SERINE PROTEASES EXPRESSED IN THE RODENT HIPPOCAMPUS: IDENTIFICATION OF TWO NOVEL GENES

Benjamin J. Davies

Ph.D.

The University of Edinburgh

1997
I declare that this work is my own, except where otherwise stated.
Acknowledgements

First of all, I would like to thank my supervisors, Rick Lathe and Richard Morris for their advice, encouragement and support throughout this work.

I would also like to thank Jan Ure for help and advice with cell culture and blastocyst injections, Louise Anderson and colleagues for technical assistance during this project and Andrew Smith for the targeting cassettes.

Thanks also go to members of the Neuro group past and present for help, advice and support in times of trauma. I am particularly indebted to Ben Pickard for introducing me to the pleasures of laboratory life, Muriel Steel for listening to my many rants and Melville Richardson who has taught me pretty much all I know and fed me excessive cake. I must apologise to Ken Rose for my chaotic bench, which I am well aware has affronted his orderly nature for many years.

Further thanks go to John Mason: his continuing support throughout this work has kept me on the straight and narrow. Honourable mentions must also go to Ailis Drysdale for violent secretarial duties, Jackie Wood for being so understanding and an assortment of CGR people who have provided welcome distractions.

Yet more thanks go to family and friends especially Margo, Craig, Kerstin, Andy and most importantly Chloe who have made Edinburgh such an enjoyable place to be.
Contents

Declaration
Acknowledgement
Table of Contents
Abbreviations
Abstract

Chapter I Introduction ........................................ 1

A - THE PHYSIOLOGY AND BIOCHEMISTRY OF MEMORY
1.1 Neuropsychological studies on amnesia
1.2 Lesion studies in rats
1.3 The structure of the hippocampal formation
1.4 Synaptic plasticity
1.5 Translation and transcription requirements of LTP
1.6 The role of hippocampal synaptic plasticity in learning and memory
1.7 Transgenic approaches to cognition
1.8 Hippocampus specific genes

B - SERINE PROTEOLYSIS IN BRAIN FUNCTION
1.9 Structural plasticity
1.10 Plasminogen activator system
1.11 Plasminogen activator expression in the developing brain
1.12 Plasminogen activator expression in the adult brain
1.13 Plasminogen activators in synaptic plasticity and memory
1.14 Plasminogen in the central nervous system
1.15 Multiple trypsin-like serine protease in the CNS
1.16 Thrombin and protein nexin-1
1.17 Extracellular matrix interactions
1.18 Serine proteases and synaptic transmission
1.19 Neuropeptides and growth factor processing
1.20 Summary

C - SERINE PROTEOLYSIS IN BRAIN DYSFUNCTION
1.21 Proteolysis and Alzheimer’s disease
1.22 The role of serine protease in neuronal cell death
1.23 Serine proteases in microglial function
1.24 Summary

D - THE STRUCTURE AND FUNCTION OF SERINE PROTEASES

E - AIMS OF THE PROJECT

Chapter II Identification of Hippocampal Serine Proteases

2.1 Multigene families
2.2 Homology probing and degenerate PCR
2.3 Considerations for degenerate primer design
2.4 Serine protease degenerate primers design
2.5 Amplification of serine proteases
2.6 Enriching for new serine proteases
2.7 Two novel serine protease sequences
2.8 Summary of serine proteases amplified from rat hippocampus
2.9 Serine proteases amplified from mouse hippocampus
2.10 An introduction to the serine proteases identified
2.11 Brain serine proteases not identified in this study
2.12 Discussion of degenerate PCR amplification
2.13 Improvements to the techniques used
2.14 Summary

Chapter III Expression Analysis of Serine Protease

3.1 The search for hippocampus specific genes
3.2 Northern analysis of serine protease expression
3.3 Relative abundance of serine proteases
3.4 In situ hybridization analysis of protease expression
3.5 Identification of a hippocampus specific transcript
3.6 Reverse-transcription PCR analysis of brain protease expression
3.7 RT-PCR analysis of thrombin expression
3.8 RT-PCR analysis of u-PA expression
3.9 RT-PCR analysis of plasminogen expression
3.10 Hepatocyte Growth Factor expression in brain
3.11 Summary

Chapter IV Molecular Characterisation ...................... 96
of BSP-2

4.1 Screening of a rat hippocampus cDNA library
4.2 The sequencing of BSP-2 cDNA
4.3 The BSP-2 open reading frame is disrupted
4.4 BSP-2 sequence shows homology with serine proteases
4.5 BSP-2 may represent a trypsin-like protease
4.6 Discussion
4.7 Summary

Chapter V Molecular Characterisation ...................... 105
of BSP-1

5.1 Screening of a rat hippocampus cDNA library
5.2 A BSP-1 cDNA constructed from 5' and 3' RACE products
5.3 The sequencing of BSP-1 cDNA
5.4 BSP-1 shows homology with trypsin-like proteases
5.5 BSP-1 is the rat orthologue of the mouse protease, neuropsin
5.6 Identification of BSP-1 genomic structure using long-range PCR
5.7 Discussion
5.8 Summary

Chapter VI Targeted disruption of ......................... 122
the BSP-1 locus

6.1 The rationale behind the targeting experiment
6.2 Transgenic experimentation with BSP-1
6.3 Homologous recombination in ES cells
6.4 BSP-1 targeting strategy
6.5 Construction of the BSP-1 targeting vector
6.6 Targeting of the BSP-1 locus
6.7 Derivation of BSP-1 mutant mice
6.8 Analysis of reporter expression
6.9 Discussion
6.10 Summary

Chapter VII Materials and Methods .................................. 146

7.1 Materials
7.2 Standard stock solutions used for molecular biology
7.3 Standard solutions for bacterial culture
7.4 Standard solutions for embryonic stem cell culture
7.5 General cloning techniques
7.6 Preparation of competent E. coli cells for transformation
7.7 Transformation of competent E. coli cells
7.8 Small scale preparation of plasmid DNA - “plasmid mini-preps”
7.9 Large scale preparations of plasmid DNA - “plasmid maxi-preps”
7.10 Preparation of genomic DNA
7.11 Total RNA extraction from mammalian tissue
7.12 Poly A+ mRNA purification
7.13 cDNA synthesis
7.14 Polymerase Chain Reaction
7.15 Long Range Polymerase Chain Reaction
7.16 Rapid amplification of cDNA ends (RACE) PCR
7.17 Cloning of PCR products
7.18 Dideoxy-termination sequencing of double-stranded DNA
7.19 Southern blot
7.20 Northern blot
7.21 Colony lifts
7.22 Radiolabelling of DNA
7.23 Hybridization conditions
7.24 Screening of bacteriophage lambda libraries
7.25 In vivo excision of λZAPII
7.26 Riboprobe synthesis
7.27 In situ hybridization
7.28 Reverse transcription PCR analysis
7.29 Culturing of embryonic stem cells
7.30 Transfection of embryonic stem cells
7.31 DNA extraction from embryonic stem cell colonies
7.32 Freezing of embryonic stem cells
7.33 Thawing of embryonic stem cells
7.34 Production of chimeras
7.35 Genotyping of mice
7.36 Brain dissections
7.37 X-gal whole mount staining of embryos

Chapter VIII Discussion ........................................... 176

8.1 Serine proteases within the brain
8.2 Serine proteases identified within the rodent hippocampus
8.3 Alternative targets for the plasminogen activators in brain
8.4 Two novel trypsin-like serine proteases
8.5 BSP-1 deficient mice
8.6 Directing gene expression to the hippocampus
8.7 Concluding remarks

References ................................................................. 188

Appendix A ................................................................. 214
### Abbreviations

Standard abbreviations for nucleotides and both the single letter and three-letter abbreviations for amino acids were used throughout the text.

Standard abbreviations and symbols recommended by the IUPAC-IUB Commission on Biochemical Nomenclature have generally been used.

Non-standard abbreviations are listed below.

<table>
<thead>
<tr>
<th>Abbreviation (single letter)</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT</td>
<td>α-antichymotrypsin</td>
</tr>
<tr>
<td>AMPA</td>
<td>α-aminooxydsalosan-4-propionate</td>
</tr>
<tr>
<td>APP</td>
<td>Amyloid precursor protein</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair(s)</td>
</tr>
<tr>
<td>BSEP</td>
<td>Brain serine protease</td>
</tr>
<tr>
<td>C2</td>
<td>Complement 2</td>
</tr>
<tr>
<td>CaMKII</td>
<td>Calcium/calmodulin-dependent protein kinase II</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>CRE</td>
<td>cAMP response element</td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP response element-binding protein</td>
</tr>
<tr>
<td>CTRL</td>
<td>Chymotrypsin-like protein</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNPase</td>
<td>Deoxyribonuclease</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular Matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>Ethylene glycol bis (β-aminoethyl ether)-N,N,N',N'-tetraacetic acid</td>
</tr>
<tr>
<td>ES</td>
<td>Embryonic stem</td>
</tr>
<tr>
<td>GABA</td>
<td>Gamma-aminobutyric acid</td>
</tr>
<tr>
<td>HGF</td>
<td>Hepatocyte Growth Factor</td>
</tr>
<tr>
<td>HPRT</td>
<td>Hypoxanthine phosphoribosyltransferase</td>
</tr>
<tr>
<td>HSPG</td>
<td>Heparin sulphate proteoglycan</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>HSV*&lt;i&gt;tk&lt;/i&gt;</td>
<td>Herpes Simplex Virus thymidine kinase</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropylthio β-D-galactoside</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal ribosome entry site</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase(s)</td>
</tr>
<tr>
<td>KPI</td>
<td>Kuntiz protease inhibitor</td>
</tr>
<tr>
<td>LIF</td>
<td>Leukeamia inhibitory factor</td>
</tr>
<tr>
<td>LTD</td>
<td>Long-term depression</td>
</tr>
<tr>
<td>LTP</td>
<td>Long-term potentiation</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>NGF</td>
<td>Nerve growth factor</td>
</tr>
<tr>
<td>NMDA</td>
<td>N-methyl-D-aspartate</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>PA</td>
<td>Plasminogen activator</td>
</tr>
<tr>
<td>PAI</td>
<td>Plasminogen activator inhibitor</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PKA</td>
<td>Protein kinase A</td>
</tr>
<tr>
<td>PKC</td>
<td>Protein kinase C</td>
</tr>
<tr>
<td>PN-1</td>
<td>Protein nexin 1</td>
</tr>
<tr>
<td>Poly A+</td>
<td>Polyadenylate</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid amplification of cDNA ends</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>RB</td>
<td>Rest of brain (brain with hippocampus removed)</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcription</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>t-PA</td>
<td>Tissue plasminogen activator</td>
</tr>
<tr>
<td>u-PA</td>
<td>Urokinase plasminogen activator</td>
</tr>
<tr>
<td>X-gal</td>
<td>5-bromo-4-chloro-3-indolyl β-D-galactoside</td>
</tr>
</tbody>
</table>
Abstract

Serine proteases in the central nervous system modulate developmental and synaptic plasticity and are implicated in the pathophysiology of Alzheimer's disease. This thesis aims to characterise the spectrum of serine proteases expressed in the brain. Degenerate primers were designed from regions conserved among the major chymotrypsin clan of serine proteases; these were used for the polymerase chain reaction amplification of family members represented in cDNA from adult rat hippocampus. 10 different members of the family were uncovered. The most abundant products corresponded to tissue plasminogen activator (t-PA) and RNK-Met-1, a lymphocyte protease not previously reported in brain. Evidence is provided to suggest that the major t-PA substrate, plasminogen, is absent from brain, arguing that the target for t-PA in brain is unlikely to be plasminogen. Other enzymes represented include elastase IV, proteinase 3, complement C2, Hageman factor, chymotrypsin B, chymotrypsin-like protein and two novel family members, BSP-1 (brain serine protease -1) and BSP-2.

Full length sequences of BSP-1 and BSP-2 are reported and sequence motifs and homologies suggest they represent trypsin-like proteases. The expression patterns for each of the serine proteases amplified, as determined by Northern and in-situ hybridization analysis, are presented. BSP-2 is expressed in the hippocampus and the cerebral cortex whereas BSP-1 is confined in its expression to the CA fields of the hippocampus.

A gene restricted in expression to the hippocampus has considerable potential for transgenic experimentation into the role of this brain region. Directing expression of suitably modified components of the synaptic signalling pathway within this brain region would allow the relationship between hippocampal synaptic plasticity and learning and memory events to be assessed. To investigate the feasibility of such an approach, the murine BSP-1 locus has been targeted with an IRES-lacZ reporter cassette. Preliminary results indicate reporter expression is active from the murine BSP-1 locus.

Besides the reporter integration, a region of the BSP-1 gene containing the catalytic aspartic acid residue, vital for enzymatic activity, has been deleted. This knock-out of gene function will allow the role of this particular hippocampus serine protease to be assessed though analysis of the homozygous mutant phenotype.
Chapter I

Introduction

The anatomical and functional complexity of the brain allows exploration from within diverse scientific disciplines. Historically, the brain has been investigated in purely metaphysical terms with minimal reference to the underlying anatomy and physiology. An alternative approach has addressed brain function through analysis of neuronal physiology and connectivity without reference to psychological processes.

Müller, at the beginning of this century made an early attempt to explain psychological phenomena in physiological terms with his theories concerning memory storage. The initial introspective dissociation of memory into long-term and short-term stores was hypothesized to reflect two different underlying physiological processes. These ideas were extended by Hebb in his “dual-trace theory”. A short-term store was proposed as a tentative and disruptable process in which information is temporarily held in vulnerable reverberating circuits of neural activity. This activity in turn produces invulnerable structural changes which take time to develop and results in a stable long-term store.

Modern genetic technologies are finally making it possible to marry physiological and psychological investigative approaches, allowing cognitive processes to be explained in molecular terms. The use of transgenic animals allows the effects of perturbed neuronal physiology to be assessed at a cognitive level.

This thesis addresses processes that may underlie aspects of memory storage and proposes a new tool for the molecular-genetic dissection of cognition. This introduction begins with an overview of the physiology and biochemistry of memory and discusses transgenic methodologies and their shortcomings. The latter sections review the structural phenomena accompanying learning and memory events, focusing on the role of one particular class of proteolytic enzymes, the serine proteases in mediating morphological change. The role of these proteolytic enzymes in brain dysfunction is also discussed.
The existence of pathologies where memory is affected allows a dissection of this cognitive process in terms of the underlying anatomy. These studies reveal the hippocampus to be of central importance in learning and memory events. Furthermore, the discovery that the hippocampus displays robust synaptic plasticity indicates that neurons within this brain region have the capacity to coordinate information necessary for the formation of long-lasting memory traces.

1.1 Neuropsychological studies on amnesia

In the late 1950s, Scoville and Milner reported a patient, H.M., who had undergone surgical removal of the medial temporal lobe bilaterally to relieve severe epilepsy. Post-operatively, he suffered severe anterograde amnesia, i.e. selective amnesia for recently acquired information with intelligence and memory of events preceding the surgical procedure left substantially intact. Furthermore the deficit was specific for declarative memory, memory of events and facts, but procedural memory, memory of skills and habit, was unaffected (Milner, 1972). Due to the severity of the lesion, removing much of the amygdala, hippocampus and adjacent cortical regions, the exact lesion responsible for this deficit was hard to determine. A second well-studied patient, R.B., suffered a similar form of amnesia as a result of an ischemic event which occurred during open-heart surgery. Subsequent post-mortem analysis revealed a selective lesion of the CA1 field of the hippocampus bilaterally, extending the full rostral-caudal extent of the formation (Zola-Morgan et al., 1986). Various considerations suggest the underlying cause of amnesia in the above patients was hippocampal damage. Other similar patients have been assessed with high-resolution magnetic resonance imaging to reveal the underlying lesion and again hippocampal abnormalities appear to correlate with anterograde amnesia (Press et al., 1989).

Animal models of human amnesia have been investigated to confirm the relative importance of temporal lobe structures in learning and memory. Animals with large medial temporal lobe lesions have severe memory impairments when tested with a delayed non-matching to sample paradigm, a test of declarative memory (Mishkin, 1978). Studies attempting to determine the exact nature of the lesion responsible for such a deficit have been confused by the inaccuracy of surgical procedures. Initial studies implicated both the hippocampus and the amygdala (Mishkin, 1978) but subsequent studies revealed that the amygdala lesion induced deficit was due to
damage to adjacent cortical regions and not to the amygdala itself (Zola-Morgan et al., 1989a). Further lesion studies revealed these important adjacent cortical areas to be the perirhinal and parahippocampal cortices (Zola Morgan et al., 1989b). These studies agree a role for the hippocampus and adjacent cortices in aspects of memory formation, however the relative contribution of each structure remains unresolved.

1.2 Lesion studies in rats

Electrophysiological investigations of the rat hippocampus revealed the presence of pyramidal cells in the CA fields which respond solely or maximally when the rat is situated in a particular spatial location. Once established the firing patterns of such place cells are particularly stable and are only altered by radical changes in environment (O'Keefe and Dostrovsky, 1971). This discovery led to a general hypothesis that the hippocampus is able to represent a non-topographic spatial map of a particular environment and thus acts as a distributed memory system. Lesions to the hippocampal formation would thus be expected to cause an impairment in the ability to perform tasks that require a spatial representation of an environment.

Rats with lesions to the entorhinal cortex, removing the majority of hippocampal inputs and rats with lesions to the main output pathways of the hippocampus were tested in a radial arm maze. These groups of animals had severe and lasting learning impairments in comparison to control groups (Olton et al., 1978). Another study tested animals with hippocampal lesions in a water maze (Morris et al., 1982). This test of spatial memory involves a rat swimming in opaque water in search of a hidden underwater platform. In repeated trials, the time taken to locate the platform decreases. The only sensory cues are provided by the spatial arrangement of objects around the testing laboratory, so the ability to perform this task is dependent on intact spatial representation. Animals with hippocampal lesions have distinct learning deficits compared to sham operated animals.

Conventional lesions are performed by aspiration, electrolysis or thermocoagulation and can cause damage to extra-hippocampal structures and projections. Such lesions also damage the local vasculature, resulting in unpredictable ischemic events. Furthermore, the hippocampus is particularly susceptible to seizure events and so manipulating the formation in this manner could result in widespread effects on other brain regions. A new generation of lesion experiments have been performed with the fibre sparing ibotenic acid. This agent causes damage to neuronal cells at the site of
delivery and leaves nearby fibre tracts intact. Rats with ibotenate lesions of the hippocampus have been tested in the water maze paradigm described above (Morris et al., 1990). The results were qualitatively similar to the effects seen with conventional lesions studies, yet quantitatively different. The greater effects on behaviour that occur following conventional lesions can therefore be attributed to damage to extra-hippocampal structures and projections.

Ibotenate lesion studies reveal that certain other types of memory remain intact in hippocampus lesioned rats. Lesioned animals perform as controls in a rodent adapted delayed non-matching to sample task, although performance of lesioned animals declined more quickly when the delay was increased (reviewed in Jarrard et al., 1993). The ability to perform this task however is at odds with the primate models of amnesia and suggests that rats with hippocampal damage can learn non-spatial/object recognition tasks.

Lesion studies thus demonstrate the role of the hippocampal formation in declarative memory processes is evolutionarily conserved as similar deficits are seen in rodents and primates. The results from lesioned rats suggest that the learning deficit seen is specific to tasks relying on spatial representations, whereas the deficit in primates appears not to have this spatial bias. This preferential hippocampal processing of space in rodents perhaps reflects an evolutionary specialisation reflecting the importance of spatial representation in the foraging behaviour characteristic of rodents. More recent studies, however, indicate that the hippocampus may mediate complex non-spatial memory such as relational representations (reviewed in Eichenbaum, 1996), suggesting a more general role for the hippocampal formation in non-modality specific learning in line with the proposed role in primates.

In lesion studies, memories of preoperative events are spared suggesting the hippocampus not to be the store of information in the brain. Rather, it appears that this brain region plays a role in aspects of memory consolidation through the formation of contextual associations. Lesion studies suggest that memory traces are stored in a distributed memory system within the neocortex, exemplified by Lashley's theories of mass action and equipotentiality, where a learning deficit is proportional to the size of the lesion no matter where in the cortex the lesion is placed.
1.3 The structure of the hippocampal formation

Neuropsychological studies thus highlight the hippocampal formation as essential for aspects of memory consolidation. Theories contending that the formation collates information and forms contextual associations are supported by the anatomy of this brain region as the hippocampus appears to lie within a circuit of information flow, receiving diverse cortical inputs.

The hippocampal formation comprises a pair of elongated structures with its long axis bending in a C shape from the septal nuclei rostro-dorsally to the temporal lobe caudo-ventrally. The formation consists of four main regions, entorhinal cortex, dentate gyrus, hippocampus proper (which is subdivided into three regions, CA3, CA2 and CA1) and the subicular complex (which includes the subiculum, the presubiculum and the parasubiculum) (Figure 1.1).

![Figure 1.1 - The rat hippocampus](image)

Micrograph of a horizontal rat brain section showing the dentate gyrus (DG), cornu Ammonis field 1-3 (CA1-3), subiculum (S), presubiculum (PrS), parasubiculum (PaS) and entorhinal cortex (EC)

Early research into the connectivity of the formation suggested the principal excitatory pathways follow the above order and are organised in a lamellar fashion (Anderson et al., 1971). More recent studies have revealed that most of the major intrinsic
connections are much more divergent, with individual neurons projecting along an extent of the septal-temporal axis (Amaral and Witter, 1989). Other pathways linking non-adjacent regions of the formation have also been characterised, revealing that the stereotyped tri-synaptic circuit is a considerable oversimplification of hippocampal connectivity (Amaral, 1993).

The entorhinal cortex receives information from neocortical association areas and is the origin of a strong projection into the dentate gyrus and hippocampus, the perforant path. Layer II neurons of the entorhinal cortex project to the molecular layer of the dentate gyrus and directly to the CA3 and CA2 fields. Projections from layer III reach the CA1 field and the subiculum. There are also intrinsic projections within the entorhinal cortex linking deep layers with more superficial layers.

The dentate gyrus granule cells give rise to the mossy fibres that collateralise in the dentate polymorphic layer before entering the CA3 field. Cells of the polymorphic layer give rise to association connections which are extensively divergent along the septal-temporal axis within the dentate gyrus. CA3 pyramidal cells give rise to the Schaffer collaterals which diverge along the long axis and terminate in the CA1 field. Proximal CA3 pyramidal neurons (i.e. adjacent to the dentate gyrus) project to more distal regions of the CA1 near the subicular border whereas distal CA3 neurons project to more proximal regions of the CA1 field. CA3 pyramidal neurons also give rise to extensive association projections within area CA3 itself.

CA1 pyramidal neurons project in a columnar fashion to the subiculum and weaker projections are directed to the deep layers of the entorhinal cortex. The subiculum is the major subcortical output field of the formation with prominent projections directed towards neocortical areas, the mammillary bodies of the hypothalamus, the thalamus, the nucleus accumbens, the amygdala and the lateral septum. The subicular neurons also project to the presubiculum and parasubiculum which, in turn, direct strong projections to layers IV and V of the entorhinal cortex respectively.

Hippocampal connectivity thus reflects a flow of information from neocortical association areas, through the hippocampus and back to the neocortex. The hippocampus is thus well placed to coordinate activity from many areas of the brain, reflecting different sensory modalities. The connections made to the hypothalamus and the amygdala, regions implicated in emotion, autonomic and visceral function, suggest the formation has a role coordinating motivated behaviour. Ultimately the
limbic circuits return to the neocortex which is considered the store of processed information.

1.4 Synaptic plasticity

Hebb and Konorski, in the late 1940s, proposed a potential storage mechanism with a memory trace represented by changes in synaptic efficacy between cells. They proposed a coincidence detection rule, whereby a synapse linking two neurons is strengthened if the two neurons are conjointly active. The neural connectivity of the hippocampus is ideally suited for this type of coincidence detection as inputs from many regions of the brain descend upon the formation. Interest was fueled by the discovery that particular regimes of stimulation could lead to an enduring strengthening of connections between neurons of the hippocampus (Bliss and Lømo, 1973), a phenomenon described as Long-Term Potentiation (LTP).

The potentiation refers to the ability of an isolated stimulus to bring about a more rapid and sustained depolarisation of the postsynaptic cell. Since the initiation of an action potential requires a threshold of depolarisation to be exceeded, LTP causes an enhancement of synaptic transmission. Other regimes of stimulation can lead to a depression in synaptic transmission, the phenomenon of Long-Term Depression (LTD) (Mulkey and Malenka, 1992). Furthermore, a depressed pathway can be potentiated and conversely depression can be induced in a potentiated pathway, indicating the two phenomena are reversible and share related underlying mechanisms. It has become clear that these two phenomena are extremes of a spectrum of long lasting changes in the efficacy of synaptic transmission.

Synaptic plasticity has been studied most extensively in region CA1, where glutamatergic inputs from CA3 neurons synapse with dendrites of CA1 neurons. The following discussion is restricted to this synapse. Other hippocampal regions exhibit LTP and LTD, but the mechanisms are not necessarily the same (see later).

LTP and LTD are both dependent on the activation of the N-Methyl-D-Aspartate (NMDA) receptor, a post-synaptic glutamate receptor that gates Ca\(^{2+}\) ions. NMDA receptor activation requires postsynaptic depolarisation, which relieves a Mg\(^{2+}\) block on channel function. This dual requirement of ligand binding and postsynaptic depolarisation for the NMDA-receptor activation is the molecular basis of the
coincidence detection mechanism, involved in the induction of long-lasting changes in synaptic transmission.

The activation of NMDA receptors leads to influx of Ca\(^{2+}\) into the postsynaptic cell. Microfluorometric measurement has allowed Ca\(^{2+}\) within individual neurons to be visualised (Regehr and Tank, 1991). Following receptor activation, a widespread accumulation of Ca\(^{2+}\) was seen as well as a transient component lasting seconds, spatially localised to the dendritic area of activated afferents. The transient component was attenuated by the specific NMDA receptor antagonist, 2-amino-5-phosphonovanerate (APV) and may represent the initial activator for the downstream signaling pathways active in synaptic plasticity. The widespread accumulation was considered to reflect influx of Ca\(^{2+}\) from other sources, including influx through voltage sensitive calcium channels and release from intracellular stores, the later being essential for the induction of LTP (Harvey and Collingridge, 1992). Ca\(^{2+}\) release from intracellular stores is linked to the activation of metabotropic glutamate receptors and accordingly metabotropic receptor antagonists impair the induction of LTP (Bashir et al., 1993).

Ca\(^{2+}\) in the postsynaptic cell can activate a variety of postsynaptic signaling cascades through a variety of different enzymes. Inhibitor studies have demonstrated the importance of Ca\(^{2+}\) activated kinases and phosphatases in the postsynaptic neuron. The Ca\(^{2+}\) dependent protein kinase C (PKC) has been implicated as loading the postsynaptic cell with a specific inhibitor peptide attenuated LTP (Wang et al., 1992). Furthermore addition of activators of PKC and the catalytic subregion itself potentiated postsynaptic responses (Malenka et al., 1986; Hu et al., 1987). A transgenic strain of mice lacking the brain specific gamma isoform of PKC have impaired LTP (Abeliovich et al., 1993a). The Ca\(^{2+}/\)calmodulin dependent protein kinase, CaMKII has also been implicated through the use of specific inhibitors (Malinow, 1989) and the subcellular localisation of this protein in the postsynaptic density indicates its abundant presence in the synapse (Kennedy et al., 1983). A role for this kinase was also revealed by a transgenic study as mice lacking the \(\alpha\) subform of this enzyme appear to have reduced LTP (Silva et al., 1992a).

PKC and CaMKII can respond to changing Ca\(^{2+}\) concentrations in the postsynaptic cell but the NMDA receptor dependent Ca\(^{2+}\) increase detected in LTP is only transient. Both these enzymes have the capacity to modulate their activity by becoming Ca\(^{2+}\)
independent, mediating a long-lasting increase in kinase activity which outlasts the initial Ca\textsuperscript{2+} signal.

Both enzymes are able to phosphorylate glutamate receptor subunits, providing a direct way in which synaptic transmission could be modulated. The \( \alpha \)-amino-3-hydroxy-5-methylisoxisole-4-propionate (AMPA) receptor subunit, GluR1 can be phosphorylated at serine-627 by CaMKII which leads to increased conductance (McGlade-McCulloch, 1993). Indeed, the cytoplasmic loops of the amino acid receptor subunit super-family contain several potential serine-threonine kinase phosphorylation sites (Raymond et al., 1993). PKC and CaMKII has also been shown to phosphorylate various presynaptic proteins, such as synapsin I (Nayak et al., 1996) and GAP43 (Gianotti et al., 1992). Evidence therefore exists for the involvement of presynaptic and postsynaptic modifications following synaptic plasticity.

The cysteine protease calpain has been implicated in the biochemistry underlying LTP (Lynch and Baudry, 1984). Calpain is Ca\textsuperscript{2+} dependent so the NMDA-receptor mediated Ca\textsuperscript{2+} influx could regulate the activity of this intracellular protease. Indeed, activation of calpain was seen following the administration of glutamate analogues (Siman and Noszek, 1988). Accordingly, inhibitors of calpain were found to block LTP (del-Cerro et al., 1990; Denny et al., 1990). Many intracellular calpain substrates exist and this protease has been implicated in structural changes following LTP. With regard to the signaling cascades, calpain is able to cleave PKC, liberating the regulatory domain and generating a Ca\textsuperscript{2+} independent activity known as PKM (Suzuki et al., 1992). Thus calpain activated by the NMDA-receptor dependent Ca\textsuperscript{2+} transient can lead to long lasting effects on intracellular kinases.

While the cAMP dependent protein kinase A (PKA) is implicated in late stages of LTP (see section 1.6), inhibition of this kinase also reduces the magnitude of earlier LTP phases. Inclusion of a phosphatase inhibitor with the PKA inhibitor rescues this effect on LTP suggesting that PKA has a role in regulating phosphatase activity in postsynaptic spines (Blitzer et al., 1995). The postsynaptic protein phosphatase 1 (PP1) is able to dephosphorylate CaMKII reverting the enzyme back to its Ca\textsuperscript{2+} dependent state. PKA is able to activate the PP1 inhibitor, inhibitor 1 which then limits the activity of this postsynaptic phosphatase. Since LTP is associated with an increase in protein kinase activity, LTD was hypothesized to be linked to phosphatase activity in the postsynaptic cell. Accordingly, okadaic acid and calyculin A, both
specific inhibitors of protein phosphatase 1 and 2A, completely block LTD in the hippocampus (Mulkey et al., 1993).

Both forms of synaptic plasticity thus require Ca\(^{2+}\) influx through the activated NMDA receptor and this Ca\(^{2+}\) activates a signaling cascade which leads to predominant kinase activity in the case of potentiation and phosphatase activity in the case of depression. The same basic induction conditions thus lead to opposing effects on protein phosphorylation. Lisman, (1989), stressed the quantitative differences in Ca\(^{2+}\) influx required to elicit the two phenomena with a mathematical simulation that modeled the Ca\(^{2+}\) induced phosphorylation cascades. He stressed the importance of the phosphorylation state of CaMKII in determining the efficacy of synaptic transmission, correlating phosphorylation with the potentiated state. It is the phosphorylation of this enzyme which allows its activity to become Ca\(^{2+}\) independent. Quantitative differences in the levels of Ca\(^{2+}\) influx required to elicit LTD and LTP were subsequently found (Cummings et al., 1996).

In agreement with Lisman’s model, the Ca\(^{2+}\) independent activity of CaMKII was seen to increase following LTP (Fukunaga et al., 1993) and retroviral transfection of hippocampal slices with a constitutively active CaMKII has been shown to result in potentiated transmission which occludes LTP (Pettit et al., 1994). Conflicting evidence comes from a transgenic study which expressed a constitutively active form of αCaMKII from the brain specific αCaMKII promoter (Mayford et al., 1995). An increased tendency for synaptic depression was seen and a correlation between the increase in Ca\(^{2+}\) independent CaMKII activity and the ability to induce depression was suggested. Supporting this theory, LTD can only be elicited in young animals and this ability negatively correlates with the amount of Ca\(^{2+}\) independent CaMKII activity which declines in age (Molloy and Kennedy, 1991). Despite these controversies, all studies agree that CaMKII is a key player in postsynaptic signaling cascades and that its regulation of constitutive activity by serine/threonine kinases and phosphatases is important in long lasting changes in synaptic transmission.

1.5 Translation and transcription requirements of LTP

PKA activity is widely held to be important in the late phase of LTP. Inhibitors of PKA block this long-lasting phase of LTP and PKA activators lead to a potentiation which lasts several hours and occludes LTP (Frey et al., 1993). NMDA-receptor mediated Ca\(^{2+}\) influx can increase levels of cAMP in the postsynaptic cell through
activation of the Ca\textsuperscript{2+} dependent adenylate cyclase I. Activation of adenylate cyclase coupled metabotropic glutamate or dopamine receptors also increase levels of cAMP. Such increases in cellular cAMP levels have been recorded following the induction of LTP (Chetkovich et al., 1991). A role for PKA in long-lasting LTP was further suggested as transgenic mice deficient in the C\textbeta1 catalytic subunit of this enzyme have attenuated LTP (Qi et al., 1996). Also overexpression of a dominant negative form of the regulatory subunit of PKA eliminates the late phase of LTP (Abel et al., 1997).

This late form of LTP is also blocked by protein synthesis inhibitors. For example, anisomycin caused LTP to decay to baseline within 2-3 hours with early LTP relatively unaffected (Krug et al., 1984). PAGE analysis of brain perfusates before and after induction of LTP showed distinct changes in protein composition implying such translational changes (Fazeli et al., 1993). Inhibition of transcription also leads to attenuated long-lasting LTP; potentiated responses decline to baseline within 2-4 hours (Nguyen et al., 1994). Inhibitors of transcription and translation also block the PKA activator induced potentiation, indicating a role for cAMP in mediating gene expression changes in the late phases of LTP (Frey et al., 1993; Nguyen et al., 1994).

High frequency stimulation of a synapse leads to a dramatic transcriptional upregulation of many immediate early genes including zif268, c-fos, c-jun and jun B, encoding transcription factors that are thought to contribute to the expression of late response genes (Cole et al., 1989). Pharmacological manipulations inducing seizure have proved particularly useful in the elucidation of stimulus-induced transcriptional changes in the hippocampus. Kainic acid induced seizures are correlated with a dramatic increase in neurotrophin transcripts (Timmusk et al., 1993) as well as secretory vesicle proteins and tyrosine phosphatases (Nedivi et al., 1993). A study on metrazol-induced seizure suggested roles for the immediate early genes, c-fos and zif-268 as well as transcripts such as tissue plasminogen activator (Qian et al., 1993). Changes in the levels of various protein kinases, such as PKC and CaMKII isoforms correlate with LTP (Thomas and Hunt, 1996) indicating that the signal transduction cascades may well be transcriptionally modulated in the course of potentiation.

One of the immediate early genes upregulated upon kainic acid induced seizure encodes CREM (Nedivi et al., 1993), a modulator of cAMP induced transcription through its interaction with the cAMP element binding protein (CREB). CREB is a transcription factor of the leucine zipper family which binds promoter regions as a dimer at the cAMP response element (CRE) activating transcription. The CRE is
found in the promoter regions of many genes, for example genes encoding neurotrophins and their receptors, neurotransmitter synthesis proteins, cytoskeletal proteins and cell adhesion molecules (Montminy et al., 1990). CREB activity is modulated by phosphorylation at serine-133 which lies within a PKA phosphorylation motif. Further PKA phosphorylation sites along with CaMKII and PKC sites are also found in the CREB activating domain indicating protein kinase mediated regulation of these transcriptional changes (Dash et al., 1991). Antibodies specific for phosphorylated CREB have detected increases in this species correlating with LTP. CREB has been shown to be absolutely required for the late phase of LTP as transgenic mice lacking brain enriched isoforms of CREB have impaired long-lasting LTP (Bourtchuladze et al., 1994). Furthermore, transgenic mice expressing β-galactosidase under the control of a cAMP response element, show induction of reporter expression by stimuli that generate long-lasting LTP within the hippocampus (Impey et al., 1996).

The requirement for de novo gene expression in consolidating long-term changes in synaptic efficacy is thus mediated, in part, though CREB mediated transcription. It appears, the increase in cAMP levels activates PKA which is then able to phosphorylate and activate CREB. Transcription from CRE containing promoters is then induced. This involvement of cAMP, PKA and CREB in mediating required transcriptional changes in synaptic plasticity appears to reflect a general strategy conserved across the species barrier. Studies into the biochemistry underlying long-term memory models in *Aplysia Californica* have revealed a similar dependence upon these signalling molecules and a requirement for transcription mediated by a CREB homologue (Kaang et al., 1993). Long term memory in *Drosophila* was also found to be dependent on CREB mediated transcription (Yin et al., 1994; 1995).

An involvement of gene transcription and protein synthesis in the maintenance of the late phases of LTP raises the question of how synapse specificity is achieved. Newly synthesised proteins are available to all synapses on a particular neuron, yet potentiation is only maintained at previously activated synapses. A synaptic tag has been envisaged which is present at potentiated synapses and interacts with molecules synthesised in the cell body (Sossin, 1996). Experimental evidence for such a synaptic tag has recently been reported (Frey and Morris, 1997).

Synaptic plasticity in the hippocampus, specifically LTP, can be partitioned into distinct phases. After the inducing stimulus, a short phase of potentiation is revealed
that is insensitive to serine/threonine kinase inhibitors. This is followed by a phase of kinase dependence so early maintenance of LTP is achieved through covalent modification of existing proteins. After this, a phase of LTP requiring translation and transcription occurs indicating the lasting change in synaptic transmission requires specially synthesized products. The different phases of LTP can also be distinguished by particular regimes of stimulation. Typically, the first two phases of LTP which last 1-3 hours can be elicited with a single train of 100 Hz but three or more trains of tetanic stimulation are necessary for the longer-lasting LTP which can last as long as 6-10 hours in hippocampal slice preparations (Huang and Kandel, 1994).

Different synapses in the classic tri-synaptic circuit have different mechanisms of synaptic plasticity. The mossy fibre/CA3 pathway elicits robust LTP (Zalutsky et al., 1990) and LTD (Kobayashi et al., 1996), but potentiation at this synapse appears to be NMDA independent as it can be induced in the presence of APV. The requirement for transcription and translation for long-lasting potentiation at this synapse however suggests that certain phases of LTP are common (Huang et al., 1994).

1.6 The role of hippocampal synaptic plasticity in learning and memory

NMDA receptor mediated synaptic plasticity is particularly attractive as a mechanism which underlies aspects of memory. As has been described, activation of the NMDA receptor requires both neurotransmitter release from the presynaptic cell and depolarisation of the postsynaptic cell. The conjunction of pre- and post-synaptic activation thus leads to a long-lasting change in synaptic efficacy - the Hebbian synapse. As such, long-lasting associations can be established within this brain region from diverse cortical inputs, a process thought to underlie the consolidation of memory traces. Associations once established are then passed on through the limbic circuits back to neocortical areas. Several different approaches have been used to formally establish a link between this tantalising phenomenon and memory/learning processes.

Many researchers have sought changes in synaptic physiology in vivo during or following certain behaviour (Green and Greenough, 1986) with conflicting success. Interest focused on a report of observable changes in synaptic physiology following brief exploration of a novel environment (Sharp et al., 1989). However, these changes were not long-lasting and were subsequently shown to be attributable to temperature variations in the brain associated with motor activity (Moser et al., 1993).
The fact that concrete changes are barely observable is perhaps of no surprise as any induced changes would presumably occur in only a select number of neurons.

Other researchers have attempted to saturate LTP and have found deficits in learning new tasks (McNaughton et al., 1986). Such studies have been questioned as it may be impossible to saturate all synapses and hippocampal physiology and indeed functions in other brain regions would be distinctly abnormal following the harsh regimes of stimulation necessary.

LTP induction usually involves tetanic stimulation but such high frequency activation is far from being physiologically relevant. Cholinergic neurons elicit an oscillatory activity in the hippocampus during exploratory and learning behaviour known as theta rhythm (O'Keefe, 1993). Artificial stimulation at an interval frequency corresponding to this endogenous rhythm induces potentiation (Larsen et al., 1986) and bidirectional changes in synaptic efficacy can be induced depending upon the phase of endogenous rhythm (Huerta and Lisman, 1995). These studies indicate that LTP and LTD like phenomena may be elicited in vivo by this theta rhythm during learning behaviour.

Pharmacological agents which inhibit synaptic plasticity have been shown to cause learning deficits. The NMDA antagonist, APV which blocks LTP and LTD induction impaired performance in the water maze, a task considered to be dependent upon the hippocampus and require spatial learning (Morris et al., 1986). Such studies appear to provide convincing evidence of a link between LTP and learning. However, NMDA receptors are widely distributed throughout the central nervous system and have many roles in neural function. NMDA antagonists have been shown to cause hyperactivity, ataxia, stereotypy and deficits in motor task induction and control (reviewed in Cain, 1997). Furthermore, normal coding and transmission of information of visual, somatosensory and pain stimuli require NMDA receptor function. The deficits described may therefore reflect subtle abnormalities in sensory and motor tasks rather than genuine learning deficits. Accordingly, recent studies have shown that intact hippocampal dependent learning can occur even when LTP is fully blocked (Bannerman et al., 1995; Saucier and Cain, 1995). These studies indicated that pretraining protects against the otherwise detrimental effects of NMDA receptor blockade. The reason for this protection, however, is still not understood.
Lesion studies, pharmacological and electrophysiological manipulations have thus suggested a role for hippocampal LTP in learning and memory processes. However, the lack of specificity of these techniques has confused such approaches. Transgenic technology allows mutations to be introduced into mice, to disrupt specific aspects of neuronal function in a precise, well characterised manner. The resulting mutated line of mice can be analysed biochemically, electrophysiologicaly and behaviourally in an attempt to establish causal links between molecular events and higher cognitive abilities.

A whole host of postsynaptic and presynaptic proteins have been mutated and many correlations between deficits in synaptic plasticity and learning have been found (reviewed in Grant et al., 1994; Chen and Tonegawa, 1997). For example, mice deficient in γPKC (Abeliovich et al., 1993a; 1993b) and αCaMKII (Silva et al., 1992a; 1992b) have deficits in both LTP and spatial learning.

Other evidence suggestive of a causal link between LTP and learning and memory relates to the distinct phases of potentiation. Long-term memory is dependent upon protein synthesis (reviewed in Squire and Davis, 1981) as is the long-lasting form of LTP (Krug et al., 1984). This correlation is extended by studies into transgenic mice with perturbed long-lasting LTP. Mice overexpressing the dominant negative form of the regulatory subunit of PKA have selective deficits in long-lasting LTP and are only deficient in long-term but not short-term contextual fear conditioning (Abel et al., 1995). Long-lasting LTP seems to correlate specifically with long-term memory.

Transgenic techniques would appear, therefore, to allow learning and memory processes to be addressed at the level of individual gene function. Furthermore correlated deficits in hippocampal synaptic plasticity and learning and memory provide evidence that processes such as LTP within the hippocampus underlie these higher cognitive abilities. These approaches are however somewhat flawed. Typically, the gene under study is expressed throughout the brain and so the learning deficit may result from altered function in other brain regions, rather than from a synaptic plasticity deficit in the hippocampus. Other brain regions implicated in learning and memory may also compensate for a hidden deficit induced by disrupted function in the hippocampus.
Previous gene targeting experiments have focused on particular isoforms of molecules involved in synaptic signalling pathways. Phenotypic analysis of such strains is often confused by the compensatory roles of other brain expressed isoforms. For example, in the investigations of αCaMKII deficient mice, some potentiation could be elicited from certain slices perhaps suggesting the brain expressed βCaMKII is compensating for this loss (Silva et al., 1992a). αCaMKII deficient mice were also found to be epileptic (Butler et al., 1995) thus bringing the basis of the learning deficit into question.

A recent study has highlighted another potential problem associated with transgenic approaches to cognition. An initial dissociation between LTP in the dentate gyrus and spatial learning was described in mice with a mutated thyl gene encoding, a neuronal glycoprotein (Nosten-Bertrand et al., 1996). Dentate gyrus LTP could however be rescued by application of GABA_A antagonists in vitro and LTP in vivo was effectively normal in some unanaesthetized animals (Errington et al., 1997). This study suggests results from in vitro preparations or in vivo studies on anaesthetized animals may not serve as accurate models of the situation in the awake animal.

The mutation of a specific gene can lead to a subtle developmental abnormality which may explain the learning deficit seen. A transgenic study into the role of the tyrosine kinase, fyn revealed a learning deficit which correlated with a reported deficit in hippocampal LTP (Grant et al., 1992). However, the mice exhibit an anatomical defect in the hippocampal formation implying that the kinase has important functions during development. The basis of the learning deficit is therefore difficult to establish.

Novel molecular genetic techniques are now a keen area of research in an attempt to avoid these pleiotropic effects of conventional knock-out technology. Region-specific gene promoters have been used in transgenic animals to drive expression of transgenes to specific regions of the brain. The brain specific αCaMKII promoter was used to direct wildtype fyn specifically within the forebrain of fyn knockout mice in order to rescue the phenotype (Kojima et al., 1997). The αCaMKII promoter is only active after birth thus allowing wild type fyn to complement the mutated gene product postnatally whilst leaving the developmental effects of this mutation. The rescue of the learning phenotype thus substantiates the link between this tyrosine kinase, LTP and learning and memory.
The αCaMKII promoter has also been used to create region specific knock-out mice. Cre recombinase from bacteriophage P1 is active in mammalian cells and can catalyse recombination between specific target sequences, known as loxP sites (Sauer and Henderson, 1988). The expression of cre recombinase from the αCaMKII promoter in transgenic mice yields active expression of the recombinase in forebrain regions. When crossed to a strain of mouse harbouring loxP sites flanking essential exons of the NMDA receptor subunit NR1, recombination occurred deleting this exon only in forebrain regions (Tsien et al., 1996a). Conclusions were drawn based upon a strain of mice that fortuitously expressed the cre transgene only within the CA1 region of the hippocampus. This CA1 specific NR1 knock-out had correlated deficits in CA1 LTP and spatial learning, in the absence of any deficits in dentate gyrus LTP (Tsien et al., 1996b).

A third study has used the αCaMKII promoter to drive a constitutively active αCaMKII transgene specifically within forebrain regions. This study added a temporal regulation to the expression by using the tetracycline inducible system (Mayford et al., 1996). The tetracycline activator, tTA, binds the tet operon, activating transcription only in the absence of tetracycline or derivatives (Furth et al., 1994). Two transgenic strains were crossed, one expressing the tTA gene from the region specific αCaMKII promoter and one expressing the constitutively active transgene ubiquitously under the control of the tet operon. Transgene expression is therefore under regional control and temporal control as withdrawing tetracycline leads to expression of the constitutively active transgene. These experiments effectively confirmed the previous finding that overexpression of this transgene causes an alteration in the balance between depression and potentiation and such synaptic plasticity effects are correlated with a specific learning deficit (Bach et al., 1995). The temporal nature of transgene activation also allowed distinct phases of learning to be dissected. The expression of the transgene once learning had been established was shown to interfere with the retrieval of information implicating hippocampal function in both storage and retrieval mechanisms.

The transgenic strains resulting from such studies have been used to characterise the effects of these mutations on place cells in the hippocampus. The absence of synaptic plasticity in the CA1 specific NR1 knock-out mice correlated with decreased spatial specificity of place fields in CA1 cells (McHugh et al., 1996). The perturbed synaptic plasticity in the mice with constitutively active αCaMKII, resulted in weaker positional firing patterns of place cells and the place fields once established were substantially
less stable over time (Rotenberg et al., 1996). These studies thus link the impaired representation of space with deficits in spatial learning and synaptic plasticity.

These elegant transgenic studies provide the best evidence to date that site-specific perturbations of synaptic plasticity can lead to learning deficits. However, transgenic technology is not without its flaws. Conventional transgenics involving the random integration of the transgene lead to slightly unpredictable transgene expression. In the above studies using the αCaMKII promoter, only certain lines fully reflected the endogenous expression pattern of this kinase. Other lines had aberrant expression patterns, higher in certain regions of the brain. This lack of consistency, due to the silencing or activating effects of sequences flanking the random transgene integration, is a major flaw of this general approach as such experiments are rarely repeatable. Also transgenes often integrate in multi-copy arrays, which can lead to a variegated pattern of expression confusing analysis further (Dobie et al., 1996).

With homologous integration, mutations in specific genes can be made in a fully defined manner. However, the genetic background of the resulting mutant mice can interfere with the results. At present gene targeting technology only exists reliably with the 129 strain of mice and this strain is associated with various congenital deficits that interfere with behavioural results (Lathe, 1996). Mutations are routinely crossed onto other genetic backgrounds to avoid these strain 129 associated deficits. Different strains of mice are associated with different learning capacities and different sensitivities to pharmacological agents. Resulting animals are hybrids between different strains and are therefore of variable genetic background. Potential modifiers of synaptic plasticity events and behaviour segregate at random between and within generations, confusing the phenotypic analysis (Gerlai et al., 1996). The use of inducible transcription regulation such as the tetracycline activator system, avoids these problems, as the effects of a transgene can be monitored before and during induction of the transgene within the same animal.

The extensive characterisation of synaptic plasticity within the hippocampus has led to a bias in the study of these phenomena. The regularity of hippocampal connections along with in vitro slice technology allow synaptic plasticity to be studied easily in the formation. Synaptic plasticity is inducible in a variety of other brain regions and there is evidence to suggest that the capacity for long-lasting increases in synaptic transmission may reflect a general method of information processing in the central nervous system. For example, LTP has been characterised in the visual cortex
The expression of the constitutively active αCaMKII transgene has been used to perturb synaptic plasticity in the amygdala. Expression of this transgene, within this brain region, results in a deficit in fear conditioning, a type of learning that is dependent on the amygdala (Mayford et al., 1996). This study highlights the importance of synaptic plasticity outside of the hippocampus and suggests a more general role in learning and memory.

Despite the shortcomings of a number of different experimental techniques, there is considerable evidence to suggest that synaptic plasticity is required for certain aspects of learning and memory. The correlated disruption in place fields indicates that synaptic plasticity events within the hippocampus may also be important in the establishment of a cognitive map reflecting a representation of space. It is of course still possible however that synaptic plasticity is merely a convenient experimental artefact which allows neuronal physiology to be assayed.

1.8 Hippocampus specific genes

The new approaches of directed gene expression allow transgenic manipulation within specific brain regions. Regarding the role of the hippocampus in learning and memory, a gene restricted in its expression to the formation would be of substantial utility. Expression control elements of such a gene could be used to yield hippocampus specific perturbations of neuronal function. The phenotype of resulting animals would permit correlations between cognitive impairment and neuronal function to be drawn with greater confidence.

Other genes have previously been reported as being "hippocampal specific" or "enriched". The mineralocorticoid receptor expression is confined to all subregions of the hippocampus, septum and some brain stem nuclei although a low level of expression is detectable in all other subregions (Yau and Seckl, 1995). Hippocalcin, a calcium binding protein of the recoverin family, was reported as being exclusively expressed in the hippocampus (Kobayashi et al., 1992). The expression, although high in the hippocampus is present at low levels in all other brain regions and quite modest levels are detectable in the cerebral cortex. Embryo brain kinase (Ebk), a member of the eph/elk tyrosine kinase family, was amplified from brain cDNA using degenerate PCR (Ellis et al., 1995). In situ hybridization revealed the primary site of expression to be the hippocampus and some diffuse signal was detectable in the
adjacent subiculum. The transcript was only detectable by Northern analysis when 2µg of hippocampus PolyA+ RNA was used, implying that Ebk is expressed at low levels. The muscarinic receptor, m5, is expressed at low levels in the CA1 pyramidal neurons of the hippocampus, in the habenular nuclei and also in the hypothalamic region (Vilaró et al., 1990). HZF-3, an orphan receptor is expressed at high levels in the hippocampus but also in the cerebellum (de Ortiz et al., 1996). These genes are thus either not sufficiently specific in their expression or are expressed at too low a level to be useful for directed gene expression in the hippocampus.

The characterisation of a gene expressed at moderate levels and exclusively within the hippocampal formation, is a requirement for the limitation of the pleiotropic effects of conventional knock-out technology. One aim of this thesis has been to investigate gene expression within the hippocampus with an aim to uncovering a hippocampus specific gene. The approach used has focused upon the serine protease gene family. As well as this technological justification, the characterisation of serine proteases expressed within the hippocampus is also of inherent interest. The following sections review the involvement of these proteases within the brain.

B - SERINE PROTEOLYSIS IN BRAIN FUNCTION

Structural changes within the brain are considered to underlie aspects of the durable long-term memory trace. This sections outlines the evidence that morphological events correlate with aspects of learning and memory and reviews the involvement of serine proteases in mediating much of this structural plasticity. Evidence for the involvement of serine proteases in synaptic transmission and the processing of bioactive molecules is reviewed.

1.9 Structural plasticity

Evidence exists for structural changes in the brain following learning. Behaviour within an enriched environment has been shown to cause increases in spine density on CA1 pyramidal cells and this increase also correlated with enhanced spatial learning (Moser et al., 1994). Within the cerebellum, an increase in the number of synapses per Purkinje neuron was seen following the learning of complex motor tasks (Black et al., 1990). Many other structural changes such as alterations in the extent and continuity of the postsynaptic specializations and the geometry of opposition between
pre- and postsynaptic compartments have been correlated with learning (reviewed in Bailey and Kandel, 1993).

Structural alterations are also observed following the induction of LTP. An alteration in the density of certain classes of synaptic contacts has been seen following LTP induction in the dentate gyrus (Desmond and Levy, 1986). Another study revealed an increase in the number of synapses with segmented or perforated post-synaptic densities following LTP induction (Geinismann et al., 1991). These perforations were considered to reflect synaptogenesis indicating an LTP-induced proliferation of functionally separate transmission zones. Thus, the long-lasting increase in synaptic transmission may, in part, be mediated by structural changes (discussed in Lisman and Harris, 1993; Edwards, 1995). Activation of CREB, the cAMP dependent transcription factor implicated in long-term plasticity, was shown to mediate morphological plasticity in primary hippocampal cultures (Murphy and Segal, 1997) suggesting the transcriptional changes induced by LTP contribute to structural reorganisation.

Structural reorganisation is common to many biological processes and similar underlying processes are often involved. Although the molecular events underlying structural change are poorly defined, it appears that cell adhesion and proteolysis represent two important mechanisms. The latter is the focus of this current review.

Calpain, the Ca\(^{2+}\) activated cysteine protease implicated in LTP, has a potential role in structural plasticity. Calpain mediated cleavage of fodrin/α-spectrin and microtubule associated proteins can lead to cytoskeletal reorganisation (Siman and Noszek, 1988; Lynch and Baudry, 1984). Neural cell adhesion molecule (N-CAM) was also found to be a substrate of calpain, possibly implicating this protease in the modulation of cell adhesion (Sheppard et al., 1991). Such studies highlight the importance of intracellular proteases in mediating the intracellular modifications that structural plasticity necessitates.

Extracellular matrix (ECM) degradation is thought to be a prerequisite for any structural reorganisation and a subclass of proteolytic enzymes, the serine proteases are centrally involved in this process. The plasminogen activator system has been implicated in ECM clearance during many morphological processes, such as development, tumour metastasis, wound repair and ovulation (Vassilli et al., 1991). The expression of components of the PA system in the adult brain may suggest a role
for this system in the structural plasticity which may underlie aspects of memory consolidation.

1.10 Plasminogen activator system

Plasminogen activators (PAs) are serine proteases of tryptic specificity expressed by the majority of cell types. There are two PAs, tissue-type plasminogen activator (t-PA) and urokinase-type plasminogen activator (u-PA) which have broadly overlapping expression patterns and functions. PAs have activity towards plasminogen, activating the inert precursor and releasing plasmin, a serine protease of broad substrate specificity. Plasmin is then able to activate downstream serine proteases and metalloproteases leading to a cascade of proteolytic degradation (figure 1.2). This system is well characterised in fibrinolysis but also plays an important role mediating cell and ECM interactions (Saksela and Rifkin, 1988).

![Diagram of the plasminogen activator system](image)

Figure 1.2 - The plasminogen activator system

Both PAs are themselves regulated by proteolysis. Cleavage of u-PA and t-PA is required to generate the fully active enzyme, however t-PA retains a degree of catalytic activity in its uncleaved state. Specific inhibitors of the PAs exist; PA inhibitor-1 (PAI-1) is active against both PAs whereas PAI-2 is more active against u-PA. PAs along with plasmin are also regulated via amino-terminal domains which mediate the binding of the protease to preferred sites of action. The presence of kringle and finger
domains allow interactions with fibrin and other ECM components which modulate catalytic activity.

Specific cell binding sites exist for components of the plasminogen activator system which may act to restrict activity to the pericellular milieu. Several different cell surface binding sites have been determined for t-PA, including the mannose receptor and the low density lipoprotein receptor (reviewed in Noorman et al., 1997) though the predominant neuronal binding site is thought to be the heparin binding p30 protein, amphoterin (Parkkinen and Rauvala, 1991). Plasminogen also interacts with the cell membrane via α-enolase (Nakajima et al., 1994). u-PA interacts with a specific cell surface receptor and there is also some indication that the u-PA receptor has a role in signal transduction. Activation of the receptor expressed in epithelial cell lines resulted in morphological changes and reorganisation of the cytoskeleton (Busso et al., 1994).

1.11 Plasminogen activator expression in the developing brain

The development of the central nervous system clearly involves substantial neuroanatomical remodelling and the involvement of the plasminogen activator system is well recognized. The molecular basis of this regulated structural change may provide insights into proteolytic processes at work within the adult brain.

PA activity is present throughout brain development particularly in regions undergoing extensive morphogenesis (Vassalli et al., 1991). High plasminogen activator activity is associated with migrating neural crest cells (Valinsky and leDourain, 1985) and with the migrating granule cells of the developing cerebellum (Krystosek and Seeds, 1981a; Verrall and Seeds, 1988). This latter system has been investigated in vitro and migration of granule cells in organotypic suspension cultures was inhibited by specific inhibitors of t-PA and plasmin (Moonen et al., 1982).

PA activity thus coincides with regions of neuronal migration and process outgrowth during development. Such morphological events require localised proteolysis of the ECM and thus the PA system is implicated in this process. Indeed PA-mediated ECM clearance has been localised to the growth cone of neurons in culture (Krystosek and Seeds, 1981b; Pittman, 1985; Pittman et al., 1989; McGuire and Seeds, 1990), granule cells in sections of developing cerebellum (Moonen et al., 1982) and to proliferating Schwann cells in culture (Kalderon, 1984). Furthermore, PA expression levels in cell culture experiments correlated with the extent of process outgrowth and
cell migration have been found (Pittman and Dibenetto, 1995; Hayden and Seeds, 1996). The exogenous application of PAs (Neuman et al., 1989) can also lead to neurite outgrowth at certain concentrations.

The expression profiles of t-PA and u-PA agree with this proposed role in mediating ECM degradation during neuronal outgrowth and migration. t-PA mRNA is expressed strongly in the ventricular zones of the developing cephalic vesicles, which contain extensively migrating and proliferating neuroepithelial cells (Friedman and Seeds, 1994). High expression is also seen in the olfactory epithelium where cells are actively migrating towards the olfactory canal. Within the developing spinal cord, t-PA expression is initially restricted to the floor plate where it is considered to play a role in the guidance of commissural fibres (Sumi et al., 1992). u-PA mRNA is expressed abundantly during development, within neurons of the cerebellum, olfactory bulbs, hippocampus and cerebral cortex (Dent et al., 1993). u-PA expression is also associated with myelinating glia in developing fibre tracts. Furthermore, expression in the developing spinal cord was detected in the motor neurons of the ventral horns and the onset of expression correlated with process outgrowth (Sumi et al., 1992).

1.12 Plasminogen activator expression in the adult brain

PA activity is also present in adult brain. Using fibrin overlays, extensive activity around the cerebral meninges, choroid plexus and the granule cells of the cerebellum and the dentate gyrus was reported (Soreq and Miskin, 1981). In agreement, a further study revealed PA activity in the hippocampus and meningeal blood vessels along with some activity in the hypothalamus and amygdala (Sappino et al, 1993). Activity was unaffected by inhibitors of u-PA but was abolished by anti-t-PA antibodies. Zymography of brain homogenates revealed the major PA in brain to be t-PA (Masos and Miskin, 1996) and t-PA thus appears to be responsible for the majority of activity in adult brain.

In agreement with the observed activity, widespread t-PA expression was reported in adult mouse brain (Sappino et al., 1993) and expression by microglia and some neurons was confirmed in the mouse hippocampus (Tsirka et al, 1995; 1997). In contrast, expression in the adult rat CNS was reported to be restricted to the ventricular ependyma and the olfactory nerve layer (Dent et al., 1993). A more recent study reported additional expression in rat subcommissural organ and cerebellar Purkinje and
granule cells along with scattered, possibly glial hybridisation signals in the thalamus, brain stem, striatum, fimbria and cortex (Ware et al., 1995).

Similar species differences have been reported for the adult brain expression pattern of u-PA. Low-level expression was only detectable by sensitive RT-PCR in mouse brain (Miskin et al., 1990) with transcripts undetectable by in situ hybridization studies (Sappino et al., 1993; Kristensen et al., 1991). More recent studies detected cortical expression of u-PA in mouse brain, in agreement with the low level of activity revealed by zymographic analysis (Masos and Miskin, 1996). In contrast, in situ hybridization studies on rat brain revealed expression in hippocampal, cortical and cerebellar neurons and in the mitral cells of the olfactory bulbs (Masos and Miskins, 1996; Dent et al., 1997).

Despite these species differences, both t-PA and u-PA are expressed in rat and mouse with significant activity in structurally dynamic brain regions. The continued expression of PAs in adult brain suggests remodelling processes similar to those occurring throughout development, persist in adult brain. The involvement of PAs in the structural changes following synaptic plasticity and potentially memory processes is suggested.

1.13 Plasminogen activators in synaptic plasticity and memory

Several studies have reported transcriptional upregulation of PAs after electrical activity and learning events. t-PA induction was detected in the hippocampus immediately following LTP induction or seizure (Qian et al., 1993) and this induction was blocked by the NMDA receptor antagonist MK801. Expression peaked at 1 hour post-tetanus and lasted at least 4 hours. Furthermore, in transgenic mice harbouring a β-galactosidase reporter gene under transcriptional regulation of the t-PA promoter, enhanced reporter expression was detected in the hippocampus following seizure (Carroll et al., 1994). Expression of u-PA within the hippocampus is also upregulated following seizure. Upregulation was first detected 2-4 hours following seizure and was maintained for up to three days (Masos and Miskin, 1997). t-PA induction was also seen in the cerebellum following motor learning, a process known to induce structural changes in this brain region (Seeds et al., 1995). Interestingly, the rise in t-PA expression had a time scale similar to that of the upregulation reported in the hippocampus following the induction of LTP.
Upregulation following procedures inducing LTP suggests that t-PA may mediate some aspects of structural reorganisation associated with changes in synaptic efficacy. Supporting this theory, mice lacking a functional t-PA gene, although viable, have electrophysiological deficits. Frey et al., (1996) found LTP to be phenotypically normal in homozygous null mice but noted reduced excitability attributable to stronger GABAergic transmission. In the presence of GABAergic inhibitors, stimuli establishing long-lasting LTP in wild type and heterozygous animals only led to short lasting potentiation in animals lacking t-PA. Huang et al., (1996) studied the same line of mice and also found them to have reduced long-lasting LTP both in the presence and absence of GABAergic inhibition. This difference was attributable to different levels of tonic inhibition in the two experimental conditions used. Null mice were analysed for behavioural deficits and were found to perform significantly poorly in two-way active avoidance tasks. Spatial learning in the Morris water maze was, however, largely unaffected by the genetic lesion, perhaps due to a residual amount of long-lasting potentiation in homozygous mice. Behavioural impairments have also been reported in a line of transgenic mice overexpressing u-PA within the brain suggesting that well regulated proteolysis is a requirement for these higher cognitive abilities (Meiri et al, 1994).

The promoter region of the t-PA gene has a CRE response element and several AP-1 motifs indicating that transcription may be regulated by the CREB family of transcription factors (Leonardsson and Ny, 1997). The maintenance of long-term LTP requires de novo gene transcription, mediated, in part, by CREB. The activation of this transcription factor during LTP may contribute to the transcriptional activation of this serine protease. CREB mediated transcription has been previously correlated with structural changes within the brain (Murphy and Segal, 1997), possibly mediated by t-PA. t-PA is also regulated by release from storage vesicles (Parmer et al., 1997) and, in neurons, this release was found to be contingent upon membrane depolarisation (Gualandris et al., 1996).

The activity dependent induction of PA transcription may explain the lack of concordance in the in situ hybridization studies described above. The differences in expression patterns may reflect differing levels of arousal and activity of animals used in each study.
1.14 Plasminogen in the central nervous system

The hypothesised role of the PAs in developmental and synaptic plasticity is the induction of a proteolytic cascade of ECM degradation consequent upon plasmin activation. However, plasminogen is principally synthesized in the liver and the 90 kDa protein is considered too large to cross the blood brain barrier. Local synthesis of this serine protease within the central nervous system has therefore been hypothesized.

Plasminogen mRNA was detected in mouse brain by the sensitive RNase protection assay (Sappino et al., 1993) although the result was not presented. Plasminogen was also detected immunologically as a secreted product of cultured microglia suggesting its synthesis within the CNS (Nakajima et al., 1992a). More recently, plasminogen transcripts and protein were detected at a low level in the mouse hippocampus although the expression was considered neuronal rather than from microglia (Tsirka et al., 1997). On the basis of these studies it can be concluded that local synthesis of plasminogen does occur but the levels are generally low. In agreement, studies investigating PA activity by substrate lysis over brain sections or neurons in culture, found the activity to be dependent upon added plasminogen protein (e.g. Krystosek and Seeds, 1981a).

Transgenic mice lacking a functional plasminogen gene, are viable and, although detailed anatomical studies have not been reported, appear to have no abnormal developmental phenotype. Plasminogen-deficient mice are however predisposed to severe thrombosis in agreement with the circulatory role (Bugge et al., 1995). The detrimental effects of plasminogen deficiency are rescued by a deficiency in a functional fibrinogen gene (Bugge et al., 1996). These two studies suggest that fibrinogen is essential for the expression of all the pathologies associated with plasminogen deficiency and suggest that the only essential physiological role of plasminogen is in fibrinolysis. The role of plasminogen in brain is therefore uncertain.

In vitro studies, however, suggest a role for plasminogen in brain function. Addition of plasminogen resulted in neurite outgrowth from cortical explants (Nakajima et al., 1993) and enhanced neuronal survival in cultured mesencephalic neurons (Nagata et al., 1993). Plasminogen has also been implicated in synaptic plasticity. Perfusion of plasminogen or plasmin enhanced short term potentiation induced in rat hippocampal slices and conversely potentiation was blocked by perfusion of α2-antiplasmin
(Mitzutani et al., 1996). However, these effects could result from the addition of a general tryptic serine proteases rather than specific effects of added plasminogen.

Although it appears that plasminogen is expressed in neural tissue, the low levels contrast sharply with the abundant PA activity and thus it seems likely that other PA substrates exist. Some clearance of fibronectin over cultured sensory neurons occurred in the presence of plasmin inhibitors (McGuire and Seeds, 1990) indicating ECM degradation does not necessarily require the presence of plasminogen. Accordingly, u-PA has limited activity towards fibronectin (Quigley et al., 1987) and a 66 kDa component of the ECM in vitro (Keski-Oja and Vaheri, 1982). Furthermore u-PA has been shown to activate type IV collagenase in the absence of plasminogen (Reith and Rucklidge, 1992).

1.15 Multiple trypsin-like serine protease within the CNS

Over the last two years, several groups have reported the isolation of cDNA clones encoding novel serine proteases from CNS tissue. All these proteases are of tryptic specificity and are presumed to require proteolytic activation by tryptic proteases. The involvement of these proteases in cascades of successive activation, possibly acting downstream of PAs warrants investigation.

Neuropsin, a trypsin-like serine protease was cloned from mouse brain and in situ hybridization revealed expression of this species was restricted to the limbic system (Chen et al., 1995). Interestingly, neuropsin transcription within the hippocampus was found to be upregulated following kindling (Okabe et al., 1996).

Neurotrypsin, a multidomain trypsin-like serine protease, was cloned from mouse brain and found to be expressed in the hippocampus, cerebral cortex, amygdala, olfactory bulbs and certain brain stem nuclei (Gschwend et al., 1997). This protease is of interest as it has many regulatory domains homologous to those of the plasminogen activators, indicating a possible role in ECM degradation.

Another trypsin-like serine protease, neurosin, was recently cloned from a human adenocarcinoma cell line (Yamashiro et al., 1997). Strong expression of this species was also detected in brain. Expression of this protease within a carcinoma is suggestive of a role in the structural remodelling required by metastatic growth. Plasminogen activator-mediated ECM degradation is well characterised in structural
reorganisation during tumour invasion and, as described above, within the brain. A role for neurosin in brain structural plasticity is thus suggested.

A trypsin-like serine protease, myelencephalon-specific protease (MSP) has recently been cloned from spinal cord tissue (Scarisbrick et al., 1997). The expression of this species was predominantly restricted to the hindbrain and spinal cord. MSP mRNA levels were found to increase three fold in response to excitotoxic injury induced by kainic acid, indicating a potential role in activity dependent processes and spinal cord injury.

A trypsin-like proteolytic activity, termed erase, has been characterised in PC12 cells and is implicated in the regulation of growth cone mobility (Baird and Raper, 1995). The activity caused collapse of retinal growth cones in culture and is inhibited by specific inhibitors of trypsin-like proteases, but not thrombin inhibitors. This tryptic activity is hypothesized to represent at least one of the avoidance cues which regulate outgrowth from ganglion cells during retinal development in vivo.

Levels of protein corresponding to an unknown serine protease were found to be upregulated following the induction of LTP in the dentate gyrus in vivo (Fazeli et al., 1990). The protease was not identified although zymography revealed the 80 kDa protein to be of tryptic specificity.

In addition neuroserpin, a serine protease inhibitor of tryptic specificity has recently been identified in the central nervous system as an axonally secreted glycoprotein (Osterwalder et al., 1996). Neuroserpin is expressed widely during development and in adult expression is most prominent in the hippocampus, striatum and cerebellar Purkinje cells (Krueger et al., 1997). This expression pattern thus suggests a role for this inhibitor in the reorganization of synaptic connectivity during development and synaptic plasticity in the adult.

Finally, acetylcholinesterase, an abundant enzyme in the central and peripheral nervous system also has a trypsin-like proteolytic activity. Diisopropyl flurorophosphate (DFP), a potent inhibitor of serine proteases and esterases, was found to bind to two sites in acetylcholinesterase and binding to the tryptic site was inhibited by other serine protease inhibitors (Small and Chubb, 1988). The physiological relevance of this proteolytic activity is unknown however the enzyme is implicated in neuropeptide processing (see section 1.19).
1.16 Thrombin and protein nexin-1

A glia-derived neurite promoting factor which induced morphological differentiation in neuroblastoma cells was found to have protease inhibitory activity (Guenther et al., 1985). Subsequent cloning of the factor revealed it to be identical to the serine protease inhibitor protease nexin-1 (PN-1), secreted by a variety of different cell types (Gloor et al., 1986). PN-1 has activity against thrombin, PAs, plasmin and trypsin in vitro, forming SDS-resistant complexes which bind to the cell surface and are rapidly internalised and degraded (reviewed in Knauer and Cunningham, 1984). Protein nexins thus serve to limit extracellular proteolysis. PN-1 expression is widespread throughout development and in the nervous system is expressed early in the neuroepithelium of the closing neural tube. Adult expression is predominantly in the olfactory bulbs, the cerebellum and the striatum (Mansuy et al., 1993).

Interest has focused on the protease thrombin as PN-1 activity is restricted to this protease when bound to the cell surface or the extracellular matrix (Wagner et al., 1989a). The in vitro activities against other tryptic proteases may therefore be of limited physiological significance. Thrombin, synthesized as inactive prothrombin, is activated at sites of tissue injury and is classically involved in the first steps of the blood coagulation pathway. Prothrombin is also expressed in the developing nervous system and expression in adult is widespread, implying a more general role (Dihanich et al., 1991).

PN-1 and thrombin have been shown to cause a variety of opposing morphological effects on neurons in culture. Application of PN-1 initiated neuroblastoma differentiation, causing neurite outgrowth and upregulation of neurotransmitter biosynthetic enzymes. These effects were inhibited and reversed by application of thrombin (Gurwitz and Cunningham, 1988). PN-1 and thrombin have also been shown to induce reciprocal morphological effects on glia in culture. PN-1 addition to cultured astrocytes induced a stellate morphology whereas thrombin reverted cells back to their protoplasmic form (Cavanaugh et al., 1990). PN-1 also inhibited granule cell migration in cultured cerebellar explants and these effects were suppressed by thrombin (Lindner et al., 1988).

In general, the effects of PN-1 can be reproduced by other serine protease inhibitors. Addition of soybean trypsin inhibitor and the inhibitor leupeptin induced neurite outgrowth in cultured dorsal root ganglia (Hawkins and Seeds, 1986; 1989). Specific
inhibitors of thrombin, such as hirudin, have also been found to induce neuroblastoma differentiation (Gurwitz and Cunningham, 1988). The morphogenic effects of PN-1 may therefore be due to the inhibition of thrombin, present within the serum in culture media. Indeed serum withdrawal was sufficient to induce neuroblastoma differentiation (Gurwitz and Cunningham, 1988). Furthermore, in the absence of serum, PN-1 had no detectable effects and antibodies blocking PN-1 activity allowed differentiation induced by hirudin or serum withdrawal (Gurwitz and Cunningham, 1990).

Thrombin inhibition thus appears to mediate the observed morphogenic effects and the apparent neurotrophic effects of PN-1. Thrombin has been found to act via a specific cell surface receptor, indicating that the effects may result from an intracellular signaling pathway altering neuronal function directly. Activation of the receptor occurs due to the release of an activation peptide by thrombin cleavage (Vu et al., 1991). The thrombin receptor is expressed throughout the CNS (Weinstein et al., 1995) and synthetic thrombin receptor activation peptides have been shown to reproduce the effects of thrombin on neurons in culture (Suidan et al., 1992). The downstream second messenger pathways activated by this receptor are as yet undetermined although serine/threonine kinase inhibitors have been shown to block thrombin receptor mediated neurite retraction (Suidan et al., 1992). Many other second messenger systems have been implicated in intracellular signalling following thrombin receptor activation (Turgeon and Houenon, 1997).

Consistent with a potential role in structural plasticity within the CNS, thrombin and PN-1 are implicated in synaptic reorganisation at the neuromuscular junction (Liu et al., 1994; Zoubine et al., 1994). In newborn mammals individual skeletal muscle fibres are innervated by several motor neuron axons. During the first few weeks of postnatal life activity dependent synapse elimination occurs until each muscle fibre is innervated by a single motor axon. The neuromuscular junction therefore serves as a useful model for understanding the mechanisms regulating synapse development and plasticity.

A role for PN-1 in synaptic plasticity has been suggested by transgenic studies. Overexpression of PN-1 in mice leads to enhanced LTP whereas mice deficient in PN-1 have reduced LTP (Lüthi et al., 1997). The levels of PN-1 thus appear to correlate with the capacity for LTP. An increase in PN-1 levels would be expected to restrict
thrombin activity, thus increasing the proliferation of neurons and astrocytes and perhaps facilitating long-term increases in synaptic efficacy.

Thrombin and PN-1 were originally considered to act in a similar manner to the plasminogen activator system, mediating structural change or stability through controlled degradation of the ECM. The involvement of a specific receptor however argues against such theories and suggests that thrombin has a role in modulating neuronal function directly. Other protease activated receptors have recently been characterised including a second thrombin receptor (Ishihara et al., 1997) and a trypsin receptor (Nystedt et al., 1995). These discoveries suggest the presence of a multi-gene family of protease activated receptors which may mediate the effects of neuronal proteases.

1.17 Extracellular matrix interactions

As discussed, extracellular proteases are thought to be involved in the controlled proteolysis of the ECM during neurite outgrowth, cell migration and synaptic remodelling. In turn, many ECM molecules are able to modulate neurite outgrowth in vitro. Laminin is a potent neurite outgrowth promoting molecule and fibronectin, collagen and some proteoglycans have also been shown to possess stimulatory domains (reviewed in Kapfhammer and Schwab, 1992). Other ECM components act as inhibitors of neurite outgrowth.

The effects exerted by ECM molecules are mediated by a variety of cell surface receptors. Integrins are abundant in the developing and adult brain and bind laminins, fibronectin and collagen. These integral membrane proteins interact with the cytoskeleton and so are able to signal ECM interactions intracellularly (James et al., 1986). The proteolytic modification of the ECM can thus lead to intracellular signals that effect cellular function directly (discussed in Werb et al., 1997).

The brain has a somewhat limited ECM compared to other tissues. Basement membranes are confined to meningeal surfaces and blood vessels and the parenchyma of the CNS is filled with an amorphous matrix containing little collagen and other fibrous protein. However the majority of well characterised ECM molecules are present to some degree within the CNS (reviewed in Venstrom et al., 1993). Laminin isoforms are expressed in the developing CNS in a pattern which parallels elongating nerve fibres. Laminin can form complexes with proteoglycans which are abundant
ECM components in the brain parenchyma (Lander, 1993). Other components include fibronectin, collagens, thrombospondins and a diverse range of glycoproteins.

Recent evidence implicates ECM and PA system interactions in novel aspects of ECM function. During wound healing, smooth muscle cell migration is enhanced by the binding of the ECM component, vitronectin to a specific receptor integrin. This interaction was found to be inhibited by vitronectin forming complexes with PAI-1 (Stefansson and Lawrence, 1997). PA/PAI-1 complexes thus serve to reduce the levels of free PAI-1 which then restores the ability of vitronectin to promote cell migration. The PA system may therefore have more diverse effects on ECM than simply degradation.

Adding to the complexity of the protease/antiprotease interactions within the CNS, a family of ECM glycoproteins has recently been characterised which possess serine protease inhibitor domains indicating a direct role for the ECM itself in limiting proteolysis.

Agrin is a multidomain protein with laminin homology regions, epidermal growth factor motifs and basement membrane binding motifs. The amino terminal region of the protein has nine tandem repeat homologous to the Kazal family of protease inhibitors (Rupp et al., 1991) and functional analysis indicates that agrin can act as a potent inhibitor of trypsin, chymotrypsin and plasmin (Biroc et al., 1993). The Kazal domains are most similar to the inhibitor domains of follistatin, itself implicated in the early stages of neural development (Hemmali-Brivanlou et al., 1994). In the peripheral nervous system, the development of neuromuscular junctions is characterised by the clustering of acetylcholine receptors upon innervation of muscle tissue by peripheral nerves. Agrin is the signal for this receptor clustering and also causes changes in the distribution of acetylcholinesterase and components of the neuromuscular ECM, indicating a general role in synapse organisation and development (Nastuk and Fallan, 1993). The physiological relevance of the protease inhibitor domains is uncertain although they may have a role in limiting the activity of a protease, facilitating neurite outgrowth upon muscle innervation. Indeed, transgenic mice deficient in agrin have considerable axon overgrowth (Gautam et al., 1996).

Agrin is also implicated in the developing and adult CNS indicating a possible universal role in synapse formation and function. Indeed certain alternatively-spliced isoforms of agrin are expressed exclusively within CNS. High levels of agrin
expression have been detected in the developing brain with levels declining after birth, indicating a role in synapse formation. Expression is retained in regions of the adult brain which maintain a high degree of structural plasticity such as the hippocampus, the olfactory bulbs and the neocortex (Cohen et al., 1997). Furthermore, agrin expression was found to be upregulated following seizure implicating the protein in the structural reorganisation following activity dependent processes (O'Connor et al., 1995).

Other glycoproteins with Kazal-like inhibitory domains have been characterised within the nervous system including osteonectin/SPARC/BM-40 (Mendis et al., 1994a), SC1 (Johnston et al., 1990), testican (Bonnet et al., 1996) and QR1 (Guerman et al., 1991). These extracellular proteins have been implicated in structural remodelling during development and following injury. As discussed in previous sections, such processes require active proteolysis, possibly mediating ECM degradation. These extracellular inhibitors may therefore serve to limit the cascades of proteolysis, necessary for well regulated structural change.

1.18 Serine proteases and synaptic transmission

As discussed above, serine proteases have well characterised roles in structural remodelling which may mediate aspects of long-term changes in synaptic transmission. A direct effect of serine proteases on synaptic transmission has been suggested by a variety of studies investigating the effects of proteolysis on receptor function. Addition of plasminogen increased intracellular Ca$^{2+}$ concentrations and these effects were attenuated by application of tetrodotoxin, implicating the voltage sensitive Ca$^{2+}$ channels (Inoue et al., 1994). In the presence of tetrodotoxin, plasminogen also enhanced NMDA evoked Ca$^{2+}$ increases indicating a general modulatory effect of plasminogen on postsynaptic function.

Conflicting reports implicate trypsin in NMDA receptor function. Trypsin irreversibly inactivated NMDA receptor mediated currents in primary hippocampal cultures. The effects were specific, leaving other glutamatergic transmission intact (Allen et al., 1988). Conversely, trypsin application on rat hippocampal slices led to epileptiform activity and this hyperexcitability was unchanged for up to 60 minutes following washout (Yamada and Bilkey, 1993). The effects were again specific as they were blocked by non-competitive NMDA-R antagonists.
Proteolysis has been shown to be a regulatory mechanism for voltage sensitive calcium channels. L-type channels undergo proteolytic modification following NMDA-receptor activation (Hell et al. 1996). This cleavage removes a carboxyl terminal region of the receptor α₁ pore forming subunit leading to increased ion conductance. The dependence of this proteolytic regulation on NMDA-receptor activation indicates that intracellular proteases may contribute to aspects of long-lasting enhanced synaptic transmission.

Proteases are therefore implicated in receptor modulation. The effects seen appear to be irreversible implying perhaps specific receptor targets exist for proteolytic degradation or modification. In peripheral epithelial tissue, a trypsin-like serine protease has recently been implicated in the activation of sodium channels (Vallet et al., 1997) indicating a proteolytic mechanism of receptor modulation may be widespread.

1.19 Neuropeptide and growth factor processing

Post-translational proteolysis is a common mechanism in the synthesis of biologically active proteins and peptides. Frequently the precursor protein is processed by cleavage at the carboxyl side of pairs of basic amino acids, especially, lysine-arginine and arginine-arginine (reviewed in Loh et al., 1984). Trypsin like serine proteases cleave preferentially after arginine residues and so are implicated in the maturation of many growth factors, hormones and bioactive peptides.

Within the central nervous system, over 50 neuropeptides are characterised. Neuronal proteases would thus be expected to play an important role in the processing of neuropeptide precursors. Many such proteases with neuropeptide processing activity have been identified in brain, but the majority are members of the metalloprotease family (for example, Dauch et al., 1992; Barnes et al., 1988). One trypsin-like serine protease has been implicated in the processing of somatostatin (Waters and David, 1997) and the trypsin activity associated with acetylcholinesterase has activity towards enkephalins and substance P (Chubb et al., 1980; 1983).

Nerve growth factor is also processed after an arginine residue (Scott et al., 1983) and various serine proteases have been implicated in this activity. In the submaxillary gland, the active βNGF is present as a complex with αNGF and γNGF (Young et al., 1988). γNGF is an active trypsin-like protease and thus may have an important role in NGF processing, however, expression of this serine protease is restricted to this gland.
and is not thought to play a role in NGF processing within the central nervous system. Interestingly serine proteases of the subtilisin family, furin and PC2 have been implicated in the processing of βNGF (Bresnahan et al., 1990) and both these proteases are expressed in the CNS (Schäfer et al., 1993).

Hepatocyte growth factor (HGF) is expressed in the central nervous system (Honda et al., 1995) and can be activated by tryptic serine proteases (Mars et al., 1993). Similarly, transforming growth factor-β secreted by migrating neural crest cells can be activated by plasmin (Brauer and Yee, 1993). Epidermal growth factor is associated with a binding protein which is an active trypsin-like protease (Lundgren et al., 1984), and the binding proteins of insulin like growth factors are proteolytically regulated (Rajah et al., 1996).

Serine proteases within the CNS could thus be functioning as processing enzymes for bioactive molecules. Both neuropeptides and growth factors have been implicated in synaptic plasticity, thus serine protease could play an important role in their activity dependent regulation.

1.20 Summary

A diversity of serine proteases and their inhibitors are expressed in brain and, for many of them, a body of evidence argues for a role in neurite outgrowth and migration and the synaptic remodelling that accompanies activity-dependent processes such as LTP. Reorganisation of the ECM is thought to be central to such processes, and hence brain serine proteases such as those of the PA system are likely to play a prominent role. The precise downstream targets of the PAs are however poorly characterised within the CNS.

As well as a role in synaptic remodelling via ECM interactions, serine proteases such as thrombin and u-PA mediate their effects through interactions with specific cell surface receptors, indicating a possible role in signal transduction. Serine proteases are also implicated in the modulation of receptor function and thus may contribute to long-term changes in synaptic efficacy. This diverse family of proteolytic enzymes may also contribute to the processing of neuropeptides.
C - SERINE PROTEOLYSIS IN BRAIN DYSFUNCTION

The involvement of serine proteases in adult brain function has been described. A role for these enzymes in the structural changes underlying learning and memory processes is inferred. It is therefore of note that serine proteases are implicated in the pathogenesis of Alzheimer's disease, a condition characterised by progressive cognitive decline and memory loss. Serine proteases are also implicated in modulating neuronal susceptibility to cell death following injury and in the resulting local immune response.

1.21 Proteolysis and Alzheimer’s disease

Alzheimer's disease is the most common neurodegenerative disorder affecting one-half to two thirds of all cases of late-life cognitive impairment. The disease is characterised by progressive dysfunction and dystrophy of neurons and glia in the limbic and association cortices and in certain subcortical nuclei that project to these areas. As has been discussed, the limbic system has an important role in higher order cognitive function and limbic degeneration leads to severe cognitive decline. Entorhinal layer II neurons are affected early by this progressive dystrophy and it is these neurons that provide the hippocampus with the majority of its input. The neuroanatomical isolation of the hippocampus is considered to underlie the characteristic memory loss seen in Alzheimer patients.

Histologically, the dysfunctional brain is characterised by the deposition of amyloid plaques, significantly in the meningeal and intracortical blood vessels and also in the brain parenchyma. These extracellular plaques have a central deposit of amyloid fibres surrounded by dystrophic neurites, along with activated microglia and astrocytes. Immunocytochemical studies also reveal a diffuse accumulation of amyloid throughout the brain. The disease pathology also includes the accumulation of abnormal fibres in neuronal cell bodies. These neurofibrillary tangles are paired helical filaments composed of an abnormally phosphorylated form of the microtubule associated protein, tau. A similar pathology is associated with hereditary cerebral hemorrhage with amyloidosis, Dutch type and with the early onset dementia seen in Down's syndrome.

Amyloid plaques are predominantly composed of a 4 kDa protein, β-amyloid (Master et al., 1985). This protein is proteolytically derived from a precursor, the amyloid
precursor protein (APP), highly conserved between species and expressed in most tissues (Kang et al., 1987). The structure of APP is summarised in figure 1.3.

The structure of APP is shown in the top panel, hashed region represents a cysteine rich amino terminus of APP; the shaded area represents an anionic domain; the alternatively spliced KPI domain is shown, as are two potential N-glycosylation sites (CHO); The β-amyloid region is shown as a black box.

The processing of APP is shown in the lower panel. APP amino acid sequence surrounding the β-amyloid domain is shown in relation to the transmembrane region. Site of cleavage by the α, β, and γ secretases are indicated.

The generation of amyloidogenic subfragments requires proteolytic cleavage of the APP polypeptide. Proteolysis is also implicated in the normal function of this precursor molecule. Cleavage within the β-amyloid domain leads to the secretion of the large soluble amino terminus of the protein, αAPP. Cleavage occurs predominantly between leucine-18 and lysine-19 of the β-amyloid domain thus precluding the production of the amyloidogenic fragment (Esch et al., 1990). The enzyme responsible for this cleavage has been termed α-secretase. Amyloidogenic processing involves the cleavage of APP between methionine-596 and aspartic acid-597 by the so called β-secretase, yielding the amino terminus of β-amyloid. The action of a third protease, γ-secretase, which cleaves within the transmembrane region at one of three positions, releases β-amyloid from its precursor as a 39, 42 or 43 residue polypeptide (reviewed in Selkoe, 1991) (figure 1.3).
Familial Alzheimer's diseases is often associated with mutations at the locus encoding APP. Mutations tend to cluster around the potential amyloidogenic cleavage sites. For example, a large Swedish pedigree has a double mutation at the β-cleavage site (Mullen et al., 1992) which increasing the proportion of aberrantly cleaved APP and thus the amount of β-amyloid deposition.

Numerous studies have attempted to characterise the proteases responsible for the α, β and γ-secretase activities, and a variety of candidates have been proposed. Interest has focused on serine proteases with regard to the β-secretase activity as the cleavage event required occurs at a site of chymotryptic specificity. A specific inhibitor of serine proteases led to a 50% reduction in the levels of β-amyloid production in vitro (Citron et al., 1996). An APP processing activity was fractionated from brain homogenates and was attributed to a cathepsin G-related Ca2+ stimulated serine protease (Razzaboni et al., 1992). However this protease cleaved APP one residue upstream of the proposed β-secretase cleavage site. Cathepsin G itself, a chymotryptic lysosomal protease, was shown to cleave APP at the correct amyloidogenic site and furthermore was able to cleave a decapeptide containing the Swedish double mutation appropriately (Sahasrabudhe et al., 1993). A chymotrypsin-like mast cell protease has been shown to cleave recombinant APP at the β-secretase site and has been localised to the meningeal and intracortical blood vessels, a major site of amyloid deposition (Nelson et al., 1992). The absence of mast cells in brain parenchyma, however argues against the participation of this enzyme in the formation of neuritic plaque core deposits. Other classes of protease have been implicated as β-secretase candidates including cysteine proteases and metalloproteases.

Serine proteases are also implicated in the mechanisms that exist to limit the accumulation of β-amyloid in the undiseased brain. A specific β-amyloid degrading activity was detected in cos cell conditioned medium and was found to be sensitive to serine protease inhibitors (Qiu et al., 1996). The presence of such an activity in brain tissue was not investigated so the physiological relevance of this study is questionable.

The serine protease elastase has activity towards APP at the carboxy terminus of the β-amyloid domain (Evin et al., 1995). However numerous other enzymes have been implicated including metalloproteases and the aspartyl protease cathepsin D (Landror et al., 1994). The heterogeneity associated with γ-secretase cleavage at the carboxy terminus of β-amyloid perhaps indicates multiple enzyme involvement.
The lysosomal cathepsins have been implicated as the proteases responsible for amyloidogenic processing of APP. This is of significance as dentritic plaques are associated with extracellular lysosome deposits amongst the proximal dendrites (Cataldo et al., 1991). Furthermore, an upregulation of components of the endosomal/lysosomal system is seen in the disease (Cataldo et al., 1995; 1997). The dysregulation and abnormal extracellular localization of lysosomal acid hydrolases can perhaps account for the multiple proteolytic events that result in the production of insoluble β-amyloid.

The majority of APP proteolytic studies have been performed on peptides representing the potential cleavage sites of APP. This approach is somewhat flawed as differences between the ability to cleave peptides and the full length substrate have been found (Brown et al., 1996). Furthermore, the mast cell chymotrypsin characterised by Nelson et al. (1993) was able to cleave recombinant APP at the required position but, in accordance with the presence of multiple chymotryptic cleavage sites, extensively cleaved the precursor. However, the β-sheet configuration of amyloid fibrils may convey proteolytic resistance suggesting that only certain target sites are available as substrates.

As discussed above, chymotryptic proteases are the main candidates responsible for the β-secretase activity. Significantly, an inhibitor of chymotryptic proteases, α1-antichymotrypsin (ACT), is present in the neuritic plaques of Alzheimer disease and transcription of this gene is upregulated in the diseased brain (Abraham et al., 1988). The presence of this inhibitor may thus act to limit amyloidogenic processing of the precursor. However, ACT has a role in regulating proteases involved in inflammatory response so the upregulation of this gene may simply reflect local inflammation due to neuronal degeneration.

Serine proteases are further implicated in Alzheimer's disease as certain isoforms of APP contain a serine protease inhibitor domain (Ponte et al., 1988). The alternatively spliced exon 7 encodes for a domain which is 33-48% homologous to members of the kunitz family of serine protease inhibitors (KPI). Both KPI+ and PKI- isoforms are expressed in brain (Tanzi et al., 1988) although protein studies suggest that the KPI containing APP is predominant in brain (Van Nostrand et al., 1991).
Structurally, the inhibitor is most similar to bovine pancreatic trypsin. *In vitro* studies suggested inhibition of trypsin, chymotrypsin and plasmin occurs most strongly. Activity was also detected against elastase and proteases of the blood coagulation pathway yet the domain has no activity towards the plasminogen activators or thrombin (Kitaguchi *et al.*, 1990; Van Nostrand *et al.*, 1990).

Soluble APP was found to represent the coagulation factor XIA inhibitor (Smith *et al.*, 1990). This inhibitor is stored in the α granules of platelets and is released when stimulated by thrombin. This suggests a physiological role for peripheral APP in regulating coagulation.

The soluble secreted form of APP containing the KPI domain was also shown to be identical to nexin II (van Nostrand *et al.*, 1989), a protease inhibitor secreted by a variety of cultured extravascular cells. Protease nexins form covalent linkage with target proteases which bind to cells and are then rapidly internalized and degraded (Knauer and Cunningham, 1984). This suggests a potential role for APP in the removal of selective proteases from the extracellular environment. Since serine proteases are implicated in the processing of APP, the precursor itself could potentially modulate the activity of the secretases.

Similar to protease nexin I, APP (nexin II) has neurotrophic effects on neurons in cell culture. Low concentrations of APP were found to potentiate the neurotrophic effects of NGF, causing an increase in neurite outgrowth and branching (Milward *et al.*, 1992). However, this ability does not correlate with the presence of a KPI domain implying a protease independent mechanism. Other studies have implicated the β-amyloid protein itself in this activity. Application of fragments of β-amyloid increased neuronal survival of embryonic rat hippocampal cells (Whitson *et al.*, 1989; Yanker *et al.*, 1990). This surprising finding suggests that amyloid deposition may be linked to regeneration in response to the denervation and cellular abnormalities associated with Alzheimer’s disease. In the diseased brain, the hippocampal circuitry undergoes compensatory growth following perforant path degeneration. Associational and commissural connections, along with extrinsic septal inputs, undergo sprouting to form new synapses with denervated target cells (Geddes *et al.*, 1985).

In contrast to such theories, β-amyloid has neurotoxic effects on neurons in culture (Yanker *et al.*, 1989) which implicates the amyloid deposits directly in the degeneration of neurons in the diseased state. It appears that β-amyloid at low
concentrations is neurotrophic and at high concentrations is neurotoxic (Yanker et al., 1990).

APP is also considered to have a structural role. APP was found to be a heparan sulphate proteoglycan (HSPG) core protein (Schubert et al., 1988). HSPGs are cell surface and ECM molecules which interact strongly with other ECM components such as fibronectin, laminin and the cell-adhesion molecule, N-CAM. HSPGs have important roles in synaptic vesicles and at synaptic contacts, so the differential processing of APP in the diseased state could result in structural abnormalities leading to alterations in synaptic transmission.

Familial Alzheimer's disease is also genetically linked with the \( \varepsilon_4 \) allele of the apolipoprotein E (ApoE) (Strittmatter and Roses, 1996). One model for the involvement of ApoE is that certain cleavage products of the protein, present in cerebral spinal fluid exhibit neurotoxicity. A 22 kDa thrombin cleavage product is analogous to this fragment and \textit{in vitro} studies indicate that the ApoE\( _4 \) cleavage product is considerably more neurotoxic than the non-Alzheimer linked allele ApoE\( _3 \) derived product (Tolar \textit{et al.}, 1997). Serine proteases are thus implicated in the pathology of this familial form of Alzheimer's disease.

Serine proteases are implicated both in the constitutive and amyloidogenic processing of APP and serine protease inhibitors are found specifically in neuritic plaques. Furthermore the presence of a KPI domain in isoforms of APP suggests a role for serine protease in the as yet undetermined function of this molecule. One model which is suggested is that impairment in the processing of the APP serine protease inhibitor, leads to an alteration in the balance of proteolysis in the diseased state. The requirement for regulated proteolysis in structural changes accompanying memory has been discussed and thus, a perturbation of this regulation could explain the cognitive decline seen in Alzheimer's disease. Serine proteases may therefore provide a target for therapeutic intervention in the pathogenesis of Alzheimer's disease.

1.22 The role of serine proteases in neuronal cell death

Extracellular serine proteases have modulatory effects on neuronal cell death induced by a variety of insults. Thrombin and PN-1 have opposing effects on neuronal sensitivity to glucose deprivation induced injury. The addition of PN-1 to hippocampal neuronal cultures prior to glucose withdrawal resulted in attenuated cell
death whereas additions of thrombin exaggerated the effects of glucose deprivation (Smith-Swintosky et al., 1995a). Thrombin itself can induce neuronal cell death via the activation of the thrombin receptor (Donovan et al., 1997). Neurotoxicity induced by β-amyloid was similarly modulated by thrombin and PN-1 (Smith-Swintosky et al., 1995b), and both species are known to be present in neuritic plaques (Wagner et al., 1989b; Akiyama et al., 1992). It is interesting to note that PN-1 is extensively downregulated in Alzheimer patients (Wagner et al., 1989b), perhaps implicating this protease inhibitor in neurodegeneration. However, a contradictory study indicated that low concentrations of thrombin can have a neuroprotective effect, attenuating neuronal cell death induced by β-amyloid (Pike et al., 1996) and cerebrovascular injury (Vaughan et al., 1995). The differences in the observed effects of thrombin can be explained by different concentrations of protease used in each study, thus it appears that low concentrations of thrombin can be neuroprotective whereas high concentrations are neurotoxic. The levels of neuronal proteolysis are thus proportional to the susceptibility to neurotoxicity.

A similar conclusion can be drawn from the effects of the PA system on excitotoxic cell death. Transgenic mice deficient in t-PA are resistant to neuronal cell death induced by a overactivation of glutamatergic system in the hippocampus. Intracerebral injections of kainic acid resulted in minimal cell death in t-PA nulls, whereas extensive cell death resulted in wild type controls (Tsirka et al., 1995). These effects are not due to any developmental defect as t-PA introduced into the hippocampus in t-PA mutant animals caused regained susceptibility to excitotoxic neurodegeneration. Furthermore, infusion of PAI-1 into wild type mice led to excitotoxic resistance (Tsirka et al., 1996). Transgenic mice deficient in plasminogen are also resistant to excitotoxic cell death and infusion of α2-antiplasmin into wild type animals conferred resistance (Tsirka et al., 1997).

Neuronal protease activity is thought to mediate the structural plasticity that accompanies long term changes in synaptic transmission. The neuronal cell death accompanying excessive excitation may therefore be due to excessive proteolysis. Membrane depolarisation causes the release of t-PA from regulated secretory granules (Gualandris et al., 1996). Prolonged membrane depolarisation caused by excitotoxins or by dysregulation of membrane channels following glucose deprivation, could potentially lead to excessive release of t-PA. High levels of unchecked extracellular proteolysis could destabilize ECM, cell membrane and cytoskeleton interactions leading to cell damage and death.
Some evidence for this model is provided by the *weaver* mutant mouse which harbours a mutation in an inward rectifying K\(^+\) channel. The mutant mice are associated with cerebellar granule cell death which is thought to explain the characteristic ataxic phenotype. The alteration in a neuronal ion channel could result in enhanced depolarisation and subsequent excessive release of PAs, contributing to the granule cell death seen. In agreement with such theories, abnormally high levels of t-PA were found in granule cells of the mutant cerebellum (Murtomäki *et al.*, 1995)

1.23 Serine proteases in microglial function

There is some evidence to suggest that proteases and their inhibitors are involved in the glia responses to CNS insults. Microglia are ontogenetically related to cells of the mononuclear phagocyte lineage. Upon activation following injury or infection, they express macrophage markers and have important roles in phagocytosis, antigen presentation and tissue repair (Kreutzberg, 1996). Consistent with a phagocytic role, microglia are hypothesized to express the neutral proteases of polymorphonuclear leukocytes, including the serine proteases neutrophil elastase, proteinase 3 and cathepsin G (Perry and Gordon, 1988). Neutrophil elastase has been identified as a secretory product of rat microglia in culture (Nakajima *et al.*, 1992b).

A direct role for microglia proteolysis in neuronal cell death is suggested in retinal development. Towards the end of development, a reduction in the number of retinal ganglion cells is seen. This neuronal cell death was found to be accompanied by microglia activation and phagocytosis of cellular debris (Thanos, 1991). Injection of protease inhibitors protected the ganglia cells from neuronal cell death suggesting a role for microglial proteases in developmentally regulated cell death.

The plasminogen activators are also implicated in microglia function in neuronal cell death. Transgenic mice deficient in t-PA have attenuated microglial activation following seizure in the hippocampus (Tsirka *et al.*, 1995). Interestingly microglial activation in plasminogen deficient mice is normal, suggesting a plasminogen-independent role of t-PA in microglia activation (Tsirka *et al.*, 1996). However, plasminogen has been detected as a secretory product from microglia conditioned media (Nakajima *et al.*, 1992a).
Indications also exist for a local expression of complement components within the central nervous system. The complement components Clq and C4 have been detected in microglia by in situ hybridization (Pasinetti et al., 1992) and Northern analysis has revealed the presence of C2, C3, C5 and C9 transcripts within human cortex (Johnson et al., 1992). The local expression of complement components suggests the brain can induce complement cascades in response to injury and infection. The complement system includes three members of the serine protease family, C2, C1r and C1s.

1.24 Summary

Serine proteases are implicated in the pathogenesis of Alzheimer's disease both in the aberrant processing of APP in the disease state and in the normal function of the precursor molecule. The identification of specific proteases interacting with APP may serve as potential therapeutic targets against this prevalent disease. Serine proteases also have modulatory roles in neuronal response to injury and evidence exists for the local expression of serine proteases involved in immunological defense within the brain.

D - The Structure and Function of Serine Proteases

Serine proteases are implicated extensively within the central nervous system and are hypothesized to play diverse roles in development, synaptic plasticity and neurodegeneration. Most serine proteases are synthesized as inactive precursors or zymogens that require proteolytic cleavage of a propeptide for activation of the catalytic domain. Typically this propeptide cleavage is achieved through the action of other serine proteases. As a result, many serine proteases work in cascades of consecutive zymogen activation, allowing a high degree of control and also signal amplification. Another important source of regulation is achieved through the action of specific inhibitors and regulatory co-factors.

The serine proteases are a class of proteolytic enzymes characterised by the presence of a uniquely reactive serine side chain. At present over 20 different families of serine proteases can be distinguished and these are grouped into six clans on the basis of similarities in three dimensional structures and catalytic mechanism (Rawlings and Barrett, 1994). The proteases of the chymotrypsin, subtilisin and carboxypeptidase C clans comprise the majority of serine protease characterised in eukaroytes. Within the chymotrypsin clan there are at least 10 different families of proteases of which the
chymotrypsin family represent the majority of serine proteases in animals. The term “chymotrypsin clan” or “chymotrypsin family” was simply adopted as convention because chymotrypsin was one of the earliest determined sequences. These terms do not therefore indicate that specific clan and family members have chymotryptic specificity. Within the chymotrypsin family are tryptic and elastolytic enzymes, along with chymotryptic proteases.

Three residues, serine, histidine and aspartic acid are conserved throughout the chymotrypsin clan and are essential to the catalytic mechanism (reviewed in Kraut, 1977). In general the catalytic residues are labeled according to the sequence of chymotrypsin (Hartley, 1964). Upon binding, the scissile bond of the substrate is fixed in a specific orientation by hydrogen bonds with the enzyme backbone within the so called oxyanion binding site. These interactions distort the scissile bond towards a tetrahedral geometry expected in the transition state. The proposed catalytic mechanism is summarised in figure 1.4.

Aspartic acid-194 is considered important for zymogen activation. Cleavage of the proenzyme activation peptide generally releases a hydrophobic residue, isoleucine-16 in chymotrypsin. In the activated enzyme, the amino group of this residue interacts with aspartic acid-194 causing a conformational changes in the active site of the enzyme. The substrate binding pocket and the oxyanion binding site which are incomplete in the zymogen are completed enabling the catalytic mechanism.

Histidine-57, aspartic acid-102 and serine-195 are all essential for catalytic activity as determined by mutagenesis studies (reviewed in Laszlo, 1989). Other mutagenesis and structural studies have highlighted the importance of other enzyme residues in determining substrate specificity (Perona et al., 1995). Tryptic proteases are associated with aspartic acid at position 189 lying within the substrate binding pocket. This negatively charged residue accommodates the predominantly positively charged lysine or arginine residues of the preferred substrate. Chymotryptic enzymes cleave preferentially after large aromatic residues and are associated with a hydrophobic binding pocket with serine at position 189. Residues 216 and 226 also lie within the substrate binding pocket. Both tryptic and chymotryptic enzymes have glycine residues at these positions allowing for a deep binding cleft to accommodate the bulky substrates. Elastase generally cleaves after uncharged, non-aromatic side chains, especially alanine. Accordingly, the substrate binding pocket is occluded by large residues with valine and threonine at positions 216 and 226.
Figure 1.4 - Catalytic mechanism of serine proteases

a) The hydroxyl group of serine-195 is activated as a potent nucleophile by the nearby imidazole group of histidine-57 and nucleophilic attack of the peptide carbonyl bond results in the tetrahedral transition state. Intermediates are stabilized by electrostatic interaction provided by aspartic acid-102.

b) The tetrahedral transition complex eventually breaks down yielding an acylenzyme intermediate releasing the amino terminal group of the substrate. The acylenzyme intermediate is then hydrolyzed by essentially the reverse of the above by nucleophilic attack on the carbonyl by the second substrate, usually water.

c) This releases the carboxy-terminal group of the substrate and the free enzyme.
E - AIMS OF THE PROJECT

The characterisation of the spectrum of serine protease expressed within the central nervous system is of extreme importance in understanding the role of these enzymes in neuronal function. Besides thrombin and the plasminogen activators, few brain expressed serine proteases have been characterised. The hypothesized role in mediating structural reorganisation implies the presence of controlled proteolytic cascades yet no serine proteases have been implicated acting downstream of the plasminogen activators. The perturbed protease and inhibitor interactions underlying sporadic Alzheimer’s disease implicate an as yet unknown serine protease in the pathogenesis of this prevalent disease. Also the characterisation of protease receptors which are active in signal transduction indicates the potential for serine proteases to function as extracellular signaling molecules. Table 1.1 presents a summary of the serine proteases characterised in brain to date.

<table>
<thead>
<tr>
<th>Serine protease</th>
<th>Detection method in brain</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-PA</td>
<td>Northern/in situ hybridization</td>
</tr>
<tr>
<td>u-PA</td>
<td>In situ hybridization</td>
</tr>
<tr>
<td>Plasminogen</td>
<td>In situ hybridization</td>
</tr>
<tr>
<td>Neuropsin</td>
<td>Northern/in situ hybridization</td>
</tr>
<tr>
<td>Neurosin</td>
<td>Northern</td>
</tr>
<tr>
<td>MSP</td>
<td>RNA dot blot analysis</td>
</tr>
<tr>
<td>Neurotrypsin</td>
<td>in situ hybridisation</td>
</tr>
<tr>
<td>Thrombin</td>
<td>Northern, in situ hybridization</td>
</tr>
<tr>
<td>Neutrophil elastase</td>
<td>Immunologically detected in microglia culture media</td>
</tr>
<tr>
<td>Complement proteases</td>
<td>polyA⁺ Northern</td>
</tr>
<tr>
<td>Acetyl cholinesterase</td>
<td>Activity found in brain</td>
</tr>
</tbody>
</table>

Table 1.1 - Summary of serine proteases previously characterised in brain

As discussed above, the hippocampus is the region of the brain particularly implicated in the consolidation of certain forms of memory. The dysfunction of this brain region is also an early feature of Alzheimer’s disease. Serine proteases expressed within the hippocampal formation are therefore of particular interest.
As well as the intrinsic interest of hippocampal serine protease, any protease restricted in its expression to this brain region would be of considerable application for transgenic experimentation. The pitfalls of gene knockouts and other conventional transgenic approaches to cognition have been outlined above and progress is now being made with a new generation of brain region specific transgenic manipulations. The characterisation of a hippocampus specific gene would allow exploitation of the control elements of such a gene for directed transgene expression within the hippocampus.

The aims of the project have been to characterise the diversity of serine proteases expressed within the hippocampus and to investigate the pattern of expression of identified proteases. Any serine protease restricted in expression to the hippocampal formation could then be used for transgenic exploitation to direct gene expression specifically within this brain region.
Chapter II

Identification of Hippocampal Serine Proteases

In an attempt to identify the spectrum of serine proteases expressed in the rat hippocampus, the technique of degenerate PCR has been employed to selectively amplify members of this multigene family. The following chapter outlines the techniques used to isolate a panel of 10 serine proteases, two of which represent putative novel genes.

2.1 Multigene families

The genomes of eukaryotes contain many groups of genes with sequence homology and overlapping functions. These so called multigene families are a result of successive rounds of gene duplication, presumed to be caused by unequal crossing-over at meiosis, transposition and other recombination events. The duplicated sequences amass random mutations in the course of evolution, leading to either pseudogenes where there is loss of gene function, or functional divergence upon which evolutionary selection can act. Initially, the overall function of the ancestral gene product is preserved but the exact specificity may be altered. The resulting members of a particular gene family will thus have common sequences, reflecting the common structural and functional properties of the protein, and specific sequences reflecting the individual properties of the respective protein.

Serine proteases are one of the most extensively studied gene families due to the abundance of sequence and structural information of numerous family members. An alignment of serine protease amino acid sequences allows various conserved regions to be identified (Figure 2.1). These generally surround the residues that are responsible for the structural integrity and proteolytic mechanism of this gene family. Specifically, each of the catalytic triad residues, serine, histidine and aspartic acid lie amidst highly conserved sequences.
Figure 2.1 - An alignment of serine protease amino acid sequences

Sequences are aligned from the putative amino terminus of the catalytically active enzyme. Conserved domains are highlighted in red and the conserved domains corresponding to the catalytic triad residue regions are highlighted in blue. Sequences are human heart chymase, human granzyme 3, mouse mast cell protease-2 (MMCP-2), mouse nerve growth factor gamma (NGF\textsubscript{\textgamma}), mouse plasmin, rat kallikrein, rat thrombin, rat trypsin 1, rat tissue plasminogen activator (t-PA), rat chymotrypsin B, rat elastase 1. Catalytic residues histidine-57, aspartic acid-102 and serine-195 are shown in bold.
IIGGTECKPHSRYPMAYLEIVSNPSGS--KFGCGFILRRNFWTLAACHC---------G
IIGGVEKSHPSRFPMASIQY----GGH----HVCQGVLIDPQWLTAAHQCQY----RFTKQ
IIGGVEAOPHSRSFPMAYLAKPTKTNXS--EFGGFPILPAQFWMTAAHCRQG--------
IVGGFKENSHQFPWVAAV--YRTQ-O--LFGQGVLIDPNWTLAAC----HCRQG--------
VGGCVANPHSWPQWISL----RTRFTQ-O--FCGPGLTIPAWNWTAAHCLC---------KSRP
VGGCVNPSHNSQPFWVAAV--YTFQ-O--LFCQGVLIDPSWNWTLAAC---------T
VEGWDAEEIPGIPQWVML---FRKSPQEL-LCGASILSDRNWTLAACHCILYPPDKNE
IVGGYCTHEPSHPYVSNLSGYH---CGGSLINDQWVVSAAHCY----------G
IKGFLITRTISPAAFGSRCPEGRF-LCGGVLISCWVLAAHCFV--------ERFP
IVGGEDAIJPISWNPQVSSL-QDEFT-GF-FHPCGGLTISDDIFTVTAACHG---------K
VGGQAERLEWSPSQWISL---YQLSSEGSYHTCGGTLIRRNFWTLAACHCVS---------Q

RSITVTGLGAHN---ITEEEDTWQKLEVIQFRHPKYNTSTLH---------HDIMLLKL
QSGPTVLGAHS---LSKNEASKQTELIKIFPFSRVTSDPSQSH---NDIMLLKL
SEISVPLGAHS---INKNEEPIQIILTKEFHPYFHKFHPR QLSYGFH---NDIMLLKL
DNYWKLVGKNN----LFKDEPSAQHFRVSPIAPDFNSLRRKHRFPLESYDNSLMRLKL
EFYVKLVGNN----AHEEYIRLGDQVE1-SVAKLLEPNNR---------DIALLKL
DNYQVWLGRNN----LYEDEPFPAQRHVLSFSFPHPQFLIQDLWNHTRQGPDDYSNLMRLKL
NDLLVRIGHGRSTRYNNH----VEKISMLEKITYHPRYNWR---------NLDRILLKL
SRIQVRLGNNH-----INVLREDQFIIAIAKIKPHNYSSH------WT---------LNDIMLKL
HNLKVLGQGNNH----QTBEEIKYIVHEKFEEFFDTYD---------DIAFLQKL
SD-VVQVAFDQG----GSQENIQVILKIAQVFKNPKFMFTVRN---------DITLLKL
MTFRVVGDBN---LQSDQGETQYSQVSMHTWNSNVA--------GYDIALRKL

KE-----KASLHTLVGTLPFPSQFN----FPVPPGMRCHRANWGR-----TVGLK-PGSDTLQER
QT-----AAKLNHVKMHLHRTSKE-L-RSGTKCKVZGGA---------TDPDSLPSDSTRLE
QK-----KEALNSDVFDSLPSSSD---PIFKPGMWTADGKQ--------TGKNN-PLSTLRE
SK-----KPAITDDVPTKLPFTEE-----KLGSTCLASGWSG--------TTTPKFQDLDLKC
---SRPQTIDDDVKKPACLPSNYY----MVADRTCITYIZGWSH--------TQ-GTFGAGRLKE
SQ-----SPAQDIDDVKQVIPFE----KVGSCLASGWSG--------TPTDGLDSDLQC
KK-----PPFKPSDHYHPVCLPDQTVSSLQAGYVRQGCVNLRETTWNTINEIQPSVLQV
SS-----PVXKNARVAPVLPSCAC------PAGTQLISGWN--------TLSQGNPPLLQC
RDSDSQCAQESSSVGTACLDPDVEQ-QLPDWTECELSGTYH--------HEASFPFSRDLKE
---AQPSAFQETTSVSAVLCPNNVD------DFPPGTVCATCGWGG------TYKYNALKTEPKQLK
-----AQSVTLLNYQVLAVLPQEGT------ILANNPICYTSGR---------TRTNQLS-OITLQK

VKRLMDPQACSHFRDFDHNLC------QLCVGHPRKT-----KAFKDGDSGPGPCICGAC-----
VTVTLRNSRLCNSQSYYNGDFFITKDVMCADARKQG------KDSCDKGDSGICLCKGAC-----
VELRMDQCAEKDHNDYDQL------QVACSPSTTL------KSIQGDGSGLPVCDCGAC-----
VNLKLNPDNEKAKHIEK------VTDAMLCAEMDGG------PWDCKDGDSGICLCKGAC-----
AILPQVNERVCNVRVEYLNRR-----VKSSTECQLAGG------VDCQDSGQDGSLVCFE--KD
VNLDDLSNEAVKEHAK------VTDMLCAEMDGG------KDTCGGDSGICLACGAC-----
VNLPIVERCAK-------STRIMTNDMCAFKGVNDTRGK----DACEDGGDGPFVKMPSYYNH
VDAPVLSQDACDAAEYYPGL------ITSMSCQFLEGG------KDCSGDGGDGPCVCNGAC-----
AHVRLLPSSQASQSVTAPCAIC-----QEGDSGGPFVMKSPYHAC-----KDTSCGQDGSLVCGAC-----
---AQPQAFSETPSVSACLCPNNVD------DFPPGTVCATCGWGG------TYKYNALKTEPKQLK
-----AQSVTLLNYQVLAVLPQEGT------ILANNPICYTSGR---------TRTNQLS-OITLQK

---VAQQITVSYGRS-----DAPKPAVFTRISH-YRPWINGWIL------QAN
---VFHAIISVSYEGD-GIVATKPGYTLTTTKYQTVWIKSNLVPHTN
---VAHITAAS-SY---CGAVKAPFVTRIS-YRPWLINK"VLY------KN
---VLOQTSHGHTC-GCEPKPMGYYKTNSKFTSWWIPDTM-----AKNP
KYLOQVTSGM-LGCAKPNKPGYTTRWS-RFVDWIEREM--------RNN
---VLOQTSHGKPN-CGEPKPMGGYKTNSKFTSWMKEMK------KENPC
RWYQMGIVSGKE--GCDRNYKGPTFYTH-FRLKRNWKQV------DQHR
---QLQIGVSGYGH---CALPDNPMPGKTVF-NVFQWIQDFT------AAN
RMTLIGQIS---LGCGQKDVPGTYKT-CYLYNLWIQDNNM------KQ
VWTVAGVSGKE--SGVCSTSTAPVSVRSL-ALMPWVPQII------EAN
QYSVHGYTSFVSMGGCNVSSKPTFTVTRPST-AIYSWNNVII------AYTC

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase

Chymase
Granzyme
MMCP2
NGFy
Kallikrein
Plasmin
Thrombin
Trypsin
tPA
Chymotrypsin
Elastase
2.2 Homology probing and degenerate PCR

The overall homology that exists between different members of a gene family can be exploited to isolate novel family members. With sequence from one gene, cDNA or genomic libraries can be screened at low stringency to obtain related genes. This strategy fails where homology is significantly lower than approximately 70%.

An alternative approach is the use of homology probing to identify novel family members. Conserved amino acid sequences in a particular family can be back-translated and synthetic oligonucleotide probes designed. The degenerate nature of the genetic code allows several differing nucleotide sequences to encode the same amino acid sequence. Mixed base oligonucleotide synthesis thus can be used to create oligonucleotide probes that will hybridize to any sequence encoding a specific amino acid sequence. This approach has been successfully used to screen cDNA libraries for novel members of the protein serine/threonine kinase multigene family (Hanks, 1987). However, the degeneracy of the labelled oligonucleotide probe restricts the actual amount of labelled relevant oligonucleotide hybridizing and this amount may be too small to provide a useful signal to noise ratio (Lathe et al., 1985).

To overcome the limitations of degenerate oligonucleotide hybridization, the Polymerase Chain Reaction (PCR) can be utilised in homology probing (Gould et al., 1989). Using two regions of conserved amino acid sequence, degenerate oligonucleotide primers can be designed by back translation and used to amplify family members from a variety of DNA sources. In general, much higher degeneracies can be employed as PCR tolerates limited mismatches between primer and template. Degeneracies greater than $10^6$ have been successfully used with this technique (Gonzalez and Chan, 1989).

2.3 Considerations for degenerate primer design

A balance in the extent of degeneracy is generally advantageous as too much degeneracy can lead to the amplification of random gene sequences. However, too little degeneracy will create bias towards the selective amplification of only certain family members. Furthermore, a lower degeneracy leads to a smaller primer pool with each individual oligonucleotide being more highly represented. Amplification with smaller primer pools is more efficient and titration of primer concentrations is often not
necessary to obtain optimum results. Restricting the degeneracy is thus considered important for the design of degenerate primers.

Degeneracy can be reduced through the use of homology regions containing amino acids encoded by only one or two possible codons for example methionine and tryptophan. Accordingly, amino acids encoded by up to six possible codons, for example leucine, serine and arginine are best avoided. The use of nucleotide analogues such as inosine which is able to pair with all four traditional nucleotides, allows the degeneracy of the primers to be reduced further. Inosine is best used at positions of three to four-fold base redundancy (Knoth et al., 1988).

Consideration of the codon usage of a particular organism can help reduce the degeneracy of the primer pair. Certain codons are rarely used in comparison to others and thus the oligonucleotide design can reflect this inherent bias. For example, in rat the codon GCG, encoding alanine, is only rarely used. A codon usage table of particular organism (Wada et al., 1990) can be used to reduce the degeneracy of a primer.

Within a specific gene family, codon usage can also be biased and primer design can be adjusted accordingly. An alignment of nucleotide sequences of various published serine protease sequences reveals examples of this bias (Figure 2.2). The catalytic histidine is encoded by CAC in greater than 95% of published sequences whereas in rat generally there is little preference between CAC and CAU (Wada et al., 1990).

2.4 Serine protease degenerate primers design

Taking into account the above considerations, degenerate sequences were designed to the conserved regions representing the catalytic histidine and serine residue regions of the serine protease family. Primers were then selected from these regions on the basis of the following criteria. The minimum annealing temperature of the primer pool was calculated by considering an A or T at every degenerate position and the primers with the highest minimal annealing temperature available were chosen. A high annealing temperature was preferred in order to allow stringent amplification in an attempt to minimalise non-specific priming. The primers were designed with no degenerate sequences at their 3’ ends as PCR primers require at least three homologous nucleotides at their 3’ ends for successful priming (Sommer and Tautz, 1989).
<table>
<thead>
<tr>
<th>Amino Acids</th>
<th>Val Leu Thr Ala Ala His Leu Met Ser</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chymase</td>
<td>GTG CTG ACG GCT GCT CAT TGT ...</td>
</tr>
<tr>
<td>Granzyne</td>
<td>GTG CTG ACA GCA GGC CAC TGC ...</td>
</tr>
<tr>
<td>MMCP2</td>
<td>GTG ATG ACT GCT GCA CAC TGC ...</td>
</tr>
<tr>
<td>NGFγ</td>
<td>GTT CTC ACA GCT GCC CAC TGC ...</td>
</tr>
<tr>
<td>Plasmin</td>
<td>GTT CTG ACT GCT GCC CAC TGC ...</td>
</tr>
<tr>
<td>Kallikrein</td>
<td>GTG ATC ACA GCT GCC CAC TGC ...</td>
</tr>
<tr>
<td>Thrombin</td>
<td>GTC CTC ACT GCT GCC CAC TGC ...</td>
</tr>
<tr>
<td>Trypsin</td>
<td>GTG CTG TCT GCA GCT GCC CAC TGC ...</td>
</tr>
<tr>
<td>t-PA</td>
<td>GTG CTA TCT GCC GCC CAC TGC ...</td>
</tr>
<tr>
<td>Chymotrypsin</td>
<td>GTG GTC ACT GCT GCC CAC TGC ...</td>
</tr>
<tr>
<td>Elastase</td>
<td>GTG ATG ACC GCT GCC CAC TGC ...</td>
</tr>
<tr>
<td>Degenerate</td>
<td>GTN CTN ACN GCN GCN CAT TGC ...</td>
</tr>
<tr>
<td>Code</td>
<td>T TG C T</td>
</tr>
<tr>
<td>Primer selected</td>
<td>GTG CTC ACA GCA GCA CAC TG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acids</th>
<th>Gly Asp Ser Gly Gly Pro Leu ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chymase</td>
<td>GGA GAC TCT GGG GGC CCT CTT ...</td>
</tr>
<tr>
<td>Granzyne</td>
<td>GGT GAC TCA GGG GGC CCC TTT ...</td>
</tr>
<tr>
<td>MMCP2</td>
<td>GGA GAT TCT GGG GGA CCT TTA ...</td>
</tr>
<tr>
<td>NGFγ</td>
<td>GGT GAC TCA GGA GGC CCA CTG ...</td>
</tr>
<tr>
<td>Plasmin</td>
<td>GGC GAC AGT GGA GGA CCC CTC ...</td>
</tr>
<tr>
<td>Kallikrein</td>
<td>GGT GAC TCA GGA GGC CCC CTC ...</td>
</tr>
<tr>
<td>Thrombin</td>
<td>GGT GAC AGT GGG GGA CCT TTT ...</td>
</tr>
<tr>
<td>Trypsin</td>
<td>GGT GAC TCT GGT GGC CCC GTG ...</td>
</tr>
<tr>
<td>t-PA</td>
<td>GGT GAC TCA GGA GGC CCT CTG ...</td>
</tr>
<tr>
<td>Chymotrypsin</td>
<td>GGT GAC TCC GGT GGC CCC CTC ...</td>
</tr>
<tr>
<td>Elastase I</td>
<td>GGT GAT TCT GGA GGC CCC CTC ...</td>
</tr>
<tr>
<td>Degenerate</td>
<td>GGN GAT TCN GGN GGN GCN CTN ...</td>
</tr>
<tr>
<td>code</td>
<td>C AG T</td>
</tr>
<tr>
<td>Primer selected</td>
<td>GAC TCT GGG GGG CCG CTT</td>
</tr>
</tbody>
</table>

Figure 2.2 - Alignment of nucleotide sequence for the catalytic regions

a) Alignment of the catalytic histidine region.
b) Alignment of the catalytic serine region.
Representative serine proteases are as described in figure 2.1
The selected primers (Figure 2.3) were designed to include restriction sites for EcoRI and XbaI at their 5' ends to aid in future DNA manipulations. This primer pair amplify over several introns so amplification products from contaminating genomic DNA can be distinguished.

\[
\text{EcoRI} \quad \begin{array}{c} \text{CGGAATTCC-GTGCTCACAGCAGCACACTG} \\
\text{GT} & \text{T} & \text{C} & \text{C} \\
\end{array}
\]

\[
\text{VLTAAH} \quad \text{DIALLL} \quad \text{GDSGGPLV}
\]

\[
\text{CTGAGACCTCCIGGIGAA-GCTCTAGACG} \\
\text{T} & \text{G} & \text{T} \\
\text{G} & \text{C} & \text{C} \\
\text{XbaI}
\]

Figure 2.3 - Degenerate PCR primers

The upstream primer was designed to the catalytic histidine residue region and was tailed with an EcoRI restriction site. The downstream primer was designed to the catalytic serine residue region and was tailed with an XbaI restriction site. The catalytic aspartic acid residue region is also presented. The catalytic residues, histidine-57, aspartic acid-102 and serine-195 are shown in bold.

2.5 Amplification of serine proteases

Rat hippocampi were dissected from freshly isolated brain and RNA was prepared. Poly A+ RNA was purified and reverse transcribed using a poly-dT primer to create 1st strand cDNA. PCR was performed for 30 cycles on this material using the above two degenerate primers under the two following conditions.

<table>
<thead>
<tr>
<th></th>
<th>High Stringency</th>
<th>Low Stringency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>93°C 1 min</td>
<td>93°C 1 min</td>
</tr>
<tr>
<td>53°C 1 min</td>
<td>47°C 1 min</td>
<td></td>
</tr>
<tr>
<td>72°C 30 secs</td>
<td>72°C 30 secs</td>
<td></td>
</tr>
</tbody>
</table>

The resulting products were separated electrophoretically on a 1% agarose gel (Figure 2.4).
Figure 2.4 - Degenerate PCR amplification products

Ethidium bromide stained agarose gel of degenerate PCR amplification products from high stringency conditions (HS), low stringency conditions (LS). Size standards (kb) are a 1 kb ladder.

The amplification conditions produced a clean band of the expected 500 bp size. There appeared to be little difference between the amplification products from the two annealing conditions. The lower anneal temperature was employed to aid in the amplification of less homologous products.

The 500 bp bands from both reactions were extracted from the agarose and were digested with *EcoRI* and *XbaI* and subsequently cloned into a similarly cut pBluescript vector. An alternative cloning strategy was employed which took advantage of the frequent 3' adenosine extensions of PCR amplification products to ligate the untreated product into the pCR2 vector (Invitrogen), a commercially available vector prepared with thymidine extensions. Recombinants from both ligations were sequenced from the polylinker T7, T3 and SP6 primer sites. The entire PCR product, being about 500 bp in length, could generally be sequenced entirely using these two primer site. No subcloning was necessary to obtain full sequence information for any of the clones.
The resulting sequence information was compared to published sequences in the GenEMBL database using a BLAST search (Altshul et al., 1990). Initially, 17 recombinants were sequenced and all were found to contain amplification products from genuine serine proteases, confirming the fidelity of the approach. 12/17 recombinants were identified as the serine protease tissue plasminogen activator (t-PA) and 4/17 were identified as the serine protease rat natural killer metase (RNK-Met-1). Both sequences obtained were 100% homologous to the published rat sequences except for differences at the 5' and 3' end of the product corresponding to the primer annealing sites.

2.6 Enriching for new serine proteases

These initial results indicated that the primer pair used amplified t-PA and RNK-Met-1 with highest frequency. To enrich for other hippocampal serine proteases a screening protocol was developed.

Figure 2.5 - Colony hybridization with t-PA and RNK-Met-1 probes

Putative recombinants were transferred to nylon membrane and probed with radiolabelled t-PA and RNK-Met-1. Non hybridizing colonies were then investigated further.
Recombinants were gridded out, transferred onto nylon filters, and probed with a mixed probe corresponding to the previously amplified t-PA and RNK-Met-1 sequences. Any clones not hybridizing were assumed to be novel family members and were investigated further. An example of such a hybridization is shown in Figure 2.5.

In all, approximately 1100 recombinants were screened in this way. Approximately 80% of all colonies obtained hybridized to the tPA/RNK-Met-1 mixed probe. Of the non-hybridizing clones, about 50% contained inserts which were then sequenced and the information compared with known sequences in the GenEMBL database using a BLAST search. This represents a total cloning efficiency of about 90%. In total 105 non-hybridizing clones were sequenced and five different sequences were obtained.

Sequences were obtained which corresponded exactly to the published rat serine protease sequences elastase IV and chymotrypsin B. Another three sequences revealed high homology at the nucleotide level to published serine proteases of other species. One sequence was 88% homologous to the human serine protease chymotrypsin-related protease (CTRL). Another was 67% homologous to the human serine protease, proteinase 3 and 93% homologous to the mouse orthologue of proteinase 3. One showed high homology (86%) to a mouse serine protease, complement protein 2 (C2). These sequences were translated and compared to the published protein sequences for these serine proteases. Figure 2.6 presents the protein sequence alignments of the amplified rat sequence and the published sequence in each case. The high percentage homology argues that these sequences represent rat orthologues of the respective proteases in human and mouse.

C2 sequence was truncated in relation to the other amplified proteases. Examination of the primer binding sites reveals that the amplification of this species had occurred with only the forward primer. The catalytic histidine residue motif provided one primer anneal site as expected and a homologous sequence elsewhere in the cDNA provided the reverse primer binding site. The CTRL sequence obtained was also truncated due to the presence of an internal EcoRI site, cleaved during cloning.

2.7 Two novel serine protease sequences

Two amplified sequences showed low homology (<40%) with a variety of published serine proteases. Translation of the amplified sequences revealed the presence of all the conserved domains of the serine protease multigene family (Figure 2.7).
Figure 2.6 - Sequence alignments of Proteinase-3, CTRL and C2

a) An alignment of human, mouse and putative rat proteinase-3 sequences.
b) An alignment of human and putative rat CTRL sequences.
c) An alignment of mouse and putative rat C2 sequences.
In each case, amino acid differences in the non-rat sequences are boxed.
Specifically the regions surrounding the catalytic triad residues are intact and highly conserved, while the spacing of cysteine residues is also particularly well conserved. The latter are considered to be important for the tertiary structure of the active enzyme through the formation of disulphide bridges. These two sequences were thus considered to be putative novel serine protease sequences as they have all the relevant sequence motifs yet no high homology with any known serine protease. They were dubbed Brain Serine Proteases 1 and 2 (BSP-1 and BSP-2). Both encode a characteristic aspartic acid residue six amino acid N terminal to the reactive serine, known as position S1, or 189 in the chymotrypsin numbering system (Hartley et al., 1964). This residue lies at the base of the substrate binding pocket and is indicative of trypsic specificity. It is thought that this S1 aspartic acid neutralises the charge on the arginine or leucine residue of the preferred substrates of trypsin-like serine proteases.

```
V L S A A H C K K D K - - - - - - Y S V R L G D H S L Q K R BSP-1
V V S A A H C Y K S R - - - - - - I Q V R L G E H N I N V L Trypsin
```

Figure 2.7 - An alignment of BSP-1, BSP-2 and trypsin

The translated sequences of the two putative novel serine proteases, BSP-1 and BSP-2 are aligned with the equivalent region from mouse trypsin 2. Regions conserved in serine proteases (see figure 2.1) are boxed and the aspartic acid residue, indicative of trypsic specificity, is indicated by the arrow.
The assumed identity of these clones along with the number of clones obtained in this experiment is shown in Table 2.1

To enrich further for previously unidentified serine proteases, recombinant clones were gridded out, transferred to nylon membrane and probed with a mixed probe representing all the serine protease sequences found to date. An example of such a hybridization is shown in Figure 2.8.

<table>
<thead>
<tr>
<th>Species amplified</th>
<th>No. of clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSP-2</td>
<td>39</td>
</tr>
<tr>
<td>BSP-1</td>
<td>8</td>
</tr>
<tr>
<td>Elastase IV</td>
<td>9</td>
</tr>
<tr>
<td>Proteinase 3</td>
<td>9</td>
</tr>
<tr>
<td>Complement 2</td>
<td>4</td>
</tr>
<tr>
<td>Chymotrypsin B</td>
<td>2</td>
</tr>
<tr>
<td>CTRL</td>
<td>1</td>
</tr>
<tr>
<td>Non specific products</td>
<td>33</td>
</tr>
</tbody>
</table>

Table 2.1 - Identity of serine proteases obtained from t-PA and RNK-Met-1 screen
Putative recombinants were transferred to nylon membrane and probed with radiolabelled t-PA, RNK-Met-1, BSP-1, BSP-2, elastase IV, CTRL, chymotrypsin B, proteinase 3 and complement C2. Non hybridizing colonies were then investigated further.

In total, some 600 clones were screened in this way and 12 clones containing inserts, but failing to hybridize, were obtained. These inserts were sequenced and this information was compared to published sequences in the GenEMBL database as before. All twelve sequences represented non-specific amplification products. This latter screen suggests that the vast majority of serine protease amplifiable with the degenerate primer pair used, had previously been isolated.

2.8 Summary of serine proteases amplified from rat hippocampus

The overall frequencies at which the various serine proteases were amplified from rat hippocampus is summarised in Table 2.2.

<table>
<thead>
<tr>
<th>Serine Protease</th>
<th>Frequency obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>tPA</td>
<td>63.0%</td>
</tr>
<tr>
<td>RNK-Met-1</td>
<td>26.0%</td>
</tr>
<tr>
<td>BSP-2</td>
<td>3.9%</td>
</tr>
<tr>
<td>Elastase IV</td>
<td>0.9%</td>
</tr>
<tr>
<td>Proteinase 3</td>
<td>0.9%</td>
</tr>
<tr>
<td>BSP-1</td>
<td>0.8%</td>
</tr>
<tr>
<td>C2</td>
<td>0.4%</td>
</tr>
<tr>
<td>CTRL</td>
<td>0.1%</td>
</tr>
<tr>
<td>Chymotrypsin B</td>
<td>0.2%</td>
</tr>
<tr>
<td>Non specific products</td>
<td>3.8%</td>
</tr>
</tbody>
</table>

Table 2.2 - Identity of serine proteases obtained

2.9 Serine proteases amplified from mouse hippocampus

The use of degenerate primers for homology probing of multigene families allows the same primer pairs to be used on other species with similar codon bias. Mouse and rat
have highly similar codon bias (Wada et al., 1990) and degenerate primer pairs for each organism should be interchangeable.

To investigate whether some of the products that were amplified from rat hippocampus are conserved across the species barrier, a similar amplification was performed on mouse hippocampus cDNA. Strain 129 mice were sacrificed, hippocampi were dissected and RNA prepared as before. Poly A+ RNA was selected and reverse transcribed using a poly-dT primer, to create 1st strand cDNA for amplification. PCR was carried out as before and again products of approximately 500 bp were cloned. Ten recombinants were sequenced and compared to the rat sequence obtained in the rat hippocampus screen. The results are summarised in Table 2.3.

<table>
<thead>
<tr>
<th>Sequence Identity</th>
<th>No. of Clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>tPA</td>
<td>4</td>
</tr>
<tr>
<td>RNK-Met-1</td>
<td>1</td>
</tr>
<tr>
<td>Elastase IV</td>
<td>3</td>
</tr>
<tr>
<td>Chymotrypsin B</td>
<td>1</td>
</tr>
<tr>
<td>Hageman Factor</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2.3 - Identity of serine protease amplified from mouse hippocampal cDNA

The amplification of similar products from both rat and mouse hippocampus highlights the reproducibility of the approach used. The isolation of Hageman Factor from mouse hippocampal cDNA represents a new product not found in the rat hippocampus screen. The sequence was found to be 100% homologous to the published mouse sequence for Hageman Factor.

2.10 An introduction to the serine proteases identified

The serine proteases identified in the degenerate PCR screen will be introduced first but will be discussed with reference to their expression patterns in Chapter VIII.

t-PA is an enzyme of tryptic specificity with a remarkably narrow substrate range (Madison et al., 1995). It has a well characterised role in fibrinolysis activating its physiological substrate, plasminogen to prevent inappropriate clot formation in the vascular system. The t-PA/plasminogen system also has important roles in extracellular matrix (ECM) degradation during cell migration and tissue remodelling.
during embryonic development, ovulation, inflammation and neoplasia (reviewed in Vassalli et al., 1991). t-PA is also implicated in the processing of growth factors (for example, Mars et al., 1993). Within the brain, t-PA expression is well characterised and is implicated in the development and remodelling of synaptic structures following synaptic plasticity and learning/memory events as discussed in chapter I. The amplification of this expected serine protease sequence validates the approach used.

RNK-Met-1 is a lymphocyte protease isolated from the granules of natural killer cells (Smyth et al., 1992). Its role is hypothesized to be in the degradation of ECM proteins to aid in the in vivo trafficking of natural killer cells and cytotoxic T-lymphocyte mediated cell lysis. The enzyme belongs to the elastase subfamily of serine proteases and preferentially cleaves after methionine residues. The initial characterisation of the enzyme indicated a lack of expression in brain tissue as demonstrated by Northern analysis on total mRNA, at odds with its PCR amplification from rat hippocampus in this study. PCR detection of transcripts is far more sensitive that Northern analysis so the discrepancy may reflect a low level of RNK-Met-1 mRNA in brain tissue. This issue is addressed directly in Chapter III.

Proteinase 3, also known as myeloblastin, was isolated by subtraction of cDNAs from pools of differentiated and undifferentiated promyelocyte-like leukemia cells from the cell line HL-60 (Bories et al., 1989). Proteinase 3 transcripts are down-regulated by monocytic and granulocytic inducers of HL-60 differentiation. Proteinase 3 is an elastolytic serine protease most homologous to neutrophil elastase (Lüdemann et al., 1990) and represents the third protease of the azurophil granules of polynuclear lymphocytes. More recently, proteinase 3 was identified as an antigen in the autoimmune disorder Wegener’s granulomatosis, a systemic vasculitis complicated by myelo-monocyte proliferation (Jenne et al., 1990). Expression in brain has not previously been reported.

Elastase IV is a serine protease expressed at high levels in the pancreas and acts with other elastolytic enzymes in digestion (Kang et al., 1992). An initial study of the expression pattern by Northern analysis of total RNA revealed no detectable transcripts in brain. The detection of this species in this study by PCR would argue for a low level of expression in rat hippocampus.

Chymotrypsin B (Bell et al., 1984), another pancreatic enzyme acting in digestion was amplified in this study along with a chymotrypsin-like protease (CTRL) which was
originally isolated from the sequencing of a tight gene cluster on chromosome 16 (Larsen et al., 1993). Both of these serine proteases have not been previously described in brain.

Complement 2 (C2) is a serine protease involved in the complement cascade which leads to lysis of cellular antigens. Several proteases are involved in this cascade which regulates phagocytosis, chemotaxis of inflammatory cells and cytotoxic effects of the membrane attack complex. Local synthesis of some components of the complement system has previously been reported in brain, including the presence of C2 transcripts by Northern analysis on human cortical poly A+ RNA (Johnson et al., 1992). In Alzheimer brain, complement components have been immunologically localised to neuritic plaques, tangles and dystrophic neurons (McGeer et al., 1989); an upregulation of complement components at the mRNA level has been detected in the disease state (Johnson et al., 1996). This Alzheimer complement induction may be related to the variety of functions of the complement system in peripheral tissues such as enhanced phagocytosis and cytotoxicity. However, it is of note that normal brain expresses complement components and these may have a role independent of those functions normally associated with cytolysis and inflammatory response.

Hageman factor, also known as blood clotting factor XII, is a serine protease involved in the intrinsic pathway of blood coagulation, fibrinolysis and complement activation. The presence of Hageman factor transcripts in brain was previously demonstrated by reverse-transcription PCR (RT-PCR) analysis (Yasuhara et al., 1994)

2.11 Brain serine proteases not identified in this study

There are certain serine proteases which have previously been shown to be expressed in the central nervous system before but were not identified in the degenerate PCR screen. This could be a failure of amplification due to the limited primer degeneracy (see section 2.12) or it could represent a failure to clone the products. These non-identified proteases are outlined below and a RT-PCR expression study is described in chapter III to confirm their expression in brain.

Thrombin, a serine protease of the blood coagulation system, is mitogenic for astrocytes and contributes to the regulation of neurite outgrowth and astrocyte stellation. RT-PCR analysis and in situ hybridization studies show the presence of prothrombin transcripts in most brain regions including the hippocampus (Dihanich et
al., 1991). The failure to amplify this serine protease sequence in this study is of concern as the degenerate primer pairs used should recognize the corresponding motifs in the prothrombin sequence (see Figure 2.2).

Urokinase-type plasminogen activator (u-PA) transcripts are present in the hippocampus, cerebellum and cerebral cortex, as shown by in situ hybridization studies (Masos and Miskin, 1996; Dent et al., 1997). However, other studies have failed to detect u-PA transcripts by this sensitive technique (Sappino et al., 1993; Kristensen et al., 1991). Irrespective of the exact level of expression, however, the primer pair used in this study would be expected to recognize this species.

The presence of both plasminogen activators in rodent brain suggests the local synthesis of plasminogen, the only physiological substrate of these enzymes. Plasminogen is too large a molecule to cross the blood brain barrier without a specialised transport system. Plasminogen transcripts have been localised to the murine hippocampus by in situ hybridization studies although at a low level of expression (Tsirka et al., 1997). This low expression level contrasts with the high expression level of t-PA. This discrepancy could imply another substrate for the plasminogen activators in brain (see Chapter III).

2.12 Discussion of degenerate PCR amplification

In all, a panel of 10 different serine protease sequences have been amplified from rat hippocampus. Two of the proteases, t-PA and RNK-Met-1, represent almost 90% of the amplified sequences. At this stage it is tempting to speculate on the relative abundances of the sequences in rat hippocampus but in general the assumption that the frequency of amplification reflects the abundance of endogenous transcript is false. In a degenerate PCR, many differing primers are present which anneal to their target sequences with variable efficiency and at varying temperatures leading to variable amplification efficiencies. Furthermore the amplification of a particular species may be inhibited by dimerisation of a portion of the primers. With the design of primers used in these experiments, where the degeneracy of the primers has been reduced, there will be an inherent bias towards certain sequences. As stated previously, in the design of degenerate primers there is a balance between restricting non-specific amplification and causing a strong bias in the variety of species amplified.
Interestingly, the most abundant sequence identified was t-PA. Although few brain serine proteases have been directly assessed for expression level, Northern analysis and in situ analysis of t-PA transcripts in brain suggests a relatively high level of expression (Sappino et al., 1993; Ware et al., 1995). Thus it appears that the most readily identified sequence is likely to be the most abundant serine protease species in the rodent hippocampus. The results of experiments to address the individual expression levels of the proteases are reported in the following chapter.

The level of degeneracy of the primers in this case was quite low in comparison to other studies. A higher level of degeneracy would allow a less biased selection of family members to be amplified, although this would be accompanied by an increase in non-specific amplification products. In practice however, the primers have the ability to amplify a more diverse array of family members than the degeneracy dictates as PCR tolerates a certain degree of mismatch. Furthermore, under certain conditions T-U and T-G pairing will occur allowing a greater degree of mismatched priming (Martin, 1985). The amplification of certain previously published sequences in this study allows the degree of mismatching to be assessed. Figure 2.9 shows the published sequences of four of the rat sequences that were amplified in this study along with the primer sequences revealed in the cloned and sequenced products. In general there are between 1 and 5 mismatches per primer which indicates that the total number of sequences available to the primer pair used was significantly more than that predicted from the degeneracy. It is interesting to note that several different isolates of the same product were amplified using different degenerate primers highlighting the ability of mismatched primers to amplify. The primer anneal sites "evolve" during successive amplification cycles as, with a degenerate primer pool, different primers will bind to the newly generated products. This means that the actual primer pool available to an individual product is considerably larger.

The elastase IV product is of interest as one of the mismatches is in the last three 3’ nucleotides of the 3’ primer. In general, the last three nucleotides of any degenerate primer are required to be homologous for efficient amplification. The nucleotides of the 3’ primer represents the aspartic acid residue encoded by the codons, GAT and GAC. The GAC codon is used in over 75% of serine proteases (see figure 2.2) so this was chosen for the degenerate primer. This bias in the primer design would, in theory, restrict 25% of serine protease sequences being amplified. The amplification of a species using GAT in this position indicates that this mismatch is tolerated under the amplification conditions used, so in practice this theoretical bias is overcome.
Figure 2.9 - Primer anneal sites for the previously characterised rat sequences

The sequences shown are previously published rat sequences and the hybridizing sequences represented above and below are the upstream and downstream primer sequences obtained from the cloned amplification products. Sequence identity at a particular position is shown as a solid line or as a dotted line in cases where multiple sequences were obtained.
2.13 Improvements to the techniques used

As described above, increasing the degeneracy results in a concomitant increase in the variety of amplifiable sequences but also the amount of non-specific products. These artefacts could be limited through the use of nested PCR. The presence of a highly conserved region around the catalytic aspartic acid residue in the serine protease family could serve as a useful template for a third degenerate PCR product. In this way two rounds of PCR could be carried out, initially with highly degenerate primers built to the serine and histidine catalytic regions and subsequently with a primer pair directed against the aspartic acid catalytic region and one of the other regions. This would allow enrichment of only genuine serine protease sequences as non-specific amplification would be reduced through the use of three annealing positions.

The use of restriction enzymes built into the 5' ends of the individual primers facilitates cloning as resulting products can be digested and ligated to suitably prepared vectors. This double restriction site requirement for cloning implies that any non-specific products which have been amplified with only one of the primers will not be able to be cloned as two differing restriction sites are necessary. This double enzyme method of PCR product cloning enriches the proportion of genuine amplification products. The use of a TA vector system allows any amplification product to be cloned thus the use of this system, although greatly facilitating the cloning process, increases the amount of clonable single primer amplification artefacts. For a higher cloning rate of genuine amplification products the restriction site method of product-vector ligation is recommended. However, in this study, C2 was amplified using the forward primer although such an event is considered exceptional.

Many sequences identified in the screen have previously been characterised in various vascular components. For example, the amplification of the lymphocyte proteases, RNK-Met-1 and proteinase 3 could have been achieved from contaminating lymphocytes in the extensive vascular system of the brain. However, there are cell types within the CNS which are in common ontological lineages with various immune system components, for example microglia represent the resident brain macrophages. It is within reason therefore, that certain CNS cells may share expressed sequences with other members of the macrophage lineage. In situ hybridization analysis (see chapter III), determining the site of expression, allows the question of blood contamination versus local synthesis to be investigated. In any case, the perfusion of
brain samples with phosphate buffer saline (PBS) prior to the sacrifice should be recommended to reduce the amount of blood contaminants in the tissue.

2.14 Summary

PCR was used to investigate the spectrum of serine proteases expressed in the hippocampus. In total, a panel of 10 serine proteases have been identified including two putative novel serine proteases and some proteases not previously reported to be expressed in brain tissue. Two novel serine proteases were identified and have characteristic aspartic acid residues in the substrate binding pocket and are thus predicted to be of tryptic specificity. t-PA and RNK-Met-1 were the two proteases identified with highest frequency perhaps reflecting the abundance of these transcripts in the rat hippocampus.
Expression Analysis of Serine Proteases

Northern and \textit{in situ} hybridization analysis was used to elucidate the pattern of expression of the serine protease sequences identified in the rat hippocampus. The expression of other previously described brain serine proteases was also investigated by reverse-transcription PCR analysis.

3.1 The search for hippocampus specific genes

There have been numerous studies applying transgenic techniques to the behavioural neurosciences in an attempt to discern the role of hippocampal processes in learning and memory. The technique of gene targeting allows components of the synaptic signalling cascade to be deleted or modified at the level of the gene and the resulting animal can be assessed for behavioural, biochemical and electrophysiological deficits. As discussed in the introduction there are problems with such techniques associated with the pleiotropic effects of the gene modification. In general, previous transgenic studies have focused on receptor subtypes and components of the signalling cascade which are expressed ubiquitously throughout the brain. It is therefore hard to discern whether any behavioural phenotype seen is due solely to deficits in hippocampal function.

A gene restricted in expression to the hippocampal formation would be of considerable utility in the development of a new generation of transgenic animals addressing the role of this brain region in learning and memory processes. The endogenous promoter region of such a gene could be exploited to drive expression of transgenes, causing a perturbation of synaptic processes specifically within the hippocampus.

Beside the technological application of a hippocampus specific gene, any gene restricted in expression to a particular anatomically specialised region is of general interest with respect to the functional specialisation of the brain region. As outlined in the introduction, morphological changes within the hippocampus accompany learning and memory events and similar changes are seen following the induction of synaptic
plasticity. Any serine protease expressed specifically within the hippocampus would therefore be of considerable interest as a potential mediator of these structural changes.

3.2 Northern analysis of serine protease expression

Total RNA was prepared from various rat organs and dissected brain regions using a guanidinium thiocyanate/acid-phenol method. 20 μg of each sample was run on a 1% agarose gel in the presence of 18% formaldehyde and transferred onto nylon membranes by capillary blotting. The serine protease sequences obtained in the previous chapter were labelled by random priming and used as probes on these blots. After autoradiography, the blots were stripped and reprobed with a loading control corresponding to S26 ribosomal protein which is expressed at equal levels in all cell types (Vincent et al., 1993).

The t-PA probe hybridized to a single species at about 2.4 kb which is consistent with the published cDNA length of 2.453 kb (figure 3.1a). Strong expression was seen in all brain regions analysed. These results agree with previous Northern and in situ hybridization studies which show t-PA expression in the cerebellar cortical layers, the hippocampus, the amygdala, the olfactory bulbs and the cortical mantle (Sappino et al., 1993; Ware et al., 1995). Another in situ study, revealed a much more limited pattern of expression for t-PA in rat brain, restricted to the ventricular ependyma and the olfactory nerve layer (Dent et al., 1993). Although a detailed expression study was not performed in this study, the results of Northern analysis suggests mouse and rat expression patterns to be more similar than different. The differences in expression pattern obtained in various studies may also be attributable to the activity dependence of the gene, which is dynamically regulated in response to a variety of stimuli. Different expression patterns could then reflect different levels of activity and arousal in the different mice populations used. The lack of consistency may also reflect technical difference in signal detection, different probe regions or hybridization conditions used.

The RNK-Met-1 probe hybridized to a single species at about 0.9 kb which approximates to the published cDNA length (figure 3.1b). Strong expression was seen in the spleen with modest expression in the lung and hippocampus. Weak expression was seen in other brain regions. Previous expression analysis on RNK-Met-1 failed to detect transcripts in freshly isolated brain and spleen by Northern blot; the report did not investigate lung expression (Smyth et al., 1992). Hybridizing
Figure 3.1 - Northern analysis of t-PA and RNK-Met-1 expression

a) Multi-tissue Northern blot probed with t-PA (upper panel) and S26 loading control (lower panel).

b) Multi-tissue Northern blot probed with RNK-Met-1 (upper panel) and S26 loading control (lower panel)

Tissues analysed were: Hi, hippocampus; Ce, cerebellum; Cx, cortex and striatum; RB, remainder of brain; He, heart; Li, liver; Sp, spleen; Ki, kidney; Te, testis, Mu, skeletal muscle.
transcripts were only detected with total cellular RNA from a rat natural killer cell leukaemia or from purified adherent lymphocyte-activated killer splenocytes. The discrepancy in this study may be explained by the level of detection sensitivity. Figure 3.1b represents a 14 day exposure of the hybridized Northern blot whereas the original report concluded from a 2 day exposure. The blot is this study is therefore up to 7 times more sensitive assuming the specific activity of the probe and the total counts added were constant between the studies.

Northern analysis on total RNA was also performed using probes corresponding to elastase IV, proteinase 3, CTRL, chymotrypsin B and C2. No detectable signal was generated by any tissue including brain (not presented). As these transcripts were initially identified by a PCR technique they may represent weakly expressed transcripts. A more sensitive method of transcript detection will be required to confirm the genuine brain expression of these serine proteases (see later sections).

The novel serine proteases were also hybridized to similarly prepared Northern blots. BSP-2 robustly identified a 1.3 kb transcript in testis and brain, specifically within the hippocampus, cortex and striatum (figure 3.2a). Low level hybridization was detected elsewhere in the brain as shown by the lack of signal in the remainder of brain (R.O.B.) lane. Very low level expression was detectable in all other tissue types.

BSP-1 hybridized to a 1.3 kb transcript only within the hippocampus (figure 3.2b). No other signals in any other tissue samples were discernible following a long exposure times (up to 4 weeks). At the level of Northern analysis BSP-1 thus appeared to be confined to the hippocampus in its expression.

3.3 Relative abundance of serine proteases

Assuming a constant level of RNA transfer to the nylon membrane, these Northern blots can be compared to give an indication of the relative expression levels. t-PA, in agreement with previous studies (see section 3.2), is probably the most abundant: a strong signal is easily detected by Northern blot after a few days of autoradiography. This is in agreement with the inferred abundance of the transcript from the PCR results as this species was identified almost three times as frequently as others. A comparable RNK-Met-1 signal is seen after a two week exposure implying that the abundance of this transcript is significantly lower in comparison with t-PA. BSP-2 expression levels are similar to those of RNK-Met-1 as a comparable signal is obtained after two
Figure 3.2 - Northern analysis of BSP-2 and BSP-1 expression

a) Multi-tissue Northern blot probed with BSP-2 (upper panel) and S26 loading control (lower panel)
b) Multi-tissue Northern blot probed with BSP-1 (upper panel) and S26 loading control (lower panel)
Tissue analysed were: Br, total brain; Te, testis; Mu, muscle; Sp, spleen; Li, liver; Lu, lung; He, heart; Ce, cerebellum; Cx, cortex and striatum; Hi, hippocampus; RB, remainder of brain; Ki, kidney.
weeks of autoradiography. These two species represent the next most abundant serine proteases identified in the PCR screen.

The mostly frequently identified species after the above are elastase IV, proteinase 3 and BSP-1. No signal was detectable for the former two but a robust signal is seen with BSP-1 after four weeks autoradiography. No other transcripts were detectable for the other proteases using this technique.

Overall, there exists a correlation between expression level and the frequency of detection by degenerate PCR. This would be expected assuming no bias in the design of the primers; inherent bias might be expected because degeneracy of the primers has been reduced. The correlation observed is therefore an initial indication that the amplification was not strongly biased in favour of one particular species. The high frequency of amplification of t-PA therefore faithfully reflects the endogenous abundance of transcripts within the hippocampus.

3.4 In situ hybridization analysis of protease expression

The previous Northern analysis provides an initial indication of where transcripts are abundantly localised in the brain, but fails to identify the substructural expression pattern nor provides information regarding which cell types express the transcripts. For these purposes the technique of in situ hybridization is invaluable.

In situ analysis provides high sensitivity allowing low levels of expression to be assessed (Emson, 1993). In general two types of probe can be used: isotopically labelled oligonucleotide probes or single stranded ribonucleotide probes (riboprobes). Oligonucleotide probes are convenient; they require no cloning manipulations and they are easily designed to be specific to a particular gene family member. Riboprobes however have a greater sensitivity for detecting low abundance species and spurious hybridization is less common as these probes are larger. Riboprobes have been used in this study to maximize the sensitivity of detecting potentially weakly expressed transcripts.

All the serine protease sequences identified in the degenerate PCR vector were previously cloned into either pBluescript or pCR2. Both these vectors have RNA polymerase binding sites flanking the polylinker facilitating the synthesis of riboprobes in both sense and antisense orientations. 10 μm horizontal brain sections were
hybridized with both sense and antisense riboprobes transcribed using T7, T3 or SP6 RNA polymerases, as appropriate, in the presence of $^{35}$S UTP. After high stringency washing, sections were exposed against autoradiographic film for up to six days. The initial result on film allowed the approximate exposure time for emulsion autoradiography to be determined; one day exposure to film was deemed to be the equivalent of seven days exposure to emulsion. Initial results confirmed that, in general, expression of the serine proteases was weak and long exposure times were required for the detection of hybridizing transcripts. Film autoradiography allowed a relatively quick assessment of expression pattern with reasonable clarity but with poor cell type resolution. Only the more abundantly expressed serine proteases were exposed to emulsion autoradiography.

Anti-sense RNK-Met-1 hybridization was detected in various brain regions by film autoradiography and then slides were dipped in liquid emulsion and exposed for 12 weeks (figure 3.3). High levels of expression were seen in the dentate gyrus and the CA fields, with expression in the CA1 and CA3 regions significantly greater than in CA2 region. The localisation of silver grains over the dentate gyrus granule cells and the pyramidal cell layers of the CA fields suggest the expression of RNK-Met-1 to be neuronal in origin (figure 3.6). Expression was also seen in deep and superficial layers of the cortex including entorhinal cortex and cingulate cortex. Hybridization was also clearly detectable within the subfornical organ. No hybridization was detectable in the cerebellum, the olfactory bulbs, the thalamus, hindbrain and the striatum.

Anti-sense BSP-2 hybridization was detected by film autoradiogram and slides were then dipped and exposed for 12 weeks (figure 3.4). Hybridization was detectable in the CA fields of the hippocampus with stronger signals in CA3 and CA1 in comparison to CA2. Weak but significant hybridization was detectable in the dentate gyrus. Localisation of silver grains was clearly over the granule cells of the dentate gyrus and the pyramidal cells of the CA fields, indicating neuronal expression (figure 3.6). Strong hybridization was also seen in superficial layers of the cerebral cortex. Within the brain stem, distinct hybridization was detectable in the motor trigeminal nucleus, medullary regions and within certain reticular nuclei by comparison with a rat brain atlas (Paxinos and Watson, 1986). No hybridization was detectable in the cerebellum, the olfactory bulbs, the thalamus and the striatum.
Figure 3.3 - *In situ* hybridization study of RNK-Met-1 expression

a) Contact autoradiogram of horizontal brain section hybridized with anti-sense RNK-Met-1 riboprobe. *Hi*, hippocampus; *Cx*, cortex; *S*, subfornical organ.

b) Dark field photomicrograph of hippocampus and cortex hybridized with anti-sense RNK-Met-1 riboprobe. *DG*, dentate gyrus; *Cx*, cortex.

c) Higher magnification dark field photomicrograph of dentate gyrus and CA3 showing the expression in the polymorphic region. *DG*, dentate gyrus; *PoDG*, polymorphic cells of the dentate gyrus.

d) Low magnification dark field photomicrograph of horizontal brain section hybridized with sense RNK-Met-1 riboprobe.
Figure 3.4 - *In situ* hybridization study of BSP-2 expression

a) Contact autoradiogram of horizontal brain section hybridized with anti-sense BSP-2 riboprobe. *Hi*, hippocampus; *Cx*, cortex.
b) Dark field photomicrograph of hippocampus hybridized with anti-sense BSP-2 riboprobe. *DG*, dentate gyrus.
c) Contact autoradiogram of horizontal brain section, more ventral to a) hybridized with anti-sense BSP-2. *R*, reticular nucleus; *T*, motor trigeminal nucleus; *M*, medulla; *Hi*, hippocampus.
d) Low magnification dark field photomicrograph of horizontal brain section hybridized with sense BSP-2 riboprobe.
Anti-sense BSP-I hybridization was detected by film autoradiography specifically in the CA1 and CA3 fields of the hippocampus and also limited hybridization to the adjacent deep layers of the entorhinal cortex (figure 3.5). Slides were dipped and exposed to liquid emulsion for 12 weeks which revealed the predominant site of hybridization to be the pyramidal cell layers of the CA1 and CA3 fields (figure 3.6). Weak but significant hybridization was detected in the CA2 fields of the hippocampus and within the polymorphic cells of the dentate gyrus. No hybridization above background was discernible within the dentate gyrus granule cells or in any other brain regions examined.

Hybridization to anti-sense CTRL riboprobe was detected in all brain regions by film autoradiography. In general a low level of expression was seen. Distinct hybridization was clearly visible in the cerebellum and the dentate gyrus but this was considered only to reflect the high cell density of these regions (figure 3.7a,b). Compared with the hybridization detected using the BSP-1, BSP-2 and RNK-Met-I anti-sense probes, CTRL hybridization was at least three times as weak. Since emulsion autoradiography took up to 12 weeks to detect a clear signal for these more abundantly expressed transcripts, the exposure time necessary for CTRL was deemed to be unfeasibly long.

A C2 anti-sense riboprobe hybridized to all brain regions at low level (figure 3.7c,d). A relatively high level of expression was also detected in the olfactory bulbs and to the cerebral vascular walls. This vascular hybridization corresponds to the expected expression of C2 which occurs within the circulatory system. As with CTRL hybridization, levels were considered too low to justify emulsion autoradiography.

An elastase IV anti-sense riboprobe hybridized to cerebellum, olfactory bulbs and to the deep and superficial layers of the cerebral cortex, (figure 3.7e,f). Clear hybridization was seen to the CA1 field of the hippocampus although no hybridization was detected in the dentate gyrus or CA3. Hybridization was not detectable in the striatum and diencephalon. Again the expression levels of elastase IV was deemed too weak to allow emulsion autoradiography.

An anti-sense proteinase 3 riboprobe hybridized to most brain regions at very low levels. The signal was barely detectable after 6 weeks autoradiography but was clearly above background (data not shown)
Figure 3.5 - *In situ* hybridization study of BSP-1 expression

a) Contact autoradiogram of horizontal brain section hybridized with anti-sense BSP-1 riboprobe. *Hi*, hippocampus.

b) Dark field photomicrograph of hippocampus and adjacent cortical regions hybridized with anti-sense BSP-1 riboprobe. *DG*, dentate gyrus; *PoDG*, polymorphic cells of the dentate gyrus; *Ec*, entorhinal cortex.

c) High magnification dark field photomicrograph of hippocampus hybridized with anti-sense BSP-1 riboprobe. *DG*, dentate gyrus; *PoDG*, polymorphic cells of the dentate gyrus.

d) Low magnification dark field photomicrograph of horizontal brain section hybridized with sense BSP-1 riboprobe.
Figure 3.6 - *In situ* hybridization studies of RNK-Met-1, BSP-1 and BSP-2 in hippocampus.

High magnification view of silver grains, clustered over cells in various regions of the hippocampus. Brain sections were counterstained with neutral red.
Figure 3.7 - *In situ* hybridization studies of CTRL, C2 and elastase IV expression

All show contact autoradiograms of horizontal brain sections hybridized with anti-sense CTRL riboprobe (a and b), C2 riboprobe (c and d) and Elastase IV riboprobe (e and f).

*Hi*, hippocampus; *Ob*, olfactory bulbs; *Cb*, cerebellum; *Cx*, cortex.
Only very weak hybridization was detectable with the chymotrypsin B anti-sense riboprobe (data not shown). The sense control riboprobe reveals edge effect hybridization artefacts after prolonged autoradiography so convincing expression of this species within the brain has not been demonstrated. This could reflect particularly low levels of chymotrypsin B expression in brain.

For all the in situ carried out, a sense control riboprobe was synthesised and hybridized to adjacent sections. In all cases, apart from the above chymotrypsin B edge effect hybridization, no hybridization was seen by film autoradiography. The dipping of the sections revealed only background levels of hybridization as assessed by the density of silver grains over the brain sections matching the density of silver grains over a dipped and exposed blank slide.

In general the in situ hybridization data agrees with the data obtained through Northern analysis on dissected brain regions thus confirming the specificity of the approach. BSP-2 and RNK-Met-1 hybridized predominantly to the hippocampus and the cortex on Northern blots and in situ hybridization confirmed these brain regions to be the main sites of expression. Similarly, BSP-1 hybridized specifically to the hippocampal RNA on a Northern blot and in situ hybridization highlighted the hippocampal formation as the only significant site of expression in brain.

The expression level as determined by film autoradiography indicates RNK-Met-1, BSP-2 and BSP-1 to be the most abundant transcripts of those investigated by in situ hybridization. This is consistent with the inferred expression levels from Northern analysis. Other sequences amplified with similar frequency as BSP-1 in the original PCR are expressed far more weakly. Specifically, the expression levels of proteinase 3 and elastase IV are significantly weaker than that of BSP-1. Both elastase IV and proteinase 3 transcripts are undetectable by Northern blot and the expression of both specifies in brain can only be demonstrated convincingly by in situ hybridization.

3.5 Identification of a hippocampus specific transcript

The original degenerate PCR was performed on hippocampus cDNA in an attempt to isolate family members restricted to this brain region in their expression. Accordingly, the hippocampus is a site of expression for all the sequences tested by Northern analysis and in situ hybridization. Expression in other brain regions seems to be common among the array of expression profiles obtained. Three transcripts, RNK-
Met-1, elastase IV and BSP-2 are expressed in the hippocampus and in the cortex. Two transcripts, CTRL and C2 appear to be expressed fairly uniformly throughout all brain regions. BSP-1 transcripts however appear to be predominantly restricted to the hippocampus, specifically with the CA pyramidal cell layer. Some expression in the entorhinal cortex is also apparent but at a low level. Layer II and III entorhinal cortical neurons project into the hippocampus to synapse with the dentate gyrus granule cells and the CA pyramidal neurons. These cortical layers provide the majority of input into the hippocampal formation.

Other genes have been reported which are expressed specifically in the hippocampus as discussed in Chapter I. In most cases, these genes are expressed at low but significant levels in other brain regions and in other tissues. Also in cases where expression is restricted to the hippocampal formation, expression level is considered too low to be of practical use in directing transgene expression specifically to this brain region.

BSP-1 is thus a putative serine protease, restricted in its expression to regions of the hippocampal formation. Furthermore, the expression level is reasonably high, in that a clear signal is detectable by Northern analysis. This gene is thus a good candidate for transgenic exploitation as the endogenous transcriptional elements may, in principle, be used to direct transgene expression specifically within the hippocampal formation (see Chapter VI).

3.6 Reverse-transcription PCR analysis of brain protease expression

The previous chapter reported the identification of ten different serine protease sequences from rat hippocampus. Several other serine protease sequences, specifically thrombin, u-PA and plasminogen have been reported to be expressed in brain before but were not detected in the degenerate PCR screen. The lack of amplification of these species could reflect a bias in the primer design, although examination of the sequences involved indicates that the sequences would be expected to be recognized. It is also possible that these species were amplified and were refractory to cloning. To investigate the expression of these species in rat brain an RT-PCR analysis using specific primers was performed to confirm the previously reported expression.
RT-PCR is an extremely sensitive method of mRNA analysis which can be used to
give quantitative information concerning the levels of transcripts in a particular tissue.
Exact quantitative information is difficult to obtain as the technique involves two
sequential enzymatic steps which can proceed with variable efficiency. However,
comparative PCR in which two tissue samples are processed and analysed
simultaneously may permit comparison of relative expression levels. By varying the
number of cycles of PCR a proportionality between product yield and transcript level
can be achieved as the reaction tends to saturate at a high number of cycles. An
endogenous sequence, known to be present at constant levels throughout a series of
samples to be compared, can be used as an internal standard. Co-amplification of this
internal standard allows the integrity of the original RNA, the equal loading and the
efficiency of the reactions to be assessed directly.

In this study, the expression level of particular serine protease in brain has been
compared against a control tissue. Primers used in this study were designed to
amplify over an intron such that genomic DNA contamination products would not
confuse the analysis. Co-amplification of the hypoxanthine phopho-ribosyltransferase
(HPRT) transcript as an internal standard has been used in each case.

### 3.7 RT-PCR analysis of thrombin expression

Thrombin, a serine protease of the blood coagulation system, is produced by
maturation of a precursor polypeptide, prothrombin. Transcripts encoding
prothrombin are clearly present in brain as shown by in situ hybridization (Dihanich et
al., 1991). The major site of expression of prothrombin however is the liver. In
vitro, thrombin regulates process outgrowth in neurons and astrocyte stellation and
thrombin is thus hypothesized to play a role in structural plasticity, development and
response to injury. The lack of identification of prothrombin in the degenerate PCR
screen is therefore at odds with its well recognized role in the CNS.

Oligo-dT primed cDNA was prepared from 1 μg of rat liver and brain poly A+ RNA
and PCR was performed using thrombin and HPRT specific primers for 20, 25 and 30
cycles. The primers used were designed to amplify over an intron in both cases, such
that any amplification product from genomic DNA could be distinguished. Amplification
products were separated on a 1.5% agarose gel, blotted and probed with a
confirmed thrombin sequence (figure 3.8a).
The 650 bp HPRT amplification products are present at equal levels in both liver and brain samples confirming the equal efficiency of amplification and loading. The 400 bp prothrombin amplification product suggests that cognate transcripts are abundant in liver. The Southern blot of the PCR products revealed a low level of prothrombin transcripts present in brain, confirming local synthesis of this protease in the CNS. The presence of prothrombin transcripts in the starting material used for the degenerate PCR screen thus suggests that PCR amplification of this species was inefficient or that amplified thrombin sequences were for some reason unclonable.

3.8 RT-PCR analysis of u-PA expression

Plasminogen activators, secreted by a variety of cell types, have a role in fibrinolysis, cell migration and tissue remodelling and repair. They act by activating their substrate plasminogen to plasmin, a protein with a broad substrate range and capable of dissolving blood clots and degrading ECM components and basement membranes. Within the brain, plasminogen activator activity is present throughout development particularly in areas undergoing extensive remodelling (Sappino et al., 1993). The detection of PA activity in normal, mature brain suggests an important role for these enzymes in neuronal function. In rodents t-PA is the predominant plasminogen activator species as determined by in situ hybridization, Northern analysis and immunological and zymographical analysis in brain homogenates and sections. Reports of u-PA expression in adult brain are somewhat conflicting. Reports have concluded a lack of u-PA expression in mouse brain based on Northern analysis and the more sensitive in situ hybridization analysis (Kristensen et al., 1991; Sappino et al., 1993). Another more recent in situ hybridization study, however, has detected expression of u-PA in mouse brain (Masos and Miskin, 1996). Expression in the rat brain appears more widespread and transcripts have been detected in the hippocampus, olfactory bulb, cerebellum and cerebral cortex by in situ hybridization (Masos and Miskin, 1996; Dent et al., 1993).

Brain and kidney oligo-dT primed cDNA was synthesised and PCR was performed using u-PA specific primers and HPRT control primers. Amplification products from 20, 25 and 30 cycles were electrophoretically separated, Southern blotted and probed with a confirmed u-PA probe (figure 3.8b).

The 400 bp HPRT amplification product was equally represented in both brain and kidney products reflecting the equality of amplification efficiency and loading. The
Figure 3.8 - RT-PCR analysis of prothrombin and u-PA transcripts in rat brain

a) Ethidium bromide stained prothrombin amplification products for the number of cycles indicated are shown (top panel) and Southern hybridization of the products with a prothrombin probe are shown (bottom panel).
b) Ethidium bromide stained u-PA amplification products for the number of cycles indicated are shown (top panel) and Southern hybridization of the products with a u-PA probe are shown (bottom panel).
Samples are Br, brain; Li, liver; Ki, kidney. Numbers refer to number of PCR cycles performed.
500 bp u-PA amplification product was present at high levels in kidney samples reflecting abundant u-PA expression in this tissue. In comparison only a very low level of u-PA expression was detectable in brain tissue by Southern analysis of amplification products. This brain expression concords with the previous reports of u-PA expression in rat brain.

3.9 RT-PCR analysis of plasminogen expression

The presence of both plasminogen activators in brain tissue raises the question of local synthesis of their only physiological substrate, plasminogen. Plasmin is a broad spectrum protease of trypsic specificity, synthesized and secreted by the liver, and is present at high concentrations in plasma and other extracellular fluids. Circulating plasminogen would not be expected to cross the blood-brain barrier unless compromised during certain pathological conditions hence plasminogen is hypothesized to be expressed locally in brain tissue. Plasminogen was previously detected immunologically in microglia cultures (Nakajima et al., 1992a) and a recent in situ hybridization study in mouse has demonstrated the presence of plasminogen transcripts in the hippocampus (Tsirka et al., 1997).

PCR using plasminogen and HPRT specific primers was performed on rat liver and brain oligo-dT primed eDNA. The amplification products after 20, 25 and 30 cycles are presented in figure 3.9a. These were Southern blotted and hybridized to a confirmed plasminogen probe.

The 400 bp HPRT amplification product demonstrated a slight overloading of the brain samples in comparison to the liver samples. Liver plasminogen amplification was robust reflecting the abundant expression of plasminogen transcripts in this tissue. In comparison, brain amplification was absent. The higher sensitivity of detection revealed by Southern blotting of the amplification products suggested that plasminogen transcripts are not present in brain tissue.

Plasminogen transcripts have been detected by in situ hybridization on mouse brain slices (Tsirka et al., 1997). In an attempt to verify these results and to investigate possible species differences between rat and mouse, the above analysis was performed on mouse brain and liver samples. Amplification products are presented in figure 3.9b. As with rat, the liver expressed high levels of plasminogen transcript as indicated by a robust 300 bp amplification product. No amplification product was
Figure 3.9 - RT-PCR analysis of plasminogen expression in rodent brain

Ethidium bromide stained amplification products for the number of cycles indicated are shown (top panel in each case) and Southern hybridization of the products with a plasminogen probe are shown (bottom panel in each case).

a) Rat brain (Br) and liver (Li)
b) Mouse brain (Br) and liver (Li)

Numbers refer to numbers of PCR cycles performed.
obtained from mouse brain suggesting the absence of plasminogen transcripts in this tissue.

The failure to detect plasminogen transcripts by this very sensitive detection method is at odds with the previously reported expression study. The results may be reconciled if plasminogen is an activity regulated transcript as is the case with t-PA and u-PA. The lack of expression in this study may therefore reflect a different levels of activity and arousal in the rat and mouse samples used. The use of multiple samples from different animals however argues against this explanation. Overall it is reasonable to conclude that if plasminogen is expressed in brain tissue, it is expressed at very low levels. This raises the question of why its activator, specifically t-PA, is expressed at high levels in the brain. One possible hypothesis is that t-PA has another physiological substrate, expressed in brain at higher levels.

3.10 Hepatocyte Growth Factor expression in brain

Tissue plasminogen activator has a strikingly narrow substrate range. The only confirmed physiological substrate in vivo is a single peptide bond of plasminogen (Arg^{560} - Val^{561}). This highly restricted substrate specificity is in contrast to the broad specificity of other related serine proteases such as trypsin. Part of this specificity is due to a ternary interaction between t-PA, plasminogen and a cofactor, fibrin which binds to the finger domain at the amino terminus of t-PA. However, the isolated protease domain alone retains this high specificity for plasminogen so some inherent structural features displayed by native plasminogen must be responsible (Madison et al., 1995).

Hepatocyte Growth Factor (HGF) has many of the structural motifs associated with plasminogen including the kringle domains, however the enzyme no longer has a fully functional catalytic triad and is not an active protease. Consistent with this homology, HGF requires proteolytic activation at a specific arginine/valine peptide bond i.e. a site of tryptic specificity, converting the single chain inactive form into the mitogenically active double chain form.

HGF acts through its receptor c-met and has mitogenic, motogenic and morphogenic activities in a variety of cell types. In the central nervous system, HGF and its receptor are expressed in the hippocampus, the cerebral cortex, the amygdala and the olfactory bulb (Honda et al., 1995; Jung et al., 1994) and HGF promotes neuronal
survival and neurite outgrowth in primary cultures of hippocampal neurons (Honda et al., 1995) and cortical neurons (Hamanoue et al., 1996).

To confirm this previously reported brain expression and to assess the relative expression levels, RT-PCR analysis was performed on rat brain and liver oligo-dT primed cDNA. PCR was performed using HGF and HPRT specific primers. The amplification products after 20, 25 and 30 cycles were Southern blotted and hybridized to a confirmed HGF probe (figure 3.10).

The 400 bp HPRT amplification product demonstrated the equal loading and efficiency of amplification in both brain and liver samples. The presence of the expected 300 bp HGF amplification product in both samples confirms the presence of HGF transcripts in liver and brain. Indeed, the expression level was comparable in each case indicating that the brain is a significant site of expression for this growth factor.

The absence of significant levels of plasminogen transcripts along with the presence of significant levels of the related transcript, HGF, suggests that HGF may serve as the substrate for the plasminogen activators in brain. Plasminogen activators have been shown to activate the latent growth factor in vitro (Mars et al., 1993) and in one study, antibodies against t-PA blocked the brain extract mediated activation of HGF (Thewke and Seeds, 1996). However, a specific HGF activator present in serum was cloned from a liver cDNA library (Miyazawa et al., 1993). Interesting, this so called HGF activator is closely homologous to Hageman Factor, the blood clotting factor amplified from brain in the degenerate PCR screen and consequently Hageman Factor or blood clotting Factor XII has also been shown to activate HGF (Shimomura et al., 1995). Analysis of kinetic parameters for HGF activation suggests there exist two different pathways for HGF activation, that performed by Factor XII and its homologue is the predominant mechanism in serum, whereas HGF activation in tissues is achieved by the plasminogen activators (Naldini et al., 1995). The tight regulation for tissue HGF activation is provided by a mechanism where the latent growth factor binds to u-PA and is cleaved and thus activated at the cell surface. This stoichiometric activation would allow the tight control of extracellular activation of pro-HGF necessary within tissues such as the brain. This independent mechanism in tissue adds credence to the theory that the plasminogen activators are responsible for HGF activation in vivo.

The colocalisation of t-PA and HGF transcripts to adjacent cell types has been demonstrated within the olfactory bulb (Thewke and Seeds, 1996), arguing for a
Figure 3.10 - RT-PCR analysis of HGF transcripts in rat brain

Ethidium bromide stained amplification products for the number of cycles indicated are shown (top panel) and Southern hybridization of the products with a HGF probe are shown (bottom panel).

Samples are Br, brain; Li, liver. Numbers refer to number of PCR cycles performed.
potential interaction of these two molecules. HGF has been described as a guidance and survival factor for the developing nervous system (Ebens et al., 1996) so HGF has the ability to act as a neurotrophic factor. HGF and c-met receptor are expressed during brain development and transcript levels increase following ischemic injury (Honda et al., 1995). The expression of these species in the adult brain, along with the plasminogen activators, indicates a role in normal brain function. The plasminogen activators therefore are not only involved in an extracellular proteolytic cascade leading to structural reorganisation but may also activate latent HGF which acts in the guidance and survival of neuronal cells. The high levels of expression of HGF within the hippocampus and the induction of t-PA seen in this brain region following LTP suggests a role for this growth factor downstream of t-PA, possibly in the structural plasticity accompanying learning and memory processes.

3.11 Summary

<table>
<thead>
<tr>
<th>Protease</th>
<th>Expression pattern with brain</th>
<th>Detection method</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-PA</td>
<td>All brain regions tested</td>
<td>Northern</td>
</tr>
<tr>
<td>RNK-Met-1</td>
<td>Hippocampus, cortex</td>
<td>Northern; in situ</td>
</tr>
<tr>
<td>BSP-1</td>
<td>CA1, CA2, CA3, entorhinal cortex</td>
<td>Northern; in situ</td>
</tr>
<tr>
<td>BSP-2</td>
<td>Hippocampus, cortex</td>
<td>Northern; in situ</td>
</tr>
<tr>
<td>Elastase IV</td>
<td>CA1, cortex, cerebellum</td>
<td>in situ</td>
</tr>
<tr>
<td>CTRL</td>
<td>All brain regions</td>
<td>in situ</td>
</tr>
<tr>
<td>Complement 2</td>
<td>Brain vasculature, weak expression</td>
<td>in situ</td>
</tr>
<tr>
<td>Proteinase 3</td>
<td>Widespread weak expression</td>
<td>in situ</td>
</tr>
<tr>
<td>Chymotrypsin B</td>
<td>Widespread weak expression</td>
<td>in situ</td>
</tr>
<tr>
<td>Thrombin</td>
<td>Weak expression</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>u-PA</td>
<td>Weak expression</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>Plasminogen</td>
<td>No expression detected</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>HGF</td>
<td>Moderate expression</td>
<td>RT-PCR</td>
</tr>
</tbody>
</table>

Table 3.1 - Summary of expression studies for brain serine proteases

Expression analysis of the brain proteases amplified in the degenerate PCR screen reveals an array of overlapping expression patterns (Table 3.1). t-PA and RNK-Met-1 are expressed at highest levels followed by the two novel proteases BSP-1 and BSP-2. BSP-1 expression is restricted to the CA fields of the hippocampus and is thus a good
candidate for transgenic exploitation to aid in the investigation into the role of this brain region in learning and memory. The expression in brain of previously described serine proteases, not cloned or amplified in the degenerate screen, is confirmed although no plasminogen transcripts were detected by RT-PCR in mouse or rat brains. The expression of HGF, a possible alternative plasminogen activator substrate, in brain is described.
Chapter IV

Molecular characterisation of BSP-2

The following chapter describes the isolation of a full length cDNA for BSP-2 and comments on the significance of motifs in the predicted amino acid sequence.

4.1 Screening of a rat hippocampus cDNA library

To further characterise serine protease BSP-2, a full length cDNA cloning was performed. A rat hippocampus cDNA library was previously prepared in the laboratory from oligo-dT primed cDNA (Stapleton et al., 1995). The cDNAs were directionally cloned into the λZAPII vector (Stratagene) using a NotI site built into the oligo-dT primer and EcoRI adaptors. Some $1 \times 10^6$ plaques were plated, duplicate lifts were taken and hybridized with probe corresponding to the cloned BSP-2 PCR product. Only one duplicate hybridization signal was detected. The corresponding plug was titred, plated at low density and rehybridized. Many clear duplicated hybridization signals were detected on secondary lifts. Well-separated plaques were picked and the pBluescript phagemid was in vivo excised using the ExAssist helper phage system (Stratagene). The resulting phagemid was cut with NotI and EcoRI which released an insert of approximately 1300 bp. This corresponds to the size of the BSP-2 transcript seen by Northern analysis (see Chapter II). The BSP-2 PCR sequence hybridized to a Southern blot of the 1300 bp insert released from the vector, confirming the cloned cDNA as genuine.

The phagemid was sequenced using the T7 and T3 primers in the pBluescript polylinker and the resulting sequence information confirmed the insert to be a serine protease sequence by the presence of characteristic motifs in the translated sequence (see chapter II).

4.2 The sequencing of BSP-2 cDNA

To obtain accurate sequence information it was necessary to sequence the cDNA in both orientations. The manual method of sequencing used can only guarantee 250-300
bp of readable sequence per reaction so it was necessary to subclone the BSP-2 cDNA. The purified 1300 bp \textit{NotI} and \textit{EcoRI} cDNA fragment obtained from the phagemid was thus further digested with \textit{Sau3AI}, \textit{RsaI} and \textit{PvuII}. The resulting fragments were cloned into the \textit{BamHI} or \textit{EcoRV} sites of pBluescript as appropriate and sequenced manually using the T7 and T3 primer sites in the polylinker. Based on this sequence information, two oligonucleotide primers were designed and used to prime sequence reactions to verify regions of the sequence and complete the contig. Figure 4.1 shows the contig map obtained for the BSP-2 cDNA along with the positions of the restriction sites used in the subcloning. The final sequence along with the predicted peptide sequence of the protein is presented in figure 4.2.

4.3 The BSP-2 open reading frame is disrupted

Translation of the cDNA sequence revealed an open reading frame of 299 residues but with an out-of-frame initiating methionine residue. Two upstream methionine codons in the context of a translation initiating methionine (Kozak, 1991) are present but neither are in the same frame as the predicted peptide sequence. Due to the high GC content of the region this result was considered to be a sequencing artefact. To expand any ambiguous compression regions, the sequencing of this 5' end of the cDNA was repeated using the dGTP analogues, dITP and 7-deaza-dGTP. The sequence information from these reactions, however, agreed with the original sequence, confirming the absence of an initiating methionine residue.

The absence of an in-frame methionine was hypothesized to result from an artefact in the cDNA synthesis or cloning. A frame-shift by the reverse transcriptase or a random mutation caused by the propagation of the cDNA in bacteriophage or in the pBluescript phagemid could result in this phenomenon.

In an attempt to resolve this question the genomic sequence corresponding to the 5' end of BSP-2 was examined. Primers were designed flanking this region and PCR was performed from rat genomic DNA. The resulting amplification product was approximately 700 bp long, 450 bp longer than predicted, due to the presence of an intron (figure 4.3a). This genomic BSP-2 fragment was cloned into the pCR2 vector and sequenced using the T7 and SP6 primers in the polylinker.

The sequence information obtained using dGTP, dITP and 7-deaza-dGTP confirmed these initiating methionine residue to be out of frame with the predicted peptide.
Figure 4.1 - Contig map for the sequencing of BSP-2

BSP-2 cDNA is shown diagramatically above, the scale is in base pairs. Restriction sites used for subcloning are indicated - Sau3A I (S), RsaI (R), PvuII (P). Arrows indicate individual sequencing reactions. BSP-2 5' and BSP2- 3' are specifically designed primers.
Figure 4.2 - BSP-2 cDNA and the sequence of the encoded polypeptide

The two potential Kozak initiation consensus sequences and a potential polyadenylation signal are underlined. The putative signal sequence is boxed and the proposed propeptide is shaded. Catalytic histidine, aspartic acid and serine residues are underlined. Residues contributing to substrate specificity are circled.
sequence of the serine protease. If this sequence information is reliable, one possible explanation is that BSP-2 may represent a transcribed pseudogene that would not be expected to be translated into a mature enzyme due to the absence of a complete open reading frame. However, dysfunctional pseudogenes are often characterised by the presence of multiple mutations and stop codons in all three reading frames. In this case, the only aberrant aspect of the sequence is at the 5' end reflecting the initiation methionine. The rest of the open reading frame is entirely intact and all the serine protease motifs are present (see below).

To investigate the conservation of this aberrant reading frame, the same genomic region, representing the 5' region of the BSP-2 gene, was amplified from a different strain of rat, sprague-dawley. This product was cloned into the pCR2 vector and was sequenced using the T7 and SP6 primer anneal sites in the polylinker. No sequence differences were identified between the two strains of rat, and thus the out-of-frame methionine was preserved. To investigate the conservation of this anomaly across species, attempts were made to clone the corresponding sequence from mouse, however no amplification product was obtained using the primers on mouse genomic DNA and brain cDNA. Whether the lack of an initiating in-frame methionine reflects a sequencing artefact or a genuine anomaly remains uncertain.

For the purpose of the analysis of the novel sequence, the downstream methionine, encoded at position 89-91, is considered to be the correct initiation signal of the cDNA. Initiation from the upstream methionine encoded at position 54-56, would yield a product without a signal sequence as determined by the presence of polar residues within the putative amino-terminal region (see figure 4.2).

4.4 BSP-2 sequence shows homology with serine proteases

The sequence obtained for BSP-2 is approximately 30% homologous to members of the serine protease family at the nucleotide level. The nucleotide sequence is most homologous to mast cell tryptase (35%).

At the amino acid level, BSP-2 is approximately 35% homologous to the mast cell tryptases and 25% homologous to trypsinogen and tissue kallikrein. The primary sequence has all the structural features characteristic of a serine protease. A hydrophobic signal sequence is discernible and, using the von Heijne algorithm
Figure 4.3 - BSP-2 5' genomic cloning

a) The 5' region of the cDNA and genomic sequences showing the position of the primers and the intron. 5' untranslated region is indicated as a shaded box. The scale is in base pairs.

b) An alignment of part of the cDNA sequence and the sequenced genomic PCR fragments from Fisher and Sprague-Dawley strain. Intron sequence is boxed. The putative upstream initiation codon, methionine-1, and alanine-3 are highlighted, signifying the open reading frame shift.
(1986), the cleavage site is predicted to occur after the 22nd amino acid. One potential N-glycosylation site is present at position 62.

Serine proteases are almost invariably synthesized as inactive precursors that are activated by cleavage of a pro-peptide. The new amino terminal residue is frequently hydrophobic, commonly isoleucine, but may often be leucine, valine or methionine. The new amino terminal α-amino group is thought to form a salt bridge with the aspartic acid adjacent to the catalytic serine residue, position 194 in the chymotrypsin numbering system (Hartley, 1964). This results in a conformational change in the enzyme such that the substrate specificity pocket and the oxyanion hole, incomplete in the zymogen, are completed allowing full enzyme activity. In the majority of serine proteases activated by trypsin like proteases this hydrophobic residue lies within a consensus which is R/K I/V I/V G G. Cleavage occurs after the arginine or lysine residue, a trypsin target, releasing the new α-amino group of the hydrophobic residue for interactions with aspartic acid-194. In the BSP-2 sequence this consensus is present as R V V G G arguing that the new amino-terminal of the mature, active enzyme is valine-42 of the cDNA polypeptide sequence.

4.5 BSP-2 may represent a trypsin-like protease

Within the BSP-2 polypeptide sequence, all the characteristic motifs of the serine protease family are conserved including the three catalytic triad regions. The position of the cysteine residues relative to the catalytic residues, thought to be related to the tertiary structure of the enzyme, is also preserved. As outlined in chapter II, the presence of an aspartic acid residue at position 228 adjacent to the catalytic serine region is indicative of tryptic specificity. This residue lies within the substrate binding pocket and is thought to neutralise the charge on the arginine or lysine residue of the substrate. In the mature enzyme, the glycine residues at positions 255 and 265 in BSP-2 lie in the substrate binding pocket. The small size of the side-chains of these residues allows for a deep binding cleft and is consistent with the proposed tryptic specificity. A large substrate binding pocket is required for the large preferred substrate residues arginine and lysine. An alignment of the translated BSP-2 sequence with other tryptic proteases is presented in figure 4.4.
Figure 4.4 - Alignment of BSP-2 with trypsin-like serine proteases

Sequences are aligned from the putative amino terminus of the catalytically active enzyme. Sequences are mouse mast cell protease-6 (MMCP-6), mouse trypsin 1 (Trypsin), rat tissue kallikrein 1 (Kallikrein), mouse tryptase 2 (Tryptase).

4.6 Discussion

The sequence information suggests BSP-2 to be an extracellular protease of tryptic specificity which represents a novel member of the serine protease family. The sequence is most homologous at the amino-acid level to the trypsin enzymes of mast cells. Mast cells express an abundance of serine proteases which are involved in diverse functions such as peptide hormone processing, inflammatory response and parasite expulsion. The low homology between this subfamily and BSP-2 is considered not to be of any functional significance but merely to reflect homology with other trypsin proteases. The expression pattern of BSP-2 is not consistent with a mast-cell site of expression (see chapter III).

The BSP-2 sequence has a proenzyme activation consensus sequence which suggests the BSP-2 zymogen is activated by another trypsin serine protease. The recent characterisation of multiple trypsin serine proteases within the brain indicates that these...
related proteases may function in catalytic cascades of sequential zymogen activation. These cascades may act downstream of t-PA activation and thus may play important roles in the synaptic remodelling associated with hippocampal function.

The lack of an in-frame methionine could suggest that this sequence is transcribed and not translated and, as such, could be classified as a pseudogene. The preservation of all serine protease conserved domains along with an open-reading frame devoid of stop codons however argues against this. Despite the sequencing of this product manually with several nucleotide analogues the most likely explanation is that the anomaly represents a sequencing artefact possibly caused by the GC rich content of the upstream region. Further investigation is warranted to establish the cause of this anomaly.

4.7 Summary

A full length BSP-2 cDNA has been cloned which encodes a novel tryptic serine protease. The predicted amino-acid sequence however lacks an in-frame initiating methionine residue and this anomaly appears to be conserved across difference strains of rats.
The following chapter describes the construction of a cDNA representing the full coding region of BSP-1. The translated amino acid sequence is discussed in relation to the serine protease multigene family. The genomic structure of BSP-1 is also described.

5.1 Screening of a rat hippocampus cDNA library

The rat hippocampus cDNA library (Stapleton et al., 1995) used in the BSP-2 screen was used to screen for a BSP-1 cDNA. The same duplicate filters representing 1x10^6 individual plaques were stripped and hybridized to the cloned BSP-1 amplification product generated in the original degenerate PCR screen. No duplicated hybridization signals were obtained after prolonged autoradiography.

Two further rat brain cDNA libraries were available. One, previously constructed in our laboratory (Stapleton et al., 1995), was prepared from rat brains from which the hippocampus had been removed. This so called, "rest of brain" cDNA library is a oligo-dT primed library directionally cloned into the λZAPII vector (Stratagene). The in situ hybridization analysis of BSP-1 expression described in the previous chapter, revealed the hippocampus to be the predominant site of expression of this serine protease. However, hybridization signals were detectable outside the hippocampus, from certain layers of the entorhinal cortex. Accordingly, this library constructed from all brain material apart from the hippocampus may contain a cDNA encoding BSP-1. The other cDNA library was a commercially available rat total brain cDNA library (Clontech), a non-directionally cloned oligo-dT primed cDNA library.

As a preliminary experiment, aliquots of the libraries were taken and the inserts were amplified using T3 and T7 primers annealing to the polylinker. An extension time of 1 min was used to ensure that the cDNA of BSP-1, which was expected to be 1300 bp long, as determined by Northern analysis (see chapter III), was fully extended. The resulting amplified inserts were Southern blotted and hybridized to labelled, random primed BSP-1 amplification product. Despite prolonged exposure times, no
hybridization signal was detectable from either library. These libraries were not investigated further.

5.2 A BSP-1 cDNA constructed from 5' and 3' RACE products

As an alternative to conventional library screening, the technique of rapid amplification of cDNA ends (RACE) was used to obtain sequence corresponding to the upstream and downstream regions of the BSP-1 amplification product, obtained in the original screen. Two nested reverse primers, for 5' RACE, were designed to the 5' end of the sequence and a forward primer for amplification of the 3' sequence was designed towards the 3' end of the sequence.

The lack of representation of BSP-1 cDNA in the libraries investigated may reflect a block at the reverse transcription phase of library construction. It is known that secondary structure within certain mRNA species can impede the progression of the reverse transcriptase enzyme. To minimize the formation of inhibitory secondary structures, all subsequent cDNA syntheses were performed in the presence of 10% DMSO. The reverse transcription reaction was allowed to progress for one hour at 42°C followed by a short 50°C heat pulse after which fresh enzyme was added. The rationale behind this heat pulse was to melt any secondary structures which would cause the reverse transcriptase to stall. Fresh enzyme is then added to replace the potentially denatured stalled enzyme. This modification is described in Steel et al., (1998).

Poly A+ RNA was prepared from rat hippocampus and was reverse transcribed using an oligo-dT primer to yield 1st strand cDNA. 3' PCR was performed on this material using the forward primer and an oligo-dT-XhoI primer. A PCR product of approximately 600 bp was obtained. The authenticity of this product was confirmed by hybridization to the original BSP-1 amplification product in Southern analysis. This 3' sequence was cloned into the pCR2 vector and sequenced using the T7 and SP6 primers of the polylinker.

Poly A+ RNA was then reverse transcribed using the most 3' of the two nested reverse primers and tailed using terminal transferase with dATP. Second strand synthesis was then performed by annealing an oligo-dT primer with a 5' extension of unique sequence and extension was performed with Klenow fragment. A subsequent round of PCR was performed using a primer designed to anneal to the 5' extension on the
oligo-dT primer. This allows a more specific second round amplification as only genuine double stranded cDNAs can be amplified. Resulting products were cloned into pCR2 vector and recombinants were gridded out and probed with the BSP-1 product. One positive was obtained which contained an amplification product of approximately 450 bp. This clone was sequenced using the T7 and SP6 sites of the polylinker.

The sequence information allowed new primers to be designed corresponding to the most 5' and 3' regions of the BSP-1 cDNA. A full cDNA was then amplified using these primers from oligo-dT primed 1st strand cDNA which yielded a product just under 1 kb in length. This was then cloned into the pCR2 vector. The construction of the BSP-1 cDNA is summarised in figure 5.1.

5.3 The sequencing of BSP-1 cDNA

The use of RACE cloning for the isolation of cDNA sequences involved multiple rounds of PCR. This may mean that mutations may arise in the cDNA sequence. The enzyme used in the above manipulations was Taq polymerase which has no inherent proof-reading ability and has an appreciable mutation rate. To minimize such mutations from successive rounds of PCR, the final cDNA was obtained from a single PCR reaction from rat hippocampus cDNA. This amplification product was then sequenced anew.

The cDNA insert was released from the vector by digestion with ApaI and SacI. The released insert thus represents the cDNA obtained flanked by regions of polylinker. This insert was then digested with Sau3AI and TaqI and products were cloned into the compatible BamHI and ClaI sites in pBluescript. These subcloned fragments were then sequenced using the T7 and T3 primers from the polylinker. Based on this sequence information, two oligonucleotide primers were designed and used to prime sequence reactions to complete the contig and verify regions of sequence. A contig map of the sequence is presented in figure 5.2. The final sequence along with the predicted amino acid sequence of the protein is shown in figure 5.3.
Figure 5.1 - The RACE cloning of BSP-1 cDNA

a) Cloning of the 5' end of BSP-1 cDNA
i. Rat hippocampus poly A+ RNA was reverse transcribed using RACE1 primer
ii. 1st strand cDNA was tailed with adenosine residues and second strand synthesis was performed using a linker-oligo-dT primer.
iii. PCR was performed using RACE2 and the linker primer.
iv. The resulting 5' region of BSP-1 cDNA to which an upstream primer, rB5 was designed

b) Cloning of the 3' end of BSP-1 cDNA
i. Rat hippocampus poly A+ RNA was reverse transcribed using oligo-dT primer.
ii. PCR was performed using RACE3 and the oligo-dT primer.
iii. The resulting 3' region of BSP-1 cDNA to which an upstream primer, rB3 was designed.
PCR with rB3 and rB5 primers was performed on 1st strand oligo-dT primed cDNA to yield the entire BSP-1 cDNA
BSP-1 cDNA is shown diagramatically above, the scale is in base pairs. Restriction sites used for subcloning are indicated - *Sau3A* I (S), *Taq* I (T). Arrows indicate individual sequencing reactions. BSP-1 5' and BSP-1 3' are specifically designed primer.
Figure 5.3 - BSP-1 cDNA and the sequence of the encoded polypeptide

Kozak initiation consensus sequence and a potential polyadenylation signal are underlined. The putative signal sequence is boxed and the proposed propeptide is shaded. Catalytic histidine, aspartic acid and serine residues are underlined. Residues contributing to substrate specificity are circled.
5.4 BSP-1 shows homology with trypsin-like proteases

The nucleotide sequence obtained for BSP-1 is most homologous to the trypsin subfamily of serine proteases. The level of homology is around 40% which suggest that BSP-1 represents a novel trypsin-like serine protease. The predicted amino acid sequence is 42% homologous to trypsinogen I and members of the kallikrein gene family.

The nucleotide sequence has all the expected motifs of a translated cDNA including an initiating methionine codon in the context of a Kozak sequence (encoded at position 146-148). There is also a polyadenylation sequence at position 947. As with many such enzymes there is a hydrophobic signal sequence which is predicted to be cleaved after position 28 in the polypeptide chain, using the von Heijne algorithm (1986). One potential N-glycosylation site is present at position 110.

Serine proteases are synthesized as inactive precursors and require proteolytic activation to release an inhibitory peptide from the amino terminus. This results in a conformational change in the enzyme which allows the substrate binding domain to be completed, yielding the active enzyme. The majority of serine proteases activated by trypsin-like serine proteases have a pro-peptide activation consensus which is R/K I/V I/V G G (see chapter IV). Cleavage thus occurs after the arginine or lysine residue to release the active enzyme. Within the BSP-1 amino acid sequence, the motif K I L E G aligns with the above consensus in other trypsin-like serine protease. Although not being a strong consensus match, cleavage after the lysine residue would release a new amino terminal hydrophobic isoleucine residue which is characteristic of this subfamily. The amino terminal residue of the mature, processed enzyme is thus predicted to be isoleucine-33.

All three catalytic triad motifs are well conserved in this sequence, along with the spacing of cysteine residues and other motifs characteristic of serine proteases. The presence of an aspartic acid residue at positions 206, as well as glycine residues at positions 229 and 239, is suggestive of BSP-1 having tryptic specificity as expected. The aspartic acid residue, lies at the base of the substrate binding pocket and the negative charge on the carboxyl group is thought to neutralise the positive charge on the lysine or arginine residue of the substrate. The glycine residues, being of small size allow large side chains such as lysine and arginine to be accommodated.
The full sequence information for BSP-1 thus suggests the enzyme to be a trypsin like serine protease. An alignment of the translated BSP-1 sequence with other tryptic serine proteases is presented in figure 5.4.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILEGQECKPHSQPWTALFQGERLVCVGVLQGDRVW/LTAHCKK--DKYSVRLGDSHLQKRDEPEQE</td>
<td>BSP-1</td>
</tr>
<tr>
<td>VHSG--PCDTSHPYQAALITSQLICGGVLIHPLMV/LTAHCKK--PNLQVFGLGHNRLQRESSSEQ</td>
<td>Neurosin</td>
</tr>
<tr>
<td>TVGGTYTCPEH5VYPQVSL--NSYHPCQGGLINQQMVVAHSCYK--SRIQVRLHGENNVLBDEQF</td>
<td>Trypsin</td>
</tr>
<tr>
<td>VVGGYNCEHSNSQP6QNVAVYYPYEILCQGGLVIDSPWTVTAHCAT--DNYQVWLGNNLYEDPFPAQM</td>
<td>Kallikrein</td>
</tr>
<tr>
<td>IIQGREGVQPHSRFP6ASNYQIRCQHICQGGLIKPQVLYLTAHCSRGSHSPTVLGALHSQKNEPEQ</td>
<td>Tryptase</td>
</tr>
<tr>
<td>IQVARSIQHPFCF-------SSNP-EDHSDMVMLRLQNSALGDKVPIEL--AMLCPKVGQKCIS</td>
<td>BSP-1</td>
</tr>
<tr>
<td>SSVVRVAPIVHDY-------AASS--HDDQILIARLPKSELIQPLPL--ERDSCANTTSCIL</td>
<td>Neurosin</td>
</tr>
<tr>
<td>INAAKIIKHHPYS-------SWT--LNNDIMLIKLSSPVKLARNAPVAL--PSACATPAGQCLIS</td>
<td>Trypsin</td>
</tr>
<tr>
<td>RLVSQVFPHGNGNQDLIWNHTRQPGDYNSDLMLHLHPSQPADTDGKVIDIL--PIEPEKVGSTCLAS</td>
<td>Kallikrein</td>
</tr>
<tr>
<td>FEIKEF1PFSGFKSG-------TNDMLIKRLTAELNKHQLLLHLSNRYTRDKCQVT</td>
<td>Tryptase</td>
</tr>
<tr>
<td>GWGTVEIPFQFLEMENTNAEQVYQSNQSCER--------YPQKITEGVCAGSSN-GADTCQGDSGPLL</td>
<td>BSP-1</td>
</tr>
<tr>
<td>GWG--KQDGFQDPTQIQCAYHLVRECEFA--------YPQQTQQMLCAGDEKYKDSQGDSGPLL</td>
<td>Neurosin</td>
</tr>
<tr>
<td>GWGNTLSCNGNPLIDLCQVADPLSQAQCEAA--------YPGETSSMICVGFLEGKGRDSCQGDSGPPV</td>
<td>Trypsin</td>
</tr>
<tr>
<td>GWGSIKTPGDLLESQNVWTDLSNREKAEV--------KKEEVTQMLCACEDGDKNTCQGDSQGPI</td>
<td>Kallikrein</td>
</tr>
<tr>
<td>GWGSTKPDVLTSTDLQEVTVTISRRKRCNWSYHNFVITKDMIAGDGRGERDSGCGGGLI</td>
<td>Tryptase</td>
</tr>
</tbody>
</table>

Figure 5.4 - Alignment of BSP-1 with trypsin-like serine proteases

Sequences are aligned from the putative amino terminus of the catalytically active enzyme. Sequences are mouse neurosin, mouse trypsin 1 (Trypsin), rat tissue kallikrein (Kallikrein), mouse tryptase 2 (Tryptase).

5.5 BSP-1 is the rat orthologue of the mouse protease, neuropsin

During the characterisation of BSP-1, a highly homologous serine protease was cloned from mouse brain cDNA (Chen et al., 1995). Given the name neuropsin, this amino acid sequence is 92% homologous to the BSP-1 sequence and thus is likely to represent the mouse orthologue of this protease. The published mouse cDNA is 1333 bp long, compared with the 974 bp of rat cDNA sequence obtained above. The difference lies in a longer 5' untranslated region in the mouse cDNA. The Northern blot of BSP-1 indicated a transcript size of approximately 1.3 kb, so the rat cDNA obtained is truncated at the 5' end. Figure 5.5 presents an alignment of the two rodent sequences.
Figure 5.5 - BSP-1 represents the rat orthologue of neuropsin

a) Representation of BSP-1 and neuropsin cDNAs showing the difference in the 5’ untranslated region. Open reading frame is indicated as a box.

b) Alignment of BSP-1 and neuropsin coding region. Amino-acid differences are boxed in the neuropsin sequence.
The published mouse cDNA sequence greatly facilitated the isolation of mouse genomic sequences corresponding to BSP-1 as regions of the mouse gene could be amplified by PCR (see below).

5.6 Identification of BSP-1 genomic structure using long-range PCR

For the exploitation of BSP-1 in transgenic animals, a knowledge of the genomic structure of BSP-1 is a prerequisite. The construction of a targeting vector for the manipulation of this gene requires the isolation and characterisation of genomic sequences. Conventionally this is achieved through the construction and screening of genomic libraries, however a quicker and simpler approach takes advantage of the new technology of long-range PCR polymerases, now commercially available. Primers designed to a particular mouse cDNA can be used to amplify corresponding regions of genomic sequence assuming moderate sized introns.

The enzyme used in this study was the Expand Long Range PCR system (Boehringer), a mixture of *Taq* polymerase and *Pwo* polymerase. This cocktail allows the amplification of longer products (up to 20 kb) than conventional *Taq* polymerase allows. Also the inclusion of a polymerase with proof-reading ability increases the fidelity of the reaction. Error rates are at least three fold lower in comparison to *Taq* enzymes alone (Barnes *et al.*, 1986).

Within the serine protease family, genomic structure is relatively well conserved. An alignment of genomic sequences reveals that certain intron and exon boundaries generally lie in homologous positions in related family members. This reflects the gene duplication events that caused the expansion of such gene families. Figure 5.6 shows the genomic structure of several subfamilies of serine proteases and the position of conserved intron/exon boundaries in relation to the amino acid sequence. Particularly conserved intron/exon boundaries lie immediately 3' to the region encoding the catalytic histidine residue and immediately 5' to the region encoding the catalytic serine residue in all family members. Chymotrypsin-like serine proteases are encoded on seven exons whereas elastase-like enzymes are encoded on eight exons. The presence of four exons was assumed for BSP-1 on the basis of conserved organisation of trypsin-like serine proteases.
Figure 5.6 - Genomic structure of serine proteases

The positions of introns and exons in four representative serine protease genomic sequences is presented schematically, sequences are not to scale. Exons are shown as boxes, introns as lines. The conserved amino acid sequences neighbouring the intron/exon boundary 3' to the catalytic histidine and 5' to the catalytic serine are shown. BSP-1 sequence is indicated above the alignment and the positions of the mapping primers for long-range PCR are displayed.
PCR primers were designed flanking these two conserved intron/exon boundaries and primers were designed at the 5' and 3' ends of the cDNA. For long range PCR, it is beneficial for primers to be relatively long, so regions of at least 24 nucleotides were selected. Furthermore, to reduce the possibility of artefactual amplification, the annealing temperature was raised to above 60°C. The position of the primers in the cDNA sequence is shown schematically in figure 5.7a and the primer sequences are listed in chapter 8 and Appendix A.

The primers, although originally designed in pairs, were used in all combinations for long-range PCR from strain 129 genomic DNA. Amplification products are shown in figure 5.7b. A comparison of the size of the genomic amplification products compared against the hypothesized sizes obtained from cDNA amplification, allows the size of the intervening introns to be calculated (table 5.1 and 5.2).

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>Predicted product size (exact)</th>
<th>Observed product size (estimate)</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5-C3</td>
<td>1127 bp</td>
<td>6150 bp</td>
<td>5020 bp = Intron I+II+III+IV</td>
</tr>
<tr>
<td>C5-A3</td>
<td>641 bp</td>
<td>1500 bp</td>
<td>860 bp = Intron I+II</td>
</tr>
<tr>
<td>C5-B3</td>
<td>1020 bp</td>
<td>6050 bp</td>
<td>5030 bp = Intron I+II+III+IV</td>
</tr>
<tr>
<td>A5-A3</td>
<td>131 bp</td>
<td>450 bp</td>
<td>320 bp = Intron II</td>
</tr>
<tr>
<td>A5-B3</td>
<td>510 bp</td>
<td>5000 bp</td>
<td>4500 bp = Intron II+III+IV</td>
</tr>
<tr>
<td>B5-B3</td>
<td>83 bp</td>
<td>1450 bp</td>
<td>1370 bp = Intron IV</td>
</tr>
<tr>
<td>A5-C3</td>
<td>617 bp</td>
<td>5050 bp</td>
<td>4430 bp = Intron II+III+IV</td>
</tr>
<tr>
<td>B5-C3</td>
<td>190 bp</td>
<td>1550 bp</td>
<td>1360 bp = Intron IV</td>
</tr>
</tbody>
</table>

Table 5.1 - Size of BSP-1 genomic amplification products

| Intron I | 540 bp |
| Intron II| 320 bp |
| Intron III| 2560 bp |
| Intron IV| 1370 bp |

Table 5.2 - Predicted BSP-1 intron sizes
Figure 5.7 - BSP-1 genomic mapping by long-range PCR

a) Primer anneal sites for genomic mapping.
b) Long range PCR products from genomic 129 DNA.
The amplification products corresponding to introns II and IV were cloned into the pCR 2 vector and sequenced using the T7 and SP6 primer anneal sites in the polylinker to obtain 5' and 3' sequence information. By comparison with the cDNA sequence, the exact position of the intron and exon boundaries was elucidated (figure 5.8a).

The amplification product corresponding to the longest product (C5-C3) was cloned into pCR 2. This plasmid was then mapped using a variety of restriction endonuclease digests in single, double and triple combinations. The resulting fragments were electrophoretically separated and Southern blotted. Hybridization of these blots with labelled amplification products corresponding to intron II and intron IV allowed the position of restriction sites to be discerned. The known location of several sites in the mouse cDNA sequence allowed an accurate map of restriction sites to be constructed (figure 5.8b).

Using this map, regions were subcloned to allow the identification of the position of introns I and III. At the 5' end of the gene, a 500 bp SacI fragment was subcloned into pBluescript and was sequenced using the T7 and T3 polylinker primer sites to obtain the position of intron I. To obtain the position of intron III, the cloned C5-C3 amplification product was sequenced using a forward primer 53, designed to the 3' end of the predicted end of exon 3. The positions of these introns in the sequence are shown in figure 5.8a. All the sequence and mapping information is consistent with the assumption of five exons interrupted by four introns (figure 5.8b).

5.7 Discussion

The BSP-1 amino acid sequence is approximately 40% homologous to rat trypsinogen I and to a variety of members of the kallikrein family. Kallikreins are serine protease involved in the processing of various peptide hormones and growth factors. However, within the identified kallikrein family, the serine proteases are approximately 70% homologous, and so BSP-1 probably does not represent a novel member of this subfamily. The homology seen probably reflects the fact that the kallikreins are all trypsin-like serine proteases.
Figure 5.8 - Position of introns in the mouse BSP-1 sequence

Figure 5.9 - Genomic map of the murine BSP-1 locus

Exons are numbered and shaded. Restriction sites are marked: BamHI (B), SacI (S), XbaI (X), EcoRV (V), KpnI (K), EcoRI (E).
As with BSP-2, the predicted site of proteolytic cleavage leading to the activation of the zymogen occurs after a lysine residue, a site of tryptic specificity. This suggests the possibility that both these novel trypsin-like serine proteases function within a proteolytic cascade of successive zymogen activation.

The lack of representation of BSP-1 in the three libraries screened indicates that either the cDNA is expressed at low levels in brain tissue or the transcript is an inefficient template for reverse transcription. The former seems unlikely as this species was detectable by in situ hybridization and Northern analysis indicating moderate levels of expression. The latter is more likely as successful 5' RACE cloning of the sequence necessitated the use of 10% DMSO and heat pulses which have been found to aid reverse transcription of sequences with abundant secondary structure (Steel et al., 1998). The cDNA obtained is approximately 1 kb which is 300 bp shorter than the species identified by Northern analysis and the corresponding mouse orthologue, neuropsin.

The genomic organization of the mouse BSP-1 gene has been elucidated through the use of long-range PCR amplification. This represents an efficient technique for the mapping of genes. With limited cDNA sequence, primers can be designed and used to amplify the equivalent genomic regions giving a quick indication of the location of introns. The efficient cloning of PCR products representing genomic fragments also aids in the genomic mapping. Such techniques, assuming moderate intron size, obviate the need for genomic library screening to obtain gene structure information. However, upstream sequences corresponding to the transcriptional control regions remain inaccessible with this technique.

The homology to trypsin-like serine proteases is reflected in the genomic organisation of the BSP-1 gene. Comparison with genomic structures of the catalytic region of serine proteases, revealed that BSP-1 is most similar to the trypsinogen organisation of five exons and four introns. Furthermore, the intron/exon boundary positions correspond to the conserved positions predicted from the genomic structure of trypsin subfamily members. This adds further evidence for BSP-1 being a member of this subfamily; chymotrypsin like proteases are encoded on seven exons and elastase like proteases are typically encoded on eight exons (see figure 5.6). The conservation of intron and exon boundary positions reflects the gene duplication events that have caused the expansion of this family (Ohta, 1989). These intron positions therefore reflect the genomic organisation of the ancestral gene. Interestingly, they seem to
correspond to functional aspects of the proteins as each catalytic triad region is encoded on a separate exon. This idea of exons representing functional domains in proteins supports the "exon shuffling" hypothesis of protein evolution (Rogers, 1985).

5.8 Summary

A cDNA corresponding to the full coding region of BSP-1 was cloned using 3' and 5' RACE. The translated amino acid sequence reveals BSP-1 to be a trypsin-like serine protease. The publication of the mouse orthologue of this gene, facilitated the isolation and characterisation of the mouse genomic structure of BSP-1. Based on comparisons with other serine protease genomic structures, BSP-1 has a structure most similar to the trypsin-like subfamily.
Chapter VI

Targeted disruption of the BSP-1 locus

The following chapter describes the generation of a line of mice deficient in BSP-1 function. An IRES-lacZ reporter was introduced into the murine BSP-1 locus via homologous recombination in embryonic stem cells and a region of the gene required for proteolytic function was deleted. Preliminary analysis of β-galactosidase expression in the resulting heterozygous mutant embryo is reported.

6.1 The rationale behind the targeting experiment

The technology of gene targeting in embryonic stem (ES) cells allows defined mutations to be introduced into specific genes in the mouse genome. The role of a particular gene may then be explored through the phenotypic assessment of the resulting mutant. For studies into the central nervous system, this technology is of particular utility as a mutant mouse strain can be assessed electrophysiologically, biochemically and behaviourally allowing the relationship between neuronal function and higher cognitive functions such as learning and memory to be investigated.

Much transgenic experimentation has focused upon the hippocampus, in an attempt to analyse the role of this brain region in learning and memory. For example, mice harbouring a mutation in γPKC have deficits in hippocampal LTP and spatial learning (Abeliovich et al., 1993a; 1993b). This correlation supports the theory that hippocampal synaptic plasticity may underlie learning and memory (see chapter I). However, the gamma isoform of PKC is expressed outside the hippocampal formation so the learning deficit observed may be due to disrupted function in other brain regions.

Directing transgenic intervention exclusively within the hippocampus would allow the relationship between neuronal processes within this brain region and learning and memory to be analysed with greater confidence. There are two principal ways of achieving this goal; both rely upon the exploitation of an endogenous gene, restricted in its expression to the hippocampus.
The elements controlling the expression of such a gene can be used to direct transgene expression to the hippocampus in conventional transgenic mice. A construct consisting of the endogenous promoter fused to a transgene is injected into the pronucleus of a fertilized egg, allowing its integration at random into the host genome. Stable expression of the transgene in the resulting animals therefore duplicates the expression pattern of the endogenous gene.

Gene targeting can be employed to direct the integration of the transgene into the coding region of the endogenous gene. The use of an internal ribosome entry site (IRES) allows the inserted transgene to be translated from the regionally expressed message. IRES sequences were isolated from a picornavirus and have been previously used in a transgenic context to create fully translated dicistronic messages (Mountford et al., 1994).

The latter approach is favoured as the targeted nature of the insertion yields a single well-defined transgene integration. The random integration of the first method can lead to adjacent sequences causing ectopic expression or silencing of the transgene and the integration of multiple arrays can lead to a variegated pattern of expression (Dobie et al., 1996). This method also necessitates the isolation of the full complement of genomic sequences required to reproduce the transcriptional regulation of the endogenous gene which may be prohibitively large.

6.2 Transgenic experimentation with BSP-1

Previous chapters have described the characterisation of BSP-1, a novel serine protease which is expressed within the hippocampus CA fields and the entorhinal cortex. This gene is therefore an attractive candidate for transgenic exploitation in an attempt to direct transgene expression uniquely within the hippocampal formation.

The expression analysis of BSP-1 was performed in rat. To verify that the expression pattern of BSP-1 in mouse is similar, in situ hybridization analysis was performed on mouse brain (Figure 6.1). Specific hybridization was detected in the CA regions of the hippocampus, with a stronger signal in the CA1 and CA3 fields than in the CA2 field. Low level hybridization was also detected within the entorhinal cortex. The expression pattern in mouse, thus agrees with rat, and indicates that the murine BSP-1 locus is suitable for transgenic experimentation.
Figure 6.1 - *In situ* hybridization analysis of BSP-1 expression in mouse brain

a) Horizontal mouse section counterstained with neutral red
b) Contact autoradiogram of adjacent mouse section hybridized with labelled BSP-1
As a study into the feasibility of employing BSP-1 to direct transgene expression to the hippocampal formation, an initial experiment is described using an IRES-β-galactosidase reporter to target the murine locus. The activity of the β-galactosidase gene in the resulting heterozygous animals can then be assessed to gain an insight into the potential for expressing transgenes from this particular locus.

The restricted expression pattern of BSP-1 suggests an important role for this novel serine protease in hippocampal function, possibly in mediating aspects of structural plasticity which is well characterised within this brain region. Gene targeting allows the function of this gene to be explored by the generation of a line of mice deficient in BSP-1 function. The targeting vector described has been engineered to cause a deletion of endogenous sequences vital for the predicted serine protease function. The resulting animals will thus be deficient in an active BSP-1 protease and the phenotype of homozygous animals may provide information concerning the role of this specifically expressed serine protease.

6.3 Homologous recombination in ES cells

Embryonic stem cell lines are established from the inner cell mass of peri-implantation mouse blastocysts. These inner cell mass cells are pluripotent and, upon differentiation, contribute to all lineages of the embryo. Once isolated they can be cultured indefinitely in an undifferentiated state through the use of mitotically inactivated fibroblast feeder layers (Evans and Kaufman, 1981) or by the supplementation of media with leukaemia-inhibiting factor (LIF) (Smith et al., 1988). DNA can be introduced into ES cells via a variety of transfection techniques although most studies use electroporation which yields a relatively high transfection efficiency. The neomycin phosphotransferase (neo) gene which confers resistance to the antibiotic geneticin (G418) is commonly used to select cells integrating the exogenous DNA.

The directed targeted disruption of specific genes can be achieved via homologous recombination between the targeting vector and the endogenous gene. Targeting vectors are thus designed with two regions of homologous sequences flanking the required mutation to be introduced, along with a selectable marker. Homologous recombination is thought to initiate in one homology arm via single strand exchange and proceeds through the non-homologous insert by branch migration. The recombination intermediate is resolved in the other homology arm (Ellis and Bernstein, 1989).
The frequency of homologous recombination per cell is low, so a variety of selection techniques are used to enrich for cells in which homologous recombination has occurred. One method, known as positive and negative selection (Mansour et al., 1988), involves the construction of a targeting vector with a negative selectable marker flanking one of the homology arms. The Herpes Simplex virus thymidine kinase (HSVtk) gene which confers sensitivity to ganciclovir is commonly used. Homologous recombination where the initiation and resolution events occur within the homology arms, leads to a loss of the HSVtk gene; conversely non-homologous integration, where the recombination events occur outside of the region of homology leads to the integration of the HSVtk gene into the genome. Selection with ganciclovir thus enriches for homologous recombination events.

G418/ganciclovir resistant colonies are screened to detect restriction fragment length polymorphism caused by the recombination event by Southern analysis using probes hybridizing to sequences of the gene lying outside the region of homology. These probes will not hybridize to the targeting vector and so only homologous recombination events will be detected.

Recombinant ES cells can be injected into the blastocoele cavity of a host blastocyst and upon reimplantation into the uterus of a recipient female, the chimeric blastocyst can develop to term. Injected cells develop with the same developmental capacity as the host inner cell mass cells and so can contribute to all tissue types including the germ layer. Chimerism is assessed in the resulting progeny through coat colour markers; the ES cell line and the host blastocyst are of differing coat colour genotype. Most ES cell lines are of XY genotype and so fertile chimeras with a germ layer colonised by ES cells are male as XY ES cells masculinize a XX host blastocyst. Chimeric males are mated and the resultant progeny are assessed for germline transmission by coat colour markers and are then genotyped by analysis of DNA, obtained from tail biopsy.

6.4 BSP-1 targeting strategy

In chapter V, the characterisation of the murine BSP-1 locus is described. This was achieved through the use of long-range PCR rather than conventional isolation of genomic clones from libraries. A similar approach was chosen to obtain homology arms for a targeting vector. This greatly simplifies the construction of the vector, as
homology arms can be obtained simply on the basis of cDNA sequence information. Furthermore, restriction enzyme sites can be introduced on PCR primers which greatly facilitates the construction of the vector.

One consideration in using PCR generated homology arms is the rate of mutation associated with the polymerase. The frequency of homologous recombination has been shown to decrease if non-isogenic DNA is used for the targeting vector, implying that small mismatches in the targeting arms can lead to abortive recombination (Te Riele et al., 1992). A relatively low mutation rate is therefore desirable if the fidelity of the homology arms is to be maintained. The use of a proof reading thermostable polymerase greatly reduces the rate of mutation during amplification (Barnes et al., 1986).

The reporter/selection cassette used in this study was a gift from Dr Andrew Smith (Niels et al., 1996). At the 5' end there are three stop codons each in a different reading-frame, preventing the translation of the 3' end of the transcript. Downstream of these is an IRES-β-galactosidase-SV40 polyA cassette which allows the reporter β-galactosidase to be translated from the targeted transcript. A neomycin phosphotransferase (neo) gene driven by a herpes simplex virus MC1 promoter and fused to a polyadenylation signal is included which allows transfected cells to be selected in G418. A tandem repeat of the HSVtk gene driven by the MC1 promoter was used to select against non-homologous recombination (Smith et al., 1995).

A targeting strategy was designed, replacing a region of exon III corresponding to the catalytic aspartic acid residue with the above reporter/selection cassette (Figure 6.2). As discussed in the chapter I, the three catalytic triad residues are vital for enzymatic function of serine proteases. Since the catalytic mechanism of serine proteases is well characterised, deletion of one of these residues is sufficient to cause an effective null mutation. There is therefore no risk of any functionality being restored to the mutated gene product through illegitimate splicing around the selection cassette or by read-through translation.

The deletion of 190 bp of exon III, allows a 3' homology arm of 4500 bp and a 5' homology arm of 1200 bp. The length of homology arms is considered to be proportional to the rate of homologous recombination (Deng and Capecchi, 1992) so in general long homology arms are desirable. Homology arms of less than 1 kb generally lead to abortive recombination (Thomas et al., 1992). The relatively short 5'
Figure 6.2 - Targeting of the BSP-1 gene by homologous recombination

Structure of the BSP-1 targeting vector, the wild-type BSP-1 locus and the targeted BSP-1 allele are shown. Exons are indicated as numbered black boxes. The catalytic triad residues, histidine, aspartic acid and serine are indicated by asterisks. Two tandem copies of the thymidine kinase selection cassette (MCltk) are incorporated for negative selection against non-homologous integration. The positive selection cassette consists of stop codons in each frame (TAG³), an internal ribosome entry sequence (IRES), a β-galactosidase gene fused to a SV40 polyadenylation signal (lacZ pA) and a neomycin phosphotransferase gene driven by the thymidine kinase MC1 promoter and fused to a SV40 polyadenylation signal (MC1neo pA).
homology arm in the BSP-1 targeting strategy is therefore sufficient although probably not optimal for homologous recombination.

6.5 Construction of the BSP-1 targeting vector

Long-range PCR was performed from strain 129 DNA using primers designed from the published BSP-1 orthologue, neuropsin cDNA sequence. Figure 6.3 presents a schematic representation of the amplification and the PCR products obtained. *SalI* sites were incorporated into the PCR primers used to amplify the 5' homology arm and *NotI* sites were incorporated into the 3' homology arm primers. Both PCR products were cloned into the pCR2 vector using the 3' deoxyadenosine extensions generated by the polymerase. The 5' and 3' homology arms were then released from the plasmids by digestion with *SalI* and *NotI* respectively to release the arms with restriction enzyme overhangs.

The reporter/selection cassette obtained as a *BamHI* fragment was cloned into the *BamHI* site of pBluescript II KS (Stratagene) to create plasmid pBSPlko1.1. The *SalI* 5' homology arm fragment was then cloned into the *SalI* site in pBSPlko1.1 to create plasmid pBSPlko1.2. The orientation was confirmed by sequencing using the T3 primer anneal site in the polylinker. The *NotI* 3' homology arm was then cloned into the *NotI* site in pBSPlko1.2 and the orientation confirmed by sequencing using the T7 primer anneal site in the polylinker, creating plasmid pBSPlko1.3. The thymidine kinase negative selection cassette was obtained as a *XhoI* / *SalI* fragment and was cloned into the *XhoI* site in pBSPlko1.3, creating plasmid pBSPlko1.4. This created the finished targeting plasmid with a unique *XhoI* site for linearization. Figure 6.4a presents a schematic representation of the construction of this targeting vector and a map of pBSPlko1.4 is shown in figure 6.4b.

The 5' and 3' probes used for screening G418/ganciclovir were also generated by PCR using specific primers annealing to the regions external to the homology arms. For the 5' probe, PCR was performed with primers, P55 and P53 yielding a 330 bp fragment. For the 3' probe, PCR was performed with primers, P35 and C3, yielding a 150 bp fragment. These were cloned into pCR2 vector and verified by sequencing. Both probes hybridized to discrete bands on genomic Southern blots (not presented). The screening strategy is summarized in figure 6.5.
Figure 6.3 - PCR amplification of homology arms

a) The location of homology arm PCR primers in BSP-1 genomic locus. Exons are indicated as black boxes and the position of the catalytic residues histidine, aspartic acid and serine is indicated.
b) Amplification products corresponding to the 5' and 3' homology arms.
Figure 6.4 - Construction of the BSP-1 targeting vector

a) The polylinker of pBluescript is shown with the insertions at individual sites. The negative selection cassette dimer was cloned into the XhoI site; The 5' homology arm was cloned into the Sall site; the reporter/selection cassette was cloned into the BamHI site; the 3' homology arm was cloned into the NotI site.

b) A map of the pBSP-1ko1.4
Figure 6.5 - Screening strategy for homologous recombination

Wild type allele (upper panel) and targeted allele (lower panel) are shown. 5' and 3' probes are shown as black bars which lie outside of the regions of homology used for the targeting vector (dotted line) and HindIII sites are shown. Correct integration at the 5' end should yield a 2.6 kb Hind III restriction fragment and correct integration at the 3' end should yield a 14 kb restriction fragment. Both the 5' and 3' probes recognize the same 12 kb restriction fragment in the wild type allele.
6.6 Targeting of the BSP-1 locus

150 μg of pBSP1k1.4 was linearised with XhoI, phenol/chloroform extracted twice, chloroform extracted, precipitated and resuspended in 100 μl PBS. The linearised construct was then electroporated into 1x10⁸ E14-TG-2a subclone IV ES cells, and plated onto twenty gelatinised 9 cm dishes. Two control plates were set up with 5x10⁶ unelectroporated ES cells. After two days in culture, the cells were placed under selection with fresh media containing 100 μg/ml G418 and 2.5 μM ganciclovir. During selection the media was changed daily and all media was supplemented with LIF to prevent differentiation. After five days in selection colonies were visible on the experimental plates. Colonies were left to expand for 3-4 days before being picked, trypsinised and plated in individual gelatinised wells on 24 well plates. Out of a total of 2000 colonies, 198 were picked and 184 colonies expanded.

When cells were confluent, individual clones were passaged onto three separate 24 well plates. Two of the plates were allowed to reach 70% confluency before freezing on 24 well plates to allow a duplicated store of individual clones. The remaining plate was left until confluent, before harvesting the DNAs. 120 individual colonies were analysed by genomic Southern blot.

Approximately 10 μg of genomic DNA was digested with HindIII, electrophoretically separated on 0.7% agarose and transferred to nylon membranes by capillary blotting. The 5' probe generated by PCR was labelled by random priming and was hybridized to the blots. This 5' probe hybridized to a 12 kb fragment corresponding to the non-targeted allele and a 1.8 kb fragment corresponding to the targeted allele (see figure 6.5). Out of 120 clones analysed, two gave a band pattern corresponding to a targeted allele (figure 6.6). These two clones were then analysed using the 3' probe to ensure homologous recombination had occurred at both ends of the reporter/selection cassette. Southern blots were prepared as before from these two clones and hybridized to the PCR generated 3' probe labelled by random priming. This 3' probe hybridized to a 12 kb fragment corresponding to the non-targeted allele and a 14 kb fragment corresponding to the correctly targeted allele (figure 6.6). Out of 120 clones screened, two were therefore correctly targeted which represented a homologous recombination efficiency of 1.7%.
Figure 6.6 - Confirmation of homologous recombination

Southern analysis of ES cell colony DNA with 5' probe (left panel) and 3' probe (right panel). Samples are wild type DNA (wt), clone number 51 and clone number 55.
The clones were further characterised by long-range PCR to verify that the integration of the reporter/selection cassette had occurred. The primers used to map the genomic BSP-1 locus were used to PCR amplify regions of the gene into which the reporter/selection cassette was inserted. Amplification with primer C5 and B3 yielded a 6 kb product from the non-targeted allele and a 11 kb product from the targeted allele. Amplification with primer A5 and C3 yielded a 5 kb product from the non-targeted allele and a 10 kb product from the targeted allele. (figure 6.7). A Southern blot of the amplification products was probed with the neo gene from the reporter/selection cassette. This probe selectively hybridized to the larger amplification products from the targeted allele as expected (figure 6.7). This confirmed the correct integration of the report/selection cassette in both clones.

6.7 Derivation of BSP-1 mutant mice

The successfully targeted clones, numbers 51 and 55, were thawed and expanded. Between 12 and 15 ES cells were injected into the blastocoele cavity of expanded C57BL/6 strain blastocysts. The chimeric embryos were allowed to recover for a few hours before being transferred to pseudopregnant animals. The resulting progeny were assessed for ES cell contribution by coat colour. Table 6.1 summarises the results of the ES cell transfers and examples of the chimeras obtained are shown in figure 6.8.

<table>
<thead>
<tr>
<th>Clone</th>
<th>Transfers</th>
<th>Pups Born</th>
<th>Pups Weaned</th>
<th>Male chimeras</th>
</tr>
</thead>
<tbody>
<tr>
<td># 51</td>
<td>40</td>
<td>10</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td># 55</td>
<td>30</td>
<td>11</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 6.1 - Results of blastocyst transfers

Three of the best chimeras from each line were crossed to C57BL/6 females for test breeding. Contribution of the implanted cells to the germ line was assessed by agouti coat colour in the resulting progeny and individuals were then genotyped by genomic Southern analysis. DNA was prepared from tail biopsy, 10 μg was digested with EcoRV and separated electrophoretically on 0.7% agarose gels. Restriction fragments were transferred to nylon membrane by capillary blotting and hybridized to a TaqI genomic fragment. This TaqI fragment is a 500 bp restriction fragment obtained from digestion of the C3-C5 PCR product (see chapter V), corresponding to the 3' of intron...
Figure 6.7 - PCR confirmation of homologous recombination

a) Wild type allele (upper panel) and targeted allele (lower panel). Position of PCR primers recognising a targeting event at the 5' end are shown in black (C5-B3); primers recognising targeting at the 3' end are shown in gray (A5-C3). The dotted line indicates the extent of the homology arms.

b) PCR products obtained from amplification of wild-type DNA (wt) and DNA from ES cell clones numbers 51 and 55.

b) Southern blot of b, probed with a labeled neo fragment
Figure 6.8 - Examples of chimeras obtained

a) Chimeric male from line 51
b) Chimeric male from line 55
2 and the 5' of exon 3. This probe hybridized to a 3 kb non-targeted wild type allele and to a 1.8 kb fragment representing the targeted allele. The genotyping strategy is summarised in figure 6.9a and a representative blot is presented in figure 6.10a. Germ line transmission of the mutation was confirmed for both clones (table 6.2). All fertile chimeras transmitted the ES cell derived complement through the germ line and around 50% of the resulting agouti pups carried a single copy of the mutated BSP-1 locus as expected.

<table>
<thead>
<tr>
<th>Clone no.</th>
<th>Test Cross No.</th>
<th>Coat colour</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>agouti black</td>
<td>++/-</td>
</tr>
<tr>
<td>#51</td>
<td>640</td>
<td>7 0</td>
<td>4 3</td>
</tr>
<tr>
<td></td>
<td>641</td>
<td>0 0</td>
<td>- -</td>
</tr>
<tr>
<td></td>
<td>642</td>
<td>11 0</td>
<td>7 4</td>
</tr>
<tr>
<td>#55</td>
<td>643</td>
<td>5 0</td>
<td>2 3</td>
</tr>
<tr>
<td></td>
<td>644</td>
<td>0 0</td>
<td>- -</td>
</tr>
<tr>
<td></td>
<td>645</td>
<td>7 0</td>
<td>5 2</td>
</tr>
</tbody>
</table>

Table 6.2 - Results of chimera test crosses

A PCR analysis was also optimized to facilitate the genotyping of the animals. Two reverse primers were designed, one to the region of the cDNA that was deleted by the targeting event, ΔD3 and one to 5' sequence of the reporter/selection cassette, SC3 (figure 6.9b). These were used in conjunction with the genomic mapping primer, A5. A 650 bp product was amplified from the wild type copy of BSP-1 and a 450 bp product was amplified from the mutant copy of BSP-1. Wild-type, heterozygous and homozygous mutant animals can be recognised by the presence or absence of these bands. A representative PCR screen is presented in figure 6.10b.

6.8 Analysis of Reporter Expression

After germ-line transmission of the mutation has been established, the fertile transmitting chimeras were set up with MFI strain mice for the collection of embryos. Embryos from 10.5 dpc and 12.5 days post coitum (dpc) females were collected and stained as whole mounts with X-gal (figure 6.11).
Figure 6.9 - Genotyping strategy

a) Southern analysis. Wild type allele is recognized by a 2.8 kb EcoRV restriction fragment (top panel) and a targeted allele is recognized by a 1.8 kb EcoRV restriction fragment (bottom panel). Probe is shown as a black bar.

b) PCR analysis. Wild-type allele yields a 650 bp PCR amplification product with the common upstream primer (A5) and a primer specific to a region of exon III (ΔD3) that is deleted in the targeted allele (upper panel). Targeted allele, yields a 450 bp amplification product with the common upstream primer (A5) and a primer specific to a region of the selection cassette (SC3) (lower panel).
Figure 6.10 - Genotype analysis of F1

Example of Southern analysis of tail DNA from F1 progeny (upper panel). Wild-type allele is indicated by a 2.8 kb *EcoRV* restriction fragment. Targeted allele is indicated by a 1.8 kb *EcoRV* restriction fragment.

Example of PCR analysis of tail DNA from F1 progeny (lower panel). Wild type allele is indicated by a 650 bp amplification product. Targeted allele is indicated by a 450 bp amplification product.
No reporter expression was observable in any of the twelve 10.5 dpc embryos stained. The embryos were not genotyped and so the presence of the transgene in the progeny was not confirmed. Statistically however, 50% of the progeny would be expected to harbour the transgene and so the absence of staining in all embryos suggests a lack of reporter expression at this stage. As is often seen in whole mount staining, some low-level staining was found to develop slowly in the residual yolk sack in all the embryos and thus was considered background.

Reporter expression was observable in 7/11 of the 12.5 dpc embryos stained. The presence of non-staining embryos served as a useful control for background staining as these embryos were considered to be wild-type. Diffuse staining was observed within the head, possibly within the developing brain. Clear staining was observed in discrete regions of the developing eye, in the outer layer of the optic cup and within developing layers of the retina and lens. Staining was also strong in the primordia of the pinnar of the ear. Clear staining was detectable delineating the somites, while the developing spinal cord was apparently devoid of staining.

Staining within the developing brain and retina suggests a neuronal site of expression and reporter activity amidst the somites may reflect peripheral nerve staining. However, the whole mount procedure does not allow the identity of the expressing cells to be discerned, so few conclusions can be drawn with confidence.

This preliminary analysis of reporter expression suggests β-galactosidase is functional from the BSP-1 locus. A detailed analysis of reporter expression on sectioned embryos is warranted to elucidate the cell types expressing the reporter transgene. Too few animals are available for sacrifice at present, although the analysis of reporter expression in later stages of embryogenesis, postnatal brain development and adult is a priority. The analysis of reporter expression in the adult brain is crucial for an assessment of the use of the BSP-1 locus for directed transgene expression within the hippocampal formation.

6.9 Discussion

The targeting of the BSP-1 locus was achieved somewhat unconventionally with homology arms generated by PCR. PCR has been previously used to generate targeting constructs for the disruption of the renin-1 gene (Clark et al., 1997). In this study a proof reading polymerase was used to amplify homology arms from a
Figure 6.11 - Whole mount β-galactosidase reporter expression analysis

a) Representative 10.5 dpc embryo showing lack of staining in all tissues.
b) Dissected 12.5 dpc heads from putative heterozygote (left) and putative wild-type (right) embryos.
c) Dissected 12.5 dpc head showing staining diffusely throughout the head, within the developing eye and around the primodia of the pinnar of the ear.
d) Reporter expression within the 12.5 dpc eye.
e) Dissected 12.5 dpc head, higher magnification of c
f) Reporter expression in decapitated 12.5 dpc embryo
g) Reporter expression within the somites of a 12.5 dpc embryo
bacteriophage P1 clone containing the entire mouse locus. In this study, homology arms have been amplified directly from genomic DNA on the basis of cDNA sequence information. This method facilitates the construction of targeting vectors although the application of this method depends upon the size of intervening introns as some may be prohibitively large to allow direct PCR.

The targeting frequency obtained was under 2% which is fairly low for the selection method used. A thymidine kinase negative selection cassette lying outside the homology arms is considered to select against non-homologous recombination. The selection cassettes used in this experiment were used previously in the targeted disruption of the \( \textit{whn} \) locus (Nehls \textit{et al.}, 1996). In this study almost 90% of colonies obtained were homologous integrants. In general it is meaningless to compare the targeting efficiencies of different loci as every targeting construct is different and some loci may simply be more accessible for homologous recombination, however the low frequency of homologous recombination seen in this study may be explained by the short 5' homology arm used. Although homologous recombination in ES cells has been achieved using smaller 5' homology arms, for example the disruption of the Hox 2.6 locus was achieved with one homology arm of under 500 bp (Hasty \textit{et al.}, 1991), it is not considered optimal for efficient gene targeting. Furthermore, despite the proof reading activity of the \( \textit{Pwo} \) enzyme used for amplification, some mutations would have been incorporated into the homology arms and these introduced mismatches are considered to reduce the frequency of homologous recombination (Te Riele \textit{et al.}, 1992).

The introduction of an IRES element linked to a promoter-less \( \beta \)-galactosidase generates a dicistronic mRNA in which the reporter protein translation is linked to the BSP-1 gene expression. At the time of writing, the transgenic lines have just been established and as such, there are too few adult animals to justify sacrificing for an investigation into the adult expression pattern. An initial indication that reporter expression is functional in transgenic animals is provided by the embryonic X-gal staining patterns. Reporter expression was found to be absent at 10.5 dpc and abundant reporter expression was detected at 12.5 dpc. This expression appears to be widespread in the developing nervous system. The developmental expression pattern of BSP-1 was not determined within this study but a developmental expression profile for the mouse orthologue, neuropsin has been published (Suzuki \textit{et al.}, 1995). Expression of neuropsin is first detectable in 12 dpc mice in agreement with the analysis of reporter expression in this study. Although detailed expression patterns
were not reported, the reported expression seen in the BSP-1 transgenic mice is consistent with the published neuropsin expression pattern.

Initial indications therefore reveal that the β-galactosidase reporter expression is reflecting the endogenous pattern of BPS-I expression. Reporter expression in adult mice is yet to be investigated but adult reporter expression within the CA fields of the hippocampus and the entorhinal cortex is expected. This locus is therefore of considerable application for transgenic exploitation. Potentially any transgene could be cloned downstream of an IRES element and then targeted into the BSP-1 locus using the above homology arms. Expression of this transgene would then be restricted to the hippocampus CA fields and adjacent entorhinal cortex. A transgene causing an alteration in synaptic function could be used to generate mice with disruptions uniquely within certain neurons of the hippocampal formation. The behavioural or electrophysiological analysis of the resulting animals would provide insights into the role of hippocampal processes in higher order cognitive function.

A similar outcome was reported using a transgenic construct which directed Cre recombinase expression to the forebrain from the αCAMKII promoter (Tsien et al., 1996). Crossing these animals with a transgenic strain harbouring a ubiquitously expressed construct requiring Cre mediated recombination of loxP sites for the activation of a β-galactosidase reporter gene, revealed expression only within CA1 region of the hippocampus. This rather convoluted approach relies on the serendipitous expression levels of the αCAMKII being sufficient to drive Cre mediated recombination only within this subregion of the hippocampus. The approach used in this study involves β-galactosidase expressed in a defined and reproducible manner from a single well characterised integration.

One consideration however results from the developmental expression revealed by reporter activity in 12.5 dpc embryos. The fact that BSP-1 is expressed developmentally is cause for concern for its application in transgenic animals. The use of this locus to drive expression of transgenes to perturb hippocampal function in the adult will lead to expression of the transgene during nervous system development. Subtle developmental abnormalities may therefore result possibly contributing to the phenotype. Correlations drawn between hippocampal dysfunction and cognitive impairment would therefore be questionable.
Irrespective of the potential of this gene in transgenic experimentation, the targeting of the gene in this experiment is expected to yield an inactive gene. As such resulting homozygous animals will be deficient in BSP-1 activity. Any phenotype resulting will therefore allow the role of BSP-1 to be determined.

The abundant expression of BSP-1 in the developing nervous system is of interest in terms of its potential role in the adult brain. Structural plasticity in the adult hippocampus mimics many developmental processes such as neurite outgrowth and synaptogenesis. Many genes expressed during developmental processes are maintained in regions of the adult brain such as the hippocampus, that retain the capacity for structural modification. A role for BSP-1 in structural plasticity is therefore suggested by its widespread developmental expression and its restricted adult expression pattern specifically within the hippocampus.

6.10 Summary

The BSP-1 locus has been disrupted with an IRES-β-galactosidase cassette using a PCR targeting strategy. An initial assessment of reporter expression in embryos reveals β-galactosidase is being expressed from the locus. A region of the gene required for gene function has been deleted and the resulting line of mice may allow the role of this serine protease to be investigated through the phenotypic analysis of homozygous mutants.
Chapter VII

Materials and Methods

7.1 Materials

Unless otherwise stated, all chemicals used were of analytical grade and were supplied by BDH (AnalaR), Sigma or Fisher scientific. Ethanol was supplied by Hayman Ltd and Tris base was supplied by Boehringer Mannheim. Bacteria media components were supplied by DIFCO laboratories and electrophoresis grade agarose was provided by GIBCO BRL.

Enzymes, unless otherwise stated were supplied by Boehringer Mannheim. Synthetic oligonucleotides were synthesized by Oswel UK Ltd. Radiolabelled nucleotides were supplied by Amersham and New England Nuclear and X-ray film was supplied by Kodak.

7.2 Standard stock solutions used for molecular biology

Ampicillin - 50 mg/ml in water

Denhardt’s Reagant (100x) - 2% w/v Ficoll (Type 400, Pharmacia)
                                          2% w/v Polyvinylpyrrolidone
                                          2% w/v Bovine serum albumin (BSA)

DTT (Dithiothreitol) - M DTT in water

EDTA (Ethylenediaminetetraacetic acid) - 0.5 M Na₂EDTA
                                         pH 8.0 with NaOH

EGTA (Ethyleneglycol-bis(β-aminoethyl ether)-N,N,N′N′-tetraacetic acid) - 0.5 M Na₂EGTA
                                                                          pH 8.0 with NaOH
<table>
<thead>
<tr>
<th>Buffer Name</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gel loading buffer</td>
<td>- 0.25% w/v Orange G dye 30% w/v Glycerol in water</td>
</tr>
<tr>
<td>IPTG (Novabiochem)</td>
<td>- 20% w/v in water</td>
</tr>
<tr>
<td>(Isopropylthio β-D-galactoside)</td>
<td></td>
</tr>
<tr>
<td>10x MOPS</td>
<td>- 0.4 M MOPS pH 7.0</td>
</tr>
<tr>
<td>(3-[N-Morpholino]-propane-sulfonic acid)</td>
<td></td>
</tr>
<tr>
<td>PBS</td>
<td>- 0.137 M NaCl</td>
</tr>
<tr>
<td>(Phosphate buffered saline)</td>
<td></td>
</tr>
<tr>
<td>M Sodium Phosphate buffer</td>
<td>- 70.9 g Na₂HPO₄</td>
</tr>
<tr>
<td></td>
<td>4 ml 85% Orthophosphoric acid</td>
</tr>
<tr>
<td></td>
<td>Water to 1000 ml</td>
</tr>
<tr>
<td>20x SSC</td>
<td>- 3 M NaCl</td>
</tr>
<tr>
<td>(Standard saline citrate)</td>
<td></td>
</tr>
<tr>
<td>STE</td>
<td>- 0.1 M NaCl</td>
</tr>
<tr>
<td></td>
<td>10 mM Tris.Cl pH 8.0</td>
</tr>
<tr>
<td></td>
<td>1 mM EDTA pH 8.0</td>
</tr>
<tr>
<td>1x TAE</td>
<td>- 40 mM Tris Acetate, 1 mM EDTA</td>
</tr>
<tr>
<td>50x TAE</td>
<td>- 242g Tris Base</td>
</tr>
<tr>
<td></td>
<td>57.1 ml glacial acetic acid</td>
</tr>
<tr>
<td></td>
<td>100 ml 0.5 M Na₂EDTA pH 8.0</td>
</tr>
<tr>
<td></td>
<td>Water to 1000 ml</td>
</tr>
<tr>
<td>1x TBE</td>
<td>- 45 mM Tris Borate, 1 mM EDTA</td>
</tr>
</tbody>
</table>
10x TBE  
- 108g Tris Base  
55 g Boric acid  
40 ml 0.5 M Na₂EDTA pH 8.0  
Water to 1000 ml

TE buffer  
- 10 mM Tris.Cl pH 7.5  
1 mM Na₂EDTA pH 8.0

Tetracycline  
- 5 mg/ml in ethanol

Tris.Cl  
- 1 M Tris base pH with HCl  
(2-amino-2-hydroxymethyl-1,3-propanediol)

X-Gal (Biochem)  
- 20 mg/ml in dimethylformamide  
(5-bromo-4-chloro-3-indolyl β-D-galactoside)  
Storage at -20°C in dark

7.3 Standard solutions for bacterial culture

Maltose  
- 20% w/v in water

SM buffer  
- 100 mM NaCl  
10 mM MgSO₄  
50 mM Tris.Cl pH 7.5  
0.01% gelatin

LB medium  
(Leuria-Bertani)  
- 10 g Bacto-tryptone  
5 g Bacto-yeast extract  
10 g NaCl  
pH adjusted to 7.0 with NaOH  
Water to 1000 ml

2X TY medium  
- 16 g Bacto-tryptone  
10 g Bacto-yeast extract  
5 g NaCl  
pH adjusted to 7.0 with NaOH  
Water to 1000 ml
Agar - Bacto-agar added to LB medium
15 g/l for bottom agar
7 g/l for top agar
(Agarose used for plaque transfers)

7.4 Standard solutions for embryonic stem cell culture

ES cell culture medium - 340 ml Ultra High Purity (UHP) water
40 ml 10x GMEM (GIBCO)
13.2 ml 7.5% NaHCO₃
4 ml 100x Non-essential amino acids (GIBCO)
8ml 200 mM Glutamine/100 mM NaPyruvate
400 μl β-mercaptoethanol
40 ml Foetal calf serum (GIBCO)

PBS - As above

Trypsin - 0.025% w/v Trypsin (DIFCO)
0.0375% w/v Na₂EDTA
1% w/w Chicken serum (Flowlabs)
Made up in PBS

Freezing mixture - ES cell culture medium supplemented with
an extra 10% Foetal calf serum (GIBCO)
10% Dimethylsulfoxide

PB1 - 822 mg NaCl
21 mg KCl
300 mg Na₂HPO₄
20 mg KH₂PO₄
104 mg Glucose
4.5 mg NaPyruvate
6.2 mg Penicillin
1.4 mg CaCl₂
1 mg MgCl₂
2 ml 0.5% Phenol red
UHP water to 100 mls
7.5 General cloning techniques

Restriction enzyme digests were performed with buffers supplied by Boehringer Mannheim. The conditions recommended for DNA digestion by the supplier were followed. Digestion products were analysed by electrophoresis using agarose gels at appropriate concentrations in TAE or TBE buffers.

Specific restriction fragments were recovered from agarose gels using the Qiaex gel extraction kit II (Qiagen) or, from low melt agarose gels, using GELase digesting preparation (Epicentre Technologies).

Removal of 5' terminal phosphate groups from linearised plasmid DNA with identical or compatible termini was required to prevent recircularisation of the plasmid DNA in subsequent ligation steps. Following restriction digestion of plasmid DNA, 1 unit of calf intestinal alkaline phosphatase was added directly to the digest mix and incubated for 15 minutes at 37°C. EDTA was added to a final concentration of 5 mM and the reaction mixture heated for 10 minutes at 75°C to inactivate the phosphatase enzyme. The reaction mixture was then phenol/chloroform extracted, ethanol precipitated and the DNA resuspended in TE buffer.

Ligations were performed in the presence of 66 mM Tris.Cl pH 7.5, 5 mM MgCl$_2$, 1 mM DTT and 1 mM ATP. The reaction was left to proceed for a minimum of 2 hours at 15°C.

7.6 Preparation of competent *E. coli* cells for transformation

1. A fresh overnight culture of XL1 Blue strain cells, grown in 12.5 μg/ml tetracycline was diluted 1:100 in LB and grown at 37°C with shaking until mid-exponential phase (OD$_{600}$=0.5).
2. Cells were immediately chilled on ice and harvested by centrifugation at 4000 rpm for 7 minutes at 4°C.
3. Cells were resuspended gently in a half volume of ice cold 50 mM CaCl$_2$ and left on ice for 15 minutes.
4. Cells were centrifuged as before and gently resuspended in 1/10 volume of ice cold storage buffer (100 mM KCl, 50 mM CaCl$_2$, 10 mM KAcetate, 10% glycerol).

5. 500 µl aliquots were then dispensed into prechilled tubes and stored at -70°C.

7.7 Transformation of competent *E. coli* cells

1. 1-5 µl of ligation was added to 100 µl of thawed competent cells and left on ice for 20 minutes.
2. The cells were then subjected to a 90 second heat-shock at 42°C and returned to ice.
3. 400 µl of LB medium was added and the cells were then incubated while shaking for 30 minutes at 37°C to allow the expression of β-lactamase to confer ampicillin resistance.
4. The cells were spread onto bacterial agar plates containing 100 µg/ml ampicillin and left to grow overnight at 37°C.
5. Where blue/white selection was necessary, prior to plating the bacterial agar plates were treated with 4 µl of IPTG and 30 µl of X-Gal stock solutions.

7.8 Small scale preparation of plasmid DNA - “plasmid mini-preps”

1. The colony of interest was grown overnight in 1.5 ml LB medium in a 37°C shaker.
2. Cells were harvested by centrifugation for 3 minutes at 13,000 rpm in a 1.5 ml Eppendorf tube.
3. The pellet was resuspended in 100 µl of lysis buffer (25 mM Tris.Cl pH 8.0, 10 mM EDTA, 50 mM glucose).
4. 200 µl of freshly made 0.2 M NaOH/1% SDS was added and mixed by inversion.
5. 150 µl 3 M KAcetate/5 M acetic acid (100 ml = 60 ml 5 M KAcetate, 11.5 ml glacial acetic acid, 28.5 ml water) was added, mixed by vortex and stored on ice for 5 minutes.
6. The resulting precipitate was removed by 5 min centrifugation at 13000 rpm and 400 µl of supernatant was transferred into a clean Eppendorf tube.
7. The DNA was phenol/chloroform extracted and then precipitated by the addition of 800 µl of ethanol.
8. After 5 minutes on ice the DNA was pelleted by centrifugation at 13000 rpm for 10 minutes, washed in 70 % ethanol and then air dried.
9. The DNA was resuspended in 50 µl of TE containing 20 µg/ml of RNase A

7.9 Large scale preparations of plasmid DNA - "plasmid maxi-preps"

1. A 10 ml overnight culture of the appropriate colony was used to seed a 250 ml culture.
2. The cells were harvested by centrifugation at 4000 rpm for 15 minutes at 4°C.
3. The bacterial pellet was then resuspended in 9 ml of lysis buffer (see above).
4. 1 ml of freshly prepared lysozyme (10 mg/ml in 10 mM Tris.Cl pH 8.0) was added.
5. 20 ml of 0.2 M NaOH/1% SDS was added, mixed by inversion and stored at room temperature for 5 minutes.
6. 10 ml of ice cold 3 M KAcetate/5 M acetic acid was added and mixed by shaking.
7. The bacterial lysate was then centrifuged as before and the rotor allowed to stop without braking.
8. The supernatant was removed and filtered through glass wool.
9. The plasmid DNA was then precipitated with 0.6 volumes of isopropanol.
10. The DNA was recovered by centrifugation at 4000 rpm for 15 minutes at room temperature, washed with 70% ethanol and then resuspended in 10 ml TE buffer.

The plasmid DNA was then purified by CsCl-ethidium bromide gradient.

11. To the above DNA preparation 10.8 g of CsCl was added and dissolved followed by 650 µl of 10 mg/ml ethidium bromide.
12. This mixture was then centrifuged at 4000 rpm for 15 minutes and the supernatant loaded into a polyallomer quick-seal ultracentrifuge tube (16x76mm) and sealed avoiding air pockets.
13. Centrifugation was performed at 55,000 rpm overnight at 20°C in a Beckman NVT65 rotor.
14. The band corresponding to plasmid DNA (see Sambrook et al., 1989 for details) was removed and the ethidium bromide removed by extraction with CsCl saturated isopropanol.
15. TE buffer was added to 4 ml and the DNA precipitated with 8 ml ethanol.
16. The pellet was then washed in 70% ethanol, resuspended in 4 ml TE buffer, phenol/chloroform extracted twice, chloroform extracted and reprecipitated.
17. The pellet was washed in 70% ethanol and resuspended at 1 µg/ml in TE buffer.

7.10 Preparation of genomic DNA

1. 1 g of fresh or frozen tissue, generally liver, was homogenized in 15 ml 10 mM Tris.Cl pH 7.5, 10 mM EDTA, 100 mM NaCl.
2. 15 ml of 10 mM Tris.Cl pH 7.5, 10 mM EDTA, 100 mM NaCl, 2% SDS was added.
3. Proteinase K was added to 50 µg/ml and incubated at 37°C overnight.
4. The lysate was phenol/chloroform extracted several times and sodium acetate pH 5.2 added to 200 mM.
5. An equal volume of isopropanol was added and genomic DNA was lifted out with a pasteur pipette and redissolved in 10 ml TE buffer.
6. RNase A was added to 20 µg/ml and incubated for 30 minutes at 37°C.
7. SDS was added to 0.5%, proteinase K was added to 50 µg/ml and the mixture was incubated for 30 minutes at 37°C.
8. After phenol/chloroform extraction the DNA was precipitated as before and resuspended in 2 ml TE buffer.

7.11 Total RNA extraction from mammalian tissue

The method described in a variant of that described by Chomczynski et al. (1987). All water used in the following protocol was pretreated with dimethyl pryrocarbonate (DEPC) which chemically inactivates contaminating RNase. DEPC was added to 0.1% and dissolved by vigorous shaking. The treated water was left overnight before being autoclaved.

1. 1 g of tissue was homogenized in 9 ml of solution D (4 M guanidinium thiocyanate, 25 mM NaCitrate, 0.5% sarcosyl, 0.72% β-mercaptoethanol) for 30 seconds using an Ultra-Turrax T25 homogeniser at high speed.
2. The homogenate was stored on ice and 1 ml 2 M NaAcetate pH 4.0 was added and mixed by inversion.
3. To this, 10 ml of water-saturated phenol was added and mixed by inversion.
4. 2 ml Chloroform/isoamyl alcohol (25:1 ratio mixture) was added and the mixture shaken for 10 seconds before being transferred to a 35 ml Corex centrifuge tube.
5. The mixture was left on ice for 15 minutes and the centrifuged at 10000g for 20 minutes at 4°C in a Sorval HB-4 rotor.
6. The aqueous layer (containing the RNA) was removed from the organic layer and interface (containing the DNA and protein), mixed with an equal volume of isopropanol and left to precipitate for 1 hour at -20°C.
7. The tubes were spun as before and the pellet washed once in 70% ethanol.
8. The RNA was resuspended in 3 ml of solution D and reprecipitated with an equal volume of isopropanol for 1 hour at -20°C.
9. The tubes were spun at 10000g for 10 minutes and the RNA washed with 70% ethanol.
10. The tubes were spun for 5 minutes at 10000g to collect the pellet.
11. The pellet was resuspended in DEPC treated water and the concentration determined by absorbance at 260 nm. The ratio of absorbances at 260 nm/280 nm was used to estimate the purity of the sample. A ratio exceeding 1.8 indicated the RNA was largely uncontaminated by protein.

7.12 Poly A+ mRNA purification

1. Total RNA was resuspended in binding buffer (0.5 M NaCl, 10 mM Tris.Cl pH 7.4, 1 mM EDTA).
2. A column was prepared with oligo-dT cellulose (Boehringer Mannheim) and was equilibrated with the binding buffer.
3. After heat denaturation for 5 minutes at 65°C, the RNA was snap cooled on ice and transferred to the column. The flow through was reapplied to the column.
4. The column was then washed with several bed volumes of binding buffer.
5. The bound mRNA was eluted from the oligo-dT cellulose column using 0.5 ml aliquots of TE buffer.
6. The RNA content of the fractions collected was assessed by ethidium bromide spot assay and RNA positive fractions were precipitated and resuspended in binding buffer.
7. The column was washed thoroughly with TE buffer and re-equilibrated with binding buffer.
8. Steps 3 to 6 were then repeated to further purify the mRNA.
9. RNA positive fractions after the second round of extraction were precipitated, resuspended and combined.
7.13 cDNA synthesis

1. 1 µg of poly A+ RNA and 0.5 µg of oligo-dT primer were mixed in a total volume of 9 µl. 
2. The mixture was heat denatured for 5 minutes at 65°C and rapidly cooled on ice. 
3. The following reagents were added on ice to create a total volume of 20 µl. 
   - 1 µl RNA Guard (Pharmacia) 
   - 4 µl 5x Reaction Buffer (250 mM Tris.Cl pH 8.3, 375 mM KCl, 15 mM MgCl₂) 
   - 2 µl 0.1 M DTT 
   - 1 µl 10 mM mixed dNTPs 
   - 1 µl Superscript II reverse transcriptase (GIBCO BRL) 
4. The reaction was incubated for 1 hour at 42°C.

7.14 Polymerase Chain Reaction

In general PCR reactions were set up as follows on ice.

10 µl buffer (100 mM Tris.Cl pH 9, 500 mM KCl, 1% Triton X-100) 
6 µl 25 mM MgCl₂ 
5 µl 5 mM mixed dNTPs 
~40 pM of each primer 
~10 ng of plasmid DNA / ~500 ng of genomic DNA 
0.7 µl Taq DNA polymerase (Promega) 
Water to 100 µl

Reactions were overlayed with a drop of mineral oil to prevent evaporation and the reaction was carried out in a PCR thermocycler for typically 30 cycles. 

The conditions used were dependent on the nature of template DNA and the primers involved. In general, the oligonucleotide primers were designed with similar melting temperatures (GC content) and the anneal temperature was selected five degrees below this. Melting temperature was determined as described in Breslauer et al., (1986). Denaturation was performed at 94°C for 30 seconds or 1 minute where genomic DNA was used as the template. Extension was performed at 72°C with the incubation time varied according to the expected product, assuming 1000 nucleotides per minute synthesis.
Each PCR has differing primer annealing, melting temperature and primer specificity characteristics and the concentration of the various components can be altered to improve the overall yield, fidelity and specificity of the reaction. In general the above conditions which reflect a MgCl₂ concentration of 1.5 μM and a dNTP concentration of 250 μM were used initially and these were varied where the amplification required optimization. The concentration of dNTPs was varied between 20-250 μM and the magnesium ion concentration was adjusted from 0.5 μM to 2.5 μM in attempts to optimise the amplification in each case.

7.15 Long Range Polymerase Chain Reaction

The Expand Long Template PCR System (Boehringer Mannheim) was used in this study to amplify regions of genomic DNA up to 12 kb in length. The enzyme used is a mixture of Taq polymerase and the Pwo polymerase which has 3'-5' exonuclease activity and results in a lower mutation rate than standard Taq amplification (Barnes et al., 1994). Longer primers, 24-34 nucleotides are recommended with balanced melting temperatures near 63-68°C. These primers permit the use of higher annealing temperatures to enhance reaction specificity. This is critical as the amplification of long targets will be compromised by the preferential amplification of shorter non-specific products. Higher dNTP concentrations (350-500 μM) and magnesium ion concentrations (1.75-2.25 μM) are recommended and extension cycles are performed at 68°C. With each cycle of PCR the extension time is increased which allows the long amplification products of the reaction to fully extend when reagents are becoming limited.

The reaction was set up using two separate mixes, one with the dNTPs, template and primers and the other with the buffer and enzyme. They were prepared separately and then mixed together before the start of the reaction. This separation of the enzyme away from the primers and target DNA prevents any degradation which may occur due to the 3'-5' exonuclease activity of the Pwo polymerase. Reactions were set up as follows.

The two master mixes were added together and overlaid with a drop of mineral oil.
Master Mix 1
1.75 μl 10 μM mixed dNTPs
300 nM downstream primer
300 nM upstream primer
500 ng of genomic DNA
Water to 25 μl

Master Mix 2
5 μl 10x buffer with 17.5 mM MgCl₂
0.75 μl enzyme mix
Water to 25 μl

The reaction conditions used were typically as follows:

1x  Denature template for 2 minutes at 92°C
10x Denature template for 10 seconds at 92°C
Anneal primers at specified temperature
Elongation at 68°C

20x Denature template for 10 seconds at 92°C
Anneal primers at specified temperature
Elongation at 68°C + cycle elongation of 20 seconds/cycle

1x  Prolonged elongation time to ensure full length products.

Elongation times recommended are shown in table 7.1

<table>
<thead>
<tr>
<th>Elongation time (min)</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR fragment length (kb)</td>
<td>3</td>
<td>6</td>
<td>10</td>
<td>20</td>
</tr>
</tbody>
</table>

Table 7.1 - Elongation times recommended for long-range PCR

Primers used for the mapping of the murine BSP-1 locus and the amplification of homology arms and targeting probes are shown in table 7.2. The position of these sequences within the mouse cDNA sequence are shown in appendix A.

<table>
<thead>
<tr>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>A5 5'-dCGGAATTCCGACTGATCTGTGGGGGTGTCCTGGTGTG-3'</td>
</tr>
<tr>
<td>A3 5'-dCGAGATTCGATCGAGCAGGCGATCTCTGCTGCTGCT-3'</td>
</tr>
<tr>
<td>B5 5'-dCGGAATTCCGACTGATCGAGCAGGCGATCTCTGCTGCT-3'</td>
</tr>
<tr>
<td>B3 5'-dCGAGATTCGATCGAGCAGGCGATCTCTGCTGCTGCT-3'</td>
</tr>
<tr>
<td>C5 5'-dCGGAATTCCGACTGATCGAGCAGGCGATCTCTGCTGCT-3'</td>
</tr>
<tr>
<td>C3 5'-dCGAGATTCGATCGAGCAGGCGATCTCTGCTGCTGCT-3'</td>
</tr>
</tbody>
</table>
7.16 Rapid amplification of cDNA ends (RACE) PCR

5' sequence

This method obviates the need to screen libraries to obtain sequences upstream of a known sequence of cDNA. Reverse transcription is primed from the known sequence to create single stranded cDNA. This is then tailed with dATP nucleotides and an oligo-dT primer with a 5' extension of unique sequence is then used to prime the second strand synthesis to create double stranded cDNA. PCR is then performed from nested primers in the known sequence to a primer annealing to the 5' extension added during second strand synthesis. This allows more specific amplification as only genuine double stranded cDNAs can be amplified. The protocol used is reported in (Steel et al., 1998). A schematic diagram of the protocol is shown in figure 5.1. Primer sequences are shown in table 7.3

1. First strand cDNA was synthesised from rat poly A+ hippocampus RNA as described in section 7.13 using a reverse primer, RACE1, designed to the region of known sequence. The reverse transcription was performed in the presence of 10% DMSO to remove secondary structure.
2. The reaction was given a 50°C heat pulse at the end of the hour incubation, re-equilibrated to 42°C, fresh enzyme was added and the reaction left for another 30 minutes incubation.
3. The RNA was removed from the hybrid by the addition of 1.2 µl M NaOH followed by incubation for 10 minutes at 37°C
4. The sample was then neutralised by the addition of 6 µl 0.9 M Tris.Cl pH 7.8.
5. Microdialysis was then performed on 0.1 μm pore filters (Millipore) in TE buffer for several hours to removed excess primer.

6. The dialysed sample was then tailed with dATP using terminal transferase.
   - 11 μl recovered 1st strand cDNA
   - 3.5 μl 5x buffer (M potassium cacodylate, 125 mM Tris.Cl pH 6.6, 1.25 mg/m BSA)
   - 1 μl 25 mM CoCl₂
   - 1 μl 2 mM dATP
   - 1 μl Terminal transferase

7. The reaction was incubated for 5 minutes at 37 °C and then stopped by heat inactivation for 2 minutes at 65°C.

8. Second strand synthesis was performed using primer M2541 an oligo-dT primer with 5' extension of unique sequence.
   - 16.5 μl Tailed cDNA
   - 2 μl 10x buffer (100 mM Tris.Cl pH 7.5, 100 mM MgCl₂, 0.5 M NaCl)
   - 0.5 μl 10 mM dNTPs
   - 1 μl Oligo-dT (10 ng/μl)

This mixture was incubated for 10 minutes at room temperature to allow the primer to anneal before the addition of 1μl (2U) of klenow enzyme.

9. The reaction was incubated for 30 minutes at 37°C and then stopped by heat inactivation for 2 minutes at 95°C.

10. Microdialysis was then performed for several hours as before.

11. PCR was performed using all the recovered double stranded cDNA as template. The primers used were the second reverse primer, RACE2 and primer M2542, a primer annealing to the 5’ extension of the oligo-dT primer, M2541.

12. The products were run out on a 1% low melt agarose gel and products exceeding 400 bp in length were gel extracted.

13. PCR products were cloned into the pCR 2 vector as described in section 7.17

14. Recombinants containing 5' BSP-1 sequence were identified by colony hybridization (see section 7.21)

3' sequence

1. First strand cDNA was synthesised as described above using an oligo-dT primer, DT, tailed with an XhoI restriction site.

2. PCR was performed using a forward primer, RACE 3, designed to known BSP-1 sequence and the oligo-dT primer, DT.
3. Products were cloned into the pCR 2 vector as described in section 7.17 and confirmed as genuine by hybridization.

Complete cDNA

1. Using the sequences obtained above, primers rB5 and rB3 were designed to the extreme 5’ and 3’ ends of the BSP-1 cDNA.
2. First strand cDNA was synthesized as described above polyA+ RNA and used as the template for PCR with primers rB5 and rB3.
3. The product obtained was cloned into the pCR 2 vector

<table>
<thead>
<tr>
<th>Primer sequence</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>RACE 1</td>
<td>5’-dAGTTCTCTTGAGGGCTGGTGACA-3’</td>
</tr>
<tr>
<td>RACE 2</td>
<td>5’-dTCTTGCTCGGTACATCTCTT-3’</td>
</tr>
<tr>
<td>RACE 3</td>
<td>5’-dAAAAGAGATGAGCCCGAGCAAGAG-3’</td>
</tr>
<tr>
<td>DT</td>
<td>5’-dGTACCGCTGAG(T)15-3’</td>
</tr>
<tr>
<td>M2541</td>
<td>5’-dGGTGTGAATTCAAGCTTGG(T)12-3’</td>
</tr>
<tr>
<td>M2542</td>
<td>5’-dGGTGTGAATCTCAAGCTTGG-3’</td>
</tr>
<tr>
<td>rB5</td>
<td>5’-dCATCCTCCAGCAAGACTCAAGACA-3’</td>
</tr>
<tr>
<td>rB3</td>
<td>5’-dTGGTTTATTGACGCCCATCC-3’</td>
</tr>
</tbody>
</table>

Figure 7.3 - Primers used for the RACE cloning of BSP-1

7.17 Cloning of PCR products

Restriction enzyme sites were frequently incorporated as 5’ extensions on PCR primers. The resulting products could then be digested with the appropriate enzymes and products cloned directly into similarly prepared vectors (see 7.5).

An alternative method of PCR product cloning relies upon the non template-dependent terminal transferase activity of Taq polymerase which adds a single adenosine to the 3’ ends of the amplification products. These 3’ adenosine overhangs allows efficient cloning when ligated with a vector prepared with complementary thymidine overhangs. Amplification products were cloned into the pCR2 vector (Invitrogen) which came supplied linearised with thymidine overhangs. Ligations were performed as described in section 7.5.
7.13 Dideoxy-termination sequencing of double-stranded DNA

The sequencing protocol is based on Sambrook et al., (1989) with modifications reported in Winship et al., (1989). The sequencing reactions were performed using the Sequenase 2.0 kit (United States Biochemicals).

1. Approximately 2 μg of plasmid DNA with 20 ng (0.5 pmol) of oligonucleotide primer were denatured by incubation in 200 mM NaOH for 5 minutes at room temperature.

2. The mixture was ethanol precipitated, washed in 70% ethanol and the pellet resuspended in 10 μl of sequencing buffer (10% DMSO, 40 mM Tris.Cl pH 7.5, 20 mM MgCl₂, 50 mM NaCl).

3. To the mixture the following was added for the labelling reaction.
   - 1 μl 0.1 M DTT
   - 0.4 μl labelling mix (1.5 μM each of dGTP, dTTP and dCTP)
   - 0.5 μl α-³⁵S dATP (5 μCi)
   - 3.25 U T7 sequenase DNA polymerase (USB)
   - Water to 15.5 μl

4. The labelling reaction was incubated at room temperature for 5 minutes.

5. Four 3.5 μl aliquots from each sequencing reaction were placed in individual wells of a 96-well plate. To each of the four aliquots 2.5 μl of one particular dideoxy-termination solution (8 μM dideoxynucleotide, 80 μM of each of the other three deoxynucleotides, 10% DMSO, 50 mM NaCl) was added such that termination reactions at all four bases were represented for each plasmid.

6. After 5 minutes incubation at 37°C the reaction was stopped by the addition of 4 μl of gel running buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF).

7. The samples were denatured for 10 minutes at 80°C and then electrophoresed through a 0.2 mm, 6% denaturing acrylamide gel (National Diagnostics) in TBE.

8. After electrophoresis the gel was fixed in 10% methanol/10% acetic acid for 10 minutes and transferred to Whatman 3MM paper and dried at 80°C under vacuum.

9. Sequencing products were visualised by autoradiography using Kodak Biomax-MR film.

The oligonucleotides used for the sequencing of BSP-1 and BSP-2 cDNAs are shown in table 7.4
<table>
<thead>
<tr>
<th></th>
<th>Oligonucleotide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSP-1-5'</td>
<td>5'-dAAAAGAGATGAGCCCGAGCAAGAG-3'</td>
</tr>
<tr>
<td>BSP-1-3'</td>
<td>5'-dAGTTCTCTTGAGGGCTGGTGACA-3'</td>
</tr>
<tr>
<td>BSP-2-5'</td>
<td>5'-dTCCAGCAGTGTTGATGTC-3'</td>
</tr>
<tr>
<td>BSP-2-3'</td>
<td>5'-dTACACCAGCCTCTAGCTCA-3'</td>
</tr>
</tbody>
</table>

Table 7.4 - Oligos used for the sequencing of BSP-1 and BSP-2

7.19 Southern blot

1. After electrophoresis gels were denatured for 30 minutes in two changes of 1.5 M NaCl, 0.5 M NaOH.
2. Neutralisation was performed in several changes of 1.5 M NaCl, M Tris.Cl pH 7.4 for 45 minutes.
3. The gel was rinsed briefly in 2x SSC before being placed inverted on 2 sheets of Whatman 3MM paper supported above a bath of 20x SSC with which the ends of the paper were in contact.
4. Hybond N nylon membrane (Amersham) was cut to size, and placed on top of the gel, followed by three sheets of Whatman 3MM paper also cut to size.
5. A 5-8 cm stack of absorbant towels were then placed on the blot under a weight.
6. After transfer overnight the filter was rinsed in 2x SSC and baked at 80°C for two hours prior to hybridization.

7.20 Northern blot

1. RNA samples (20 μg) were precipitated in ethanol and resuspended in 12 μl of sample buffer (1x MOPS, 50% formamide, 18% formaldehyde).
2. Samples were then denatured for 10 minutes at 60°C and snap cooled on ice.
3. 3μl of gel loading buffer containing 1 μg/ml ethidium bromide was added.
4. Samples were electrophoresed in a 1% agarose gel containing 1x MOPS and 18% formaldehyde.
5. After electrophoresis the gel was rinsed in water before blotting.
6. The capillary blot was set up as for Southern blotting (section 7.19).
7.21 Colony lifts

1. Colonies were patched onto a fresh agar plate in a grid formation.
2. After growth the agar was chilled at 4°C before transfer.
3. A Hybond N nylon filter (Amersham) was placed onto the plate and left for 1 minute for the colonies to adhere.
4. The filter was then carefully peeled away and was denatured on Whatman 3MM paper soaked in 1.5 M NaCl, 0.5 M NaOH for 2 minutes.
5. The filters was transferred to Whatman 3MM paper soaked in 1.5 M NaCl, M Tris.Cl pH 7.4 and neutralised for 5 minutes.
6. The filters were then washed in 2x SSC to remove all cellular debris before being baked for 2 hours at 80°C prior to hybridisation.

7.22 Radiolabelling of DNA

1. 25 ng of DNA was resuspended in 19.4 µl of water, boiled for 5 minutes and then transferred to ice.
2. The following reagents were added.
   - 6 µl oligo labelling buffer (see below)
   - 1.2 µl 10 mg/ml enzyme grade BSA
   - 4 µl α-32P dCTP (40µCi)
   - 0.4 µl Klenow enzyme
3. The reaction was incubated for 45 minutes at 37°C
4. The reaction mixture was made up to 100 µl with STE buffer and spun through a 1 ml Sephadex G-50 (Pharmacia Biotech) spin column equilibrated with STE buffer to remove unincorporated nucleotides.

Oligo labelling buffer is a mixture of solutions A, B and C in a ration of 2:5:3.

Solution A - 1.21 M Tris.Cl pH 8.0
   - 0.121 M MgCl2
   - 1.74 % v/v β-Mercaptoethanol
   - 0.5 mM dATP/dTTP/dGTP

Solution B - 2 M HEPES pH 6.6 (N-[2-hydroxyethyl]-piperasine N’-(2-ethane-sulfonic acid)
Solution C - Hexa-deoxyribonucleotides dissolved to 90 O.D. units/ml
3 mM Tris.Cl pH 7.0
0.2 mM EDTA

7.23 Hybridization conditions

Northern blots, Southern blots, plaque and colony lifts were hybridized following the conditions described by Church and Gilbert, (1984). Filters were stripped by boiling 0.1% SDS.

1. Filters were pre-wet in 2x SSC and pre-hybridized for approximately an hour at 68°C.
2. The radiolabelled probe was heat denatured and added to the hybridization chamber.
3. Following overnight hybridization, non-specific hybridizing probe was removed with three successive 20 minute washes with pre-warmed wash solution.
4. After washes, filters were wrapped in Saran wrap and exposed to Kodak X-OMAT-AR film for the required period of time.

Hybridization buffer - 0.5 M Sodium phosphate pH 7.2
1 mM EDTA
7% SDS
1% BSA Fraction V

Wash Solution - 40 mM Sodium phosphate pH 7.2
1 mM EDTA
1% SDS

7.24 Screening of bacteriophage lambda libraries

1. Plating cells were initially prepared. A 1:50 diluted culture of an XL1-Blue strain of *E. coli* was grown to mid-log phase (OD$_{600}$=0.5) in LB supplemented with 0.2% maltose and 10 mM MgSO$_4$.
2. The cells were chilled on ice, centrifuged at 4000 rpm for 5 minutes and then resuspended in an equal volume of ice-cold 10 mM MgSO$_4$.
3. 1.3 ml cells were infected with 50,000 recombinant phage diluted in SM and incubated for 15 minutes at 37°C.
4. 13 ml of 0.7% agarose in LB, supplemented with 10 mM MgSO₄ (melted and warmed at 45°C) was added, mixed gently by inversion and then poured evenly over a dried 15 cm agar plate.

5. Once set, the plates were inverted and incubated at 37°C until the appearance of the bacterial lawn (7-10 hours). Plaques were grown until they were touching each other but not confluent. The plates were then chilled at 4°C.

6. Replica lifts were performed as described in section 7.21. The first lift was left for 1 minute and the second lift was left for 2 minutes.

7. Hybridisation was performed in bags as described in section 7.23.

7.25 In vivo excision of λZAPII

1. Approximately 10⁵ plaque forming units in 100 µl of SM buffer were added to 200 µl XL1 blue plating cells and 1 µl of ExAssist helper bacteriophage (Stratagene). This was mixed and incubated for 15 minutes at 37°C.

2. 3 ml of 2X TY medium was added and cells were incubated for 3 hours in a shaking 37°C incubator.

3. Cells were lysed by 20 minute incubation in a 70°C water bath and cellular debris removed by centrifugation for 5 minutes at 4000 rpm.

4. 5 µl of supernatant was used to infect 200 µl of SOLR strain plating cells (made as in 1) and incubated for 15 minutes at 37°C.

5. The resulting cells were plated on agar plates containing ampicillin and incubated overnight at 37°C. Colonies contained the excised cDNA insert within pBluescript KS plasmid vector.

7.26 Riboprobe synthesis

1. 2 µg of plasmid DNA was digested with an appropriate restriction enzyme to create a linear DNA template which would allow transcription of the insert in the appropriate orientation from the RNA polymerase binding sites in the polylinker.

2. After digestion, the template was phenol/chloroform extracted, ethanol precipitated and resuspended in 2 µl of water.

3. The following was added

   0.75 µl 200 mM DTT
   0.75 µl 2 mg/ml BSA
   2.25 µl 3.3 mM ATP/GTP/CTP
   0.5 µl RNA Guard (Pharmacia)
Solution C - Hexa-deoxyribonucleotides dissolved to 90 O.D. units/ml
3 mM Tris.Cl pH 7.0
0.2 mM EDTA

7.23 Hybridization conditions

Northern blots, Southern blots, plaque and colony lifts were hybridized following
the conditions described by Church and Gilbert, (1984). Filters were stripped by
boiling 0.1% SDS.

1. Filters were pre-wet in 2x SSC and pre-hybridized for approximately an hour at
68°C.
2. The radiolabelled probe was heat denatured and added to the hybridization
chamber.
3. Following overnight hybridization, non-specific hybridizing probe was removed
with three successive 20 minute washes with pre-warmed wash solution.
4. After washes, filters were wrapped in Saran wrap and exposed to Kodak X-
OMAT-AR film for the required period of time.

Hybridization buffer - 0.5 M Sodium phosphate pH 7.2
1 mM EDTA
7% SDS
1% BSA Fraction V

Wash Solution - 40 mM Sodium phosphate pH 7.2
1 mM EDTA
1% SDS

7.24 Screening of bacteriophage lambda libraries

1. Plating cells were initially prepared. A 1:50 diluted culture of an XL1-Blue
strain of E. coli was grown to mid-log phase (OD\textsubscript{600}=0.5) in LB supplemented
with 0.2% maltose and 10 mM MgSO\textsubscript{4}.
2. The cells were chilled on ice, centrifuged at 4000 rpm for 5 minutes and then
resuspended in an equal volume of ice-cold 10 mM MgSO\textsubscript{4}.
3. 1.3 ml cells were infected with 50,000 recombinant phage diluted in SM and
incubated for 15 minutes at 37°C.
4. 13 ml of 0.7% agarose in LB, supplemented with 10 mM MgSO4 (melted and warmed at 45°C) was added, mixed gently by inversion and then poured evenly over a dried 15 cm agar plate.
5. Once set, the plates were inverted and incubated at 37°C until the appearance of the bacterial lawn (7-10 hours). Plaques were grown until they were touching each other but not confluent. The plates were then chilled at 4°C.
6. Replica lifts were performed as described in section 7.21. The first lift was left for 1 minute and the second lift was left for 2 minutes.
7. Hybridisation was performed in bags as described in section 7.23.

7.25 In vivo excision of λZAP II

1. Approximately 10⁵ plaque forming units in 100 μl of SM buffer were added to 200 μl XL1 blue plating cells and 1 μl of ExAssist helper bacteriophage (Stratagene). This was mixed and incubated for 15 minutes at 37°C.
2. 3 ml of 2X TY medium was added and cells were incubated for 3 hours in a shaking 37°C incubator.
3. Cells were lysed by 20 minute incubation in a 70°C water bath and cellular debris removed by centrifugation for 5 minutes at 4000 rpm.
4. 5 μl of supernatant was used to infect 200 μl of SOLR strain plating cells (made as in 1) and incubated for 15 minutes at 37°C.
5. The resulting cells were plated on agar plates containing ampicillin and incubated overnight at 37°C. Colonies contained the excised cDNA insert within pBluescript KS plasmid vector.

7.26 Riboprobe synthesis

1. 2 μg of plasmid DNA was digested with an appropriate restriction enzyme to create a linear DNA template which would allow transcription of the insert in the appropriate orientation from the RNA polymerase binding sites in the polylinker.
2. After digestion, the template was phenol/chloroform extracted, ethanol precipitated and resuspended in 2 μl of water.
3. The following was added

   0.75 μl 200 mM DTT
   0.75 μl 2 mg/ml BSA
   2.25 μl 3.3 mM ATP/GTP/CTP
   0.5 μl RNA Guard (Pharmacia)
6.25 µl α-35S UTP (250 µCi)
1.5 µl 10x Transcription buffer (400 mM TrisHCl pH 8.0, 60 mM MgCl₂, 100mM DTT, 20 mM spermidine hydrochloride)
1 µl T7/SP6/T3 RNA polymerase as appropriate

4. The reaction was incubated for 1 hour at 37°C.
5. 2 x 1 µl DNase was added and incubated for 10 minutes at 37°C to digest the template.
6. The reaction was made up to 100 µl with water and phenol/chloroform extracted.
7. The probe was purified away from unincorporated ribonucleotide on a 1 ml Sephadex G50 column.
8. 4 µl of M DTT was added and the probe stored at -70°C.

7.27 *In situ* hybridization

**Section Pretreatment**

Unless stated otherwise, all treatments were performed in coplin jars at room temperature.

1. 10 µm horizontal cryo-sections were cut and mounted on TESPA (3-aminopropyltriethoxysilane) coated slides. Coating of slides was performed as described in Rentrop *et al.*, (1986).
2. The sections were fixed on in 4% w/v paraformaldehyde in PBS for 15 minute on ice.
3. Sections were rinsed in two changes of PBS for 5 minutes.
4. Sections were treated with 20 µg/ml proteinase K in 50 mM Tris.Cl pH 7.4, 5 mM EDTA for 1 minute.
5. The proteinase was then inhibited by a 5 minute wash in 0.2% w/v glycine in PBS.
6. The sections were rinsed in two changes of PBS for 5 minutes.
7. Sections were fixed again in 4% paraformaldehyde in PBS for 10 minutes on ice.
8. After rinsing in PBS the sections were acetylated in a bath of 0.25% acetic anhydride, 0.1 M triethanolamine pH 8.0, 0.8% w/v NaCl for 10 minutes.
9. Sections were then dehydrated in 50% ethanol in 0.8% w/v NaCl for 1 minute, 70% ethanol in water for 5 minutes and then quickly through 85%, 95% and 100% ethanol in water for 1 minute each.
10. Sections were rinsed again in 100% ethanol before being immersed in chloroform for 5 minutes.
11. After two rinses in 100% ethanol and a rinse in 95% ethanol in water the sections were allowed to air dry before hybridization.

Hybridization

1. 400 μl of hybridization buffer was placed over the sections and Parafilm coverslips were added.
2. Prehybridization was carried out in a perspex box humidified with a tissue soaked in 50% formaldehyde, 5x SSC. The box was sealed with tape and incubated at 55°C for 2 hours.
3. The riboprobe was prepared, as described in section 7.26, and diluted in hybridization buffer such that the specific activity was 1000 cpm/μl. DTT was added to a final concentration of 10 mM.
4. Probe solutions were denatured at 80°C for 5 minutes.
5. 300 μl of probe solution was applied over the sections and new coverslips were added avoiding air bubbles.
6. Hybridization was carried out overnight at 55°C in a sealed, humidified perspex box.

Hybridization buffer -
- 50% deionised formamide (Kodak)
- 0.3 M NaCl
- 20 mM Tris.Cl pH 8
- 5 mM EDTA
- 10 mM NaPO₄ pH 8
- 10% dextran sulphate
- 1x Denhardt’s solution
- 0.5 mg/ml yeast RNA

Prior to use, this mixture was degassed in a vacuum drier.

Washing of hybridized sections

1. Coverslips were removed by placing the slides vertically in a bath of 5x SSC, 10 mM DTT for 20 minutes at 55°C.
2. A high stringency wash was performed in 50% formamide, 2x SSC, 0.1 M DTT for 30 minutes at 65°C.
3. 3x Low stringency washes were performed in 0.5 M NaCl, 10 mM Tris.Cl pH 7.5, 5 mM EDTA for 10 minutes at 37°C.
4. RNase A was added to a concentration of 20 μg/ml and incubated for 30 minutes at 37°C.
5. A 10 minute low stringency wash preceded another high stringency wash for 30 minutes at 65°C.
6. Two final washes, one 2x SSC and then 0.1x SSC were performed for 10 minutes each at room temperature.
7. The sections were then dehydrated through increasing concentrations of ethanol in 0.3 M ammonium acetate before air dying.

**Autoradiography**

All manipulations were carried out under safelight conditions

1. Slides were initially exposed to Kodak Biomax film to provide an initial indication of the pattern and level of expression. For high resolution detection, NTB2 photographic liquid emulsion (Kodak) was used.
2. The emulsion was melted in a 43°C water bath and diluted 1:1 with water.
3. Once melted the emulsion was poured into a slide mailer and bubbles removed by dipping blank slides.
4. Each hybridized slide was dipped evenly into the emulsion vertically for 1-2 seconds, drained vertically and then wiped to remove excess emulsion.
5. Slides were stored horizontally overnight in a light tight box.
6. Once dry, the slides were transferred into a plastic slide box containing silica gel and sealed with tape. The box was wrapped in aluminium foil and stored at 4°C for exposure.
7. After exposure the slides were allowed to equilibriate to room temperature for several hours.
8. The slides were placed in D19 developer (Kodak) for 5 minutes at 15°C
9. After a brief rinse with water the slides were fixed in Amfix (Champion Photochemistry) for 5 minutes at 15°C.
10. The slides were then rinsed in a large volume of running water for 10 minutes.

**Counterstaining and Mounting of Sections**

1. The wet sections were placed in 1% neutral red stain and stained for 5 minutes.
2. The sections were destained in several changes of ethanol and treated with Histoclear (National Diagnostics).
3. Sections were mounted in DPX and covered with a glass coverslip.
4. Silver grains were visualised by bright field high power microscopy or by dark field low power microscopy.

7.28 Reverse Transcription PCR analysis

1. 1 µg poly A+ RNA from whole brain and liver (kidney as appropriate) was reverse transcribed as described, using oligo-dT primer.
2. 1 µl of the 20 µl reaction was used as the template for PCR using HPRT specific primers (5’-dGGGGGCTATAAGTTCTTTGCTGAC-3’, 5’-dCTTTTCCACTTT-CGCTGATGACAC-3’) for 20 cycles.
3. Products were compared and quantities of starting templates normalised.
4. Multiplex PCR was performed using the above quantities of cDNA with HPRT primers and specific serine protease primers (see table 7.5) for 20, 25 and 30 cycles. (94°C 1 minute, 55°C 1 minute, 72°C 30 seconds)
5. Products were electrophoretically separated on 1.5% agarose and transferred to nylon membrane by capillary blotting.
6. Filters were hybridized with labelled PCR product corresponding to thrombin, u-PA, plasminogen or HGF, as appropriate.

<table>
<thead>
<tr>
<th>Protease</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prothrombin</td>
<td>5’-dCAGGCCTGACATCAACTCCACCAC-3’</td>
</tr>
<tr>
<td></td>
<td>5’-dCACCACGCGCCCTCCTCATCC-3’</td>
</tr>
<tr>
<td>u-PA</td>
<td>5’-dTGGGGCTGCTTGCACTGTTCG-3’</td>
</tr>
<tr>
<td></td>
<td>5’-dTGAGCTGTGCTTGACTGTCG-3’</td>
</tr>
<tr>
<td>Plasminogen</td>
<td>5’-dGGAGTACTGTGAGATTCGTTCCGTC-3’</td>
</tr>
<tr>
<td></td>
<td>5’-dGTGGTGAAGCACCAGGGG-3’</td>
</tr>
<tr>
<td>HGF</td>
<td>5’-dAGGCAGAGGAGGACAAGCAGACAAGCAG-3’</td>
</tr>
<tr>
<td></td>
<td>5’-dTTCAGCCCCAGCACAATTCAGA-3’</td>
</tr>
</tbody>
</table>

Table 7.5 - Primers used for RT-PCR analysis
7.29 Culturing of embryonic stem cells

All tissue culture manipulations were performed inside a laminar flow sterile hood (Jencons Nuaire). All flasks and pipettes were sprayed with 70% industrial methylated spirits prior to use. Cells were grown in a humidified incubator (Heraeus) with 6% CO₂ at 37°C. Maintenance of ES cell cultures was carried out in tissue culture grade plastic flasks (Corning). All flasks were gelatinised by covering the base with a 0.1% solution of gelatin in PBS for 5 minutes after which it removed by aspiration. Tissue culture medium, trypsin solution and PBS were pre-warmed to 37°C before use.

ES cells were cultured in media containing Leukaemia Inhibitory Factor (LIF). LIF was prepared by transient transfection of murine or human LIF expression plasmids in COS-7 cells using the method described in Smith et al., (1991). Serial dilutions of LIF were tested on ES cells in 24 well plates. The minimum dilution required to keep ES cells undifferentiated was determined and a working concentrations corresponding to 100 times this value was used. LIF used in this study was prepared by D. Colby and D. Rout, Centre for Genome Research.

Confluent cultures were passaged to fresh flasks at a density of approximately 1:10. The medium was removed and the cells rinsed in PBS, covered with trypsin and incubated at 37°C for up to 5 minutes. The cells were agitated to assist their lifting from the base and their disaggregation until a suspension was achieved. The trypsin was neutralised by the addition of fresh medium and pipetted to create a single cell suspension.

7.30 Transfection of embryonic stem cells

1. Four large 175 cm² flasks of near confluent ES cells were trypsinised and harvested by centrifugation for 5 minutes at 1300 rpm.
2. The cells were resuspended and counted in a haemocytometer.
3. 1x10⁸ cells were resuspended in 600 μl PBS
4. 150 μg of DNA in 100 μl of PBS was added and mixed gently with the cells.
5. The mixture was loaded into a cuvette and electroporated at 0.8 V, 3.0 μF with a time constant between 0.1 and 0.2 seconds.
6. Cells were diluted into 20 ml of media and plated at 5x10⁶ cells/9 cm dish.
7. Selection was added 48 hours after electroporation.
8. Colonies were picked with a Gilson 20 μl pipette tip loaded with 10 μl of trypsin solution to help disperse the colony. The cells were then transferred into a well on a 96 well plate containing 50 μl of trypsin solution and incubated for up to 5 minutes.
9. Cells were agitated by vigorous aspiration using a pipette and transferred to individual wells on a 24 well plate containing 1.5 ml of media.
10. Cells were grown to near confluency and then split for freezing and analysis of genomic DNA.

7.31 DNA extraction from embryonic stem cell colonies

1. Media was aspirated from a confluent well on a 24 well plate.
2. 500 μl of lysis buffer (100 mM Tris.Cl pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 100 μg/ml proteinase K) was added and incubated overnight at 37°C.
3. The lysate was transferred to an Eppendorf tube and 500 μl isopropanol was added and mixed.
4. After 5 minutes rotating, the DNA was collected by centrifugation for 5 minutes at 13000 rpm.
5. The pellet was then washed in 70% ethanol and resuspended in 100 μl TE buffer.

7.32 Freezing embryonic stem cells

24 well plate method
The method described is after Ure et al., 1996.

1. 24 well plates were checked daily and when the majority of the wells had reached up to 70% confluency the cells were frozen directly on the plate.
2. Media was aspirated and replaced with 400 μl of freezing mixture.
3. Plates were packed in polystyrene boxes packed with styrofoam and transferred to a -20°C freezer to allow gentle freezing.
4. The following day, plates were transferred to a -80°C freezer for long-term storage.
Freezing vials of cells

1. Confluent cells in a 25 cm² flask were trypsinised and resuspended in 10 ml culture media to achieve a single cell suspension.
2. Cells were harvested by centrifugation for 5 minutes at 1200 rpm.
3. Cells were resuspended in 1.5 ml of freezing mix and 0.5 ml aliquots of the suspension were transferred to individual cryotubes (Nunc).
4. Cryotubes were placed in a -80°C freezer for overnight storage before being transferred to a liquid nitrogen cell bank (Cryotechnics).

7.33 Thawing of embryonic stem cells

24 well plate method

1. Plates were removed from the -80°C freezer and gently warmed by hand.
2. While still frozen, 2 ml of warmed media was added.
3. When the ice crystals has melted, the well contents was aspirated leaving approximately 0.5 ml of media covering the recovering cells.
4. 2 ml of fresh media was added and the plates incubated overnight.
5. The following day, the entire media was replaced.

These precautions are necessary to remove traces of DMSO from the media. The entire well contents cannot be removed as the cells have a tendency to detach from the well bottom when recently thawed.

Thawing vials of cells

1. Frozen vials were removed from storage and warmed through in a 37°C water bath.
2. When all ice crystals had melted, the cells were transferred to 10 ml of fresh media.
3. Cells were harvested by centrifugation for 5 minutes at 1200 rpm and resuspended in fresh media.
4. Cells were then recentrifuged, resuspended and then transferred to a gelatinised 25 cm² flask.
5. The following day, the media was replaced.
Thawing of cells was generally performed last thing at night and the media was replaced early the following morning.

7.34 Production of chimeras

Blastocyst injection and implantations were performed by Mrs J. Ure, Centre for Genome Research.

1. Blastocysts were flushed from C57BL/6 mice on the fourth day of pregnancy using PB1 medium supplemented with 10% foetal calf serum.
2. Blastocysts were transferred to ES cell media and incubated in a humidified incubator with 6% CO₂ at 37°C to allow unexpanded embryos to expand.
3. When required for injection, blastocysts were transferred to small hanging drops of PB1 medium supplemented with 10% foetal calf serum on a siliconised coverslip suspended over a manipulation chamber flooded with liquid paraffin (Boots).
4. Trypsinised cells were placed in a hanging drop and the chamber was refrigerated for at least 10 minutes.
5. Injections were performed with a rounded holding pipette of internal diameter 20 μm, and a heat polished injection pipette of internal diameter 15 μm. These were formed using an electrode puller (Campden Instruments), and cut to size and heat-polished using a microforge (Alcatel).
6. Holding and injection pipettes were operated by manipulators (Narishige) and connected to injectors (Narishige) by paraffin-filled plastic tubing.
7. The micromanipulation assembly was mounted underneath a IMT2 image corrected microscope (Olympus)
8. Between 10 and 15 cells were injected into each blastocysts and operated embryos were allowed to recover before being transferred to recipient mothers.
9. Pseudopregnant recipients were of an F1 stock from matings between C57BL/6 and CBA strains, previously mated 2.5 days previously with vasectomised DBA males.
10. Up to 12 embryos were transferred unilaterally to the uterus of the recipient females.

The contribution of the injected ES cells in the resulting pups was assessed by coat colour. The appearance of sandy, agouti or white patches amongst the host black hairs indicated successful incorporation of ES cell derived cells.
Germline transmission was ascertained by test breeding of chimeric males with C57BL/6 strain mice. The generation of light coloured offspring indicated successful transmission.

7.35 Genotyping of mice

1. 0.5 cm tail was digested overnight at 55°C in 100 mM Na\textsubscript{2}EDTA, 50 mM Tris.Cl pH8.0, 0.5% SDS.
2. Lysates were extracted with phenol/chloroform
3. DNA was precipitated with the addition of 50 µl NaAcetate pH 5.2 and 1 ml ethanol, removed with a gilson pipette tip and dissolved in 100 µl TE.
4. 20 µl samples were digested with the appropriate restriction enzyme for Southern analysis.
5. 0.5 µl samples were used as the template for 30 cycles of PCR using the primers A5 (see table 7.2), SC3 (5'-dCCCGGGATCATATCAGGCTGGGGCAGCTG) and ΔD3 (5'-dCCGAGGCCTC-CAGATCATGCATCCCATACAA) (94°C 1 minute, 60°C 1 minute, 72°C 45 seconds)

7.36 Brain dissections

Rats were sacrificed by carbon dioxide asphyxiation and subsequent cervical dislocation. After decapitation, the top of the skull was exposed and bone plates overlying were removed to allow removal of the brain in an intact state. The hippocampus were exposed by peeling back the two cortical hemispheres from the midline and all corpus striatum tissue was removed. The hippocampus could then be detached and removed from the brain.

Cortex and striatum, hippocampus, cerebellum, and remainder of brain were collected and immediately frozen in liquid nitrogen before being stored at -70°C.

Brains from which sections were to be cut, were frozen by burial of the intact brain in powdered dry ice. The use of liquid nitrogen causes the tissue to shatter during sectioning.
Using watchmakers forceps, embryos were liberated from the decidua, the yolk sac pierced and the embryos separated from the placenta. Embryos were transferred to PBS prior to staining.

1. Embryos were rinsed once in 0.1 M sodium phosphate buffer at room temperature
2. Embryos were fixed in 0.1 M sodium phosphate pH 7.4, 2 mM MgCl₂, 5 mM EGTA for 15-20 minutes.
3. 3x 15 minute washes were performed in 0.1 M sodium phosphate pH 7.4, 2 mM MgCl₂, 0.02% v/v Nonoidet-P40, 0.01% v/v deoxycholate.
4. Embryos were stained overnight at 37°C in the solution below.

0.1 M Sodium phosphate pH 7.4
5 mM Potassium ferrocyanide
5 mM Potassium ferricyanide
1 mg/ml X-Gal
2 mM MgCl₂
0.02% v/v NP40
0.01% v/v deoxycholate
Chapter VIII

Discussion

8.1 Serine proteases within the brain

Brain serine proteases are thought to mediate processes involved in neurite outgrowth, cell migration and synaptic remodelling. Proteolytic enzymes are considered central to the degradation of ECM components that such structural changes demand. Other proteases mediate effects on neuronal function directly through interaction with specific cell surface receptors. Serine proteases are further implicated in neuronal function in the modulation of neurotransmitter receptors and in the processing of growth factors and neuropeptides. Furthermore, this diverse family of proteolytic enzymes also has important roles in brain dysfunction both in the pathogenesis of Alzheimer's disease and in neuronal response to injury.

Despite the potential importance of serine proteases within the CNS, the spectrum of family members expressed within this brain region remains unknown. This thesis has focused on the characterisation of serine proteases expressed within the hippocampus, a region of the brain implicated in learning and memory processes and in which synaptic plasticity is well characterised. The characterisation of hippocampal serine proteases may identify key regulators of the structural plasticity that accompanies long-term changes in synaptic efficacy. The identification of a serine protease restricted in its expression to the hippocampal formation may also allow the potential for more specific transgenic investigations into the role of this brain region in cognitive processes.

8.2 Serine proteases identified within the rodent hippocampus

To address the spectrum of serine proteases expressed in the rat hippocampus a PCR approach was used. The homology that exists between family members was used to design degenerate primers for the selective amplification of serine proteases from rodent hippocampal cDNA. In total, 10 serine proteases was isolated, the majority of which represent known transcripts previously uncharacterised within the brain. These are discussed individually below.
t-PA was isolated from the screen at high frequency, representing approximately 60% of cloned sequences. This concurs with previous studies reporting abundant expression of this serine protease within the mouse hippocampal formation (Sappino et al., 1993; Tsirka et al., 1995). In rat brain a more restricted expression pattern was reported (Dent et al., 1993; Ware et al., 1995), suggesting a possible species difference. Northern analysis of rat brain regions in this study, however, indicates widespread expression in rat in agreement with the mouse expression pattern. The t-PA gene is induced as an immediate early gene following processes contingent upon neuronal activity, such as in memory tasks and synaptic plasticity (Qian et al., 1993; Seeds et al., 1995). This plasminogen activator is considered to initiate a cascade of proteolytic events leading to the degradation of extracellular matrix components necessitated by structural change. The rapid induction of this species, coupled with the instability associated with the t-PA transcript, may explain the lack of consistency in expression studies. The isolation of this expected serine protease at high frequency validates the degenerate PCR approach used.

RNK-Met-1, a chymotryptic protease of rat natural killer cells, represents a quarter of the serine protease sequences amplified in this screen. This protease has not been characterised in the brain before. Northern analysis on total RNA in this study clearly demonstrated the presence of transcripts in the cortex and hippocampus of rat brain, indicating a moderate level of expression. In situ hybridization analysis confirmed this expression pattern and revealed the predominant site of expression to be neuronal. The contribution of this species from contaminating lymphocytes in the brain vasculature can thus be ruled out.

Both the hippocampus and the cerebral cortex retain the capacity for structural plasticity in adults. Expression of RNK-Met-1 within these areas suggests a role for this protease in this structural reorganisation, possibly in cascades of extracellular matrix degradation. The role of structural modification within the hippocampus is considered important for aspects of memory consolidation whereas morphogenesis within the cortex may reflect memory storage.

The protease is of chymotryptic specificity and preferentially cleaves after methionine residues. The amyloid precursor protein is aberrantly processed after a methionine residue to yield the amino terminus of the β-amyloid fragment. RNK-Met-1 is therefore a candidate β-secretase and may play a role in the pathogenesis of
Alzheimer's disease. Furthermore, abundant expression in the entorhinal cortex is apparent, a brain region susceptible to early dysfunction in Alzheimer's disease.

Elastase IV was amplified in this screen and expression was detected in the hippocampal CA fields, the cerebellum and specific cortical layers. This serine protease was originally isolated from pancreas and has a well characterised role, along with other elastases, in digestive processes. Elastase activity has been detected in microglia conditioned media (Nakajima et al., 1992b) but activity was attributable to neutrophil elastase. Expression of this pancreatic enzyme within brain has not been reported previously. Possible roles include degradation of extracellular matrix components, mediating aspects of the structural plasticity that have been documented within the hippocampus, cortex and cerebellum. In vitro studies suggest elastases are capable of inactivating PN-1, so this enzyme may have a role in the regulation of protease inhibition (Nick et al., 1990). With respect to the processing of APP, elastase has activity towards the carboxy terminus of the β-amyloid domain and thus may be considered a potential γ-secretase (Evin et al., 1995). Elastase has activity against plasminogen releasing a low molecular weight form of activated enzyme (Machovich and Owen, 1989). The presence of this processed species in microglia conditioned media has been reported, although the elastase responsible was considered neutrophil elastase (Nakajima et al., 1993). A potential role as a plasminogen activator is thus suggested, however the expression of plasminogen within the central nervous system is unclear (see below).

A serum calcium-decreasing factor, caldecrin, has been isolated from porcine and rat pancreas. The caldecrin cDNA sequence indicates that this enzyme is likely to be identical to elastase IV (Tomomura, 1995). The serum calcium-decreasing activity was found to require activation of the pro-enzyme sequence but was not dependent upon proteolysis. This study highlights another function for proteases and indicates possible functions for certain serine proteases, independent of proteolysis.

C2. The amplification of complement component, C2, is in agreement with previous studies reporting the local expression of the complement system within brain. Specifically C2 transcripts were previously detected by poly A+ Northern analysis (Johnson et al., 1992). In this present study, in situ hybridization analysis revealed the majority of C2 transcripts to be localised within the brain vasculature. At the resolution obtained in this analysis, expression in the brain parenchyma is not conclusive. However, complement components have been immunologically detected
in both neurons and microglia (Johnson et al., 1992; Pasinetti et al., 1992). C2 is one of three proteases involved in cascades of protein activation, leading to the formation of the lytic complex for the lysis of cellular antigens. Various complement derived activation peptides are also involved in inflammatory response. The local expression of complement components suggests the brain is able to mount an immunological response following injury, disease and infection. Accordingly, an induction of complement components is seen in Alzheimer's disease and components are localised to neuritic plaques, neurofibrillary tangles and degenerating neurites (McGeer et al., 1989).

The expression of complement component in the normal brain may suggest roles for these proteins in processes other than immune response. In support of this, mice deficient in complement component C5 have altered glutamate receptor responses (Pasinetti et al., 1996).

CTRL, a chymotryptic serine protease, has not previously been reported in brain. Expression in pancreas has previously been demonstrated by Northern analysis, suggesting a role in digestive processes (Reseland et al., 1997). In this present study low level expression was detectable in most brain regions by the more sensitive in situ hybridization analysis indicating a role for this protease in brain function. Chymotryptic enzymes are implicated in the pathogenesis of Alzheimer's disease, releasing the amino terminus of the β-amyloid domain. The protease inhibitor domain of APP also has activity against chymotrypsin indicating a role for chymotryptic enzymes in normal APP function within the brain.

Chymotrypsin B, another pancreatic enzyme, has been isolated from both rat and mouse hippocampus. In situ hybridization studies indicated a very low level of expression with hybridization signals barely above background. The expression of this species in brain has not been reported previously.

Proteinase 3. Microglia are thought to express the neutral proteases of polynuclear lymphocytes including the serine proteases neutrophil elastase, cathepsin G and proteinase 3. The isolation of proteinase 3 from brain tissue is therefore not unexpected. In situ hybridization revealed diffuse, low level expression. The presence of this enzyme in brain has not been detected before. Microglial proteases are thought to be involved in the phagocytic role associated with activated microglia following infection and neuronal cell death.


Hageman factor (blood clotting factor XII), was isolated from mouse hippocampus in this screen. The expression pattern of this protease has not been investigated so the local synthesis of this species is not conclusive. The presence of a blood clotting factor could indicate contamination of the hippocampal material used with components of brain vasculature. Further expression analysis is necessary to determine whether this protease is a genuinely expressed within the brain parenchyma. A previous study detected Hageman factor in the brain, but immunoreactivity was localised to the residual serum of capillaries (Yasuhara et al., 1994). However, some immunoreactivity was associated with the dendritic plaques of Alzheimer’s disease, perhaps indicating local brain synthesis. Hageman factor is implicated in complement activation and since components of this system are expressed in brain tissue, local synthesis of this protease may be expected. Another potential role for this protease in brain is the activation of latent growth factors. A serine protease highly homologous to Hageman factor was found to activate hepatocyte growth factor (Shimomura et al., 1995).

8.3 Alternative targets for the plasminogen activators in brain

8-PA within the brain is thought to initiate a proteolytic cascade, most probably mediating structural remodelling consequent upon neuronal activity. In support of this view, t-PA transcription is upregulated following neuronal activity (Qian et al., 1993; Carroll et al., 1994) and mice lacking t-PA are deficient in the long-term maintenance of LTP (Huang et al., 1996; Frey et al., 1996). Surprising, clones encoding plasminogen, the only recognised in vivo target for t-PA, were not obtained in the degenerate PCR screen.

The lack of amplification could represent primer bias although examination of the plasminogen sequence reveals the presence of conserved binding sites. Other previously identified serine proteases with well recognised roles in the CNS were also not identified, suggesting a degree of primer bias in the screen. However, a comparative RT-PCR analysis using specific primers confirmed the presence of thrombin and u-PA transcripts but failed to detect plasminogen transcripts in either rat or mouse brain.

A variety of previous studies have examined the presence of plasminogen within the central nervous system and conflicting results have been obtained. The presence of
plasminogen transcripts and protein within the mouse hippocampus has been reported
(Tsirka et al., 1997) and plasminogen protein has been immunologically detected as a
secreted product of microglia (Nakajima et al., 1992a). Analysis of plasminogen
deficient mice suggests a functional role in brain, as such mice are resistant to
excitotoxic cell death. However, differences in phenotype between plasminogen
deficient and t-PA deficient mice have been reported: t-PA null mice have attenuated
microglia activation whereas activation in plasminogen null mice is as wild type
(Tsirka et al., 1997). This study suggests plasminogen independent function for t-PA.

The absence or low level of plasminogen expression in brain suggests that the major
target for the abundant PAs is unlikely to be plasminogen. If, as widely assumed, the
role of t-PA is the initiation of extracellular proteolysis, in the absence of plasminogen,
how might this be achieved?.

Firstly, in vitro studies have suggested that PAs may cleave components of the ECM
directly, in the absence of plasminogen (Quigley et al., 1987; Keski-Oja and Vageri,
1982)

Secondly, other proteases besides plasminogen may function as PA targets in brain. A
brain specific plasminogen variant could possibly mediate the effects of the PAs in
brain. However, preliminary experiments addressing this possibility have failed to
detect such a molecule (not presented). Alternatively, some of the brain serine
proteases identified in this study may serve as alternative substrates for PA activation,
although the restricted substrate specificity of PAs indicates this to be unlikely.

Thirdly, the PAs themselves have also been shown to interact with specific cell surface
binding sites, suggesting a role in signal transduction. PAs could thus mediate their
effects on proteolysis perhaps through transcriptional upregulation of brain serine
proteases rather than through direct proteolytic activation.

Finally, the PAs have also been implicated in the activation of hepatocyte growth factor
(HGF) and macrophage stimulating protein (MSP). Both HGF and MSP have a
plasminogen-like domain organisation and require cleavage for activation (Donate et
al., 1994). These growth factors are serine protease family members but are not
proteolytically active so cannot contribute directly to the proposed ECM degradation,
accompanying structural change.
However, HGF has been found to have neurotrophic effects on neuronal cultures (Hamanoue et al., 1996), enhancing cell survival and neurite outgrowth. This growth factor could thus act downstream of t-PA, mediating aspects of structural plasticity. RT-PCR analysis, in this study, robustly identified HGF transcripts in brain, indeed the levels in brain were found to be similar to those in liver. This is consistent with previous reports of the local synthesis of this growth factor within the central nervous system (Thewke and Seeds, 1996). Indeed, evidence exists for a functional coupling with t-PA, as antibodies against t-PA attenuated HGF activation by brain homogenates (Thewke and Seeds, 1996) and both HGF and t-PA were upregulated following ischemic injury (Honda et al., 1995).

Although uncharacterised within the CNS, MSP is an effector of macrophage chemotaxis and phagocytosis (Skeel et al., 1991) and thus would be expected to modulate microglial function. This growth factor could mediate the effects of t-PA on microglial activation (see above).

8.4 Two novel trypsin-like serine proteases

Two novel members of the serine protease family have been characterised in this study. Both BSP-1 and BSP-2 represent members of the trypsin subfamily of serine proteases as indicated by conserved sequence motifs. Both enzymes are considered to be activated by trypsin-like proteases due to the presence of lysine and arginine residues in the putative pro-enzyme activation sequence.

BSP-2 is most similar to mast cell protease-6 although the homology is not sufficient for functional conclusions to be drawn. The isolated cDNA for this novel protease lacks an initiating methionine in-frame and this anomaly was found to be conserved in the isolated genomic sequences from two strains of rat. BSP-2 is not thought to represent a pseudogene as the sequence contains no other frame-shift mutations and the all structural motifs are conserved within the cDNA sequence. The most likely explanation of this anomaly is a robust sequencing artefact although this demands further investigation.

BSP-2 is expressed within the hippocampus and the cerebral cortex. The expression of a protease in these regions, as with RNK-Met-1, suggests a role in structural plasticity. Expression is also detectable within certain hindbrain nuclei, specifically within the reticular nuclei and the motor trigeminal nucleus. The role of a specifically expressed protease in these regions is unclear.
BSP-1 is most similar to the kallikrein family of serine proteases involved in the processing of bioactive molecules. The homology that exists, however, is low compared to relationships within this family indicating the functional significance of this homology is unclear. BSP-1 expression is restricted to the CA fields of the hippocampus and the entorhinal cortex. Such a restricted expression pattern is suggestive of an important role in the functional specialisation of the hippocampal formation.

BSP-1 represents the rat orthologue of the mouse serine protease, neuropsin, which was characterised during this study (Chen et al., 1995). The expression pattern of neuropsin was also reported and transcripts were localised within the CA fields of the hippocampus. Expression was also reported in the lateral amygdaloid nucleus and the septal nuclei although expression within the entorhinal cortex was not reported. Broadly, the expression patterns of neuropsin and BSP-1 correspond, thus supporting BSP-1 as the rat orthologue of neuropsin. BSP-1 expression within the amygdaloid nucleus was not investigated in this study. Transcription of neuropsin was found to be upregulated following kindling indicating a potential role in hippocampal activity-dependent processes as its expression pattern suggests (Okabe et al., 1996).

8.5 BSP-1 deficient mice

To address the possible role of BSP-1 in brain function, a line of mice deficient in BSP-1 has been derived. A region of exon III, corresponding to the catalytic aspartic acid residue, was deleted and replaced with a reporter/selection cassette through homologous recombination in ES cells. Lines of mice have been derived from two clones, however no homozygous animals are available at present.

The deletion of a region of exon III, containing a portion of the gene essential for enzymatic function, precludes any functionality being restored at the locus through illegitimate splicing events or read-through translation. The analysis of mice deficient in BSP-1 function may provide insights into the role of this specifically expressed protease. This BSP-1 transgenic line represents one of the first experiments to cause a genetic lesion uniquely within the hippocampal formation. As such, correlations between hippocampal dysfunction and any behavioural phenotype seen can be drawn with confidence.
The expression of BSP-1 within the hippocampal formation suggests that this protease may mediate aspects of structural plasticity reflecting memory consolidation within this brain region. BSP-1 deficient mice may therefore be expected to have deficits associated with the long-term consolidation of memory. Similar to mice deficient in t-PA, BSP-1 deficient mice may have selective deficits for the long-lasting phases of LTP. Alternatively, since hippocampal gene expression reflects aspects of developmental expression, BSP-1 may play a role in embryogenesis (see below). As such, a developmental phenotype may be expected.

The two lines of BSP-1 transgenic mice are being expanded and are currently being bred onto the 129 and C57BL/6 genetic backgrounds. If homozygotes are viable, phenotypic assessment will include detailed anatomical analysis, electrophysiological studies of synaptic transmission and plasticity and an assessment of cognitive function. Furthermore, the involvement of this serine protease in the modulation of susceptibility to excitotoxic cell death, warrants an investigation.

The inconsistencies of genetic background in transgenic studies involving cognition, have confused analysis of phenotype, particularly with respect to learning abilities (Gerlai, 1996). To standardise such analyses, the guidelines of the Banbury Conference on Genetic Background in Mice, (1997) will be adopted.

Other future work necessitates the raising of antibodies against BSP-1 protein. These are necessary for confirmation of deficient expression in homozygous mutant mice. Immuno-techniques will also allow localisation of the wildtype BSP-1 protein to be determined and provide a confirmation of the expression pattern at the protein level. Furthermore, protein biochemical analysis may permit the processing of BSP-1 and APP interactions to be investigated.

8.6 Directing gene expression within the hippocampus

Directed gene expression within the hippocampus is a goal of current molecular neurobiological research into learning and memory. The study of transgenic animals with perturbed hippocampal synaptic plasticity allows an assessment of correlated cognitive decline. However, such approaches are fraught with difficulty. The expression of mutant alleles in other brain regions and throughout development may contribute to the phenotype. Thus inferences concerning the role of specific hippocampal processes in cognitive function cannot be drawn with confidence.
The characterisation of a gene restricted in its expression to the hippocampal formation is of considerable utility for transgenic experimentation. The transcriptional control elements of such a gene could be used to direct transgene expression uniquely within the hippocampus. Alternatively, the endogenous hippocampus specific gene can be targeted with a transgene downstream of an IRES sequence. Dicistronic translation of the endogenous transcript would yield transgene expression specifically within this brain region. The latter approach has been used in this study to express a β-galactosidase gene downstream of an IRES sequence under the control of BSP-1.

β-galactosidase expression, as determined by X-gal staining, is active from the locus and abundant expression is seen at 12.5 dpc embryos. Although only a preliminary analysis has been carried out, reporter expression appears to be widely distributed within the developing nervous system, possibly within the brain and the retina. Clear expression was also detected delineating the somites which may reflect expression in extending peripheral nerves. This unexpected pattern of expression suggests a role for BSP-1 in neuronal process outgrowth during development. Correlated expression, widespread within the developing nervous system and within the hippocampal formation in adult adds credence to the hypothesis that BSP-1 may play a role in structural plasticity within this brain region.

Analysis of adult heterozygous animals is a priority to assess whether the β-galactosidase reporter is faithfully reflecting the endogenous expression pattern of BSP-1. An initial indication is provided by the published expression pattern of neuropsin, the BSP-1 orthologue. Expression by in situ hybridization was first detected in the developing brain at 12 dpc (Suzuki et al., 1995), in agreement with the reporter expression seen in BSP-1 transgenic animals. One potential problem with the resulting transgenic mice is that the neo selection cassette remains within the locus. The presence of strong transcriptional control elements, i.e. the MCI promoter, within the locus may alter the transcriptional regulation of the endogenous gene leading to silencing or ectopic expression. The reported selection cassette used in this study has, however been used in previous studies and β-galactosidase expression has been shown to mimic the endogenous targeted gene (e.g. Nehls et al., 1996).

In terms of the utility of this locus for the expression of other transgenes, this apparent widespread developmental expression is disadvantageous. The use of this locus to perturb neuronal function through the expression of dominant-negative transgenes or
inhibitors, may lead to developmental abnormalities which could contribute to the phenotype. The contribution of perturbed hippocampal processes to the phenotype seen would be unclear.

Modern molecular-genetic techniques now allow temporal control of the expression of transgenes. The tetracycline inducible system has been used in transgenic context to express the constitutively active αCaMKII transgene in a temporally and spatially restricted manner (Mayford et al., 1996). A modified system, the reverse tetracycline activator allows expression of the transgene only in the presence of tetracycline (Gossen et al., 1995). Another inducible system, the ecdysone transcriptional activation system of Drosophila has been adapted for efficient function in mammalian cells (No et al., 1996). The application of these systems with the BSP-1 locus would allow the adult expression pattern to be exploited without developmental complications.

The BSP-1 locus could then be used for the expression of any transgene to modify neuronal function exclusively within this brain region. Alternatively, the locus could be used to express the Cre recombinase. Crossing such mice with strains with loxP sites flanking specific genes, could allow region specific knock-outs as has been achieved with the NMDA receptor subunit, NR1, within the hippocampus (Tsien et al., 1996a). Another use of this locus would be the directed expression of genes conveying conditional cell ablation, such as E. coli nitroreductase (Clark et al., 1997b), for a new generation of specific lesion studies.

8.7 Concluding remarks

A panel of ten serine proteases has been isolated from rodent hippocampus, many of which have not been characterised within the central nervous system before. Northern and in situ analysis reveals overlapping expression patterns for these proteases with the majority expressed within the hippocampus and the cortex. These areas are associated with structural plasticity and these proteases may therefore mediate aspects of this morphogenesis. Two novel members of this gene family, BSP-1 and BSP-2 were isolated and full length cDNAs obtained. Sequence information indicates that both are proteases of tryptic specificity.

BSP-1 expression was found to be restricted to the hippocampal formation and thus is of interest with regards the function of this brain region. The expression profile is also
of potential use for transgenic experimentation as directed gene expression within the hippocampus is the goal of current molecular-genetic research into the function of this brain region in cognition. To investigate the utility of BSP-1 in directed transgene expression, a reporter under the control of an IRES sequences has been introduced into the locus. Preliminary studies indicate reporter expression is active from the locus and that BSP-1 is expressed abundantly during embryogenesis. Analysis of reporter expression throughout development and within adult brain will allow an assessment of the use of this locus for directed transgene expression. To assess BSP-1 function, a transgenic strain of mice has been derived lacking an essential region of the gene. Phenotypic analysis of homozygous mice deficient in BSP-1 function may unveil the role of this serine protease \textit{in vivo}.

Serine proteases are abundantly expressed within the brain and may serve to mediate the structural reorganisation which correlates with learning and memory events. The application of transgenic technologies will allow the involvement of these proteases in such cognitive abilities to be addressed. The BSP-1 transgenic strain may serve as a useful tool in the exploration of long-term memory function.


212

Appendix A

Map of neuropsin cDNA (mouse BSP-1)

GACACCGAGGGAAGTGCTCCGGGCTTCCACCGAGTCCGAGTGACCCCGCCCCTTGCATTCTGGAAGGTGAGGCC

Exon 1

AGAGGTCCCCAGACACGGACCTCAGGCGCAGGGAGGTCCCTTTCTCTGAGCCCAGGACCCTCCCACCCCCAGGCTCACA

Exon 1

TTTTTCTCTCAAGATCTTAAGCTCCTTTAGCTCTCTCCCAGGACGTTGGAGTCACAGCCTAGATCTTT

Exon 1

TCCTCAATCTCAAGTGGGGAACATCTTTATATATTGCTGGATCCCAACACAGCAAGCTCCTCCACCTCCCCCCACCTAAAACT

Exon 1

GGGATCTAGAAGCTGCCCCATTAGCTTTCTCAGGCCCCCTAGCTCATCTCCCAGCAGTCTCCGCGAAGACCAT

Exon 1

TCCTCCCCAGTCCCAGACAACCAGATCTCAGGCTCTCTGCTTCAGATCTTT

Exon 1

GCCACCATGGGACGGCCCCCCACCCTGTGCAATCCAGCTGGATCCTTCTGCTTCTGTTCATGGGAGCGTGGGCAGGGCT

Exon 1

CACCAGAGCTCAGGGCTCCAAGATCCTGGGAAGGTCGAGAGTGTATACCCCACTCCCAGGCTTGGCAGGCAGCCTTGTTCC

Exon 2

AGGGCAGAGGACGCTGATCTGTGGGGGTGTCCTGGTTGGAGACAGATGGGTCCTCACGGCAGCCCACTGCAAAAAACAGAAG

Exon 2

GCCACCATGGGACGGCCCCCCACCCTGTGCAATCCAGCTGGATCCTTCTGCTTCTGTTCATGGGAGCGTGGGCAGGGCT

Exon 2

AGGGCAGAGGACGCTGATCTGTGGGGGTGTCCTGGTTGGAGACAGATGGGTCCTCACGGCAGCCCACTGCAAAAAACAGAAG

Exon 2

Gin Gly Glu Arg Leu Ile Cys Gly Gly Val Leu Val Gly Asp Arg Trp Val Leu Thr Ala Ala His Cys Lys Lys Gin Lys

Exon 2

Gin Gly Glu Arg Leu Ile Cys Gly Gly Val Leu Val Gly Asp Arg Trp Val Leu Thr Ala Ala His Cys Lys Lys Gin Lys

Exon 2

Thr Arg Ala Gin Gly Ser Lys Ile Leu Glu Gly Arg Gly Cys Ile Pro His Ser Gin Pro Trp Gin Ala Ala Leu Phe

Exon 2

Thr Arg Ala Gin Gly Ser Lys Ile Leu Glu Gly Arg Gly Cys Ile Pro His Ser Gin Pro Trp Gin Ala Ala Leu Phe

Exon 2

Met Gly Arg Pro Pro Pro Cys Ala Ile Gin Pro Trp Ile Leu Leu Leu Phe Met Gly Ala Trp Ala Gly Leu

Exon 1

55 primer

P55

C5 primer

P53

214
Neuropsin nucleotide and translated amino acid sequences are shown (Chen et al, 1995). Position of intron/exon boundaries are indicated. The position of PCR primers used for genomic mapping, targeting vector construction, screening and genotyping are shown.