kb fragment from the *Xba* I digest (Figures 6.1a and 6.1b).

Since it was known that: i) *Xba* I cuts the polylinker sites in both arms of the λ2001 vector (see Chapter 2 Figure 2.5 for vector map), ii) an *Xba* I restriction site is present in the published mouse Vipr2 cDNA sequence (GeneBank accession number D28132, (Inagaki *et al.*, 1994)) at base pair 133, a site which lies 103 bp upstream from the 3' end of clone ESD1 and is located in exon 2 of the *Vipr2* gene, iii) the size of intron 1 of the mouse *Vipr2* gene was estimated to be ~2.2 kb; it seemed likely that the 5-6 kb *Xba* I restriction fragment that had been identified using the Q12 derived probe would contain several kilobases of genomic DNA lying upstream of Q12, i.e. the putative promoter region of the *Vipr2* gene. Subcloning of this *Xba* I restriction fragment into pGEM7Z(+) was carried out using the "shotgun" subcloning method described in Chapter 3, and two clones with inserts of the expected size (clones D1X2 and D1X8; Figure 6.2) were selected. Sequence analysis of the D1X2 and D1X8 insert ends, using T7 and SP6 primers (Promega) revealed that the 3' end of the insert in clone D1X8 contained sequence that corresponded to the region directly upstream of the exon 2 *Xba* I site, whereas the sequence obtained from the 3' end of D1X2 was equivalent to the intron 2 sequence that was found at the 3' end of the ESD1 λ clone insert. Both subclones also generated sequence data when tested with the *Vipr2* gene exon 1 primer 8484, thus generating genomic sequence that extended upstream of Q12. However, although both D1X2 and D1X8 appeared to contain the region of interest, the presence of an additional 103 bp of DNA at the 3' end of D1X2 indicated that this subclone contained a partial digestion product from the *Xba* I digest of ESD1 DNA, and D1X2 was not used in subsequent studies.

6.3.2 Construction of a basic restriction map of subclone D1X8

A series of double digests carried out on D1X8 using the restriction enzyme *Xba* I in combination with other restriction enzymes, revealed that *Sal* I cut the D1X8 insert

Figure 6.1a

Restriction enzyme digests of DNA from clones ESC4 and ESD1, separated on an agarose gel in preparation for Southern blotting.

(A ruler has been photographed alongside the gel so that if necessary the positions of the molecular weight markers can be estimated on the corresponding autoradiograph of the hybridised blotting membrane).

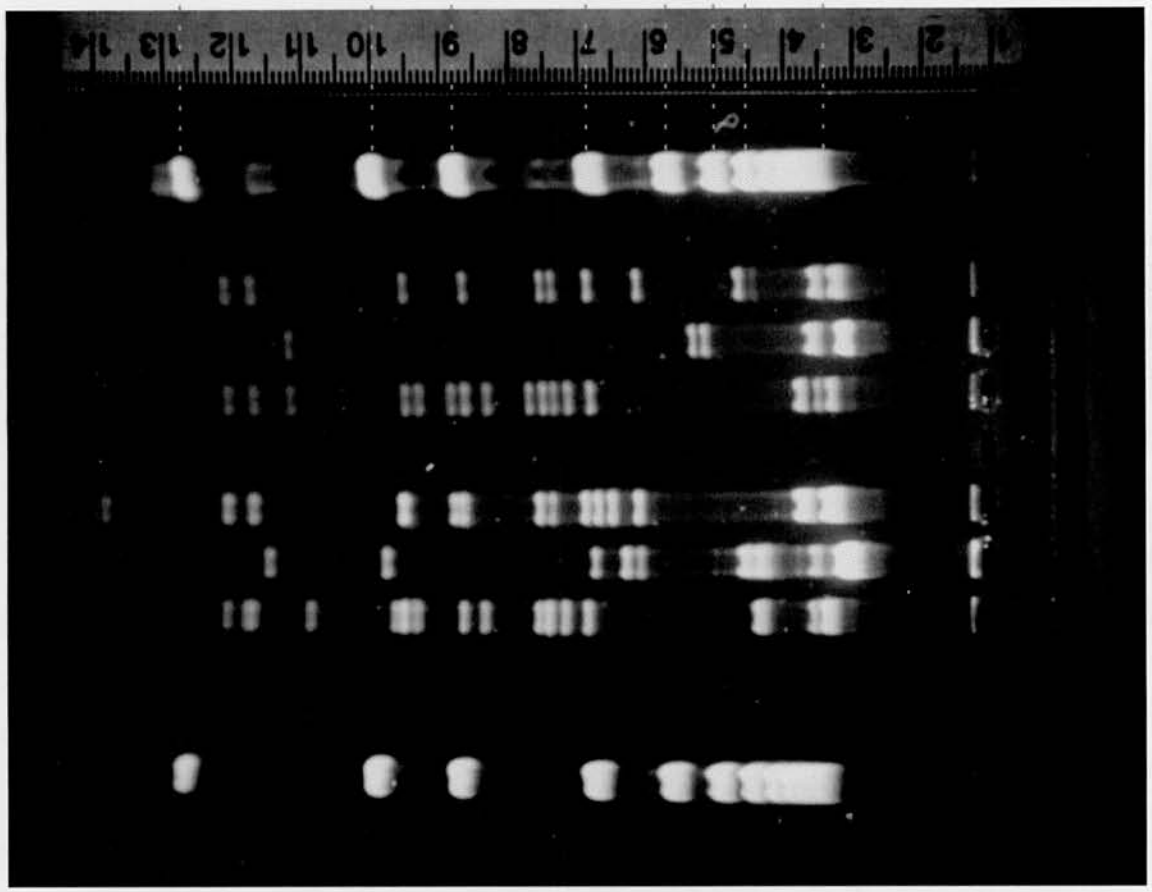
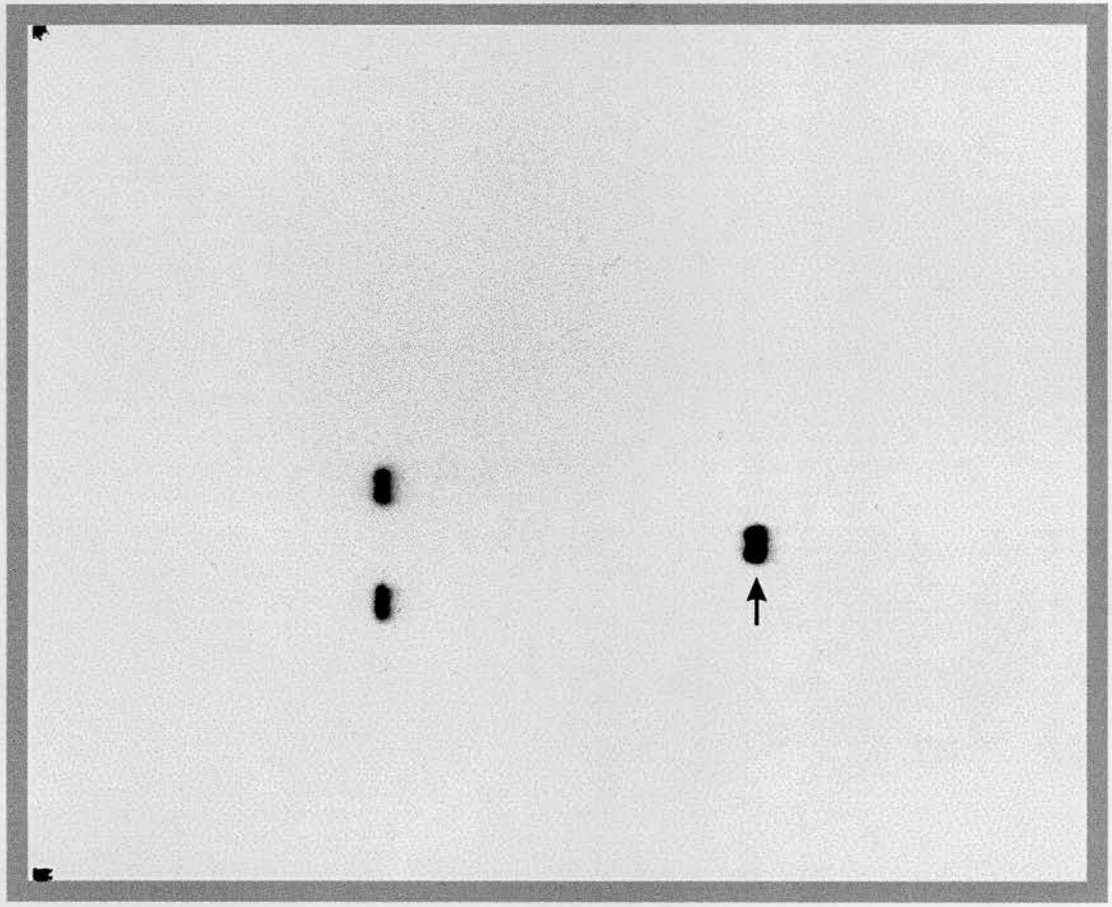

1.02 kb
1.64 kb
2.04 kb
3.05 kb
4.07 kb
5.09 kb
6.11 kb
12.2 kb
1 kb ladder
ESC4 *Pst* I
ESC4 *Xba* I
ESC4 *Pst* I/*Xba* I
ESD1 *Pst* I
ESD1 *Xba* I
ESD1 *Pst* I/*Xba* I
1 kb ladder

Figure 6.1b

Hybridisation pattern observed when a Southern blot taken from the gel of restriction digested ESC4 and ESD1 DNA shown in Figure 6.1a is incubated with the radiolabelled Vipr2 mouse 5' RACE product (Q12) -derived cDNA probe

A 5-6 kb *Pst* I/*Xba* I fragment that hybridised to the probe is arrowed.

(Blotting membrane is shown at actual size)



1 kb ladder

ESC4 *Pst* I

ESC4 *Xba* I

ESC4 *Pst* I/*Xba* I

ESD1 *Pst* I

ESD1 *Xba* I

ESD1 *Pst* I/*Xba* I

1 kb ladder

Figure 6.2: Subcloning of *Xba* I fragments from ESD1

(suclones D1X2 and D1X8 appeared to have single inserts of the desired size, 5-6 kb)

1 kb ladder
D1X1 Xba I
D1X2 Xba I
D1X4 Xba I
D1X8 Xba I
D1X17 Xba I
ESD1 Xba I
1 kb ladder
12.2 kb
7.13 kb
6.11 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb

once, to produce two fragments (~4 kb and ~2 kb) in an *Xba* I /Sal I digest. This *Sal* I site was chosen as the starting point for the construction of a restriction map, and the orientation of the two *Xba* I/*Sal* I restriction fragments relative to the SP6 and T7 polylinker sites of the pGEM7Z(+) vector was determined by *Xho* I/*Sal* I double digestion of D1X8. *Xho* I cuts in the SP6 polylinker region of pGEM7Z (+), but not in the T7 polylinker region (see Figure 2.2 for vector map). The release of an ~2 kb *Xho* I/*Sal* I restriction fragment (Figure 6.3), therefore demonstrated that the *Sal* I site in D1X8 lies ~2 kb downstream from the SP6 polylinker/insert junction, and ~4 kb upstream from the exon 2 *Xba* I site that lies at the T7 polylinker/insert junction. Subsequent single and double restriction digests of D1X8 using the enzymes *Sal* I, *Pst* I, *Nco* I, and *Xho* I, then allowed the positions of restriction sites for these enzymes, relative to those for *Sal* I and *Xho* I, to be established. The preliminary restriction map of D1X8 that was derived from these results is shown in Figure 6.4.

6.3.3 Sequence analysis of the putative promoter region of the mouse *Vipr2* gene

The D1X8 insert was sequenced by primer walking, beginning with the sequences obtained using the SP6 primer (5' vector/insert junction) and the internal primer 8484 (exon 1). Successive internal primers were designed to bind to the end of the D1X8 sequences that had been revealed by the previous sequencing primers, and sequencing was continued until an overlap between the two contigs of sequence data was observed (see Figure 6.4). Several sequencing reactions were set up for each of the sequencing primers, and a consensus sequence was derived by comparison of the sequences obtained from the individual reactions. However, it should be noted that for each region of DNA that was sequenced, the data results from sequencing of only one strand of the D1X8 DNA. In total, the sequence data obtained spanned 3325 bp, 3193 bp of which lay upstream of the 5' end of the Vipr2 5' RACE product derived sequence. The overall consensus sequence for the putative promoter region of the *Vipr2* gene, that was obtained from D1X8, is shown in Figure 6.5. Although the high GC content of the proximal region of the putative *Vipr2* promoter was apparent even from visual inspection of the sequence data, the computer-based sequence analysis programs Genejockey II (Biosoft) and the Genetics Computer Group program "Mapplot" (HGMP Resource Centre, Cambridge, UK) were used to confirm this, and to examine the possibility that a CpG island may be present in this region. In contrast to the overall GC content of the ~3.2 kb consensus sequence, which was estimated to be 49%, the GC content of the 300 bp lying directly upstream of the 5' RACE product derived sequence was ~65%. This region of high GC content extends downstream beyond the

Figure 6.3: Restriction digest to determine the orientation of the D1X8 insert.

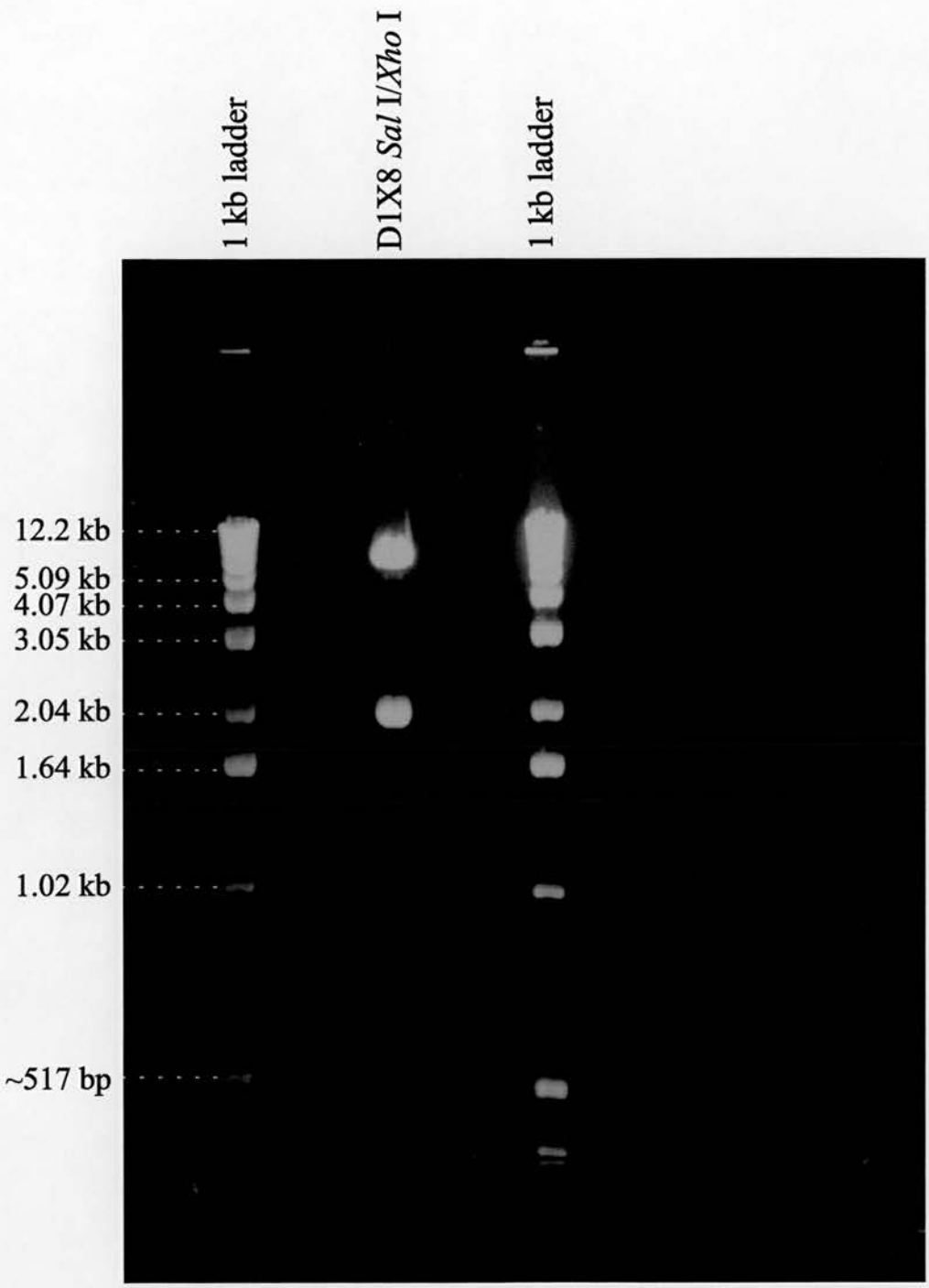

1 kb ladder
D1X8 Sal I/Xho I
1 kb ladder
12.2 kb
5.09 kb
4.07 kb
3.05 kb
2.04 kb
1.64 kb
1.02 kb
~517 bp

Figure 6.4: Preliminary restriction map of subclone D1X8

(The positions of the sequencing primers that were subsequently used to determine the sequence of the putative promoter region of the *Vipr2* gene, are also shown, and their direction of extension is indicated by a small arrow)

Subclone D1X8

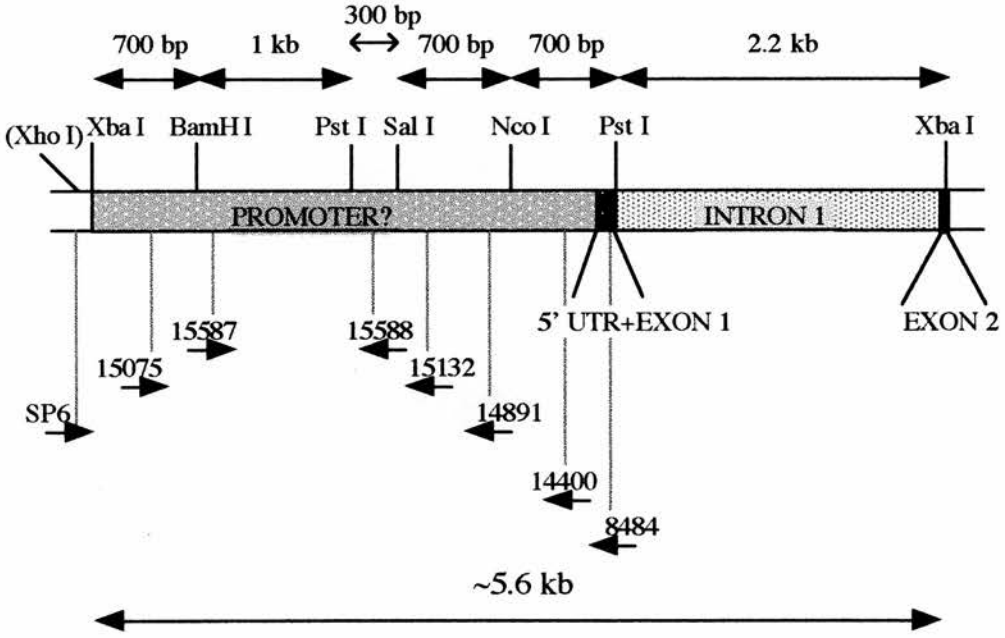


Figure 6.5: Sequence data obtained from the 5' flanking region of the mouse *Vipr2* gene

The sequence shown opposite is a consensus sequence (obtained from sequencing of the ESD1 *Xba* I subclone D1X8, see Figure 6.4) which includes the putative promoter region of the mouse *Vipr2* gene. The sequences shown in blue represent the 5' untranslated region of the *Vipr2* gene. Published untranslated sequence (GeneBank accession number D28132; Inagaki *et al.*,1994) is shown in dark blue, whereas the unpublished untranslated sequence identified from the Vipr2 5' RACE product is shown in light blue. The transcription start site is shown in red and the remainder of exon 1 is shown in green.

IUPAC codes for DNA
A or C = M; A or G = R; A or T = W; C or G = S; C or T = Y; G or T = K;
C, G or T = B; A, G, or T = D; A, C, or T = H; A, C, or G =V; Any = N

```
   1 TCTAGAAAAA CTGAGCCTGT GGTATGGGTG GGTCAGACCT GATAGTGAGG TGAGACACCC
  61 ATGAAAACCT GCTCCATTCA CTGACTACCT CTGTAAAACA AAAAAAAAAG ACACAAAGGC
 121 TGGCCTTTCA GATTTGTATA ATAAAATAAA TTACAAGGTT GTTTTCCCCA GCTTTGACAG
 181 GATATTAGGT AAAACTCATT TTCACTTTAA AGACTCATAA CAATATGAAA TGCATATAAA
 241 TCTTACATTT AGTTTGCTTT GCAGGGTATT AGCAGGAGTC AGGAGGAGCC ACCAACCCAT
 301 CATCTGATGC GTTTTTGCAA AGGAGATTTG CCCTTTACAA ATCAGTTGCG TGTTGTCTTC
 361 CTTCTGTGCA CATGAGAGGT GCAGAAGCCT TGGAGACAAC CGCAGCTGTA GCCCCTAATG
 421 GGTCGAGTGT CTCTCATAAC CACCCCTCCA GGCACAGCAA ATAGATTCAG TTTATGTACT
 481 CCCCTGAAAT TGCTCCCCTT ATTACAGGGC GGTGTCAGAA GGAAGTGCTG GGAAAGAACT
 541 GTCGTTTTGT AAAAGTAAGA CAAACAGAGA TGAAACAATC GCTTTCATTC TTTGTTGTTG
 601 TTGAAGATTT TATGTTGAAC ACCAGGCTCT TGTAGATACG AAGGCAGCAG CACTTCTGAG
 661 GGGGAGTCAG CGGATGTTAC TGGGTATTGT CTTGTCAGCT CCTGAGGATC CTCTTCTTAT
 721 GTGCTCATAC AGGCTTCTAA TGTCTCTGAT AAGCTTCATC ACTGGACCCA CTCCATGTGG
 781 AAGGTGCTAG CTGTTTGGCC GTCATTGCTG ACCCAAAACT TCACTGGCAA AAAAGCAAAG
 841 GGACTCTGTG ACCACATCCT AAGAGAACAG AGGGTTAGCT AAAATGGCTT GTGAAGAACT
 901 CTGCCCCAAA TGCTGCTGCT GCTGCTGCTT TGGGTATGAG AAGCCTCCGT GATAGTGAGG
 961 ACTGAAACCC TARAGTCTGG AAACCAGAGC AGCCACTCAA GTGGAAGGTG CCAGGAAACT
1021 GCCCAAGAGG ATGTCAGTGG GGAAGGGATA GGAAGGCCAG ACATCAGTAG GTAGAGATGA
1081 AACGGGGCTA ARAGCAGGTC TGGAGGATGC TGGAGGGCAC AACACAAGGC TGAACTCAGG
1141 TTCATATCCA GGCTGTGAAC AATGTGAAGT CCCATCAGAA CCAAGCAGTC ACCATAGGCA
1201 CAAGAGTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC
1261 TCTCCWWACC CCAATTCAGA ARAATGGAGA GACCTTGAAA GACATGACCA GTCTTTTTAG
1321 GACCTTCGYT CYTCCAGGAT ACCCCCAGGA AAACAAGTCC YTCTTGTCTC CCCGTGCCTG
1381 TATCTCATCC TTGTCAGGAC TATCACTTCA TCTCATCATA TATGCCATTG TGTGGTTGGA
1441 ACCCTGGCGT CTCACCCAGT TTTGACCAAT AACATCCGGA CGTGACTTTA GAATGCATCC
1501 CTCACAACAA TTCAGGTGTA TCAGCTCTCC TAATTCTAAT TGTAACCCTC TGGGATGGAG
1561 CAATGGCAGG GGTCGATGTT GTGATTTACA AATGTGCAGG TGGATCAGGT TCACTGGCTG
1621 AAACAGTGGA CAGGGGTTGT GGACAAAGCC AGAGCCCAAG CTTCCTGGCC CCAAGCTGCA
1681 GACCTAGTGC CCAGACCTTC TCCTCCATAA AATCCCCTGA TGGACAAACA CAGGCAGATG
1741 CACAGATAAT GGAAGGGTGA CAGAAAACAG ATAGAAACCT AGGTCCAAGG GTACACGGGG
1801 CTCTGATGTC ATGGAGCACA TGATGAGCTC AGTGGGGAGA AAGATAATGG GCAGATAGAC
1861 ACCCCAAAGT GTGAAGGATA CATTTTGTAC TATTTCAGAC TTCTGACTTT GGGCTTGAAA
1921 TTATATYTGA ATAAGAAGTT TAAAACTTGC CTGCTTTACT TAATCGAGTA CTATTGTYTG
1981 CTGGTAGCTT TYTGCTYTTC CCTAGACTGG TCTGTCGACA GACAGCCAAA CCATAGCTGA
2041 ACAGTGGCAG TGTYTCCACA AGCCTAGTGT ACAATGAGCC ACACACGTGT CACCTGAGGG
2101 TTTGGGGTGC CGATGAGCAG ACCCACTTTT TTTTTCTTTT TTTTTTTTTC CGTTTTTTAT
2161 TAGGTATTTA GCTCATTTAC ATTTCCAATG CTATACCAAA AGTCCCCCCA TACCCAACCC
2221 AACCCCACTC CCNTACCCAC CCACTCCCCT TTTTTGGCCC TGGCGTTCCC CTGTACTGGG
2281 GCATATAAAG TTTGCGTGTC CAATGGGCCT CTCTTTCCAG TGATGGCCGC CTAGGCCATC
2341 TTTTGATACA TATGCAGCTA GAGTCAAGAG CTCCGGGGTA CTGGTTAGTT CATAATGTTG
2401 TTCCACCTAT AGGGTTGCAG ATCCCTTTAG CTTCTTGGGT ACTTTCTCTA GCTCCTCCAT
2461 TGGGGGCCCT GTGATCCATC CAATAGCTGA CTGTGAGCAT CCACTTCTGT GTTTGCTAGG
2521 CCCAGGCATA YTCTGACAAG AGACAGCTAT ATCAGGGTCC TTTCAGCATA ATCTTGCTAG
2581 TGTATGCAAT GGTGGAGCAG ACCCACTTTT GCGTTGATGA AGTCTCTGCT TTCTGTTCAA
2641 CTTGGCCTTA GGCTCTCAAG GGGAGCCATG CAAGAGCTTT CTCTTTTCTC TCTGCTCACA
2701 CCAGTGCAGA CCTAACCATG GTCGTCCAGC CACTTTCTGC ACTTTTTCCT TGTGCTGCAC
2761 ACAGGCTTCG CATTTGCACA GGGAGTTGGT GCCAGATGCT ACCTCAGCGC TCAAGTTCCC
2821 TTACCTGGAA TTGGGCAATG CCCTCTGTGG GTGAGCATCT GCTTACTTGA GCCACCAGCT
2881 CTCCAAAGGC TTGAACATGG CAGCCCACAG TGGGGTCGCA GAGGAGGGAG GCAGTACAGC
2941 TCGGTTAGCA GAGGGAGCCT GGGAGGAAGG AGCGCGAGAG AGCAGAGCTG ACAAGGGAAG
3001 TACCCAGGCC AGGAGGGGGT CGGGGTGAGG CCCAGGGTCC TGCACTTCAG AGGGAAGTAG
3061 GGGTGGAAGG AGGGACGTTT GGACAGAGAT CTGGTGGACC GGAGTGCCAG AGAGAGACTC
3121 CCCAAAAAGG AGAAAGGGGC GCGAGGCGGG GACTGGNCCG AGGGGCGGGN CCGGGACTTG
3181 CAGCCTGGAA GGCGGAGAGG GCGATAGGCG CGAGACTGAG GAAATCCGCG CTGGGAGGCC
3241 CGGAGCAGGG GACCGTGCTG CTGAGGCGCC AAGGACCGAG GCGGCACGCT GAGCCCAAGA
3301 TGAGGGCGTC GGTGGTGCTG ACCTGCTACT GCTGGTTGCT GGTGCGG
```

putative promoter region, through the 5' untranslated sequence, exon 1, and into intron 1 of the *Vipr2* gene; a pattern that is frequently found in association with CpG islands (Cross & Bird, 1995). Analysis of the distribution of CpG dinucleotides within the D1X8 consensus sequence (Figure 6.6), revealed that a CpG island is present at the putative 5' end of the *Vipr2* transcribed sequence, beginning around 160 bp upstream of the proposed translational start site (base pair 3140), and continuing at least to the end of the D1X8 consensus sequence, which ends 375 bp into intron 1 (base pair 3722). A search of the DNA sequence databases using the BLASTN program (National Centre for Biotechnology Information, Bethesda, USA), with the putative promoter sequence of the *Vipr2* gene as a template, revealed the presence of an *L1*-type repeat sequence in the region from -705 to -1177 bp, but did not identify any other extensive regions of homology between the putative promoter region of the *Vipr2* gene and the promoter regions of other genes.

6.3.4 Primer extension experiments

Due to the GC rich nature of the sequence lying directly upstream of the *Vipr2* translation start site, Superscript II was selected for use in primer extension experiments as this enzyme is stable at temperatures of up to 55°C. Initial primer extension experiments carried out using primer extension method 1 (see Methods) on olfactory bulb RNA failed to produce any detectable extension products, although the primer successfully primed the dideoxy sequencing reaction from the D1X8 plasmid. Subsequent attempts at primer extension, carried out using the Promega AMV Reverse Transcriptase Primer Extension System (Promega) on poly A^+ mRNA from mouse AtT20 cells, also failed to produce any detectable extension products from *Vipr2* transcripts. However, using this method (primer extension method 2; see Methods) an extension product of 87 bp was obtained from the control primer and RNA that were provided with the kit, suggesting that the problems that were encountered may be specific to the detection of extension products from *Vipr2* transcripts, rather than being the result of a more general technical problem.

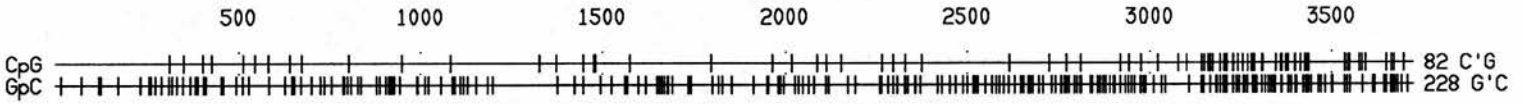
6.3.5 Identification of possible transcription factor binding sites

The transcription factor databases: TRANSFAC 3.0 including MatInspector 2.0 (GBF -Braunschweig, Germany; (Wingender *et al.*, 1997)), and TFD (National Institute of Health, Bethesda, USA), were used to search for possible transcription factor binding sites within the putative promoter region of the *Vipr2* gene. The analysis programs that are linked to these databases produced somewhat different results when presented with the D1X8 consensus sequence, due in part to differences in the content of the individual

Figure 6.6: A map of the relative distribution of CpG and GpC dinucleotide sites within the putative promoter region of the mouse *Vipr2* gene

(Linear) MAPPLOT of: mus5pr ck: 421, 1 to: 3722 January 31, 1997 16:36.

Enzyme Data: enzyme.dat



Enzymes that do not cut:

NONE

databases, and whether matrix consensus sequences or sites from individual gene sequences were used for binding site identification. These differences, together with the inherent bias towards short recognition sequences that occurs in these analysis programs, meant that a considerable amount of personal discretion was necessary in the interpretation of these results. The length and degeneracy of each potential recognition site were taken into consideration when deciding which sites to include in the overall picture of the promoter region at this stage. Transcription factor binding sites of 5 bp or less have generally been excluded from the results presented here even if there is no degeneracy in the recognition sequence, and binding sites that are 6 bp long have usually only been included if they have well-defined recognition sequences with little or no degeneracy, although there are some exceptions. Longer recognition sequences for transcription factors were also excluded despite their length if they displayed unusually high levels of degeneracy, but in each case every effort was made to check the *Vipr2* sequence against the individual gene sequences from which the consensus recognition site had been derived. Searches were also carried out for some transcription factor binding site sequences that were not represented in the databases, particularly if the known functions or expression patterns of the factors in question suggested a possible role in the control of *Vipr2* gene expression.

The results of the search for transcription factor binding sites in the region of the *Vipr2* gene that lay upstream of the sequence obtained from the 5' RACE product indicated that at least two potential TATA boxes and one pair of potential Sp1 binding sites were present within this sequence. The pair of potential Sp1 binding sites begin at -157 bp and -138 bp from the translation start site and lie within the CpG island described above, the latter of the two potential Sp1 sites ending 24 bp upstream of the 5' RACE product sequence and 11 bp upstream of a potential cap site. In contrast, the potential TATA box sequences begin at -3065 bp and -1016 bp relative to the translational start site of the *Vipr2* gene. Although neither of these TATA box consensus sites have a CCAAT box situated 30-40 bp upstream of the TATA site, a CCAAT box is found ~100 bp upstream of the proximal (-1016 bp) TATA box.

Putative transcription factor binding sites that were identified upstream of the distal (-3065 bp) TATA box sequence included possible sites for: activator protein 3 (AP3), the pituitary factor (Pit-1) (Peers *et al.*, 1990), a factor that binds to a T/A rich sequence in the growth hormone gene (GH-CSE) (Ye *et al.*, 1988), a component of the NF-CLEO complex (that binds to the promoter regions of several interleukin genes) (Miyatake *et al.*, 1991), and the Sry-related testis-specific factor Sox-5 (Denny *et al.*, 1992). Interestingly the TATA box itself appears to overlap with a potential binding site for the

POU (pit-oct-unc) factor Brn-2 (Li *et al.*, 1993). Possible transcription factor binding sites that lie upstream of the proximal (-1016 bp) TATA box sequence include potential sites for: a cAMP responsive factor, several E-box binding proteins (Fisher & Goding, 1992), Brn-2, activator protein 2 (AP2) , and serum response factor (SRF) (Treisman *et al.,* 1992). A putative cap-box sequence is found within the context of a sequence which corresponds to the Yin Yang 1 (YY1) consensus (Shrivastava & Calame, 1994), ~50 bp downstream from this TATA box. Finally, the GC rich putative promoter region that lies directly upstream of the 5' untranslated region identified by RACE PCR, contains possible binding sites for: mammary activating factor (MAF) (Mink *et al.*, 1992), the transcriptional silencer tramtrack (Ttk 69) (Guo *et al.*, 1995; Read & Manley, 1992), activating protein (AP2), lymphocyte factor-1 (Lyf-1) (Lo *et al.*, 1991), E-box binding proteins, and signal transducer and activator proteins (STATs) (Seidel *et al.*, 1995), interleukin-6 responsive factors (IL-6 RE-BP, NF-IL6) (Hocke *et al.*, 1992), and an interferon response factor (ISGF2) (Tanaka *et al.*, 1993). The positions of the sites described above and other putative transcription factor binding sites within the DNA fragment are shown in Figure 6.7.

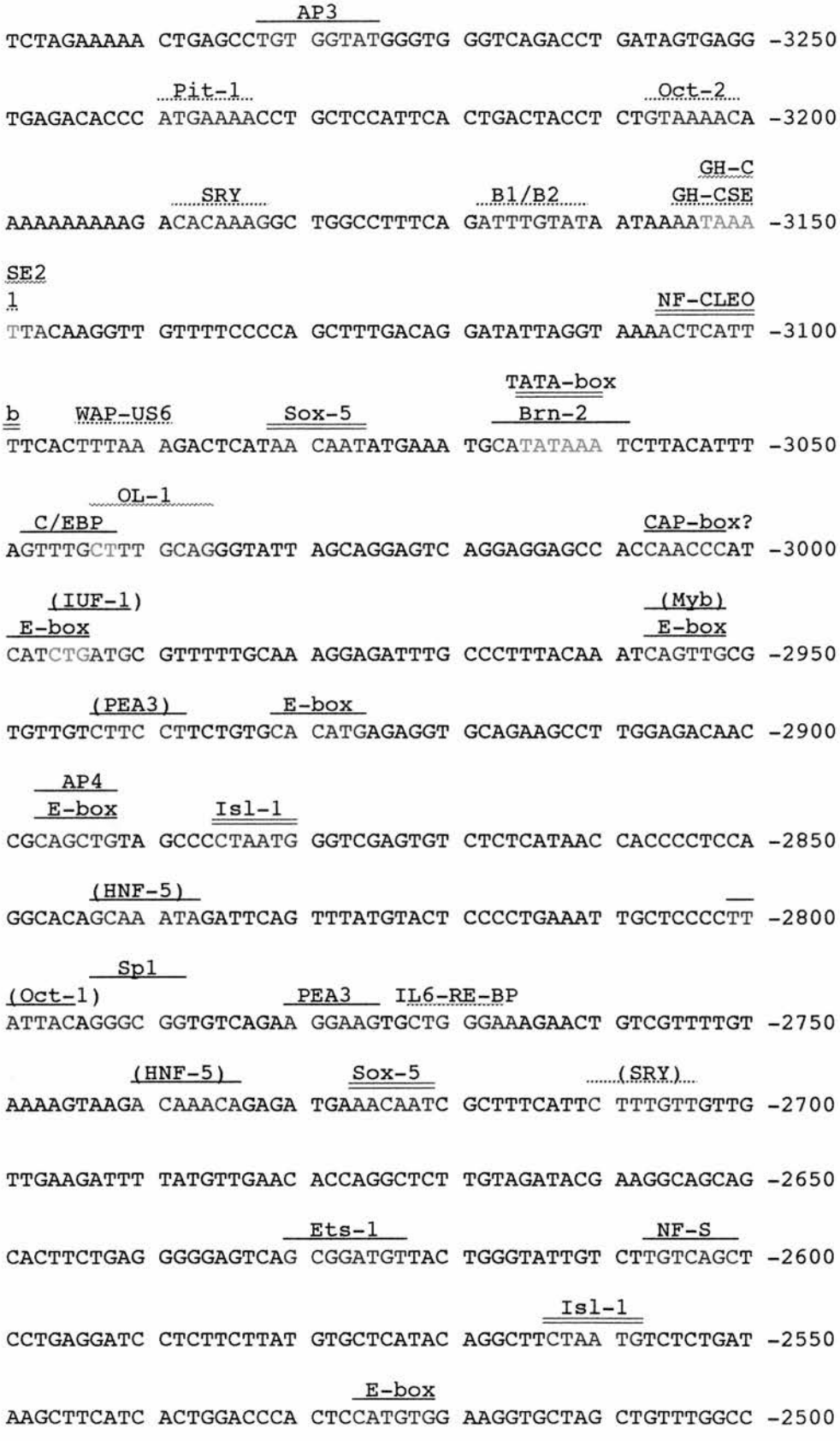
6.3.6 Search for extended regions of homology

Selected searches for extended regions of homology between sites in the putative *Vipr2* promoter and the promoter sequences of other genes that are expressed in *Vipr2* expressing cells or tissues, were carried out using the Genetics Computer Group program 'Wordsearch' (HGMP Resource Centre, Cambridge, UK). In most cases the stimulus for these comparisons came from the identification in the *Vipr2* gene of one or more potential transcription factor binding sites (during the computer database searches described above), which were identical to, or displayed a very high level of similarity to, a site for a tissue specific or hormonally regulated transcription factor in another gene. These sites included: the putative isl-1 sites of the *Vipr2* gene, which prompted a comparison of the *Vipr2* and insulin gene promoter sequences; and the possible binding site for component b of the NF-CLE0 complex, which led to a comparison of the *Vipr2* promoter with the promoter regions of the T-cell expressed interleukin-4 (IL-4) and interleukin-5 (IL-5) genes. However in some cases, the routes by which the regions of similarity with the putative *Vipr2* promoter were identified were less straightforward. For example, the similarity between the *Vipr2* promoter sequence and the binding site for a pituitary specific factor PP1 in the rat pro-opiomelanocortin gene, was noticed largely by chance during a survey of papers on pituitary specific gene expression. Consequently, the results presented here do not in any way represent a comprehensive search for regions of homology between the putative *Vipr2* promoter and the promoter

Figure 6.7: Putative transcription factor binding sites within the 5' flanking region of the *Vipr2* gene

(Putative binding sites are shown in green or dark blue, and overlaps between any two given sites are usually shown in light blue/aqua. TATA box and CCAAT box consensus sequences are shown in red unless they form part of a binding site overlap. Sites that match exactly or show a high level of homology, to a well defined consensus sequence are double underlined; those that match slightly less well defined sequences or have one or two mismatches are single underlined; and sites that represent transcription factor consensus sequences that are highly degenerate have been indicated with a dotted line. The sequence shown in dark blue directly upstream of the ATG start site is the published 5' untranslated sequence of *Vipr2*, whereas the extended sequence shown in aqua, from -53 to -106, represents the sequence derived from the Vipr2 5' RACE product Q12)

Putative promoter region of the *Vipr2* gene



GTCATTGCTG ACCCAAAACT TCACTGGCAA AAAAGCAAAG GGACTCTGTG -2450

(Myc) (YY1)
ACCACATCCT AAGAGAACAG AGGGTTAGCT AAAATGGCTT GTGAAGAACT -2400

triplet repeat
E-box
CTGCCCCAAA TGCTGCTGCT GCTGCTGCTT TGGGTATGAG AAGCCTCCGT -2350

E-
IL6-RE-BP TTF-1
GATAGTGAGG ACTGAAACCC TARAGTCTGG AAACCAGAGC AGCCACTCAA -2300

box
PEA3 Ets-1
GTGGAAGGTG CCAGGAAACT GCCCAAGAGG ATGTCAGTGG GGAAGGGATA -2250

PEA3 IUF-1
GGAAGGCCAG ACATCAGTAG GTAGAGATGA AACGGGGCTA ARAGCAGGTC -2200

Ets-1
TGGAGGATGC TGGAGGGCAC AACACAAGGC TGAACTCAGG TTCATATCCA -2150

IUF-1
Sox-5
GGCTGTGAAC AATGTGAAGT CCCATCAGAA CCAAGCAGTC ACCATAGGCA -2100

(GAGA) (GAGA) (GAGA)
CAAGAGTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC -2050

Pit-1
(GAGA) CCAAT
TCTCTCTCTC TCTCCWWACC CCAATTCAGA ARAATGGAGA GACCTTGAAA -2000

PEA3
GACATGACCA GTCTTTTTAG GACCTTCGYT CYTCCAGGAT ACCCCCAGGA -1950

AAACAAGTCC YTCTTGTCTC CCCGTGCCTG TATCTCATCC TTGTCAGGAC -1900

TATCACTTCA TCTCATCATA TATGCCATTG TGTGGTTGGA ACCCTGGCGT -1850

CCAAT (Ets-1) (Pit-1)
CTCACCCAGT TTTGACCAAT AACATCCGGA CGTGACTTTA GAATGCATCC -1800

E-box
Pit-1
Sox-5 (C/EBP) IL6-
CTCACAACAA TTCAGGTGTA TCAGCTCTCC TAATTCTAAT TGTAACCCTC -1750

NF-IL6
RE-BP (AP2) E-box E-box
TGGGATGGAG CAATGGCAGG GGTCGATGTT GTGATTTACA AATGTGCAGG -1700

TGGATCAGGT TCACTGGCTG AAACAGTGGA CAGGGGTTGT GGACAAAGCC -1650

(Ets-1)
AGAGCCCAAG CTTCCTGGCC CCAAGCTGCA GACCTAGTGC CCAGACCTTC -1600

(IUF-1) (HNF-5) E-box
TCCTCCATAA AATCCCCTGA TGGACAAACA CAGGCAGATG CACAGATAAT -1550

GGAAGGGTGA CAGAAAACAG ATAGAAACCT AGGTCCAAGG GTACACGGGG -1500

CARE
(IUF-1) E-box c-Jun
CTCTGATGTC ATGGAGCACA TGATGAGCTC AGTGGGGAGA AAGATAATGG -1450

GCAGATAGAC ACCCCAAAGT GTGAAGGATA CATTTTGTAC TATTTCAGAC -1400

WAP-US6
TTCTGACTTT GGGCTTGAAA TTATATYTGA ATAAGAAGTT TAAAACTTGC -1350

CTGCTTTACT TAATCGAGTA CTATTGTYTG CTGGTAGCTT TYTGCTYTTC -1300

CCTAGACTGG TCTGTCGACA GACAGCCAAA CCATAGCTGA ACAGTGGCAG -1250

E-box E-box
TGTYTCCACA AGCCTAGTGT ACAATGAGCC ACACACGTGT CACCTGAGGG -1200

TTTGGGGTGC CGATGAGCAG ACCCACTTTT TTTTTCTTTT TTTTTTTTTC -115

(Brn-2)
(Oct-1) CCAAT
CGTTTTTTAT TAGGTATTTA GCTCATTTAC ATTTCCAATG CTATACCAAA -1100

AP2
AP2 Yi?
AGTCCCCCCA TACCCAACCC AACCCCACTC CCNTACCCAC CCACTCCCCT -1050

SRE TATA-box C
TTTTTGGCCC TGGCGTTCCC CTGTACTGGG GCATATAAAG TTTGCGTGTC -1000

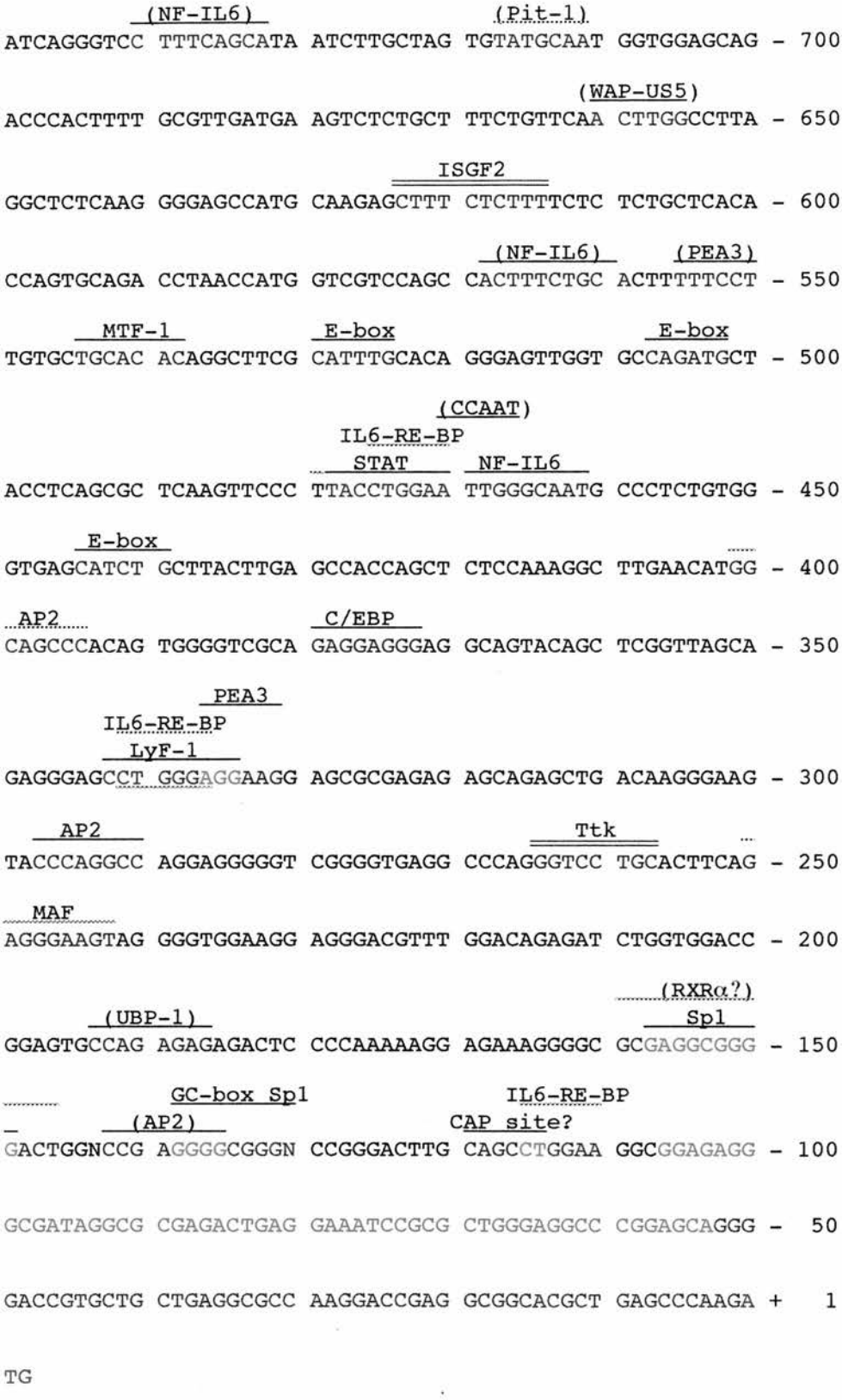
CAP-box?
CAAT (IL6-RE-BP) YY1 E-
CAATGGGCCT CTCTTTCCAG TGATGGCCGC CTAGGCCATC TTTTGATACA - 950

box
TATGCAGCTA GAGTCAAGAG CTCCGGGGTA CTGGTTAGTT CATAATGTTG - 900

TTCCACCTAT AGGGTTGCAG ATCCCTTTAG CTTCTTGGGT ACTTTCTCTA - 850

(CCAAT) CCAAT
GCTCCTCCAT TGGGGGCCCT GTGATCCATC CAATAGCTGA CTGTGAGCAT - 800

HNF-5
CCACTTCTGT GTTTGCTAGG CCCAGGCATA YTCTGACAAG AGACAGCTAT - 750


(NF-IL6)
(Pit-1)
ATCAGGGTCC TTTCAGCATA ATCTTGCTAG TGTATGCAAT GGTGGAGCAG - 700
(WAP-US5)
ACCCACTTTT GCGTTGATGA AGTCTCTGCT TTCTGTTCAA CTTGGCCTTA - 650
ISGF2
GGCTCTCAAG GGGAGCCATG CAAGAGCTTT CTCTTTTCTC TCTGCTCACA - 600
(NF-IL6)
(PEA3)
CCAGTGCAGA CCTAACCATG GTCGTCCAGC CACTTTCTGC ACTTTTTCCT - 550
MTF-1
E-box
E-box
TGTGCTGCAC ACAGGCTTCG CATTTGCACA GGGAGTTGGT GCCAGATGCT - 500
(CCAAT)
IL6-RE-BP
STAT
NF-IL6
ACCTCAGCGC TCAAGTTCCC TTACCTGGAA TTGGGCAATG CCCTCTGTGG - 450
E-box
GTGAGCATCT GCTTACTTGA GCCACCAGCT CTCCAAAGGC TTGAACATGG - 400
AP2
C/EBP
CAGCCCACAG TGGGGTCGCA GAGGAGGGAG GCAGTACAGC TCGGTTAGCA - 350
PEA3
IL6-RE-BP
LyF-1
GAGGGAGCCT GGGAGGAAGG AGCGCGAGAG AGCAGAGCTG ACAAGGGAAG - 300
AP2
Ttk
TACCCAGGCC AGGAGGGGGT CGGGGTGAGG CCCAGGGTCC TGCACTTCAG - 250
MAF
AGGGAAGTAG GGGTGGAAGG AGGGACGTTT GGACAGAGAT CTGGTGGACC - 200
(RXRα?)
(UBP-1)
Sp1
GGAGTGCCAG AGAGAGACTC CCCAAAAAGG AGAAAGGGGC GCGAGGCGGG - 150
GC-box Sp1
IL6-RE-BP
(AP2)
CAP site?
GACTGGNCCG AGGGGCGGGN CCGGGACTTG CAGCCTGGAA GGCGGAGAGG - 100
GCGATAGGCG CGAGACTGAG GAAATCCGCG CTGGGAGGCC CGGAGCAGGG - 50
GACCGTGCTG CTGAGGCGCC AAGGACCGAG GCGGCACGCT GAGCCCAAGA + 1
TG

regions of other genes, but instead comprise a very restricted selection of similarities between the putative *Vipr2* promoter and the promoter regions of some genes whose products have physiological functions/distribution patterns that may be linked to those of the VIP_2 receptor.

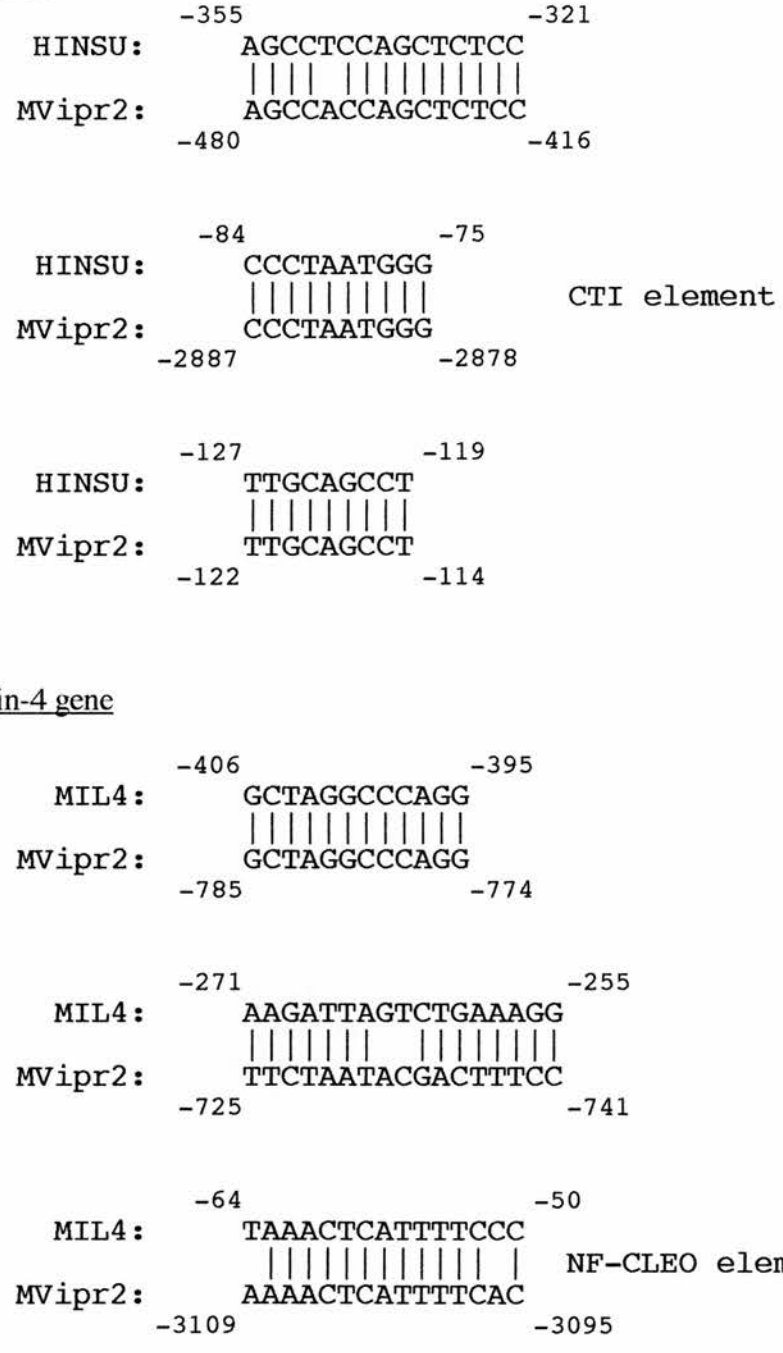
A total of 10 regions of sequence similarity were identified using the "Wordsearch" program (Figure 6.8). Three regions of similarity were identified between the *Vipr2* sequence and the promoter region of the human insulin gene, one of which represented an extended region of homology betweeen one of the *Vipr2* gene putative isl-1 sites and a known transcription factor binding site (CT1 element) of the insulin gene (Petersen *et al.*, 1994). Several regions of the *Vipr2* sequence were found to display a high degree of similarity with sequences from the promoter regions of the mouse IL-4 and IL-5 genes. These included: the sequence encompassing the putative binding site for the b-component of the NF-CLE0 complex, which was found to exhibit strong similarity to the NF-CLE0 (a and b) elements of both the IL-4 and IL-5 genes (but particularly that of the IL-4 gene, (Miyatake *et al.*, 1991); Figure 6.8); two other regions of the IL-4 gene's promoter which were found to be similar to sequences that lay within the *L1*-repeat sequence of the *Vipr2* gene; and a shorter (8bp) region of the mouse IL-5 gene which displayed sequence identity to a site in the distal region of the putative *Vipr2* promoter. Single regions of sequence similarity with the putative *Vipr2* promoter were found in the promoter regions of the *VIP*, pro-opiomelanocortin, and epidermal growth factor receptor genes, in the form of: a previously identified transcription factor binding site (STAT) site in the promoter of the *VIP* gene (Symes *et al.*, 1994); an element composed of adjacent Sp1/pituitary factor (PP1) site in the promoter of the rat pro-opiomelanocortin gene (Liu *et al.*, 1995) ; and a composite Sp1/thyroid hormone receptor•retinoid X receptor element in the promoter of the human epidermal growth factor receptor gene (Xu *et al.*, 1993).

6.3.7 Identification of potential translational- and splice-factor binding sites

Having identified 3 potential core promoter regions within the 5' region of the *Vipr2* gene, a search for alternative translational start sites (upstream ATGs), and potential splice site sequences, was undertaken in an attempt to gain some indication of whether alternative promoter usage might occur within the *Vipr2* gene, and to examine the possibility that previously unidentified introns could be present within the D1X8 sequence. In total, 23 ATG sequences were found within the ~3.2 kb sequence (Figure 6.9), but only 6 of these sites were located within a strong translational context (RNNatgG; (Kozak, 1996)). No translational starts were found within 400 bp of the

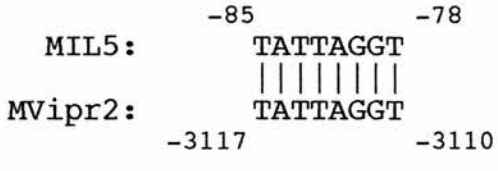
Figure 6.8: Selected regions of homology between the *Vipr2* putative promoter region and the promoter regions of other genes

Human insulin gene

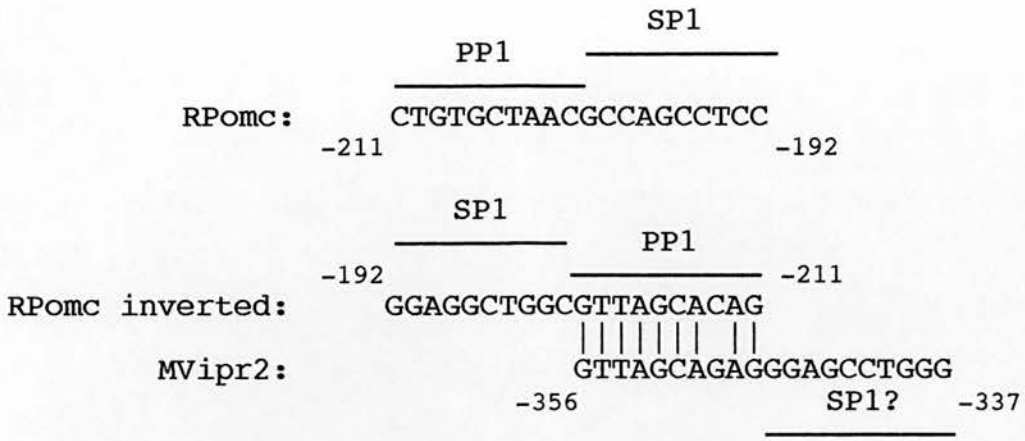


Mouse interleukin-4 gene

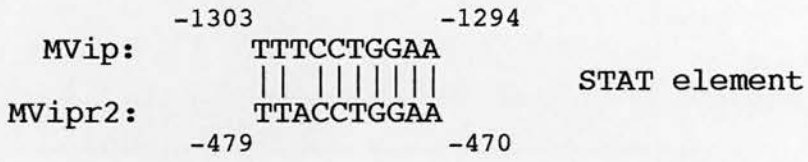
Mouse interleukin-5 gene



Rat pro-opiomelanocortin gene



Mouse VIP gene



Epidermal growth factor receptor gene

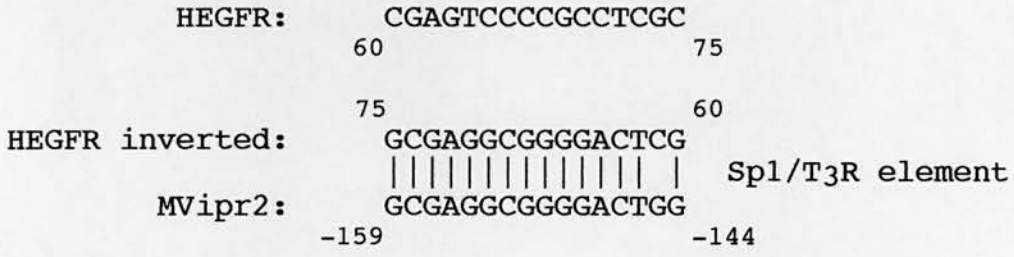


Figure 6.9: A search for potential translational start sites and splice sites within the putative promoter region of the *Vipr2* gene

(ATG sequences that are in a relatively strong translational context are shown in dark blue, and those that are in a very strong context are marked with an asterisk. ATG sequences that are in a weaker translational context are shown in light blue. Potential 3' splice sites (acceptor sites) are shown in pink, and potential 5' splice sites (donor sites) are shown in green. Sites that match the acceptor and donor consensus sequences relatively weakly are indicated by a question mark. The regions shown in light and dark blue directly upstream of the ATG start site, i.e., -1 to -106 are 5' untranslated sequence)

Putative promoter region of the *Vipr2* gene

```
TCTAGAAAAA CTGAGCCTGT GGTATGGGTG GGTCAGACCT GATAGTGAGG -3250

TGAGACACCC ATGAAAACCT GCTCCATTCA CTGACTACCT CTGTAAAACA -3200

AAAAAAAAAG ACACAAAGGC TGGCCTTTCA GATTTGTATA ATAAAATAAA -3150

                ?
TTACAAGGTT GTTTTCCCCA GCTTTGACAG GATATTAGGT AAAACTCATT -3100

TTCACTTTAA AGACTCATAA CAATATGAAA TGCATATAAA TCTTACATTT -3050

AGTTTGCTTT GCAGGGTATT AGCAGGAGTC AGGAGGAGCC ACCAACCCAT -3000

CATCTGATGC GTTTTTGCAA AGGAGATTTG CCCTTTACAA ATCAGTTGCG -2950

TGTTGTCTTC CTTCTGTGCA CATGAGAGGT GCAGAAGCCT TGGAGACAAC -2900

                                                  ?
CGCAGCTGTA GCCCCTAATG GGTCGAGTGT CTCTCATAAC CACCCCTCCA -2850

                                                  ?
GGCACAGCAA ATAGATTCAG TTTATGTACT CCCCTGAAAT TGCTCCCCTT -2800

ATTACAGGGC GGTGTCAGAA GGAAGTGCTG GGAAAGAACT GTCGTTTTGT -2750

AAAAGTAAGA CAAACAGAGA TGAAACAATC GCTTTCATTC TTTGTTGTTG -2700

TTGAAGATTT TATGTTGAAC ACCAGGCTCT TGTAGATACG AAGGCAGCAG -2650

CACTTCTGAG GGGGAGTCAG CGGATGTTAC TGGGTATTGT CTTGTCAGCT -2600

CCTGAGGATC CTCTTCTTAT GTGCTCATAC AGGCTTCTAA TGTCTCTGAT -2550

AAGCTTCATC ACTGGACCCA CTCCATGTGG AAGGTGCTAG CTGTTTGGCC -2500

GTCATTGCTG ACCCAAAACT TCACTGGCAA AAAAGCAAAG GGACTCTGTG -2450

                                  *
ACCACATCCT AAGAGAACAG AGGGTTAGCT AAAATGGCTT GTGAAGAACT -2400

CTGCCCCAAA TGCTGCTGCT GCTGCTGCTT TGGGTATGAG AAGCCTCCGT -2350
```

GATAGTGAGG ACTGAAACCC TARAGTCTGG AAACCAGAGC AGCCACTCAA -2300

GTGGAAGGTG CCAGGAAACT GCCCAAGAGG ATGTCAGTGG GGAAGGGATA -2250

GGAAGGCCAG ACATCAGTAG GTAGAGATGA AACGGGGCTA ARAGCAGGTC -2200

TGGAGGATGC TGGAGGGCAC AACACAAGGC TGAACTCAGG TTCATATCCA -2150

GGCTGTGAAC AATGTGAAGT CCCATCAGAA CCAAGCAGTC ACCATAGGCA -2100

CAAGAGTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC -2050

*

TCTCTCTCTC TCTCCWWACC CCAATTCAGA ARAATGGAGA GACCTTGAAA -2000

GACATGACCA GTCTTTTTAG GACCTTCGYT CYTCCAGGAT ACCCCCAGGA -1950

AAACAAGTCC YTCTTGTCTC CCCGTGCCTG TATCTCATCC TTGTCAGGAC -1990

TATCACTTCA TCTCATCATA TATGCCATTG TGTGGTTGGA ACCCTGGCGT -1850

CTCACCCAGT TTTGACCAAT AACATCCGGA CGTGACTTTA GAATGCATCC -1800

CTCACAACAA TTCAGGTGTA TCAGCTCTCC TAATTCTAAT TGTAACCCTC -1750

TGGGATGGAG CAATGGCAGG GGTCGATGTT GTGATTTACA AATGTGCAGG -1700

TGGATCAGGT TCACTGGCTG AAACAGTGGA CAGGGGTTGT GGACAAAGCC -1650

AGAGCCCAAG CTTCCTGGCC CCAAGCTGCA GACCTAGTGC CCAGACCTTC -1600

*

TCCTCCATAA AATCCCCTGA TGGACAAACA CAGGCAGATG CACAGATAAT -1550

GGAAGGGTGA CAGAAAACAG ATAGAAACCT AGGTCCAAGG GTACACGGGG -1500

*

CTCTGATGTC ATGGAGCACA TGATGAGCTC AGTGGGGAGA AAGATAATGG -1450

GCAGATAGAC ACCCCAAAGT GTGAAGGATA CATTTTGTAC TATTTCAGAC -1400

TTCTGACTTT GGGCTTGAAA TTATATYTGA ATAAGAAGTT TAAAACTTGC -1350

CTGCTTTACT TAATCGAGTA CTATTGTYTG CTGGTAGCTT TYTGCTYTTC -1300

CCTAGACTGG TCTGTCGACA GACAGCCAAA CCATAGCTGA ACAGTGGCAG -1250

TGTYTCCACA AGCCTAGTGT ACAATGAGCC ACACACGTGT CACCTGAGGG -1200

TTTGGGGTGC CGATGAGCAG ACCCACTTTT TTTTTCTTTT TTTTTTTTTC -1150

CGTTTTTTAT TAGGTATTTA GCTCATTTAC ATTTCCAATG CTATACCAAA -1100

AGTCCCCCCA TACCCAACCC AACCCCACTC CCNTACCCAC CCACTCCCCT -1050

TTTTTGGCCC TGGCGTTCCC CTGTACTGGG GCATATAAAG TTTGCGTGTC -1000

CAATGGGCCT CTCTTTCCAG TGATGGCCGC CTAGGCCATC TTTTGATACA - 950

TATGCAGCTA GAGTCAAGAG CTCCGGGGTA CTGGTTAGTT CATAATGTTG - 900

TTCCACCTAT AGGGTTGCAG ATCCCTTTAG CTTCTTGGGT ACTTTCTCTA - 850

GCTCCTCCAT TGGGGGCCCT GTGATCCATC CAATAGCTGA CTGTGAGCAT - 800

CCACTTCTGT GTTTGCTAGG CCCAGGCATA YTCTGACAAG AGACAGCTAT - 750

ATCAGGGTCC TTTCAGCATA ATCTTGCTAG TGTATGCAAT GGTGGAGCAG - 700

ACCCACTTTT GCGTTGATGA AGTCTCTGCT TTCTGTTCAA CTTGGCCTTA - 650

GGCTCTCAAG GGGAGCCATG CAAGAGCTTT CTCTTTTCTC TCTGCTCACA - 600

*

CCAGTGCAGA CCTAACCATG GTCGTCCAGC CACTTTCTGC ACTTTTTCCT - 550

TGTGCTGCAC ACAGGCTTCG CATTTGCACA GGGAGTTGGT GCCAGATGCT - 500

ACCTCAGCGC TCAAGTTCCC TTACCTGGAA TTGGGCAATG CCCTCTGTGG - 450

*

GTGAGCATCT GCTTACTTGA GCCACCAGCT CTCCAAAGGC TTGAACATGG - 400

CAGCCCACAG TGGGGTCGCA GAGGAGGGAG GCAGTACAGC TCGGTTAGCA - 350

GAGGGAGCCT GGGAGGAAGG AGCGCGAGAG AGCAGAGCTG ACAAGGGAAG - 300

TACCCAGGCC AGGAGGGGGT CGGGGTGAGG CCCAGGGTCC TGCACTTCAG - 250

AGGGAAGTAG GGGTGGAAGG AGGGACGTTT GGACAGAGAT CTGGTGGACC - 200

GGAGTGCCAG AGAGAGACTC CCCAAAAAGG AGAAAGGGGC GCGAGGCGGG - 150

GACTGGNCCG AGGGGCGGGN CCGGGACTTG CAGCCTGGAA GGCGGAGAGG - 100

GCGATAGGCG CGAGACTGAG GAAATCCGCG CTGGGAGGCC CGGAGCAGGG - 50

GACCGTGCTG CTGAGGCGCC AAGGACCGAG GCGGCACGCT GAGCCCAAGA + 1

TGAGGGCGTC GGTGGTGCTG ACCTGCTACT GCTGGTTGCT GGTGCGG + 48

previously identified identified start site (+1). However, strong potential translational sites were found at -403 and -582 bp, each of which could represent the beginning of a short ORF (encoding 43 and 50 amino acids respectively). The search for potential splice sites within the D1X8 sequence resulted in the identification of several potential 5' and 3' sites (Figure 6.9), although the degree to which these sites corresponded to consensus splice site sequences (5' MAG↓GTRAGT; 3' Y_{11}NYAG↓G; (Mount, 1982)) varied considerably.

6.4 Discussion

6.4.1 General characteristics of the mouse *Vipr2* gene's putative promoter region

The proximal region of the putative Vipr2 promoter is GC rich and CpG rich

As a first step towards characterising the promoter region(s) of the mouse *Vipr2* gene, a 3.2 kb fragment of the upstream region of the mouse *Vipr2* gene has been isolated and sequenced. Like the promoter region of the *VIPR1* gene (Sreedharan *et al.*, 1995), and other members of the secretin/glucagon family of receptor genes, the putative promoter region of the *Vipr2* gene has a proximal region that is GC-rich and contains binding sites for the ubiquitous transcription factor Sp1 (located at -138 bp and -157 bp). Although analysis of the methylation status of this region of *Vipr2* gene has not been carried out, the high CpG content of the sequence beginning about 160 bp upstream of the proposed translational start site and extending into the first intron, strongly suggests that an unmethylated CpG island is present at the 5' end of the *Vipr2* gene. This would place the *Vipr2* gene amongst the ~40% of genes isolated to date that have been found to possess both CpG islands and tissue-specific patterns of expression (reviewed by (Cross & Bird, 1995)). Moreover, as CpG islands are associated with the region surrounding the first exon of a gene (Cross & Bird, 1995), the location of this CpG island within the *Vipr2* gene adds weight to the identity of the proposed transcriptional start site (at -106 bp) which lies within the CpG rich sequence. The published promoter sequence (814 bp) of the human *VIPR1* gene (Sreedharan *et al.*, 1995) also appears to contain a CpG island.

Tissue specific promoter or enhancer sequences may play a role in the regulation of Vipr2 gene expression

Considering the complex patterns of tissue- and cell- specific expression that are observed for transcripts of the *Vipr1* and *Vipr2* genes (Sheward *et al.*, 1995; Usdin *et al.*, 1994; Wei & Mojsov, 1996), there appears to be a remarkable absence of putative binding sites for tissue-restricted transcription factors within the GC rich promoter region of either the *VIPR1* gene (Sreedharan *et al.*, 1995) or to a slightly lesser extent the *Vipr2* gene, when these sequences are subjected to searches for known transcription factor binding sites. Possible explanations for these findings include: the presence of binding sites for as yet unidentified tissue-specific transcription factors; the presence of novel binding sites for known tissue-specific transcription factors; tissue specific enhancer regions that lie upstream of and interact with the GC rich promoter regions; the presence of alternative promoter regions within the genes; or any combination of these possibilities. Alternative promoter usage has been identified in at least one other gene encoding a member of this subfamily of G-protein coupled receptors, the mouse parathyroid hormone/parathyroid hormone related peptide (PTH/PTHrP) receptor gene which has two alternative promoter regions and three 5' untranslated exons (McCuaig *et al.*, 1995). The proximal promoter region P2, which is linked to untranslated exon U3, is GC rich and appears to be used in all tissues examined, whereas the distal promoter region P1, which is linked to untranslated exon U1 and spliced to a second untranslated exon U2, is not GC rich, contains a CCAAT box, and is used predominantly in the kidney (McCuaig *et al.*, 1995). However, there are indications that this type of process may also occur in other members of the secretin/glucagon receptor gene sub-family, as there is evidence to suggest that alternative splicing of untranslated 5' exons is utilised in a tissue specific fashion in the rat glucagon receptor gene (Maget *et al.*, 1996b), and alternative splicing of untranslated 5' exons within the rat *Pacapr1* gene has also been reported (Chatterjee *et al.*, 1996a). In both the mouse PTH/PTHrP receptor gene and the rat *Pacapr1* gene a consensus 3' splice acceptor sequence is located in the region immediately upstream of the ATG start site (76 bp upstream in the case of the *Pacapr1* gene, and 45 bp upstream in the PTH/PTHrP receptor gene; (Chatterjee *et al.*, 1996a; McCuaig *et al.*, 1995)). Two TATA box consensus sequences were identified in the *Vipr2* gene ~1 kb and ~3 kb upstream of the ATG start site, but the distance of these putative core promoter sequences from the ATG start site would suggest that in order for these putative promoter regions to be functional these sequences and/or any 5' untranslated exons associated with them, must be spliced into place upstream of the ATG start site. However no consensus 3'

splice acceptor sites were found in the region immediately upstream of the ATG start site in the *Vipr2* gene, where the most proximal potential 3' splice site that was identified was found 596 bp upstream of the translation start site. Thus if alternative promoter usage is a feature of the *Vipr2* gene, the exact mechanisms by which this occurs are not yet apparent. Interestingly a TATA box consensus sequence has also been identified within the promoter region of the *VIPR1* gene (at -310) although its position relative to the identified transcriptional start site in HT29 cells led the authors to conclude that the TATA box sequence was unlikely to be functional in this cell line (Sreedharan *et al.*, 1995). Nevertheless, despite uncertainties over the functional status of the TATA box sequences in distal region of the D1X8 sequence, the presence of several potential binding sites for tissue-restricted transcription factors within the areas surrounding these sites indicates that regions upstream of the GC rich proximal promoter sequence may well play a role in the control of the *Vipr2* gene in some cells, either as alternative promoter regions, or by acting as enhancer sequences which interact with the proximal core promoter sequences.

6.4.2 Tissue-restricted transcription factors that may bind to sites in the 5' regulatory region of the mouse *Vipr2* gene

Conserved lymphokine element 0 binding factors may play a role in regulation Vipr2 expression within the immune system

One potential binding site that is of particular interest with regard to the possible role of tissue-specific alternative promoter or upstream enhancer sequences, is a site at - 3109 bp to - 3095 bp that displays a high level of similarity to the conserved lymphokine element 0 (CLE0). CLE0 sites are found in the promoter regions of several lymphokine genes that are expressed in T helper 2 (Th2) lymphocytes, including those of the mouse and human genes encoding IL-4, IL-5, and GM-CSF (Miyatake *et al.*, 1991). The identity of the transcription factors that bind to the CLE0 elements of these genes *in vivo* is not entirely clear, although several factors including activator protein 1 (AP-1), nuclear factor of activated T cells (NF-ATp), and the Ets protein Elf-1 are candidates (Jenkins *et al.*, 1995; Siegel *et al.*, 1995). However, in all of these promoter regions the CLE0 element is located within 40 bp of the TATA box (Miyatake *et al.*, 1991), appears to be required for the maximal activation of the promoters by phorbol myristate acetate (PMA) and calcium ionophores (Jenkins *et al.*, 1995), and is thought to be involved in co-ordinating the stimulation of transcription of these genes in antigen-activated T-cells (Miyatake *et al.*, 1991).

Within the immune system VIP is involved in several processes, including the regulation of cytokine expression (see Chapter 1; reviewed by (Ganea, 1996)). VIP decreases IL-2 and IL-10 production at the transcriptional level (Ganea *et al.*, 1996; Ganea & Sun, 1993), decreases IL-4 production at a post transcriptional level (probably through its effects on IL-2) (Sun & Ganea, 1993; Wang *et al.*, 1996), and increases IL-5 production (Mathew *et al.*, 1992). Although the individual roles of the VIP_1 and VIP_2 receptors in some of these processes have not yet been ascribed, current reports indicate that the VIP_1 receptor may be responsible for the effects on IL-10 levels, whilst the VIP_2 receptor is thought to be responsible for the down regulation of IL-2 (the primary cytokine responsible for T-cell proliferation following T-cell receptor activation) (Delgado *et al.*, 1996b).

Therefore the presence of a site within the *Vipr2* gene, that displays a high level of homology to the CLE0 element (13/15 base match with the mouse IL-4 gene CLE0 element; Figure 6.8) and is centred 37 bp upstream of the distal (-3065 bp) TATA box sequence is intriguing though not entirely surprising in light of the proposed functions of VIP in the immune system, and could be interpreted as supporting the possibility that this TATA box might be functional under certain circumstances. Further indications that the CLE0-like sequence within the *Vipr2* gene may be functional are provided by recent work carried out by Delgado *et al.*, who have shown that *Vipr2* gene transcription in $CD4^+$ and $CD8^+$ T-cells is only induced following T-cell activation, and requires two signals (anti-CD3 and PMA), a requirement which the authors noted was similar to the dual signal requirement for cytokine production (Delgado *et al.*, 1996b).

It may also be significant that the 8 bp sequence TATTAGGT which lies directly adjacent to the 5' end of the putative CLE0 element in the *Vipr2* gene, is also found near to the CLE0 element in the promoter region of the mouse IL-5 gene. In the mouse IL-5 gene this octamer sequence is situated 22 bp upstream of the CLE0 element and spans the region from -83 to -76 bp (Siegel *et al.*, 1995). Although the transcription factor(s) that bind to this region of the mouse IL-5 gene have not been identified, a DNase I footprint is detected between -88 and -77 bp in the presence of nuclear extracts from the mouse thymoma cell line EL-4, and a reporter construct containing the region from -76 bp to +1 of the IL-5 promoter showed only ~40-50% of the activity observed in constructs containing the -91 bp to +1 region following stimulation with Bt_2cAMP and PMA (Siegel *et al.*, 1995). Two other regions of the *Vipr2* gene situated upstream of the putative CLE0 element, also demonstrated high levels of homology with sequences from the promoter regions of lymphokine genes (Figure 6.8), as did other

areas of the putative *Vipr2* promoter, but the importance of these sequences, if any, within either the lymphokine or *Vipr2* promoter regions, has yet to be demonstrated.

Nevertheless, although functional analysis of this distal region of the *Vipr2* promoter will obviously be required to substantiate a role for this area of the putative promoter region in the cell-specific expression of the *Vipr2* gene, it is tempting to speculate that the CLE0-like element and possibly other sites within the upstream region of the *Vipr2* gene may contribute to the co-induction of transcription of *Vipr2* and lymphokine genes following antigen stimulation of T-cells.

Brain-2-like factors as candidates for a role in the control of Vipr2 expression within the CNS

Other possible binding sites for tissue-restricted transcription factors, that were of interest with regard to potential functions of the VIP_2 receptor, included a site that conformed to the consensus binding sequence for the POU transcription factor Brain-2 (Brn-2) and overlapped the distal (-3065 bp) TATA box sequence of the putative *Vipr2* promoter. Brn-2, which belongs to the class III group of POU transcription factors, is found mainly within the CNS (He *et al.*, 1989), although its presence in melanocytes (Angus *et al.*, 1995; Eisen *et al.*, 1995) and in small cell lung cancer cells (Schreiber *et al.*, 1992) has also been reported. The distribution patterns of individual POU-III transcription factors within the developing brain have led to the proposal that these factors are involved in the differentiation of neuronal subtypes (Alvarez-Bolado *et al.*, 1995; He *et al.*, 1989). However, these transcription factors are also expressed in a more restricted pattern in adult brain (He *et al.*, 1989). Evidence that Brn-2 is important for neuronal development initially came from studies of embryonal carcinoma cells, which indicated that Brn-2 is required for but is not sufficient for neuronal differentiation (Fujii & Hamada, 1993), and subsequent gene targeting studies have shown that Brn-2 is essential for the normal development of the hypothalamic-pituitary axis (Nakai *et al.*, 1995; Schonemann *et al.*, 1995). In homozygous Brn-2-null mutant mice, the paraventricular nucleus (PVN) and supraoptic nucleus (SON) of the hypothalamus are absent, and the posterior lobe of the pituitary is hypoplastic (Nakai *et al.*, 1995; Schonemann *et al.*, 1995). The precise cause of these defects (which first start to become apparent at ~E12.5) is not known, although the inability of Brn-$2^{-/-}$ mutants to activate the transcription of certain neuropeptide genes (including oxytocin, arginine vasopressin, CRF) in the PVN, and SON, or to form correct axonal projections from the hypothalamus to the posterior lobe of the pituitary, appear to be contributing factors (Nakai *et al.*, 1995; Schonemann *et al.*, 1995).

Vipr1 mRNA has not been reported to be present in the hypothalamus, but *Vipr2* transcripts have been detected in several nuclei including the PVN of the hypothalamus (particularly the posterior magnocellular region and the parvocellular division of the periventricular region), and the supraoptic nucleus (Sheward *et al.*, 1995; Usdin *et al.*, 1994). Thus, *Vipr2* is expressed within regions of the hypothalamus where *Brn-2* is also expressed, and it seems entirely possible that Brn-2 could play an important role in the hypothalamic expression of the *Vipr2* gene. Target sites for Brn-2 that have been identified to date include several high affinity binding sites within the CRF gene's promoter (Peng *et al.*, 1993), and a possible binding site in the promoter of the calcitonin/CGRP gene (Schreiber *et al.*, 1992). However, it is interesting that although the *Brn-2* knockout studies indicate that Brn-2 may be involved in the regulation of oxytocin and vasopressin gene expression, the authors noted that these peptides are not found in neuronal precursor cells at E12.5 in wild type mice, and go on to suggest that the Brn-2 target genes essential for the survival of the migrating hypothalamic neurons "might include genes encoding a receptor for neurotrophic factors or a protein with anti-apoptotic activity" (Nakai *et al.*, 1995). The *Vipr2* gene, which is expressed in mouse embryos from at least as early as E9.5 (Sheward *et al.*, 1996b), and whose product binds ligands (VIP and PACAP) that have been shown to be involved in the promotion of neuronal survival (Arimura *et al.*, 1994; Festoff *et al.*, 1996; Pincus *et al.*, 1994; Tanaka & Koike, 1994), would be well placed to fulfil this type of function within the developing hypothalamus.

In addition to the potential Brn-2 binding site at -3067 bp, a second potential site for Brn-2 binding within the putative promoter region of the *Vipr2* gene was found at -1119 bp. This site lies ~100 bp upstream of the more proximal TATA box sequence, within a region that also contains possible sites for serum- and cAMP- responsive factors. Thus if these areas of the *Vipr2* gene do function as alternative promoter regions rather than enhancer sequences, it is conceivable that Brn-2 might block binding of transcription factors to the distal TATA sequence by competing for binding to this site, whilst binding of Brn-2 upstream of the more proximal TATA sequence might possibly promote transcriptional activation from this region. However, it is also important to stress that Brn-2 displays considerable flexibility in its DNA binding specificity, and can bind to bipartite core CAT and A/TAAT motifs that are separated by 0, 2 or 3 bp (Li *et al.*, 1993). Therefore the probability of a sequence that fits the Brn-2 consensus binding sequence, occuring by chance within a given region of DNA is much higher than that for an element such as the CLE0-like sequence that was described earlier, and consequently the suggestion that these sites in the *Vipr2* gene may be functional binding sites for Brn-2 is currently very tentative.

Transcription factors that may be involved in the pituitary expression of the Vipr2 gene

As *Vipr2* transcripts are detected in the pituitary, particularly in the anterior region where a dramatic increase in the level of *Vipr2* expression is seen during pregnancy (Usdin *et al.*, 1994), it was interesting to find that several possible sites for the anterior pituitary specific transcription factor Pit-1 (Bodner *et al.*, 1988; Simmons *et al.*, 1990) were present at various locations in the putative promoter region of the *Vipr2* gene. These sites do not conform to the proposed Pit-1 consensus sequence TATNCAT (Nelson *et al.*, 1988) but do display a high level of similarity, or in several cases identity with Pit-1 binding sites found in the human prolactin gene promoter (Peers *et al.*, 1990). Pit-1 is found in somatotrope, lactotrope, and thyrotrope cells of the pituitary, and is essential for the determination and maintenance of these cell types (Li *et al.*, 1990). However, although Pit-1 is required for the transcriptional activation of several genes encoding peptide hormones and receptors in the pituitary, including the genes for growth hormone and prolactin (Mangalam *et al.*, 1989), thyroid stimulating hormone β (Haugen *et al.*, 1993), and the thyroid hormone β2 receptor (Wood *et al.*, 1996b), Pit-1 alone cannot account for the cell-specific regulation of the transcription of these genes within the pituitary (reviewed by (Rhodes *et al.*, 1994)).

One site that may be involved in the cell-specific regulation of *Vipr2* gene transcription in the pituitary was identified by comparison of the putative *Vipr2* promoter with the promoter region of the rat pro-opiomelanocortin gene. The site in question is located in the GC rich region of the putative *Vipr2* promoter, and is very similar (1/10 bp mismatch in the form of a C→G substitution; Figure 6.8) to a binding site for a chemically unidentified transcription factor named PP1 that plays a major role in the pituitary cell-specific expression of the pro-opiomelanocortin gene (Liu *et al.*, 1995). Furthermore, the pro-opiomelanocortin gene PP1 site to which the *Vipr2* sequence shows homology, appears to be aided in its transcriptional activation of the pro-opiomelanocortin gene by two Sp1 sites one of which lies directly adjacent to the PP1 site, and although the putative PP1 site within the *Vipr2* gene is in an inverted orientation relative to that in the pro-opiomelanocortin gene promoter, it was found that the *Vipr2* PP1 site is also situated directly adjacent to a GC rich sequence that may act as an Sp1 binding site (2/10 bp mismatches with the pro-opiomelanocortin gene Sp1 site, both G↔C substitutions). Transcripts from the pro-opiomelanocortin gene and *Vipr2* gene are both found in the AtT20 pituitary cell line (Liu *et al.*, 1995; Rawlings *et al.*, 1995). However, the apparent restriction of PP1 expression to cells that express the pro-opiomelanocortin gene (Liu *et al.*, 1995), suggests that although PP1 may be involved in cell specific expression of *Vipr2* in some pituitary cell types, other cell-

specific factors must be responsible for directing *Vipr2* gene expression in cell lines such as GH_4C_1 where *Vipr2* transcripts are found (Rawlings *et al.*, 1995) but PP1 is not detected (Liu *et al.*, 1995).

Isl-1-like factors and other E-box binding factors

The promoter region of the insulin gene (which in adults is active only in pancreatic β-cells; (Giddings *et al.*, 1985)) has been extensively studied (reviewed by (Dumonteil & Philippe, 1996), and was thus an ideal template with which to begin to search the putative *Vipr2* promoter for possible binding sites for pancreatic transcription factors. Of the three regions of strong sequence similarity that were identified in the *Vipr2* and insulin promoter sequences (Figure 6.8), the discovery that the putative islet-1 (isl-1) binding site (CTAATG) at ~ -2885 bp (identified in the previous computer database search) displayed an extended homology to the CTI box (CCCTAATGGG) of the human insulin gene, was the most interesting at present, as transcription factor binding to this sequence has been demonstrated in the insulin gene (Petersen *et al.*, 1994). The LIM homeobox transcription factor isl-1 was originally identified in pancreatic islet cells (Karlsson *et al.*, 1990), and is involved in the islet-specific transcription of several genes including the rat proglucagon gene (Wang & Drucker, 1995) and the amylin gene (Wang & Drucker, 1996). However isl-1 has subsequently been found in other endocrine tissues including the kidney, thymus, and pituitary, and also in the nervous system (Dong *et al.*, 1991; Thor *et al.*, 1991) where its functions are thought to include a role in motor neuron development (reviewed by (Lumsden, 1995)). Two potential isl-1 sites exist in the putative *Vipr2* promoter (just over 300 bp from one another), but as the isl-1 like binding sites (known as CT sites) in the insulin gene are now thought to bind the islet-specific homeobox protein insulin promoter factor 1/somatostatin transcription factor 1 (IPF-1/STF-1) (Petersen *et al.*, 1994), and at least 12 different homeobox genes are known to be expressed in pancreatic beta cells (Rudnick *et al.*, 1994), the range of factors that might be capable of binding to the putative isl-1 sites of the *Vipr2* gene is likely to be quite large. Nevertheless, in both the rat and the human insulin gene the CTI motif lies a short distance downstream of a sequence GCCATCTG known as GCI (Boam *et al.*, 1990) which contains the consensus E-box binding site (CANNTG) for the basic-helix-loop-helix family of transcription factors (Blackwell & Weintraub, 1990), and it is interesting that the *Vipr2* putative isl-1 site that is identical to the CTI motif also lies just downstream of an E-box sequence. The E-box sequence near the '*Vipr2* CTI site' is not the same as the GCI sequence of the insulin gene, but as synergistic interactions between helix-loop-helix proteins and IPF-1/STF-1 are involved in the regulation of insulin gene transcription (Peers *et al.*, 1994), the presence

in the *Vipr2* gene of adjacent sites for members of these groups of transcription factors may indicate that similar types of regulatory mechanisms are operating in these genes.

In addition to the E-box sequence mentioned above, several other E-box sequences were identified within the putative promoter region of the *Vipr2* gene. Due to the degenerate nature of the E-box consensus sequence, these sequences will occur by chance on average once every 256 bp. Thus the identification of a total of 18 potential sites within 3.2 kb is not in itself very significant, although the presence of two 'double E-box sites' at -1216 and -1711 may suggest that the E-box binding basic-helix-loop-helix family of transcription factors might be important regulators of *Vipr2* gene transcription. Tissue restricted E-box binding proteins that have been identified in tissues and cells where the *Vipr2* gene is expressed include: myogenic transcription factors such as MyoD and myogenin (Weintraub *et al.*, 1991; Wright *et al.*, 1989); the neuron-specific factors such as the HEN (Brown & Baer, 1994; Brown *et al.*, 1992), and MASH (Guillemot, 1995) subgroups; NeuroD which is found in the nervous system and also in the pancreas (Lee *et al.*, 1995b; Naya *et al.*, 1995); and the slightly more widely distributed E-box factor TAL1 which is thought to be involved in T-cell leukaemia, but is also expressed in the developing brain, hematopoietic cells, and in some endothelial cells (Hsu *et al.*, 1994). The basic-helix-loop-helix (bHLH) proteins bind to E-box sequences as either homodimers or heterodimers, and as each bHLH protein is thought to interact principally with one half of the E-box sequence, different bHLH pairs display binding preferences for different E-box sequences (Blackwell & Weintraub, 1990). However the differences in sequence specificity between the various E-box binding factors are not clear cut, and it is often difficult to predict which bHLH pairs might be capable of binding to a given E-box sequence. This seemed to be particularly true for the putative *Vipr2* promoter as two of the E-box binding factors that were of particular interest with regard to the tissue-specific expression of the *Vipr2* gene, TAL1 and NeuroD, appear to bind to similar E-box sequences. Thus although it was interesting to see that the preferred core binding sequence (5' CAGATG 3') for TAL1/E2A heterodimers (Hsu *et al.*, 1994) occurs at several sites within the putative *Vipr2* promoter (including the two most proximal E-box sites), it is equally possible that these sites may represent regulatory sites for NeuroD or other related E-box factors, as the opposite strand of this sequence (5' CATCTG 3') corresponds to the sequence of the GC-I site of the insulin gene which has been reported to bind heterodimers of the bHLH factors E47 and Beta 2 (the hamster homologue of NeuroD; (Naya *et al.*, 1995)).

A site that may be involved in transcriptional repression of the Vipr2 gene

Although research into the factors and mechanisms underlying transcriptional repression is still in its infancy in comparison with that carried out on transcriptional activators, the importance of gene-specific transcriptional repressors in the establishment and maintenance of tissue and cell specific patterns of gene expression has been clearly established (reviewed by (Gray & Levine, 1996b)). Only one putative binding site whose sequence is thought to be recognised exclusively by a sequence-specific transcriptional repressor protein was found within the *Vipr2* gene. However , this site is of considerable interest for two reasons: i) it is found within 160 bp of the proposed transcriptional start site, a distance which might be consistent with short range repression of the core transcription complex (Gray & Levine, 1996a; Gray & Levine, 1996b); ii) although the site in question is a recognition site for a factor that has only been identified in *Drosophila* to date, this factor which is known as Tramtrack (Ttk) is involved in the establishment and maintenance of cell fates within the nervous system (Badenhorst *et al.*, 1996; Guo *et al.*, 1995), and also appears to play a role in the regulation of patterning in early embryonic development (Pritchard & Schubiger, 1996). In the nervous system Ttk is expressed in support cells but not in neurons, and is thought to suppress the expression of neuronal genes in non-neuronal cells. Overexpression of Ttk in the sensory organ precursor cells of the PNS stimulates neuronal to nonneuronal shifts in cell fate, whilst underexpression promotes a shift in the opposite direction (Guo *et al.*, 1995).

Transcriptional repressors that are known to be involved in the control of neuronal gene expression in mammalian cells include the neuron-restrictive silencing factor (NRSF) for which binding sites or potential binding sites have been identified in over 15 neuronal genes (Schoenherr *et al.*, 1996; Wood *et al.*, 1996a), and also in a few non-neuronal genes (Schoenherr *et al.*, 1996). However, the binding sites for this factor do not appear to be related to those that bind Ttk, and to date no putative NRSF binding sites have been detected within the *Vipr2* gene .

6.4.3 Transcription factors and sites that may be involved in elevating or repressing *Vipr2* gene transcription in response to certain physiological changes

There is increasing evidence to suggest that although *Vipr2* transcripts are generally present at relatively low levels even in tissues where expression of this gene is greatest (Usdin *et al.*, 1994), *Vipr2* gene transcription can be quite dramatically induced in

response to certain physiological changes. For example, in addition to the induction of *Vipr2* gene expression following T-cell activation (discussed previously), it is interesting to note that in unactivated T-cells stimulation by VIP results in induction of *Vipr2* gene transcription, but produces no significant change in the levels of *Vipr1* transcripts (Delgado *et al.*, 1996b). In the pituitary, elevated levels of *Vipr2* mRNA are found during pregnancy (Usdin *et al.*, 1994). There are indications that developmental regulation of *Vipr2* mRNA levels may occur in the testes (Krempels *et al.*, 1995; Usdin *et al.*, 1994), and preliminary studies of *Vipr2* mRNA levels in the suprachiasmatic nucleus have revealed circadian fluctuations in receptor mRNA levels which appear to precede photic cues (C. Cohen, unpublished data).

Sites that could be involved in the induction of *Vipr2* gene transcription under some of these conditions include, the putative: cAMP responsive element (CRE), serum response element (CArG box), Sox-5 and Sry-type sites, and a STAT site.

cAMP responsive factors

The presence of a putative CRE in the *Vipr2* gene is not surprising, as both the *VIP* gene and the *VIPR1* gene also appear to contain CRE sites within their promoter regions (Sreedharan *et al.*, 1995). In the case of the *VIPR1* and *Vipr2* genes it is not yet known whether these proposed sites are functional, but it nevertheless seems likely that these elements might be involved in the regulation of the relative levels of peptide and receptor gene transcription. The sequence of the putative CRE in the *Vipr2* gene (TGATGTCA), is particularly interesting because although it does not conform to the consensus CRE binding sequence, it is identical to the cAMP responsive element-like element (CARE) 4 site in the P2 promoter of the cAMP response element modulator (*CREM*) gene (Molina *et al.*, 1993). This homology may be relevant to the regulation of *Vipr2* gene expression especially in the suprachiasmatic nucleus, as in the *CREM* promoter this sequence can also bind the inducible cAMP early repressor protein (ICER) which blocks cAMP inducible transcription from CRE sites (Molina *et al.*, 1993). Transcripts encoding ICER are produced from the P2 cAMP inducible promoter of the *CREM* gene, and display a circadian pattern of expression within the pineal gland where ICER production increases dramatically at night (Stehle *et al.*, 1993). When ICER was first discovered studies of its distribution showed that it is found predominantly in neuroendocrine tissues, particularly in the pineal gland (where it has been implicated in mediating circadian clock-directed modulation of hormone synthesis; (Foulkes *et al.*, 1996)), the pituitary, and the adrenal gland (Stehle *et al.*, 1993). However, more recent reports indicate that it is also expressed in other cells of the neuroendocrine and immune systems, including: hypothalamic magnocellular

neurons after osmotic stimulation (Luckman & Cox, 1995), the thyroid gland (Lalli & Sassone-Corsi, 1995), suprachiasmatic nucleus (Stehle *et al.*, 1996), and human T-cells (where increases in cAMP levels are paralleled by ICER induction which in turn leads to the inhibition of calcineurin-mediated stimulation of IL-2 expression) (Bodor *et al.*, 1996).

The observation that ICER is capable of binding to sequences that diverge significantly from the canonical CRE consensus (TGACGTCA), has led to the proposal that many cAMP inducible genes could potentially be regulated by this protein in T-cells (Bodor *et al.*, 1996), and this possibility obviously also applies to the neuroendocrine system. However there have been suggestions that ICER, and other CRE activator and repressor proteins, do not display equal affinity for all CRE sites (Foulkes *et al.*, 1996). Hopefully in the near future characterisation of the affinities of the different CRE binding proteins for CRE-like sites will help in the identification of potential binding sites for these factors. Nevertheless from current knowledge, the sequence identity of the potential CRE site in the *Vipr2* promoter and the CARE 4 site of the *CREM* gene, which also extends 3 bp 5' of the proposed binding site, would appear to strongly support the possibility of ICER regulation of *Vipr2* expression through this site, whilst the ability of ICER to bind to sequences that diverge from the CRE consensus could suggest that a sequence TGAGCTCA 12 bp downstream of the putative *Vipr2* CRE, might conceivably also be involved in cAMP mediated regulation of this gene.

Serum responsive factors

Serum response element binding factors have been implicated in the regulation of the expression of two categories of genes: i) the regulation of immediate early gene transcription in response to mitogen stimulation (Hipskind & Nordheim, 1991; Karagianni & Tsawdaroglou, 1994; Latinkic & Lau, 1994; Latinkic *et al.*, 1991), and ii) the regulation of myogenic genes during muscle development (Soulez *et al.*, 1996). However, as VIP induces immediate-early gene expression (Hisanaga *et al.*, 1993; Vaccarino *et al.*, 1993), VIP receptor levels are modulated by serum factors (Bellan *et al.*, 1992), and *Vipr2* transcripts are present in skeletal muscle (Wei & Mojsov, 1996), the putative *Vipr2* SRE that we have identified could potentially fall into either or both of these categories. The basis for the specificity of SRE activated transcription in each system appears to result from the interaction of SRF with other factors, that bind near to the SRE sequence. These "SRE accessory factors" are thought to include Ets-family ternary complex factors (TCFs) in the case of growth factor induced gene expression (Marais *et al.*, 1993; Price *et al.*, 1996), C/EBPβ and TCFs in serum stimulated

transcriptional activation (Sealy *et al.*, 1997), and muscle-specific factors at the SREs of muscle-specific genes (Galvagni *et al.*, 1997; Groisman *et al.*, 1996). Of these accessory proteins, the most extensively studied factors are the TCFs, which are phosphorylation targets of the MAP kinase signalling pathways (Cahill *et al.*, 1996; Marais *et al.*, 1993), and bind to the Ets consensus sequence (C/A) (C/A) GGA (A/T) which is found just upstream of the SRE in several genes, including the prototypical SRE containing *c-fos* gene (Treisman *et al.*, 1992). From our *Vipr2* gene sequence data, the *Vipr2* putative SRE does not appear to belong to the subgroup of SREs that possess adjacent upstream sites for known TCFs . However, the presence in the *Vipr2* gene of the part of the central TCF binding sequence (Hipskind *et al.*, 1991) CAGG (in the opposite orientation) 1bp downstream of the SRE sequence, makes it difficult to completely exclude the possibility of TCF-SRF interactions in this gene, as it seems likely that only a small proportion of the factors that can interact with SRFs have been identified to date.

Cytokine regulated transcription factors

Cytokines, which include the interleukins, interferons, tumour necrosis factor, and colony stimulating factors, play an important role in the immune system, where they are involved in a wide range of functions (reviewed by (Belardelli, 1995)). However, many of these factors also appear to act as signalling molecules in the neuro-immune axis, and are synthesised within the CNS where they are involved in the regulation of homeostatic and developmental processes in the brain, including the control of neuronal growth and survival (reviewed by (Merrill & Jonakait, 1995)), a process in which VIP has also been implicated (Festoff *et al.*, 1996; Muller *et al.*, 1995; Pincus *et al.*, 1994; Tanaka & Koike, 1994). Potential binding sites for IL-6 regulated transcription factors (NF-IL6 sites or IL-6_REs; (Akira *et al.*, 1995)) were found at several positions throughout the putative promoter region of the *Vipr2* gene. However within the proximal GC rich region of the putative *Vipr2* promoter two additional sites were identified that could represent binding sites for other cytokine responsive transcription factors.

The first of these sites, is a possible binding site (at -612 to -624 bp) for the transcriptional activator interferon regulatory factor-1 (IRF-1, also known as interferon-stimulated gene factor-2 or ISGF-2; (Pine *et al.*, 1990)), which competes with its antagonistic repressor IRF-2 for binding to DNA sites that conform to the consensus sequence (G/C)(A)AAA(N)$_{2-3}$AAA(G/C)(T/C) (Tanaka *et al.*, 1993). IRF-1 was initially characterised as a possible regulator of the promoter regions of the type I interferon genes (Fujita *et al.*, 1989; Pine *et al.*, 1990). However, the discovery that

the expression of transfected IRF-1 constructs can lead to a decrease or arrest in cell growth (Kirchhoff *et al.*, 1993), indicates that the effects of IRF-1 and IRF-2 extend beyond the transcriptional regulation of the interferon system, and is consistent with the recent identification of putative IRF-1/IRF-2 binding sites in the promoter regions of several genes that are involved in the regulation of cell growth and apoptosis (Tanaka *et al.*, 1993; Taniguchi *et al.*, 1995).

The second site is located ~150 bp downstream of the IRF-1 site (at -479) and although it was originally detected as a possible IL-6 RE site in the computer database search for binding sites, this site was subsequently found to display a high level of similarity (1/10 bp mismatch in the form of a T/A substitution) to a STAT binding site that lies within the 180 bp cytokine responsive element (CyRE) of the *VIP* gene. The promoter region of the *VIP* gene has not been characterised in great detail to date, but considerable work has been carried out on the CyRE, which was identified on the basis of its ability to mediate transcriptional activation of the *VIP* gene in response to leukaemia inhibitory factor (LIF), ciliary neurotrophic factor (CNTF) and oncostatin M (Symes *et al.*, 1994). The STAT site in the *VIP* gene CyRE is required for but alone is not sufficient for CNTF/LIF-induced transcriptional activation of the *VIP* gene (Symes *et al.*, 1995a; Symes *et al.*, 1995b), which also requires the presence of 3 C/EBP-related sites that are found downstream of the STAT site (Symes *et al.*, 1995b). Thus, it is interesting that the putative STAT site within the *Vipr2* gene is directly juxtaposed to a site that fits the C/EBP consensus sequence for NF-IL6 binding T(T/G)NNGNAA(T/G) (reviewed in (Akira *et al.*, 1995)), although the exact sequence of this site differs from those found in the *VIP* gene's CyRE. Similar patterns of quite closely positioned C/EBP and STAT sites are found in other IL-6 responsive genes (Dalmon *et al.*, 1993; Kordula & Travis, 1996), and although the basis of the interactions between these elements is not yet entirely clear it seems that they may represent a point of convergence between the MAP kinase and JAK-STAT signalling pathways (Akira *et al.*, 1995).

The specificity of STAT binding to DNA has been shown to be dependent at least in part on the spacing between the core TT and AA half site sequences (Seidel *et al.*, 1995), and in this respect it is possible that the putative *Vipr2* STAT site and the *VIP* gene STAT site may differ in their ability to respond to the STAT signals induced by some cytokines. In contrast to the *VIP* gene STAT site TTCCTGGAA (Symes *et al.*, 1994) which has the 5bp spacing that appears to be associated with general STAT binding (Seidel *et al.*, 1995), the *Vipr2* gene's putative STAT sequence TTACCTGGAA has a 6bp spacing that is characteristic of genes that selectively bind

the STAT complexes activated by IL-4 (Hou *et al.*, 1994; Seidel *et al.*, 1995). This potential selectivity of the putative *Vipr2* STAT site could be interesting as the NF-CLE0 element described earlier in this chapter is found in the promoter region of genes that are expressed in Th2 cells (including the *IL-4* gene and possibly the *Vipr2* gene), while one of the functions of IL-4 is the promotion of Th2 lineage expansion over that of Th1 (Kopf *et al.*, 1993; Noble *et al.*, 1993), a function which requires the presence of the IL-4-induced transcription factor STAT6 (Kaplan *et al.*, 1996).

6.4.4 The *Vipr2* gene's transcription start site(s)

Problems that are generally associated with use of primer extension assays to identify transcriptional start sites include, difficulties in obtaining extension through regions that have a high GC content, and problems with detecting RNA transcripts that are present at very low levels. Unfortunately, it is likely that both of these factors contributed to the failure of the primer extension experiments described in this chapter. In particular, low levels of *Vipr2* transcripts were probably a major factor in the problems encountered, since *Vipr2* transcripts do not appear to be very abundant, even in the olfactory bulb (which relative to other tissues has high levels of *Vipr2* mRNA), and if heterogeneous transcriptional start sites are present in the gene (as has been found in the mouse PTH/PTHrP receptor gene; (McCuaig *et al.*, 1994)) this may compound the difficulties of detecting the extension products.

Possible alternative approaches that were considered included: i) attempting to induce higher levels of *Vipr2* gene expression in cells prior to mRNA isolation, as it appears that dramatic increases in *Vipr2* transcript levels can be induced under certain physiological conditions; or ii) using nuclease protection assays (which are thought to be more sensitive than primer extension methods in the detection of low level transcripts) in order to determine the transcription start site of the *Vipr2* gene. However, as the factors that are responsible for increasing *Vipr2* transcript levels have not yet been established, nuclease protection assays seemed to be the most promising of the two approaches, and while this thesis was being compiled, this method has been succesfully used in our laboratory to establish the transcription start site of the *Vipr2* gene in AtT20 cells and in mouse olfactory bulb. The results of these assays place the transcription start site of the *Vipr2* gene a short distance (27 bp) upstream of the Vipr2 Q12 5' RACE product described earlier, and indicate that in the RNA sources examined *Vipr2* gene transcripts have a 5' untranslated region of 133 bp that is encoded by the same exon as the*Vipr2* gene's ATG translational start site. The results in no way exclude the possibility that tissue-specific alternative promoter usage and/or alternative 5' untranslated exons may exist within the *Vipr2* gene. However, they do suggest that

although upstream enhancer-like sequences may interact with the proximal GC-rich putative promoter sequence, this latter sequence (which includes binding sites for Sp1) is likely to represent the major region responsible for the initiation of *Vipr2* gene transcription in mouse olfactory bulb and AtT20 cells.

6.4.5 Future directions

At present this analysis of the putative promoter region of the *Vipr2* gene is very provisional, as there is no functional data to support a role for any of the transcription factors discussed above, in the regulation of *Vipr2* gene expression. Nevertheless, several interesting sites within the *Vipr2* gene's 5' region have been identified that could act as very useful starting points for further studies.

The best way forward at the moment would probably be to subclone the proximal GC rich region of the promoter and examine its ability to activate the transcription of reporter genes in cell lines where the *Vipr2* gene is normally expressed. Having established whether or not this region on its own can mediate *Vipr2* gene transcription, additional reporter gene constructs could then be used to examine the possible role of more distal sequences and the two upstream TATA box sequences.

Chapter 7

General overview

The studies presented in this thesis have resulted in the isolation of mouse genomic clones that span the entire *Vipr2* gene with the exception of part of intron 4, and also contain a considerable amount of the *Vipr2* gene's 5' flanking region. Through analysis of these λ clones, 12 introns were identified within the *Vipr2* gene's coding sequence, and following sequencing of the gene's exon/intron boundaries, a variety of methods were successfully used to determine the sizes of all but one of these introns. Several of the *Vipr2* gene's introns were considerably larger than had been expected and this caused some problems, particularly in the case of intron 4, where a gap in the contig of *Vipr2* clones still remains. Nevertheless, even with an estimated minimum size of 19 kb for intron 4 (of which ~12 kb has been subcloned), the *Vipr2* gene appears to span at least 59 kb of genomic DNA, and of those members of the secretin/glucagon receptor gene subfamily that have been characterised to date, only the calcitonin receptor gene (Zolnierowicz *et al.*, 1994) is larger.

A bacteriophage P1 clone containing at least part of the human *VIPR2* gene was also isolated, and both this clone and a subcloned restriction fragment from the 3' end of the mouse gene were then used to determine the chromosomal localisation of the VIP_2 receptor gene in mouse and human. The mouse *Vipr2* gene was found to map to the telomeric band of mouse chromosome 12 (12F2), and the human *VIPR2* gene was localised to the human chromosomal region 7q36.3, a result that contrasts with previous reports which indicate that human 7q36 is syntenic with mouse chromosome 5 (Marigo *et al.*, 1995; Martin *et al.*, 1990; Matthes *et al.*, 1993; Ohshima *et al.*, 1995). However, as the *VIPR2* gene is the most distal gene that has been mapped on 7q so far, it appears that this localisation may define a previously unidentified region of synteny between the subtelomeric area of human 7q and mouse chromosome 12.

The localisation of the human *VIPR2* gene to 7q36.3, together with the proposed role of VIP in the control of embryonic growth (Gressens *et al.*, 1993; Gressens *et al.*, 1994), prompted the examination of the *VIPR2* gene as a candidate gene for the brain developmental disorder holoprosencephaly. Initial results from the hybridisation of the *VIPR2* P1 probe to cell lines from patients who had 7q36 deletions and holoprosencephaly (HPE3), demonstrated that in each case only one copy of the *VIPR2* gene was present (a finding which led to the redefinition of the minimal critical region for the *HPE3* locus). However, further collaborative studies, in which two copies of

the *VIPR2* gene were found in patients who had 7q36 translocations and mild HPE phenotypes, indicated that the *VIPR2* gene could not be the sole gene responsible for the HPE3 phenotype, and it has subsequently been reported that mutations affecting the Sonic Hedgehog gene (*SHH)* are the cause of HPE3 (Belloni *et al.*, 1996; Roessler *et al.*, 1996; Roessler *et al.*, 1997).

During this work, the position of the *VIPR2* gene within a series of YAC and cosmid contigs from 7q36 was determined, and the 5' end of the *VIPR2* gene's coding region was found to be situated within ~250 kb of the 7q telomere. This proximity to the telomere, raises some interesting questions with regard to both the potential susceptibility of the *VIPR2* gene to chromosomal damage/deletion, and the possibility that the telomeric position of the gene might be important for the correct regulation of its expression (as has previously been suggested by Riethman *et al.*, 1993, following the mapping of a CpG island, which we now suspect to be the 5' end of the *VIPR2* gene, to a position ~195 kb from the 7q telomere). Telomeric/subtelomeric rearrangements and/or deletions are frequently found in tumour cells (see Meltzer *et al.*, 1993), and have also been detected in some patients who have idiopathic mental retardation (Flint *et al.*, 1995; Wilkie, 1993). However, very recently, as a result of the high resolution now attainable with FISH, and the isolation of probes for all human telomeres (see Yung, 1996), many additional cases of previously undetectable interstitial subtelomeric deletions/rearrangements, cryptic telomeric translocations, and possible telomere capture, have been identified (Flint *et al.*, 1996; Johannesson *et al.*, 1997; Tosi *et al.*, 1996; Wenger *et al.*, 1997), and it is likely that more will follow. As a single copy gene that lies only a few hundred bp from the telomere, the *VIPR2* gene should prove to be a useful marker for the detection of chromosomal deletions or translocations involving the terminal region of 7q, and although the effects, if any, that the telomeric location of the *VIPR2* gene has on the gene's regulation, will probably be difficult to determine, it would nevertheless be interesting to try to examine whether telomeric/subtelomeric rearrangements lead to alterations in *VIPR2* expression levels.

The phenotypic effects of monosomy for the *VIPR2* locus are not yet clear, and the contribution, if any, of *VIPR2* gene deletion to the phenotypes observed in some HPE3 patients remains to be established. Limitations in access to cell lines and cytogenetic/phenotype descriptions of cases of HPE, make it very hard to assess whether certain phenotypic abnormalities might be more common in cases where there is a terminal deletions of 7q that encompasses *VIPR2*, as opposed to an interstitial deletion, mutation, or translocation breakpoint that affects only *SHH*. Nevertheless, even from the published information currently available (Belloni *et al.*, 1996; Belloni *et*

al., 1995; Collins *et al.*, 1993; Grass *et al.*, 1995; Gurrieri *et al.*, 1993; Muenke *et al.*, 1994; Roessler *et al.*, 1997), it is clear that the HPE3 phenotype is extremely variable, and it is often difficult to draw a line between HPE microforms and abnormalities that may be caused by the mutation or deletion of other genes on distal 7q.

Despite the close linkage between the *SHH/HPE3* locus and the site of the sacral agenesis locus described by Lynch *et al.* 1995, the identity or otherwise of the AD sacral agenesis locus and the *HPE3* gene (*SHH*) has not yet been determined. Indeed, in a recent report describing isolated sacral agenesis in association with deletion 7q36.1-qter, it was suggested that the phenotypes observed in 7q terminal deletion cases could be more easily explained if the genes at the *HPE3* and AD sacral agenesis loci on 7q36 were contiguous rather than allelic (Savage *et al.*, 1997). However, although it seems that if *SHH* is responsible for both HPE3 and sacral agenesis, the explanation of the phenotypic variability is likely to be complex, it is nevertheless a very real possibility in light of the wide-ranging effects of SHH during development (Hammerschmidt *et al.*, 1997).

A problem that has frequently arisen during attempts to correlate the known functions of VIP and PACAP with possible functions of the VIP_2 receptor, particularly those related to the developmental effects of VIP, has been the scarcity of experimental evidence to support the proposed ability of the VIP_2 receptor to signal via non-cAMP mediated signalling pathways. In contrast to PACAP type I receptor gene transcripts where splice variation occurs within the region encoding the 3rd intracellular loop (Spengler *et al.*, 1993), no clear patterns of potential sites of splice variation were identified within the corresponding region (intron 10) of the mouse *Vipr2* gene during this study. Nevertheless, the recent characterisation of a PACAP type I receptor variant that stimulates calcium influx by activation of L-type calcium channels (Chatterjee *et al.*, 1996b); the demonstration that in rat bone marrow-derived stromal cells, in which only the *Vipr2* transcripts can be detected by RT-PCR, VIP and PACAP appear to stimulate IL-6 production through a dual signal transduction pathway involving both cAMP and IP_3 (Cai *et al.*, 1997); and the demonstration that the GIP receptor can activate the MAP kinase signal transduction cascade through two different pathways, which differ in their sensitivity to the phosphatidylinositol 3-kinase inhibitor wortmannin (Kubota *et al.*, 1997); strongly indicate that the range of signalling pathways to which members of the secretin glucagon family can couple may be much wider than previously envisaged, and suggest that the non-cAMP signalling mechanisms reported to be involved in mediating the growth effects of VIP, need not necessarily imply that additional VIP receptor genes

exist. Other recent discoveries that may be of relevance regarding the very high affinity VIP binding sites that have been observed on rat cortical astrocytes (see Oláh, *et al.*, 1994 and Gozes *et al.*, 1991), include reports of additional peptide variants in some tissues, e.g. miniglucagon, which is produced from glucagon by proteolytic cleavage at the surface of target cells and has been shown to be active at picomolar concentrations, although its effects often oppose those of glucagon (Bataille *et al.*, 1996). Therefore, although to date, the cAMP-mediated signalling pathways of the VIP_1 and VIP_2 receptors have predominated in the literature, the clarification of any additional pathways through which these receptors signal, and in particular whether or not the VIP_2 receptor can signal via IP_3-linked pathways, is likely to be one of the most critical aspects of studies of these receptors in the immediate future.

The 5' flanking sequence of the mouse *Vipr2* gene, was not isolated until near the end of the studies presented here, and consequently only a preliminary analysis of the its sequence was possible in the time available. Nevertheless, even from this initial analysis, it is clear that the putative promoter region of the *Vipr2* gene is similar to the promoter regions of several other members of the secretin/glucagon receptor family (Aino *et al.*, 1995; Burcelin *et al.*, 1995; McCuaig *et al.*, 1995; Sreedharan *et al.*, 1995; Tsai-Morris *et al.*, 1996) in that it contains Sp1 sites as opposed to a TATA box, and is GC and CpG rich. Of the possible binding sites for transcription factors that were identified, there were several that are of particular interest in light of either known or suspected functions of the VIP_2 receptor.

These include:

i) a conserved lymphokine element 0 (CLE0) site, which in lymphokine genes is involved in stimulating gene transcription in T-cells following antigen activation (Miyatake *et al.*, 1991), a stimulus which has also appears to induce T-cell *Vipr2* gene transcription (Delgado *et al.*, 1996b).

ii) a site that is identical in sequence to the CTI element (a binding site for isl-1-like factors) in the human insulin gene (Petersen *et al.*, 1994), and in light of the proposed role of the VIP_2 receptor in mediating the effects of VIP and PACAP on glucose-stimulated insulin secretion (Bertrand *et al.*, 1996; Inagaki *et al.*, 1994), may be involved in regulating the expression of the *Vipr2* gene in pancreatic islets.

iii) a region of similarity between the proximal promoter region of *Vipr2* gene and that of the rat pro-opiomelanocortin gene, which (although in the opposite orientation in the *Vipr2* gene) represented the apparent conservation of a composite element consisting of sites for Sp1 and the pituitary specific transcription factor PP1 (Liu *et al.*, 1995).

iv) a putative cAMP response element, which is identical in sequence to a site in the CREM gene that has previously been shown to be capable of binding the inducible cAMP early repressor 'ICER' (Molina *et al.*, 1993).

The possible role of ICER in the regulation of *Vipr2* gene expression is especially intriguing with regard to the proposed role of VIP in the suprachiasmatic nucleus (SCN), as it has recently been reported that a light induced upregulation of CREM (and hence ICER) mRNA occurs in the SCN during the second half of the night (Stehle *et al.*, 1996). Interestingly, a peak in the level of *Vipr2* mRNA (which is expressed biphasically) within the SCN, is also observed during the latter part of the night (Cagampang *et al.*, submitted). At present the temporal relationship between the peak in *Vipr2* mRNA and that of ICER (which might be predicted to result in the downregulation of *Vipr2* mRNA levels) is not known, but this might prove to be an important line of enquiry for future work. At the molecular level, further studies of the role of the putative *Vipr2* cAMP response element and the other proposed binding sites for transcription factors (described above and in Chapter 6), are likely to take the form of reporter construct experiments using a set of nested deletions from the *Vipr2* gene's 5' flanking sequence (and cell lines in which the *Vipr2* gene is usually expressed), to try to establish the relative contributions that individual regions make to the transcriptional activation of the gene. Having identified specific regions of interest, band shift and footprinting assays with appropriate cell extracts could then be used to characterise the transcription factor binding sites present in these areas. Other approaches that are likely to be useful include the ectopic expression of transcription factors in cell lines that express *Vipr2*. Indeed, within the last few months a paper has been published in which the effects of ectopic expression of ICER on cell morphology and adrenocorticotrophic hormone (ACTH) secretion in the pituitary corticotroph AtT20 cell line are described (Lamas *et al.*, 1997). The *Vipr2* gene is also expressed in AtT20 cells, and has recently been implicated in mediating the transcriptional stimulation of the pro-opiomelanocortin gene (which encodes ACTH) by VIP and PACAP (Aoki *et al.*, 1997). The possible conservation of a pituitary-specific transcription factor binding site between the *Vipr2* and pro-opiomelanocortin genes (see iii), might therefore be important in co-ordinating or inter-regulating the expression of the two genes in some pituitary cell types, and as ICER does not affect the transcription of the pro-

opiomelanocortin gene directly, but acts through the transcriptional regulation of a prohormone convertase gene (Lamas *et al.*, 1997), AtT20 cells could be a very interesting system in which to begin to examine the interrelationships between ICER, *Vipr2*, and pro-opiomelanocortin gene expression.

The ~5.6 kb *Xba* I fragment that was isolated and subcloned from the 5' end of the *Vipr2* gene (see Chapter 6) has subsequently been used by other members of our laboratory to produce a *Vipr2* replacement construct for use in the generation of *Vipr2*-null mice. Viable fertile lines of mice that were heterozygous for the *Vipr2* replacement sequences were successfully generated, and at least one homozygous *Vipr2*-knockout mouse has now been produced, although the phenotype associated with loss of the VIP_2 receptor in mice has not yet been studied in detail. Some caution should be observed when attempting to draw parallels between the functions of the VIP_2 receptor in mouse and those in human, as preliminary analysis of the 5' flanking region of the human *VIPR2* gene (carried out by E. Lutz in our laboratory) indicates that significant differences may exist between the promoter regions of the mouse and human genes. Nevertheless, it is anticipated that the physiological, pharmacological and behavioural studies of the *Vipr2*-knockout mice will contribute substantially to our understanding of the role of the VIP_2 receptor in mediating the known effects of VIP and PACAP, and may also uncover previously unsuspected functions of these peptides and their receptors.

REFERENCES

Abe, H., Engler, D., Molitch, M.E., Bollinger-Gruber, J. & Reichlin, S. (1985). Vasoactive intestinal peptide is a physiological mediator of prolactin release in the rat. *Endocrinology* **116**, 1383-1390.

Ahren, B., Alumets, J., Ericsson, M., Fahrenkrug, J., Fahrenkrug, L., Hakanson, R., Hedner, P., Loren, I., Melander, A., Rerup, C. & Sundler, F. (1980a). VIP occurs in intrathyroidal nerves and stimulates thyroid hormone secretion. *Nature* **287**, 343-345.

Ahren, B., Alumets, J., Ericsson, M., Fahrenkrug, J., Fahrenkrug, L., Hakanson, R., Hedner, P., Loren, I., Melander, A., Rerup, C. & Sundler, F. (1980b). VIP occurs in intrathyroidal nerves and stimulates thyroid hormone secretion. *Nature* **287**, 343-345.

Aino, H., Hashimoto, H., Ogawa, N., Nishino, A., Yamamoto, K., Nogi, H., Nagata, S. & Baba, A. (1995). Structure of the gene encoding the mouse pituitary adenylate cyclase-activating polypeptide receptor. *Gene* **164**, 301-304.

Aiyar, N., Rand, K., Elshourbagy, N.A., Zeng, Z., Adamou, J.E., Bergsma, D.J. & Li, Y. (1996). A cDNA encoding the calcitonin gene-related peptide type I receptor. *Journal of Biological Chemistry* **271**, 11325-11329.

Akira, S., Nishio, Y., Tanaka, T., Inoue, M., Matsusaka, T., Wang, X.-J., Wei, S., Yoshida, N. & Kishimoto, T. (1995). Transcription factors NF-IL6 and APRF involved in gp130-mediated signaling pathway. *Annals of the New York Academy of Sciences* **762**, 15-27.

Albrandt, K., Brady, E.M.G., Moore, C.X., Mull, E., Sierzega, M.E. & Beaumont, K. (1995). Molecular cloning and functional expression of a third isoform of the human calcitonin receptor and partial characterisation of the calcitonin receptor gene. *Endocrinology* **136**, 5377-5384.

Aliakbari, J., Sreedharan, S.P., Turck, C.W. & Goetzl, E.J. (1987). Selective localization of vasoactive intestinal peptide and substance P in human eosinophils. *Biochemical and Biophysical Research Communications* **148**, 1440-1445.

Altiere, R.J. & Diamond, L. (1984). Relaxation of cat tracheobronchial and pulmonary arterial smooth muscle by vasoactive intestinal peptide: lack of influence by peptidase inhibitors. *British Journal of Pharmacology* **82**, 321-328.

Alvarez-Bolado, G., Rosenfeld, M.G. & Swanson, L.W. (1995). Model of forebrain regionalization based on spatiotemporal patterns of POU-III homeobox gene expression, birthdates, and morphological features. *The Journal of Comparative Neurology* **355**, 237-295.

Amizuka, N., Lee, H.S., Kwan, M.Y., Arazani, A., Warshawsky, H., Hendy, G.N., Ozawa, H., White, J.H. & Goltzman, D. (1997). Cell-specific expression of the parathyroid hormone (PTH)/PTH-related peptide receptor gene in kidney from kidney-specific and ubiquitous promoters. *Endocrinology* **138**, 469-481.

Anand, P., Gibson, S.J., McGregor, G.P., Blank, M.A., Ghatei, M.A., Bacarese-Hamilton, A.J., Polak, J.M. & Bloom, S.R. (1983). A VIP-containing system concentrated in the lumbosacral region of human spinal cord. *Nature* **305**, 143-145.

Angus, J., Thomson, F., Murphy, K., Baker, E., Sutherland, G.R., Parsons, P.G. & Sturm, R.A. (1995). The *brn-2* gene regulates the melanocytic phenotype and tumorigenic potential of human melanoma cells. *Oncogene* **11**, 691-700.

Aoki, Y., Iwasaki, Y., Katahira, M., Oiso, Y. & Saito, H. (1997). Regulation of the rat proopiomelanocortin gene expression in AtT-20 cells. II: Effects of the pituitary adenylate cyclase-activating polypeptide and vasoactive intestinal polypeptide. *Endocrinology* **138**, 1930-1934.

Arimura, A. & Shioda, S. (1995). Pituitary adenylate cyclase activating polypeptide (PACAP) and its receptors: neuroendocrine and endocrine interaction. *Frontiers in Neuroendocrinology* **16**, 53-88.

Arimura, A., Somogyvari-Vigh, A., Weill, C., Fiore, R.C., Tatsuno, I., Bay, V. & Brenneman, D.E. (1994). PACAP functions as a neurotrophic factor. *Annals of the New York Academy of Sciences* **739**, 228-243.

Atkin, N.B. & Baker, M.C. (1993). Chromosome 7q deletions: observations on 13 malignant tumors. *Cancer Genetics and Cytogenetics* **67**, 123-125.

Badenhorst, P., Harrison, S. & Travers, A. (1996). End of the line? Tramtrack and cell fate determination in *Drosophila*. *Genes to Cells* **1**, 707-716.

Barajas, L., Sokolski, K.N. & Lechago, J. (1983). Vasoactive intestinal polypeptide-immunoreactive nerves in the kidney. *Neuroscience Letters* **43**, 263-269.

Barr, M., Hanson, J.W., Currey, K., Sharp, S., Toriello, H., Schmickel, R.D. & Wilson, G.N. (1983). Holoprosencephaly in infants of diabetic mothers. *Journal of Pediatrics* **102**, 565-568.

Bataille, D., Blache, P. & Bergeron, F. (1996). Endoprotease regulation of miniglucagon production. *Annals of the New York Academy of Sciences* **805**, 1-8.

Batiuk, T.D., Kung, L. & Halloran, P.F. (1997). Evidence that calcineurin is rate-limiting for primary lymphocyte activation. *Journal of Clinical Investigation* **100**, 1894-1901.

Baud, V., Chissoe, S.L., Viegaspequignot, E., Diriong, S., Nguyen, V.C., Roe, B.A. & Lipinski, M. (1995). EMR1, and unusual member in the family of hormone receptors with 7 transmembrane segments. *Genomics* **26**, 334-344.

Bedell, M.A., Jenkins, N.A. & Copeland, N.G. (1996). Good genes in bad neighbourhoods. *Nature Genetics* **12**, 229-232.

Beechey, C.V., Kirk, M. & Searle, A.G. (1980). A reciprocal translocation induced in an oocyte and affecting fertility in male mice. *Cytogenetics and Cell Genetics* **27**, 129-146.

Belardelli, F. (1995). Role of interferons and other cytokines in the regulation of the immune response. *APMIS* **103**, 161-179.

Bellan, C., Fabre, C., Secchi, J., Marvaldi, J., Pichon, J. & Luis, J. (1992). Modulation of the expression of the VIP receptor by serum factors on the human melanoma cell line IGR39. *Experimental Cell Research* **200**, 34-40.

Belloni, E., Muenke, M., Roessler, E., Traverso, G., Siegel-Bartelt, J., Frumkin, A., Mitchell, H.F., Donis-Keller, H., Helms, C., Hing, A.V., Heng, H.H.Q., Koop, B., Martindale, D., Rommens, J.M., Tsui, L.-C. & Scherer, S.W. (1996). Identification of *Sonic hedgehog* as a candidate gene responsible for holoprosencephaly. *Nature Genetics* **14**, 353-356.

Belloni, E., Scherer, S.W., Siegel-Bartelt, J., Frumkin, A., Donis-Keller, H., Helms, C., Hing, A.V., Roessler, E., Barnoski, B., Mitchell, H., Muenke, M. & Tsui, L.-C. (1995). Characterization of holoprosencephaly minimal critical region at 7q36. *American Journal of Human Genetics* **57**, 1487-1487.

Ben-Arie, N., Lancet, D., Taylor, C., Khen, M., Walker, N., Ledbetter, D.H., Carrozzo, R., Patel, K., Sheer, D., Lehrach, H. & North, M.A. (1994). Olfactory receptor gene cluster on human chromosome 17: possible duplication of an ancestral receptor repertoire. *Human Molecular Genetics* **3**, 229-235.

Bertrand, G., Puech, R., Maisonnasse, Y., Bockaert, J. & Loubatières-Mariani, M.M. (1996). Comparative effects of PACAP and VIP on pancreatic endocrine secretions and vascular resistance in rat. *British Journal of Pharmacology* **117**, 764-770.

Birren, B. & Lai, E. (1994). Rapid pulsed field separation of DNA molecules up to 250 kb. *Nucleic Acids Research* **22**, 5366-5370.

Bishop, A.E., Polak, J.M., Green, I.C., Bryant, M.G. & Bloom, S.R. (1980). The location of VIP in the pancreas of man and rat. *Diabetologia* **18**, 73-78.

Bitar, K.N. & Makhlouf, G.M. (1982). Relaxation of isolated gastric smooth muscle cells by vasoactive intestinal peptide. *Science* **216**, 531-533.

Blackwell, K.T. & Weintraub, H. (1990). Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **250**, 1104-1110.

Boam, D.S.W., Clark, A.R. & Docherty, K. (1990). Positive and negative regulation of the human insulin gene by multiple trans-acting factors. *The Journal of Biological Chemistry* **265**, 8285-8296.

Bodner, M., Castrillo, J.L., Theill, L.E., Deerinck, T., Ellisman, M. & Karin, M. (1988). The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. *Cell* **55**, 505-518.

Bodor, J., Spetz, A.L., Strominger, J.L. & Habener, J.F. (1996). cAMP inducibility of transcriptional repressor ICER in developing and mature human T-lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 3536-3541.

Braas, K.M., Brandenberg, C.A. & May, V. (1994). Pituitary adenylate cyclase-activating polypeptide regulation of AtT-20/D16v corticotrope cell proopiomelanocortin expression and secretion. *Endocrinology* **134**, 186-195.

Braas, K.M. & May, V. (1996). Pituitary adenylate cyclase-activating polypeptides, PACAP-38 and PACAP-27, regulation of sympathetic neuron catecholamine and neuropeptide Y expression through activation of type I PACAP/VIP receptor isoforms. *Annals of the New York Academy of Sciences* **805**, 204-216.

Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978). Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 4853-4857.

Brenneman, D.E. & Gozes, I. (1996). A femtomolar-acting neuroprotective peptide. *Journal of Clinical Investigation* **97**, 2299-2307.

Brenneman, D.E., Hill, J.M., Gozes, I. & Phillips, T.M. (1996). Vasoactive intestinal peptide releases interleukin-1 from astrocytes. *Annals of the New York Academy of Sciences* **805**, 280-287.

Brenneman, D.E., Nicol, T., Warren, D. & Bowers, L.M. (1990). Vasoactive intestinal peptide: a neurotrophic release agent and an astroglial mitogen. *Journal of Neuroscience Research* **25**, 386-394.

Brick, P.L., Howlett, A.C. & Beinfeld, M.C. (1985). Synthesis and release of vasoactive intestinal polypeptide (VIP) by mouse neuroblastoma cells: modulation by cyclic nucleotides and ascorbic acid. *Peptides* **6**, 1075-1078.

Brown, L. & Baer, R. (1994). *HEN1* encodes a 20-kilodalton phosphoprotein that binds an extended E-box motif as a homodimer. *Molecular and Cellular Biology* **14**, 1245-1255.

Brown, L., Espinosa, R., Le Beau, M.M., Siciliano, M.J. & Baer, R. (1992). *HEN1* and *HEN2*: A subgroup of basic helix-loop-helix genes that are coexpressed in a human neuroblastoma. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 8492-8496.

Brownstein, B.H., Silverman, G.A., Little, R.A., Burke, D.T., Korsmeyer, S.J., Schlessinger, D. & Olson, M.V. (1989). Isolation of single-copy human genes from a library of yeast artificial chromosome clones. *Science* **244**, 1348-1351.

Bryant, M., Polak, J.M., Modlin, I., Bloom, S.R., Albuquerque, R.H. & Pearse, A.G. (1976). Possible dual role for vasoactive intestinal peptide as a gastrointestinal hormone and neurotransmitter substance. *Lancet* **1**, 991-993.

Buisson, B., Laflamme, L., Bottari, S.P., de Gasparo, M., Gallo-Payet, N. & Payet, M.D. (1995). A G protein is involved in the angiotensin AT_2 receptor inhibition of the T-type calcium current in non-differentiated NG108-15 cells. *The Journal of Biological Chemistry* **270**, 1670-1674.

Burcelin, R., Li, J. & Charron, M.J. (1995). Cloning and sequence analysis of the murine glucagon receptor encoding gene. *Gene* **164**, 305-310.

Burke, D.T., Carle, G.F. & Olson, M.V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**, 806-812.

Cagampang, F.R.A., Sheward, W.J., Harmar, A.J., Piggins, H.D. & Coen, C.W. (submitted). Circadian changes in the expression of vasoactive intestinal peptide 2 receptor mRNA in the rat suprachiasmatic nuclei. *Molecular Brain Research*

Cahill, M.A., Janknecht, R. & Nordheim, A. (1996). Signalling pathways: Jack of all cascades. *Current Biology* **6**, 16-19.

Cai, Y., Xin, X., Shim, G.-J., Mokuno, Y., Uehara, H., Yamada, T., Agui, T. & Matsumoto, K. (1997). Pituitary adenylate cyclase activating polypeptide (PACAP) and vasoactive intestinal peptide (VIP) stimulate interleukin-6 production through the third subtype of PACAP/VIP receptor in rat bone marrow-derived stromal cells. *Endocrinology* **138**, 2515-2520.

Campbell, R.M. & Scanes, C.G. (1992). Evolution of the growth hormone-releasing factor (GRF) family of peptides. *Growth Regulation* **2**, 175-191.

Candia, A.F., Hu, J., Crosby, J., Lalley, P.A., Noden, D., Nadeau, J.H. & Wright, C.V. (1992). Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. *Development* **116**, 1123-1136.

Carlquist, M., McDonald, T.J., Go, V.L., Bataille, D., Johansson, C. & Mutt, V. (1982). Isolation and amino acid composition of human vasoactive intestinal polypeptide (VIP). *Hormone and Metabolism Research* **14**, 28-29.

Carlquist, M., Mutt, V. & Jornvall, H. (1979). Isolation and characterization of bovine vasoactive intestinal peptide (VIP). *FEBS Letters* **108**, 457-460.

Cavalieri, F. (1996). Drugs that target gene expression: an overview. *Critical Reviews in Eukaryotic Gene Expression* **6**, 75-85.

Ceccatelli, S., Tsuruo, Y., Hokfelt, T., Fahrenkrug, J. & Dohler, K.D. (1988). Some blood vessels in the rat median eminence are surrounded by a dense plexus of vasoactive intestinal polypeptide/peptide histidine isoleucine (VIP/PHI) immunoreactive nerves. *Neuroscience Letters* **84**, 29-34.

Chang, D.T., López, A., von Kessler, D.P., Chiang, C., Simandi, B.K., Zhao, R., Seldin, M.F., Fallon, J.F. & Beachey, P.A. (1994). Products, genetic linkage and limb patterning activity of a murine *hedgehog* gene. *Development* **120**, 3339-3353.

Chang, E., Welch, S., Luna, J., Giacalone, J. & Francke, U. (1993). Generation of a human chromosome 18-specific YAC clone collection and mapping of 55 unique YACs by FISH and fingerprinting. *Genomics* **17**, 393-402.

Chartrel, N., Wang, Y., Fournier, A., Vaudry, H. & Conlon, J.M. (1995). Frog vasoactive intestinal polypeptide and galanin: primary structures and effects on pituitary adenylate cyclase. *Endocrinology* **136**, 3079-3086.

Chatterjee, T.K., Liu, X., Davisson, R.L. & Fisher, R.A. (1996a). Structural organization of the rat pituitary adenylate cyclase activating polypeptide (PACAP) receptor gene: evidence for alternative splicing in the 5'-untranslated region. *FASEB Journal* **10**, 2255.

Chatterjee, T.K., Sharma, R.V. & Fisher, R.A. (1996b). Molecular cloning of a novel variant of the pituitary adenylate cyclase-activating polypeptide (PACAP) receptor that stimulates calcium influx by activation of L-type calcium channels. *The Journal of Biological Chemistry* **271**, 32226-32232.

Chen, R., Lewis, K.A., Perrin, M.H. & Vale, W.W. (1993). Expression cloning of a human corticotropin-releasing-factor receptor. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8967-8971.

Cheung, C.Y. & Holzwarth, M.A. (1986). Fetal adrenal VIP: distribution and effect on medullary catecholamine secretion. *Peptides* **7**, 413-418.

Chew, L.J., Burke, Z.D., Morgan, H., Gozes, I., Murphy, D. & Carter, D.A. (1997). Transcription of the vasoactive intestinal peptide gene in response to glucocorticoids: differential regulation of alternative transcripts is modulated by a labile protein in anterior pituitary. *Molecular and Cellular Endocrinology* **130**, 83-91.

Chiang, C., Litingtung, Y., Lee, E., Young, K.E., Corden, J.L., Westphal, H. & Beachy, P.A. (1996). Cyclopia and defective axial patterning in mice lacking *Sonic hedgehog* gene function. *Nature* **383**, 407-413.

Christophe, J. (1993). Type I receptors for PACAP (a neuropeptide even more important than VIP?). *Biochemica et Biophysica Acta* **1154**, 183-199.

Christophe, J., Cauvin, A., Vervisch, E., Buscail, L., Damien, C., Abello, J., Gourlet, P. & Robberecht, P. (1990). VIP receptors in human SUP-T1 lymphoblasts. *Digestion* **46 (supplement 2)**, 148-155.

Christophe, J., Svoboda, M., Lambert, M., Waelbroeck, M., Winand, J., Dehaye, J.-P., Vandermeers-Piret, M.-C., Vandermeers, A. & Robberecht, P. (1986). Effector mechanisms of peptides of the VIP family. *Peptides* **7 Suppl. 1**, 101-107.

Christophe, J., Vandermeers-Piret, M.-C., Vandermeers, A. & Robberecht, P. (1993). From lizard helodermin to mammalian helodermin-preferring and PACAP- preferring VIP receptors. *Biomedical Research* **14**, 53-60.

Cohen, M.M., Jr. (1989). Perspectives on holoprosencephaly: Part I. Epidemiology, genetics, and syndromology. *Teratology* **40**, 211-235.

Collins, A.L., Lunt, P.W., Garrett, C. & Dennis, N.R. (1993). Holoprosencephaly: a family showing dominant inheritance and variable expression. *Journal of Medical Genetics* **30**, 36-40.

Collins, J. & Hohn, B. (1978). Cosmids: a type of plasmid gene-cloning vector that is packageable *in vitro* in bacteriophage λ heads. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 4242-4246.

Coulter, C.L., Leech, R.W., Schaefer, G.B., Scheithauer, B.W. & Brumback, R.A. (1993). Midline cerebral dysgenesis, dysfunction of the hypothalamic-pituitary axis, and fetal alcohol effects. *Archives of Neurology* **50**, 771-775.

Cousins, L. (1983). Congenital anomalies among infants of diabetic mothers. *American Journal of Obstetrics and Gynecology* **147**, 333-338.

Couvineau, A., Rouyer-Fessard, C., Darmoul, D., Maoret, J.-J., Carrero, I., Ogier-Denis, E. & Laburthe, M. (1994). Human intestinal VIP receptor: cloning and functional expression of two cDNA encoding proteins with different N-terminal domains. *Biochemical and Biophysical Research Communications* **200**, 769-776.

Coy, D.H., Jiang, N.-Y., Fuselier, J. & Murphy, W.A. (1996). Structural simplification of potent growth hormone-releasing hormone analogs: implications for other members of the VIP/GHRH/PACAP family. *Annals of the New York Academy of Sciences* **805**, 149-158.

Craig, J.M. & Bickmore, W.A. (1994). The distribution of CpG islands in mammalian chromosomes. *Nature Genetics* **7**, 376-382.

Cross, S.H. & Bird, A.P. (1995). CpG islands and genes. *Current Opinion in Genetics and Development* **5**, 309-314.

Cunningham, L.A. & Holzwarth, M.A. (1988). Vasoactive intestinal peptide stimulates adrenal aldosterone and corticosterone secretion. *Endocrinology* **122**, 2090-2097.

Curran, M.E., Splawski, I., Timothy, K.W., Vincent, G.M., Green, E.D. & Keating, M.T. (1995). A molecular basis for cardiac arrythmia: HERG mutations cause long QT syndrome. *Cell* **80**, 795-803.

D'Agata, V., Cavallaro, S., Stivala, F., Travali, S. & Canonico, P.L. (1996). Tissue-specific and developmental expression of pituitary adenylate cyclase-activating polypeptide (PACAP) receptors in rat brain. *European Journal of Neuroscience* **8**, 310-318.

D'Eustachio, P. & Riblet, R. (1996). Mouse Chromosome 12. *Mammalian Genome* **6**, S221-S231.

Dalmon, J., Laurent, M. & Courtois, G. (1993). The human β fibrinogen promoter contains a hepatocyte nuclear factor 1-dependent interleukin-6-responsive element. *Molecular and Cellular Biology* **13**, 1183-1193.

Davoren, J.B. & Hsueh, A.J. (1985). Vasoactive intestinal peptide: a novel stimulator of steroidogenesis by cultured rat granulosa cells. *Biology of Reproduction* **33**, 37-52.

De, A., Brown, D.G., Gorman, M.A., Carr, M., Sanderson, M.R. & Freemont, P.S. (1994). Crystal structure of a disulphide-linked "trefoil" motif found in a large family of putative growth factors. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 1084-1088.

de la Fuente, M., Delgado, M., del Rio, M., Garrido, E., Leceta, J., Hernanz, A. & Gomariz, R.P. (1994). Vasoactive intestinal peptide modulation of adherence and mobility in rat peritoneal lymphocytes and macrophages. *Peptides* **15**, 1157-1163.

Dean, M. (1996). Polarity, proliferation and the *hedgehog* pathway. *Nature Genetics* **14**, 245-247.

Delgado, M., Garrido, E., Martinez, C., Leceta, J. & Gomariz, R.P. (1996a). Vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptides (PACAP27) and PACAP38) protect CD4+ CD8+ thymocytes from glucocorticoid-induced apoptosis. *Blood* **87**, 5152-5161.

Delgado, M., Martinez, C., Johnson, M.C., Gomariz, R.P. & Ganea, D. (1996b). Differential expression of vasoactive intestinal peptide receptors 1 and 2 (VIP-R1 and VIP-R2) mRNA in murine lymphocytes. *Journal of Neuroimmunology* **68**, 27-38.

Della, N.G., Papka, R.E., Furness, J.B. & Costa, M. (1983). Vasoactive intestinal peptide-like immunoreactivity in nerves associated with the cardiovascular system of guinea-pigs. *Neuroscience* **9**, 605-619.

Delporte, C., Poloczek, P., de Neef, P., Vertongen, P., Ciccarelli, E., Svoboda, M., Herchuelz, A., Winand, J. & Robberecht, P. (1995). Pituitary adenylate cyclase activating polypeptide (PACAP) and vasoactive intestinal peptide stimulate two signalling pathways in CHO cells stably transfected with the selective type I PACAP receptor. *Molecular and Cellular Endocrinology* **107**, 71-76.

Demetrick, D.J., Zhang, H. & Beach, D.H. (1994). Chromosomal mapping of human CDK2, CDK4, and CDK5 cell cycle kinase genes. *Cytogenetics and Cell Genetics* **66**, 72-74.

DeMyer, W. & Zeman, W. (1963). Alobar holoprosencephaly (arhinencephaly) with median cleft lip and palate: clinical, electroencephalographic and nosologic considerations. *Confinia Neurologica* **23**, 1-36.

DeMyer, W., Zeman, W. & Palmer, C.G. (1964). The face predicts the brain: diagnostic significance of median facial anomalies for holoprosencephaly (arhinencephaly). *Pediatrics* **34**, 256-263.

Denny, P., Swift, S., Brand, N., Dabhade, N., Barton, P. & Ashworth, A. (1992). A conserved family of genes related to the testis determining gene, *SRY*. *Nucleic Acids Research* **20**, 2887.

Dickinson, K.E.J., Schachter, M., Miles, C.M.M., Coy, D.H. & Sever, P.S. (1986). Characterization of vasoactive intestinal peptide (VIP) receptors in mammalian lung. *Peptides* **7**, 791-800.

Dickinson, T., Fleetwood-Walker, S.M., Mitchell, R. & Lutz, E.M. (1997). Evidence for roles of vasoactive intestinal polypeptide (VIP) and pituitary adenylate cyclase activating polypeptide (PACAP) receptors in modulating the responses of rat dorsal horn neurons to sensory inputs. *Neuropeptides* **31**, 175-185.

Dietrich, W.F., Miller, J., Steen, R., Merchant, M.A., Damronboles, D., Husain, Z., Dredge, R., Daly, M.J., Ingalls, K.A., Oconnor, T.J., Evans, C.A., Deangelis, M.M., Levinson, D.M., Kruglyak, L., Goodman, N., Copeland, N.G., Jenkins, N.A., Hawkins, T.L., Stein, L., Page, D.C. & Lander, E.S. (1996). A comprehensive genetic-map of the mouse genome. *Nature* **380**, 149-152.

Dillon, J.S., Tanizawa, Y., Wheeler, M.B., Leng, X.H., Ligon, B.B., Rabin, D.U., Yoo-Warren, H., Permutt, M.A. & Boyd, A.E. (1993). Cloning and functional expression of the human glucagon-like peptide-1 (GLP-1) receptor. *Endocrinology* **133**, 1907-1910.

Dimaline, R., Reeve, J.R., Shively, J.E. & Hawke, D. (1984). Isolation and characterization of rat vasoactive intestinal peptide. *Peptides* **5**, 183-187.

Dodd, J., Kelly, J.S. & Said, S.I. (1979). Excitation of CA1 neurones of the rat hippocampus by the octacosapeptide, vasoactive intestinal polypeptide. *British Journal of Pharmacology* **66**, 125P.

Dong, J., Asa, S.L. & Drucker, D.J. (1991). Islet cell and extrapancreatic expression of the LIM domain homeobox gene *isl*-1. *Molecular Endocrinology* **5**, 1633-1641.

Donnelly, D. (1997). The arrangement of the transmembrane helices in the secretin receptor family of G-protein-coupled receptors. *FEBS Letters* **409**, 431-436.

Dumonteil, E. & Philippe, J. (1996). Insulin gene: organisation, expression and regulation. *Diabetes and Metabolism* **22**, 164-173.

Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J.A. & McMahon, A.P. (1993). Sonic hedgehog, a member of a family of putative signalling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417-1430.

Ehrhart-Bornstein, M., Bornstein, S.R., Scherbaum, W.A., Pfeiffer, E.F. & Holst, J.J. (1991). Role of vasoactive intestinal peptide in a neuroendocrine regulation of the adrenal cortex. *Neuroendocrinology* **54**, 623-628.

Eiden, L.E., Nilaver, G. & Palkovits, M. (1982). Distribution of vasoactive intestinal polypeptide (VIP) in the rat brain stem nuclei. *Brain Research* **231**, 472-477.

Eisen, T., Easty, D.J., Bennett, D.C. & Goding, C.R. (1995). The POU domain transcription factor Brn-2: elevated expression in malignant melanoma and regulation of melanocyte-specific gene expression. *Oncogene* **11**, 2157-2164.

Ekker, S.C., Ungar, A.R., Greenstein, P., von Kessler, D.P., Porter, J.A., Moon, R.T. & Beachy, P.A. (1995). Patterning activities of vertebrate *hedgehog* proteins in the developing eye and brain. *Current Biology* **5**, 944-955.

Eng, J., Du, B.H., Raufman, J.P. & S., Y.R. (1986). Purification and amino acid sequences of dog, goat and guinea pig VIPs. *Peptides* **7, Suppl. 1**, 17-20.

Eng, J., Yu, J., Rattan, S. & Yalow, R.S. (1992). Isolation and amino acid sequences of opossum vasoactive intestinal polypeptide and cholecystokinin octapeptide. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 1809-1811.

Ernström, U., Gafvelin, G. & Mutt, V. (1995). Rescue of thymocytes from cell death by vasoactive intestinal peptide. *Regulatory Peptides* **57**, 99-104.

Escalada, J., Cacicedo, L., Ortego, J., Melian, E. & Sanchez-Franco, F. (1996). Prolactin gene expression and secretion during pregnancy and lactation in the rat: role of dopamine and vasoactive intestinal peptide. *Endocrinology* **137**, 631-637.

Fahrenkrug, J. (1979). Vasoactive intestinal polypeptide: measurement, distribution and putative neurotransmitter function. *Digestion* **19**, 149-169.

Fantes, J.A., Bickmore, W.A., Fletcher, J.M., Ballesta, F., Hanson, I.M. & Van Heyningen, V. (1992). Submicroscopic deletions at the WAGR locus, revealed by nonradioactive *in situ* hybridization. *American Journal of Human Genetics* **51**, 1286-1294.

Feinberg, A.P. & Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**, 6-13.

Feliu, J.E., Mojena, M., Silvestre, R.A., Monge, L. & Marco, J. (1983). Stimulatory effect of vasoactive intestinal peptide on glycogenolysis and gluconeogenesis in isolated rat hepatocytes: antagonism by insulin. *Endocrinology* **112**, 2120-2127.

Festoff, B.W., Nelson, P.G. & Brenneman, D.E. (1996). Prevention of activity-dependent neuronal death: vasoactive intestinal polypeptide stimulates astrocytes to secrete the thrombin- inhibiting neurotrophic serpin, protease nexin-I. *Journal Of Neurobiology* **30**, 255-266.

Filizola, M., Cartenì-Farina, M. & Perez, J.J. (1997). Conformational study of vasoactive intestinal peptide by computational methods. *Journal of Peptide Research* **50**, 55-64.

Fink, T. & Weihe, E. (1988). Multiple neuropeptides in nerves supplying mammalian lymph nodes: messenger candidates for sensory and autonomic neuroimmunomodulation ? *Neuroscience Letters* **90**, 39-44.

Fisher, F. & Goding, C.R. (1992). Single amino acid substitutions alter helix-loop-helix protein specificity for bases flanking the core CANNTG motif. *The European Molecular Biology Organisation Journal* **11**, 4103-4109.

Flaws, J.A., DeSanti, A., Tilly, K.I., Javid, R.O., Kugu, K., Johnson, A.L., Hirshfield, A.N. & Tilly, J.L. (1995). Vasoactive intestinal peptide-mediated suppression of apoptosis in the ovary: Potential mechanisms of action and evidence of a conserved antiatretogenic role through evolution. *Endocrinology* **136**, 4351-4359.

Flint, J., Rochette, J., Craddock, C.F., Dode, C., Vignes, B., Horsley, S.W., Kearney, L., Buckle, V.J., Ayyub, H. & Higgs, D.R. (1996). Chromosomal stabilisation by a subtelomeric rearrangement involving two closely related Alu elements. *Human Molecular Genetics* **5**, 1163-1169.

Flint, J., Wilkie, A.O., Buckle, V.J., Winter, R.M., Holland, A.J. & McDermid, H.E. (1995). The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nature Genetics* **9**, 132-140.

Forsgren, S. (1989). Vasoactive intestinal polypeptide-like immunoreactivity in the bovine heart: high degree of coexistence with neuropeptide Y-like immunoreactivity. *Cell and Tissue Research* **256**, 125-135.

Foster, N. & Lee, D.L. (1996). Vasoactive intestinal polypeptide-like and peptide histidine isoleucine-like proteins excreted/secreted by *Nippostrongylus brasiliensis*, *Nematodirus battus* and *Ascaridia galli*. *Parasitology* **113**, 287-292.

Foulkes, N.S., Borjigin, J., Snyder, S.H. & Sassone-Corsi, P. (1996). Transcriptional control of circadian hormone synthesis via the CREM feedback loop. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 14140-14145.

Fournier, A., Saunders, J.K. & St-Pierre, S. (1984). Synthesis, conformational studies and biological activities of VIP and related fragments. *Peptides* **5**, 169-177.

Francis, F., Zehetner, G., Höglund, M. & Lehrach, H. (1994). Construction and preliminary analysis of the ICRF human P1 library. *Genetic Analysis Techniques and Applications* **11**, 148-157.

Frischauf, A.-M., Lehrach, H., Poustka, A. & Murray, N. (1983). Lambda replacement vectors carrying polylinker sequences. *Journal of Molecular Biology* **170**, 827-842.

Fujii, H. & Hamada, H. (1993). A CNS-specific POU transcription factor, Brn-2, is required for establishing mammalian neural cell lineages. *Neuron* **11**, 1197-1206.

Fujita, T., Kimura, Y., Miyamoto, M., Barsoumian, E.L. & Taniguchi, T. (1989). Induction of endogenous IFN-α and INF-β genes by a regulatory transcription factor, IRF-1. *Nature* **337**, 270-272.

Galvagni, F., Lestingi, M., Cartocci, E. & Oliviero, S. (1997). Serum response factor and protein-mediated DNA bending contribute to transcription of the dystrophin muscle-specific promoter. *Molecular and Cellular Biology* **17**, 1731-1743.

Ganea, D. (1996). Regulatory effects of vasoactive intestinal peptide on cytokine production in central and peripheral lymphoid organs. *Advances in Neuroimmunology* **6**, 61-74.

Ganea, D., Martinez, C., Delgado, M., Sun, W. & Gomariz, R.P. (1996). Vip and Pacap-38 Inhibit Il-10 Production In Murine T-Lymphocytes. *Faseb Journal* **10**, 426-426.

Ganea, D. & Sun, L. (1993). Vasoactive-Intestinal-Peptide Down-Regulates the Expression Of Il-2 But Not Of Ifn-Gamma From Stimulated Murine T-Lymphocytes. *Journal Of Neuroimmunology* **47**, 147-158.

Gaylinn, B.D., Harrison, J.K., Zysk, J.R., Lyons, C.E., Lynch, K.R. & Thorner, M.O. (1993). Molecular cloning and expression of a human anterior pituitary receptor for growth hormone-releasing hormone. *Molecular Endocrinology* **7**, 77-84.

Gaytan, F., Martinez-Fuentes, A. J., Garcia-Navarro, F., Vaudry, H. & Aguilar, E. (1994). Pituitary adenylate cyclase-activating peptide (PACAP) immunolocalization in lymphoid tissues of the rat. *Cell and Tissue Research* **276**, 223-227.

Georg, B., Wulff, B.S. & Fahrenkrug, J. (1994). Characterization of the effects of retinoic acid on vasoactive intestinal polypeptide gene expression in neuroblastoma cells. *Endocrinology* **135**, 1455-1463.

Gespach, C., Bawab, W., Chastre, E., Emami, S., Yanaihara, N. & Rosselin, G. (1988). Pharmacology and molecular identification of vasoactive intestinal peptide (VIP) receptors in normal and cancerous gastric mucosa in man. *Biochemical and Biophysical Research Communications* **151**, 939-947.

Gibson, S.J., Polak, J.M., Anand, P., Blank, M.A., Morrison, J.F., Kelly, J.S. & Bloom, S.R. (1984). The distribution and origin of VIP in the spinal cord of six mammalian species. *Peptides* **5**, 201-207.

Giddings, S.J., Chirgwin, J. & Permutt, M.A. (1985). Evaluation of rat insulin messenger RNA in pancreatic and extrapancreatic tissues. *Diabetologia* **28**, 343-347.

Gomariz, R.P., Delgado, M., Naranjo, J.R., Mellstrom, B., Tormo, A., Mata, F. & Leceta, J. (1993). VIP gene expression in rat thymus and spleen. *Brain, Behaviour, and Immunity* **7**, 271-278.

Gomariz, R.P., Leceta, J., Garrido, E., Garrido, T. & Delgado, M. (1994). Vasoactive intestinal peptide (VIP) mRNA expression in rat T and B lymphocytes. *Regulatory Peptides* **50**, 177-184.

Gorn, A.H., Lin, H.Y., Yamin, M., Auron, P.E., Flannery, M.R., Tapp, D.R., Manning, C.A., Lodish, H.F., Krane, S.M. & Goldring, S.R. (1992). Cloning, characterization, and expression of a human calcitonin receptor from an ovarian carcinoma cell line. *Journal of Clinical Investigation* **90**, 1726-1735.

Gotoh, E., Yamagami, T., Yamamoto, H. & Okamoto, H. (1988). Chromosomal assignment of human VIP/PHM-27 gene to 6q26-q27 region by spot blot hybridization and *in situ* hybridization. *Biochemistry International* **17**, 555-562.

Gottschall, P.E., Tatsuno, I.& Arimura, A. (1994). Regulation of interleukin-6 (IL-6) secretion in primary cultured rat astrocytes: synergism of interleukin-1 (IL-1) and pituitary adenylate cyclase activating polypeptide (PACAP). *Brain Research* **637**, 197-203.

Gourlet, P., Vandermeers, A., Vandermeers-Piret, M.C., De Neef, P. & Robberecht, P. (1996a). Addition of the (28-38) peptide sequence of PACAP to the VIP sequence modifies peptide selectivity and efficacy. *International Journal of Peptide and Protein Research* **48**, 391-396.

Gourlet, P., Vertongen, P., Vandermeers, A., Vandermeers-Piret, M.C., Rathe, J., De Neef, P., Waelbroeck, M. & Robberecht, P. (1997). The long-acting vasoactive intestinal polypeptide agonist RO 25-1553 is highly selective of the VIP2 receptor subclass. *Peptides* **18**, 403-408.

Gourlet, P., Vilardaga, J.P., De Neef, P., Waelbroeck, M., Vandermeers, A. & Robberecht, P. (1996b). The C-terminus ends of secretin and VIP interact with the N-terminal domains of their receptors. *Peptides* **17**, 825-829.

Gozes, I., McCune, S.K., Jacobson, L., Warren, D., Moody, T.W., Fridkin, M. & Brenneman, D.E. (1991). An antagonist to vasoactive intestinal peptide affects cellular functions in the central nervous system. *Journal of Pharmacology and Experimental Therapeutics* **257**, 959-966.

Gozes, I., Nakai, H., Byers, M., Avidor, R., Weinstein, Y., Shani, Y. & Shows, T.B. (1987). Sequential expression in the nervous system of C-MYB and VIP genes, located in human chromosomal region 6q24. *Somatic Cell and Molecular Genetics?* **13**, 305-313.

Gozes, I. & Tsafriri, A. (1986). Detection of vasoactive intestinal peptide-encoding messenger ribonucleic acid in the rat ovaries. *Endocrinology* **119**, 2606-2610.

Grass, F.S., Spence, J.E., Saikevych, J.A., Patel, J., Brasington, C., Pencarinha, D., Tilley, D. & Ulm, J. (1995). 7q Deletions and variations in the expression of a holoprosencephaly chromosome region. *American Journal of Human Genetics* **57**, 637-637.

Gray, J.X., Haino, M., Roth, M.J., Macguire, J.E., Jensen, P.N., Yarme, A., Stetler-Stevenson, M.A., Siebenlist, U. & Kelly, K. (1996). CD97 is a processed, 7-transmembrane, heterodimeric receptor associated with inflammation. *Journal of Immunology* **157**, 5438-5447.

Gray, S. & Levine, M. (1996a). Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in *Drosophila*. *Genes and Development* **10**, 700-710.

Gray, S. & Levine, M. (1996b). Transcriptional repression in development. *Current Opinion in Cell Biology* **8**, 358-364.

Graziano, M.P., Hey, P.J., Borkowski, D., Chicchi, G.G. & Stader, C.D. (1993). Cloning and functional expression of a human glucagon-like peptide-1 receptor. *Biochemical and Biophysical Research Communications* **196**, 141-146.

Gremlich, S., Porret, A., Hani, E.H., Cherif, D., Vionnet, N., Froguel, P. & Thorens, B. (1995). Cloning, functional expression, and chromosomal localization of the human pancreatic islet glucose-dependent insulinotropic polypeptide receptor. *Diabetes* **44**, 1202-1208.

Gressens, P., Hill, J.M., Gozes, I., Fridkin, M. & Brenneman, D.E. (1993). Growth factor function of vasoactive intestinal peptide in whole cultured mouse embryos. *Nature* **362**, 155-158.

Gressens, P., Hill, J.M., Paindaveine, B., Gozes, I., Fridkin, M. & Brenneman, D.E. (1994). Severe microcephaly induced by blockade of vasoactive intestinal peptide function in the primitive neuroepithelium of the mouse. *Journal Of Clinical Investigation* **94**, 2020-2027.

Grigoriou, M., Kastrinaki, M.C., Modi, W.S., Theodorakis, K., Mankoo, B., Pachnis, V. & Karagogeos, D. (1995). Isolation of the human MOX2 homeobox gene and localization to chromosome 7p22.1-p21.3. *Genomics* **26**, 550-555.

Groisman, R., Masutani, H., Leibovitch, M.P., Robin, P., Soudant, I., Trouche, D. & Harelbellan, A. (1996). Physical interaction between the mitogen-responsive serum response factor and myogenic basic-helix-loop-helix proteins. *Journal of Biological Chemistry* **271**, 5258-5264.

Gronenborn, A.M., Bovermann, G. & Clore, G.M. (1987). A 1H-NMR study of the solution conformation of secretin. Resonance assignment and secondary structure. *FEBS Letters* **215**, 88-94.

Gudermann, T., Kalkbrenner, F. & Schultz, G. (1996). Diversity and selectivity of receptor-G protein interaction. *Annual Review of Pharmacology and Toxicology* **36**, 429-59.

Guénet, J.L., Simon-Chazottes, D., de Robertis, E. & Bluni, M. (1995). The mouse goosecoïd gene (Gsc) maps to the telomeric part of mouse chromosome 12. *Mammalian Genome* **6**, 816-817.

Guillemot, F. (1995). Analysis of the role of basic-helix-loop-helix transcription factors in the development of neural lineages in the mouse. *Biology of the Cell* **84**, 3-6.

Guo, M., Bier, E., Jan, L.Y. & Jan, Y.N. (1995). *tramtrack* acts downstream of *numb* to specify distinct daughter cell fates during asymmetric cell divisions in the Drosophila PNS. *Neuron* **14**, 913-925.

Gurrieri, F., Trask, B.J., Vandenengh, G., Krauss, C.M., Schinzel, A., Pettenati, M.J., Schindler, D., Dietzband, J., Vergnaud, G., Scherer, S.W., Tsui, L.-C. & Muenke, M. (1993). Physical mapping of the holoprosencephaly critical region on chromosome 7q36. *Nature Genetics* **3**, 247-251.

Guyer, M.S. & Collins, F.S. (1995). How is the Human Genome Project doing, and what have we learned so far. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 10841-10848.

Hahm, S.H. & Eiden, L.E. (1996). Tissue-specific expression of the vasoactive intestinal peptide gene requires both an upstream tissue specifier element and the 5' proximal cyclic AMP-responsive element. *Journal of Neurochemistry* **67**, 1872-1881.

Hamann, J., Eichler, W., Hamann, D., Kerstens, H.M.J., Poddighe, P.J., Hoovers, J.M.N., Hartmann, E., Strauss, M. & van Lier, R.A.W. (1995). Expression cloning and chromosomal mapping of the leukocyte activation antigen CD97, a new seven-span transmembrane molecule of the secretin receptor superfamily with an unusual extracellular domain. *Journal of Immunology* **150**, 1942-1950.

Hamann, J., Hartmann, E. & Lier, R.A.W. (1996). Structure of the human CD97 gene: exon shuffling has generated a new type of seven-span transmembrane molecule related to the secretin receptor superfamily. *Genomics* **32**, 144-147.

Hammerschmidt, M., Brook, A. & McMahon, A.P. (1997). The world according to *hedgehog*. *Trends in Genetics* **13**, 14-21.

Hammond, P.J., Khandan-Nia, N., Withers, D.J., Jones, P.M., Ghatei, M.A. & Bloom, S.R. (1997). Regulation of anterior pituitary galanin and vasoactive intestinal peptide by oestrogen and prolactin status. *Journal of Endocrinology* **152**, 211-219.

Hammond, P.J., Talbot, K., Chapman, R., Ghatei, M.A. & Bloom, S.R. (1993). Vasoactive intestinal peptide, but not pituitary adenylate cyclase-activating peptide, modulates the responsiveness of the gonadotroph to LHRH in man. *Journal of Endocrinology* **137**, 529-532.

Hannibal, J., Ding, J.M., Chen, D., Fahrenkrug, J., Larsen, P.J., Gillette, M.U. & Mikkelsen, J.D. (1997). Pituitary adenylate cyclase-activating peptide (PACAP) in the retinohypothalamic tract: a potential daytime regulator of the biological clock. *Journal of Neuroscience* **17**, 2637-2644.

Harmar, T. & Lutz, E. (1994). Multiple receptors for PACAP and VIP. *Trends in Pharmacological Sciences* **15**, 97-99.

Hartschuh, W., Reinecke, M., Weihe, E. & Yanaihara, N. (1984). VIP-immunoreactivity in the skin of various mammals: immunohistochemical, radioimmunological and experimental evidence for a dual localization in cutaneous nerves and merkel cells. *Peptides* **5**, 239-245.

Hashimoto, H., Ishihara, T., Shigemoto, R., Mori, K. & Nagata, S. (1993). Molecular cloning and tissue distribution of a receptor for pituitary adenylate cyclase-activating polypeptide. *Neuron* **11**, 333-342.

Hatziioannou, A.G., Krauss, C.M., Lewis, M.B. & Halazonetis, T.D. (1991). Familial holoprosencephaly associated with a translocation breakpoint at chromosomal position 7q36. *American Journal of Medical Genetics* **40**, 201-205.

Haugen, B.R., Wood, W.M., Gordon, D.F. & Ridgway, E.C. (1993). A thyrotrope-specific variant of Pit-1 transactivates the thyrotropin-beta promoter. *Journal of Biological Chemistry* **268**, 20818-20824.

He, X., Treacy, M.N., Simmons, D.M., Ingraham, H.A., Swanson, L.W. & Rosenfeld, M.G. (1989). Expression of a large family of POU-domain regulatory genes in mammalian brain development. *Nature* **340**, 35-42.

Heiskanen, M., Peltonen, L. & Palotie, A. (1996). Visual mapping by high resolution FISH. *Trends in Genetics* **12**, 379-382.

Helm, G., Ottesen, B., Fahrenkrug, J., Larsen, J.J., Owman, C., Sjoberg, N.O., Stolberg, B., Sundler, F. & Walles, B. (1981). Vasoactive intestinal polypeptide (VIP) in the human female reproductive tract: distribution and motor effects. *Biology of Reproduction* **25**, 227-234.

Heutink, P., Zguricas, J., van Oosterhout, L., Breedveld, G.J., Testers, L., Sandkuijl, L.A., Snijders, P.J.L.M., Weissenbach, J., Lindhout, D., Hovius, S.E.R. & Oostra, B.A. (1994). The gene for triphalangeal thumb maps to the subtelomeric region of chromosome 7q. *Nature Genetics* **6**, 287-292.

Hill, J.M., Agoston, D.V., Gressens, P. & McCune, S.K. (1994). Distribution of VIP mRNA and two distinct VIP binding sites in the developing rat brain: relation to ontogenic events. *Journal of Comparative Neurology* **342**, 186-205.

Hill, J.M., Harris, A. & Hilton-Clarke, D.I. (1992). Regional distribution of guanine nucleotide-sensitive and guanine nucleotide-insensitive vasoactive intestinal peptide receptors in rat brain. *Neuroscience* **48**, 925-932.

Hill, J.M., McCune, S.K., Alvero, R.J., Glazner, G.W., Henins, K.A., Stanziale, S.F., Keimowitz, J.R. & Brenneman, D.E. (1996). Maternal vasoactive intestinal peptide and the regulation of embryonic growth in the rodent. *Journal of Clinical Investigation* **97**, 202-208.

Hill, R.E., Shaw, P.H., Barth, R.K. & Hastie, N.D. (1985). A genetic locus closely linked to a protease inhibitor gene complex controls the level of multiple RNA transcripts. *Molecular and Cellular Biology* **5**, 2114-2122.

Hing, A.V., Helms, C. & Donis-Keller, H. (1993). VNTR and microsatellite polymorphisms within the subtelomeric region of 7q. *American Journal of Human Genetics* **53**, 509-517.

Hing, A.V., Helms, C., Slaugh, R., Burgess, A., Wang, J.C., Herman, T., Dowton, S.B. & Doniskeller, H. (1995). Linkage of preaxial polydactyly type-2 to 7q36. *American Journal of Medical Genetics* **58**, 128-135.

Hinson, J.P., Ho, M.M., Vinson, G.P. & Kapas, S. (1996). Vasoactive intestinal peptide is a local regulator of adrenocortical function. *Endocrine Research* **22**, 831-838.

Hipskind, R.A. & Nordheim, A. (1991). Functional dissection *in vitro* of the human c-fos promoter. *Journal of Biological Chemistry* **266**, 19583-19592.

Hipskind, R.A., Rao, V.N., Mueller, C.G.F., Reddy, E.S.P. & Nordheim, A. (1991). Ets-related protein Elk-1 is homologous to the c-fos regulatory factor p62TCF. *Nature* **354**, 531-534.

Hisanaga, K., Sagar, S.M., Koistinaho, J., Hicks, K.J. & Sharp, F.R. (1993). VIP-induced stellation and immediate early gene expression in astrocytes: effects of dexamethasone. *Neuroscience Letters* **156**, 57-60.

Hocke, G.M., Barry, D. & Fey, G.H. (1992). Synergistic action of interleukin-6 and glucocorticoids is mediated by the interleukin-6 response element of the rat α_2 macroglobulin gene. *Molecular and Cellular Biology* **12**, 2282-2294.

Hofman, M.A., Zhou, J.N. & Swaab, D.F. (1996). No evidence for a diurnal vasoactive intestinal polypeptide (VIP) rhythm in the human suprachiasmatic nucleus. *Brain Research* **722**, 78-82.

Hohmann, E.L., Elde, R.P., Rysavy, J.A., Einzig, S. & Gebhard, R.L. (1986). Innervation of periosteum and bone by sympathetic vasoactive intestinal peptide-containing nerve fibers. *Science* **232**, 868-871.

Holzwarth, M.A. (1984). The distribution of vasoactive intestinal peptide in the rat adrenal cortex and medulla. *Journal of the Autonomic Nervous System* **11**, 269-283.

Hoshino, M., Li, M., Zheng, L.Q., Suzuki, M., Mochizuki, T. & Yanaihara, N. (1993). Pituitary adenylate cyclase activating peptide and vasoactive intestinal polypeptide: differentiation effects on human neuroblastoma NB-OK-1 cells. *Neuroscience Letters* **159**, 35-38.

Hosoya, M., Kimura, C., Ogi, K., Ohkubo, S., Miyamoto, Y., Kugoh, H., Shimizu, M., Onda, H., Oshimura, M., Arimura, A. & Fujino, M. (1992). Structure of the human pituitary adenylate cyclase activating polypeptide (PACAP) gene. *Biochimica et Biophysica Acta* **1129**, 199-206.

Hosoya, M., Onda, H., Ogi, K., Masuda, Y., Miyamoto, Y., Ohtaki, T., Okazaki, H., Arimura, A. & Fujino, M. (1993). Molecular cloning and functional expression of rat cDNAs encoding the receptor for pituitary adenylate-cyclase activating polypeptide (PACAP). *Biochemical and Biophysical Research Communications* **194**, 133-143.

Hou, J., Schindler, U., Henzel, W.J., Ho, T.C., Brasseur, M. & McKnight, S.L. (1994). An interleukin-4-induced transcription factor: IL-4 Stat. *Science* **265**, 1701-1706.

Hsu, H.-L., Huang, L., Tsan, J.T., Funk, W., Wright, W.E., Hu, J.-S., Kingston, R.E. & Baer, R. (1994). Preferred sequences for DNA recognition by the TAL1 helix-loop-helix proteins. *Molecular and Cellular Biology* **14**, 1256-1265.

Huang, M. & Rorstad, O.P. (1990). PHI preferentially binds to VIP receptors in normal rat tissues. *Peptides* **11**, 1015-1020.

Hunter, K.W., Riba, L., Schalkwyk, L., Clark, M., Resenchuk, S., Beeghly, A., Su, J., Tinkov, F., Lee, P., Ramu, E., Lehrach, H. & Housman, D. (1996). Toward the construction of integrated physical and genetic maps of the mouse genome using interspersed repetitive sequence PCR (IRS-PCR) Genomics. *Genome Research* **6**, 290-299.

Inagaki, N., Yoshida, H., Mizuta, M., Mizuno, N., Fujii, Y., Gonoi, T., Miyazaki, J. & Seino, S. (1994). Cloning and functional characterization of a 3rd pituitary adenylate cyclase-activating polypeptide receptor subtype expressed in insulin-secreting cells. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 2679-2683.

Ishibashi, K., Sasaki, S., Saito, F., Ikeuchi, T. & Marumo, F. (1995). Structure and chromosomal localization of a human water channel (AQP3) gene. *Genomics* **27**, 352-354.

Ishihara, T., Nakamura, S., Kaziro, Y., Takahashi, T., Takahashi, K. & Nagata, S. (1991). Molecular cloning and expression of a cDNA encoding the secretin receptor. *EMBO Journal* **10**, 1635-1641.

Ishihara, T., Shigemoto, R., Mori, K., Takahashi, K. & Nagata, S. (1992). Functional expression and tissue distribution of a novel receptor for vasoactive intestinal polypeptide. *Neuron* **8**, 811-819.

Itoh, N., Obata, K., Yanaihara, N. & Okamoto, H. (1983). Human prepro-vasoactive intestinal polypeptide contains a novel PHI-27-like peptide, PHM-27. *Nature* **304**, 547-549.

Jeftinija, S., Murase, K., Nedeljkov, V. & Randic, M. (1982). Vasoactive intestinal polypeptide excites mammalian dorsal horn neurons both in vivo and in vitro. *Brain Research* **243**, 158-164.

Jelinek, L.J., Lok, S., Rosenberg, G.B., Smith, R.A., Grant, F.J., Biggs, S., Bensch, P.A., Kuijper, J.L., Sheppard, P.O., Sprecher, C.A., O'Hara, P.J., Foster, D., Walker, K.M., Chen, L.H.J., McKernan, P.A. & Kindsvogel, W. (1993). Expression cloning and signalling properties of the rat glucagon receptor. *Science* **259**, 1614-1616.

Jenkins, F., Cockerill, P.N., Bohmann, D. & Shannon, M.F. (1995). Multiple signals are required for function of the human granulocyte-macrophage colony-stimulating factor gene promoter in T cells. *The Journal of Immunology* **155**, 1240-1251.

Jensen, R.T., Tatemoto, K., Mutt, V., Lemp, G.F. & Gardner, J.D. (1981). Actions of a newly isolated intestinal peptide PHI on pancreatic acini. *American Journal of Physiology* **241**, G498-G502.

Jiang, C., Atkinson, D., Towbin, J.A., Splawski, I., Lehmann, M.H., Li, H., Timothy, K., Taggart, R.T., Schwartz, P.J., Vincent, G.M., Moss, A.J. & Keating, M.T. (1994). Two long QT syndrome loci map to chromosomes 3 and 7 with evidence for further heterogeneity. *Nature Genetics* **8**, 141-147.

Jobert, A.-S., Fernandes, I., Turner, G., Coureau, C., Prie, D., Nissenson, R.A., Friedlander, G. & Silve, C. (1996). Expression of alternatively spliced isoforms of the parathyroid hormone (PTH)/PTH-related peptide receptor messenger RNA in human kidney and bone cells. *Molecular Endocrinology* **10**, 1066-1076.

Johannesson, T., Ehlers, S. & Wahlstrom, J. (1997). Complex chromosome rearrangement involving chromosomes 1, 4, and 16 revealed by fluorescence *in situ* hybridization. *Clinical Genetics* **51**, 281-285.

Jonveaux, P., Daniel, M.T., Martel, V., Oksenhendler, E., Maarek, O. & Berger, R. (1996). Isochromosome 7q is a primary, non random chromosomal abnormality in hepatosplenic T-gamma/delta lymphoma. *British Journal Of Haematology* **93**, 1245-1245.

Joun, H., Lanske, B., Karperien, M., Qian, F., Defize, L. & Abou-Samra, A. (1997). Tissue-specific transcription start sites and alternative splicing of the parathyroid hormone (PTH)/PTH-related peptide (PTHrP) receptor gene: a new PTH/PTHrP receptor splice variant that lacks the signal peptide. *Endocrinology* **138**, 1742-1749.

Jüppner, H., Abou-Samra, A.-B., Freeman, M.W., Kong, X.-F., Schipani, E., Richards, J., Kolakowski, L.F., Hock, J., Potts, J.T., Kronenberg, H.M. & Segre, G.V. (1991). A G protein-linked receptor for parathyroid hormone and parathyroid hormone-related peptide. *Science* **254**, 1024-1026.

Kaku, K., Tsuchiya, M., Tanizawa, Y., Okuya, S., Inoue, Y., Kaneko, T. & Yanaihara, N. (1986). Circadian cycles in VIP content and VIP stimulation of cyclic AMP accumulation in the rat pineal gland. *Peptides* **7 Suppl 1**, 193-195.

Kalter, H. (1993). Case reports of malformations associated with maternal diabetes: history and critique. *Clinical Genetics* **43**, 174-179.

Kapas, S. & Clark, A.J. (1995). Identification of an orphan receptor gene as a type I calcitonin gene-related peptide receptor. *Biochemical and Biophysical Research Communications* **217**, 832-838.

Kaplan, M.H., Schindler, U., Smiley, S.T. & Grusby, M.J. (1996). Stat6 is required for mediating responses to IL-4 and for the development of Th2 cells. *Immunity* **4**, 313-319.

Karagianni, N. & Tsawdaroglou, N. (1994). The c-fos serum response element (SRE) confers negative response to glucocorticoids. *Oncogene* **9**, 2327-2334.

Karlsson, O., Thor, S., Norberg, T., Ohlsson, H. & Edlund, T. (1990). Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a
Cys-His domain. *Nature* **344**, 879-882.

Karn, J., Matthes, H.W.D., Gait, M.J. & Brenner, S. (1984). A new selective phage cloning vector, λ2001, with sites for *Xba* I, *Bam*H I, *Hind* III, *Eco*R I, *Sst* I and *Xho* I. *Gene* **32**, 217-224.

Kasper, S., Popescu, R.A., Torsello, A., Vrontakis, M.E., Ikejiani, C. & Friesen, H.G. (1992). Tissue-specific regulation of vasoactive intestinal peptide messenger ribonucleic acid levels by estrogen in the rat. *Endocrinology* **130**, 1796-1801.

Kasson, B.G., Lim, P. & Hsueh, A.J. (1986). Vasoactive intestinal peptide stimulates androgen biosynthesis by cultured neonatal testicular cells. *Molecular and Cellular Endocrinology* **48**, 21-29.

Kato, I., Suzuki, Y., Akabane, A., Yonekura, H., Tanaka, O., Kondo, H., Takasawa, S., Yoshimoto, T. & Okamoto, H. (1994). Transgenic mice overexpressing human vasoactive intestinal peptide (VIP) gene in pancreatic beta cells. Evidence for improved glucose tolerance and enhanced insulin secretion by VIP and PHM-27 *in vivo*. *Journal of Biological Chemistry* **269**, 21223-21228.

Kato, Y., Iwasaki, Y., Iwasaki, J., Abe, H., Yanaihara, N. & Imura, H. (1978). Prolactin release by vasoactive intestinal polypeptide in rats. *Endocrinology* **103**, 554-558.

Kenakin, T. (1995). Agonist-receptor efficacy II: agonist trafficking of receptor signals. *Trends in Pharmacology* **16**, 232-238.

Kimata, H. & Fujimoto, M. (1995). Induction of IgA1 and IgA2 production in immature human fetal B cells and pre-B cells by vasoactive intestinal peptide. *Blood* **85**, 2098-2104.

Kimmerly, W.J., Kyle, A.L., Lustre, V.M., Martin, C.H. & Palazzolo, M.J. (1994). Direct sequencing of terminal regions of genomic P1 clones: a general strategy for the design of sequence-tagged site markers. *Genetic Analysis Techniques and Applications* **11**, 117-128.

Kimura, S., Ohshige, Y., Lin, L., Okumura, T., Yanaihara, C., Yanaihara, N. & Shiotani, Y. (1994). Localization of pituitary adenylate cyclase-activating polypeptide (PACAP) in the hypothalamus-pituitary system in rats: light and electron microscopic immunocytochemical studies. *Journal of Neuroendocrinology* **6**, 503-507.

Kirchhoff, S., Schaper, F. & Hauser, H. (1993). Interferon regulatory factor-1 (IRF-1) mediates cell-growth inhibition by transactivation of downstream target genes. *Nucleic Acids Research* **21**, 2881-2889.

Kofman-Alfaro, S., Speed, R.M., Boyle, S. & Chandley, A.C. (1994). Condensation behaviour of the human X chromosome in male germ cells and Sertoli cells examined by fluorescence *in situ* hybridization. *Chromosome Research* **2**, 439-444.

Kong, X.-F., Schipani, E., Lanske, B., Joun, H., Karperien, M., Defize, L.H.K., Jüppner, H., Potts, J.T., Segre, G.V., Kronenberg, H.M. & Abou-Samra, A.B. (1994). The rat, mouse and human genes encoding the receptor for parathyroid hormone and parathyroid hormone-related peptide are highly homologous. *Biochemical and Biophysical Research Communications* **200**, 1290-1299.

Kononen, J., Paavola, M., Penttila, T.L., Parvinen, M. & Pelto-Huikko, M. (1994). Stage-specific expression of pituitary adenylate cyclase-activating polypeptide (PACAP) mRNA in the rat seminiferous tubules. *Endocrinology* **135**, 2291-2294.

Kopf, M., Le Gros, G., Bachmann, M., Lamers, M.C., Bluethmann, H. & Köhler, G. (1993). Disruption of the murine IL-4 gene blocks Th2 cytokine responses. *Nature* **362**, 245-248.

Kordula, T. & Travis, J. (1996). The role of Stat and C/EBP transcription factors in the synergistic activation of rat serine-protease inhibitor-3 gene by interleukin-6 and dexamethasone. *Biochemical Journal* **313**, 1019-1027.

Kozak, M. (1996). Interpreting cDNA sequences: some insights from studies on translation. *Mammalian Genome* **7**, 563-574.

Krejs, G.J., Fordtran, J.S., Fahrenkrug, J., Schaffalitzky de Muckadell, O.B., Fischer, J.E., Humphrey, C.S., O'Dorisio, T.M., Said, S.I., Walsh, J.H. & Shulkes, A.A. (1980). Effect of VIP infusion on water and ion transport in the human jejunum. *Gastroenterology* **78**, 722-727.

Krempels, K., Usdin, T.B., Harta, G. & Mezey, E. (1995). PACAP acts through VIP type 2 receptors in the rat testis. *Neuropeptides* **29**, 315-320.

Kubota, A., Yamada, Y., Yasuda, K., Someya, Y., Ihara, Y., Kagimoto, S., Watanabe, R., Kuroe, A., Ishida, H. & Seino, Y. (1997). Gastric inhibitory polypeptide activates MAP kinase through the wortmannin-sensitive and -insensitive pathways. *Biochemical and Biophysical Research Communications* **235**, 171-175.

Kumar, S., White, D.L., Takai, S., Turczynowicz, S., Juttner, C.A. & Hughes, T.P. (1995). Apoptosis regulatory gene NEDD2 maps to human chromosome segment 7q34-35, a region frequently affected in haematological neoplasms. *Human Genetics* **95**, 641-644.

Kunz, J., Scherer, S.W., Klawitz, I., Soder, S., Du, Y.-Z., Speich, N., Kalff-Suske, M., Heng, H.H.Q., Tsui, L.-C. & Grzeschik, K.-H. (1994). Regional localization of 725 human chromosome 7-specific yeast artificial chromosome clones. *Genomics* **22**, 439-448.

Laburthe, M., Couvineau, A., Gaudin, P., Maoret, J.-J., Rouyer-Fessard, C. & Nicole, P. (1996). Receptors for VIP, PACAP, secretin, GRF, glucagon, GLP-1, and other members of their new family of G protein-linked receptors: structure-function relationship with special reference to the human VIP-1 receptor. *Annals of the New York Academy of Sciences* **805**,

Laitinen, A., Partanen, M., Hervonen, A., Pelto-Huikko, M. & Laitinen, L.A. (1985). VIP like immunoreactive nerves in human respiratory tract. Light and electron microscopic study. *Histochemistry* **82**, 313-319.

Lalli, E. & Sassone-Corsi, P. (1995). Thyroid-stimulating hormone (TSH)-directed induction of the CREM gene in the thyroid-gland participates in the long-term desensitization of the TSH receptor. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 9633-9637.

Lam, K.S. (1991). Vasoactive intestinal peptide in the hypothalamus and pituitary. *Neuroendocrinology* **53 Suppl 1**, 45-51.

Lamas, M., Molina, C., Foulkes, N.S., Jansen, E. & Sassone-Corsi, P. (1997). Ectopic ICER expression in pituitary corticotroph AtT20 cells: effects on morphology, cell cycle, and hormonal production. *Molecular Endocrinology* **11**, 1425-1434.

Lanske, B., Karaplis, A.C., Lee, K., Luz, A., Vortkamp, A., Pirro, A., Karperien, M., Defize, L.H.K., Ho, C., Mulligan, R.C., Abou-Samra, A.-B., Jüppner, H., V., S.G. & Kronenberg, H.M. (1996). PTH/PTHrP receptor in early development and Indian hedgehog-related bone growth. *Science* **273**, 663-666.

Larsen, J.J., Otteson, B., Fahrenkrug, J. & Fahrenkrug, L. (1981). Vasoactive intestinal polypeptide (VIP) in the male genitourinary tract: concentration and motor effect. *Investigative Urology* **19**, 211-213.

Larsson, L.I., Fahrenkrug, J., Schaffalitzky, D., Muckadell, O., Sundler, F., Hakanson, R. & Rehfeld, J.R. (1976). Localization of vasoactive intestinal polypeptide (VIP) to central and peripheral neurons. *Proceedings of the National Academy of Sciences of the United States of America* **73**, 3197-3200.

Latinkic, B.V. & Lau, L.F. (1994). Transcriptional activation of the immediate-early gene pip92 by serum growth factors requires both Ets and CArG-like elements. *Journal of Biological Chemistry* **269**, 23163-23170.

Latinkic, B.V., O'Brien, T.P. & Lau, L.F. (1991). Promoter function and structure of the growth factor-inducible immediate early gene cyr61. *Nucleic Acids Research* **19**, 3261-3267.

Leceta, J., Martinez, C., Delgado, M., Garrido, E. & Gomariz, R.P. (1996). Expression of vasoactive intestinal peptide in lymphocytes: a possible endogenous role in the regulation of the immune system. *Advances in Neuroimmunology* **6**, 29-36.

Lee, A.T., Plump, A., DeSimone, C., Cerami, A. & Bucala, R. (1995a). A role for DNA mutations in diabetes-associated teratogenesis in transgenic embryos. *Diabetes* **44**, 20-24.

Lee, J.E., Hollenberg, S.M., Snider, L., Turner, D.L., Lipnick, N. & Weintraub, H. (1995b). Conversion of *Xenopus* ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science* **268**,
Lelièvre, V., Becq-Giraudon, L., Meunier, A.-C. & Muller, J.-M. (1996). Switches in the expression and function of PACAP and VIP receptors during phenotypic interconversion in human neuroblastoma cells. *Neuropeptides* **30**, 313-322.

Lengauer, C., Henn, T., Onyango, P., Francis, F., Lehrach, H. & Weith, A. (1994). Large-scale isolation of human 1p36-specific P1 clones and their use for fluorescence *in situ* hybridization. *Genetic Analysis Techniques and Applications* **11**, 140-147.

Li, P., He, X., Gerrero, M.R., Mok, M., Aggarwal, A. & Rosenfeld, M.G. (1993). Spacing and orientation of bipartite DNA-binding motifs as potential functional determinants for POU domain factors. *Genes & Development* **7**, 2483-2496.

Li, S., Crenshaw, E.B., Rawson, E.J., Simmons, D.M., Swanson, L.W. & Rosenfeld, M.G. (1990). Dwarf locus mutants, which lack three pituitary cell types, result form mutations in the POU domain gene, Pit-1. *Nature* **347**, 528-533.

Liaw, C.W., Lovenberg, T.W., Barry, G., Oltersdorf, T., Grigoriadis, D.E. & De Souza, E.B. (1996). Cloning and characterization of the human corticotropin-releasing factor-2 receptor complementary deoxyribonucleic acid. *Endocrinology* **137**, 72-77.

Lin, H.Y., Harris, T.L., Flannery, M.S., Aruffo, A., Kaji, E.H., Gorn, A., Kolakowski, L.F., Lodish, H.F. & Goldring, S.R. (1991). Expression cloning of an adenylate cyclase-coupled calcitonin receptor. *Science* **254**, 1022-1024.

Lin, S.-C., Lin, C.R., Gukovsky, I., Lusis, A.J., Sawchenko, P.E. & Rosenfeld, M.G. (1993). Molecular basis of the *little* mouse phenotype and implications for cell type-specific growth. *Nature* **364**, 208-213.

Litwin, D.K., Wilson, A.K. & Said, S.I. (1992). Vasoactive intestinal polypeptide (VIP) inhibits rat alveolar macrophage phagocytosis and chemotaxis *in vitro*. *Regulatory Peptides* **40**, 63-74.

Liu, B., Mortrud, M. & Low, M.J. (1995). DNA elements with AT-rich core sequences direct pituitary cell-specific expression of the pro-opiomelanocortin gene in transgenic mice. *Biochemical Journal* **312**, 827-832.

Lo, K., Landau, N.R. & Smale, S.T. (1991). LyF-1, a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific genes. *Molecular and Cellular Biology* **11**, 5229-5243.

Lok, S., Kuijper, J.L., Jelinek, L.J., Kramer, J.M., Whitmore, T.E., Sprecher, C.A., Mathewes, S., Grant, F.J., Biggs, S.H., Rosenberg, G.B., Sheppard, P.O., O'Hara, P.J., Foster, D.C. & Kindsvogel, W. (1994). The human glucagon receptor encoding gene: structure, cDNA sequence and chromosomal localization. *Gene* **140**, 203-209.

Lorén, I., Emson, P.C., Fahrenkrug, J., Bjorklünd, A., Alumets, J., Håkanson, R. & Sundler, F. (1979). Distribution of vasoactive intestinal polypeptide in the rat and mouse brain. *Neuroscience* **4**, 1953-1976.

Lorén, I., Tornqvist, K. & Alumets, J. (1980). VIP (vasoactive intestinal polypeptide)-immunoreactive neurons in the retina of the rat. *Cell and Tissue Research* **210**, 167-170.

Lovenberg, T.W., Liaw, C.W., Grigoriadis, D.E., Clevenger, W., Chalmers, D.T., De Souza, E.B. & Oltersdorf, T. (1995). Cloning and characterization of a functionally distinct corticotropin-releasing factor receptor subtype from rat brain. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 836-840.

Luckman, S.M. & Cox, H.J. (1995). Expression of inducible cAMP early respressor (ICER) in hypothalamic magnocellular neurons. *Molecular Brain Research* **34**, 231-238.

Lumsden, A. (1995). A 'LIM code' for motor neurons? *Current Biology* **5**, 491-495.

Lund, P.K., Goodman, R.H., Dee, P.C. & Habener, J.F. (1982). Pancreatic preproglucagon cDNA contains two glucagon-related coding sequences arranged in tandem. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 345-349.

Lundberg, J.M., Fahrenkrug, J., Hokfelt, T., Martling, C.R., Larsson, O., Tatemoto, K. & Anggard, A. (1984). Co-existence of peptide HI (PHI) and VIP in nerves regulating blood flow and bronchial smooth muscle tone in various mammals including man. *Peptides* **5**, 593-606.

Lutz, E.M., Sheward, W.J., West, K.M., Morrow, J.A., Fink, G. & Harmar, A.J. (1993). The VIP_2 receptor: molecular characterisation of a cDNA encoding a novel receptor for vasoactive intestinal peptide. *FEBS Letters* **334**, 3-8.

Lynch, S.A., Bond, P.M., Copp, A.J., Kirwan, W.O., Nour, S., Balling, R., Mariman, E., Burn, J. & Strachan, T. (1995). A gene for autosomal dominant sacral agenesis maps to the holoprosencephaly region at 7q36. *Nature Genetics* **11**, 93-95.

Macdonald, R., Barth, K.A., Xu, Q., Holder, N., Mikkola, I. & Wilson, S.W. (1995). Midline signalling is required for Pax gene regulation and patterning of the eyes. *Development* **121**, 3267-3278.

MacNeil, D.J., Occi, J.L., Hey, P.J., Strader, C.D. & Graziano, M.P. (1994). Cloning and expression of a human glucagon receptor. *Biochemical and Biophysical Research Communications* **198**, 328-334.

Maget, B., Tastenoy, M. & Svoboda, M. (1994). Sequencing of eleven introns in genomic DNA encoding rat glucagon receptor and multiple alternative splicing of its mRNA. *FEBS Letters* **351**, 271-275.

Maget, B., Tastenoy, M. & Svoboda, M. (1996a). Study of the promoter of the rat glucagon gene: 1/ Tissue specific polymorphism of the 5' end of glucagon receptor mRNA. *Regulatory Peptides* **64**, 116.

Maget, B., Tastenoy, M. & Svoboda, M. (1996b). Study of the promoter of the rat glucagon gene: 2/ Cloning of the promoter domain. *Regulatory Peptides* **64**, 117.

Magistretti, P.J. (1990). VIP neurons in the cerebral cortex. *Trends in Pharmacological Sciences* **11**, 250-254.

Magistretti, P.J., Hof, P.R., Martin, J.-L., Dietl, M. & Palacios, J.M. (1988). High- and low-affinity binding sites for vasoactive intestinal peptide (VIP) in the rat kidney revealed by light microscopic autoradiography. *Regulatory Peptides* **23**, 145-152.

Mangalam, H.J., Albert, V.R., Ingraham, H.A., Kapiloff, M., Wilson, L., Nelson, C., Elsholtz, H. & Rosenfeld, M.G. (1989). A pituitary POU domain protein, Pit-1, activates both growth-hormone and prolactin promoters transcriptionally. *Genes and Development* **3**, 946-958.

Marais, R., Wynne, J. & Treisman, R. (1993). The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell* **73**, 381-393.

Marigo, V., Roberts, D.J., Lee, S.M.K., Tsukurov, O., Levi, T., Gastier, J.M., Epstein, D.J., Gilbert, D.J., Copeland, N.G., Seidman, C.E., Jenkins, N.A., Seidman, J.G., McMahon, A.P. & Tabin, C. (1995). Cloning, expression, and chromosomal location of *SHH* and*IHH* : two human homologues of the *Drosophila* segment polarity gene *Hedgehog*. *Genomics* **28**, 44-51.

Marrero, M.B., Schieffer, B., Paxton, W.G., Heerdt, L., Berk, B.C., Delafontaine, P. & Bernstein, K.E. (1995). Direct stimulation of Jak/STAT pathway by the angiotensin II AT_1 receptor. *Nature* **375**, 247-250.

Martin, G.R., Richman, M., Reinsch, S., Nadeau, J.H. & Joyner, A. (1990). Mapping of the two mouse *engrailed*-like genes: close linkage of *En-1* to *dominant hemimelia (Dh)* on chromosome 1 and of *En-2* to *hemimelic extra-toes (Hx)* on chromosome 5. *Genomics* **6**, 302-308.

Martinez, C., Delgado, M., Gomariz, R.P. & Ganea, D. (1996). Vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide-38 inhibit IL-10 production in murine T-lymphocytes. *Journal of Immunology* **156**, 4128-4136.

Martínez-Frías, M.L., Bermejo, E., García, A., Galán, E. & Prieto, L. (1994). Holoprosencephaly associated with caudal dysgenesis: a clinical-epidemiologic analysis. *American Journal of Medical Genetics* **53**, 46-51.

Maruno, K., Absood, A. & Said, S.I. (1995). VIP inhibits basal and histamine-stimulated proliferation of human airway smooth muscle cells. *American Journal of Physiology: Lung Cellular and Molecular Physiology* **12**, L1047-L1051.

Mathew, R.C., Cook, G.A., Blum, A.M., Metwali, A., Felman, R. & Weinstock, J.V. (1992). Vasoactive intestinal peptide stimulates T lymphocytes to release IL-5 in murine schistosomiasis masoni infection. *Journal of Immunology* **148**, 3572-3577.

Matsui, T., Hirai, M., Wakita, M., Hirano, M. & Kurosawa, Y. (1993). Expression of novel homeobox-containing gene that maps to chromosome 7q36.1 in hematopoietic cells. *FEBS Letters* **322**, 181-185.

Matthes, H., Boschert, U., Amlaiky, N., Grailhe, R., Plassat, J.-L., Muscatelli, F., Mattei, M.-G. & Hen, R. (1993). Mouse 5-hydroxytryptamine$_{5A}$ and 5-hydroxytryptamine$_{5B}$ receptors define a new family of serotonin receptors: cloning, functional expression, and chromosomal localization. *Molecular Pharmacology* **43**, 313-319.

Mayo, K.E. (1992). Molecular cloning and expression of a pituitary-specific receptor for growth hormone-releasing hormone. *Molecular Endocrinology* **6**, 1734-1744.

McCuaig, K.A., Clarke, J.C. & White, J.H. (1994). Molecular cloning of the gene encoding the mouse parathyroid hormone/parathyroid hormone-related peptide receptor. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5051-5055.

McCuaig, K.A., Lee, H.S., Clarke, J.C., Assar, H., Horsford, J. & White, J.H. (1995). Parathyroid hormone/parathyroid hormone related peptide receptor gene transcripts are expressed from tissue-specific and ubiquitous promoters. *Nucleic Acids Research* **23**, 1948-1955.

McRory, J. & Sherwood, N.M. (1997). Two protochordate genes encode pituitary adenylate cyclase-activating polypeptide and related family members. *Endocrinology* **138**, 2380-2390.

Meltzer, P.S., Guan, X.-Y. & Jeffrey, M.T. (1993). Telomere capture stabilizes chromosome breakage. *Nature Genetics* **4**, 252-255.

Merrill, J.E. & Jonakait, G.M. (1995). Interactions of the nervous and immune systems in development, normal brain homeostasis, and disease. *FASEB Journal* **9**, 611-618.

Mertens, F., Johansson, B., Hoglund, M. & Mitelman, F. (1997). Chromosomal imbalance maps of malignant solid tumors: a cytogenetic survey of 3185 neoplasms. *Cancer Research* **57**, 2765-2780.

Metwali, A., Blum, A., Mathew, R., Sandor, M., Lynch, R.G. & Weinstock, J.V. (1993). Modulation of T-lymphocyte proliferation in mice infected with Schistosoma-Mansoni: VIP suppresses mitogen-induced and antigen-induced T-cell proliferation possibly by inhibiting IL-2 production. *Cellular Immunology* **149**, 11-23.

Mills, J.L., Baker, L. & Goldman, A.S. (1979). Malformations in infants of diabetic mothers occur before the seventh gestational week. *Diabetes* **28**, 292-293.

Mink, S., Härtig, E., Jennewein, P., Doppler, W. & Cato, A.C.B. (1992). A mammary cell-specific enhancer in mouse mammary tumor virus DNA is composed of multiple regulatory elements including binding sites for CTF/NFI and a novel transcription factor, mammary cell-activating factor. *Molecular and Cellular Biology* **12**, 4906-4918.

Miyata, A., Arimura, A., Dahl, R.D., Minamino, N., Uehara, A., Jiang, L., Culler, M.D. & Coy, D.H. (1989). Isolation of a novel 38 residue-hypothalamic polypeptide which stimulates adenylate cyclase in pituitary cells. *Biochemical and Biophysical Research Communications* **164**, 567-574.

Miyata, A., Jiang, L., Dahl, R.D., Kitada, C., Kubo, K., Fujino, M., Minamino, N. & Arimura, A. (1990). Isolation of a neuropeptide corresponding to the N-terminal 27 residues of the pituitary adenylate cyclase activating polypeptide with 38 residues (PACAP38). *Biochemical and Biophysical Research Communications* **170**, 643-648.

Miyatake, S., Shlomai, J., Arai, K.-I. & Arai, N. (1991). Characterization of the mouse granulocyte-macrophage colony-stimulating factor (GM-CSF) gene promoter: nuclear factors that interact with an element shared by three lymphokine genes- those for GM-CSF, interleukin-4 (IL-4), and IL-5. *Molecular and Cellular Biology* **11**, 5894-5901.

Mizuki, N., Kimura, M., Ohno, S., Miyata, S., Sato, M., Ando, H., Ishihara, M., Goto, K., Watanabe, S., Yamazaki, M., Ono, A., Taguchi, S., Okumura, K., Nogami, M., Taguchi, H., Ando, A. & Inoko, H. (1996). Isolation of cDNA and genomic clones of a human Ras-related GTP-binding protein gene and its chromosomal localization to the long arm of chromosome 7, 7q36. *Genomics* **34**, 114-118.

Modi, W.S., Jaye, M. & O'Brian, S.J. (1987). Chromosomal localization of a cDNA clone for the human B1 laminin chain. *Cytogenetics and Cell Genetics* **46**, 663.

Molina, C.A., Foulkes, N.S., Lalli, E. & Sassone-Corsi, P. (1993). Inducibility and negative autoregulation of CREM: an alternative promoter directs the expression of ICER, an early response repressor. *Cell* **75**, 875-886.

Momany, F.A. & Bowers, C.Y. (1996). Speculations on the mechanism of hormone-receptor interactions of the secretin/glucagon family of polypeptide hormones derived from computational structural studies. *Annals of the New York Academy of Sciences* **805**, 172-181.

Morice, A., Unwin, R.J. & Sever, P.S. (1983). Vasoactive intestinal peptide caused bronchodilatation and protects against histamine-induced bronchoconstriction in asthmatic subjects. *Lancet* **2**, 1225-1227.

Morichon-Delvallez, N., Delezoide, A.-L. & Vekemans, M. (1993). Holoprosencephaly and sacral agenesis in a fetus with a terminal deletion 7q36-7qter. *Journal of Medical Genetics* **30**, 521-524.

Morrow, J.A., Lutz, E.M., West, K.M., Fink, G. & Harmar, A.J. (1993). Molecular cloning and expression of a cDNA encoding a receptor for pituitary adenylate cyclase activating polypeptide (PACAP). *FEBS Letters* **329**, 99-105.

Mount, S.M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Research* **10**, 459-472.

Muenke, M. (1994). Holoprosencephaly as a genetic model for normal craniofacial development. *Seminars in Developmental Biology* **5**, 504/1-504/9.

Muenke, M., Bone, L.J., Mitchell, H.F., Hart, I., Walton, K., Hall-Johnson, K., Ippel, E.F., Dietz-Band, J., Kvaløy, K., Fan, C.-M., Tessier-Lavigne, M. & Patterson, D. (1995). Physical mapping of the holoprosencephaly critical region in 21q22.3, exclusion of *SIM2* as a candidate gene for holoprosencephaly, and mapping of *SIM2* to a region of chromosome 21 important for Down syndrome. *American Journal of Human Genetics* **57**, 1074-1079.

Muenke, M., Gurrieri, F., Bay, C., Yi, D.H., Collins, A.L., Johnson, V.P., Hennekam, R.C.M., Schaefer, G.B., Weik, L.-A., Lubinsky, M.S., Daack-Hirsch, S., Moore, C.A., Dobyns, W.B., Murray, J.C. & Price, R.A. (1994). Linkage of a human brain malformation, familial holoprosencephaly, to chromosome 7 and evidence for genetic heterogeneity. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 8102-8106.

Muller, J.M., Lelievre, V., Becqgiraudon, L. & Meunier, A.C. (1995). VIP as a cell-growth and differentiation neuromodulator role in neurodevelopment. *Molecular Neurobiology* **10**, 115-134.

Mutt, V. (1988). Vasoactive intestinal polypeptide and related peptides: isolation and chemistry. *Annals of the New York Academy of Sciences* **527**, 1-19.

Nakai, S., Kawano, H., Yudate, T., Nishi, M., Kuno, J., Nagata, A., Jishage, K., Hamada, H., Fujii, H., Kawamura, K., Shiba, K. & Noda, T. (1995). The POU domain transcription factor Brn-2 is required for the determination of specific neuronal lineages in the hypothalamus of the mouse. *Genes & Development* **9**, 3109-3121.

Nathanson, I., Widdicombe, J.H. & Barnes, P.J. (1983). Effect of vasoactive intestinal peptide on ion transport across dog tracheal epithelium. *Journal of Applied Physiology* **55**, 1844-1848.

Naya, F.J., Stellrecht, C.M. & Tsai, M.J. (1995). Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. *Genes and Development* **9**, 1009-1019.

Nelson, C., Albert, V.R., Elsholtz, H.P., Lu, L.I.-W. & Rosenfeld, M.G. (1988). Activation of cell-specific expression of rat growth hormone and prolactin genes by a common transcription factor. *Science* **239**, 1400-1405.

Nilsson, A. (1975). Structure of the vasoactive intestinal octacosapeptide from chicken intestine. The amino acid sequence. *FEBS Letters* **60**, 322-326.

Noble, A., Kemeny, D.M. & Staynov, D.Z. (1993). Generation of Th1- and Th2-like cells *in vitro* is regulated by IL-4 and IFN-gamma. *Journal of Immunology* **150**, A274.

O'Brien, S.J., Womack, J.E., Lyons, L.A., Moore, K.J., Jenkins, N.A. & Copeland, N.G. (1993). Anchored reference loci for comparative genome mapping in mammals. *Nature Genetics* **3**, 103-112.

Ogata, H. & Podolsky, D.K. (1997). Trefoil peptide expression and secretion is regulated by neuropeptides and acetylcholine. *American Journal of Physiology* **273**, G348-G354.

Ogi, K., Miyamoto, Y., Masuda, Y., Habata, Y., Hosoya, M., Ohtaki, T., Masuo, Y., Onda, H. & Fujino, M. (1993). Molecular cloning and functional expression of a cDNA encoding a human pituitary adenylate cyclase activating polypeptide receptor. *Biochemical and Biophysical Research Communications* **196**, 1511-1521.

Ohkubo, S., Kimura, C., Ogi, K., Okazaki, K., Hosoya, M., Onda, H., Miyata, A., Arimura, A. & Fujino, M. (1992). Primary structure and characterization of the precursor to human pituitary adenylate cyclase activating polypeptide. *DNA and Cell Biology* **11**, 21-30.

Ohshima, T., Nagle, J.W., Pant, H.C., Joshi, J.B., Kozak, C.A., Brady, R.O. & Kulkarni, A.B. (1995). Molecular cloning and chromosomal mapping of the mouse cyclin-dependent kinase 5 gene. *Genomics* **28**, 585-588.

Okamoto, S., Okamura, H., Miyake, M., Takahashi, Y., Takagi, S., Akagi, Y., Fukui, K., Okamoto, H. & Ibata, Y. (1991). A diurnal variation of vasoactive intestinal peptide (VIP) mRNA under a daily light-dark cycle in the rat suprachiasmatic nucleus. *Histochemistry* **95**, 525-528.

Okamoto, S., Okamura, H., Terubayashi, H., Akagi, Y., Okamoto, H. & Ibata, Y. (1992). Localization of vasoactive intestinal peptide (VIP) messenger RNA (mRNA) in amacrine cells of rat retina. *Current Eye Research* **11**, 711-715.

Oláh, Z., Lehel, C., Anderson, W.B., Brenneman, D.E. & Agoston, D.v. (1994). Subnanomolar concentration of VIP induces the nuclear translocation of protein kinase C in neonatal rat cortical astrocytes. *Journal of Neuroscience Research* **39**, 355-363.

Ornoy, A., Zaken, V., Abir, R., Yaffe, P. & Raz, I. (1995). Effects on rat embryos of culture in serum of women with gestational diabetes. *Toxicology In Vitro* **9**, 643-651.

Ottesen, B., Hansen, B., Fahrenkrug, J. & Fuchs, A.R. (1984). Vasoactive intestinal peptide (VIP) stimulates oxytocin and vasopressin release from the neurohypophysis. *Endocrinology* **115**, 1648-1650.

Overhauser, J., Mitchell, H.F., Zackai, E.H., Tick, D.B., Rojas, K. & Muenke, M. (1995). Physical mapping of the holoprosencephaly critical region in 18p11.3. *American Journal of Human Genetics* **57**, 1080-1085.

Pan, H.-Q., Wang, Y.-P., Chissoe, S.L., Bodenteich, A., Wang, Z., Iyer, K., Clifton, S.W., Crabtree, J.S. & Roe, B.A. (1994). The complete nucleotide sequences of the sacBII kan domain of the P1 pAd10-sacBII cloning vector and three cosmid cloning vectors: pTCF, svPHEP and LAWRIST16. *Genetic Analysis Techniques and Applications* **11**, 181-186.

Pantaloni, C., Brabet, P., Bilanges, B., Dumuis, A., Houssami, S., Spengler, D., Bockaert, J. & Journot, L. (1996). Alternative splicing in the N-terminal extracellular domain of the pituitary adenylate cyclase-activating polypeptide (PACAP) receptor modulates receptor selectivity and relative potencies of PACAP-27 and PACAP-38 in phospholipase-C activation. *Journal of Biological Chemistry* **271**, 22146-22151.

Pawelzik, H., Dodt, H.U. & Zieglgansberger, W. (1992). Actions of vasoactive intestinal polypeptide (VIP) on neocortical neurons of the rat in vitro. *Neuroscience Letters* **147**, 167-170.

Peers, B., Leonard, J., Sharma, S., Teitelman, G. & Montminy, M.R. (1994). Insulin expression in pancreatic islet cells relies on cooperative interactions between the helix loop helix factor E47 and the homeobox factor STF-1. *Molecular Endocrinology* **8**, 1798-1806.

Peers, B., Voz, M.L., Monget, P., Mathy-Hartert, M., Berwaer, M., Belayew, A. & Martial, J.A. (1990). Regulatory elements controlling pituitary-specific expression of the human prolactin gene. *Molecular and Cellular Biology* **10**, 4690-4700.

Pence, J.C. & Shorter, N.A. (1993). The autocrine function of vasoactive intestinal peptide on human neuroblastoma cell-growth and differentiation. *Archives Of Surgery* **128**, 591-595.

Peng, L., Xi, H. & Rosenfeld, M.G. (1993). Transcriptional activation of the corticotropin releasing hormone gene by POU-domain protein Brn-2. *Journal of Cellular Biochemistry* **S17A**, 200.

Perrin, M.H., Donaldson, C.J., Chen, R., Lewis, K.A. & Vale, W.W. (1993). Cloning and functional expression of a rat brain corticotropin releasing factor (CRF) receptor. *Endocrinology* **133**, 3058-3061.

Petersen, H.V., Serup, P., Leonard, J., Michelsen, B.K. & Madsen, O.D. (1994). Transcriptional regulation of the human insulin gene is dependent on the homeodomain protein STF1/IPF1 acting through the CT boxes. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 10465-10469.

Phillis, J.W., Kirkpatrick, J.R. & Said, S.I. (1978). Vasoactive intestinal polypeptide excitation of central neurons. *Canadian Journal of Physiology and Pharmacology* **56**, 337-340.

Pierce, J.C. & Sternberg, N.L. (1992a). Using bacteriophage P1 system to clone high molecular weight genomic DNA. *Methods in Enzymology* **216**, 549-574.

Pierce, J.C. & Sternberg, N. (1992b). *The bacteriophage P1 cloning system.* In: Techniques for the Analysis of Complex Genomes, (eds). Academic Press Ltd, pp. 39-58.

Pikkarainen, T., Eddy, R., Fukushima, Y., Byers, M., Shows, T., Pihlajaniemi, T., Saraste, M. & Tryggvason, K. (1987). Human laminin B1 chain: a multidomain protein with gene (LAMB1) locus in the q22 region of chromosome 7. *Journal of Biological Chemistry* **262**, 10454-10462.

Pincus, D.W., DiCicco-Bloom, E. & Black, I.B. (1994). Trophic mechanisms regulate mitotic neuronal precursors: role of vasoactive intestinal peptide (VIP). *Brain Research* **663**, 51-60.

Pine, R., Decker, T., Kessler, D.S., Levy, D.E. & Darnell, J.E. (1990). Purification and cloning of interferon-stimulated gene factor 2 (ISGF2): ISGF2 (IRF-1) can bind to the promoters of both beta interferon- and interferon-stimulated genes but is not a primary transcriptional activator of either. *Molecular and Cellular Biology* **10**, 2448-2457.

Pisegna, J.R. & Wank, S.A. (1993). Molecular cloning and functional expression of the pituitary adenylate cyclase-activating polypeptide type-I receptor. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 6345-6349.

Polak, J.M., Pearse, A.G.E., Garaud, J.C. & Bloom, S.R. (1974). Cellular localization of a vasoactive intestinal peptide in the mammalian and avian gastrointestinal tract. *Gut* **15**, 720-724.

Pompa, J.J., Smitherman, T.C., Bedotto, J.B., Eichhorn, E.J., Said, S.I. & Dehmer, G.J. (1990). Direct coronary vasodilation induced by intracoronary vasoactive intestinal peptide. *Journal of Cardiovascular Pharmacology* **16**, 1000-1006.

Poole, S.J., Law, M.L., Kao, F.-T. & Lau, Y.-F. (1989). Isolation and chromosomal localization of the human *En-2* Gene. *Genomics* **4**, 225-231.

Price, M.A., Hill, C. & Treisman, R. (1996). Integration of growth-factor signals at the c-fos serum response element. *Philosophical Transactions of the Royal Society of London Series B- Biological Sciences* **351**, 551-559.

Pritchard, D.K. & Schubiger, G. (1996). Activation of transcription in *Drosophila* embryos is a gradual process mediated by the nucleocytoplasmic ratio. *Genes and Development* **10**, 1131-1142.

Probst, J.C., Landgraf, R. & Behl, C. (1997). Expression of the trefoil polypeptide ITF in PC12 cells. *Biochemistry and Molecular Biology International* **42**, 425-432.

Quick, M., Iversen, L.L. & Bloom, S.R. (1978). Effect of vasoactive intestinal peptide (VIP) and other peptides on cAMP accumulation in rat brain. *Biochemical Pharmacology* **27**, 2209-2213.

Rawlings, S.R., Piuz, I., Schlegel, W., Bockaert, J. & Journot, L. (1995). Differential expression of pituitary adenylate cyclase-activating polypeptide / vasoactive intestinal polypeptide receptor subtypes in clonal pituitary somatotrophs and gonadotrophs. *Endocrinology* **136**, 2088-2098.

Read, D. & Manley, J.L. (1992). Alternatively spliced transcripts of the Drosophila tramtrack gene encode zinc finger proteins with distinct DNA binding specificities. *EMBO Journal* **11**, 1035-1044.

Rhodes, S.J., DiMattia, G.E. & Rosenfeld, M.G. (1994). Transcriptional mechanisms in anterior pituitary cell differentiation. *Current Opinion in Genetics and Development* **4**, 709-717.

Riethman, H.C., Spais, C., Buckingham, J., Grady, D. & Moyzis, R.K. (1993). Physical analysis of the terminal 240 kb of DNA from human chromosome 7q. *Genomics* **17**, 25-32.

Rivera-Pérez, J.A., Mallo, M., Gendron-Maguire, M., Gridley, T. & Behringer, R.R. (1995). *goosecoid* is not an essential component of the mouse gastrula organizer but is required for craniofacial and rib development. *Development* **121**, 3005-3012.

Robb, S., Cheek, T.R., L., H.F., Hall, L.M., Midgley, J.M. & Evans, P.D. (1994). Agonist-specific coupling of a cloned *Drosophila* octopamine/tyramine recepor to multiple second messenger systems. *European Molecular Biology Organisation Journal* **13**, 1325-1330.

Robberecht, P., Cauvin, A., Gourlet, P. & Christophe, J. (1990). Heterogeneity of VIP receptors. *Archives Internationales de Pharmacodynamie et de Therapie* **303**, 51-66.

Robberecht, P., Gourlet, P., Vertongen, P. & Svoboda, M. (1996). Characterization of the VIP receptor from SUP T1 lymphoblasts. *Advances in Neuroimmunology* **6**, 49-57

Robberecht, P., Tatemoto, K., Chatelain, P., Waelbroeck, M., Delhaye, M., Taton, G., De Neef, P., Camus, J.C., Heuse, D. & Christophe, J. (1982). Effects of PHI on vasoactive intestinal peptide receptors and adenylate cylcase activity in lung membranes. A comparison in man, rat, mouse and guinea pig. *Regulatory Peptides* **4**, 241-250.

Robberecht, P., Vervisch, E., De Neef, P., Coy, D.H. & Christophe, J. (1989). Cholera toxin pretreatment of the human lymphoblastic cell line SUP T_1 affects the selectivity of helodermin/VIP receptors. *Regulatory Peptides* **26**, 177-177.

Robberecht, P., Waelbroeck, M., Deneef, P., Tastenoy, M., Gourlet, P., Cogniaux, J. & Christophe, J. (1988). A new type of functional VIP receptor has an affinity for helodermin in human SUP-T1 lymphoblasts. *FEBS Letters* **228**, 351-355.

Roberts, G.W., Woodhams, P.L., Polak, J.M. & Crow, T.J. (1984). Distribution of neuropeptides in the limbic system of the rat: the hippocampus. *Neuroscience* **11**, 35-37.

Robinson, L.J., Weremowicz, S., Morton, C.C. & Michel, T. (1994). Isolation and chromosomal localization of the human endothelial nitric oxide synthase (NOS3) gene. *Genomics* **19**, 350-357.

Roessler, E., Belloni, E., Gaudenz, K., Jay, P., Berta, P., Scherer, S.W., Tsui, L.-C. & Muenke, M. (1996). Mutations in the human *Sonic Hedgehog* gene cause holoprosencephaly. *Nature Genetics* **14**, 357-360.

Roessler, E., Ward, D.E., Gaudenz, K., Belloni, E., Scherer, S.W., Donnai, D., Siegel-Bartelt, J., Tsui, L.-C. & Muenke, M. (1997). Cytogenetic rearrangements involving the loss of the *Sonic Hedgehog* gene at 7q36 cause holoprosencephaly. *Human Genetics* **100**, 172-181.

Rosa, F., Piazza-Hepp, T. & Goetsch, R. (1994). Holoprosencephaly with 1st trimester topical tretinoin. *Teratology* **49**, 418-419.

Rudnick, A., Ling, T.Y., Odagiri, H., Rutter, W.J. & German, M.S. (1994). Pancreatic beta cells express a diverse set of homeobox genes. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 12203-12207.

Said, S.I. (1988). Vasoactive intestinal peptide in the lung. *Annals of the New York Academy of Sciences* **527**, 450-464.

Said, S.I. & Mutt, V. (1970). Polypeptide with broad biological activity: isolation from small intestine. *Science* **169**, 1217-1218.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Samson, W.K., Said, S.I. & McCann, S.M. (1979). Radioimmunologic localization of vasoactive intestinal polypeptide in hypothalamic and extrahypothalamic sites in the rat brain. *Neuroscience Letters* **12**, 265-269.

Sasaki, K., Dockerill, S., Adamiak, D.A., Tickle, I.J. & Blundell, T. (1975). X-ray analysis of glucagon and its relationship to receptor binding. *Nature* **257**, 751-757.

Savage, N.M., Maclachlan, N.A., Joyce, C.A., Moore, I.E. & Crolla, J.A. (1997). Isolated sacral agenesis in a fetus monosomic for 7q36.1→qter. *Journal of Medical Genetics* **34**, 866-868.

Scarbrough, K., Harney, J.P., Rosewell, K.L. & Wise, P.M. (1996). Acute effects of antisense antagonism of a single peptide neurotransmitter in the circadian clock. *American Journal of Physiology: Regulatory Integrative and Comparative Physiology* **39**, R 283-R 288.

Schaffer, M.M., Carney, D.N., Korman, L.Y., Lebovic, G.S. & Moody, T.W. (1987). High affinity binding of VIP to human lung cancer cell lines. *Peptides* **8**, 1101-1106.

Schanen, N.C., Scherer, S.W., Tsui, L.-C. & Francke, U. (1996). Assignment of the 5-hydroxytryptamine (serotonin) receptor 5A gene (HTR5A) to human-chromosome band 7q36.1. *Cytogenetics and Cell Genetics* **72**, 187-188.

Schebalin, M., Said, S.I. & Makhlouf, G.M. (1977). Stimulation of insulin and glucagon secretion by vasoactive intestinal peptide. *American Journal of Physiology* **232**, E197-E200.

Schell, U., Wienberg, J., Köhler, A., Bray-Ward, P., Ward, D.E., Wilson, W.G., Allen, W.P., Lebel, R.R., Sawyer, J.R., Campbell, P.L., Aughton, D.J., Punnett, H.H., Lammer, E.J., Kao, F.-T., Ward, D.C. & Muenke, M. (1996). Molecular characterization of breakpoints in patients with holoprosencephaly and definition of the HPE2 critical region 2p21. *Human Molecular Genetics* **5**, 223-229.

Scherer, S.W., Otulakowski, G., Robinson, B.H. & Tsui, L.-C. (1991). Localization of the human dihydrolipoamide dehydrogenase gene (DLD) to 7q31-32. *Cytogenetics and Cell Genetics* **56**, 176-177.

Scherer, S.W., Tompkins, B.J.F. & Tsui, L.-C. (1992). A human chromosome 7-specific genomic DNA library in yeast artificial chromosomes. *Mammalian Genome* **3**, 179-181.

Schinzel, A. (1984). In: Catalogue of unbalanced chromosome aberrations in man, (eds). De Gruyter, Berlin, New York, pp. 604-609.

Schipani, E., Karga, H., Karaplis, A.C., Potts, J.T., Kronenberg, H.M., Segre, G.V., Abou-Samra, A.B. & Jüppner, H. (1993). Identical cDNAs encode a human renal and bone PTH/PTHrP receptor. *Endocrinology* **132**, 2157-2165.

Schoenherr, C.J., Paquette, A.J. & Anderson, D.J. (1996). Identification of potential target genes for the neuron-restrictive silencer factor. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 9881-9886.

Schonemann, M.D., Ryan, A.K., McEvilly, R.J., O'Connell, S.M., Arias, C.A., Kalla, K.A., Li, P., Sawchenko, P.E. & Rosenfeld, M.G. (1995). Development and survival of the endocrine hypothalamus and posterior pituitary gland requires the neuronal POU domain factor Brn-2. *Genes & Development* **9**, 3122-3135.

Schreiber, E., Himmelmann, A., Malipiero, U., Tobler, A., Stahel, R. & Fontana, A. (1992). Human small cell lung cancer expresses the octamer DNA-binding and nervous system-specific transcription factor N-Oct-3 (Brain-2). *Cancer Research* **52**, 6121-6124.

Schröck, E., Blume, C., Meffert, M.C., Dumanoir, S., Bersch, W., Kiessling, M., Lozanowa, T., Thiel, G., Witkowski, R., Ried, T. & Cremer, T. (1996). Recurrent gain of chromosome arm 7q in low-grade astrocytic tumors studied by comparative genomic hybridization. *Genes Chromosomes & Cancer* **15**, 199-205.

Schütz, B.R., Scheurlen, W., Krauss, J., du Manoir, S., Joos, S., Bentz, M. & Lichter, P. (1996). Mapping of chromosomal gains and losses in primitive neuroectodermal tumors by comparative genomic hybrdization. *Genes, Chromosomes & Cancer* **16**, 196-203.

Sealy, L., Malone, D. & Pawlak, M. (1997). Regulation of the c-fos serum response element by C/EBP beta. *Molecular and Cellular Biology* **17**, 1744-1755.

Searle, A.G., Edwards, J.H. & Hall, J.G. (1994). Mouse homologues of human hereditary disease. *Journal of Medical Genetics* **31**, 1-19.

Segre, G.V. & Goldring, S.R. (1993). Receptors for secretin, calcitonin, parathyroid hormone (PTH)/PTH-related peptide, vasoactive intestinal peptide, glucagonlike peptide 1, growth hormone-releasing hormone, and glucagon belong to a newly discovered G-protein-linked receptor family. *Trends in Endocrinology and Metabolism* **4**, 309-314.

Seidel, H.M., Milocco, L.H., Lamb, P., Darnell, J.E., Stein, R.B. & Rosen, J. (1995). Spacing of palindromic half sites as a determinant of selective STAT (signal transducers and activators of transcription) DNA binding and transcriptional activity. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 3041-3045.

Sexton, P.M., Houssami, S., Hilton, J.M., O'Keeffe, L.M., Center, R.J., Gillespie, M.T., Darcy, P. & Findlay, D.M. (1993). Identification of brain isoforms of the rat calcitonin receptor. *Molecular Endocrinology* **7**, 815-821.

Sheward, W.J., Lutz, E.M. & Harmar, A.J. (1995). The distribution of vasoactive intestinal peptide$_2$ receptor messenger RNA in the rat brain and pituitary gland as assessed by *in-situ* hybridization. *Neuroscience* **67**, 409-418.

Sheward, W.J., Lutz, E.M. & Harmar, A.J. (1996a). Expression of PACAP receptors in the early mouse embryo as assessed by reverse transcription PCR and *in situ* hybridisation. *Neuroscience Letters* **216**, 45-48.

Sheward, W.J., Lutz, E.M. & Harmar, A.J. (1996b). The presence of messenger RNA for subtypes of the vasoactive intestinal peptide receptor during development of the mouse embryo. *Journal of Physiology (London)* **491.P**, 96P-97P.

Shimatsu, A., Kato, Y., Matsushita, N., Katakami, H., Yanaihara, N. & Imura, H. (1981). Immunoreactive vasoactive intestinal polypeptide in rat hypophysial portal blood. *Endocrinology* **108**, 395-398.

Shivers, B.D., Gorcs, T.J., Gottschall, P.E. & Arimura, A. (1991). Two high affinity binding sites for pituitary adenylate cyclase-activating polypeptide have different tissue distributions. *Endocrinology* **128**, 3055-3065

Shrivastava, A. & Calame, K. (1994). An analysis of genes regulated by the multi-functional transcriptional regulator Yin Yang-1. *Nucleic Acids Research* **22**, 5151-5155.

Siebert, J.R., Cohen, M.M., Jr., Sulik, K.K., Shaw, C.-M. & Lemire, R.J. (1990). *Holoprosencephaly: An overview and atlas of cases.* Wiley-Liss, New York.

Siegel, M.D., Zhang, D.-H., Ray, P. & Ray, A. (1995). Activation of the interleukin-5 promoter by cAMP in murine EL-4 cells requires the GATA-3 and CLE0 elements. *The Journal of Biological Chemistry* **270**, 24548-24555.

Simmons, D.M., Voss, J.W., Ingraham, H.A., Holloway, J.M., Broide, R.S., Rosenfeld, M.G. & Swanson, L.W. (1990). Pituitary cell phenotypes involve cell-specific Pit-1 messenger RNA translation and synergistic interactions with other classes of transcription factors. *Genes and Development* **4**, 695-711.

Simonneaux, V., Ouichou, A. & Pevet, P. (1990). Vasoactive intestinal peptide stimulates melatonin release from perifused pineal glands of rats. *Journal of Neural Transmission* **79**, 69-79.

Singh, G., Supp, D.M., Schreiner, C., McNeish, J., Merker, H.-J., Copeland, N.G., Jenkins, N.A., Potter, S.S. & Scott, W. (1991). *legless* insertional mutation: morphological, molecular, and genetic characterization. *Genes and Development* **5**, 2245-2255.

Sorg, O. & Magistretti, P.J. (1991). Characterization of glycogenolysis elicited by vasoactive intestinal peptide, noradrenaline and adenosine in primary cultures of mouse cerebral cortical astrocytes. *Brain Research* **563**, 227-233.

Soulez, M., Rouviere, C.G., Chafey, P., Hentzen, D., Vandromme, M., Lautredou, N., Lamb, N., Kahn, A. & Tuil, D. (1996). Growth and differentiation of C2 myogenic cells are dependent on serum response factor. *Molecular and Cellular Biology* **16**, 6065-6074.

Spangelo, B.L., Isakson, P.C. & MacLeod, R.M. (1990). Production of interleukin-6 by anterior pituitary cells is stimulated by increased intracellular adenosine. *Endocrinology* **127**, 403-409.

Spengler, D., Waeber, C., Pantaloni, C., Holsboer, F., Bockaert, J., Seeburg, P.H. & Journot, L. (1993). Differential signal transduction by five splice variants of the PACAP receptor. *Nature* **365**, 170-175.

Spurr, N.K. (1996). Mapmakers - progress in the Human Genome Project. *Biochemical Society Transactions* **24**, 285-289.

Sreedharan, S.P., Huang, J.X., Cheung, M.C. & Goetzl, E.J. (1995). Structure, expression, and chromosomal localization of the type-I human vasoactive intestinal peptide receptor gene. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 2939-2943.

Sreedharan, S.P., Patel, D.R., Huang, J.-X. & Goetzl, E.J. (1993). Cloning and functional expression of a human neuroendocrine vasoactive intestinal peptide receptor. *Biochemical and Biophysical Research Communications* **193**, 546-553.

Sreedharan, S.P., Patel, D.R., Xia, M., Ichikawa, S. & Goetzl, E.J. (1994). Human vasoactive intestinal peptide$_1$ receptors expressed by stable transfectants couple to two distinct signaling pathways. *Biochemical and Biophysical Research Communications* **203**, 141-148.

Stehle, J.H., Foulkes, N.S., Molina, C.A., Simonneaux, V., Pévet, P. & Sassone-Corsi, P. (1993). Adrenergic signals direct rhythmic expression of transcriptional repressor CREM in the pineal gland. *Nature* **365**, 314-320.

Stehle, J.H., Pfeffer, M., Kuhn, R. & Korf, H.W. (1996). Light-induced expression of transcription factor ICER (inducible cAMP early repressor) in rat suprachiasmatic nucleus is phase-restricted. *Neuroscience Letters* **217**, 169-172.

Sternberg, N. (1990). A bacteriophage P1 cloning system for the isolation, amplification and recovery of DNA fragments as large as 100 kbp. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 103-107.

Sternberg, N., Smoller, D. & Braden, T. (1994). Three new developments in P1 cloning: increased cloning efficiency, improved clone recovery, and a new P1 mouse library. *Genetic Analysis Techniques and Applications* **11**, 171-180.

Stoffel, M., Espinosa, R., Trabb, J.B., Lebeau, M.M. & Bell, G.I. (1994). Human type I pituitary adenylate cyclase activating polypeptide receptor (ADCYAP1R): localization to chromosome band 7p14 and integration into the cytogenetic, physical, and genetic map of chromosome 7. *Genomics* **23**, 697-699.

Straub, S.G. & Sharp, G.W.G. (1996). A wortmannin-sensitive signal transduction pathway is involved in the stimulation of insulin release by vasoactive intestinal polypeptide and pituitary adenylate cyclase-activating polypeptide. *Journal of Biological Chemistry* **271**, 1660-1668.

Streck, R.D., Rajaratnam, V.S., Fishman, R.B. & Webb, P.J. (1995). Effects of maternal diabetes on fetal expression of insulin-like growth factor and insulin-like growth factor binding protein mRNAs in the rat. *Journal of Endocrinology* **147**, R5-R8.

Sulik, K.K., Dehart, D.B., Rogers, J.M. & Chernoff, N. (1995). Teratogenicity of low doses of all-trans retinoic acid in presomite mouse embryos. *Teratology* **51**, 398-403.

Sulik, K.K. & Johnston, M.C. (1982). Embryonic origin of holoprosencephaly: interrelationship of the developing brain and face. *Scanning Electron Microscopy* **Pt 1**, 309-322.

Sun, L. & Ganea, D. (1993). Vasoactive-Intestinal-Peptide Inhibits Interleukin-(Il)-2 and Interleukin-Il-4 Production Through Different Molecular Mechanisms In T-Cells Activated Via the T-Cell Receptor Cd3 Complex. *Journal Of Neuroimmunology* **48**, 59-70.

Sun, Y., Rao, M.S., Landis, S.C. & Zigmond, R.E. (1992). Depolarization increases vasoactive intestinal peptide- and substance P-like immunoreactivities in cultured neonatal and adult sympathetic neurons. *Journal of Neuroscience* **12**, 3717-3728.

Suzuki, Y., McMaster, D., Lederis, K. & Rorstad, O.P. (1984). Characterization of the relaxant effects of vasoactive intestinal peptide (VIP) and PHI on isolated brain arteries. *Brain Research* **322**, 9-16.

Svoboda, M., Tastenoy, M., Van Rampelbergh, J., Goossens, J.-F., De Neef, P., Waelbroeck, M. & Robberecht, P. (1994). Molecular cloning and functional characterization of a human VIP receptor from SUP-T1 lymphoblasts. *Biochemical and Biophysical Research Communications* **205**, 1617-1624.

Symes, A.J., Corpus, L. & Fink, J.S. (1995a). Differences in nuclear signaling by leukemia inhibitory factor and interferon-γ: the role of STAT proteins in regulating vasoactive intestinal peptide gene expression. *Journal of Neurochemistry* **65**, 1926-1933.

Symes, A.J., Lewis, S.E., Corpus, L., Rajan, P., Hyman, S.E. & Fink, J.S. (1994). STAT proteins participate in the regulation of the vasoactive intestinal peptide gene by the ciliary neurotrophic factor family of cytokines. *Molecular Endocrinology* **8**, 1750-1763.

Symes, A.J., Rajan, P., Corpus, L. & Fink, J.S. (1995b). C/EBP-related sites in addition to a STAT site are necessary for ciliary neurotrophic factor-leukemia inhibitory factor-dependent transcriptional activation by the vasoactive intestinal peptide cytokine response element. *Journal of Biological Chemistry* **270**, 8068-8075.

Szele, F.G. & Pritchett, D.B. (1993). High affinity agonist binding to cloned 5-hydroxytryptamine$_2$ receptors is not sensitive to GTP analogs. *Molecular Pharmacology* **43**, 915-920.

Takahashi, R., Valeika, S.R. & Glass, K.W. (1992). A simple method of plasmid transformation of *E. coli* by rapid freezing. *BioTechniques* **13**, 711-715.

Tanaka, N., Kawakami, T. & Taniguchi, T. (1993). Recognition DNA sequences of interferon regulatory factor 1 (IRF-1) and IRF-2, regulators of cell growth and the interferon system. *Molecular and Cellular Biology* **13**, 4531-4538.

Tanaka, S. & Koike, T. (1994). Vasoactive intestinal peptide suppresses neuronal cell death induced by nerve growth factor deprivation in rat sympathetic ganglion cells in vitro. *Neuropeptides* **26**, 103-111.

Tang, H., Sun, L., Xin, Z. & Ganea, D. (1996). Down-regulation of cytokine expression in murine lymphocytes by PACAP and VIP. *Annals of the New York Academy of Sciences* **805**, 768-778.

Tang, H., Welton, A. & Ganea, D. (1995). Neuropeptide regulation of cytokine expression: effects of VIP and Ro 25-1553. *Journal of Interferon and Cytokine Research* **15**, 993-1003.

Taniguchi, T., Harada, H. & Lamphier, M. (1995). Regulation of the interferon system and cell growth by the IRF transcription factors. *Journal of Cancer Research and Clinical Oncology* **121**, 516-520.

Tatsuno, I., Morio, H., Tanaka, T., Uchida, D., Hirai, A., Tamura, Y. & Saito, Y. (1996). Pituitary adenylate cyclase-activating polypeptide (PACAP) is a regulator of astrocytes: PACAP stimulates proliferation and production of interleukin 6 (IL-6), but not nerve growth factor (NGF), in cultured rat astrocyte. *Annals of the New York Academy of Sciences* **805**, 482-488`.

Tatsuno, I., Somogyvari-Vigh, A. & Arimura, A. (1994). Developmental changes of pituitary adenylate-cyclase activating polypeptide (PACAP) and its receptor in the rat brain. *Peptides* **15**, 55-60.

Taylor, A.I. (1968). Autosomal trisomy syndromes: a detailed study of 27 cases of Edwards' syndrome and 27 cases of Patau's syndrome. *Journal of Medical Genetics* **5**, 227-252.

Teclemariam-Mesbah, R., Kalsbeek, A., Pevet, P. & Buijs, R.M. (1997). Direct vasoactive intestinal polypeptide-containing projection from the suprachiasmatic nucleus to spinal projecting hypothalamic paraventricular neurons. *Brain Research* **748**, 71-76.

Theriault, Y., Boulanger, Y. & St-Pierre, S. (1991). Structural determination of the vasoactive intestinal peptide by two-dimensional H-NMR spectroscopy. *Biopolymers* **31**, 459-464.

Thor, S., Ericson, J., Brännström, T. & Edlund, T. (1991). The homeodomain LIM protein Isl-1 is expressed in subsets of neurons and endocrine cells in the adult rat. *Neuron* **7**, 881-889.

Thorens, B. (1992). Expression cloning of a pancreatic β-cell receptor for the gluco-incretin hormone, glucagon-like peptide-1. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 8641-8645.

Tornell, J., Carlsson, B. & Hillensjo, T. (1988). Vasoactive intestinal peptide stimulates oocyte maturation, steroidogenesis, and cyclic adenosine 3',5'-monophosphate production in isolated preovulatory rat follicles. *Biology of Reproduction* **39**, 213-220.

Tosi, S., Harbott, J., Haas, O.A., Douglas, A., Hughes, D.M., Ross, F.M., Biondi, A., Scherer, S.W. & Kearney, L. (1996). Classification of deletions and identification of cryptic translocations involving 7q by fluorescence *in situ* hybridization (FISH). *Leukemia* **10**, 644-649.

Treisman, R., Marais, R. & Wynne, J. (1992). Spatial flexibility in ternary complexes between SRF and its accessory proteins. *The EMBO Journal* **11**, 4631-4640.

Tsai-Morris, C.H., Buczko, E., Geng, Y., Gamboa-Pinto, A. & Dufau, M.L. (1996). The genomic structure of the rat corticotropin releasing factor receptor: a member of the class II G protein-coupled receptors. *Journal of Biological Chemistry* **271**, 14519-14525.

Tsui, L.-C., Donnis-Keller, H. & Grzeschik, K.-H. (1995). Report of the 2nd International Workshop on Human Chromsome 7 Mapping 1994. *Cytogenetics and Cell Genetics* **71**, 1-21.

Tsujii, T., Ishizaka, K. & Winters, S.J. (1994). Effects of pituitary adenylate cyclase-activating polypeptide on gonadotropin secretion and subunit messenger ribonucleic acids in perifused rat pituitary cells. *Endocrinology* **135**, 826-833.

Tsukurov, O., Boehmer, A., Flynn, J., Nicolai, J.P., Hamel, B.C.J., Traill, S., Zaleske, D., Mankin, H.J., Yeon, H., Ho, C., Tabin, C., Seidman, J.G. & Seidman, C. (1994). A complex bilateral polysyndactyly disease locus maps to chromosome 7q36. *Nature Genetics* **6**, 282-286.

Uddman, R., Alumets, J., Hakanson, R., Loren, I. & Sundler, F. (1980). Vasoactive intestinal peptide (VIP) occurs in nerves of the pineal gland. *Experientia* **36**, 1119-1120.

Usdin, T.B., Bonner, T.I. & Mezey, E. (1994). Two receptors for vasoactive intestinal polypeptide with similar specificity and complementary distributions. *Endocrinology* **135**, 2662-2680.

Usdin, T.B., Mezey, E., Button, D.C., Brownstein, M.J. & Bonner, T.I. (1993). Gastric inhibitory polypeptide receptor, a member of the secretin- vasoactive intestinal peptide receptor family, is widely distributed in peripheral organs and the brain. *Endocrinology* **133**, 2861-2870.

Vaccarino, F.M., Hayward, M.D., Le, H.N., Hartigan, D.J., Duman, R.S. & Nestler, E.J. (1993). Induction of immediate early genes by cyclic AMP in primary cultures of neurons from rat cerebral cortex. *Molecular Brain Research* **19**, 76-82.

Vertongen, P., Camby, I., Darro, F., Kiss, R. & Robberecht, P. (1996). VIP and pituitary adenylate cyclase activating polypeptide (PACAP) have an antiproliferative effect on the T98G human glioblastoma cell line through interaction with VIP_2 receptor. *Neuropeptides* **30**, 491-496.

Vertongen, P., Woussen-Colle, M.-C., Cauvin, A., Robberecht, P. & Christophe, J. (1992). Distribution of PACAP-38 and PACAP-27 in rat brain. *Biomedical Research* **13 Suppl. 2**, 377-382.

Volz, A., Goke, R., Lankat-Buttgereit, B., Fehmann, H.C., Bode, H.P. & Goke, B. (1995). Molecular cloning, functional expression, and signal transduction of the GIP receptor cloned from a human insulinoma. *FEBS Letters* **373**, 23-29.

Vortkamp, A., Lee, K., Lanske, B., Segre, G.V., Kronenberg, H.M. & Tabin, C.J. (1996). Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein. *Science* **613-622**,
Wang, H.Y., Xin, Z.C., Tang, H. & Ganea, D. (1996). Vasoactive intestinal peptide inhibits IL-4 production in murine T- cells by a post-transcriptional mechanism. *Journal of Immunology* **156**, 3243-3253.

Wang, M. & Drucker, D.J. (1995). The LIM domain homeobox gene *isl-1* is a positive regulator of islet cell-specific proglucagon gene transcription. *The Journal of Biological Chemistry* **270**, 12646-12652.

Wang, M. & Drucker, D.J. (1996). Activation of amylin gene transcription by the LIM domain homeobox gene *isl-1*. *Molecular Endocrinology* **10**, 243-251.

Wang, Y. & Conlon, J.M. (1993). Neuroendocrine peptides (NPY, GRP, VIP, somatostatin) from the brain and stomach of the alligator. *Peptides* **14**, 537-579.

Wang, Y. & Conlon, J.M. (1995). Purification and structural characterization of vasoactive intestinal polypeptide from the trout and bowfin. *General and Comparative Endocrinology* **98**, 94-101.

Waschek, J.A., Ellison, J., Bravo, D.T. & Handley, V. (1996). Embryonic expression of vasoactive intestinal peptide (VIP) and VIP receptor genes. *Journal Of Neurochemistry* **66**, 1762-1765.

Watanabe, T., Shimamoto, N., Takahashi, A. & Fujino, M. (1995). PACAP stimulates catecholamine release from adrenal medulla: a novel noncholinergic secretagogue. *American Journal of Physiology* **269**, E903-E909.

Wei, X. & Sulik, K.K. (1993). Pathogenesis of craniofacial and body wall malformations induced by ochratoxin A in mice. *American Journal of Medical Genetics* **47**, 862-871.

Wei, Y. & Mojsov, S. (1996). Tissue specific expression of different human receptor types for pituitary adenylate cyclase activating polypeptide and vasoactive intestinal polypeptide: implications for their role in human physiology. *Journal of Neuroendocrinology* **8**, 811-817.

Weihe, E., Nohr, D., Michel, S., Muller, S., Zentel, H.J., Fink, T. & Krekel, J. (1991). Molecular anatomy of the neuro-immune connection. *International Journal of Neuroscience* **59**, 1-23.

Weihe, E., Reinecke, M. & Forssmann, W.G. (1984). Distribution of vasoactive intestinal polypeptide-like immunoreactivity in the mammalian heart. Interrelation with neurotensin- and substance P-like immunoreactive nerves. *Cell and Tissue Research* **236**, 527-540.

Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T.K., Turner, D., Rupp, R., Hollenberg, S., Zhuang, Y. & Lassar, A. (1991). The *myoD* gene family: nodal point during specification of the muscle cell lineage. *Science* **251**, 761-766.

Weitzel, J.N. & Patel, J. (1994). A single P1 clone bearing three genes from human chromosome 11p15.5: *HRC1 HRAS1*, and *RNH*. *Genetic Analysis Techniques and Applications* **11**, 165-170.

Wenger, S.L., Sell, S.L., Painter, M.J. & Steele, M.W. (1997). Inherited unbalanced subtelomeric translocation in a child with 8p- and Angelman syndromes. *American Journal of Medical Genetics* **70**, 150-154.

West, A.P., McKinnell, C., Sharpe, R.M. & Saunders, P.T.K. (1995). Pituitary adenylate cyclase activating polypeptide can regulate testicular germ cell protein synthesis *in vitro*. *Journal Of Endocrinology* **144**, 215-223.

Wilkie, A.O. (1993). Detection of cryptic chromosomal abnormalities in unexplained mental retardation: a general strategy using hypervariable subtelomeric DNA polymorphisms. *American Journal of Human Genetics* **53**, 688-701.

Wilkie, T.M., Chen, Y., Gilbert, D.J., Moore, K.J., Yu, L., Simon, M.I., Copeland, N.G. & Jenkins, N.A. (1993). Identification, chromosomal localisation, and genome organization of mammalian G-protein-coupled receptors. *Genomics* **18**, 175-184.

Wingender, E., Kel, A.E., Kel, O.V., Karas, H., Heinemeyer, T., Dietze, P., Knüppel, R., Romaschenko, A.G. & Kolchanov, N.A. (1997). TRANSFAC, TRRD and COMPEL: towards a federated database system on transcriptional regulation. *Nucleic Acids Research* **25**, 265-268.

Winter, R.M. (1996). What's in a face? *Nature Genetics* **12**, 124-129.

Wood, I.C., Roopra, A. & Buckley, N.J. (1996a). Neural specific expression of the m4 muscarinic acetylcholine receptor gene is mediated by a RE1/NRSE-type silencing element. *The Journal of Biological Chemistry* **271**, 14221-14225.

Wood, W.M., Dowding, J.M., Bright, T.M., McDermott, M.T., Haugen, B.R., Gordon, D.F. & Ridgway, E.C. (1996b). Thyroid hormone receptor β2 promoter activity in pituitary cells is regulated by Pit-1. *The Journal of Biological Chemistry* **271**, 24213-24220.

Wray, V., Kakoschke, C., Nokihara, K. & Naruse, S. (1993). Solution structure of pituitary adenylate cyclase activating polypeptide by nuclear magnetic resonance spectroscopy. *Biochemistry* **32**, 5832-5841.

Wright, W.E., Sassoon, D.A. & Lin, V.K. (1989). Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* **56**, 607-617.

Wulff, B., Moller Knudsen, S., Adelhorst, K. & Fahrenkrug, J. (1997). The C-terminal part of VIP is important for receptor binding and activation, as evidenced by chimeric constructs of VIP/secretin. *FEBS Letters* **413**, 405-408.

Wynshaw-Boris, A. (1996). Model mice and human diseases. *Nature Genetics* **13**, 259-260.

Xia, M., Gaufo, G.O., Wang, Q., Sreedharan, S.P. & Goetzl, E.J. (1996a). Transduction of specific inhibition of HuT 78 human T-cell chemotaxis by type I vasoactive intestinal peptide receptors. *Journal of Immunology* **157**, 1132-1138.

Xia, M., Sreedharan, S.P. & Goetzl, E.J. (1996b). Predominant expression of type II vasoactive intestinal peptide receptors by human T lymphoblastoma cells: transduction of both Ca^{2+} and cyclic AMP signals. *Journal of Clinical Immunology* **16**, 21-30.

Xu, J., Thompson, K.L., Shephard, L.B., Hudson, L.G. & Gill, G.N. (1993). T_3 receptor suppression of Sp1-dependent transcription from the epidermal growth factor receptor promoter via overlapping DNA-binding sites. *The Journal of Biological Chemistry* **268**, 16065-16073.

Yamada, G., Mansouri, A., Torres, M., Stuart, E.T., Blum, M., Schultz, M., De Robertis, E.M. & Gruss, P. (1995a). Targeted mutation of the murine *goosecoid* gene results in craniofacial defects and neonatal death. *Development* **121**, 2917-2922.

Yamada, Y., Hayami, T., Nakamura, K., Kaisaki, P.J., Someya, Y., Wang, C.-Z., Seino, S. & Seino, Y. (1995b). Human gastric inhibitory polypeptide receptor: cloning of the gene (GIPR) and cDNA. *Genomics* **29**, 773-776.

Yamagami, T., Ohsawa, K., Nishizawa, M., Inoue, C., Gotoh, E., Yanaihara, N., Yamamoto, H. & Okamoto, H. (1988). Complete nucleotide sequence of human vasoactive intestinal peptide/PHM-27 gene and its inducible promoter. *Annals of the New York Academy of Sciences* **527**, 87-102.

Yang-Feng, T.L., Xue, F.Y., Zhong, W.W., Cotecchia, S., Frielle, T., Caron, M.G., Lefkowitz, R.J. & Francke, U. (1990). Chromosomal organization of adrenergic receptor genes. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 1516-1520.

Ye, Z.-S., Forman, B.M., Aranda, A., Pascual, A., Park, H.-Y., Casanova, J. & Samuels, H.H. (1988). Rat growth hormone gene expression. *The Journal of Biological Chemistry* **263**, 7821-7829.

Yung, J.-F. (1996). New FISH probes - the end in sight. *Nature Genetics* **14**, 10-12.

Yuwiler, A. (1983). Vasoactive intestinal peptide stimulation of pineal serotonin-N-acetyltransferase activity: general characteristics. *Journal of Neurochemistry* **41**, 146-153.

Zhong, Y. & Pena, L.A. (1995). A novel synaptic transmission mediated by a PACAP-like neuropeptide in *Drosophila*. *Neuron* **14**, 527-536.

Zolnierowicz, S., Cron, P., Solinas-Toldo, S., Fries, R., Lin, H.Y. & Hemmings, B.A. (1994). Isolation, characterization, and chromosomal localization of the porcine calcitonin receptor gene. *The Journal of Biological Chemistry* **269**, 19530-19538.

LOCUS: CONTIG **5'-INTRON 1**

```
   1 tctagaaaaa ctgagcctgt ggtatgggtg ggtcagacct gatagtgagg tgagacaccc
  61 atgaaaacct gctccattca ctgactacct ctgtaaaaca aaaaaaaaag acacaaaggc
 121 tggcctttca gatttgtata ataaaataaa ttacaaggtt gttttcccca gctttgacag
 181 gatattaggt aaaactcatt ttcactttaa agactcataa caatatgaaa tgcatataaa
 241 tcttacattt agtttgcttt gcagggtatt agcaggagtc aggaggagcc accaacccat
 301 catctgatgc gtttttgcaa aggagatttg ccctttacaa atcagttgcg tgttgtcttc
 361 cttctgtgca catgagaggt gcagaagcct tggagacaac cgcagctgta gcccctaatg
 421 ggtcgagtgt ctctcataac caccctcca ggcacagcaa atagattcag tttatgtact
 481 cccctgaaat tgctcccctt attacagggc ggtgtcagaa ggaagtgctg ggaaagaact
 541 gtcgttttgt aaaagtaaga caaacagaga tgaaacaatc gctttcattc tttgttgttg
 601 ttgaagattt tatgttgaac accaggctct tgtagatacg aaggcagcag cacttctgag
 661 ggggagtcag cggatgttac tgggtattgt cttgtcagct cctgaggatc ctcttcttat
 721 gtgctcatac aggcttctaa tgtctctgat aagcttcatc actggaccca ctccatgtgg
 781 aaggtgctag ctgtttggcc gtcattgctg acccaaaact tcactggcaa aaaagcaaag
 841 ggactctgtg accacatcct aagagaacag agggttagct aaaatggctt gtgaagaact
 901 ctgccccaaa tgctgctgct gctgctgctt tgggtatgag aagcctccgt gatagtgagg
 961 actgaaaccc taragtctgg aaaccagagc agccactcaa gtggaaggtg ccaggaaact
1021 gcccaagagg atgtcagtgg ggaagggata ggaaggccag acatcagtag gtagagatga
1081 aacggggcta aragcaggtc tggaggatgc tggagggcac aacacaaggc tgaactcagg
1141 ttcatatcca ggctgtgaac aatgtgaagt cccatcagaa ccaagcagtc accataggca
1201 caagagtctc tctctctctc tctctctctc tctctctctc tctctctctc tctctctctc
1261 tctccwwacc ccaattcaga araatggaga gaccttgaaa gacatgacca gtctttttag
1321 gaccttcgyt cytccaggat acccccagga aaacaagtcc ytcttgtctc cccgtgcctg
1381 tatctcatcc ttgtcaggac tatcacttca tctcatcata tatgccattg tgtggttgga
1441 accctggcgt ctcacccagt tttgaccaat aacatccgga cgtgacttta gaatgcatcc
1501 ctcacaacaa ttcaggtgta tcagctctcc taattctaat tgtaaccctc tgggatggag
1561 caatggcagg ggtcgatgtt gtgatttaca aatgtgcagg tggatcaggt tcactggctg
```

LOCUS: CONTIG **5'-INTRON 1 continued**

1621 aaacagtgga caggggttgt ggacaaagcc agagcccaag cttcctggcc ccaagctgca
1681 gacctagtgc ccagaccttc tcctccataa aatcccctga tggacaaaca caggcagatg
1741 cacagataat ggaagggtga cagaaaacag atagaaacct aggtccaagg gtacacgggg
1801 ctctgatgtc atggagcaca tgatgagctc agtggggaga aagataatgg gcagatagac
1861 accccaaagt gtgaaggata cattttgtac tatttcagac ttctgacttt gggcttgaaa
1921 ttatatytga ataagaagtt taaaacttgc ctgctttact taatcgagta ctattgtytg
1981 ctggtagctt tytgctyttc cctagactgg tctgtcgaca gacagccaaa ccatagctga
2041 acagtggcag tgtytccaca agcctagtgt acaatgagcc acacacgtgt cacctgaggg
2101 tttggggtgc cgatgagcag acccactttt tttttctttt ttttttttc cgttttttat
2161 taggtattta gctcatttac atttccaatg ctataccaaa agtcccccca tacccaaccc
2221 aaccccactc ccntacccac ccactcccct tttttggccc tggcgttccc ctgtactggg
2281 gcatataaag tttgcgtgtc caatgggcct ctctttccag tgatggccgc ctaggccatc
2341 ttttgataca tatgcagcta gagtcaagag ctccggggta ctggttagtt cataatgttg
2401 ttccacctat agggttgcag atccctttag cttcttgggt actttctcta gctcctccat
2461 tgggggccct gtgatccatc caatagctga ctgtgagcat ccacttctgt gtttgctagg
2521 cccaggcata ytctgacaag agacagctat atcagggtcc tttcagcata atcttgctag
2581 tgtatgcaat ggtggagcag acccactttt gcgttgatga agtctctgct ttctgttcaa
2641 cttggcctta ggctctcaag gggagccatg caagagcttt ctcttttctc tctgctcaca
2701 ccagtgcaga cctaaccatg gtcgtccagc cactttctgc acttttttcct tgtgctgcac
2761 acaggcttcg catttgcaca gggagttggt gccagatgct acctcagcgc tcaagttccc
2821 ttacctggaa ttgggcaatg ccctctgtgg gtgagcatct gcttacttga gccaccagct
2881 ctccaaaggc ttgaacatgg cagcccacag tggggtcgca gaggagggag gcagtacagc
2941 tcggttagca gagggagcct gggaggaagg agcgcgagag agcagagctg acaagggaag
3001 tacccaggcc aggaggggggt cggggtgagg cccagggtcc tgcacttcag agggaagtag
3061 gggtggaagg agggacgttt ggacagagat ctggtggacc ggagtgccag agagagactc
3121 cccaaaaagg agaaaggggc gcgaggcggg gactggnccg aggggcgggn ccgggacttg
3181 cagcctggaa ggcGGAGAGG GCGATAGGCG CGAGACTGAG GAAATCCGCG CTGGGAGGCC

LOCUS: CONTIG **5'-INTRON 1 continued**

```
3241 CGGAGCAGGG GACCGTGCTG CTGAGGCGCC AAGGACCGAG GCGGCACGCT GAGCCCAAGA
3301 TGAGGGCGTC GGTGGTGCTG ACCTGCTACT GCTGGTTGCT GGTGCGGgtg agtgaacccg
3361 cgcccacttc ccgaagaggc gcatcctgca gtggaccgcg gggctcacct ccgagtccca
3421 gcacgagcgg gcggagcgtg gggccagggg agacaatcca gagcagaatg cttcctactg
3481 agccttacca aaccctgggc tgctggactc aggggaattt tcctattccc accgccgttg
3541 cgggttcgcc agaagtgggg ttgaagttct cctcgacccc cgcccctggg acgacctggg
3601 tcctggtctg gcagtgctcc tggagttggg aacaccccta gcacccagcg agaagagtgg
3661 ccgctgcccg ggccacctgc tctgactttg cttccgatag aaactgctgt tactctgtgg
3721 tt
```

LOCUS: CONTIG **EXON 2**

```
  1 acccacagcc actaagagtt aacttgcttt acttataagc aaaccataga ctatgaaatg
 61 ggagctcttc aattctgaga gatgaacatt tctagcagcc cattgctctt aagttagtaa
121 taacaacata gaatgtcagt ttactctcct gctcagaaaa agtattctat tggattttat
181 gtttgtttca caaaataaat agtaaagttt ttcttctttt ttttttaata accacatcag
241 gaaaatgaaa tagctgattt gaaatgaatt tcagcttcaa agaaatttag aaaggtgaag
301 cgttggatct tcctgtagat ttttgcttaa ccctctcttg tgtcattttt agttacttct
361 atttctttca cttgaacagG TGAGCAGCAT CCATCCAGAA TGTCGCTTTC ATCTAGAAAT
421 ACAAGAAGAA GAGACAAAAT GTGCAGAGCT GCTAAGCAGC CAAACGGAGA ATCAGAGAGg
481 taggtttggg gattaagtct gggagtggga tcctagctgt ctatacttct cagccccttg
541 aggggttgga ggtcacctac accttaaaaa aaaatagaac cacttgctac tgttctggtg
601 ggccacaagc aaaactaatg tagccttgtg acagtgggga tttcagcaaa ccttttaaaa
661 gaaaacctca atgctttaca ccaacagttt ccatggcttg ctgaggtaac tgttcctaac
721 ataatgttat ggctctggaa tgctctgaaa aggtgactct agttcccagg gaataatagc
781 acatgtctga attcttggct ctaaaatatt agctatcctt gaagttggtc catctaagaa
841 taatagaaac ttatcagaag caaa
```

LOCUS: CONTIG **EXON 3**

```
  1 cacaaagaaa tcacatttga agtgatagat acactaattc attcatcctt caatgtatac
 61 attgtatcaa acakctcatc taccccataa gtatatataa ttattagtca tccgtttaaa
121 aataaatagg ccttttttaaa aacgacacag tgtacattga attattcaca cgttttcaac
181 cacaggcctt cttcaaaagc aaaaagtgag cacacaagct gctcttacct caacagcgct
241 gtgtgagtgt ctcagggaaa aggcttgggc aggaggctgt ctaacagcag tattcttcct
301 gcggactgtt cagacttgct tattctcctg ccacccttct aagcagtgag ttctcatggc
361 ccatgatttg cacctaattc tttgcctctc cttctagCCT GCAGCGGTGT CTGGGACAAC
421 ATCACATGCT GGCGCCCGGC AGACGTTGGG GAAACTGTCA CAGTGCCCTG CCCCAAAGTA
481 TTCAGCAATT TCTACAGCAG ACCAGgtatt tgcattgatt cctaggttta caggggtttt
541 cctctcagcc gtcagttttt cccaaaagaa gcaagctggt ccttgctact tttgtagcta
601 tattcttact ccctagtccc aagggactat gttgggtgag ttctttaagg aagcaggaat
661 acttcagcct ctggagccta tgtaggcaag caaggcagat acagagggga aaacgggatg
721 ataaactctc aggggacaaa gcaggaaaag tgttcacgag tgatagtgga gatgaggg
```

LOCUS: CONTIG **EXON 4**

```
  1 actaagggag gacttctggg caaccttagt cactttctga ttagggcttc agcttttcat
 61 atctagcttc cmaggaccca acatgcgaaa tttactttcc agGAAACATA AGCAAAAACT
121 GCACTAGCGA TGGATGGTCA GAGACATTTC CAGATTTCAT AGATGCGTGT GGCTACAACG
181 ACCCCGAGGA TGAGAGTAAG gtaggtgccc gacccgccac acaagctcca ctcgtgggca
241 ctgcaaacag acaacattct ctatgggtct ccttacctct gtagcagcgc agcttctcga
301 tccacacttg gtctctgata tgctctggca tttgcgattt ggcgactcag cataatttgg
361 gaggcccagg aaggaaggct ttccttacct gtcagcatgt ctggatggga gagactcctg
421 tgagtgggtt attcctgcca cagactagct cttctgggac cctctgaaca ccacatcctt
481 tgtaccatct ggatctcttc atcccattta tagagttg
```

LOCUS: CONTIG **EXON 5**

```
  1 agttccattc atctcagcct tcctttgtct tggccctaaa ggtagccaat acagttcctt
 61 tggataaaaa aattttaaca gaagtcttcc actcgcttat agtggcttgt aagagaagat
121 aggtgttcaa atattgaagg tccttctagg ctcaagatgg aggtagaatt ataaaatgac
181 cctgtctcct gtcttccagc acttggcatg caagagagag agagagagag ggagagagag
241 agagagagag tcctttgtga ctgaaatctc aagtgtcctg tcccccatgc atttggctag
301 cagcttacac atgtgttttt tcatttcttt gtagATCTCG TTTTATATTT TGGTGAAGGC
361 CATTTATACC TTGGGCTACA GTGTTTCTCT GATGTCTCTT ACAACAGGAA GCATAATTAT
421 CTGCCTCTTC AGgtaaggaa caaaaaaaat gagcaattct gcaaagatga tgagaagaac
481 aaacttcact agaggttatt tggtaactca gaggtttaga ctcttgaccc tctgtgtacc
541 aacctataga gaatccttgg ggctagagag atagatcagc atttaaaagt acttgctgtt
601 ggacatgcta ggattcagtt cccaacaccc atctgtgtca catggtcatc tgtaactcca
661 gttacaagga atctcttgct ttcttctggg ctctttgggc tcccatgcac acatgataca
721 cataaactca ctc
```

LOCUS: CONTIG **EXON 6**

```
  1 ttgagaccaa gttttaacct gagaaatggg tgaggtcaca gctgtggctc tgtctcagag
 61 tagaagcatg aagaccacaa gctacagcag aagatgagga taatgggaca cccacattca
121 tgggggaaac aaacacagtt gactgcaggc tcaaagcagc ttctgggtgg cttttgcatt
181 atttattcat actgagttcc aacctaaggg gacaatctga aaagaaagga gagctcatgt
241 gtgtcatgat gcaaactgaa aatgaacaaa aaactgtgat aattgaagtt aaataaatgg
301 taaacttcac gagaagaaag accatagatg ataattaaac atcagtttct gactctccca
361 cttcagGAAG CTGCACTGCA CAAGGAACTA CATCCACCTG AACCTGTTCC TCTCCTTCAT
421 GCTGAGAGCC ATCTCTGTGC TGGTCAAGGA CAGCGTGCTC TACTCCAGCT CAGGTCTACT
481 GCGCTGCCAC GACCAGCCAG CCTCCTGGgt aagactatcc tgagggcaag atctctgggt
541 atggggagac cctagtatct cccaaccttt ttgctctcag tggggcagca ggatcatagt
601 ccttctcacg cctgagcaat ggaaagaagc tggatgggtt caggcctgtg tcagctcgca
661 agaccagggg tccaccagaa ggctgactgg tgtctagcca ctgtgtgggg ctgggaaatt
721 aaaagactct aggccctggt agccaatgct cacagcagga gcacgtgaga aatcctttcg
781 tagatgagca tacactttag tctaaatctt ccagattcca gggaatcctt ctacccgtt
841 cctacagaaa atcaactc
```

LOCUS: CONTIG **EXON 7**

```
  1 aacaaatggc tgccagaaac tccttgccta tactttcatc agagaatcag aggctccctg
 61 cagatggtcc ctggcattgt gtatacaaag ggttgtgcct cagtataaca cattttggc
121 ctctgccagg ccaagatatt aacaatgtct gtccaagtca gcacgctggc tttgctttac
181 aacttatcca aacagatcgc tgccccgcct accataacta tttttaccaa ctggcatgtt
241 ctctcactgt ctctggttcc agGTTGGCTG CAAGCTCAGC CTGGTATTCT TCCAGTACTG
301 TATCATGGCA AACTTCTACT GGCTTCTGGT GGAGGGTCTC TACCTGCACA CCCTCCTGGT
361 AGCCATCCTT CCTCCCAGCA GGTGCTTCCT GGCCTACCTT CTGATCGGAT GGGgtaaggg
421 cttcatattc tccatcttga ttcctatctc ttttgggcaa aatatgtagg tttctgtgca
481 aaaccctact ccagggsata ctccatggtc accttgggtg cactgagggt tccattgatg
541 tcaacaag
```

LOCUS: CONTIG **EXON 8**

```
  1 gagtaggctg acaggcacac ctaactcctg ctgggagccc agcctttcc agagcaggaa
 61 gggatgaaag cttgcatgac ttagggagta aaacgcacag cccctcagac aacaatagag
121 ccagtgagcg tgtctctctc ttacttggcc tctctgtctt caacagatag gtttcagatt
181 ctcggctttc agatatattt gtcctggctg cttagtaaaa tataagtgca aaccttcaca
241 tattgatgaa gacattaagt gagaacttcc atatggtcct ttgattgaaa attcctcttt
301 acatgttctg gttcatgagt accaggggtg gtgcttctca ttgatgaaca ggaccaaatc
361 tgatgggtct ctccacaGCA TCCCCAGTGT GTGTATAGGT GCATGGACAG CAACTCGCCT
421 CTCTTTAGAA GACACAGGgt aagtgctata tgtgcatcca tgcctgtaaa acaaggcctg
481 ctggtactgg tagaatatct gagagctaaa gaaacttgta ccattagaga caaaagactt
541 ccagccagtc aaggcatttc taggcaacat ttagcacccc gtgattaaaa gacatgtaga
601 tgatggtcat gggcaagact cactacaagt ctcagtattt agaagatggc ccaaatgaga
661 ggtgaaaatt aagccatgtc taagtttagc taaaatcggg acaaaccagg cagaatgagt
721 gatatttttc atgactcttt ggtgtcccca agagagccaa gcagaaacaa gcaattttac
781 cgggttgggg ggttgggggt tcacggggga aatgcattaa gatcaggcta tgttggctgc
841 ctgga
```

LOCUS: CONTIG **EXON 9-INTRON 12**

```
   1 tggagataaa ccctactgaa agtcaagacc acttcaatca ttttcaaggc acatttttaa
  61 ttctttgaga cagagatata aacaatttgg aggactatgt tttcaaaagg aagaaaagag
 121 ccttcatggc atccagagca agtgagagca catgccttga attaactaac atcacaactg
 181 tgagacttac cagcccagga gctgcactcc gtgctacctc gccattagca gactgtcagg
 241 ttgactcacc caggaaagcc gcttcactct acatgcaggg cagttagatc tagttggagg
 301 tgactgtgca atcagagtct gattggcaca ttgagaccga gaaaggggaa agttggttct
 361 gcaagcttca tcctcaccca aacagttctg cccagagcct agtttatatg ttgagtcaga
 421 gtgacaccag tttctttctc tctaacagTT GCTGGGACAC AAACGACCAC AGCATCCCCT
 481 GGTGGGTCAT TCGGATGCCC ATTCTAATTT CTATTGTAgt aagtagctgc ctttggagag
 541 gctatgctct gcttttcttt tgtgtacctg taccatacag agtatagcca ccctgcttca
 601 ctgtgtgcaa gtggctacca acttcctaca gggccccatc acctccagag aaacacccag
 661 ggaagagtgg actccaaggc aataagacca gaaggattct agacttggag gagattaaaa
 721 tagtctgctg cccatttcca actgctcagt taaccacaac tttatgcagt agtaccctca
 781 atgcccacac gaggtcctct ttggtgcagt gggcagacct agtacctggg gtagatggtg
 841 gcaaaatcct atagatgggt ggtcagggat aagatgtaaa ggtgctcaca gatcacacag
 901 tcctttcttg tctctgtgca gGTCAACTTT GCCCTCTTCA TCAGCATTGT AAGGATCTTA
 961 CTTCAGAAAC TAACTTCTCC AGATGTTGGT GGCAATGACC AGTCACAGTA CAAgtgagta
1021 tctgcttgtc cagcatgaga aatgtgcagc aaagttcatg gagagtacag cctaatgcca
1081 gtatacagaa ggtcctcagg gtacctaggt ggccctctaa ggctcaacgt gcactgacac
1141 tggagtctgc cggcccagca ctcacagggg ccaagaaacc acactgatag gatgatggtg
1201 ctaaaaataa gaacattctc tcatgtttac catcatgcat agagccttca tgtgagagtg
1261 ttccttaaat gtgtgtcctt ttctgtgata tcacagcctg acatctarct ccctttaacc
1321 acacattgct cttcgtatgt ctgtcctccc aaccagacct tggggctagt ccacacaagc
1381 acagctggcc ttttgtgtct aaacctgctt ttagcaataa tctgacaagt ggtgtgtcac
1441 tttgactggc acagggtcat gaatgccacc agctaatttg tggtaatcca aggaatcagt
1501 aggttaccca taattcacaa aggaaaaata tatcccagct caggtgtgat gtgaacatga
1561 tttcagcgat gttgtatttg gtgcagctca gagcctgggc tcagtttggt gatgctgtat
```

LOCUS: CONTIG **EXON 9-INTRON 12 continued**

```
1621 ttggtacagg tcaggagcct gggttcagat gatcctgtgc tggctaagat ctgcatcacc
1681 ctgggaagtg atcagctccc agccagttgg aagttaacac ctgcacccag gctgcttttc
1741 atagaaaaaa aaaaatatgc agcgtgatgt gggtgctgag ccaccaggat ggctgaacag
1801 agaggggtct tctctgtatg ctgtctcctc tcctgtaaaa tggaatggat gttttgcctg
1861 ttgggtttga gaaacaactg aaagacattc agtgcctggc acagaagaat agagaaagga
1921 cgtggtgggc atttaagtct tagttgtgtc atgctgtgca caaaatacaa gaactaaggt
1981 tctctttgtg gaagaaacct gccaggccac acataacaag gtataagctg ccacatggtg
2041 cctgatcaga cataatgggt ctcatgcatc ccctctgcag ccctggaaag cagggcgggc
2101 attcttagtc gcctttcaat tggatttgga ttttcggcaa atcctgttta gtgcctgtgc
2161 tgacctttc tctgggctcc tggggagtgt tagtgaagaa ggtctttgga tggatacttt
2221 tattttggaa acaattatat agggtaacga agtctgtgct cttccagggt gatgagaagg
2281 gcagagggct gtttcacccc ttgggttgta gcaaagaaag actaaagtcc ttcctctcct
2341 acaggctttc ctaggacaga tgtgtagata ggtcatgtaa atgtcaggac ctggcagctt
2401 ccatatataa ccttagctaa aggaattgtg ctcaagagcc tgctgaccca ccctgctggg
2461 cacagagcct tgagtgtgtg catttgcaag gaagcagctg cagtgacatc cggttggtag
2521 agatacctac cacttgcacc atgaccatga cagaaagtac caggcttctg tgcaggttta
2581 tcaaatggga aaaaccctgc ccttgatccc caaccctgat agaacccaac atgttggagg
2641 aaggcacaca tctctagaat ttgcatactc tgtgtcttag agttaagact gtgataaaac
2701 ttcacgacca aaaagcaagt taggtaggaa agagtgtatg tgacttacat ttccacagca
2761 ctattcctta ctgcaagaac tcaggatagg agctcaaaca gggcaggatc ctggaggcag
2821 gagctgatgc agaagccttg gagggtatgg agggtactgc ttactggctt gctcagcctg
2881 cttgcttgct tgcttgcttg cttgcttgct tgcttgcttg cttttctctt cttttctttt
2941 ctttttgtt tgttgggttg ggttggtttg gtttggtttg gttttcgaga cagcgtttct
3001 ctgtgtagtg tagtcctggc tctcctggaa ctcactctgt agaccaggct gacctggaac
3061 tcagaaattt gcctcccaag tgctgggatt aaaggcgagt gccaccacgg ccggccctca
3121 gcatgctttc ttatagaacc catgaccatc agcccagggg ttgcctcact cactgggctg
3181 ggcacttccc cattgatcac taattaagaa aatgccttac aacgtgatct tttggaggca
```

LOCUS: CONTIG **EXON 9-INTRON 12 continued**

3241 gttcctcaat tgagggtccc tcctttcaga taactctagc ttgtgtggac ttggaaggga
3301 aacatttgaa ataagtgaga gaaagttttc tcctgtgtta gaaggtatct tagttagggt
3361 tttactgctg tgaacagaca ccatgaccaa agcaactctt aaaaaggaca tttaattcgg
3421 gctggcttac aggttcagag gttcagtcaa ttatcatcaa ggcaggaaca tggcagcatc
3481 caggcaagca tgatactgga ggaactaaga gttctacaac ttcatctgga gactgctagc
3541 agaatactgg cttccagaca gctaggatga gggtcttaaa gcccaagttc acagtgacac
3601 acctactcta agggggccac accttyttat cctgccaytc cctgggytaa gcatatacaa
3661 accatcacag aaggtaatga aatgaaaaag aggaacagga accaggagga gcytcatgcy
3721 tcccctagct gaatctagag gagataggct gtgtgcttag tgatagtaag aagccacgag
3781 agggcttccg catgcctgtg taccctccag ctggctgcta agatgagcac tcactcctta
3841 gcagttggag gcacctgggg atggtaagct agatacctgg tccctagtgt ggggctaata
3901 ggacaggata aggacagcag acatcactag caactggaca gaaattcaga ttttcagact
3961 gttgcccaac tcttcctgtt tcataattct gatgaggcct ggaatttgag ctgccacaag
4021 gcctccacac ggtttcaact cagtagaagt ctaggaagct ttggctgaga taactcagat
4081 tccaggaaaa actgatccat aatgactggt gcctatatcc gtaatggaag aacccccccta
4141 tatgtttttg gggatgctgc atgcaatccc ccccccccagt tctgagaatg gaatccaggg
4201 tcctgctcct ttccagtccc accagaacta tctgtccaag agtagtgtgg tccagcagtt
4261 ggcacttgtg acaaagaaca gtaaatttcc ctgtaagctg gtgagccata caatatgccc
4321 tgtaggaatt tcatcttcat gagtaaactt gcatgtccac atgtgaaaat tctgcccaag
4381 cgtgcaaaga agcgggttca caggcacaaa aagagagtaa gcatttgtat attccctgac
4441 tgagtcagcc tctcaaatct ctcatgatct tgggtgggag aattgagagg cacaaggcat
4501 catgtccttc tccttttact agGAGGCTTG CCAAGTCCAC ACTGCTGCTA ATCCCCCTGT
4561 TTGGCGTCCA CTACATGGTG TTTGCTGCCT TCCCTATTGG CATCTCATCC ACATACCAGA
4621 TCCTGTTTGA GTTATGTGTT GGTTCCTTCC AGgtaagtat gggaggtgct atgggtgccg
4681 caaactgtgg gcagctctga cactctgtct ctttgcagGG CCTGGTGGTA GCAGTTCTAT
4741 ACTGCTTCCT GAACAGTGAG gtaagtcctt tgtcagctgg aatttggcac tgacgttggg
4801 tcaggaattg gccccatcat caacctctgc ccagctatgc ccccatgcaa ccaagtccaa

LOCUS: CONTIG **EXON 9-INTRON 12 continued**

```
4861 acttggtttc ctaaagcccc tcttagcaat acttccagtc cttccagcaa aattgagcct
4921 ctgtcctcca cccaactgcc ccttcatgat caggtattta tttagttatt tagttagtta
4981 ctgtatatga gtgctctatc tgcatgtaca catgcatgtc agaagaggtc atcagatctc
5041 agtat
```

LOCUS: CONTIG **INTRON 12 + EXON 13 + 3'UTR**

```
   1 gcctgaggat tacttctctc ctggagccag gaatgggcat gggccagact gctcatcctg
  61 tggctactag agaaaagctc taaaaaccac accagctgtt tgctatctag tggttgaata
 121 cacgctacat gctgctttgg aggattcacg agttcctgca tttcaagctg actcttcact
 181 aacacgctca tccccagGTA CAGTGTGAAC TGAAAAGAAG ATGGCGAGGC CTGTGCCTGA
 241 CCCAAGCTGG GAGCCGGGAC TACCGGCTGC ACAGCTGGTC CATGTCCCGG AATGGCTCAG
 301 AAAGTGCCCT ACAGATACAC CGTGGCTCCC GCACCCAGTC CTTCCTGCAG TCAGAGACTT
 361 CAGTCATTTA GCTGTGTCCC TTGTACAGAG CTGTCAGTGC TGCTGGGTTT GACATATGTT
 421 TGCTGGATTC CTCTGCTGCC CCAGTGTCTG GTGCCTTATT GGTTCAGCCC TGGTCCTTAA
 481 CTCTGATCCT CATGTGTAAC TTGATTGAAA CACCAATTAT TATTGTCAAA CTCTAGCCTT
 541 TAAGCCATCT TCTTCATAAT ATGGCTCAGC CATATTCTAC TTTCAAAGAG AGCAAGGAAG
 601 CCAGGTGGCT GTGAACATCA AAACTGGATT CTAGTACCTA TGAAGAGAAC AAGGAGGAAG
 661 ATTCCAGTTC CCCTACTGCC TCCTGTCAGG GGGAACCAGG TCCAGAAACA GCTGAGCATC
 721 ACCACCCTGT GACACCCAAC GCATGCCTAT GGATGTCTCT GAAATTCCTC CTCAGTCTAT
 781 GGCATACCAT AGAAACAGTT GAGAGCAACC TTCTCATCCT AGAAGCTTCC TGGTCCCCCT
 841 AGCCTTGAGT TGTGTCCTGT GGAACTACAT CTGGACAGCC ACACCCTTGC CATTCCCTGG
 901 ACCAAAGATA AGTCAAAGAG GATATTCATG TTTCATTGGC CAAACCAGGA ACTCATATCC
 961 TCCTGAGGCA CTTAGCATCT GCTGCTTCCT ACAAGTGGTG AGCAGCTCTG GATCCCATGT
1021 TGTCATGAAG GCCCTTGTGA CTTGTCCATT CAGCCCAGGG TTTGTTCTCT AGCCAATAGA
1081 AAACTGTTGC ATTGTTAAAG AAAATCGCCA TCTTAAAAAC TGTTTAAAAA GAGAATTGTT
1141 TTATCAAAGT ATAAAACAGC ATCAAGTATT CCTACCCGAT ACTTGCCAAT AACAATCCAT
1201 ATTTTTCCGA TGTTTTTGCA ACTGTGGGCA AGCAGACCTG AGGAAATGCT TGTCGTGGTT
1261 GGAAAAGCAG CTGAAGACAG AGAGTGCCTT TCCCTGCAGT GCTCAGAGAC CCTGGAAGAC
1321 TATGTCGTGG TCCAGTGCAA GGACTGTAGG TTCTGTACCT AGTGTTCTGG CACAGCCTAT
1381 GTAACTTCCC CCCAACACAC ACACACACAC ACACACACAC ACACACACCG ACAGCAACGG
1441 AATTCCGTTC
```

GENOMICS 37, 345–353 (1996)
ARTICLE NO. 0569

Chromosomal Localization in Mouse and Human of the Vasoactive Intestinal Peptide Receptor Type 2 Gene: A Possible Contributor to the Holoprosencephaly 3 Phenotype

M. MACKAY,* J. FANTES,† S. SCHERER,‡ S. BOYLE,† K. WEST,* L.-C. TSUI,‡ E. BELLONI,‡ E. LUTZ,* V. VAN HEYNINGEN,† AND A. J. HARMAR*,[1]

*MRC Brain Metabolism Unit, Royal Edinburgh Hospital, Morningside Park, Edinburgh, EH10 5HF, United Kingdom; †MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, United Kingdom; and ‡Department of Genetics, Research Institute, The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada

Received May 29, 1996; accepted August 9, 1996

The neuropeptides vasoactive intestinal peptide (VIP) and pituitary adenylate cyclase activating polypeptide (PACAP) have been shown to act on a wide range of tissue and cell types, both in the central nervous system and in the periphery. Two distinct receptors for VIP, the VIP receptor type 1 (VIPR1) and the VIP receptor type 2 (VIPR2), have recently been cloned, each of which binds PACAP and VIP with equal affinity. We report here the chromosomal mapping of the human and mouse VIPR2 genes by fluorescence *in situ* hybridization. The VIPR2 gene maps to the human chromosomal region 7q36.3 and to the F2 region of mouse chromosome 12. Our localization of the human gene places it in the region where the locus for the craniofacial defect holoprosencephaly type 3 (HPE3) maps. Further mapping experiments, carried out on cell lines derived from patients with HPE or HPE microforms and associated 7q deletions, have led us to redefine the distal extent of the HPE3 minimal critical region, originally characterized by Gurrieri *et al.* (1993, *Nature Genet.* 3: 247–251.) The VIPR2 gene lies within this new HPE3 minimal critical region. Our results suggest that deletion of the VIPR2 gene is not the sole factor responsible for the HPE3 phenotype. However, it is possible that monosomy at the VIPR2 locus may contribute to the phenotype observed in many cases of HPE3.

INTRODUCTION

Vasoactive intestinal peptide (VIP), a member of the glucagon/secretin family of peptides, was first isolated over 20 years ago. It functions as a neuroendocrine hormone and neurotransmitter, activating adenylate cyclase (Magistretti and Schorderet, 1984; Quik *et al.*, 1978) through G-protein-coupled receptors (Ishihara *et al.*, 1992; Lutz *et al.*, 1993). VIP has a number of actions in the periphery, including vasodilatation, stimulation of electrolyte secretion, and smooth muscle relaxation (Gozes and Brenneman, 1989). A more recently discovered member of this family, pituitary adenylate cyclase activating polypeptide (PACAP), displays 60% sequence homology to VIP (Miyata *et al.*, 1989). PACAP has a distinct distribution in the central and peripheral nervous systems, where it is thought to function as a neurotransmitter. Possible peripheral targets for PACAP include the adrenal medulla, the pancreas, and the reproductive tract (Shivers *et al.*, 1991).

It is only in the past 4 years, however, following the cloning of receptors through which these peptides act, that we have been able to begin to examine the routes by which VIP and PACAP achieve their diverse effects. Three receptors that bind PACAP have been cloned to date, two of which also bind VIP with an affinity approximately equal to that with which they bind PACAP. The receptors have been designated the PACAP receptor type 1 (ADCYAP1R1), the VIP type 1 receptor (VIPR1), and the VIP type 2 receptor (VIPR2). The PACAP receptor type 1 (Ogi *et al.*, 1990; Spengler *et al.*, 1993) has a marked preference for PACAP, whereas VIPR1 (Ishihara *et al.*, 1992; Sreedharan *et al.*, 1993) and VIPR2 (Inagaki *et al.*, 1994; Lutz *et al.*, 1993; Usdin *et al.*, 1994) will bind both VIP and PACAP with approximately equal affinity. However, VIPR1 and VIPR2 show only about 50% similarity to each other at the amino acid level and exhibit very significant differences in their distribution within the brain. VIPR1 mRNA is most abundant in the cerebral cortex and hippocampus (Ishihara *et al.*, 1992; Usdin *et al.*, 1994), whereas VIPR2 mRNA is found principally in the thalamus and suprachiasmatic nucleus (Sheward *et al.*, 1995; Usdin *et al.*, 1994). The PACAP receptor type 1 and VIPR1 genes have been mapped to human chromosomal regions 7p14 (Stoffel *et al.*, 1994) and 3p22 (Sreedharan *et al.*, 1995), respectively.

[1] To whom correspondence should be addressed at the MRC Brain Metabolism Unit, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF, United Kingdom. Telephone: (+44) 131 537 6533. Fax: (+44) 131 537 6110.

0888-7543/96 $18.00

We report here the chromosomal localizations of the human and mouse VIPR2 genes. Using fluorescence *in situ* hybridization (FISH), we have mapped the VIPR2 gene to the human chromosomal region 7q36.3 (using both genomic and cDNA probes) and to mouse chromosome 12 region F2. This places the human VIPR2 locus close to the locus for the craniofacial abnormality holoprosencephaly type 3 (HPE3), a defect of embryonic development that is characterized by incomplete cleavage of the forebrain (reviewed by Muenke, 1994). We have examined cell lines derived from patients with HPE or HPE microforms associated with chromosomal deletions in the distal region of 7q. Our results redefine the distal extent of the HPE3 minimal critical region, originally characterized by Gurrieri *et al.* (1993), and show that the VIPR2 gene lies within the new minimal critical region.

MATERIALS AND METHODS

Isolation of a mouse genomic clone. A mouse genomic DNA library, in the λ vector EMBL3 (Stratagene), was screened using an [α-^{32}P]dCTP-labeled cDNA probe corresponding to bases 711 to 1192 of the rat VIPR2 cDNA sequence (Lutz *et al.,* 1993). One positive clone was isolated, and a 4-kb *Bam*HI restriction fragment, containing the final three introns and exons of the mouse VIPR2 gene, was subcloned into pBluescript (Stratagene) for subsequent use in FISH. Plasmid DNA for use in FISH was prepared using a Qiagen plasmid maxi kit.

Isolation of a human genomic clone. A gridded human genomic DNA library, constructed in the bacteriophage P1 cloning vector pAd10-SacBII (Francis *et al.,* 1994), was screened using an [α-^{32}P]-dCTP-labeled probe encompassing bases 1–947 of the published rat VIPR2 cDNA sequence (Lutz *et al.,* 1993). One positive clone was identified. P1 clone DNA for use in FISH was prepared by alkaline lysis followed by CsCl banding and RNase treatment (Sambrook *et al.,* 1989).

The identity of the P1 clone was confirmed by PCR amplification of a 1.2-kb fragment from the 3′ end of the VIPR2 gene using the primers 4334 (5′-CCAGGTATGGGGTTTAGTGGAC) and 4335 (5′-GCCGTCCTCTACTGTTTCCTGA), which flank the last intron of the human VIPR2 gene. After denaturation at 94°C for 5 min, 30 cycles of PCR were performed under the following conditions: 1 min at 94°C, 1 min at 57°C, 1.5 min + 4 sec per cycle at 72°C. A final extension at 72°C for 10 min was followed by incubation at 4°C. Sequence analysis of the PCR product after subcloning into pGEM-T vector demonstrated that the product contained genomic sequence encoding bases 1276–1640 of the published human cDNA sequence (Svoboda *et al.,* 1994) and intervening intron sequence, as expected.

Cell lines. Details of the cell lines used in this study (obtained from NIGMS Human Genetic Mutant Cell Repository) are given in Table 1.

Chromosomal mapping by FISH. Metaphase preparations were made from short-term peripheral blood cultures, lymphoblastoid or fibroblast cultures, and mouse ES cells by standard methods.

Probes. P1 clones, cDNAs, cosmids, and YACs were labeled by nick-translation with biotin-16–dUTP or digoxigenin-16–dUTP. Chromosome 7 paint was produced by *Alu*-PCR of DNA from somatic cell hybrids containing a single human chromosome (Kofman-Alfaro *et al.,* 1994). Amplified DNA was labeled with biotin-16–dUTP or digoxigenin-16–dUTP by nick-translation. Chromosome 13 paint labeled with biotin was obtained from Cambio. YAC HSC7E207 contains the DNA markers D7S392 and D7S550, and YAC HSC7E526 contains D7S68 (Kunz *et al.,* 1994). Cosmid 29f5 contains marker D7S168. YAC HTY3006 is a 275-kb chromosome 7 telomeric YAC, originally isolated by H. Riethman (Riethman *et al.,* 1993).

Hybridization and detection. Hybridization conditions were as described in Fantes *et al.* (1992). After single-probe hybridizations biotin was detected by successive layers of avidin–FITC, biotinylated anti-avidin, and avidin–FITC. When two probes, labeled with either biotin or digoxigenin, were hybridized simultaneously, the signals were detected using successive layers of FITC-conjugated sheep anti-digoxigenin, FITC-conjugated anti-sheep antibody plus avidin–Texas Red, biotinylated anti-avidin, and avidin–Texas Red. The VIPR2 gene was localized on prometaphase chromosomes produced by early S phase synchronization with methotrexate of PHA-stimulated peripheral blood cultures, followed by release with bromodeoxyuridine (BrdU). In this case hybridization signals were detected by successive layers of avidin–Texas Red, biotinylated anti-avidin, and avidin–Texas Red. A final incubation with FITC-conjugated anti-BrdU gave a replication G-banding pattern (Craig and Bickmore, 1994). The chromosomes were counterstained with DAPI. Texas Red hybridization signals were visualized simultaneously with FITC signals through a Chroma Pinkel #1 filter block fitted to a Zeiss Axioplan microscope, with a Photometrics CCD camera and Digital Scientific software.

Use of YAC and cosmid clones to determine the position and orientation of the human VIPR2 gene relative to the 7q telomere. YAC clones HSC7E526 and HSC7E1357 were isolated from a chromosome 7-specific library prepared from human–hamster somatic cell hybrid lines (Scherer *et al.,* 1992), E145A7 was isolated from the St. Louis total human genomic YAC library (Brownstein *et al.,* 1989), and HTY146 (locus D7S427) was isolated from a telomeric YAC library that was described previously (Riethman *et al.,* 1993). The cosmid clones 2, 3 (D7S591), 6 (D7S592), 14, 19, 30 (D7S590), 48 (D7S593), and 49 were derived from a library made from YAC clone HTY146 and have been described previously (Riethman *et al.,* 1993). Another 22 cosmid clones (8d1, 14e10, 66e9, 77e7, 79b6, 79g2, 79g3, 81e9, 107g9, 116h3, 122e11, 124g2, 126g6, 134a8, 135c5, 137g2, 164g2, 215a5, 222f1, 230e6, 245b9, 250h9) were isolated from the Lawrence Livermore National Laboratory chromosome 7-specific library by hybridization screening with a human VIPR2 cDNA probe corresponding to the entire coding region (13 exons) of the published sequence (Svoboda *et al.,* 1994). Cosmids 8d1, 77e7, 79g3, and 137g2 were shown to contain primers 4334 and 4335 and therefore encompassed the 3′ end of the gene.

RESULTS

Localization of the Mouse VIPR2 Gene

FISH analysis on metaphase preparations from mouse ES cells using a 4-kb *Bam*HI restriction fragment of the mouse VIPR2 gene as a probe showed that the gene is located on chromosome 12 band F2. The identity of the chromosome to which the probe hybridized was confirmed by carrying out a colocalization experiment using the VIPR2 probe and a cosmid containing the *Spi-2* gene (Fig. 1b). The *Spi-2* gene has been mapped to mouse chromosome 12 by interspecific backcross mapping (Hill *et al.,* 1985) and by FISH.

Localization of the Human VIPR2 Gene

FISH analysis using the human VIPR2 gene P1 clone as a probe showed that the human VIPR2 gene is located on 7q36.3 (Fig. 1a). No reproducible hybridization to any other chromosomal regions was observed. The human VIPR2 cDNA was also mapped by FISH to 7q36.3 (data not shown).

Our location of the human VIPR2 gene, close to the locus for HPE3, prompted us to attempt to refine our

TABLE 1

Cell Lines Used, Cytogenetic Abnormalities, and FISH Results with the VIPR2 Probe

Cell line	Phenotype[a]	Reported cytogenetic abnormalities	Comments and references	FISH results with VIPR2
GM03240	MC, MDF	46,XY,del(7)(q34 > qter) deletion of 7q34–qter	Gurrieri *et al.*, 1993	One copy of VIPR2 on normal 7q
GM00657	MC, MR, CP, small chin	46,XY,t(7;18) (7pter > 7q34::18q12.2 > 18qter;18pter > 18q12.2::7q36 > 7qter) mat deletion of 7q34–q36	Punnett *et al.*, 1979 Gurrieri *et al.*, 1993	One copy of VIPR2 on normal 7q
GM07216	HPE	46,XX,−7,+der(7)t(3;7) (7pter > 7q36::3p21.3 > 3pter) deletion of 7q36; duplication of 3p21.3 > pter	Kurtzman *et al.*, 1987 Gurrieri *et al.*, 1993	One copy of VIPR2 on normal 7q
GM10064	HPE	46,XX,−7,+der(7)t(7;13) (7pter >7q34::13q12.3 > 13qter) pat deletion of 7q34–qter; duplication of 13q12.3 > qter	Gurrieri *et al.*, 1993	One copy of VIPR2 on normal 7q One copy of VIRP2 on normal 13p
GM10067	Normal	46,XY,t(7;13)(7pter > 7q34::13q12.3 > 13qter; 13pter > 13q12.2::7q34 > 7qter) balanced translocation between 7 and 13	Father of GM10064	One copy of VIPR2 on normal 7q One copy of VIPR2 on normal 13p One copy of VIPR2 on 13q der

[a] MC, microcephaly; MDF, mildly dysmorphic facial features; MR, mental retardation; CP, cleft palate.

localization of the gene by mapping its position and orientation relative to the 7q telomere. FISH experiments indicated that a region of overlap existed between our VIPR2 P1 clone and the 7q telomeric YAC HSC7E526 (Fig. 1c). We therefore used primers (4334 and 4335) that flank the last intron of the human gene to screen by PCR (see Materials and Methods) a subset of a collection of YAC clones that have previously been mapped to 7q36 (Kunz *et al.*, 1994). These primers amplified a product of the expected size (1.2 kb) from three YAC clones (HSC7E526, HSC7E1357, and E145A7) that were also known to contain D7S593 (H. Donis-Keller, unpublished data). The D7S593 locus (Hing *et al.*, 1993) was derived from the YAC HTY146, which has previously been shown to encompass the telomeric end of the chromosome (Riethman *et al.*, 1993). Therefore, it was assumed that HSC7E526, HSC7E1357, and E145A7 overlap with HTY146 to form a contig of clones extending inward from the telomere as is shown in Fig. 2. YAC clone HSC7E526 has also previously been shown to contain D7S68 (Kunz *et al.*, 1994).

Having established that the VIPR2 gene lay directly adjacent to the 7q telomere, we proceeded to refine further the localization of the VIPR2 gene within this region. A fragment of the human VIPR2 cDNA that encompasses all 13 known coding exons was hybridized against the genomic DNA of 8 cosmids that spanned the length of the HTY146 telomeric YAC clone (Riethman *et al.*, 1993). Cosmid clones 48 (D7S593) and 14 gave positive hybridization results, which according to the map of Riethman *et al.* (1993) indicated that a portion of the gene resided approximately 200 kb from the end of the chromosome. The same cDNA probe was used to identify 22 additional cosmids from a chromosome 7-specific cosmid library. The relationship of these clones with respect to each other and cosmids 48 and 14 was determined by comparing the extent of overlap of restriction enzyme and hybridization patterns. A set of 8 representative cosmids that encompassed approximately 100 kb and the entire known coding region of the VIPR2 gene is shown in Fig. 2. From this analysis the entire VIPR2 gene could be positioned within about 350 kb of the terminal end of chromosome 7q, and the orientation of the gene along the chromosome (Fig. 2) was 7cen–D7S68–(3′-VIPR2-5′)–D7S593–D7S591–D7S592–D7S590–D7S594–7qtel. The VIPR2 gene is currently the closest known gene to the telomere of 7q.

Localization of the VIPR2 Gene in HPE3 7q Deletion Cases

Previous studies (Gurrieri *et al.*, 1993) have led to the identification of cell lines containing chromosomal deletions on 7q, the breakpoints of which defined the minimal critical region for HPE3. Using FISH, we localized the VIPR2 gene in cell lines from four of the patients described by Gurrieri *et al.* (1993) and in one unaffected relative. Details of the cell lines used (obtained from NIGMS Human Genetic Mutant Cell Repository) are given in Table 1.

The breakpoints in GM03240 and GM00657 had been used to define the proximal and distal boundaries of the HPE3 minimal critical region (Gurrieri *et al.*, 1993). We found that only one copy of the VIPR2 gene was present (on the normal chromosome 7) in each of these cell lines (GM00657 shown in Fig. 1d; data for GM03240 not shown). These results placed the VIPR2 gene within the HPE3 minimal critical region. The deletion in GM07216 extends to the telomere from a point proximal to the breakpoint in GM03240. In this cell line there was only a single copy of the VIPR2 gene on the normal chromosome 7 (data not shown), consistent with the results obtained from GM03240 and GM00657.

In the cell line GM10064 two copies of the VIPR2 gene were present, one on the normal chromosome 7 and the other on the normal (paternal) 13p (Fig. 1f).

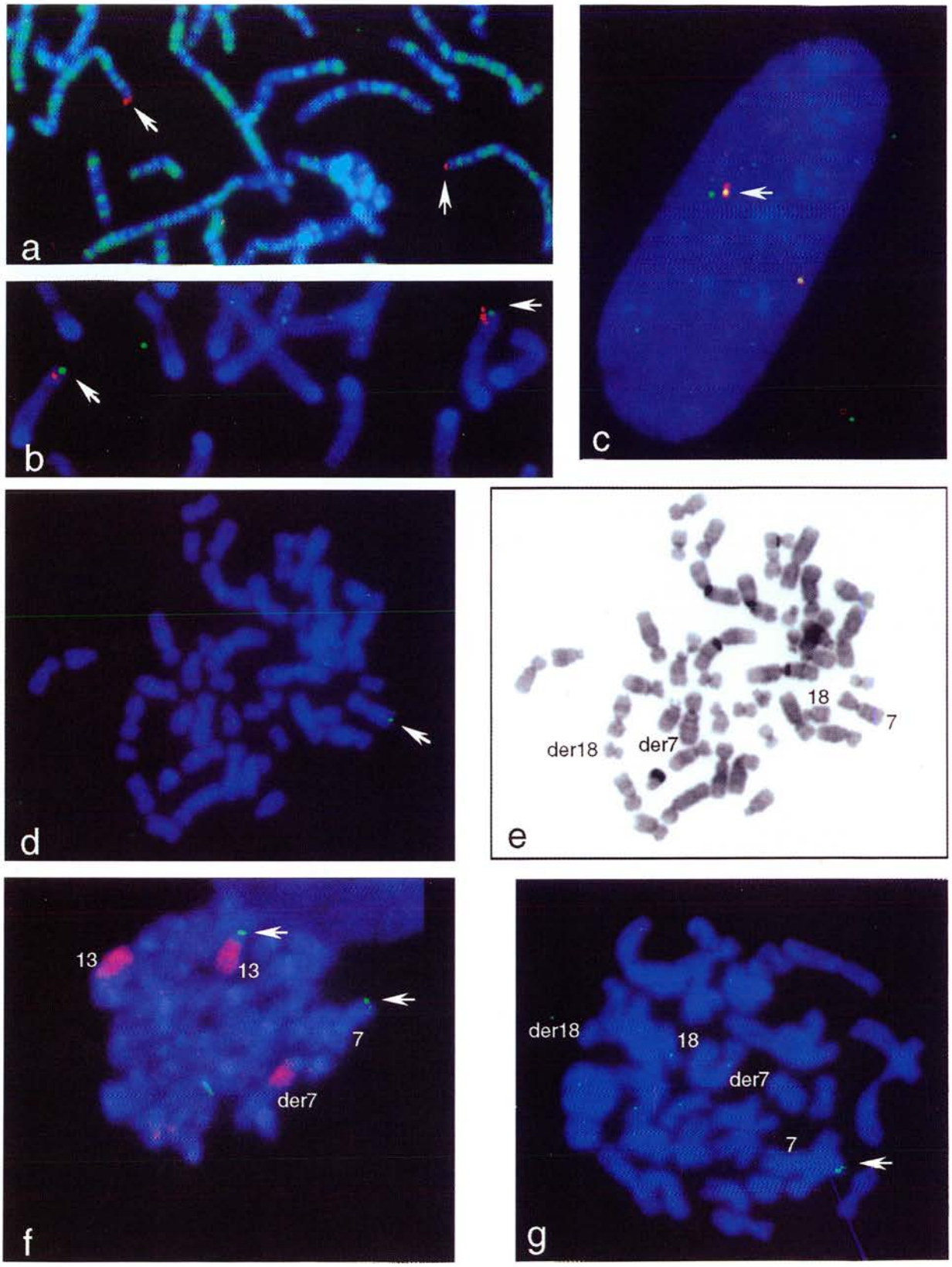


FIG. 1. Localization of the VIPR2 gene in human and mouse cell lines by FISH. (**a**) The VIPR2 P1 clone (red, arrowed) hybridized to human prometaphase chromosomes. A late replication G-banding pattern (green) is seen superimposed on DAPI (blue)-stained chromosomes. VIPR2 maps to band 7q36.3. (**b**) The VIPR2 mouse genomic clone (green, arrowed) and *Spi-2* cosmid (red) hybridized to mouse chromosomes. Signal from the *Spi-2* cosmid identifies chromosome 12; VIPR2 maps to band 12F2. (**c**) Hybridization of the VIPR2 P1 clone (orange, arrowed), YAC HSC7E526 (red), and YAC HSC7E207 (green) to human fibroblast nuclei. The orange signal overlaps with the red signal in all nuclei examined. (**d**) Hybridization of the VIPR2 P1 clone (green) to a metaphase from the cell line GM00657. The normal chromosomes

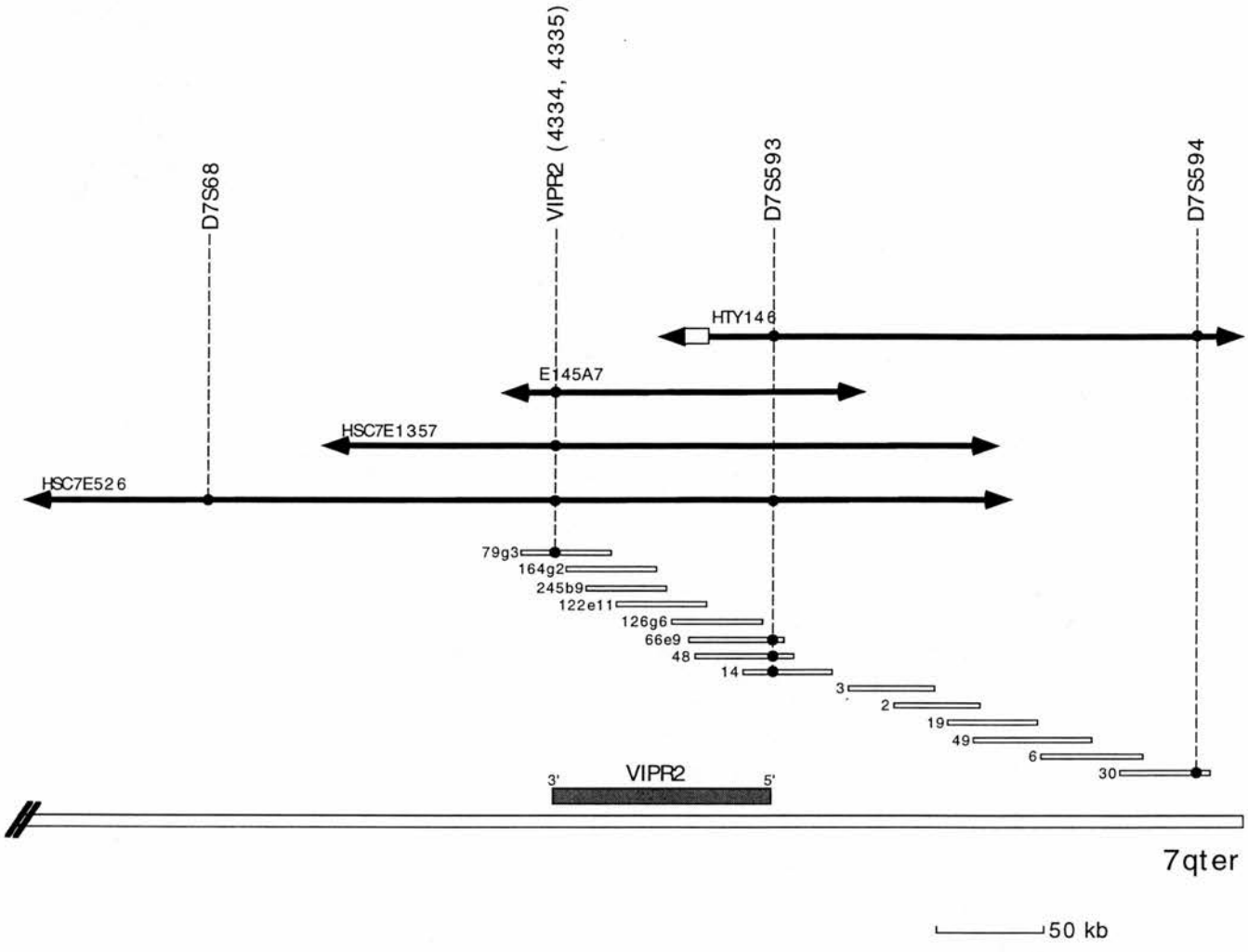


FIG. 2. Localization of the VIPR2 gene to the terminal end of chromosome 7q. The positions of the VIPR2 gene, D7S68, D7S593, and D7S594 among four YAC clones and 14 cosmids that encompass the terminal portion of the long arm of human chromosome 7 are shown. The 5′ end of the gene was located approximately 200 kb from the end of the chromosome, and the gene extended centromeric about 100 kb. VIPR2 (4334, 4335) are oligonucleotide primers specific for the last intron of the 3′ end of the gene.

Chromosome 7 and chromosome 13 paints were used to confirm this result. We also examined a cell line (GM10067) derived from the father of GM10064. This cell line, which carried a balanced translocation t(7;13), had three copies of the VIPR2 gene: one on the normal chromosome 7, one on the translocated chromosome 13 derivative, and one on the otherwise apparently normal 13p. The phenotype of GM10067 was normal. However, the relevance of the GM10064 cell line to chromosome 7-linked HPE is uncertain, as studies using chromosome 7 and chromosome 13 paints revealed that in addition to two copies of the VIPR2 gene, most of chromosome 13 was present in three copies in GM10064 (Fig. 1f), raising the possibility that HPE in this patient may be due to trisomy 13. Trisomy 13 is one of the most common cytogenetic abnormalities associated with HPE, and reports suggest that 70% of trisomy 13 cases have HPE (Taylor, 1968), compared to ~50% of del(7)(q36) cases (Overhauser *et al.*, 1995). As there was no way of establishing which of these chromosomal abnormalities was responsible for HPE in this patient, and it was statistically more likely to be due to trisomy 13, cell line GM10064 was excluded from our study of HPE3.

---

7 and 18 and the der7 and der18 are labeled in (e), a black and white enhanced DAPI image. Only one copy of VIPR2 (arrowed in d) on the normal 7, is present in this cell line. (**f**) Hybridization of the VIPR2 P1 clone (green) and a chromosome 13 paint (red; appears pink in figure) to a metaphase from the cell line GM10064. Signals from VIPR2 are found on the normal chromosome 7 and on 13p (arrowed). The 13 paint identifies both chromosome 13s and additional 13 material on the der7. (**g**) Hybridization of cosmid 29f5 (green) to a metaphase from the cell line GM00657. Hybridization with a chromosome 18 paint (image not shown) identifies the der7, 18, and der18 chromosomes. Only one copy of cosmid 29f5(D7S168) is present on the normal chromosome 7 in this cell line.

Reconciliation of the Cytogenetic and Physical Mapping Data: Redefinition of the Distal Extent of the HPE3 Minimal Critical Region

The HPE3 minimal critical region defined by Gurrieri *et al.* (1993) spans the region of 7q36 from marker D7S392 to D7S292. Our mapping experiments using the YAC and cosmid contigs from distal 7q indicated that the VIPR2 gene was situated close to the marker D7S593, which is at least 1.5 Mb telomeric to the proposed breakpoint in GM00657 that defines the distal boundary of the HPE3 minimal critical region (S. Scherer, unpublished data; Gurrieri *et al.,* 1993). This suggested either that the deletion in cell line GM00657 extended much further toward the telomere than previously thought or that a complex rearrangement had occurred in this cell line resulting in a second deletion distal to that originally characterized. To investigate these possibilities, cosmids containing D7S168 (cos29f5), VIPR2 (cos79g3 and cos126g6), and D7S594 (cos30) were used as FISH probes on this cell line (see Fig. 3 for location). A single copy of each of these cosmids, and a single copy of the 275-kb 7q telomeric YAC HTY3006 (Riethman *et al.,* 1993), were found on the normal chromosome 7 (the result obtained with the cosmid containing D7S168 is shown in Fig. 1g). The order of markers on distal 7q has been reported as cen–D7S550–D7S392 (proximal HPE3 critical region boundary)–D7S468–D7S104–D7S168–D7S292 (distal HPE3 critical region boundary)–D7S559–D7S68–D7S427–D7S594 (Tsui *et al.,* 1995). Although we cannot rule out the possibility that a segment of 7q36 between D7S168 and the VIPR2 gene remains, our results clearly show that significant areas of the region distal to D7S292, including the VIPR2 gene and the 7q telomere, are deleted in this cell line. These results redefine the distal extent of the HPE3 minimal critical region.

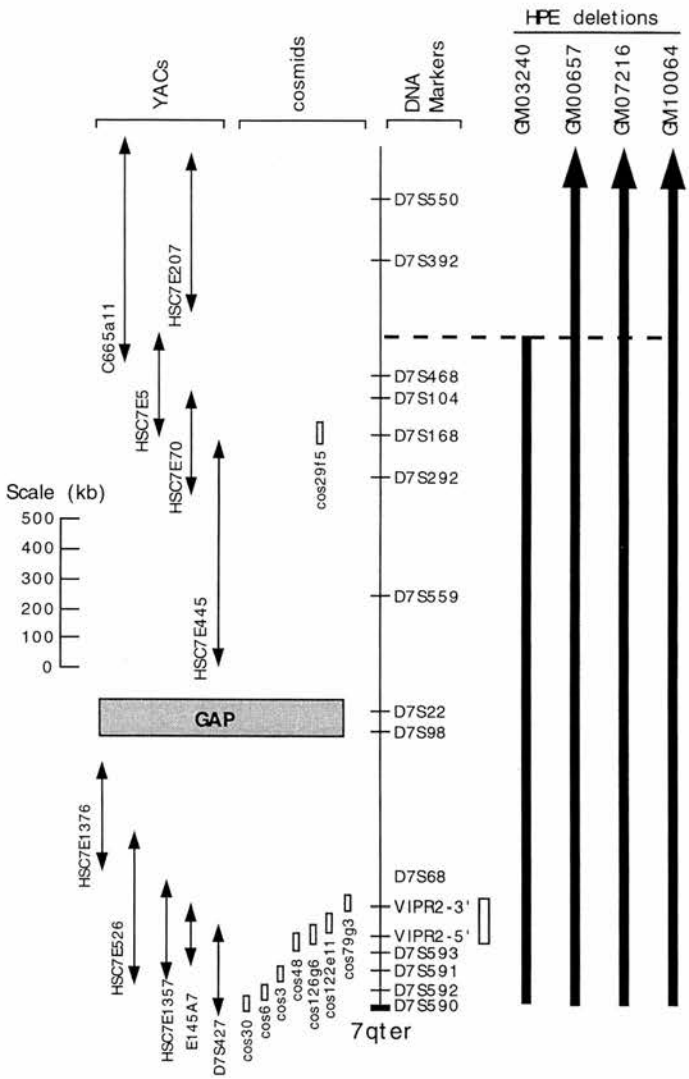


FIG. 3. Physical map of the terminal region of 7q36 that is deleted in HPE3. The order of the DNA markers at the extreme end of 7q36 as determined by YAC and cosmid contig analysis and linkage mapping is shown. The schematic diagram is drawn to the best possible scale based on all available data. A YAC and cosmid contig extending in from the telomere, which is encompassed by D7S427 (also called YAC HTY146 or Y3006, Riethman *et al.,* 1993), is shown (see Fig. 2 for a higher resolution map). VIPR2 is located in this region. Extending centromeric, a gap in the YAC map exists, but linkage mapping has positioned the D7S22 and D7S98 markers in this interval (Tsui *et al.,* 1995; and Helen Donis-Keller, unpublished data). The YAC contig then extends centromeric approximately 2 Mb (E. Belloni, unpublished data). The vertical bars on the right indicate the region in the four HPE3 patients that has been determined to be deleted in this study. The proximal deletion breakpoint in GM03240, which was precisely mapped between D7S392 and D7S468 (Gurrieri *et al.,* 1993), defines the proximal boundary of the critical region. The distal boundary of the HPE3 critical region is now defined as the 7q telomere. The translocation breakpoints of two unrelated HPE3 patients with apparently balanced rearrangements have been positioned between D7S392 and D7S104 (Belloni *et al.,* 1995).

DISCUSSION

We have mapped the human VIPR2 gene to 7q36.3 and demonstrated that it lies within the HPE3 minimal critical region.

HPE is a relatively common developmental abnormality in humans, occurring at a frequency of 1/16,000 live births and 1/250 conceptuses (Cohen, 1989). Phenotypic findings in patients with HPE vary widely, with the degree of facial malformation generally corresponding to the extent of the brain abnormalities (DeMyer *et al.,* 1964). In the most severe form of HPE, alobar HPE, a single brain ventricle is seen, which encompasses both the telencephalon and the diencephalon (DeMyer and Zeman, 1963). The olfactory bulb and tracts are absent, and the pituitary may be absent or hypoplastic. There is often abnormal development of the optic tracts and underdevelopment of the thalami. Microcephaly is very common. Facial abnormalities found in cases of alobar HPE include cyclopia, severe ocular hypotelorism, and cebocephaly (Siebert *et al.,* 1990).

Our interest in the VIPR2 gene as a potential contributor to the HPE phenotype stems from a number

of observations: (i) Gressens *et al.* (1993) have reported a dramatic effect of VIP on the growth of the CNS of mouse embryos in culture. Administration of a VIP receptor antagonist (Gozes *et al.,* 1991) during E9–E11 (the period during which subdivision of the prosencephalon begins) blocks these growth effects and induces severe microcephaly (Gressens *et al.,* 1994). (ii) In adult rats, VIPR2 mRNA is abundant in many of the structures that can be affected in HPE, including the pituitary, olfactory bulb, thalamus, and hypothalamus (Lutz *et al.,* 1993; Sheward *et al.,* 1995; Usdin *et al.,* 1994). Unfortunately no studies on the prenatal distribution of VIPR2 have been published to date. (iii) HPE has been found to occur at a higher incidence in children of diabetic mothers (Barr *et al.,* 1983). The VIPR2 gene is expressed at high levels in pancreatic β-cells (Inagaki *et al.,* 1994), where it is thought to be involved in the amplification of glucose-stimulated insulin secretion (Inagaki *et al.,* 1994; Straub and Sharp, 1994). (iv) Deletion of 18p has also been implicated in holoprosencephaly (HPE4), and a minimal critical region on distal 18p11.3 has recently been defined (Overhauser *et al.,* 1995). The gene encoding PACAP (an alternative endogenous ligand for VIPR2) maps to 18p11.32 within this minimal critical region (Chang *et al.,* 1993).

Our results from analysis of HPE3-associated 7q deletions have shown that the cytogenetic rearrangements in the cell lines GM00657 and GM10064, used by Gurrieri *et al.* (1993) to define the distal boundary of the HPE3 minimal critical region, are more complex than originally thought. The new HPE3 minimal critical region extends to the telomere of 7q and includes the VIPR2 gene locus. VIPR2 can therefore be considered as a candidate gene for HPE3.

The HPE phenotype is extremely variable, and this makes it difficult to distinguish patients who have mild forms of HPE3 from those who may have phenotypes associated with the deletion/rearrangement of other genes in the 7q36 region. Even among families with autosomal dominant HPE3, phenotypic expression within a family can vary from alobar HPE through to obligate carriers who have a normal phenotype (Muenke *et al.,* 1994).

Although the VIPR2 gene was shown to be monosomic in all of the HPE3-associated deletion patients examined, two unrelated patients with mild HPE symptoms and balanced translocations mapping between D7S392 and D7S104 (Belloni *et al.,* 1995) were both found to have two copies of the VIPR2 gene (data not shown). These findings suggest that the VIPR2 gene is unlikely to be the sole gene responsible for the HPE phenotype. However, one of the translocation patients is the mother of four children. Two of the children carry the same balanced translocation as the mother, and the other two have inherited an unbalanced translocation resulting in the deletion of one copy of the terminal region of 7q36 (Gurrieri *et al.,* 1993; Hatziioannou *et al.,* 1991). The children with the balanced translocation display a mild HPE phenotype, whereas those with the deletion have more severe HPE (Gurrieri *et al.,* 1993; Hatziioannou *et al.,* 1991). This seems to suggests that the gene(s) responsible for the more severe form of HPE3 lies distal to the translocation breakpoint. Our current knowledge of the expression pattern of the VIPR2 gene, the proposed involvement of at least one of the receptor's ligands in embryogenesis, and in particular the relatively high levels of VIPR2 mRNA found in olfactory bulb make the VIPR2 gene a strong contender for this role.

Our localization of the mouse VIPR2 gene to the telomeric region of mouse chromosome 12 was unexpected. There is no previously defined region of synteny between human 7q36 and mouse chromosome 12, although synteny has been demonstrated between other regions of human chromosome 7 and mouse chromosome 12 (Tsui *et al.,* 1995). HPE can be induced in mouse embryos by teratogens (Sulik *et al.,* 1995; Wei and Sulik, 1993); however, the apparent absence of HPE-like phenotypes in mouse genetic mutants mapping to chromosome 12 or to any other mouse chromosomes may reflect developmental differences between the two species.

Another midline defect, sacral agenesis, maps within the HPE critical region on 7q36 and, like HPE, has been found to occur at a higher incidence in children of diabetic mothers (Lynch *et al.,* 1995). It is not clear whether a different gene is responsible for this defect (Lynch *et al.,* 1995). Further work will be necessary to assess the involvement of the VIPR2 gene in HPE3. However, it seems likely that a delicate balance between the various components of the VIP/PACAP signaling pathway is required for normal brain development and that deletion of the VIPR2 gene may contribute to the phenotypic symptoms seen in many cases of HPE3.

ACKNOWLEDGMENTS

We thank Muriel Lee for assistance with the FISH localization of the mouse VIPR2 gene and Pat Malloy for culture of the human fibroblast cell lines. Helen Donis-Keller (University of Washington, St. Louis) allowed us access to unpublished mapping data. YAC HTY3006 was kindly provided by Anna Jauch (University of Heidelberg, Institute of Human Genetics, Heidelberg) and Harold Riethmann (Wistar Institute, Philadelphia). The *Spi-2* cosmid was provided by Robert Hill (MRC Human Genetics Unit, UK). We also thank Norma Brearley for careful preparation of the manuscript. This study was supported by the Medical Research Council (UK) and by a grant from the Canadian Genome Analysis and Technology Program (to L.-C.T.). E.B. is supported by the Telethon Italia.

REFERENCES

Barr, M., Hanson, J. W., Currey, K., Sharp, S., Toriello, H., Schmickel, R. D., and Wilson, G. N. (1983). Holoprosencephaly in infants of diabetic mothers. *J. Pediatr.* **102:** 565–568.

Belloni, E., Scherer, S. W., Siegel-Bartelt, J., Frumkin, A., Donis-Keller, H., Helms, C., Hing, A. V., Roessler, E., Barnoski, B., Mitchell, H., Muenke, M., and Tsui, L.-C. (1995). Characterization of the holoprosencephaly minimal critical region at 7q36. *Am. J. Hum. Genet.* **57:** A256.

Brownstein, B. H., Silverman, G. A., Little, R. A., Burke, D. T., Korsmeyer, S. J., Schlessinger, D., and Olson, M. V. (1989). Isolation of single-copy human genes from a library of yeast artificial chromosome clones. *Science* **244:** 1348–1351.

Chang, E., Welch, S., Luna, J., Giacalone, J., and Francke, U. (1993). Generation of a human chromosome 18-specific YAC clone collection and mapping of 55 unique YACs by FISH and fingerprinting. *Genomics* **17:** 393–402.

Cohen, M. M., Jr. (1989). Perspectives on holoprosencephaly. Part I. Epidemiology, genetics, and syndromology. *Teratology* **40:** 211–235.

Craig, J. M., and Bickmore, W. A. (1994). The distribution of CpG islands in mammalian chromosomes. *Nature Genet.* **7:** 376–382.

DeMyer, W., and Zeman, W. (1963). Alobar holoprosencephaly (arhinencephaly) with median cleft lip and palate. Clinical nosologic and electroencephalographic considerations. *Conf. Neurol.* **23:** 1–36.

DeMyer, W., Zeman, W., and Palmer, C. G. (1964). The face predicts the brain: Diagnostic significance of median facial anomalies for holoprosencephaly (arhinencephaly). *Pediatrics* **34:** 256–263.

Fantes, J. A., Bickmore, W. A., Fletcher, J. M., Ballesta, F., Hanson, I. M., and Van Heyningen, V. (1992). Submicroscopic deletions at the WAGR locus, revealed by nonradioactive in situ hybridization. *Am. J. Hum. Genet.* **51:** 1286–1294.

Francis, F., Zehetner, G., Hoglund, M., and Lehrach, H. (1994). Construction and preliminary analysis of the ICRF human P1 library. *Genetic Anal. Techniques Appl.* **11:** 148–157.

Gozes, I., and Brenneman, D. E. (1989). VIP: Molecular biology and neurobiological function. *Mol. Neurobiol.* **3:** 202–235.

Gozes, I., McCune, S. K., Jacobson, L., Warren, D., Moody, T. W., Fridkin, M., and Brenneman, D. E. (1991). An antagonist to vasoactive intestinal peptide affects cellular functions in the central nervous system. *J. Pharmacol. Exp. Ther.* **257:** 959–966.

Gressens, P., Hill, J. M., Gozes, I., Fridkin, M., and Brenneman, D. E. (1993). Growth factor function of vasoactive intestinal peptide in whole cultured mouse embryos. *Nature* **362:** 155–158.

Gressens, P., Hill, J. M., Paindaveine, B., Gozes, I., Fridkin, M., and Brenneman, D. E. (1994). Severe microcephaly induced by blockade of vasoactive intestinal peptide function in the primitive neuroepithelium of the mouse. *J. Clin. Invest.* **94:** 2020–2027.

Gurrieri, F., Trask, B. J., Vandenengh, G., Krauss, C. M., Schinzel, A., Pettenati, M. J., Schindler, D., Dietzband, J., Vergnaud, G., Scherer, S. W., Tsui, L. C., and Muenke, M. (1993). Physical mapping of the holoprosencephaly critical region on chromosome 7q36. *Nature Genet.* **3:** 247–251.

Hatziioannou, A. G., Krauss, C. M., Lewis, M. B., and Halazonetis, T. D. (1991). Familial holoprosencephaly associated with a translocation break at chromosomal position 7q36. *Am. J. Med. Genet.* **40:** 201–205.

Hill, R. E., Shaw, P. H., Barth, R. K., and Hastie, N. D. (1985). A genetic locus closely linked to a protease inhibitor gene complex controls the level of multiple RNA transcripts. *Mol. Cell. Biol.* **5:** 2114–2122.

Hing, A. V., Helms, C., and Donis-Keller, H. (1993). VNTR and microsatellite polymorphisms within the subtelomeric region of 7q. *Am. J. Hum. Genet.* **53:** 509–517.

Inagaki, N., Yoshida, H., Mizuta, M., Mizuno, N., Fujii, Y., Gonoi, T., Miyazaki, J.-I., and Seino, S. (1994). Cloning and functional characterization of a third pituitary adenylate cyclase-activating polypeptide receptor subtype expressed in insulin-secreting cells. *Proc. Natl. Acad. Sci. USA* **91:** 2679–2683.

Ishihara, T., Shigemoto, R., Mori, K., Takahashi, K., and Nagata, S. (1992). Functional expression and tissue distribution of a novel receptor for vasoactive intestinal polypeptide. *Neuron* **8:** 811–819.

Kofman-Alfaro, S., Speed, R. M., Boyle, S., and Chandley, A. C. (1994). Condensation behaviour of the human X chromosome in male germ cells and Sertoli cells examined by fluorescence in situ hybridization. *Chromosome Res.* **2:** 439–444.

Kunz, J., Scherer, S. W., Klawitz, I., Soder, S., Du, Y.-Z., Speich, N., Kalff-Suske, M., Heng, H. H. Q., Tsui, L.-C., and Grzeschik, K.-H. (1994). Regional localization of 725 human chromosome 7-specific yeast artificial chromosome clones. *Genomics* **22:** 439–448.

Kurtzman, D. N., van Dyke, D. L., Rich, C. A., and Weiss, L. (1987). Duplication 3p21 → 3pter and cyclopia. *Am. J. Med. Genet.* **27:** 33–37.

Lutz, E. M., Sheward, W. J., West, K. M., Morrow, J. A., Fink, G., and Harmar, A. J. (1993). The VIP_2 receptor: Molecular characterisation of a cDNA encoding a novel receptor for vasoactive intestinal peptide. *FEBS Lett.* **334:** 3–8.

Lynch, S. A., Bond, P. M., Copp, A. J., Kirwan, W. O., Nour, S., Balling, R., Mariman, E., Burn, J., and Strachan, T. (1995). A gene for autosomal dominant sacral agenesis maps to the holoprosencephaly region at 7q36. *Nature Genet.* **11:** 93–95.

Magistretti, P. J., and Schorderet, M. (1984). VIP and noradrenaline act synergistically to increase cyclic AMP in cerebral cortex. *Nature* **308:** 280–282.

Miyata, A., Arimura, A., Dahl, R. R., Minamino, N., Uehara, A., Jiang, L., Culler, M. D., and Coy, D. H. (1989). Isolation of a novel 38 residue-hypothalamic polypeptide which stimulates adenylate cyclase in pituitary cells. *Biochem. Biophys. Res. Commun.* **164:** 567–574.

Muenke, M. (1994). Holoprosencephaly as a genetic model for normal craniofacial development. *Dev. Biol.* **5:** 504–509.

Muenke, M., Gurrieri, F., Bay, C., Yi, D. H., Collins, A., Johnson, V. P., Hennekam, R. C. M., Schaefer, G. B., Weik, L., Lubinsky, M. S., Daack-Hirsch, S., Moore, C. A., Dobyns, W. B., Murray, J. C., and Price, R. A. (1994). Linkage of a human brain malformation, familial holoprosencephaly, to chromosome 7 and evidence for genetic heterogeneity. *Proc. Natl. Acad. Sci. USA* **91:** 8102–8106.

Ogi, K., Kimura, C., Onda, H., Arimura, A., and Fujino, M. (1990). Molecular cloning and characterization of cDNA for the precursor of rat pituitary adenylate cyclase activating polypeptide (PACAP). *Biochem. Biophys. Res. Commun.* **173:** 1271–1279.

Overhauser, J., Mitchell, H. F., Zackai, E. H., Tick, D. B., Rojas, K., and Muenke, M. (1995). Physical mapping of the holoprosencephaly critical region in 18p11.3. *Am. J. Hum. Genet.* **57:** 1080–1085.

Punnett, H. H., Kistenmacher, M. L., Miller, R. C., Greene, A. E., and Coriell, L. L. (1979). A (7;18) translocation. 46, XY. *Cytogenet. Cell Genet.* **24:** 126.

Quik, M., Iversen, L. L., and Bloom, S. R. (1978). Effect of vasoactive intestinal peptide (VIP) and other peptides on cAMP accumulation in rat brain. *Biochem. Pharmacol.* **27:** 2209–2213.

Riethman, H. C., Spais, C., Buckingham, J., Grady, D., and Moyzis, R. K. (1993). Physical analysis of the terminal 240 kb of DNA from human chromosome 7q. *Genomics* **17:** 25–32.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Scherer, S. W., Tompkins, B. J. F., and Tsui, L.-C. (1992). A human chromosome 7-specific genomic DNA library in yeast artificial chromosomes. *Mamm. Genome* **3:** 179–181.

Sheward, W. J., Lutz, E. M., and Harmar, A. J. (1995). The distribution of vasoactive intestinal $peptide_2$ receptor messenger RNA in the rat brain and pituitary gland as assessed by *in situ* hybridization. *Neuroscience* **67:** 409–418.

Shivers, B. D., Görcs, T. J., Gottschall, P. E., and Arimura, A. (1991). Two high affinity binding sites for pituitary adenylate cyclase activating polypeptide have different tissue distributions. *Endocrinology* **128:** 3055–3065.

Siebert, J. R., Cohen, M. M., Jr., Sulik, K. K., Shaw, C.-M., and Lemire, R. J. (1990). "Holoprosencephaly: An Overview and Atlas of Cases," Wiley-Liss, New York.

Spengler, D., Waeber, C., Pantaloni, C., Holsboer, F., Bockaert, J.,

Seeburg, P. H., and Journot, L. (1993). Differential signal transduction by 5 splice variants of the PACAP receptor. *Nature* **365:** 170–175.

Sreedharan, S. P., Huang, J.-X., Cheung, M.-C., and Goetzl, E. J. (1995). Structure, expression and chromosomal localization of the type I human vasoactive intestinal peptide receptor gene. *Proc. Natl. Acad. Sci. USA* **92:** 2939–2943.

Sreedharan, S. P., Patel, D. R., Huang, J. X., and Goetzl, E. J. (1993). Cloning and functional expression of a human neuroendocrine vasoactive intestinal peptide receptor. *Biochem. Biophys. Res. Commun.* **193:** 546–553.

Stoffel, M., Espinosa, R., Trabb, J. B., Lebeau, M. M., and Bell, G. I. (1994). Human type I pituitary adenylate cyclase activating polypeptide receptor (ADCYAP1R): Localization to chromosome band 7p14 and integration into the cytogenetic, physical, and genetic map of chromosome 7. *Genomics* **23:** 697–699.

Straub, S. G., and Sharp, G. W. G. (1994). Evidence for a novel mechanism of action of VIP and PACAP to stimulate insulin release. *Diabetologia* **37:** A22.

Sulik, K. K., Dehart, D. B., Rogers, J. M., and Chernoff, N. (1995). Teratogenicity of low-doses of all-trans-retinoic acid in presomite mouse embryos. *Teratology* **51:** 398–403.

Svoboda, M., Tastenoy, M., Van Rampelbergh, J., Goossens, J. F., De Neef, P., Waelbroeck, M., and Robberecht, P. (1994). Molecular cloning and functional characterization of a human VIP receptor from SUP-T1 lymphoblasts. *Biochem. Biophys. Res. Commun.* **205:** 1617–1624.

Taylor, A. I. (1968). Autosomal trisomy syndromes: A detailed study of 27 cases of Edwards' syndrome and 27 cases of Patau's syndrome. *J. Med. Genet.* **5:** 227–252.

Tsui, L.-C., Donnis-Keller, H., and Grzeschik, K.-H. (1995). Report of the 2nd International Workshop on Human Chromosome-7 Mapping 1994. *Cytogenet. Cell Genet.* **71:** 1–21.

Usdin, T. B., Bonner, T. I., and Mezey, E. (1994). Two receptors for vasoactive intestinal polypeptide with similar specificity and complementary distributions. *Endocrinology* **135:** 2662–2680.

Wei, X., and Sulik, K. K. (1993). Pathogenesis of craniofacial and body-wall malformations induced by ochratoxin-A in mice. *Am. J. Med. Genet.* **47:** 862–871.