The molecular characterisation of a gene trap integration into the Aminopeptidase O (*ApO*) gene

Richard Alan Axton

PhD

The University of Edinburgh

2006

I declare that the work presented in this thesis is my own unless otherwise stated

Richard Alan Axton

Date: 26 March 2007

Acknowledgements:

I would like to thank my supervisors Dr Lesley Forrester and Professor John Mullins for their support and guidance throughout my studies. I thank the University of Edinburgh for awarding me a staff scholarship to cover my fees and funding from Procter and Gamble, The British Heart Foundation and Leukaemia Research Fund. The John Hughes Bennett Laboratory has been a fun place to work and I would like to thank all the past and present members for their contributions. My family have been very supportive and patient with me and I thank Helen, Lewis and Corinne for this.

Abstract

The protease aminopeptidase O (*ApO*) was isolated during a gene trap screen designed to identify novel developmentally regulated genes involved in cardiovascular development. The aims of this thesis were to characterise the gene trap integration, determine the full-length open reading frame (ORF) and to investigate any functional domains that the gene contained. Additionally, germline transmission of the gene trap demonstrated expression in the vasculature of mice. Work was undertaken to investigate if there was a phenotype in the homozygous gene trap mice with particular emphasis on whether this gene was involved in the process of new blood vessel formation (angiogenesis).

Initially, 5' RACE (Rapid Amplification of cDNA Ends) generated a 206 bp sequence that aligned with expressed sequence tag data to build a contig with a partial ORF of 177 amino acids. Additional rounds of 5'RACE generated a full length open reading frame of 823 amino acids. The nucleotide and protein sequence are highly conserved between mouse, rat and human. The gene encodes a catalytic domain characteristic of the M1 family of aminopeptidases and contains a C-terminal armadillo repeat sequence and a putative nuclear localisation signal (NLS). The NLS was experimentally confirmed by transfection of eGFP-*ApO* constructs demonstrating that the signal localised the fusion protein to the nucleoli.

The gene trap ES cells were injected into blastocysts and transmitted through the germline. Generation of homozygous animals resulted in normal, healthy, viable mice with no apparent phenotype. Expression analysis by *lacZ* staining and RNA *in situ* indicated that *ApO* expression was restricted to the cardiovasculature of embryos and adult tissues. The possibility that the *ApO* may be involved in the process of angiogenesis was investigated using the aortic ring assay. Northern blot analysis has shown that there are tissue specific isoforms and RT-PCR analysis has revealed that the catalytic domain and nuclear localisation signal can be removed by alternative splicing. An antibody was raised against purified full length murine APO protein and the

immunopurified antibody tested by Western blot, immunoprecipitation and immunohistochemistry.

Table of contents

Chapter 1:

Chapter 3: Molecular characterisation of the gene trap integration GT411

Chapter 4: Expression analysis and phenotypic analysis of mice carrying the GT411 gene trap integration.

Chapter 5: Cellular localisation studies of the mouse aminopeptidase O protein

Chapter 6: The expression and purification of fusion proteins to raise an antibody against aminopeptidase O

Figures and tables

Chapter 1: Introduction

Chapter 3: Molecular characterisation of the gene trap integration GT411

Chapter 4 Expression analysis and phenotypic analysis of mice carrying the GT411 gene trap integration.

Chapter 5 Cellular localisation studies of the mouse aminopeptidase O protein

Chapter 6 : The expression and purification of fusion proteins to raise an antibody against aminopeptidase O

Appendix: Supplementary data

List of abbreviations

ACE	Angiotensin Converting Enzyme
Ang I	Angiotensin I
Ang II	Angiotensin II
Ang III	Angiotensin III
Ang IV	Angiotensin IV
APA	Aminopeptidase A
APN (CD13)	Aminopeptidase N
ApO	Aminopeptidase O gene
APO	Aminopeptidase O protein
ARM	Armadillo repeat
AT1R	Angiotensin II type 1 receptor
AT2R	Angiotensin II type 2 receptor
bFGF	Basic fibroblast growth factor
β-gal	β-galactosidase

βgeo	Fusion of *lac Z* and neomycin gene
BP	Base Pair
CP	Capillary Plexus
CT	C-terminal
DA	Dorsal Aorta
DPC	Days Post Coitus
EC	Endothelial cell
ECM	Extra Cellular Matrix
ENU	Ethylnitrosourea chemical mutagen
eGFP	enhanced Green Fluorescent Protein
ERK	Extracellular Regulated Kinase
ES cell	Embryonic Stem cell
EST	Expressed Sequence Tag
FB	Forebrain
FISH	Fluorescent *In Situ* Hybridisation
Gapdh	Glyceraldehyde 3-phosphate dehydrogenase

GST	Glutathione sepharose beads
GT	Gene trap
GT	Gene trap cell line 411
Ha	Hamster
HIS-tag	Histidine Tag
HT	Heart
Hu	Human
HUVEC	Human Umbilical Vein Endothelial Cell
IP	Immuno Precipitation
IV	Intersomitic Vessels
KB	Kilo Base
kDa	Kilo Dalton
Lac Z	Reporter gene
LB	Limb Bud
MMP	Matrix Metallo Protease

MEROPS	Protease database
MetAP2	Methionine Aminopeptidase 2
Mu	Murine
Neo	Neomycin resistance gene
NLS	Nuclear Localisation Signal
nt	nucleotide
NT	N-terminal
ORF	Open Reading Frame
PA	Poly Adenylation
PCR	Polymerase Chain Reaction
PILSAP	Puromycin insensitive leucyl specific aminopeptidase
PLAP	Placental leucine aminiopeptidase
RA	Retinoic Acid
RACE	Rapid Amplification of cDNA Ends
RAS	Renin angiotensin system
SA	Splice Acceptor

SD	Splice Donor
SHR	Spontaneous Hypertensive Rat
SMART	Simple Modular Architecture Research Tool
TIMP	Tissue Inhibitor Metallo Protease
TSP1	Thrombospondin
uPA	Urokinase Plasminogen Activator
UTR	Untranslated Region
VEGF	Vascular Endothelial Growth Factor
VEGFR2 (Flk1)	Vascular Endothelial Growth Factor Receptor 2
YS	Yolk Sac

Chapter 1:

Introduction

Chapter 1

1.0 Overview

During a screen designed to identify novel regulated genes in cardiovascular development, the gene trap GT411 ES clone was isolated. This thesis describes the characterisation of this gene trap integration and the subsequent analysis of the mouse gene that had been trapped. The gene was originally identified to be a novel aminopeptidase but during the course of the research its human orthologue was cloned and named Aminopeptidase O (*APO*) (Diaz-Perales et al., 2005). Knowledge about *APO* is still in its infancy and although the gene has been characterised with regard to structure, little is known about how it functions physiologically.

From protein sequence analysis, several domains have been predicted. Firstly, it contains a protease domain indicating that it is a member of the M1 family of aminopeptidases. These enzymes are involved in a number of important biological processes including angiogenesis, maintenance of blood pressure and inflammatory pathways (Haeggstrom, 2004; Mitsui et al., 2004; Sato, 2003). Secondly, prediction of an armadillo repeat structure suggests that *ApO* may be implicated in the transport of proteins to the nucleus (Malik et al., 1997). Furthermore, characterisation of a putative nuclear localisation signal sequence (NLS) using eGFP-*ApO* fusions demonstrated that the protein can localise to the nucleoli.

Germline transmission of the GT411 ES cells demonstrated *LacZ* reporter expression in developing embryonic and adult vasculature. This is important because the gene trap vector marks endogenous gene expression. Understanding the function of *ApO* therefore might provide insight into vascular biology in both normal developmental and pathological conditions. These could include processes such as angiogenesis and the maintenance of blood pressure as well as its deregulation in diseases such as cancer and hypertension.

1.1 Proteolytic enzymes

Proteolytic enzymes are a group of structurally and functionally diverse proteins that have the common ability to hydrolyse peptide bonds (Neurath, 1999). The specific cleavage of substrates mediated by proteases is vital to every cell and organism. Proteolytic processing events regulate the activity of many proteins, for example, the regulation of transcription factor activity and the appropriate intra or extra cellular localisation of many proteins (Lamkanfi et al., 2006; Pickart, 2004; Tuteja, 2005). Proteases control essential biological processes such as DNA replication, cell cycle progression, cell proliferation, differentiation and migration, immunological reactions, ovulation, angiogenesis, haematopoiesis and apoptosis (Adams, 2003; Haeggstrom, 2004; Roy et al., 2006; Yamazaki et al., 2004). Given the importance of these proteolytic processing events in the control of these biological processes, it is not surprising that deficiencies or alteration in the regulation of enzymes underlie important human pathologies such as arthritis, cancer, neurodegenerative and cardiovascular disease (Egeblad and Werb, 2002; Hol et al., 2005; Middleton et al., 2004; Oudit et al., 2003).

1.1.1 Enzyme classification

Several databases, such as the MEROPS and the Degradome have been established and serve as a resource for information on proteases. The MEROPS database (http://www.merops.ac.uk) includes details of approximately 3000 individual proteases from multiple organisms. They are essentially classified into six groups which are defined by the proteases catalytic activity. They are namely the aspartic, cysteine, metallo, serine, threonine and a group which are of unknown catalytic type. Within each group the enzymes are further classified into families on the basis of significant similarities between protein sequences in the part called the active site that is most directly responsible for the protease activity (Rawlings et al., 2002). The group of metalloproteases, for example, consists of approximately

75 different families and each family can have a considerable number of members.

The Degradome database (www.uniovi.es/degradome) specifically includes details on peptidases present within the human, rat and mouse genomes with particular emphasis on accession numbers and chromosome locations. The proteases are grouped into five different classes with regard to their mechanism of catalysis (Puente et al., 2003) and classified into families in accordance with the MEROPs database.

1.1.2 Aminopeptidase O (*APO*)

The human gene for *APO* was cloned in 2005 with one paper published to date (Diaz-Perales et al., 2005). It was cloned during a bioinformatic analysis specifically targeted at identifying genomic sequences containing novel aminopeptidases. The full length cDNA for this enzyme was predicted by comparison with the structure of other aminopeptidases and then verified by RT-PCR experiments on a brain cDNA library. Purification of the recombinant putative catalytic domain of APO and incubation with fluorogenic synthetic substrates demonstrated that it exhibits significant proteolytic activity with an N-terminal arginine residue, and to a lesser extent, aparagine. The enzymatic activity was strongly inhibited by o-phenanthroline, a known metalloprotease inhibitor, but was not inhibited by 4-(2-aminoethyl)-benzenesuphonyl fluoride or E-64 which are common inhibitors for serine and cysteine proteases respectively (Diaz-Perales et al., 2005).

The expression of APO in heart and testis, tissues in which the renin angiotensin system (RAS) plays important roles, prompted Diaz-Perales to investigate the putative involvement of APO in the processing of peptides in this pathway (Mitsui et al., 2004). Recombinant human APO was incubated with angiotensin III (Ang III) and although the reaction proceeded slowly it did generate angiotensin IV.

1.1.3 General introduction of aminopeptidases

Aminopeptidases are enzymes that catalyse the sequential removal of amino acids from the amino terminus of proteins or peptides (Taylor, 1993). They are metalloproteases because they require a metal ion usually zinc but colbalt is used by some in their catalytic active site such as MetAP2 (Griffith et al., 1997). They are widely distributed in the plant, bacterial, fungi and animal kingdoms (Taylor, 1993). The enzymes have broad substrate specificity and in mammals perform a wide variety of biological processes involving angiogenesis, blood pressure regulation and inflammatory immunological responses (Haeggstrom, 2004; Sato, 2004; Wright et al., 1990).

1.1.4 The M1 aminopeptidases

The M1 aminopeptidases contain the conserved catalytic domain with a zinc-binding site (HEXXHX-E18). Essentially, based upon protein structural motifs, they can be organised into four groups, namely the membrane bound, Armadillo repeat (ARM), cytosolic and secreted aminopeptidases (Figure 1.1a and 1.1b).

1.1.5 Armadillo (ARM) domain aminopeptidases

It has been proposed that the ARM aminopeptidases are a sub-family of the M1 aminopeptidases which diverged from an evolutionary ancestor to give rise to this specific group. They are aminopeptidase O, leukotreine A4 hydrolase, aminopeptidase B and aminopeptidase B-like (Diaz-Perales et al., 2005).

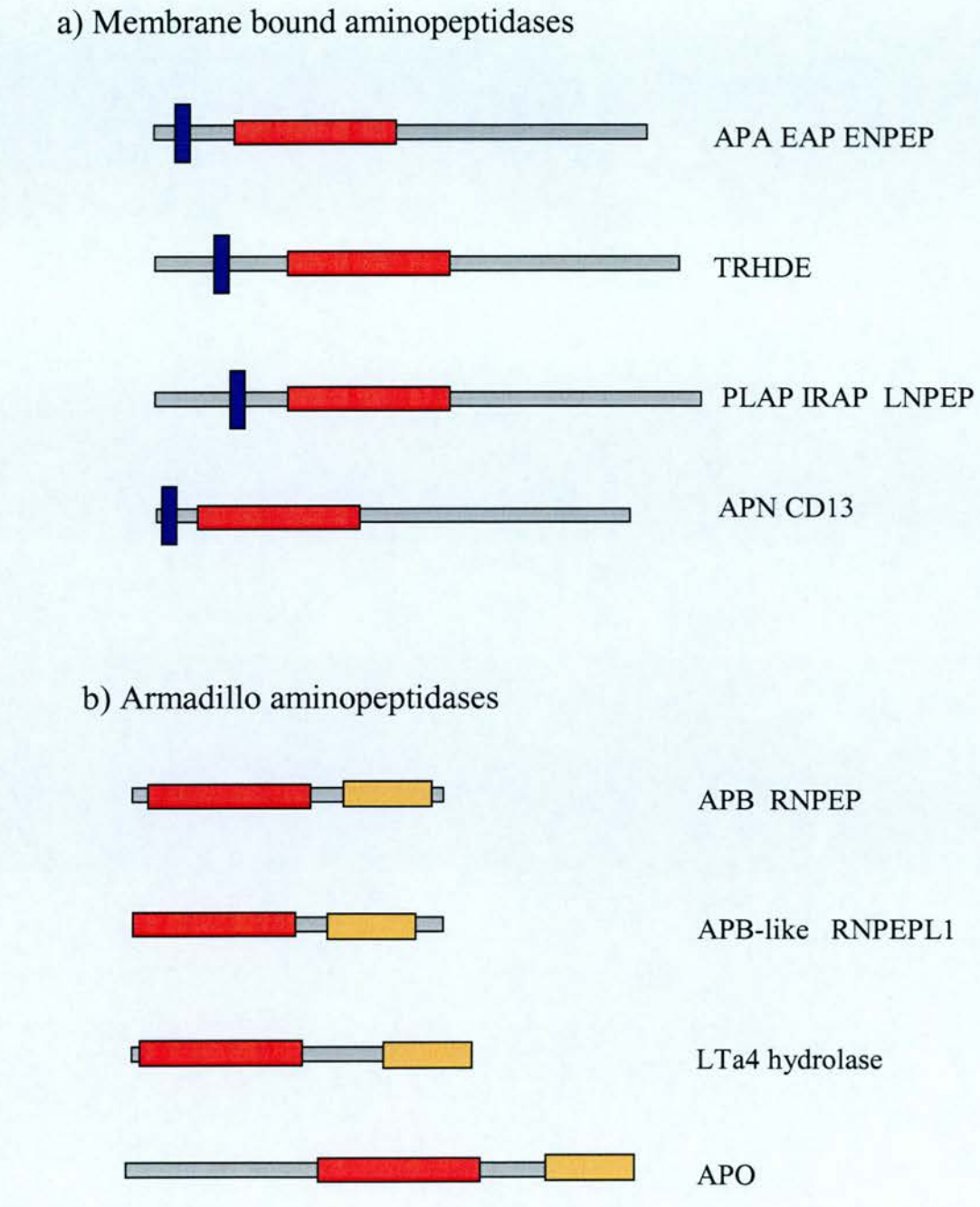
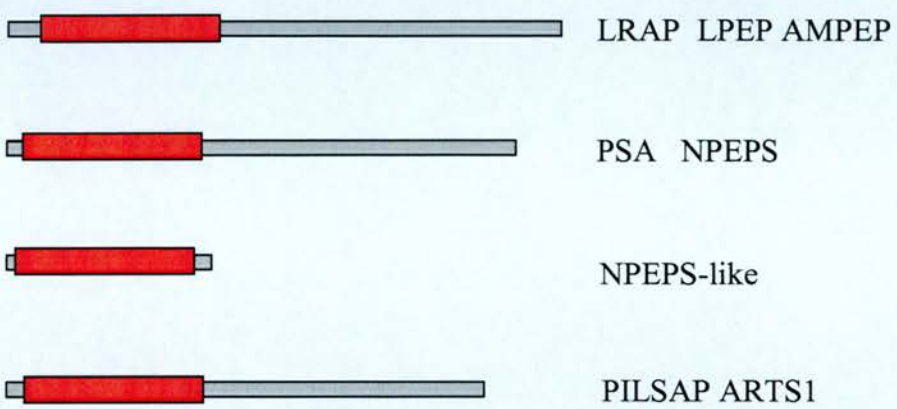


Figure 1.1a: Schematic diagram of the M1 family of aminopeptidases organised by domain structure.

c) Cytosolic aminopeptidases



d) Secreted aminopeptidases

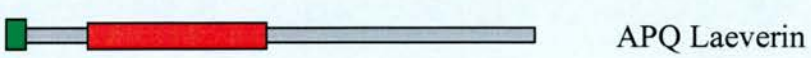


Figure 1.1b: Schematic diagram of the M1 family of aminopeptidases organised by domain structure.

Signal peptide Catalytic M1 domain

1.1.6 ARM domain

The ARM aminopeptidases contain a C-terminal armadillo repeat structure. This is a tandem copy of a degenerate imperfect repeat within a protein sequence motif that forms a conserved three-dimensional structure. The sequence motif is typically about 42 amino acids and a single ARM motif consists of three α helices (H1, H2 and H3) which interact extensively with one another to form a right handed super-helix of helices. Structurally they have been likened to HEAT repeats which consist of 2 helices (H1 and H2) (Andrade et al., 2001). The armadillo repeat sequence was first identified in *Drosophila melanogaster* segment polarity gene known as Armadillo (Peifer et al., 1993), but this structure is also present in other proteins such as its orthologue β-catenin (Huber et al., 1997), the plant plakophilins (Klymkowsky, 1999) and the family of nuclear transport proteins called the importins (Malik et al., 1997).

Importin α, a protein involved with nuclear transport, contains ARM repeats which bind to the nuclear localisation signals of other proteins to transport them to the nucleus through the nuclear pore complex (Conti and Izaurralde, 2001; Gorlich, 1997).

Although no specific function of the ARM repeat has been attributed to these enzymes, the ARM repeat structure allows for protein-protein interactions and proteins that have this structure are involved with processes such as nuclear cytoplasmic transport, maintenance and control of the cytoskeleton (Coates, 2003). The LTA4 hydrolase ARM repeat fold has been confirmed by crystallography studies (Thunnissen et al., 2001). It has recently been published that LTA4 hydrolase can localise to the nucleus of type II alveolar epithelial cells in rat, mouse and human cells (Brock et al., 2005).

1.1.7 Bi-functionality of catalytic sites

Some proteases exhibit the ability to cleave different substrates and this has led to the proposition that the catalytic activities are exerted via specific but overlapping sites. Proteases that have this ability are termed bi-functional and include enzymes such as the hepatitis C virus NS2-3 protease and the aminopeptidase leukotriene A4 hydrolase (Haeggstrom, 2004; Lorenz et al., 2006). In the case of LTA4 hydrolase the enzyme exhibits epoxide hydrolase activity for LTA4 and aminopeptidase activity (Orning et al., 1994). LTA4 hydrolase is involved in inflammatory processes as a key mediator of the 5-lipoxygenase pathway where it hydrolyses the peptide LTA4 into LTB4, a molecule that stimulates the adherence and aggregation of leukocytes to the endothelium (Haeggstrom, 2004).

Analysis of the LTA4 hydrolase protein sequence confirms that it contains the characteristic M1 aminopeptidase catalytic domain (HEXXH18E). The purified enzyme was able to hydrolyse short peptides with a preference for those with N-terminal Ala and Arg residues in standard aminopeptidase assays (Orning et al., 1994). Although it has not been experimentally verified, it is generally assumed that the aminopeptidase activity is involved in the processing of peptides related to inflammation and host defence (Haeggstrom, 2004).

From structural studies it has been determined that the two separate catalytic sites require the use of the zinc ion and some common residues, but also have distinct residue usage, such as Asp 375 in the epoxide hydrolase reaction and Glu 296 and Tyr 383 for the aminopeptidase reaction (Kester and Matthews, 1977; Rudberg et al., 2004).

1.1.8 Domain structure and chaperone folding models

Many membrane and secretory proteins have pro-domains at their N or C-termini that can undergo post translational proteolytic cleavage. It has been demonstrated that these domains act as intramolecular chaperones responsible for the correct folding of their cognate catalytic domain. Deletion of the N-terminal propeptides of the enzymes subtilisin (Ikemura et al., 1987) and alpha lytic protease (Baker et al., 1992) and the C- terminus in the protein intestinal enzyme sucrase-isomaltase (Jacob et al., 2002) results in an inactive enzyme (Silen et al., 1989; Zhu et al., 1989).

Aminopeptidase A is a type II integral membrane glycoprotein consisting of 945 amino acids. Proteolytic fragmentation studies of purified pig APA showed that the protein existed *in vivo* as two polypeptides of 107 and 45 kDa (Hesp and Hooper, 1997). N-terminal sequencing of the 107 kDa fragment demonstrated that it was the N-terminal extracellular catalytic domain of APA from residues 43-602 whereas the 45 kDa fragment corresponded to the C-terminus of APA from residues 603 to 942. In addition the rat hippocampus has a short variant of APA produced by alternative splicing lacking the C-terminal domain. This variant displayed no APA activity when produced in COS-1 cells (Lee et al., 2000). Similarly, a recombinant form of the N-terminus in the absence of the C-terminus did not have enzymatic activity and remained blocked in the endoplasmic reticulum (Ofner and Hooper, 2002a). This suggests that the C-terminus may be required for the correct folding of APA. A study was performed with the production of recombinant protein N-terminal fragment 1-598 and a separate clone coding the C-terminal fragment 595 to 945 and the proteins were expressed either separately or in trans. It was found that co-expression of the N and C-terminal fragments resulted in an active enzyme localised at the membrane.

A similar study has been conducted on the membrane bound M 1 aminopeptidase APN (CD13). It also has a C-terminal domain that promoted maturation and cell surface expression indicating that this mechanism of

protein folding is common to membrane bound aminopeptidases (Rozenfeld et al., 2004).

1.2 Vasculature

1.2.1 Vascular endothelium

The vascular endothelium is the layer of squamous epithelial cells that is in direct contact with the blood. It forms a continuous lining of the cardiovascular system (i.e. the heart and all the vessels contained within it) and is ideally positioned to monitor bloodborne or locally produced stimuli. The vascular system consists of a massive network of tubular channels carrying nutrients and waste products between blood and tissue. Vessel walls of arteries and veins consist of specialised connective tissues with elastin and collagen as their major constituents. Major arteries carry blood direct from the heart and have thick elastic walls composed mainly of elastin, while the smaller arteries are rigid containing more collagen. In contrast, veins have thinner walls, less dense connective tissue and a larger lumen than their arterial counterparts. They have fewer smooth muscles and are equipped with valves to prevent back flow of blood (Figure 1.2). Capillaries connect arteries to veins and contain only one layer of endothelial cells with an underlying basement membrane and are defined as a vessel small enough to allow a single blood cell to pass through. These thin walls permit rapid exchange of water, nutrients and metabolic fluids between blood and interstitial fluid (Rosenthal, 1999).

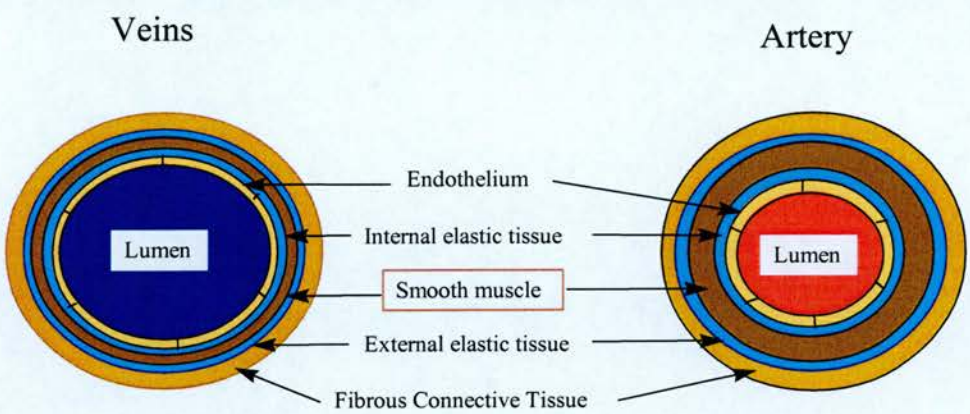


Figure 1.2: Vascular wall components of the major vessels
(Adapted from *Heart Development* Harvey and Rosenthal 1999)

1.2.2 Embryogenesis of the vasculature

The development of the embryonic vascular system is highly conserved between vertebrates. It begins with the appearance of angioblasts (endothelial precursor cells) that differentiate from the mesoderm and proliferate within previously avascular tissue. These angioblasts coalesce to form a primitive tubular network by a process termed vasculogenesis (Yancopoulos et al., 2000). This primary network, which includes some of the major vessels in the embryo such as the aorta and major veins and a homogeneous capillary network, undergoes maturation through a secondary process called angiogenesis (Figure 1.3). This is new vessel formation from pre-existing microvessels and is a complex phenomenon which requires the following sequential steps: detachment of prexisting pericytes for vascular destabilisation, breakdown of extracellular matrix (ECM), migration, proliferation, tube formation by endothelial cells and then reattachment of pericytes for vascular stabilisation (Folkman, 2002).

1.3 Disease and angiogenesis

The vasculature in the normal adult mammal is quiescent with only 1 in 10,000 endothelial cells actively dividing at any given time (Hobson and Denekamp, 1984). The exceptions to this are the highly ordered processes of wound healing and the female reproductive cycles (ovulation, menstruation, implantation and pregnancy). Endothelial cells are able in response to the appropriate stimulus, to actively proliferate to form new vessels by angiogenesis (Hanahan and Folkman, 1996).

Angiogenesis plays a significant role in the progression of many diseases, like solid tumour growth, diabetic retinopathy, rheumatoid arthritis and atherosclerosis (Brooks et al., 1998). Solid tumours promote tumour cell proliferation and invasiveness by the acquisition of new blood vessels which are required to supply nutrients and oxygen, dispose of metabolic waste products,

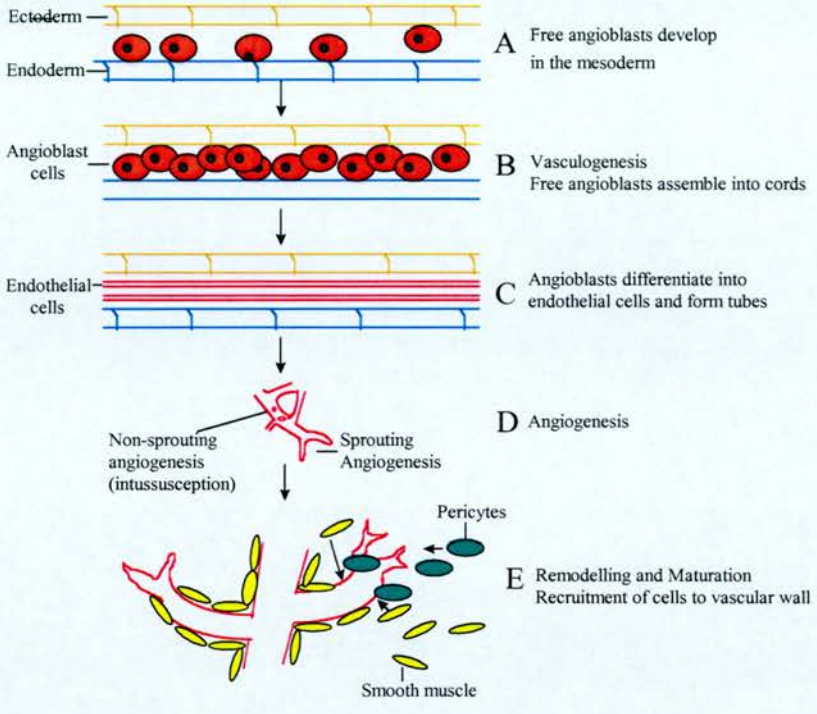


Figure 1.3: Schematic representation of the major processes involved in vascular development. (Adapted from *Heart Development* Harvey and Rosenthal 1999)

and generate paracrine stimuli. Failure by a tumour to gain a blood supply arrests growth and the tumour remains dormant and unable to expand more than a few millimetres (Carmeliet and Jain, 2000).

1.3.1 Angiogenic switch hypothesis

There is evidence that angiogenesis is regulated by both inducers and inhibitors of endothelial cell proliferation and migration. This has give rise to the 'angiogenic switch' model whereby a quiescent endothelial cell if given the correct signal can actively proliferate (Bergers et al., 2000; Hanahan and Folkman, 1996).

1.3.2Angiogenic activator

It has been observed that solid tumours can be implanted either in an avascular region, such as the cornea or on a characteristically vascularised surface, such as the chick chorioallantoic membrane, and in each case elicit the formation of new capillaries (Hanahan and Folkman, 1996). This suggested that tumours release diffusible activators of angiogenesis that could signal a quiescent vasculature to begin capillary sprouting. A number of inducers of angiogenesis have been detected such as basic fibroblast factor (bFGF) and the vascular endothelial growth factors (VEGFs), and the angiopoietins (Davis et al., 1996; Dumont et al., 1994; Suri et al., 1996)

1.3.3 Vascular endothelial growth factor (VEGF) and its receptors

Vascular endothelial growth factors (VEGFs) of which there are seven (VEGF-A, B, C, D, E, F and placental growth factor (PIGF)) are the major regulators of endothelial cell proliferation which occur during the process of vasculogenesis and angiogenesis (Meyer et al., 1999; Nicosia, 1998; Suto et al., 2005). The vascular endothelial growth factors (VEGFs) act through their receptors (VEGFR1, 2, and 3,) (de Vries et al., 1992; Folkman and Klagsbrun,

1987; Millauer et al., 1993) although some members are able to bind another group of receptors neuropillin 1 and 2 (Nrp-1 and 2) (Cebe-Suarez et al., 2006).

The VEGF mitogen causes a signal transduction cascade through binding to the VEGFR tyrosine kinase receptors involving, for example, phospholipase C-protein kinase signalling pathways which regulate the processes of angiogenesis (Shibuya, 2006). One of the consequences of this is the up-regulation of various proteolytic enzymes including the urokinase plasminogen activator and matrix metalloproteinases and aminopeptidases which are required to breakdown basement membrane and extracellular matrix (ECM) (Kato et al., 2001; Mazzieri et al., 1997)

1.3.4 Angiogenic inhibitors

The evidence for endogenous angiogenic inhibitors has come from work on a non tumoriginic hamster cell line that became tumoriginic with a mutation that inactivated a tumour suppressor gene. Comparisons of the cell lines containing or missing the suppressor gene revealed that the non tumoriginic line released high levels of a potent angiogenesis inhibitor. The inhibitory activity was purified and shown to be a truncated form of thrombospondin (TSP1) (Good et al., 1990).

In cancer patients, the removal of a carcinoma alone or in combination with other therapies may be curative. However, the removal of some tumours such as breast carcinomas and colon carcinomas, can be followed by the rapid growth of distant metastases. It has been suggested, therefore, that the primary tumour site is inhibiting the other distant tumours. It is thought that whilst a primary tumour can stimulate angiogenesis in its own vascular bed, it can inhibit angiogenesis in the vascular bed of another metastasis (O'Reilly et al., 1994a).

This hypothesis was proven when an endogenous 38 kDa protein was purified from the serum and urine of mice with Lewis lung carcinoma and shown that it was a cleavage fragment of plasminogen. The fragment was

called angiostatin and structurally consisted of the first four of five highly homologous looped multiple disulphide bonded domains, termed kringle domains (O'Reilly et al., 1994b). Purified angiostatin has been demonstrated to be a potent inhibitor of endothelial cell (EC) proliferation (Cao et al., 1996). Angiostatin prevents angiogenesis by binding to integrin alpha-v-beta3 where it inhibits interaction with matrix ligands reducing cell adhesion and attachment properties of the EC which are a necessary requirement for endothelial cell survival and migration (Tarui et al., 2001).

A key component of the angiogenic process is the involvement of proteases that are able to degrade extra cellular matrix and control cellular processes such as cell cycle and proliferation (Roy et al., 2006). Several protease systems have been demonstrated to be involved in angiogenesis and these include the matrix metallo proteases (MMPs), the 'A Disintegrin and Metalloprotease Domain family (ADAMs), urokinase plasminogen activator (uPA), and the aminopeptidases (Egeblad and Werb, 2002; Roy et al., 2006; Sato, 2003).

1.3.5 Matrix metalloproteases (MMP) and angiogenesis

MMPs are zinc dependent proteolytic enzymes belonging to family M10 and are principally involved with the breakdown of extracellular matrix. They are synthesised as latent enzymes which can be activated upon proteolytic cleavage through a cysteine switch consensus sequence in the propeptide domain (Springman et al., 1990).

Once activated, the activity of the MMPs is controlled and tightly regulated predominantly by a group of structurally related endogenous inhibitors known as tissue inhibitors of metalloproteases (TIMP 1, 2 ,3 and 4) and RECK (Moses, 1997; Takahashi et al., 1998). These specific inhibitors reportedly bind directly to MMPs and inhibit their proteolytic action (Hajitou et al., 2001). The important role of MMP-2 and MMP-9 has been demonstrated in both normal and tumour neovascularisation (Fang et al., 2000; Vu et al., 1998) and siRNA mediated target RNA degradation of MMP-

9 inhibited formation of both capillary like structures in both *in vitro* and *in vivo* models of angiogenesis (Lakka et al., 2005).

1.3.6 A Disintegrin and Metalloprotease domain (ADAMs) proteases.

These are a family of zinc dependent integral membrane and secreted glycoproteins. They contain multiple domains and like the MMPs are activated through a cysteine switch mechanism. These enzymes have diverse roles including ECM degradation, for example ADAM 10 cleaves collagen type IV (Millichip et al., 1998), and they have also been implicated in diseases such as arthritis and cancer (Black et al., 1997; Chubinskaya et al., 2001; Seals and Courtneidge, 2003).

1.3.7 Serine protease uPA

Urokinase plasminogen activator (uPA) belongs to a large serine protease family and specifically the S1 family. It is a highly specific enzyme which catalyses the membrane bound inactive proenzyme plasminogen (plg) into the active enzyme plasmin which cleaves components of the ECM (Mazar et al., 1999). UPA has its own specific cell surface receptor uPAR which greatly enhances the action of uPA on Plg (Ragno, 2006). UPAR is found localised to regions of cell contact (Limongi et al., 1995) and the leading edge of invading monocytes (Estreicher et al., 1990). An additional role of the uPAR receptor is to initiate signalling cascades that require receptor occupancy, for example, integrins represent one set of molecules that are able to interact with uPAR on the cell surface. Immunolocalisation and co-immunoprecipitation have identified uPAR in complex with several integrin families (Reinartz et al., 1995; Wei et al., 1996). Control of these proteases is by the specific inhibition to plasmin with alpha2 antiplasmin and uPA by PAI-1 and PA-I2 (Lijnen, 2001).

1.3.8 Aminopeptidases in angiogenesis and cancer

Two distinct families of aminopeptidases have been reported to be involved in angiogenesis including the M24 type 2 methionine aminopeptidase (MetAP2) and the M1 zinc dependent aminopeptidases comprising the membrane bound APA, aminopeptidase N (APN), and the cytoplasmic adipocyte-derived leucine aminopeptidase (A-LAP) also known as puromycin insensitive leucyl specific aminopeptidase (PILSAP) (Sato, 2003).

The methionine aminopeptidases (MetAPs) are a family of cytosolic metalloproteases responsible for the cleavage of the initial methionine from the N-termini of nascent proteins. Eukaryotes express two forms of MetAPs, types 1 and 2. MetAP2 has stimulated a lot of interest in relation to angiogenesis as it has been shown to be a target of the potent inhibitor of endothelial cell proliferation, TNP470, the synthetic derivative of fumagillin (Griffith et al., 1998; Griffith et al., 1997).

Evidence for APA involvement with angiogenesis has come from several observations. Firstly, that the expression of the enzyme is up-regulated in vessels of human tumours (Marchio et al., 2004) and secondly, that APA knockout mice fail to mount the expected angiogenic response to growth factors and hypoxia. Additionally, specific APA peptide inhibitors and an anti APA monoclonal antibody suppress EC migration and proliferation (Marchio et al., 2004).

Further evidence for aminopeptidase involvement in the process of angiogenesis is from the APN also known as CD13 which has been reported to be expressed in the surface of endothelial cells (EC) of new vessels and tumours but not in quiescent vessels (Fukasawa et al., 2006). Specific inhibition of APN with a monoclonal antibody and functional inhibitors (bestatin) prevented capillary network formation in endothelial cell *in vitro* assays. Furthermore, siRNA mediated interference of APN transcripts in human umbilical vein endothelial cells (HUVEC) prevented capillary tube formation on Matrigel (Fukasawa et al., 2006; Pasqualini et al., 2000). There

data support its involvement in angiogenesis and it has been presumed that it is involved in the degradation of extracellular matrix.

PILSAP is a cytoplasmic aminopeptidase and is also implicated in the process of angiogenesis. The human gene was isolated from a cDNA screen (Schomburg et al., 2000) whilst the mouse orthologue was isolated in a subtractive screen to identify genes that were up-regulated during the *in vitro* differentiation of murine embryonic stem cells into EC cells (Miyashita et al., 2002). Both PILSAP and its mouse orthologue (mPILSAP) have been found to be induced in endothelial cells following vascular endothelial growth factor (VEGF) stimulation and is expressed in ECs during angiogenesis *in vivo* (Miyashita et al., 2002; Yamazaki et al., 2004). Further supporting evidence of this proteases involvement in the process of angiogenesis was provided by an antisense oligodeoxynucleotide to mPILSAP. It was found that EC proliferation, migration and network formation both *in vitro* and *in vivo* were attenuated. It has been proposed that PILSAP may mediate its effect through cell cycle pathways as the cells were arrested in the G1 cell cycle phase (Miyashita et al., 2002; Yamazaki et al., 2004).

1.4 The renin angiotensin system (RAS)

The renin angiotensin system (RAS) regulates blood pressure, endocrine processes and behaviour at both the systemic and local tissue level when angiotensinogen is cleaved by the protease renin to generate decapeptide angiotensin I (Ang I). This is converted to angiotensin II (Ang II) by angiotensin converting enzyme (ACE) (Kim and Iwao, 2000) where it causes vasoconstriction of the endothelium increasing vascular resistance and increase in blood pressure.

Ang II can be catabolised into Angiotensin III (Ang III) by aminopeptidase A (Reaux et al., 1999). Ang III can be proteolytically converted to Angiotensin IV (Ang IV) by other aminopeptidases including placental leucine aminopeptidase (P-LAP) and aminopeptidase O (APO)

(Diaz-Perales et al., 2005; Tsujimoto et al., 1992; Wright et al., 1990) (Figure 1.4).

1.4.1 Aminopeptidases and the control of blood pressure

The evidence that APA can be regarded as an important candidate responsible for the conversion of Ang II to Ang III comes from systemic administration of purified APA into spontaneous hypertensive rats (SHRs) (Mizutani et al., 1987). This has the effect of decreasing blood pressure whilst the addition of an APA specific inhibitor (amastatin) elevates blood pressure in normotensive rats (Ahmad and Ward, 1990).

A study with APA knockout mice demonstrated the baseline systolic blood pressure was significantly elevated in the null mice when compared to wild type littermate controls. Furthermore the infusion of Ang II increased the systolic blood pressure to a greater extent in APA knockout mice (Mitsui et al., 2003). These findings support the role APA might have in the maintenance of vascular tone homeostasis.

Supporting evidence for APO involvement in the RAS was in an *in vitro* assay where recombinant human APO was incubated with angiotensin III (Ang III) and although the reaction proceeded slowly it did generate angiotensin IV. The APO enzyme assay was performed with the recombinant catalytic domain and not the entire protein. The authors acknowledge that because of this, further experiments would be required to confirm the putative activity of the enzyme under physiological conditions (Diaz-Perales et al., 2005). The vasoactive peptides of the renin angiotensin system are systemically transported in the blood and are either enzymatically cleaved by secreted enzymes or by membrane bound enzymes. Unlike APA and PLAP, the

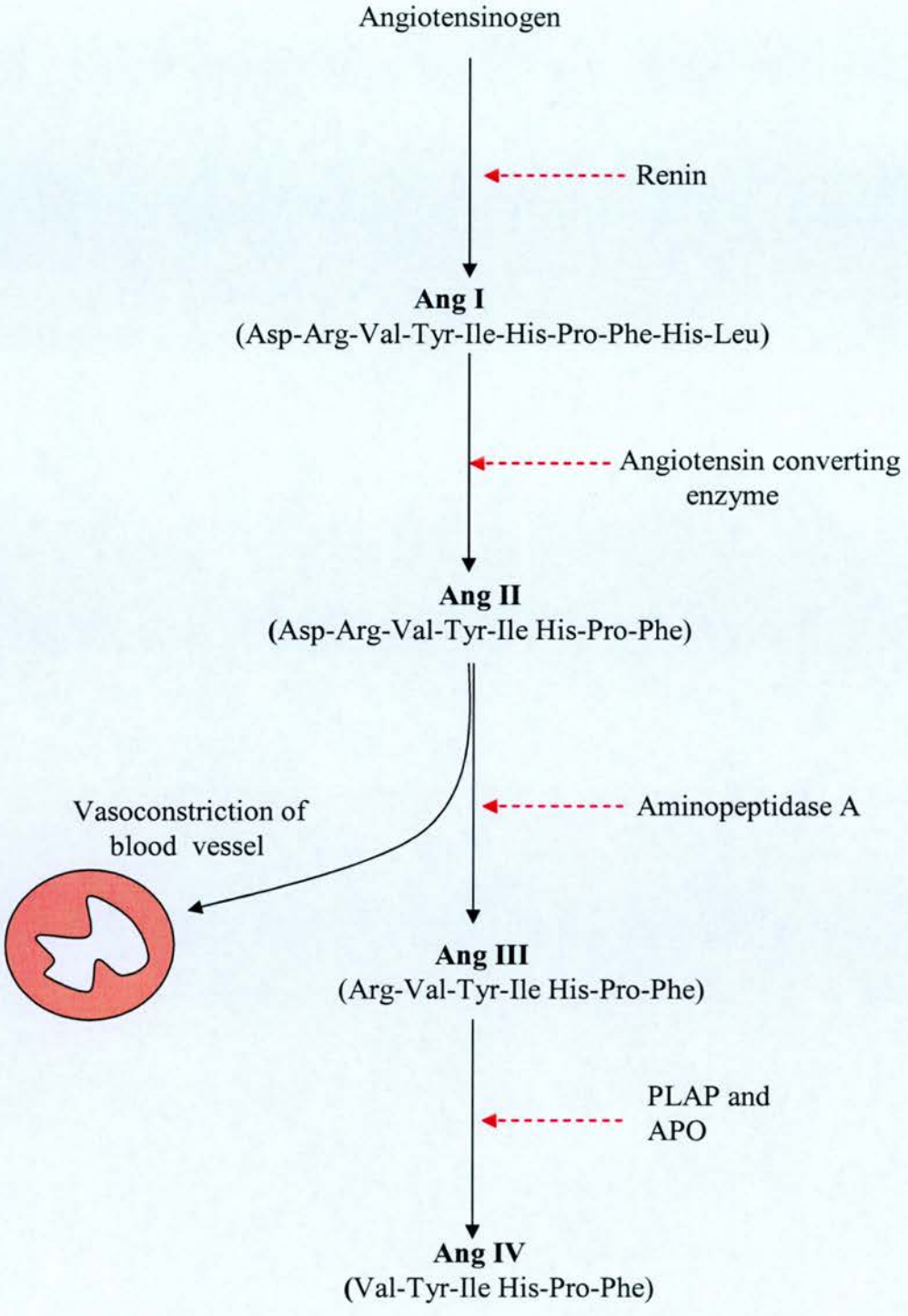


Figure 1.4: Schematic representation of the renin angiotensin pathway. Proteolytic cleavage of angiotensinogen by renin releases Ang I which is subsequently cleaved by angiotensin converting enzyme (ACE) to release the peptide Ang II which exerts its effect on the endothelium to raise blood pressure by causing vasoconstriction. Ang II is subsequently cleaved by a series of proteases that control homeostasis.

protease APO does not have a transmembrane domain or signal peptides to direct it to the cell membrane. It is therefore unlikely that this enzyme is physiologically able to participate directly in the catabolism of the systemic vasoactive peptide Ang III. Furthermore, APO enzyme assay was performed with the recombinant catalytic domain rather than the entire protein. The authors acknowledge that because of this, further experiments would be required to confirm the putative activity of the enzyme under physiological conditions ((Diaz-Perales et al., 2005). and it is therefore important to be cautious about over interpreting the significance of this data.

1.4.2 Clinical evidence for aminopeptidase involvement in blood pressure regulation

Blood pressure regulation and pathogenesis have been extensively studied in pregnancy with regard to preeclampsia. This is a hypertensive disorder that affects at least 5-8% of all pregnancies (Page, 2002). The symptoms include high blood pressure, the presence of protein in the urine, oedema, sudden weight gain, headaches and changes in vision. The condition usually occurs after 20 weeks gestation, although it can be present earlier. It is estimated that preeclampsia and other hypertensive disorders of pregnancy are responsible globally for approximately 76,000 deaths per annum (Page, 2002).

There is growing evidence from clinical studies of aminopeptidases involvement during pregnancy. Foetuses produce high levels of the vasoactive peptides such as Ang II compared to their mothers. It has been shown that maternal APA increases during normal pregnancy whilst those with preeclampsia show a variety of changes. For example, in mild preeclampsia, APA levels are elevated whilst levels in those with full blown preeclampsia APA levels are decreased (Mizutani et al., 1995; Mizutani and Tomoda, 1996).

Another M1 aminopeptidase is placental leucine aminopeptidase (P-LAP) so-called because it was believed to be expressed only in this tissue, although has now been shown to have a broad tissue distribution (Rogi et al., 1996). PLAP is believed to play a role in blood pressure regulation by hydrolysing vasopressin (AVP) and Ang III (Matsumoto et al., 2000; Tsujimoto et al., 1992). Once again most of the knowledge of this enzyme comes from studies conducted during pregnancy. Membrane bound PLAP is released into maternal blood flow via proteolytic cleavage by other proteases such as the ADAMs and the metalloproteases (Iwase et al., 2001; Ofner and Hooper, 2002b). During normal pregnancy the serum levels of PLAP gradually increases. However, in mild preeclampsia patients PLAP levels are very high whilst in severe cases there is a dramatic decrease in the levels of PLAP (Mizutani and Tomoda, 1996). These findings are not dissimilar to data acquired for APA.

1.4.3 Renin angiotensin system and cancer

There is increasing evidence that peptides Ang II and Ang III known to control blood pressure may also play vital roles in other cellular processes such as the regulation of cell proliferation, angiogenesis and inflammation (Miyajima et al., 2002; Uemura et al., 2003) and this has led to the suggestion that these molecules may be implicated in cancer (Deshayes and Nahmias, 2005).

At the molecular level Ang II and Ang III stimulate specific receptors, namely, Ang II type 1 receptor (AT1R) and Ang II type 2 receptor (AT2R) that reside on the cell surface (Nouet and Nahmias, 2000). Stimulation of AT1R has been shown to induce cell proliferation in human cancer cells (Uemura et al., 2003) (Fujimoto et al., 2001), by activating various intracellular cascades of protein kinases usually associated with growth factor stimulation. ATR1 transactivates the EGFR in prostate and breast cancer cells leading to extracellular regulated kinase (ERK) activation, phophorylation of a signal transducer and activation of a transcription factor STAT3 and

activation of protein kinase C (PKC) (Greco et al., 2003; Uemura et al., 2003) (Figure 1.5). Another aspect of Ang II and Ang III is the effect of AT1R stimulation on the upregulation of vascular endothelial growth factor (VEGF) and its receptor VEGFR2, which are major players in the process of angiogenesis (Imanishi et al., 2004; Otani et al., 2001) (Figure 1.5).

Ang IV may also be implicated in cancer as it has been demonstrated *in vitro* to increase cell proliferation of rat anterior pituitary cells (Pawlikowski and Kunert-Radek, 1997). It is another proteolytically generated peptide from the RAS pathway and it is cleaved by the aminopeptidase PLAP. These data support a more recent view of the RAS being not only involved in blood pressure regulation but also implicated in other cellular processes such as cell proliferation and angiogenesis.

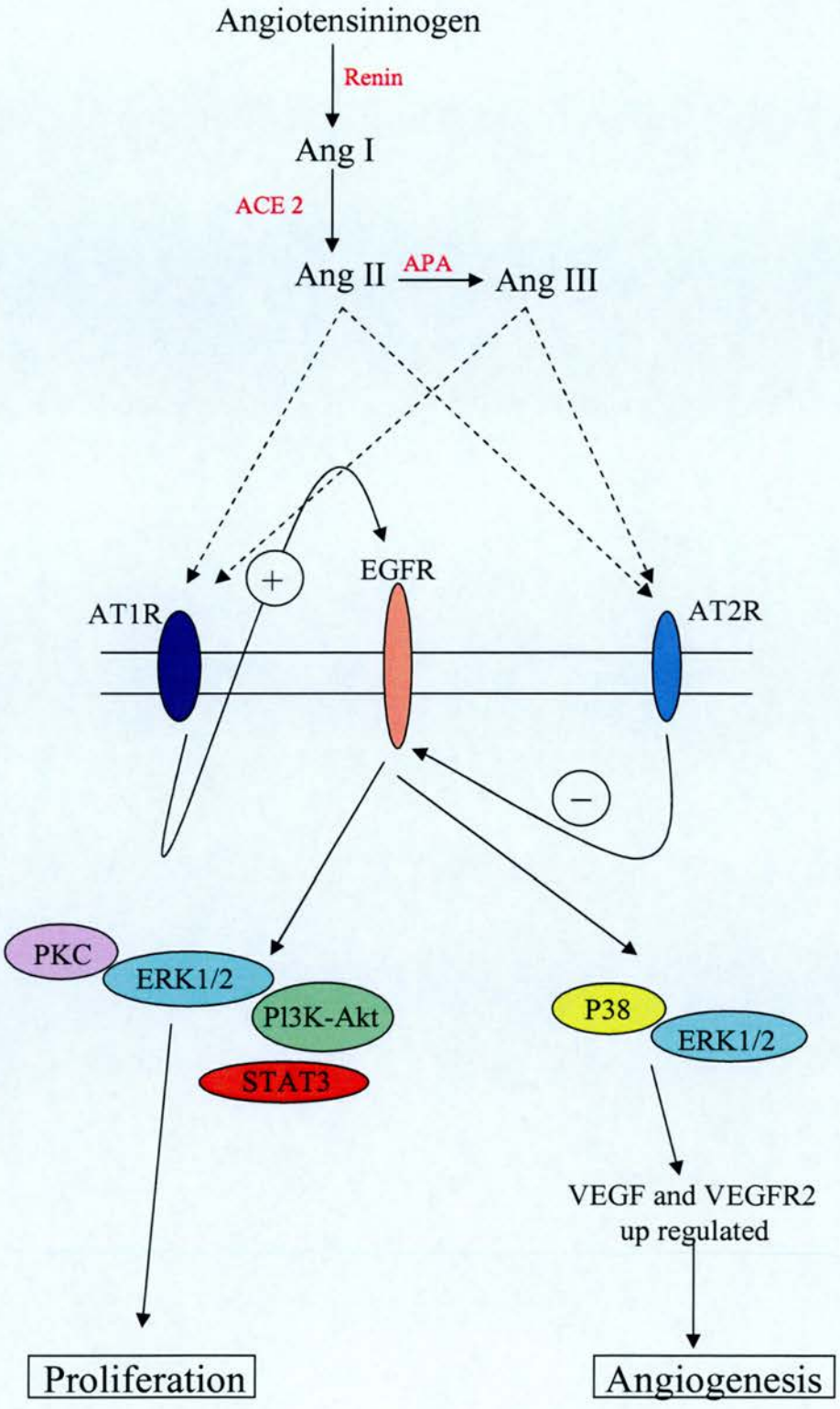


Figure 1.5: Schematic illustration of the renin angiotensin system and its cleavage products involvement with proliferation and angiogenesis.
Signal transduction cascades mediated through the binding of Ang II and Ang III at the cell membrane to receptors AT1R and AT2R where they instigate a signal transduction cascade through the EGFR receptor leading to an increase in cell proliferation and angiogenesis (adapted from Kurdi et al 2005).

1.5 Mutagenesis

1.5.1 Mutagenesis strategies in the mouse

With the human and mouse genomic sequence databases publicly available, the next major goal is to determine the function of genes. The approach taken by the mouse genetics community is to mutate every gene in the mouse genome termed 'genome saturation' and to identify any resulting phenotype.

To this end, several mutagenesis strategies have been developed which include: the use of the chemical mutagen ethylnitrosourea (ENU) (Favor and Neuhauser-Klaus, 2000; Russell et al., 1979), gene targeting (Thomas and Capecchi, 1986) and gene trapping (Gossler et al., 1989).

1.5.2 Ethylnitrosourea (ENU) Mutatgenesis

The chemical mutagen ethylnitrosourea (ENU) introduces point mutations into spermatogonial stem cells. They are used in phenotypic driven screens where mutants that have specific phenotypes are chosen (Quwailid et al., 2004) or where region specific mutations that lie on a particular genomic interval are selected for further analysis. To identify the gene that is mutated requires appropriately designed breeding schemes and these are often costly and time consuming to carry out. The use of balancer chromosome technology has been used to simplify the ENU screening strategy by reducing the amount of labour required (Yu and Bradley, 2001).

1.5.3 Gene targeting

Gene targeting involves the genetic modification of ES cells through the introduction of a vector that contains two homology arms specific to the locus of interest. Homologous recombination facilitates the replacement of the targeted allele with the vector sequence. The targeted ES cell can then be used to generate germ line transmiting mice that can be studied for a phenotype

(Thomas and Capecchi, 1986) (Doetschman et al., 1987; Thomas and Capecchi, 1987; Wong and Capecchi, 1987). To date, gene targeting has contributed approximately 3600 mutants that corresponds to about 15% coverage of the mouse genome (Skarnes, 2005).

Gene targeting is both labour intensive and time consuming since it requires the construction of the targeting vector and the screening of a large number of ES cell clones to identify the correctly targeted allele. It is unsuitable for high-throughput methodologies although recombineering strategies have been developed to assist in the making of the targeting construct (Carlson and Largaespada, 2005) (Testa et al., 2003).

1.5.4 Gene trapping

Another mutagenesis strategy, gene trapping (Evans et al., 1997), entails the random introduction into ES cells (via electroporation or retroviral infection) of a DNA vector that has a selectable marker to screen for integrations and a reporter gene that mimics the endogenous gene expression. A consequence of the integration is that the endogenous gene may potentially be mutated and the subsequent transmission of the trapped ES cell line to the germ line allows phenotypic analysis in mice (Gossler et al., 1989)

The basic principle of gene trapping involves a vector with a splice acceptor site (SA) immediately upstream of a reporter gene (*lacZ*) followed by a neomycin resistance (neo) selectable marker. The random integration of these vectors into a genomic locus downstream of a functional endogenous promoter results in a fusion transcript between the integrated gene and the *lacZ* reporter gene (Friedrich and Soriano, 1991; von Melchner et al., 1992).

The consequences of the gene trap integration are 1) the ES cell gene trap clone can be characterised *in vitro* before germ-line transmission 2) the fusion transcript between the gene and vector will mimic endogenous gene expression which can be assessed by reporter gene activity at the insertion locus and 3) the fusion transcript serves as PCR templates for 5'RACE allowing identification of the gene that has been trapped (Stanford et al.,

2001). This technique has been used extensively and lends itself to high-throughput protocols. Furthermore, the ES cell clones bearing random integrations can be stored for future use (see section 1.5.7).

1.5.5 Gene trap vector development

The basic gene trap vector consisted of a splice acceptor followed by *lacZ* reporter gene and a poly A signal. The vector also included a *neo* selection gene driven by a PGK promoter (Gossler et al., 1989). This vector was subsequently adapted to include an ATG up-stream of the *lacZ* gene which resulted in a three-fold increase in the number of βgal clones probably due to the detection of out-of-frame splicing events (Hill and Wurst, 1993). A drawback of this particular vector was that the promoter driving the neo selectable marker resulted in increased background since intergenic integrations were also isolated (Forrester et al., 1996) To circumvent this problem, the βgeo vector containing a fused *neo* to the 3' end of the *lacZ* gene was developed (Friedrich and Soriano, 1991). This had the effect of placing the *neo* selectable marker under the control of the promoter of the endogenous trapped gene, thus reducing the background intergenic integrations.

1.5.6 Technical issues and limitations

Gene trapping has been successfully used to mutate and identify genes (Zhou et al., 2004) (Torres et al., 1997). However, it is possible that the gene trap reporter may not recapitulate the endogenous gene expression and, for this reason, it should be confirmed by RNA *in situ* hybridisation with a probe specific to the endogenous gene (Deng and Behringer, 1995; Pall et al., 2004).

Gene trap integrations may occur anywhere in a gene and those that take place at the 3'end are more likely to generate a hypomorphic allele (a protein that retains its function and does not result in a phenotype) (Herrick and Cooper, 2002). It is also possible for the gene trap vector to be spliced around

generating the wild type transcript resulting in a milder or absent phenotype (Skarnes et al., 1992) (Sam et al., 1998).

Another limitation is that it is not uncommon for the gene trap vector to integrate multiple times within a genetic locus. This can complicate analysis of the integration event and cause uncertainty about whether other genes at the locus are also affected (Pall et al., 2004).

1.5.7 Large scale screens

A few centres have performed high throughput screens with the aim of generating a worldwide mutant resource and are collectively known as the International Gene-trap Consortium (IGTC). These academic groups have generated 27,000 tags, covering 32% of the genome (Skarnes et al., 2004). The IGTC aims to make the ES cell lines available to the public without restriction and this consists of The Bay Genomics Gene Trap Project (www.http:/baygenomics.ucsf.edu/) and The Gene Trap Project of the German Human Genome Project (www.http://tikus.gsf.de). A biotechnology company, Lexicon Genetics, has also created a large library of ES cell lines (OmniBank: http://www.lexicon-genetics.com) and claim to have achieved close to 60% coverage of the genome with the generation of 200,000 tags. A comparison between the IGTC and Lexicon gene trap data indicates that 20% of the trapped genes in the public database are different from Lexicon (Skarnes et al., 2004).

1.5.8 ES cell directed gene trap screens

Gene trap screening strategies have been developed to exploit the unique capacity for ES cells to differentiate *in vitro* in to different lineages enabling the selection of gene trap clones for germline transmission that are likely to be involved in the development of a particular lineage (Stanford et al., 1998). An *In vitro* pre-selection screen for genes responsive to retinoic acid (RA), the active metabolite of vitamin A was performed to enrich for insertions into

genes that are involved with embryogenesis (Forrester et al., 1996). The application of RA as a cytokine to embryos (Conlon and Rossant, 1992; Kessel and Gruss, 1991; Morriss-Kay et al., 1991) or to differentiating ES cells in culture (Kondo et al., 1992) can alter the expression of developmentally regulated genes. For example RA enhances differentiation into neural lineages but represses commitment to haematopoietic and cardiac fates (Bain et al., 1996; Wobus et al., 1997). The gene trap screen that resulted in the GT411 ES cell clone was an RA *in vitro* pre-selection screen designed to enrich for genes involved in the development of the cardiovasculature (Forrester et al., 1996).

1.5.9 The gene trap screen that isolated GT411

E14 ES cells were electroporated with the gene trap vector pGT01, 2. The vector consisted of a splice acceptor (SA) derived from the *engrailed* -2 gene, a βgalactosidase reporter gene fused to a promoterless neomycin selectable marker (βgeo) (Figure 1.6). Colonies which were resistant to G418 were picked and replica plated prior to incubation with RA. Clones which showed repression of reporter gene in response to the RA were differentiated into 'beating' cardiomyocytes, which were recorded on video and then stained again with X-gal. Those clones that demonstrated βgal staining in the 'beating' regions were selected for further analysis. The GT 411 ES cell line was one of 42 clones that were selected from this screen.

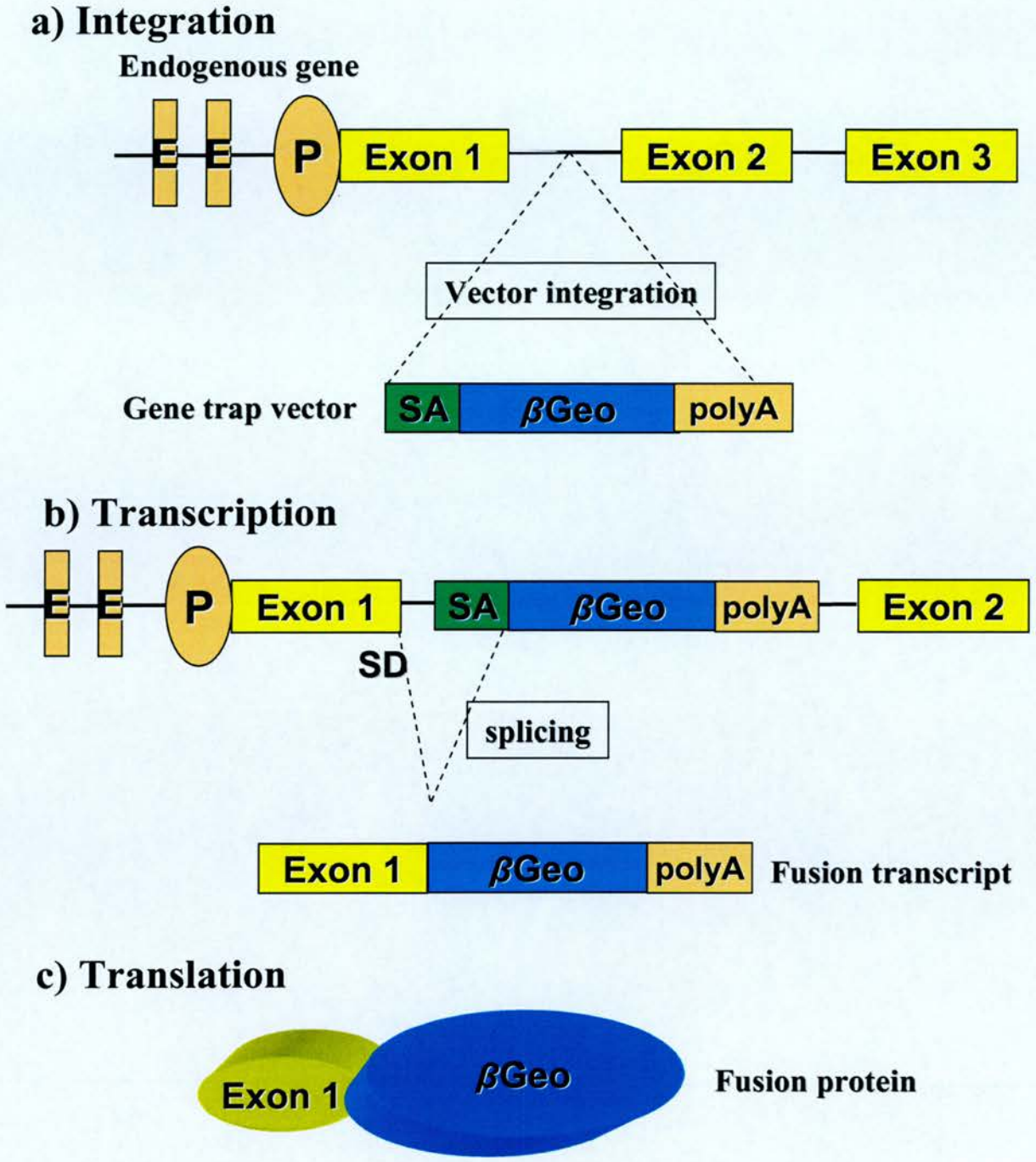


Figure 1.6: Gene trapping. a) The gene trap vector as used in this study consisted of a splice acceptor(SA), βgeo (*LacZ* reporter fused to neomycin resistance gene) and a polyadenylation signal (Poly A) integrates randomly into the intron of a gene. b) The gene trap vector is regulated by the endogenous genes promoter and enhancer sequences. The presence of the splice acceptor in the construct and the endogenous splice donor from exon 1 results in a fusion transcript between the trapped gene and vector. c) Translation of the mRNA results in a fusion protein.
E=enhancer, P= promoter, SD= splice donor.

Chapter 2:

Materials and methods

2.0 Material and methods

2.1 Molecular techniques

The preparation of standard solutions and general molecular techniques described in this section were carried out in accordance with Molecular Cloning Laboratory Manual as a source reference unless otherwise stated (Sambrook 1989).

2.1.1 Polymerase chain reaction (PCR)

This method was used for a variety of purposes ranging from 5'RACE, RT-PCR and generating full-length clones. A typical PCR reaction was as follows: 10X PCR Buffer supplemented with $MgCl_2$ (1.5mM), dNTP's (10 mM), primer A and primer B (10μM), Template DNA (100ng), Taq Polymerase (1 unit) and made up to a final volume of 50µl with ddH_2O. PCR was performed on a Peltier DNA Engine.

2.1.2 Gel purification of PCR products

Bands were excised from low melting point agarose and purified by either one of two methods; either by a spin prep method (see TA cloning) or by melting the agarose in 500µl of ddH_2O at 70°C for 10 minutes with occasional vortexing. To this mix, 1ml of a DNA binding resin (Promega) was added and incubated for 2 minutes at RT. The slurry was captured with a commercial filter and then washed with 2ml isopropanol (80%) and the filter dried by spinning in a microfuge (12,000 g) for 2 minutes. To elute the DNA of the resin through the filter, 50µl of ddH_2O was applied and incubated for 1 minute. The filter was then spun at full speed (12,000 g) for 1 minute and the eluted DNA was collected in a microfuge tube.

2.1.3 Manual cycle sequencing

PCR products were prepared for sequencing with 1μl of exonuclease I (10 units/ μl) and 1μl of shrimp alkaline phosphatase (2 units/ μl) incubated at 37°C for 15 minutes, followed by heat inactivation at 80°C for 15 minutes. The cleaned PCR product was then sequenced using the ^{33}P dideoxy terminator kit (Amersham Pharmacia Biotech) in accordance with the manufacturer's protocol.

2.1.4 Denaturing gel electrophoresis of sequencing products

An 8% glycerol tolerant gel stock comprised the following reagents; 24ml acrylamide (BDH 30% stock 3% bis-acrylamide), Urea (50g), 5ml of 20X glycerol tolerant buffer (Tris base, 216g Taurine 72g, $Na_2EDTA.H_2O$ 4g) and 100ml was made up with ddH_2O.

The Biorad gel cast system was set up according to manufacturer's guidelines. To the 8% glycerol tolerant gel mixture, 1ml of 10% ammonium persulphate and 25 μl TEMED (N,N.N',N'-tetramethylethylenediamine) was added and the mix was poured into glass plates. The sharks tooth comb was inserted and the gel was allowed to set for two hours.

The gel was pre-run in 1X glycerol tolerant buffer at 70 watts for 1 hour. To the cycle sequencing products, 4μl of stop solution (95% formamide, 20mM EDTA, 0.05% bromphenol blue and 0.05% xylene cyanol FF) was added prior to heating the samples to 70°C for 5 minutes. The wells of the gel were flushed to remove urea and then loaded with the samples; each reaction was loaded G,A,T,C. The gel was run until the bromphenol blue was near the bottom of the gel. The apparatus was dismantled and the gel adhered to 3mm Whatman paper before drying on a gel drier for two hours. Autoradiography was carried out overnight at room temperature.

2.1.5 Fluorescent ABI Sequencing

Sequencing reactions were cleaned as described above if a PCR product was used otherwise 0.5μg of plasmid DNA was sent to ICAPB Edinburgh University Sequencing Service with the appropriate primer. Sequences were returned as a text file and analysed appropriately.

2.1.6 TA cloning (Invitrogen)

PCR products were run on low melting point agarose gels containing ethidium bromide and visualised with a UV spectrophotometer. Bands were excised and placed in a SNAP microfuge tube and centrifuged at full speed (12,000 g) for 1 minute. The agarose was retained by the filter paper barrier, whilst DNA was collected in the microfuge tube. The captured DNA (2μl) was then used in the TA cloning procedure with vector PCR4-TOPO (Invitrogen) according to manufacturer's protocol.

The reaction was placed at room temperature for 5 minutes and then put on ice. The reactions were subsequently transformed into DH5 alpha bacteria (see below) spread on the appropriate antibiotic selection plates and incubated at 37°C overnight. Colonies were selected for further analysis and characterised by sequencing (see transformation procedure below).

2.1.7 Preparation of DNA inserts

PCR products were cut with the appropriate restriction enzyme/s for each particular cloning strategy. Products were run on 0.8% low melting point 1XTBE gels and bands were excised under UV light (250nm). Excess agarose was trimmed away and the band was placed in an eppendorf with 500μl of distilled water, heated to 70°C to melt the agarose for 10 minutes and vortexed on a whirlimix. The protocol used was a Promega DNA purification kit which involves adding a charged bead that binds DNA in solution and applying this to a column to capture the beads followed by washes to remove

contaminants and finally an elution step was carried out with ddH_2O to elute the DNA from the bead.

2.1.8 Restriction digest

DNA was cut with the appropriate restriction enzyme and buffer. The reaction was incubated at the appropriate temperature for the enzyme (usually 37°C) for a minimum of 4 hours. Double enzyme restriction digest for directional cloning was carried out in a similar manner, but using an enzyme buffer that gave optimal cutting for both enzymes. DNA could then be isolated by gel purification.

2.1.9 Ligations

Ligation reactions were set up with a vector to insert molarity ratio of 1:3 with a rapid DNA ligation kit (Roche). Vector DNA and insert DNA (1:3 molar ratio, respectively) were dissolved in 1x concentration DNA dilution buffer (vial 2) in a total volume of 10 µl. To this mixture, 10 µl of 2X T4 DNA ligation buffer (vial 1) was added followed by 1 µl T4 DNA ligase (vial 3) and the entire solution was incubated for 5 minutes at RT. The ligation reactions were then immediately used in bacterial transformations.

2.1.10 Transformation of bacterial cells

Frozen aliquots of competent cells (DH5 alpha or BL21) were thawed on ice prior to use. The cells were mixed with 3µl of the ligation reaction in a pre-chilled eppendorf and left on ice for 30 minutes. The cells/ligation mix was heated in a water bath set at 42°C for 40 seconds and then placed back on ice for 2 minutes. To each transformation tube, 250 µl of SOC medium was added and then the tubes were placed in an orbital shaker at 37°C at 225 rpm for 1 hour. The cells were pelleted at low speed in a centrifuge (2000 g) before being resuspended in 100µl of L-broth and spread on to LB-agar plates

with the appropriate antibiotic for positive selection. Plates were placed in a 37°C incubator overnight.

2.1.11 Glycerol stocks

For long term storage, 8% glycerol stocks of bacterial clones were made by taking 20μl of a mini prep culture to 1 ml of 8% glycerol in L broth and stored in a minus 80°C freezer.

2.1.12 Plasmid preparation (Miniprep and Maxiprep)

A single colony from a freshly streaked plate was used to innoculate 5ml of L-broth containing the appropriate antibiotic. The culture was grown overnight at 37°C in an orbital shaker (225 rpm). The cells were harvested by centrifugation 3,000 g for 10 minutes. The supernatant was discarded in accordance with local biohazard guidelines and the pellet was then used to extract plasmid DNA with the Mini-prep extraction kit (Promega) following the manufacturer's protocol.

Large scale DNA plasmid purification (Maxi-Prep) was treated in the same way as a mini prep except that after the overnight incubation, 400μl of the LB mini prep was used to seed a 400ml flask and shaken in an orbital shaker at 225rpm, 37°C for 16 hours. The cells were harvested in a centrifuge at 3000 g for 10 minutes and then the plasmid was extracted using a Quiagen Maxi prep kit in accordance with the manufacturer's handbook.

2.1.13. Sub-cloning

Plasmid vectors were cut with the restriction enzyme required for the cloning application. The linearised vector was then treated with 1 unit of shrimp alkaline phosphatase (SAP) for 15 minutes at 37°C to prevent re-circularisation. The SAP was inactivated by heating at 80°C for 15 minutes.

2.1.14 Phenol extraction and ethanol precipitation

The vector was extracted with an equal volume of phenol/chloroform and then precipitated with 0.1 volume of 3M sodium acetate pH 5.3 and 2 volumes of ethanol (-70°C) for 30 minutes. Following precipitation the DNA was pelleted by centrifugation (12,000 g) at 4°C for 20 minute, washed in 70% ethanol, air-dried and resuspended in an appropriate volume of TE.

2.1.15 Isolation of genomic DNA from tissues/cells

The tissue was chopped into small pieces with a sterile razor blade (cells are homogenised with a narrow guage needle) and placed in a microfuge tube. 500μl of lysis buffer (100mM Tris-HCl pH 8.5, 5mM EDTA, 0.2% SDS, 200mM NaCl and 100μg/ml proteinase K. This was incubated at 55°C overnight. The microfuge tube was spun at 12,000 g for 10 minutes. The supernatant was removed to a fresh microfuge tube and extracted with 500μl phenol chloroform, mixed briefly and centrifuged at 12,000 g for 1 minute. The aqueous layer was taken and mixed with an equal volume of isopropanol to precipitate the DNA. It was spun at 12,000 g for 20 minutes to pellet the DNA. The pellet was washed with 70% ethanol, air dried before resuspending in 500μl of ddH_2O. The DNA concentration was determined on spectrophotometer (OD_{260}).

2.1.16 Random primed labelling of probes

All experiments involving radioactivity were carried out in designated areas following local rules and guidelines to minimise exposure and avoid contamination. DNA fragments were purified by elution from agarose gel bands and labelled with ^{32}P for hybridisations for Northern or Southern blots.

Random-primed labelling was performed with a "High Prime" kit as recommended by Roche: 25ng of DNA was mixed with ddH_2O to a final volume of 11µl, denatured at 90°C for 10 minutes in a screw cap tube, snap cooled on ice and condensate spun down by centrifugation. On ice, 5µl high

prime mix (1 unit/µl Klenow polymerase; 0.125mM of each dATP; dGTP and dTTP; 50% v/v glycerol) and 4µl of 50µCi [alpha ^{32}P] dCTP (3000Ci/mMol) aqueous solution were added to the denatured DNA and the reaction was then incubated at 37°C for 10 minutes. The reaction was stopped by heating to 65°C for 10 minutes. After this time, 80µl of ddH_2O was added to the tube and unincorporated nucleotides were removed by centrifugation through a Sephadex G-50 column at 3000 g for 2 minutes.

2.1.17 Southern blot

Genomic DNA (10µg) was restriction digested overnight at the appropriate temperature for the enzyme used. The products were resolved on a 1% TBE agarose gel. The gel was stained and destained with ethidium bromide. A photograph was taken with a ruler along the side of the gels, so that the distance of the markers could be calibrated. The gel was placed in denaturing solution (0.5M NaOH; 1.5M NaCl) for 20 minutes and then neutralised with neutralisation solution (1.5M NaCl; 0.5M Tris.HCl pH 7.4) for 30 minutes. The gel was transferred to a membrane Hybond N+ as described in Maniatis by a wet blot capillary method. After overnight blotting the membrane was marked for orientation, rinsed in 2XSSC (pH 4.4) and baked at 80°C for 2 hours.

2.1.17.1 Prehybridisation and hybridisation

The ULTRAhyb (Ambion) was preheated in a water bath at 68°C and the buffer swirled until all precipitated material had dissolved. The membrane was prehybridised with 10ml ULTRAhyb for 30 minutes at 42°C in rotating Techne hybridisation bottles in Techne Hybridistion HB-1 ovens.

The probe was denatured for 10 minutes at 90°C, cooled on ice and centrifuged and then added to the blot with fresh hybridisation buffer. Routinely 10^7 counts per minute of a probe in a 10ml hybridisation were used for hybridising to a standard (13X11cm) midi-gel blot. It was incubated at

42°C overnight in a rotating Techne hybridisation bottle in Techne Hybridisation HB-1 oven.

2.1.17.2 Washing

The hybridisation buffer was discarded and the blot was washed twice in 0.1 %SDS, 2XSSC pH 4.5 for 5 minutes each. Then the blot was washed in 0.1%SDS, 0.1%SSC pH 7.4 at 55°C. At this point the blot was monitored with a Geiger counter to assess the background level of radiation. The washing and monitoring continued until the Geiger counter readings were very low. The membrane was wrapped in Saran plastic before being exposed to X-ray film at -70°C with an intensifying screen, initially for 24 hours. If the band signal was low the time of exposure was increased accordingly.

2.1.17.3 Stripping DNA and RNA blots

Blots were stripped in 0.01xSSC pH 4.5, 0.01% SDS at 70°C to remove bound radiolabelled probe. The buffer was changed frequently until the radioactive counts dropped to background levels. Stripped blots were then exposed to autoradiographic film for 1 week to check that the probe had been removed prior to re-hybridisation

2.1.18 Preparation of RNA

RNA is extremely susceptible to degradation by RNases. Care must be taken to avoid contamination of samples. Gloves should be warn, barrier tips used with pipettes and fresh tubes/glassware that have been treated to remove RNases.

2.1.19 Preparation of diethylpyrocarbonate (DEPC) water

1ml of DEPC was added to 1 litre of autoclaved water, mixed thoroughly and left in a 37°C water bath for 2 hours. In order to inactivate it, the DEPC

water, was autoclaved. Care must be taken when using DEPC as it is carcinogenic

2.1.20 RNA extraction

Tissues were dissected and directly frozen in liquid nitrogen. RNA isolation from tissues and cells grown as a monolayer were homogenised (Polytron) in Trizol (Invitrogen) and then total RNA was extracted following the manufacturer's protocol (in fumehood). Specifically, 1 ml of Trizol was used per 50mg of solid tissue or in the case of cells grown as a monolayer, 1 ml of Trizol was sufficient to cover a 3.5 cm diameter dish. The homogenates were incubated for 5 minutes at room temperature (RT) before the addition of 0.2 ml chloroform. Tubes were sealed and shaken vigorously for 15 seconds and then incubated for 15 minutes at RT. The tubes were centrifuged at 12,000 g in 4°C for 10 minutes. The upper phase (60%) contains the RNA fraction and this was precipitated with 0.5ml of absolute isopropanol. Samples were incubated for 10 minutes prior to centrifuging at 12,000 g for 10 minutes at 4°C. The RNA pellet was washed with 1 ml of 75% ethanol and then centrifuged at 5000 g for 5 minutes at 4°C. The pellet was allowed to air dry (care being taken not to over dry the pellet as it can be difficult to resuspend RNA) and then it was resuspended in 100μl of RNase free water.

Quality was assessed by gel electrophoresis (0.8% TBE agarose gels) and the RNA concentration determined by UV spectrophotometry.

An Oligotex mRNA mini extraction kit (Qiagen) was used in accordance with the manufacturer's guidelines to isolate messenger RNA. Samples were treated identically and diluted to the same concentration (1μg/μl). A typical reaction was 100 μl of total RNA (1μg/μl), 150 μl of dd H_2O, 250 μl of Oligotex binding buffer (OBB) and 15 μl of Oligotex resin beads.

This reaction was incubated at 70°C for 3 minutes, then at RT for 10 minutes. The Oligotex bead/RNA mix was loaded onto a column. The column was washed with wash buffer and then centrifuged at 10,000 g to remove unbound RNA and excess buffer. The mRNA was eluted from the column

with pre-warmed oligotex elution buffer (two elutions each 50μl). To determine the amount of mRNA extracted, it was assumed that 1μg of total RNA contains 5% mRNA, thus 100μg of total RNA would yield 5μg mRNA.

2.1.21 Northern blot analysis

2.1.21.1 Electrophoresis

The electrophoresis tank and gel casting apparatus and 50µl well capacity comb were cleaned thoroughly with DEPC water prior to washing in a 3% solution of hydrogen peroxide for 10 minutes. The equipment was washed again in DEPC water and dried. A 150 ml midi agarose gel was cast in a fume hood. The gel consisted of 3ml 20mM phosphate buffer (11.5 ml Na_2HPO_4 (1Molar pH 7.0) and 8.5 ml $Na_2H_2PO_4$ pH 7.0 made up to 1 litre), 2.25g of agarose, 25.5ml ultrapure formamide (Sigma) and 121 ml of DEPC water. (The formamide was added after the agarose mix had cooled).

In order to prepare RNA for loading onto the agarose gel, it was prepared as follows: 16.5µl ultrapure formamide, 5µl formaldehyde, 6µl of 20mM phosphate buffer and 15µl of mRNA (3µg). To this mix, 7µl of loading dye (bromophenol blue EDTA) was added and heated to 70°C for 10 minutes and then cooled on ice.

The RNA sample was loaded onto the formamide agarose gel and was electrophoresed at 90 volts for 3 hours at RT. The Marker lane was removed, washed in 0.1M ammonium acetate, stained in ethidium bromide (0.5mg/ml), destained and then photographed. RNA lanes were washed in distilled water and then placed in 0.05M sodium hydroxide/1.5M sodium chloride for 30 minutes. The gel was neutralised in 0.5M Tris/HCl pH7.4/1.5M NaCl for 20 minutes. The mRNA was transferred from the gel to a nylon membrane (hybond N+ Amersham/Pharmacia) in a 10XSSC (pH 4.4) wet transfer set-up overnight in a blotting sandwich (Maniatis).

The blotting apparatus was dismantled and the membrane baked in an oven at 80°C for 2 hours. The membrane was stored at room temperature prior to hybridisation (see sections Southern blotting)

2.1.22 RT-PCR

First strand cDNA synthesis was performed on 250µg mRNA using superscript Reverse Transcriptase (Invitrogen). After cDNA synthesis, the 25µl reaction volume was brought to 50µl with ddH_20. For each PCR reaction, 1µl of cDNA was used. A typical PCR condition was as follows: 2 minutes at 94°C for 1 cycle, followed by 15 seconds at 94°C, 30 seconds at 55°C and 2 minutes at 72°C for 30 cycles.

2.1.23 Rapid amplification of cDNA ends (5'RACE)

5' RACE was carried out on mRNA isolated from 11.5 dpc wild type mouse embryos. Using a Gene Racer kit (Invitrogen), cDNA was generated as described in the manual. First round RACE-PCR was carried out with P/3 and anchor primer AP1 (Invitrogen) and then nested with P/2 and anchor primer AP2 under the following 'touch down' PCR conditions using Thermozyme Taq (Invitrogen) as described in the product protocol.

1 cycle	5 cycles	5 cycles	25cycles
2min 94° C	15 sec 94°C	15 sec 94°C	15 sec 94°C
	30 sec 72°C	30 sec 70°C	30 sec 68°C
	2 min 72°C	2 min 72°C	2 min 72°C

Products were cloned into the TA cloning vector (Invitrogen) and sequenced.

2.2 Protein Methods

2.2.1 Electrophoresis and Western blotting procedures

For SDS-PAGE, proteins from tissues or cells were solubilised in sample buffer (62.5mM Tris-HCl pH6.8, 1% SDS 0.1M glycerol, ß-mercaptoethanol (0.25%) and traces of bromophenol blue before separation on 10% polyacrylamide (Laemmli, 1970).

The commercial blotting apparatus (Biorad) was arranged in accordance with the manufacturer's guidelines and transferred at 100mA for 1 hour. Western blot analysis was carried out after separation of proteins by SDS-PAGE and then the protein was transferred onto the PVDF membrane at 200ma for 1 hour in transfer buffer (Tris 3.03g, glycine 14.4g, 20% methanol made up to 1 litre with ddH_2O) that had been prepared by a brief wash with methanol (change in colour was noted from opaque white to translucent grey), followed by a wash in ddH_2O for 2 minutes and then the membrane was soaked in transfer buffer (Tris glycine 20% methanol) for 5 minutes before proceeding to blotting.

Proteins were blocked in 5% milk 0.1% Tween 20 in 1X PBS for 1hour at RT. The primary antibody was added to the membrane in a sealed bag at the appropriate dilution (example 1/1000) in 1XPBS 0.1% Tween 20 for 1 hour at RT with vigorous shaking. The membrane was washed vigorously in large volumes of 1XPBS 0.1 % Tween 20 three times for 5 minutes each. The appropriate secondary HRP conjugated antibody was diluted 1/1000 in 1XPBS 0.1% Tween 20 and added to the membrane in a sealed bag and placed on a shaker for 1 hour at RT, followed by extensive washes to remove unbound antibody prior to chemi-luminescent detection (ECL Amersham Pharmacia Biotech) as described in the manufacturer's protocol and eventually visualised on photographic film.

2.2.2 Stripping blots

The blot was incubated in stripping buffer (2% w/v SDS, 62.5mM Tris-HCl pH 6.8, 100mM β-mercaptoethanol for 30 minutes at 60°C. The membrane was washed twice in PBS for 10 minutes each. The blot was then tested with fresh ECL to determine if the stripping procedure had been successful.

2.2.3 *In vitro* translation

This was carried out following manufacturer's protocol for the T7 coupled transcription/translation system (Promega). Plasmids were linearised with the appropriate restriction site downstream of the stop codon (NOTI). The DNA was phenol extracted and ethanol precipitated prior to the *in vitro* translation.

The following reaction was set up: vector DNA (1μg/μl), radiolabelled S^{35} methionine, amino acid (minus Methionine) mix, RNase inhibitor, rabbit reticulocyte. This reaction mix was incubated for 1 hour at 30°C and then incubated at 60°C to stop the reaction.

2.2.4 Electrophoresis

(See electrophoresis section 2.2.1 as for Western blotting). After electrophoresis, the gel was soaked in a fixative solution (7% acetic acid, 7% methanol and 1% glycerol) for 30 minutes. The gel was then placed onto 3mm Whatmann paper and dried for 1 hour. The dried gel was exposed to photographic film over night at RT.

2.2.5 Protein induction

The bacterial expression vector pGEX 4T-1 containing a cloned in-frame region of the *ApO* gene was transformed into a protease deficient bacteria strain (BL21). A positive clone was selected and grown overnight in 5ml LB supplemented with 50mg/ml ampicillin in a shaking orbital incubator set at

37°C and 225 rpm. A 500 ml LB culture (50mg/ml ampicillin) was seeded with 500μl of the overnight culture and grown till in log phase (optical density OD_{600} equivalent 0.6) at 37 °C and 225 rpm. At this stage IPTG was added to the culture (final concentration 0.4mM) to induce the bacterial promoter to transcribe an N-terminal GST tagged *ApO* fragment for 3 hours.

2.2.6 Purification of GST-*ApO*

2.2.6.1 Mechanical lysis

The bacteria were pelleted by centrifugation at 3000 g for 10 minutes at 4°C. The pellet was resuspended in 5ml of PBS containing protease inhibitors (complete EDTA free cocktail, Roche), aliquotted into five 1 ml eppendorfs and kept on ice. The bacteria were lysed mechanically by sonication on ice using a micro tip with the power at mid range for 15-20 second bursts. It is important to keep the samples cold during sonication to avoid heat denaturation of proteins. When the cells have been lysed there is an observable colour change of the sample and lysis can also be confirmed by viewing a fraction of the sample under a microscope. The bacteria cells appear irregular in morphology.

2.2.6.2 Sample preparation

The removal of insoluble protein aggregates and the membrane fraction was achieved by centrifugation at full speed for 10 minutes. The upper soluble fractions were recovered (contains soluble proteins), and were diluted in 20ml binding buffer (PBS pH 7.3 (140mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.3 and protease inhibitors). A large volume is required to reduce sample viscosity and prevent the column clogging.

2.2.6.3 Purification

This stage was carried out at RT unless otherwise stated. The glutathione sepharose 4B (1 ml GSTrap FF columns Amersham Pharmacia Biotech) was connected to the peristaltic pump set at flow rate of 1 ml per minute and the column was equilibriated with 5 ml of binding buffer (140mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.3 and protease inhibitors). The sample was applied to the column. Following this the column was washed with 20 ml of binding buffer to remove unbound contaminating proteins. The elution of the bound proteins was achieved with 5mls of elution buffer (50mM Tris-HCL, 10mM reduced glutathione, pH 8.0). Fractions of the elution were collected and the concentration of the GST fusion protein was estimated by measuring the absorbance at 280nM. The GST tag can be approximated using the conversion A_{280} equivalent to 1 corresponds to approximately 0.5mg/ml.

2.2.7 Concentrating proteins

Two different methods were used:

1). Centricon spin columns (Millipore) with the appropriate molecular weight cut off (MWCO) in accordance with the manufacturer's guidelines. The protein sample was loaded into the retention vial and this was inserted into the filtrate vial and centrifuged (3000 g) for 1 hour at RT.

2). Protein was loaded into a dialysis membrane of appropriate MWCO and placed in a vessel to which high molecular weight polyethylene glycol (PEG) was poured. The PEG absorbed water through the dialysis membrane leaving the protein concentrated.

2.2.8.1 His tag purification

2.2.8.2 Protein induction

The *ApO* 2.3 kb transcript was subcloned into pET32a N-terminal His tag bacterial expression vector. This was transformed into a protease deficient bacteria strain (BL21) by heat shock method (see section 2.1.10). A positive clone was selected and grown overnight in 5ml LB supplemented with 50mg/ml ampicillin in a shaking orbital incubator set at 37°C and 225 rpm. A 500 ml LB culture (50mg/ml ampicillin) was seeded with 500μl of the overnight culture and grown until log phase (optical density OD_{600} equivalent 0.6) at 37°C and 225 rpm. At this stage IPTG was added to the culture (final concentration 1mM) to induce the His tagged APO protein expression for 5 hours.

2.2.8.3 Purification of insoluble aggregates from inclusion bodies

The cells were harvested by centrifugation at 3000 g for 10 minutes. The supernatant was decanted. The cell pellet was resuspended in 200ml PBS and then sonicated with maxi probe (set at medium setting) for 5 minutes. The sample was centrifuged 3000 g to collect inclusion bodies and cellular debris.

The supernatant was removed and the cell pellet resuspended in 25 ml of 1X binding buffer (0.5M NaCl, 20mM Tris-HCl, 5mM imidazole pH 7.9) containing 6M urea. The sample was incubated on ice for 1 hour to dissolve the protein and then centrifuged at 12,000 g (Sorval) to remove insoluble material. The supernatant was filtered through a 0.45μm membrane.

2.2.9 Dialysis

The sample was dialysed through a membrane of (molecular weight cut off 10 kDa) overnight at 4°C in 5 litres of PBS to remove the 6M urea denaturant. The sample was then concentrated by the PEG method so that the final volume was approximately 10 ml (see section 2.2.7).

2.2.10 Batch method purification

The resin was prepared by taking 1ml of the slurry (Novagen) and centrifuged to remove the supernatant. The resin was washed in 2ml ddH_2O, centrifuged at 500g for 1 minute. This step was then repeated. The resin was washed three times in 2ml of charge buffer (50mM $NiSO_4$) and then washed twice in 2ml of binding buffer (0.5M NaCl, 20mM Tris-HCl, 5mM imidazole pH 7.9).

The dialysed concentrated sample was added to the charged resin and incubated overnight on a rotating wheel at RT.

The resin was washed three times in 30ml of binding buffer and then two times in 50ml wash buffer (0.5M NaCl, 20mM Tris-HCl, 60mM imidazole pH 7.9)

The protein was eluted from the resin with elution buffer (1M imidazole, 500mM NaCl, 20mM Tris-HCl pH 7.9). After a final centrifugation at 500 g the supernatant was removed to a fresh tube.

2.2.11 Irreversible binding of purified HisTag protein to activated cyanogen bromide sepharose

The APO-His tagged protein was concentrated on a Centricon column (molecular weight cut off 10 kDa) and spun at 3000 g for 1 hour. The concentrated protein was desalted on an NAP-10 column (Amersham Pharamcia Biotech) and then diluted into coupling buffer (0.1M $NaHCO_3$, 0.5M NaCl pH 8.3).

1ml of the dried cyanogen bromide (CnBr) matrix was prepared by adding this to 5ml of 1mM HCl for 2 hours. The activated sepharose was centrifuged at 3000 g for 1 minute and then washed with three changes of coupling buffer (0.1M $NaHCO_3$, 0.5M NaCl pH 8.3). The purified His tag protein was bound to the activated CnBr overnight at 4°C on a rotating wheel.

The activated CnBr was then treated with 5 sequential washes of a high pH Tris buffer (0.1M Tris-HCl, 0.5M NaCl pH 8.0) and low pH acetate buffer (0.1M sodium acetate, 0.5M NaCl, pH 4). This had the effect of deactivating any charged unbound CnBr preventing more proteins binding irreversibly to the matrix.

The rabbit serum (2.5 ml) was diluted into 25 ml PBS and then filtered through a 0.45μM membrane. The diluted serum was added to the cyanogen bromide matrix and incubated overnight at 4°C.

The cyanogen bromide matrix was washed 6 times in 50ml of PBS to remove unbound proteins. The bound antibody was removed from the His tag APO with the addition of 2 ml elution buffer (0.1M glycine pH 2.7). The beads were spun 3000 g for 2 minutes and the supernatant transferred to a tube containing 500μl neutralisation buffer (1.5M Tris pH 8.8). The purified antibody was stored in the -20°C freezer after the addition of sodium azide (0.01%).

2.2.12 Antibody production

1mg of the purified GST-APO protein was supplied to Cell Diagnostics (Scotland), who raised an antibody in a single rabbit. The animal was bled prior to administering the antigen (preimmune) and then after the antigen was injected subcutanously at weeks 8, 12 and 16.

2.2.13 Immunoprecipitation

Cells were lifted from the tissue culture plate with trypsin (2.3.2). The cells were centrifuged at 1000 g for 5 minutes. The pellet was washed twice with PBS prior to lysing with 1ml of lysis buffer for 30 minutes at 4°C and centrifugation at full speed for 15 minutes. The supernatant was removed to a microfuge tube and pre-cleared with 20μl of protein A (1 mg/ml Roche). The microfuge was spun at 1000 g for 10 minutes and then the sample supernatant was split between two microfuge tubes (500μl each). To one tube, 5μl of antibody was added and incubated overnight at 4°C on a rotating wheel.

50μl of Protein A was added to both tubes and incubated at 4°C for two hours. The tubes were centrifuged at 1000 g for 2 minutes and then washed with 1 ml of PBS on a rotating wheel for 5 minutes. This procedure was repeated six times. 50μl of sample buffer (62.5mM Tris-HCl pH6.8, 1% SDS 0.1M glycerol, ß-mercaptoethanol 0.1%) was added to each of the tubes and they were heated at 90°C for 10 minutes tubes. The samples were then Western blotted and probed with appropriate antibodies (see section 2.2.1).

2.3 Cell culture

2.3.1 Thawing cells

Vials of frozen COS7 cells were taken directly from storage at -140°C and thawed quickly at 37°C in a waterbath. Cells were transferred to warm GMEM (Invitrogen) medium and centrifuged at 200g for 5 minutes. The medium was removed and the pellet resuspended with fresh medium prior to seeding a 25cm^2 flask. The medium was replaced after 8 hours to remove dead cells.

2.3.2 Transfection of DNA into mammalian cells

1μg of Vector DNA was diluted in 50μl of serum free Opti-MEM (Invitrogen). 5μl of Lipofectamine 2000 (Invitrogen) was mixed with 50μl of serum free Opti-MEM and incubated for 5 minutes at room temperature. After the 5 minute incubation the lipofectamine was combined with the vector DNA and incubated for 20 minutes at RT. Then the vector/lipofectamine mix was added to the cells and incubated for 48 hours at 37°C in a humidified 5% CO_2 atmosphere.

2.3.3 Cell fractionation

A nuclear and cytoplasmic extraction kit was used following the manufacturer's protocol. COS7 cells were removed from the culture dish with trypsin (0.025% trypsin, 0.1% chicken serum and 1.3mM EDTA in PBS). The cells were spun in a centrifuge at 1000 g for 5 minutes. The pellet was washed twice with PBS. To an estimated packed cell volume of 20μl in a 1.5ml microcentrifuge tube, 200μl reagent cytoplasmic extraction reagent CER I was added. The tube was vortexed for 15 seconds to fully resuspend the pellet and incubated on ice for 10 minutes. To this mix, 11μl of reagent cytoplasmic extraction reagent II (CER II) was added. The tube was vortexed for 5 second and then placed on ice for 1 minute. The tube was vortexed again for 5 second and then centrifuged at maximum speed (10,000 g) for 5 minutes. The supernatant containing the cytoplasmic fraction was transferred to a fresh tube and stored. The pellet containing the nuclei was resuspended in 100μl ice cold nuclear extraction reagent (NER) by vortexing for 15 seconds every 10 minutes over a total 40 minutes incubation time. The nuclear fraction was centrifuged at maximum speed (10,000 g) for 10 minutes. The supernatant containing the nuclear proteins was transferred to a fresh tube and stored with the cytoplasmic fraction in a -20°C freezer.

2.4 Histology

2.4.1 Animal breeding

The GT411 gene trap was transmitted through the germline by injection of C57/Bl6 blastocysts (Jennifer Nichols at ISCR). Chimaeric males were back-crossed for 3 generations to wild type MF1 out bred females before intercrossing heterozygous animals. Adult animals were genotyped by staining an ear-punch biopsy or the yolk sac respectively for βgal activity. Homozygous and heterozygous animals for the gene trap integration were distinguished by the speed of appearance and the intensity of the staining reaction. The reliability of this genotyping method was confirmed by breeding analysis.

2.4.2 Cryostat sections

Tissues were freshly dissected, placed in OTC, immediately frozen in liquid nitrogen prior to sectioning on a cryostat following which 10μm sections were cut at blade temperature at -20^{0}C. The sections were lifted onto commercially bought slides (super frost VWR), air dried at RT. For long term storage the slides were kept at -20^{0}C.

2.4.3 ßgal staining method

The βgal staining protocol requires the tissue or cells to be fixed (fix solution 0.2% glutaldehyde, 5mM EGTA pH 7.3, 2mM $MgCl_2$ in PBS pH 7.3) for 10 minutes at room temperature. The embryos were then washed three times in wash buffer (2mM $MgCl_2$; 0.01% deoxycholate; 0.02% Nonidet (NP40) in PBS pH 7.3), for 10 minutes each at room temperature. Finally, the tissues were stained in filtered X-gal stain (1mg/ml X-gal dissolved in di-methyl formamide, 5mM potassium ferrocyanide and 5mM potassium ferricyanide in wash buffer) until the desired level of staining was reached.

The reaction was stopped by several rinses in wash buffer and then samples were stored at 4°C (Gossler et al., 1989).

2.4.4 Dissection of embryos

Timed matings were set up for the collection of embryos at specific stages of gestation (day of plug designated as day 0.5). Pregnant females were culled by a schedule one method (cervical dislocation). The uterus was dissected and placed in cold phosphate buffer (PBS). The extra embryonic tissues were removed and the foetus collected.

2.4.5 Aortic ring assay (Nicosia and Ottinetti, 1990a) (Nicosia and Ottinetti, 1990b)

Adult mice were killed by CO_2 gas exposure and aorta were dissected and gently flushed with cold PBS using a hypodermic needle to remove excess blood. The aorta was cut into 1mm cylindrical sections, embedded in Matrigel (Beckton Dickinson) in accordance with manufacturer's guidelines. They were then covered with an endothelial basal medium with growth supplements provided with the media (Clonetics) and incubated at 37°C, 5% CO_2, 20% oxygen. The aortic sprouts were monitored regularly over the next 12 days using light microscopy.

2.4.6 Microscopy and photography

Cells were examined with a Leitz Labovert light or Zeiss Axiovert 25 flourescence microscope. eGFP was visualised with UV light through Zeiss filter #44. Embryo dissections were performed using a Zeiss Stemi 2000-C dissecting microscope. Cells and embryos were photographed with an Axiocam digital camera using AxisVision 3.1 software.

Chapter 3:

Molecular characterisation of the gene trap integration GT411

3.1 Introduction

The gene trap integration GT411 had been isolated from a screen to identify novel genes involved in cardiac and vascular development. Using 5'RACE and EST databases, a partial open reading frame (ORF) had been determined. This chapter describes 5'RACE and RT-PCR experiments performed to generate a full length ORF. Northern blot analysis was carried out on tissues to gain insight into the expression and availability of tissue specific transcripts. Bioinformatic analysis of the cDNA and protein sequences enabled homology and functional characterisation of the gene. The mouse and human gene were mapped and candidate disease databases searched.

3.2 Results

3.2.1 Identification of the trapped gene by 5'RACE-PCR

To identify the gene associated with the GT411 gene trap integration primers complementary to the gene-trap vector (EN2) were used to identify 206 base pairs (bp) of sequence in a 5'RACE experiment (Townley et al., 1997) (Figure 3.1a). This sequence was used to screen the DNA databases and a contig of overlapping ESTs (accession nos: AA560030, AA003012 and AA755402) was generated that coded for an ORF of 177 amino acids (Figure 3.1b) (this work was done by Dr J.Wallis).

Studies reported in this thesis began by performing further 5' RACE analysis on this sequence to obtain the full-length coding sequence of the trapped gene. Experiments were performed on random primed first strand cDNA from 11.5 day post coitus (dpc) whole embryos with PCR primer P3 and anchor primer AP1 using a Gene Racer kit (Invitrogen) (primer sequences appendix Table 2). This was followed by nested PCR with primer P2 and anchor primer AP2. Three fragments were obtained of approximate sizes 650, 750 and 500 bp, (fragment 1, 2 and 3 shown schematically in Figure 3.1c,

3.1d and 3.1e respectively) and were TA cloned (Invitrogen kit) and sequenced.

Fragment 1 and 2 were identical except that band 2 extended further 5' and the reading frame was maintained in-frame with the 3' end of the gene. Fragment 3 had a different 5' end to bands 1 and 2. Sequencing and database analysis of the genomic DNA revealed that it represented unspliced message and that the sequence at the 5' end was part of an intron (data not shown). To clarify this further, Northern blot analysis, using a 300 bp probe (N3) that was specific to the 5' end of this sequence failed to show a signal, supporting the finding that it did not represent a true transcript (data not shown).

Additional rounds of 5'RACE were carried out with primers P20 to anchor primer AP1 and nested PCR with primers P21 to anchor primer AP2 on 11.5 dpc embryo cDNA. The resulting band was TA cloned and sequenced. A putative methionine start codon with an in-frame stop codon upstream was found in this fragment and this was predicted to be the start of the translated gene (Figure 3.1f). An alternative transcript was also identified in further RT-PCR experiments (Figure 3.1g, section 3.4.2).

3.2.2 Verification of predicted transcripts by RT-PCR

To confirm the cDNA structure that was predicted by 5'RACE, extensive RT-PCR analysis was performed with gene specific primers using first strand cDNA generated from RNA isolated from day 11.5 dpc embryos (Figure 3.2a).

In most instances amplicons of the predicted size were obtained (Figure 3.2b) and subsequent sequencing of the bands confirmed their identity (data not shown). However, using primers P26 - P36, two bands were obtained. The smaller was the expected size (459 bp) and the other a larger band of 1393 bp (Figure 3.2b lane 1). Following cloning and sequencing the smaller band (459 bp) was shown to be the expected sequence (Figure 3.1f), whilst the 1393bp band contained an alternatively spliced exon (934 bp) that was in-

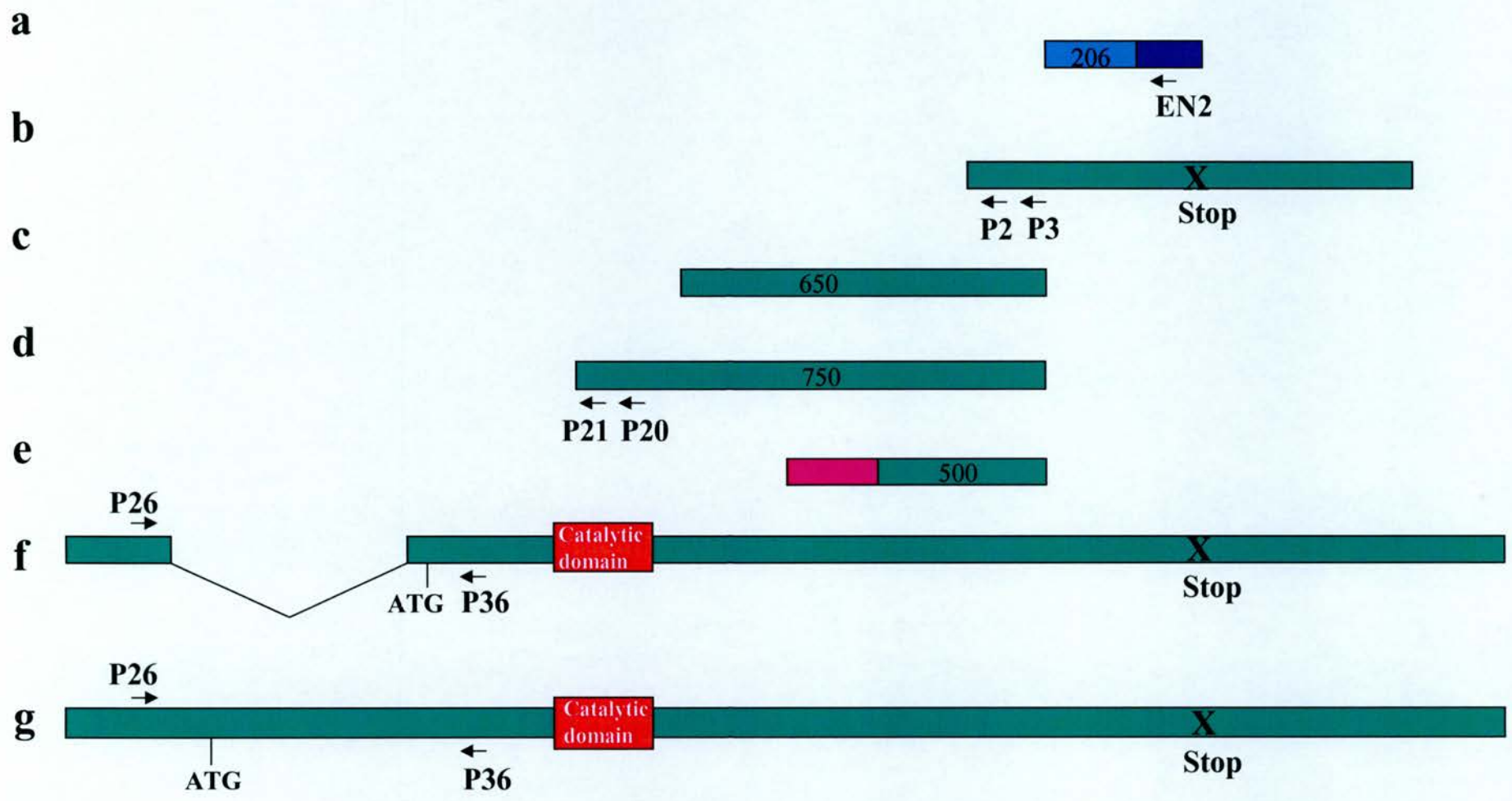


Figure 3.1: Schematic diagram of *ApO* cDNA transcripts determined by 5' RACE (Rapid Amplification of cDNA Ends).
a) Original sequence from RACE using primer EN2. b) EST contig build at start of project Jan 2001.c,d and
e) cDNA generated from 5'RACE with primers P3 and P2. Genomic intron sequence (e) shown in purple.
f) cDNA generated from 5'RACE with primers P20 and P21and g) alternative spliced 934 bp exon.
The predicted sequence that encodes the catalytic domain is shown in red (not drawn to scale).

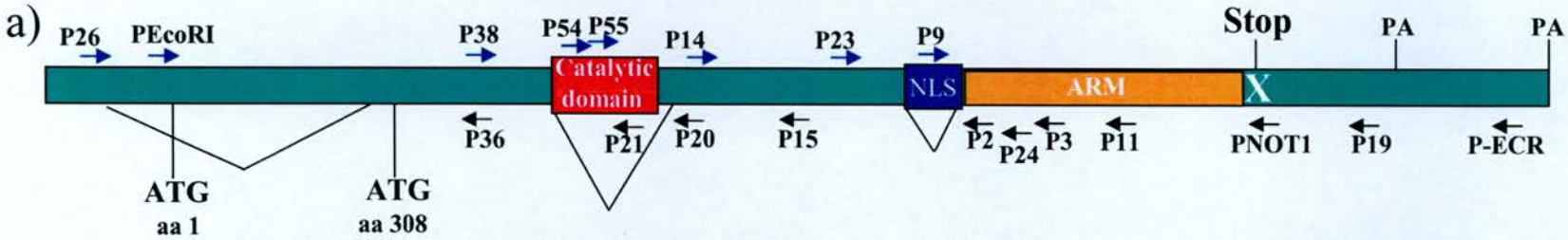
frame with the 3' end of the gene (Figure 3.1g). This sequence was highly conserved between mouse, rat and human both at the amino acid and nucleotide sequence and was further supported by EST data (Accession numbers, DV645899, BF100003 and BU512973). This new transcript would have a different putative start codon (Figure 3.1f and 3.1g).

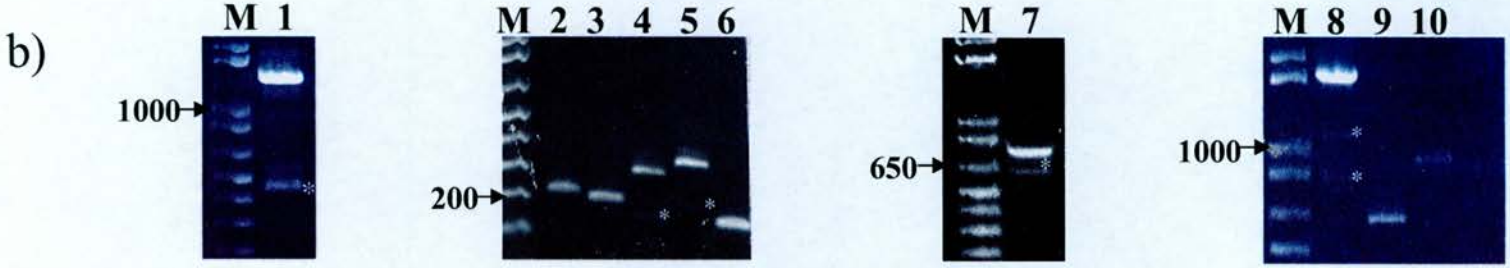
Another alternative splicing event of a 144bp exon at nucleotide position 2262 bp was observed with primer pairs P23 - P24, P23 - P3 and P14 – P11 (Figure 3.2b lanes 4, 5 and 7). Interestingly, this exon contains the predicted nuclear localisation signal and lends support to the idea of differentially spliced tissue specific transcripts affecting their distribution and function within a cell (chapter 5). Using a primer within this alternatively spliced (144 bp) exon P9 to the downstream primer, P3, a predicted single PCR product of 130bp was obtained (Figure 3.2b lane 6).

3.2.3 Generation and analysis of full-length clones

To study the gene in more detail full-length clones were generated and cloned into a range of vectors (Table 1, Appendix 1.2). A full-length clone (Figure 3.1g) was made by RT-PCR with primer pairs P26 and ECR followed by nested PCR with primer pairs PECOR1 (at the putative start codon) and PNOT1 (primer at stop codon) using adult heart cDNA. The 2.3kb PCR product was cloned into a mammalian expression vector pCDNA3.1. It was sequenced in its entirety to confirm identity and to ensure the absence of RT-PCR generated mutations. However, surprisingly, sequencing revealed that the clone was an alternatively spliced variant removing a 297bp exon. This exon encoded the putative aminopeptidase catalytic domain. The EST database also has transcripts that lacked the exon encoding the catalytic domain (AK027581, NM_032823) which supports their authenticity.

To investigate this further, PCR primers specific to the catalytic domain were used in a 3'RACE experiment (P54 – AP1 and nested P55 –





Lane	Primers		Size (bp)	Size (bp)	Size(bp)
1	P26	P36	460	1393	
2	P14	P15		230	
3	P23	P2		185	
4	P23	P24	154	298	
5	P23	P3	192	336	
6	P9	P3		130	
7	P14	P11	636	780	
8	P26	P21	842	1776	1050
9	P55	P3		656	
10	P55	PNOT1		951	

Figure 3.2: RT-PCR analysis of the mouse aminopeptidase O gene.
a) cDNA structure of *ApO* and primers used for RT-PCR (not drawn to scale). b) RT-PCR data with primer pairs across the gene (see table for specific details). Asterisks indicate alternative splicing

AP2, Gene Racer kit, Invitrogen). No new additional transcripts to those already described were detected and further analysis of the bioinformatic databases concurred with these findings.

The possibility existed that the transcript containing catalytic domain might be expressed to perform a particular biological function during a particular stage of development or process and that alternative splicing may be a mechanism to regulate protein function. Transcripts containing that sequence encoding this domain might therefore not be present in the RNA sample used. RNA was extracted from a wider range of tissues, including embryonic stem cells (ES cells), 11.5 day post coitus (dpc) embryos and adult tissues and PCR performed with primers spanning the catalytic domain (P38-P15). All the PCR products were identical and sequencing revealed that none of them contained sequence that encoded catalytic domain (data not shown).

Gene specific primers within the putative catalytic domain were then used in the following over lapping PCR reactions P26 to P21, P55 to P3 and P55 to PNOTI (Figure 3.2b lanes 8, 9 and 10). Products of the correct predicted sizes were obtained on a range of tissues (heart, lung and ES cell). Sequencing confirmed them to contain their respective part of the putative catalytic domain. Furthermore, the EST database has transcripts that contain this domain (Accession numbers, AK057450, BU512973 and BY708307).

Two faint bands were also seen at approximate position 842 bp and 1050 bp (Figure 3.2b lane 8 (marked by asterisks)). Although these bands were not sequenced the 842bp band would have been predicted since this would be the expected size of the PCR product missing the alternatively spliced 934 bp exon. The other band might represent another splice variant as yet not determined.

There are 2 polyadenylation signals (PA (AATAAA)) in the 3'UTR of the *ApO* gene. They are at position 3238 and 3822 (Appendix 1.4), and this corresponds to 490 bp and 1174 bp downstream of the termination codon respectively. Polyadenylation signals can stabilise RNA molecules and can be a mechanism for gene control such as where an immunoglobulin gene can

produce alternative transcripts by alternative RNA splicing and use of different poly A signals (Early et al., 1980b). Using PCR primers (P38 – ECR and P38 - P19)(Figure 3.2a) specific to different regions of the 3'UTR and by performing nested PCR, it did not appear that there was a preference for a particular Poly A signal with any given transcript. Furthermore, transcripts did not contain sequence encoding the catalytic domain which suggests that the poly A used does not affect the stability of a transcript with regard to the catalytic domain.

3.2.4 Genomic structure of the *ApO* locus

The genomic structure of *ApO* locus was determined by screening the mouse public database with the cDNA sequence. The entire predicted cDNA was homologous to parts of a contiguous region of chromosome 13 (AC069013). The gene structure was determined by placing the exons in order and calculating the intron sizes (Figure 3.3a and 3.3b). The exon structure was confirmed with the HGMP NIX software programme.The gene consists of 16 coding exons (conserved across species) and the entire murine genomic region spans approximately 320 kb. The intron sizes range from 100 bp, (between exon 12 and 13), to the largest at 50 kb (between 5'UTR and coding exon 1 (http://genome.ucsc.edu.)). Bioinformatic analysis has predicted CpG island several kilobases upstream of the predicted initiation codon.

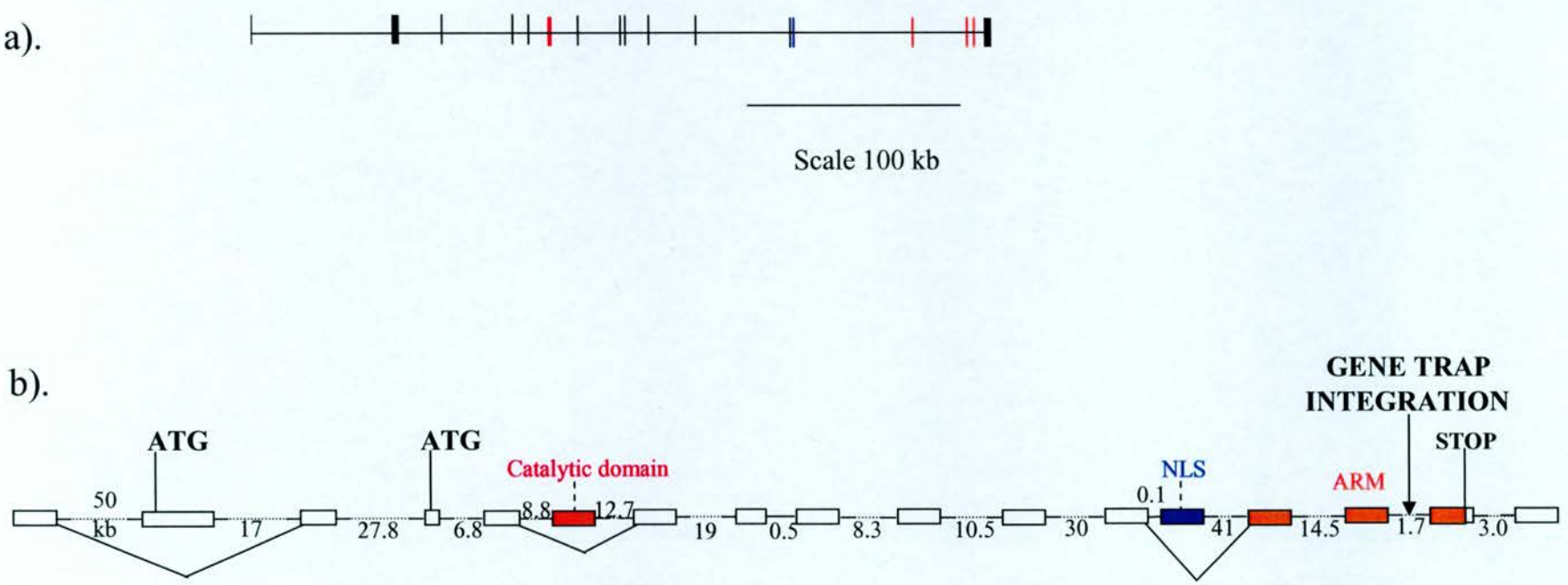


Figure 3.3. Schematic representation of the genomic structure of mouse aminopeptidase O (*ApO*).

a) Genomic structure (drawn to scale) with exons indicated as vertical lines, red line catalytic domain, blue line the nuclear localisation signal (NLS) and orange is the putative armadillo repeat (ARM).

b) Schematic diagram of *ApO* showing inton sizes, predicted start and stop sites, alternative splicing events and the regions predicted to encode functional domains. The predicted position of the gene trap vector integration in the last intron is shown.

3.2.5 Genomic integration site

It was predicted that the vector had integrated into the intron between exons 15 and 16 because the sequence originally identified by 5'RACE was identical to the 3' end of exon 15.

To determine the exact site of the genomic integration, a PCR reaction was carried out with an *ApO* gene specific primer and a gene trap vector specific primer (Exon F- EN2). The PCR product was sequenced to confirm the site of integration (data not shown). This information was used to design a genotyping assay for the GT411 gene trap mice (Appendix 1.1).

3.2.6 Northern blot analysis

Northern blot analysis was used to examine the expression of *ApO* in mRNA isolated from a range of adult tissues. Using a C-terminal probe (N1), two predominant transcripts of approximately 2.1 and 2.3 kilobases (kb) were detected in all tissues tested and it is presumed that these are the + and – catalytic domain isoforms, although they could also be the +/- 144bp splice variant. In addition other transcripts ranging in size between 3.2 kb and 4.6 kb were also detected (Figure 3.4a).

The same filter was stripped and re-hybridised with a probe (N2), specific to the alternatively spliced isoform (934bp) at the N-terminus. Some of the larger transcripts that were detected with probe N1 were also detected with probe N2 (figure 3.4b red asterisks).

Interestingly, however, some of the larger transcripts seen with the N1 probe were not detected with probe N2. For example, in heart the larger transcripts are absent indicating the possibility of additional tissue specific transcripts that as yet have not been determined (figure 3.4a and b). A *Gapdh* probe was used to control for loading.

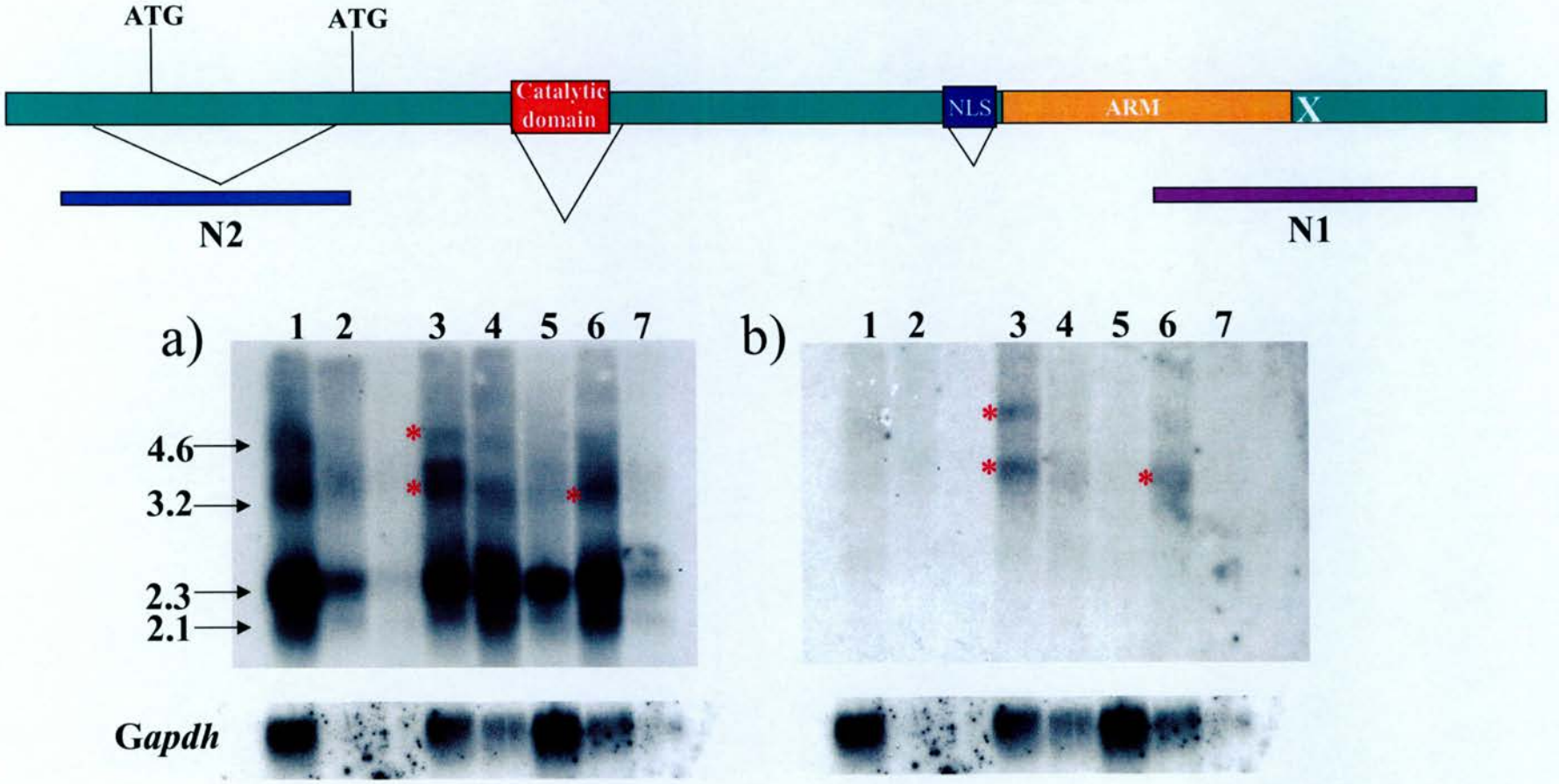


Figure 3.4: Northern blot analysis of mRNA isolated from adult tissues.
heart (1), brain (2), lung (3), liver (4), skeletal muscle (5), kidney (6), and ovary (7) using probes N1(a) and N2 (b) (indicated on schematic). Red asterisks show common bands detected with probes N1 and N2. G*apdh* loading controls.

3.2.7 Protein homology search

A bioinformatic blastP analysis with the murine APO peptide sequence has 40% homology to a protease called aminopeptidase B. This protein is involved in processing mechanisms during spermatid differentiation and in inflammatory processes (Cadel et al., 1997). The consensus catalytic domains are conserved and both have the putative armadillo repeat motif. Additionally, two other aminopeptidase proteins also have this structural motif namely, LTA4 hydrolase and aminopeptidase B-like. It has been suggested the presence of this 170 residue C-terminal domain may define a novel subfamily of aminopeptidases which have been termed ARM aminopeptidases (Diaz-Perales et al., 2005).

3.2.8 Predicted Aminopeptidase O protein structure

Using protein databases, (Simple Modular Architecture Research Tool (SMART) and other software at the Expasy site (http://www.expasy) the protein had a predicted consensus sequence typical of the zinc aminopeptidase M1 family. This sequence has the essential zinc binding site HEXXH(X) 18E (amino acid position 480 Figure 3.6) with a second glutamic acid residue (E)18 amino acids downstream (http://www.bi.bbsrc.ac.uk/Merops) (Rawlings et al., 2002).

Further analysis of the protein revealed an armadillo (ARM) repeat motif at the C-terminus identified by the SMART programme (smart.embl-heidelberg.de).

3.2. 9 Nuclear localisation signal (NLS)

A putative nuclear localisation sequence (NLS) was not predicted by bioinformatics programmes, but selected by eye on the basis that the protein sequence consists of a region of highly basic amino acid residues at position 693-703 (RRPRKRKRGKR) near to the C-terminus of the gene (description of the experimental evidence and biological relevance of the NLS are

discussed in chapter 5). Interestingly the 144bp exon that contains the putative NLS can be spliced out (Figure 3.2b lanes 4, 5 and 7) and this may be a regulatory mechanism for gene function.

3.2.10 Conservation of *ApO* between species

A comparison of the mouse APO amino acid sequence revealed that it was 72%, 95% and 38% homologous to the human, rat and puffer fish respectively. The putative peptidase domain HEXXH-E_{18} is totally conserved between all these species (Figure 3.5). An additional homologue has been found in *Xenopus tropicalis*, and although there are some gaps (missing regions), the sequence similarity is 48% between mouse and frog over the defined region (data not shown).

The putative start codon for the longer transcript, MetDIK (amino acid position 1) was considered to be the initiation codon as there was an in-frame stop codon 11 amino acids upstream. This initiation codon is conserved in mouse, human and rat and there is a good Kozak consensus sequence in the context of this initiation codon (Kozak, 2002). Two additional amino acids, Met**EP**D are present in the sequence derived from fish (Figure 3.5). In mouse, human and rat the protein would have an 823 amino acid ORF with a calculated protein mass of 90 kilo Daltons (kDa) whilst in fugu the ORF would be 825 amino acids.

The shorter major transcripts (2.1 and 2.3 kb) could use a variety of different initiation codons. Different start sites effectively lead to variant isoforms, potentially with different functions but it is not an uncommon phenomenon (Yan et al., 1998). It was considered that the start codon at position 332 MetPMPAS was a good candidate because of its proximity to a good Kozak sequence (Figure 3.5) and it is conserved between species. The protein generated from this initiation codon would be 491 amino acids (with catalytic domain) and would have a calculated protein mass of 60 kDa.

CLUSTAL W (1.83) multiple sequence alignment

```
mouse-ApO       M--DIKLDPSRDDLPLMANTSHMLVKHYILDLDVDFGNQVIEGNIVLFFG----DGNRFK
rat-ApO         M--DINLDPSRDDLPLMANTSHLLVKHYVLDLDVDFGSRVLEGNIVLFFG----DGNRCK
human-ApO       M--DIQLDPARDDLPLMANTSHILVKHYVLDLDVDFESQVIEGTIVLFLE----DGNRFK
fugu-ApO        MEPDKDLHPDTDDLPLRANTNHILVRHYVLNLAVHIDRQVISGSVVLFLEPCPVDGTNSE

mouse-ApO       NQSRSTQETFQMESEEADIFRTAEPCHVPEMDSSTFSLKMGHRECAVCGKGDQDAFDNDG
rat-ApO         KQSSSSQETFQMESEEAYILRTAEPCHVPKMDSNTFSPKMGHREFAVFGKSDQDAFDNDG
human-ApO       KQNSSIEEACQSESNKACKFGMPEPCHIPVTNARTFSSEMEYNDFAICSKGEKDTSDKDG
fugu-ApO        KDDNEIVRPIQAGGRRDGSARPIQDPNLKRESERNFGKSVSH----------KNTLDKFQ

mouse-ApO       THDNQERDSEISSSKYCCDTGNHGKRDFLLVLDCCDLSVLKVEEVDVAAVPGLEKFTK--
rat-ApO         NHDNKEHDSESSSSKYCCDTGNHGREDFLLVLDCCDLSVLKVEEVDVAAVPDLEKFTK--
human-ApO       NHDNQEHASGISSSKYCCDTGNHGSEDFLLVLDCCDLSVLKVEEVDVAAVPGLEKFTR--
fugu-ApO        VSSGATLQSSPGWEHTS-------DEDFTLVLDCCDIDVFKVEEVDVSSLSSMPDLQSGS

mouse-ApO       APKLLATP--EKLRCEIVRDLVALPADAWREQLDCYTRCSQAPGCG---ELLFDSDNWSL
rat-ApO         APNLTAAP--EKRRCEIVRDLVALPADAWREQLDCYTRCSQAPGCG---ELLFDSDKWSL
human-ApO       SPELTVVS--EEFRNQIVRELVTLPANRWREQLDYYARCSQAPGCG---ELLFDTDTWSL
fugu-ApO        SSETSAVPSWSLKSAAFVQNVISMPSSQWKQKHQLYLLSSHAPGVQDGSSLHFYRDQWSV

mouse-ApO       RIRKTGTSTPADFPRAIRIWYKTKPEGQSVAWTTDQNGRPCVYTMGSPINNRALFPCQEP
rat-ApO         QIRKKGVPTAADFPHAIRIWYKTKPEGQSVTWTSDQDGRPCVYTMGSPINNRALFPCQEP
human-ApO       QIRKTGAQTATDFPHAIRIWYKTKPEGRSVTWTSDQSGRPCVYTVGSPINNRALFPCQEP
fugu-ApO        QVRKIGISSPLEFPRALRICYETKPSGGSVRWTKDQDNRVCVYTAGSPINNRALFPCQEP

mouse-ApO       PVAMSTWQATVRAAASFVVLMSGENSAKPTPLREG-YMSWHYYVTMPMPASTFAIAVGCW
rat-ApO         PVAMSTWQATVRAAASFVVLMSGEKSAKPTPLREG-YMSWHYYVTMPMPASTFTIAVGCW
human-ApO       PVAMSTWQATVRAAASFVVLMSGENSAKPTQLWEE-CSSWYYYVTMPMPASTFTIAVGCW
fugu-ApO        PVALSTWQATIRAPCDCLVLMSGEEETPPINDEDTRFLIWNYYVTMPMPASTFTLAVGHW

mouse-ApO       TEMKPKASPPDDLMTEHSLPLSPSEVDLRYDNTCNHMEYPCRFQSASAASQDIIPYRVFA
rat-ApO         TEMKPKTSPLDDL-TEHTLPLRPSDADFRYGNTCSHMEYPCRFQSASAATQNIIPYRVFA
human-ApO       TEMKMETWSSNDLATERPFSPSEAN--FRHVGVCSHMEYPCRFQNASATTQEIIPHRVFA
fugu-ApO        QQVT-----------AEIPLESGEGMADEGIFCSHGDYPCRFTQWSARSQRAIPHRVFG

mouse-ApO       PVCLEGACQEALLWLIPSCLSAAHSVLGTHPFSRLDILVVPTNFPSLGMANPHIIFLSQS
rat-ApO         PVCLEGACREALLWLIPSCLSAAHSVLGTHPFSRLDILIVPTNFPSLGMASPHIIFLSQS
human-ApO       PVCLTGACQETLLRLIPPCLSAAHSVLGAHPFSRLDVLIVPANFPSLGMASPHIMFLSQS
fugu-ApO        PVSLL--QKAQVLKLLPGCLAAAYTVLGVHPFPRLDVLIVPAGFSSLGMASPHIIFLSQS
```

Figure 3.5: Protein sequence alignment of APO between mouse, rat, human and fugu.

The two putative start sites are shown in green and purple, the conserved catalytic domain ($HEXXH(X)_{18}E$) in red and the nuclear localisation signal (NLS) RRPRKRKRGKR in blue and armadillo repeat in orange.

3.2.11 Mapping the Aminopeptidase O gene in mouse and human

At the beginning of the project, bioinformatic mapping of the human aminopeptidase O gene was found on chromosomes, 9 and 18 and at that time it was considered important to clarify this ambiguity. The accurate chromosome location of the gene would enable phenotypic and linkage mapping searches of disease databases.

Using human genomic PCR primers HGT411/5 and HGT411/6 on a human/mouse/hamster somatic cell hybrid panel the human gene was mapped to chromosome 9 (Figure 3.6a). Additionally, with the same PCR primers, the human homolog was fine mapped on the Gene-bridge 4 human/hamster radiation hybrid panel (Gyapay et al., 1996)(Figure 3.6b). The human gene is found at q22 chromosome 9, a region syntenic to mouse chromosome 13. The experimental mapping has since been confirmed by analysis of the new published mouse (http://genome.ucsc.edu. chromosome 13 (61300451-616357741)) and human (http://genome.ucsc.edu. chromosome 9 (94568538-94928995)).

Fluorescent in situ hybridisation (FISH) analysis was performed using a labelled PAC clone (583C16) on mouse metaphase chromosomes. The *ApO* gene was mapped to chromosome 13 C1 (M.Lee MRC) (Figure 3.7).

3.2.12 Online Mendalian Inheritance of Man (OMIM)

Phenotypic and linkage mapping services have not as yet revealed any candidate human genetic diseases to this locus although this region is deleted in certain cancer types including basal cell and oesophageal carcinomas, (Lichun et al., 2004)bladder (Obermann et al., 2004; Simoneau et al., 2000)and ovarian cancer (Byrom et al., 2004).

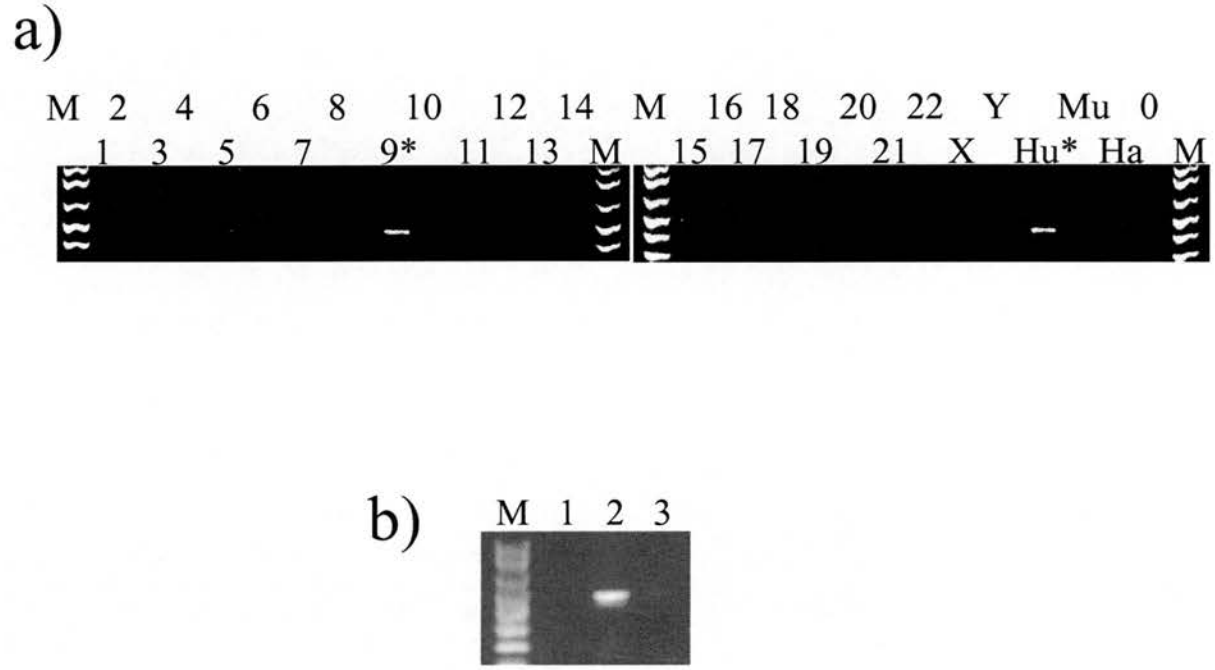


Figure 3.6: Genomic mapping of human aminopeptidase O gene.

a) PCR with primers HuP5-HuP6 on human somatic cell hybrid library. The PCR product (250bp) in lane 9 indicates that human APO is located on chromosome 9. Hu =human DNA, Mu = mouse DNA and Ha =hamster DNA. M= molecular weight marker

b) Radiation hybrid test PCR with primers HuP5-HuP6 showing human specific amplification.
(1), hamster DNA, (2) human DNA and (3) negative control.

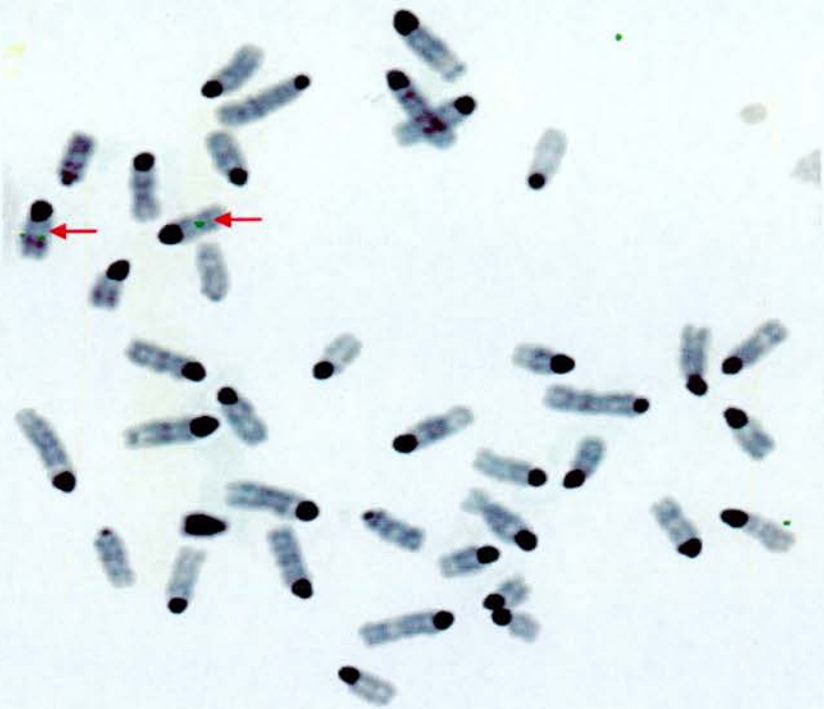


Figure 3.7 Fluorescent *in situ* hybridisation (FISH) of mouse metaphase chromosomes.

Arrows point to hybridised probe on chromosome 13 in green

3.3 Discussion

An aminopeptidase, a member of the M1 family of metallopeptidases, has been identified by gene-trapping. The ORF was determined by 5'RACE and alternative splice transcripts were demonstrated by RT-PCR. The gene is ubiquitously expressed in the range of tissues tested by Northern blot. Furthermore, analysis by Northern confirmed that there are alternative transcripts detected with probes to different regions of the gene. A particularly interesting observation was that the larger transcripts which were initially observed in the heart lung and kidney with the N1 probe were absent in the heart tissue with the N2 specific probe. This raises the possibility of tissue specific transcripts which may be driven from tissue specific promoters (Hughes and Brady, 2005a; Hughes and Brady, 2005b) or by tissue specific splicing mechanisms (Xu et al., 2002).

The human gene was recently cloned and called Aminopeptidase O (Diaz-Perales et al., 2005). The Northern data from Diaz-Pereles indicates that there are two transcripts 5kb and 7.5 kb in human tissues. The expression of these transcripts is seen in pancreas, liver, placenta, heart and testes. The Northern was hybridised with a short 377 bp probe that spans C-terminal region (nt 2101-2477). Northern analysis in this study, with a 980 bp C-terminal probe (nt 2260-3240) indicated that there are two major transcripts 2.1 kb and 2.3kb that are present in all tissues. There are also longer transcripts, ranging from 3.4 kb- 4.6 kb which are seen mainly in heart, lung, liver and kidney. There are differences between the two sets of Northern analysis, in particular the two major transcripts 2.1 and 2.3 reported in this study are not reported in human tissue (Diaz-Perales et al., 2005). This could indicate technical variation, for example, only a small portion of the blot is visible in the Diaz-Pereles figure and it is conceivable that the smaller RNA transcripts (2.1 and 2.3 kb) have not been shown. Alternatively, this may reflect variant species isoforms, for example the protease ADAM 28 in mice has reportedly three isoforms whilst the human has two. The isoforms have

different subcellular localisation and tissue expression (Howard et al., 2000; Roberts et al., 1999).

It is evident from both Northern blot analysis and RT-PCR data that *ApO* has alternatively spliced isoforms. Interestingly, there is alternative splicing of the functional domains which include the catalytic domain and NLS. The mechanism of alternative splicing is believed to occur in up to 60% of human genes (Modrek, 2002) and effectively generates more transcripts from a single gene leading to greater diversity and functional complexity (Graveley, 2001). Much of the evidence for alternative splicing has been derived from analysis of the EST databases. It is possible that sequences present within the databases are artefacts from splicing as several studies have been conducted to evaluate the authenticity of splicing variants by comparing the splicing conservation between mouse, rat and human genomes and concluded that a significant portion of alternative splicing in the EST databases is not functional and might result from aberrant splicing (Modrek, 2002; Sorek et al., 2004). Through alternative splicing there are 8 possible isoforms for the *ApO* gene (Figure 3.8). However, in reality splicing is extensively regulated (Joseph, 1995) and so it is unlikely that all of the variations depicted would be generated.

Alternative mechanisms such as the differential use of polyadenylation sites can also affect the pattern of splicing yielding polypeptides with different C-termini. For example, immunoglobin heavy chains (IGM) can either be membrane bound or secreted depending upon sequences at the C-terminus. There are two polyA signals in the IgM transcriptional unit. One of these is located within an intron and the other is at the end of the coding region. If transcription terminates at the intronic pA, the IgM is secreted whilst if transcription is terminated later to include the entire ORF which encodes a C-terminal hydrophobic membrane anchor, the protein is not secreted but sequesters in the membrane (Early et al., 1980a). Aminopeptidase O has two polyadenylation signals within the 3'UTR and although alternative transcripts with a preference for either one of the polyadenylation signals was not

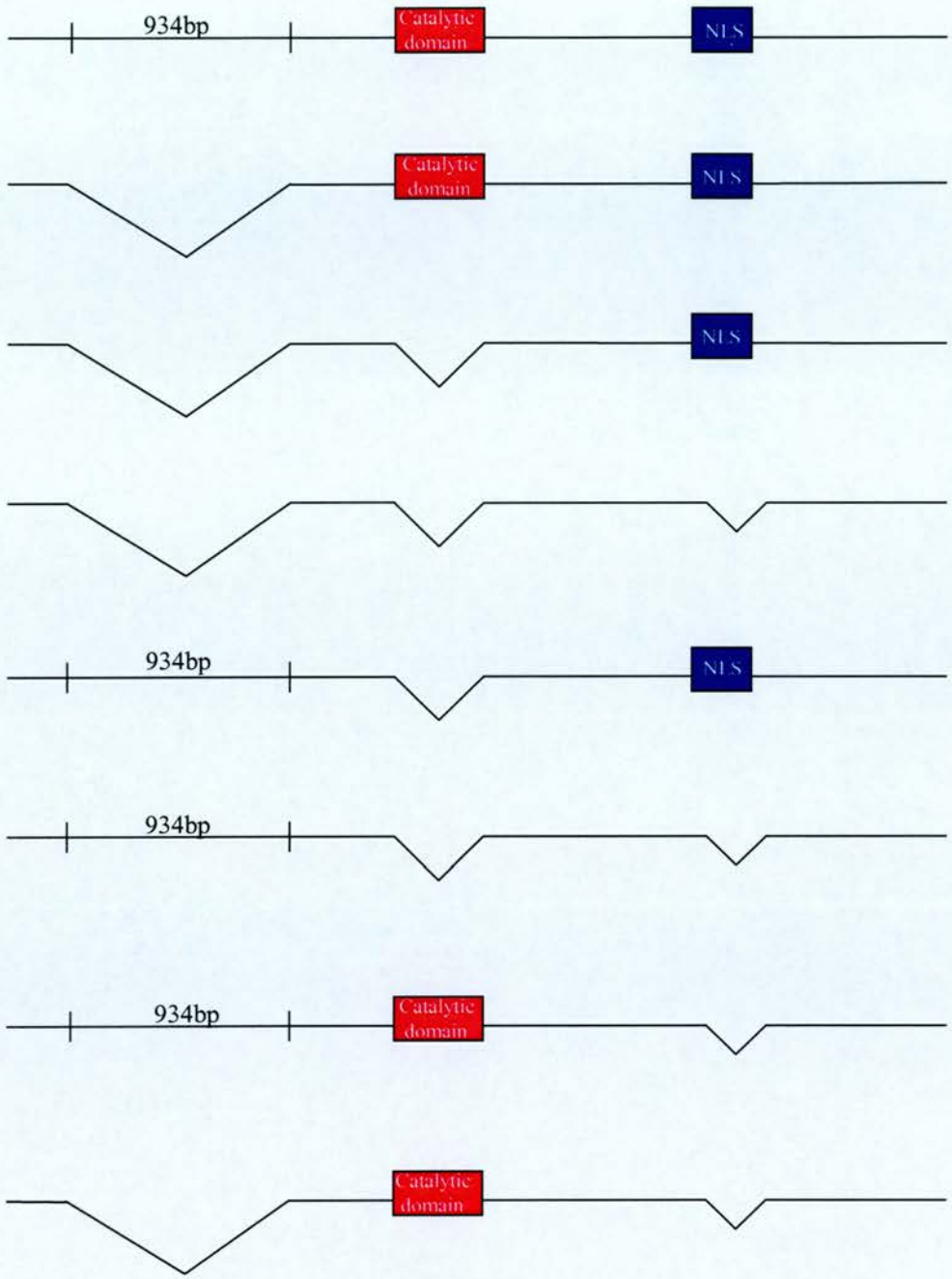


Figure 3.8: Possible alternative splicing of the *ApO* gene.
There are 8 possible isoforms that could be generated by alternative splicing of the 934 bp, catalytic domain and NLS exons.

observed it does remain a possibility and could be a mechanism of gene regulation. Translation efficiency of eukaryotic mRNAs depends upon structural features of the 5 untranslated regions (5'UTR) and the nucleotide sequences flanking the translation start codon (Kozak, 2002). At position 1 there is the putative start codon in the alternatively spliced 934 bp exon and it is predicted to be the initiation codon for the large transcripts. Shortly upstream from this start codon there is an in-frame stop codon which supports its credibility as the putative start codon for this transcript. With regard to the major predominant transcripts (2.1 and 2.3 kb) there are a number of possible start codons. The prediction that the AUG at position 332 is the authentic start codon was based upon a strong Kozak sequence and that this amino acid residue was conserved between species. Other methionine residues which could act as the initiation codon, for example, are Met 272, Met 291, Met 308, Met 324. These were excluded because they either had a poor Kozak consensus sequence or were not conserved in other species.

In most cases translation of eukaryotic mRNAs is initiated at the first available AUG of the 5'UTR, although exceptions to this rule have been described (Geballe and Morris, 1994). In this study it was shown that upstream AUG codons, although not uncommon, occur significantly less frequently than expected by pure chance. It has been hypothesised that there is positive selection against in-frame methionine residues upstream of the true start codon. These residues have been shown to affect the efficiency of translation (Geballe and Morris, 1994; Rogozin et al., 2001). It may be that *ApO* upstream AUGs work in a similar fashion.

DNA that contains secondary structure can be difficult to amplify by PCR, as it can obstruct the DNA polymerase during extension of the template. The possibility that secondary structure within the catalytic domain may prevent polymerase extension would offer a plausible explanation for the difficulties in obtaining full length transcripts with this domain (Shi and Jarvis, 2006). Although this experiment was not done, PCR additives (10% DMSO) can be used to amplify difficult templates (Benita et al., 2003).

Indeed, further evidence that secondary structure may be responsible for the difficulty in obtaining full-length transcript encoding the catalytic domain, comes from the original 5'RACE experiment. Using the primers P2 and P3 the initial 5' RACE experiment did not extend through the sequence that encodes the catalytic domain (Figure 3.1c and 3.1d).

With the advent of the entire human and mouse genomes sequence publicly available (http://www.ncbi.nlm.nih.gov) (Lander et al., 2001) and Celera (http://www.celera.com) (Venter et al., 2001), it is now possible to determine both the cDNA and genomic structure of almost any gene with increasing ease. One still has to confirm any bioinformatics prediction with bench work but the availability of the genome sequences simplifies this task greatly.

Chapter 4:

Expression analysis and phenotypic analysis of mice carrying the GT411 gene trap integration

4.1 Introduction

As a consequence of the gene trap vector integration it is predicted that the endogenous *ApO* promoter would regulate the expression of the βgeo gene. Using *LacZ* staining the endogenous gene expression was examined. In addition, the integration can obliterate normal gene function and possibly cause a mutagenic effect in germ line embryos. To this end a detailed analysis of both embryonic and adult expression by a comparative phenotypic analysis between wild type, heterozygote $ApO^{Gt411For/+}$ and homozygote $ApO^{Gt411For/Gt411For}$ mice tissues was undertaken.

4.2 Results

4.2.1 Animal breeding and phenotype analysis

The GT411 gene trap integration was transmitted through the germline by injection of the ES cells into C57/Bl6 blastocysts. Chimaeric males were backcrossed for three generations to wild type MF1 outbred females before intercrossing heterozygote $ApO^{Gt411For/+}$ animals. Adult animals and embryos were genotyped by staining ear punch biopsy or the yolk sac respectively for βgal activity. Homozygote $ApO^{Gt411For/Gt411For}$ and heterozygote $ApO^{Gt411For/+}$ were distinguished by the speed of appearance and the intensity of the staining reaction. The reliability was confirmed by breeding analysis.

4.2.2 Northern blot analysis of the gene-trap fusion transcript

To assess the transcriptional consequences of the gene trap integration, a Northern blot was performed with wild type, heterozygous $ApO^{Gt411For/+}$ and homozygous $ApO^{Gt411For/Gt411For}$ mRNA, that had been extracted from day 11.5 pc whole embryos and on 1 month old adult hearts (Figure 4.1b and c). The probe N1 spanned the gene trap integration site (Chapter 3, Figure 3.4a,). Two predominant transcripts (2.1 and 2.3kb) were present in the wild type and heterozygote $ApO^{Gt411For/+}$ but not in the homozygote $ApO^{Gt411For/Gt411For}$.

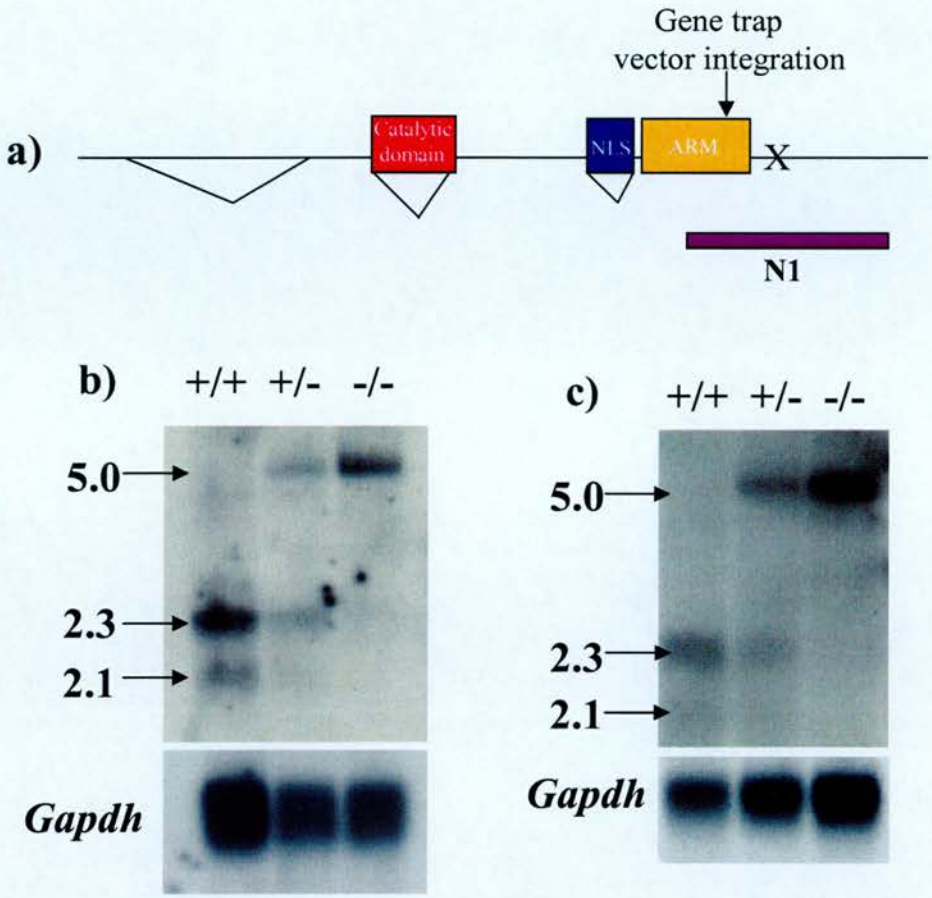


Figure 4.1: Northern blot of analysis of wild type, heterozygous *ApO* $^{Gt411For/+}$ and homozygous *ApO*$^{Gt411For/Gt411For}$ tissues.

a) Diagram of *ApO* cDNA showing gene trap integration site and position of probe (N1) used in Northern analysis (not drawn to scale).

b) Poly A+ RNA extracted from day 11.5 pc embryos with probe N1 which spans the gene trap integration site.

c) Poly A+ RNA extracted from 1 month old adult heart using probe N which spans the gene trap integration site.

G*apdh* was used as loading controls.

A 5kb transcript was observed in RNA isolated from heterozygote $ApO^{Gt411For/+}$ and homozygote $ApO^{Gt411For/Gt411For}$ tissues and was the expected size for the GT411-βgeo fusion transcript. Furthermore, the heterozygote band was approximately half the intensity of the homozygote which is expected because it would carry one copy whereas the homozygote would carry two copies. There was no evidence for splicing around the gene trap vector to produce a wild type transcript which could result in a hypomorphic allele (Voss et al., 1998). Thus from the data shown, the gene trap vector was functioning in the predicted manner. It was noted that the longer transcripts seen previously by Northern (Figure 3.4a and b, chapter 3) were not observed on these particular blots. Possible reasons for their absence could include, for example, that less RNA has been loaded on to the gel or that the transfer was less efficient.

4.2.3 Phenotypic analysis

Breeding of the gene trapped transgenic mice resulted in the generation of apparently healthy homozygous $ApO^{Gt411For/Gt411For}$ offspring at the normal Mendelian ratio. Homozygous males mated with homozygous females also resulted in apparently normal healthy offspring. This would indicate that the gene trap integration has not resulted in a gross mutant phenotype perhaps because the gene is still functional or that there is genetic redundancy with another gene able to compensate for it (Prince and Pickett, 2002). Another possibility is that there is a subtle phenotype which would be detectable with a specific assay or observation, for example the matrix metallo protease MMP2 knockout mice develop normally except for a subtle delay in growth (Itoh et al., 1997).

To assess the possibility of a cardiovascular phenotype heart and kidney were dissected and the body weight, heart and kidney weight were compared between wild type and heterozygotes $ApO^{Gt411For/+}$ and homozygotes gene trap $ApO^{Gt411For/Gt411For}$. T-test calculations indicated that there were no

significant statistical differences between the wild type and gene trap mice (Figure 4.2).

4.2.4 Embryonic expression and phenotypic analysis

Embryos were collected at day 8.5, 9.5 and 10.5 pc. They were stained for β galactosidase (βgal) activity and a comparison of the vasculature between heterozygotes $ApO^{Gt411For/+}$ and homozygotes gene trap $ApO^{Gt411For/Gt411For}$ embryos was undertaken.

Day 8.5 pc embryos have βgal expression in the dorsal aortae. There were no apparent morphological differences between them (Figure 4.3) The only notable dissimilarity between the embryos was the less intense blue staining seen in heterozygote $ApO^{Gt411For/+}$, which was attributed to there being only one copy of the βgal gene. In the experimental design, heterozygote $ApO^{Gt411For/+}$ embryos were left staining for a longer period to try to stain them equivalently. This was true also for the day 9.5 and 10.5 pc embryos (see below).

A preliminary analysis of day 9.5 pc did not indicate any substantive differences between heterozygote $ApO^{Gt411For/+}$ and homozygous $ApO^{Gt411For/Gt411For}$ gene trap embryos. The dorsal aorta appeared to develop normally and intersomitic vessels were present between somites in the normal segmented pattern. There was a highly arborised capillary network formed from dorsal sprouts from the intersomitic vessels (Figure 4.4). There was also notable staining in the limb bud, tail bud and forebrain.

βgal expression in day 10.5pc embryos had extensive staining in the capillary plexus of the head and throughout the body and endothelial staining in the intersomitic vessels. A comparison of heterozygotes $ApO^{Gt411For/+}$ and homozygote $ApO^{Gt411For/Gt411For}$ gene trap embryos did not show significant differences (data not shown).

A comparison of a βgal and riboprobe stained day 9.5 pc embryos demonstrated staining of the intersomitic vessels, tail bud and forebrain (Figure 4.5). This would indicate the *lacZ* stain does truly mimic the

endogenous expression of *ApO*. There was also strong staining in the tail bud and forebrain which was also seen by *lacZ* staining (*in situ* performed by Dr J.Wallis).

	Wild type (*n3*)	Heterozygote$^{Gt411For/+}$ (*n3*)	Homozygote $^{Gt411For/Gt411For}$ (*n3*)	P*
Body weight, g	24.3 +/- 0.8	23.7 +/- 0.81	24.2 +/-0.7	NS
Heart weight, mg	214+/-0.02	220+/- 0.02	219+/-0.004	NS
Kidney weight, mg	536+/-0.013	552+/- 0.017	550+/-0.039	NS

*P>0.05 was considered not significant

Figure 4.2 Phenotypic analysis by comparative mean values and T-test calculations of tissues.

Total body weight, heart and kidney weight of 1 month old wild type, heterozygous and homozygous mice.

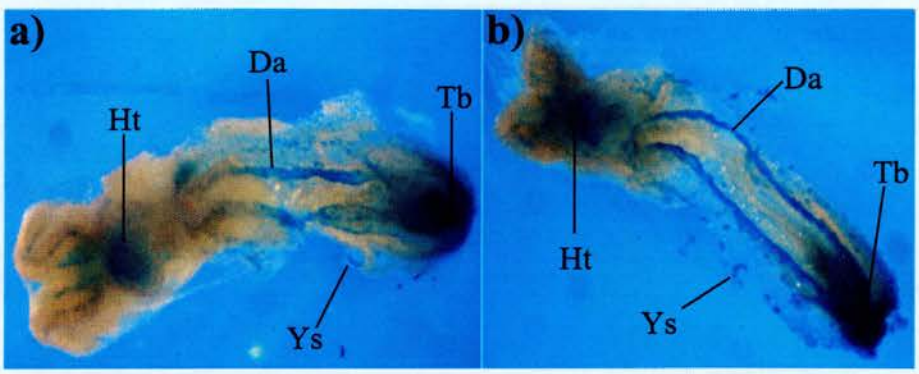


Figure 4.3: Comparative expression of βgalactosidase staining of day 8.5 pc embryos.

ApO $^{G\,t411For/+}$ (a) *ApO* $^{G\,t411For/\,G\,t411For}$ (b)

Da=dorsal aorta, Ht=heart, Tb=tail bud, Ys= yolk sac

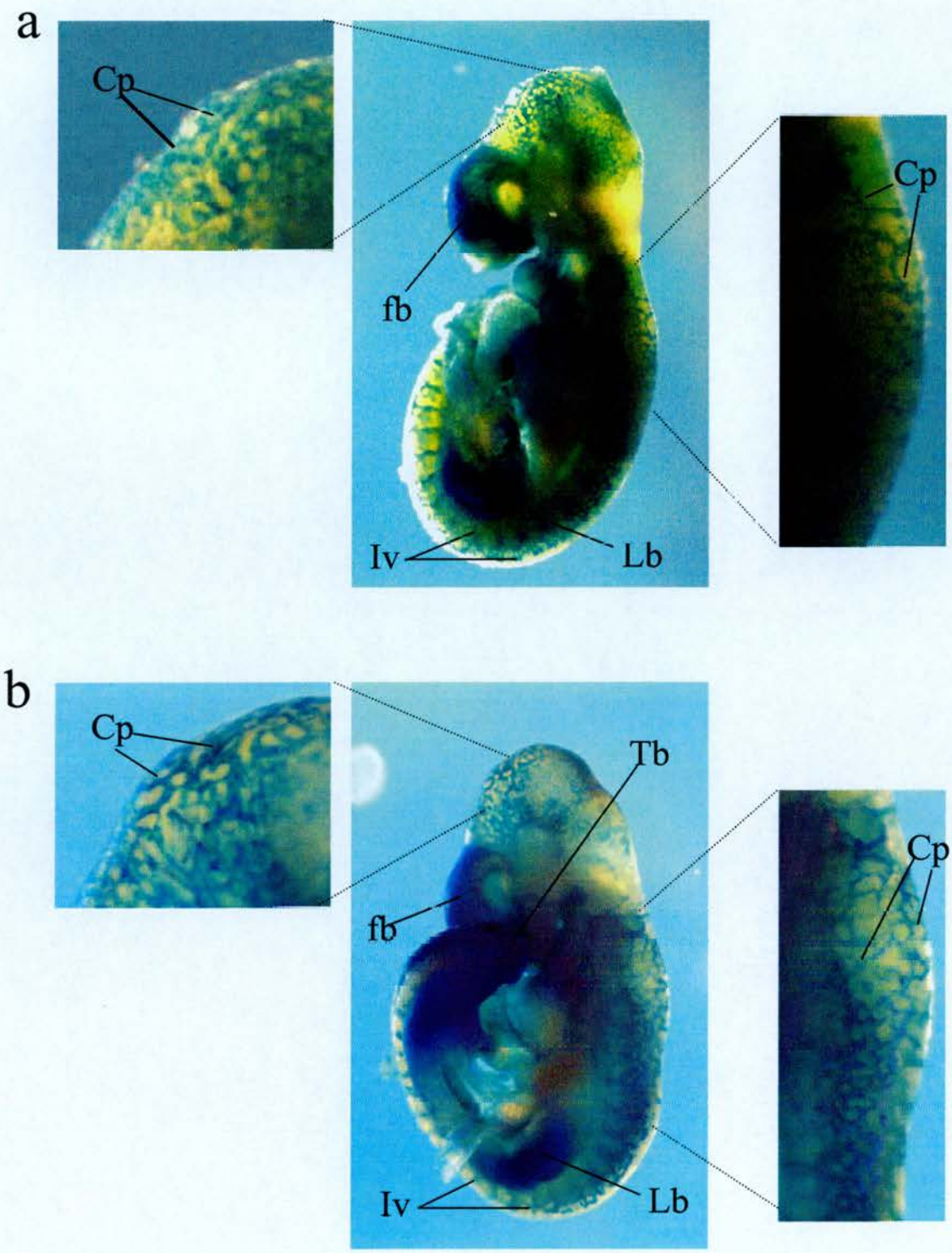


Figure 4.4: Comparative expression of day 9.5 pc gene trap mice by βgal expression.

ApO$^{G\ t411For/+}$ (heterozygote) (a), *ApO* $^{G\ t411For/\ G\ t411For}$ (homozygote)(b). Lb=limb bud, Tb=tail bud, fb=fore brain, Cp=capillary plexus, Iv=intersomitic vessel

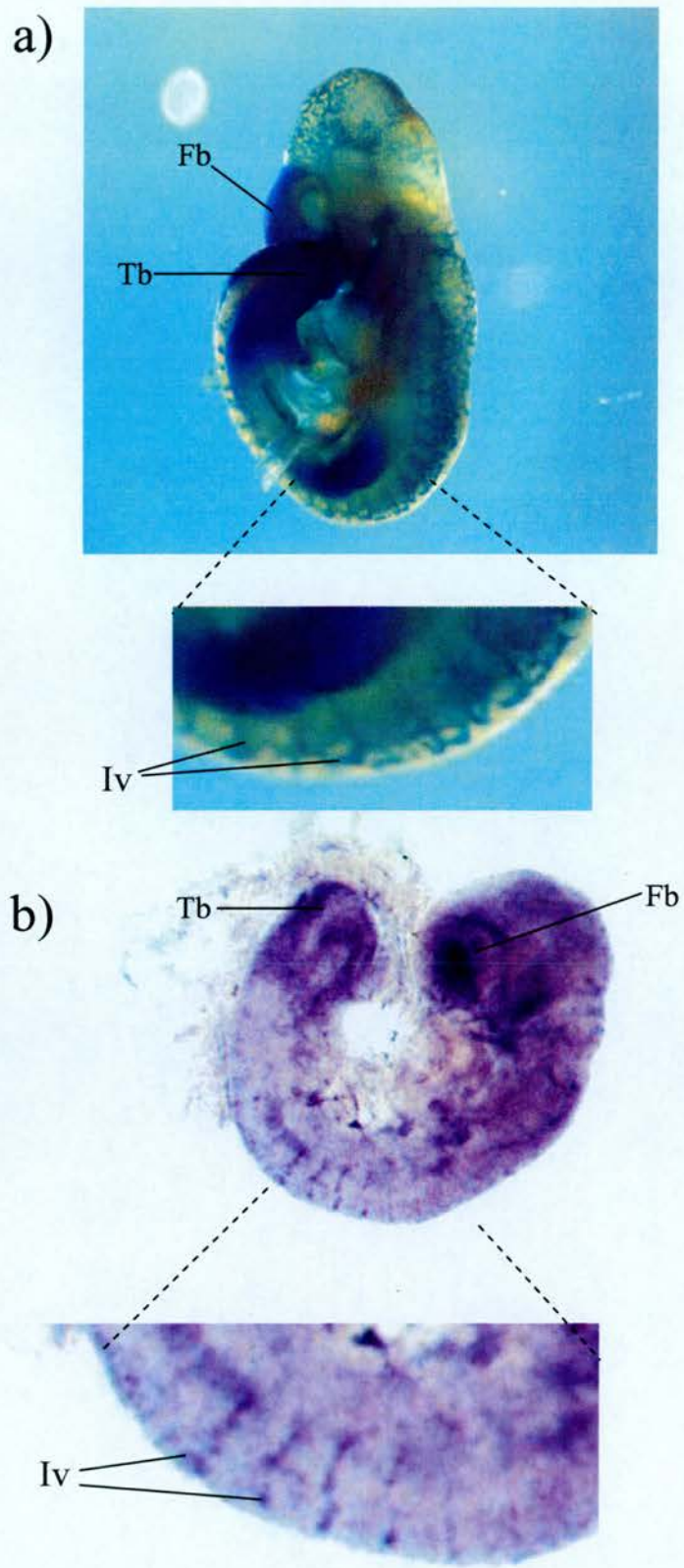


Figure 4.5: Comparison of *ApO*-GT411 fusion and endogenous *ApO* expression.

9.5.dpc βgal stained *ApO* $^{Gt411For/\ Gt411For}$ embryo (a), 9.5dpc *in situ* (ribo probe N1) (b). Iv= intersomitic vessels, Tb=tail bud, Fb=forebrain

4.2.5 Aortic ring assay

Although there was no obvious defect in the development of blood vessels in the embryo, it was felt to be worth assessing whether there was a defect in the growth of adult endothelial cells using the aortic ring assay. Growth of aortic sprouts within this three dimensional matrix were compared with aorta from wild type, heterozygous and homozygous GT411 mice. The aortic were monitored daily from day 0 to day 12.

Angiogenic sprouts first appeared on day 4 and on subsequent days microvessels were seen to elongate and undergo significant branching. Comparisons by manual processing and observation did not indicate significant differences between the wild type, heterozygote and homozygote gene trap mice. It was concluded that there were not any gross morphological differences or changes in the rates of new vessel growth or branching observed between the aortic explants (mice number n =3 for each, Figure 4.6 a, b, and c).

To confirm that these sprouting structures were new vessels, a fluorogenic labelled acetylated low density lipoprotein (Cambio), which specifically stains endothelial cells was used (Eter and Spitznas, 2002), (data not shown). Moreover, *Ap0* expression in these vessels was confirmed by both *LacZ* staining (4.6 d and e) and RT-PCR (4.7 b) with RNA purified from the aortic sprouts. The primer pairs EcoRI-P36 and EcoRI-P28 demonstrate that *ApO* transcript with the alternatively spliced 933 bp exon is expressed in the developing aortic sprouts of adult mice. The primer pair PEcoRI-P21 indicates that a transcript containing the putative catalytic domain was expressed in the developing aortic sprouts. Primer pair PEcoRI-P15 generated a single PCR band and sequencing confirmed it was the isoform minus the catalytic domain. This data concurs with the analysis of clones in section 3.2.3.

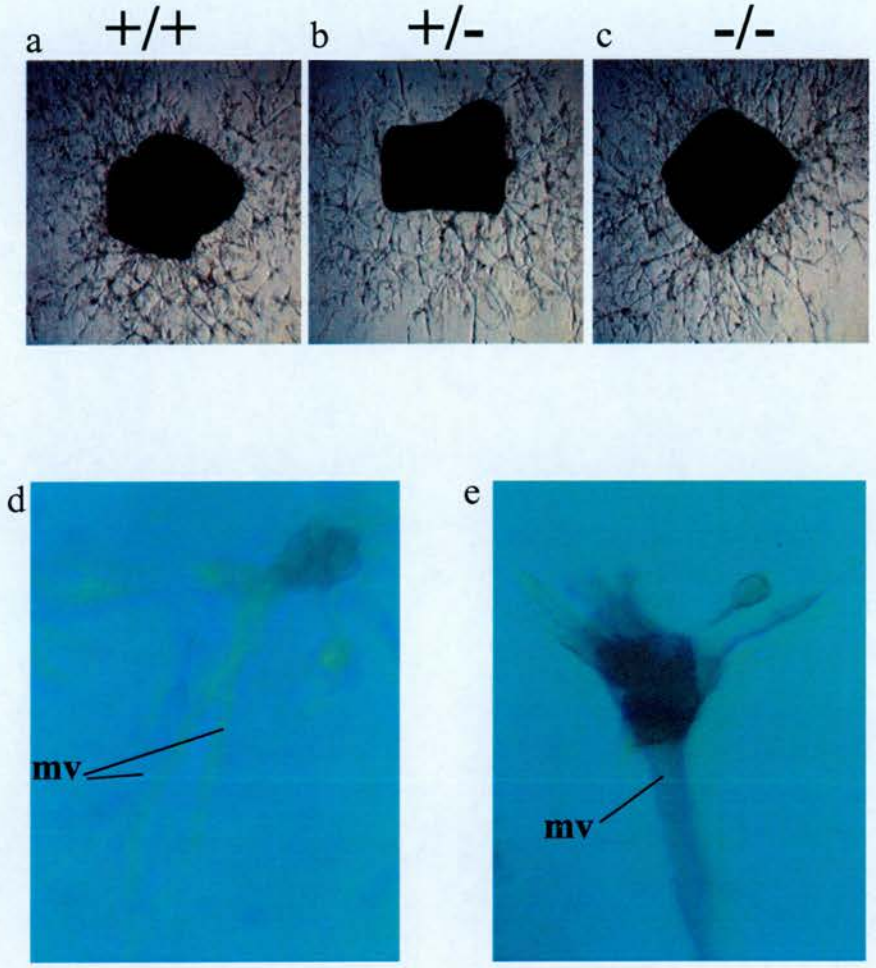
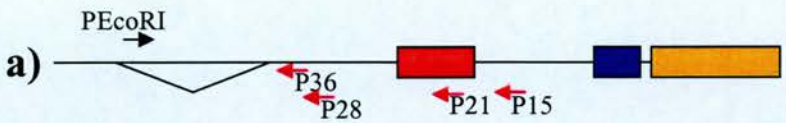
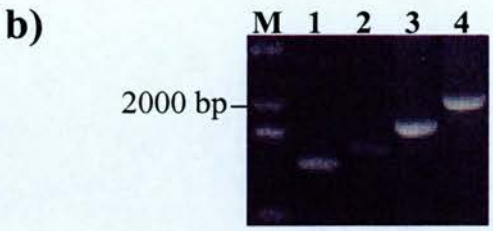


Figure 4.6: Aortic ring assay using aortae isolated from adult mice.

wild type (a), *ApO* $^{Gt411For/+}$ (b) and *ApO*$^{Gt411For/GtFor}$ (c).

βgal stained microvessels from aortic ring assay.
Wild type (d), Heterozygote *ApO* $^{Gt411For/+}$ (e).
mv= microvessel





Lane	Primers	Size bp
1	PEcoRI - P36	1230
2	PEcoRI - P28	1370
3	PEcoRI- P21	1580
4	PEcoRI - P15	1920

Figure 4.7: RT-PCR of RNA from aortic ring assay

Schematic representation of *ApO* cDNA (not to scale) with PCR primers (a).

RT-PCR on RNA purified aortic sprouts and primer table indicating sizes of fragments (b).

4.2.6 Adult expression

Cryostat sections of adult tissues were stained for βgal reporter gene activity (Figure 4.8). In heterozygous $ApO^{Gt411For/+}$ and homozygous $ApO^{Gt411For/Gt411For}$ mice, the βgal expression was seen in the vessels of heart, spleen and kidney skeletal muscle and lung. The wild-type tissues did not express the βgal reporter gene.

A comparative analysis of the tissues with the histology web site (http://www.siumed.edu/~dking2/crr/rsguide.htm) was carried out. In the heart there was expression of βgal in the vasculature and cardiomyocytes. It is possible to see the striated intercalating discs of the muscle fibre (Figure 4.8).

In the kidney βgalactosidase staining was observed in the glomeruli, interlobular arteries and afferent arterioles but not in the peritubular capillaries which would pervade the entire tissue.

The spleen has βgalactosidase staining in large irregular thin walled blood vessels called the splenic sinusoids and in other vessels too.

βgalactosidase staining of lung results in discrete irregular shaped pockets that appear to form a lumen. These are reminiscent of veins and arteries. The irregular shape may be an artefact of the crysectioning caused by destruction of the morphology during the preparation.

To verify that the βgal was labelling vessels, an *in situ* with an anti Flk-1 antibody was used on adult spleen. The Flk-1 was very similar to the βgal staining indicating that these are vessels (Figure 4.9 a and b). To confirm that the βgal staining was truly in the vessels a double labelling with anti bodies to Flk1 and βgal should be performed.

Endothelial and haematopoietic cells have a common progenitor termed the haemangioblast (Huber et al., 2004) and recent evidence also suggests that smooth muscle cells may differentiate from this common progenitor too (Ema et al., 2003). Haematopoietic tissue was tested from gene trap mice to see whether aminopeptidase O was expressed. Cytospins were made from adult bone marrow and stained for βgalactosidase activity. The finding was

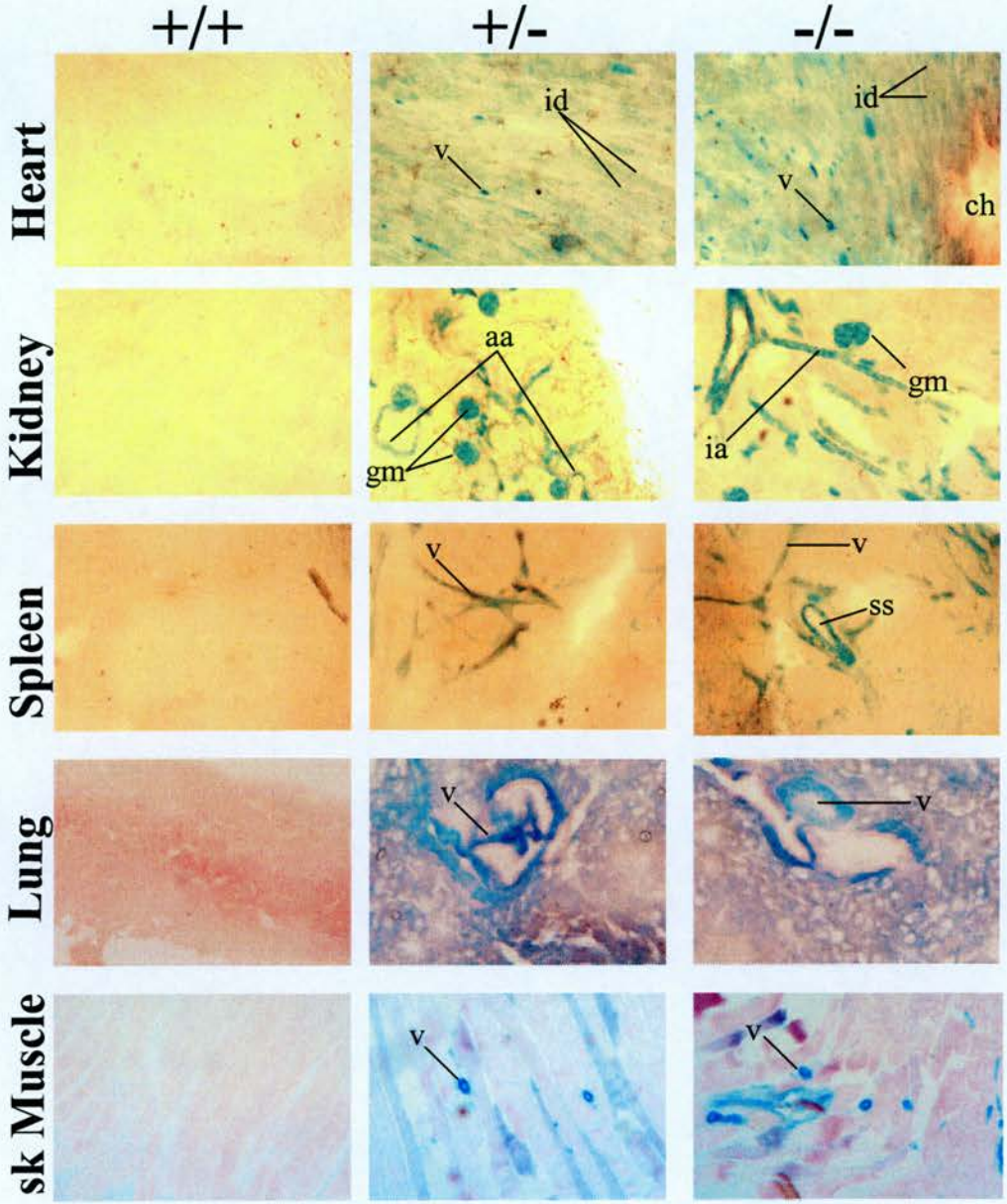


Figure 4.8: Expression of the *ApO*-GT411 gene trap from 1 month old adult tissues.

βgal stained cryosections of adult tissues from wild type, *ApO* $^{Gt411For/+}$ and *ApO* $^{Gt411For/Gt411For}$ gene trapped mice.
ch = chamber, v = vessel, gm = glomeruli, aa = afferent arteries, Ia = interlobular arteries, id = intercalated discs, ss = splenic sinusoid

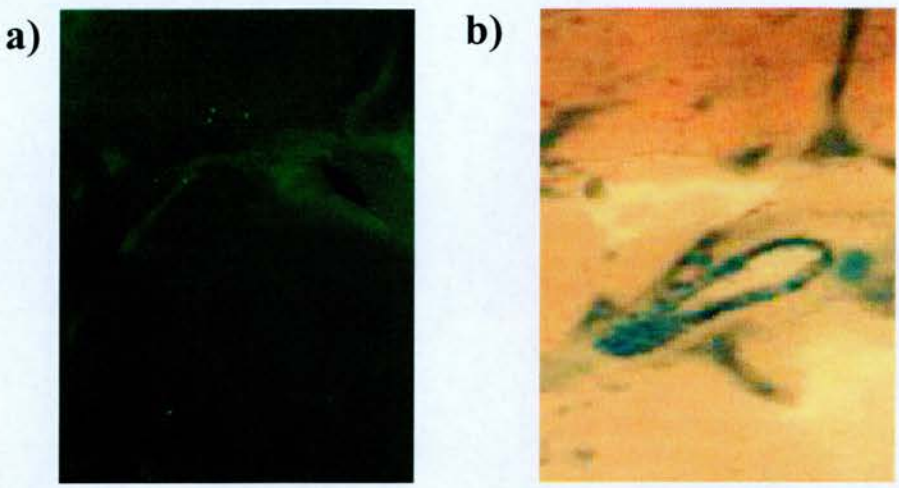


Figure 4.9: Comparative expression analysis of Flk 1 Immunohistochemistry and βgal stained spleen

a) Cryosection of adult spleen FITC labelled vessel with anti Flk-1 antibody

b) βgal stained heterozygous $^{Gt411For/+}$ cryostat section adult spleen

that the haematopoietic tissues did not have βgalactosidase activity demonstrating that *ApO* was not expressed in this tissue (data not shown).

4.3 Discussion

The *ApO* gene is expressed in ES cells, and during embryogenesis (8.5, 9.5 and 10.5 dpc embryos) and in the vasculature of a range of adult tissues. Interestingly, in heart a diffuse *lacZ* staining was observed and this was presumed to indicate cardiomyocyte staining. The prospect that the vascular expression of the aminopeptidase O gene is restricted to a subset or specific type of endothelial cell remains a possibility, as it appears to be expressed in the larger vessels but not in the capillaries. For example the ephrin receptor tyrosine kinases are expressed on different types of vessel with ephrin-B2 selectively marking arterial endothelium whilst ephrin-B4 reciprocally marks venous endothelium (Gerety et al., 1999). To determine if *ApO* is expressed in a particular endothelial cell double labelling experiments could be performed with an APO specific antibody and another antibody that marks a specific cell type.

The riboprobe whole mount *in situ* appears to recapitulate the reporter gene expression demonstrating that the β-gal activity authentically marks the expression of the endogenous gene in the developing vasculature.

Embryogenesis proceeds normally, resulting in apparently normal healthy homozygous $ApO^{Gt411For/Gt411For}$ offspring at the normal Mendelian ratio. Analysis of embryos and adult tissues did not indicate a phenotype in the morphological structure of the vasculature when comparing wild type and gene trap mice.

Furthermore, using the aortic ring assay, the aorta derived from wild type, heterozygous $ApO^{Gt411For/+}$ and homozygous $ApO^{Gt411For/Gt411For}$ gene trap animals were indistinguishable. A manual comparison of the microvessels, however, is both tedious and subjective. Computer image analysis software has since been developed which may greatly enhance the applicability of this

method and allow more accurate and detailed analysis to be undertaken in future (Blacher et al., 2001).

The integration is near to the 3' end of the gene, removing the last 46 amino acids. It is possible that the truncated fusion protein retains its normal function despite being fused to the βgal reporter gene and removing an ARM repeat. Another possibility is that there is genetic redundancy where another gene is able to alleviate the effects of a mutated gene because it has overlapping function (Nowak et al., 1997; Shastry, 1995). At the protein sequence level aminopeptidase O shows the highest sequence similarity to aminopeptidase B. To test if there is genetic redundancy between these two proteins one could generate a homozygous *ApO* and *ApB* double knockout. The assessment of a mouse phenotype might demonstrate overlapping functions of these two proteins. This approach has been taken with the Hox genes. For example *HoxB3* and *HoxB4* double knockout has a more severe phenotype than the single gene knockout for *HoxB4* suggesting that these two proteins have similar function in haematopoiesis (Bjornsson et al., 2003; Brun et al., 2004).

The Northern blots on heterozygous $ApO^{Gt411For/+}$ and homozygous $ApO^{Gt411For/Gt411For}$ gene trap embryos and adult heart tissue with the N1 probe demonstrate that the gene trap vector is not spliced out to generate wild type transcript resulting in hypomorphic allele. The presence of wild type transcript, even at low levels can ameliorate a possible phenotype (Voss et al., 1998). It is possible, however, that there are alternative transcripts which would not use the C-terminal region of the gene and would therefore not be trapped by the gene trap integration. The use of additional probes on Northern blots and more extensive 3' RACE would help to address this possibility.

The discovery that *ApO* has a specific vascular expression is interesting because of the relationship between angiogenesis and diseases such as cancer. If a tumour could be prevented from acquiring a blood supply by the action of a drug that specifically inhibits processes such as cell proliferation and angiogenesis this could be offered as a therapy in the treatment of cancer. The

development of molecules that specifically inhibit APO would be of interest as they could then be used to assess the effect on *in vitro* endothelial cell proliferation assays and on *in vivo* models of cancer. For example, methionine aminopeptidase 2 which is specifically inhibited by TNP-470 has been demonstrated to reduce endothelial cell proliferation *in vitro* (Antoine et al., 1994; Satchi-Fainaro et al., 2004). Moreover clinical trials with TNP-470 have been shown to be effective in the treatment of cervical cancer (Kudelka et al., 1998).

Chapter 5:

Cellular localisation studies of the mouse aminopeptidase O protein

5.1 Introduction

Proteins contain sequences that determine their ultimate destination, for example signal peptides of metalloproteases enable proteins to traverse membranes to be secreted from a cell (Das et al., 2003). In other circumstances a signal peptide may direct the protein across a specific membrane of an organelle as is the case for mitochondria (Rapaport, 2003). Of particular interest in the context of this thesis are the nuclear localisation signals (NLS) because a manual comparison of the murine APO protein sequence with known NLSs (Williams, 2003) indicate the presence of a putative signal sequence between residues 694-670 (Figure 3.5). Proteins which translocate to and from the nucleus do this by traversing the nuclear envelope through a proteinacous complex called the nuclear pore complex (NPC). This structure comprises over 100 different proteins (in higher eukaryotes) termed nucleoporins (NUPS) (Wente, 2000). The transport signals show considerable variation in complexity from short peptide motifs like the classical monopartite basic NLS (Kalderon et al., 1984) and the leucine-rich nuclear export sequence (Wen et al., 1995) up to large protein domains and even multicomponent signals such as those involved in U-snRNA transport (Bogerd et al., 1999).

The nucleus is a membrane bound structure which contains a number of domains or sub-compartments which include nuclear cajal bodies, nuclear speckles, transcription and replication foci, chromosomal territories and the non membrane bound proteinacous complex nucleolus (Handwerger and Gall, 2006). The latter are the only visible structure when viewed under light microscopy because of the highly refractile nature due to their denseness. Initially it was established as the ribosome factory where rRNA transcripts are processed and assembled with ribosomal proteins to form ribosomes (Moss and Stefanovsky, 2002). Recently, however, other roles have been defined which include cell cycle regulation, DNA damage repair, pre-mRNA processing (Desterro et al., 2003; Olson et al., 2002; Pederson, 1998; Tsai and McKay, 2002) and as a 'molecular sink' to retain proteins and prevent them

from interacting with downstream partners until a particular stage of cell cycle or metabolic state (Leung et al., 2003).

To investigate whether the NLS predicted from the sequence analysis of APO was functional, several approaches were taken. In the first instance, an assessment of the putative NLS protein conservation between different species was carried out. eGFP-*ApO* constructs with and without the putative NLS sequence were then cloned and the subcellular localisation assessed using the eGFP tag in transiently transfected COS 7 cells.

5.2 Results

5.2.1 Cellular localisation of APO protein

ßgal staining of the GT411 ES cell line indicated that the APO-βGAL fusion protein localised in distinct punctate regions within the cell (Figure 5.1a). To further characterise the subcellular localisation, an N-terminal eGFP-tagged full length *ApO* construct (nucleotide 278-2748, minus the catalytic domain, Appendix 1.4) was transfected into COS 7 cells. The cells were viewed with fluorescence microscopy and the eGFP-APO protein also localised to punctate regions within the nucleus which were believed to be the nucleoli (Figure 5.1b).

Owing to the difference in density between the nucleolus and the surrounding nucleoplasm, the nucleoli were readily visible under bright field (white light) phase microscopy (Figure 5.2 a) and it was these structures, when viewed with fluorescent microscopy, that fluoresced green with eGFP tagged APO fusion protein (Figure 5.2 c). Fibrillarin is a protein which primarily localises to the nucleoli (Newton et al., 2003) where it is a structural component of a small nuclear ribonucleoprotein (SnRNP) particle involved in the post transcriptional modification by methylation of rRNA (Decatur and Fournier, 2003) and it was as a positive control to mark the nucleolus. It was cloned by RT-PCR into the eGFP expression vector using the primers Pfib1-Pfib2 (Table 1, Appendix 1.2). COS 7 cells were transfected with the eGFP-

fibrillarin construct. Under white light the nucleoli were visible (Figure 5.2 b) and the nucleoli were green under fluorescent microscopy (Figure 5.2 d).

To confirm that the eGFP signal was indeed within the nucleus a virtual section of the cell was compiled. Sequential images of transfected cells were taken on a microscope with a motorised camera to generate a cross section of the cell (Figure 5.3 a). The images were computationally manipulated and represented as a side profile. In this way a three dimensional image was generated. The eGFP signal was confirmed using this technique as being within the nucleus (Figure 5.3 b, c and d).

5.2.2 Fractionation of COS 7 cells

Cytoplasmic and nuclear fractions were separated with a spin kit (Pierce) from COS 7 cells that had been transfected with the eGFP-APO fusions (*ApO* 1-stop (Minus catalytic domain) and *ApO* 552-stop). The eGFP and the eGFP-APO fusion proteins are present in both the cytoplasmic and nuclear fractions. It would appear that eGFP-APO is enriched in the nuclear fraction as determined by Western blotting (Figure 5.4).

5.2.3 How does the APO protein traffic to the nucleoli?

Manual comparison of the APO protein sequence with known nuclear localisation signals (NLS) indicated the presence of a putative NLS between residues 694-670, but interestingly standard bioinformatics programmes did not identify this domain. To test whether this putative NLS motif could guide the murine APO protein to the nucleolus, various eGFP constructs with and outwith this domain were generated and transfected into COS 7 cells (Figure 5.5).

Those cells that were transfected with eGFP-*ApO* vectors containing the putative NLS were detected in the nucleolus (Figure 5.5 a, b, e and f), whereas cells transfected with eGFP-*ApO* vectors without the NLS remained in the cytoplasm (Figure 5.5 c and d). A similar sub-cellular localization

pattern was observed with fibrillarin control when compared with the eGFP-*ApO* NLS constructs (Figure 5.5 g). These data strongly support the hypothesis that the putative APO-NLS does function as a signal to direct proteins to the nucleus and more specifically to the nucleoli.

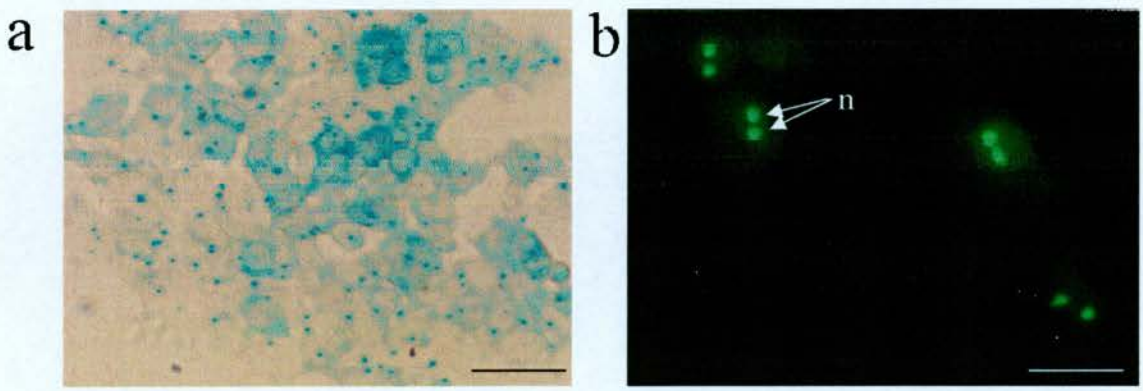


Figure 5.1: Sub-cellular localisation of APO fusion proteins

a) β-gal staining of gene trapped ES $^{Gt411For/+}$ cells.

b) COS 7 cells transfected with full-length *ApO* fused to eGFP

Arrows indicate nucleoli.= n Scale bar 10μM

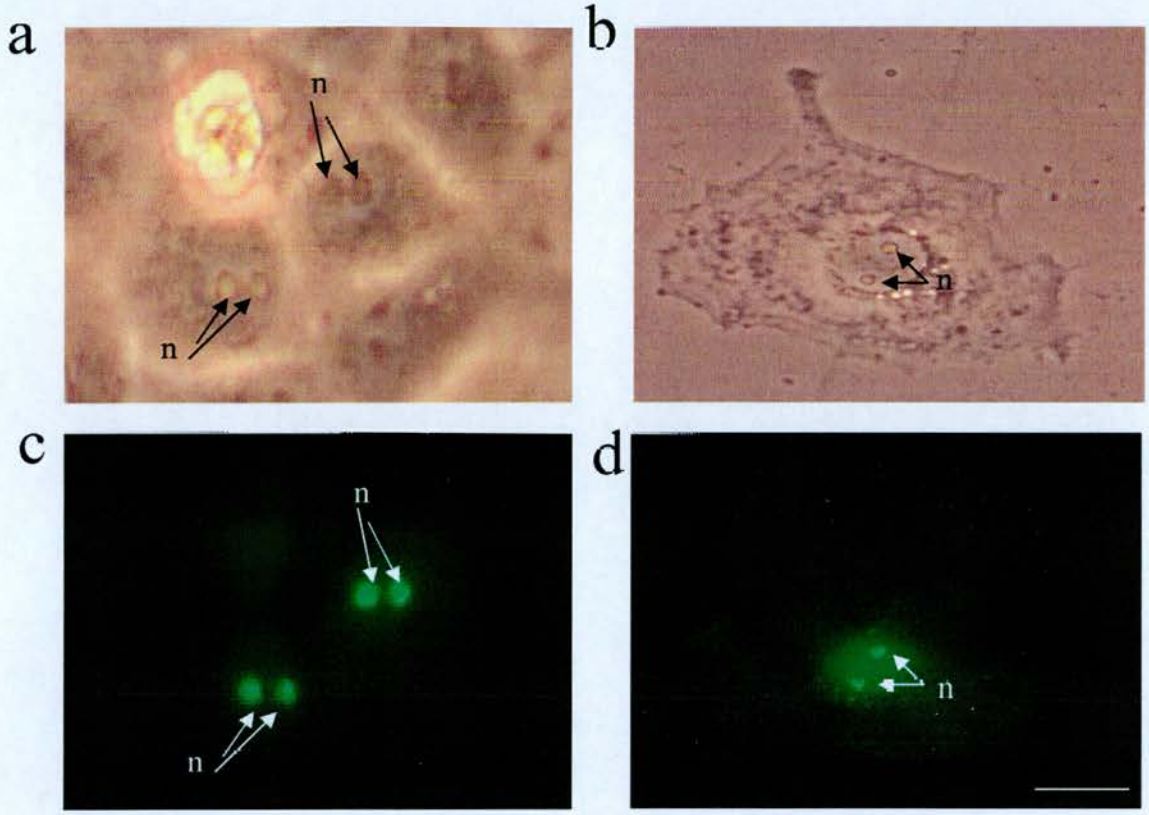


Figure 5.2: Sub-cellular localisation of APO fusion proteins and fibrillarin

Transfected COS 7 cells with eGFP-Fl-APO Bright field (a) and fluorescent microscopy (c).

COS 7 cells transfected with nucleoli specific protein fibrillarin bright field (b) and fluorescence (d).

Arrows indicate nucleoli.= n
Scale bar 10μM

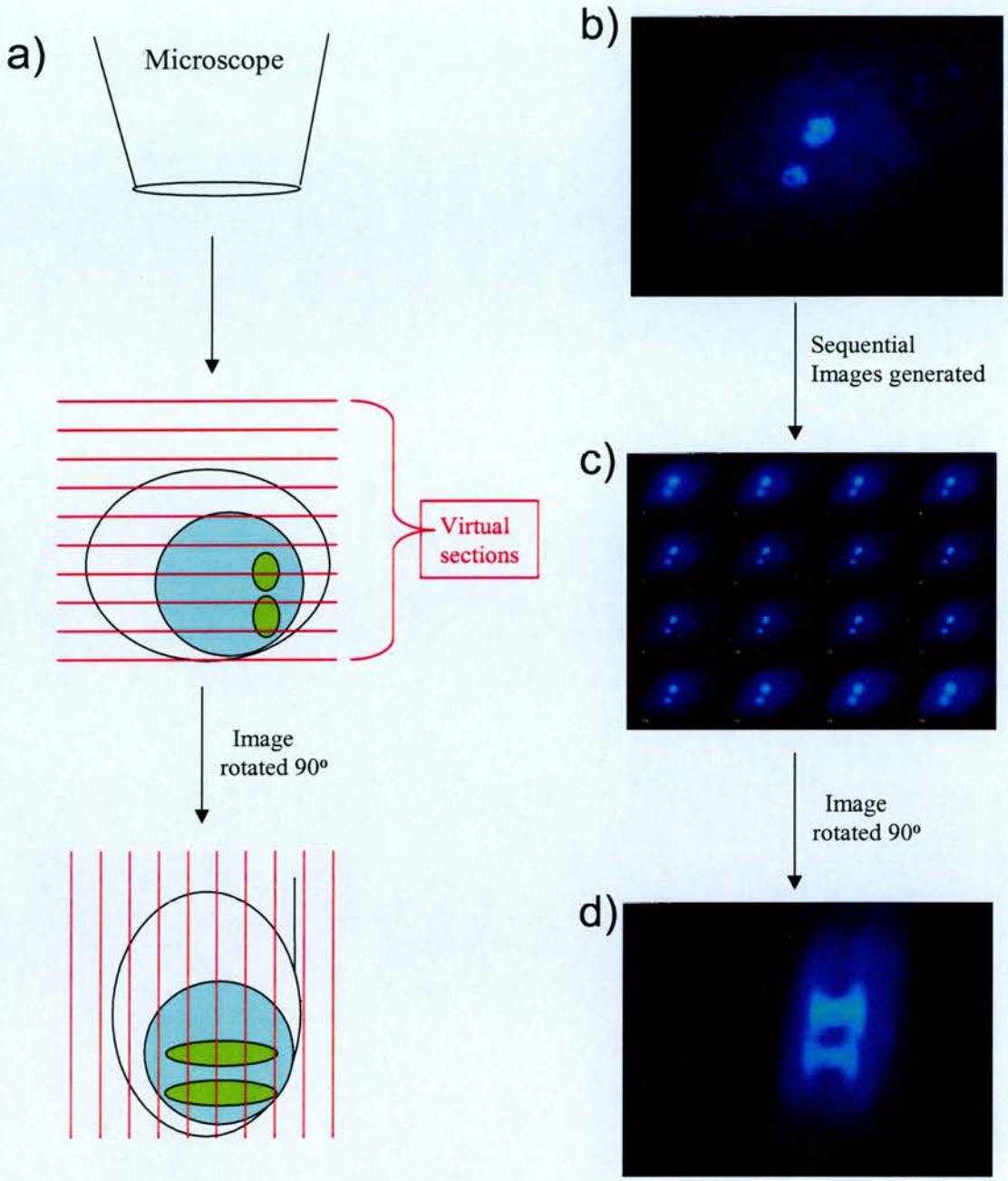


Figure 5.3: Three dimensional sectioning of eGFP-*ApO* transfected COS 7 cells.
Schematic illustration of virtual sectioning of cells (a)
Image of COS 7 cells transfected with eGFP-Fl-*ApO (1-stop)* and DAPI-stained (b).
Virtual sectioning compiled by serial photography (c)
Computer generated side profile of eGFP-Fl-*ApO (1-stop)* (d)

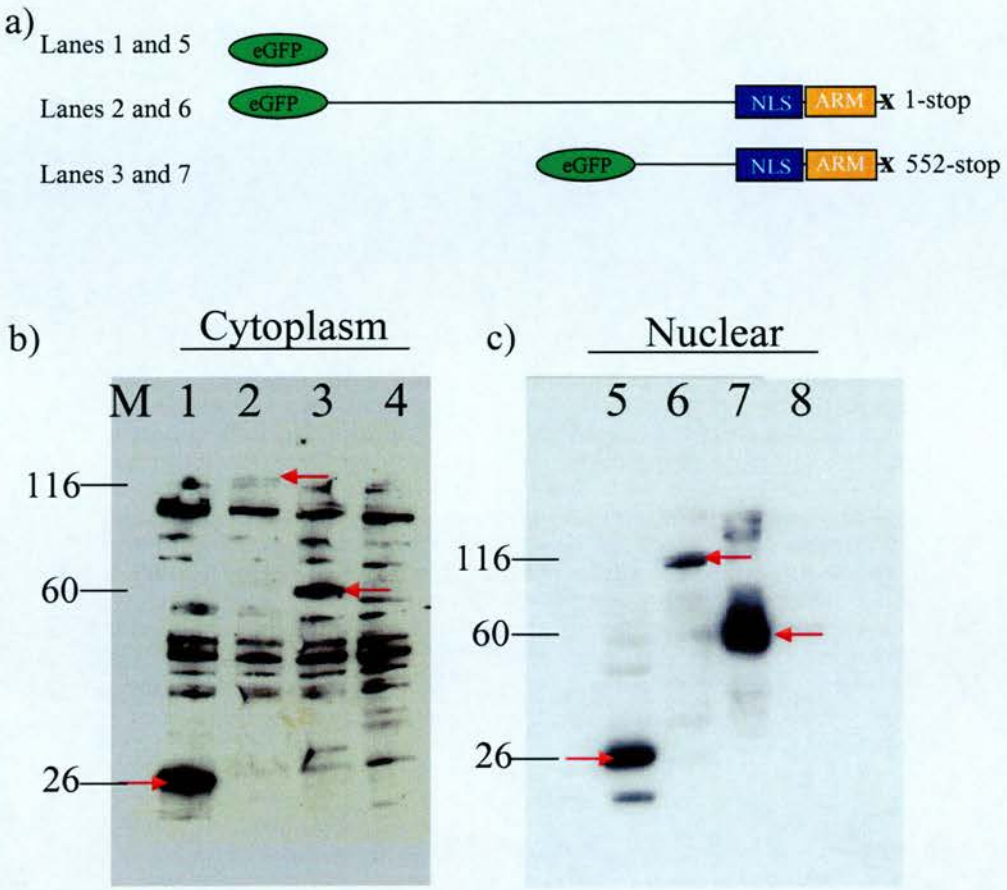


Figure 5.4: Analysis of fractionated protein lysates from transiently transfected COS 7 cells by Western blot.

a). Schematic illustration of eGFP clones transfected into COS 7 cells.

b). Cytoplasmic fractions analysed by Western blot with eGFP antibody of COS 7 cells transfected with eGFP tagged fusions. Lane 1 eGFP vector only, lane 2 eGFP-Fl-*ApO*, lane 3 eGFP-*ApO* (552-stop) and lane 4 mock transfection.

c). Nuclear fractions analysed by Western blot with eGFP antibody of COS 7 cells transfected with eGFP tagged fusions. Lane5 eGFP vector only, lane 6 eGFP-Fl-*ApO*, lane 7 eGFP-*ApO* (552-stop) and lane 8 mock transfection.
(72hr transfection). Red arrows= eGFP and eGFP fusions.

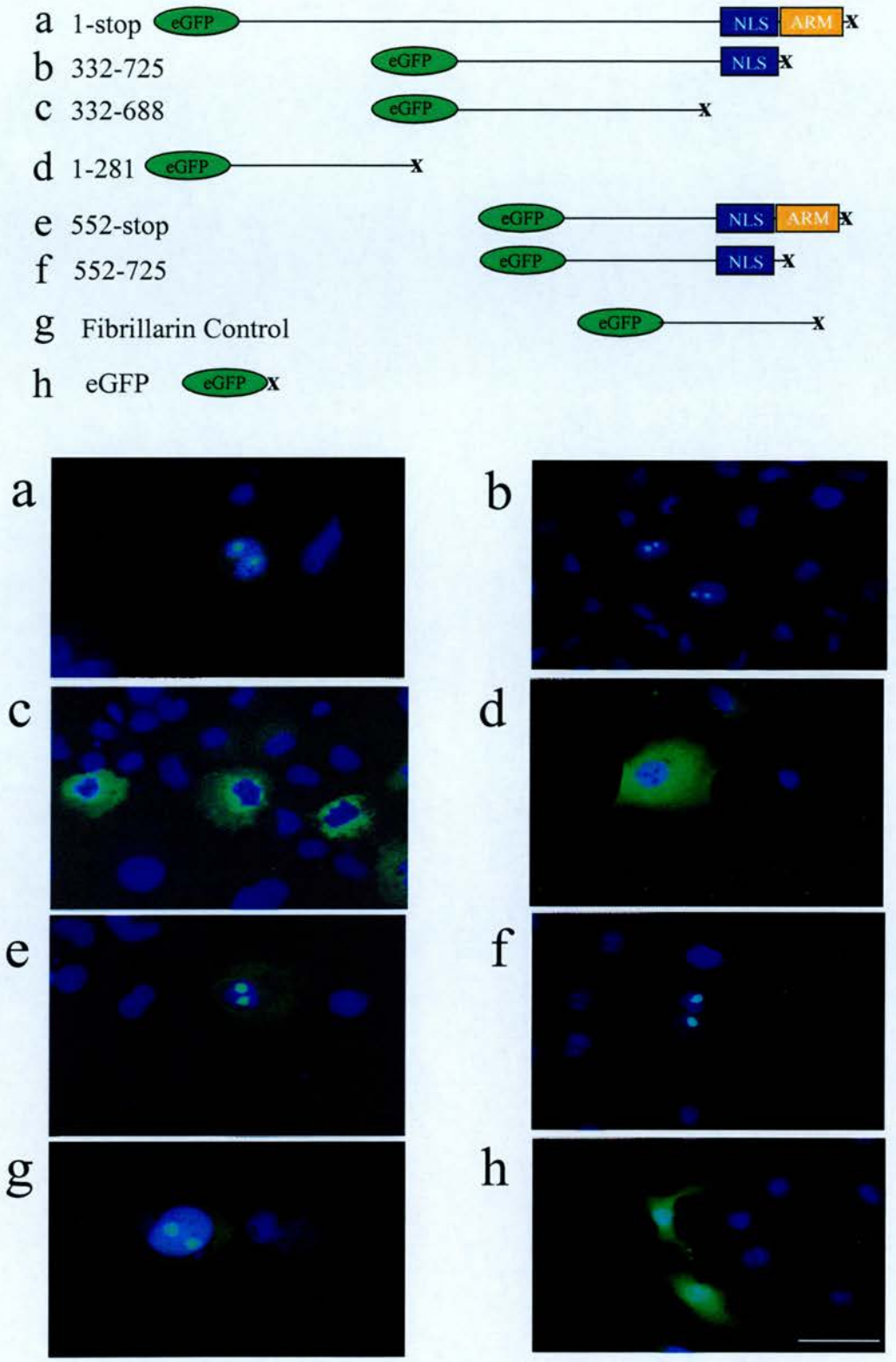


Figure 5.5: Testing putative NLS with eGFP-*ApO* fusions.
Schematic diagram of eGFP-*ApO* fusions transfected into COS 7 cells and DAPI stained COS 7 cells transfected with full length *ApO* (a) and various truncated forms of *ApO* fused to eGFP (b-f). Fibrillarin-eGFP fusion (g) and eGFP alone (h) are shown as controls. Scale bar 10μM

5.3 Discussion

Using eGFP-*ApO* reporter constructs it has been shown that this protein can localise to the nucleolus and a putative nuclear/nucleolar localisation signal sequence has been defined by deletion analysis. Recombinant eGFP-APO protein that contain the putative NLS always localise to the nucleoli, whilst constructs without the putative NLS remain in the cytoplasm. To unequivocally prove this it would be necessary to clone the putative NLS (RVNRRPRKRKRRGKR) alone in-frame with the eGFP protein and observe the fluorescent signal when transfected into COS 7 cells. One would expect the eGFP-NLS to traffic to the nucleolus if the hypothesis was correct.

Furthermore, to characterise the putative NLS signal sequence more fully, single codon mutations could be introduced by site direct mutagenesis. The mutations which prevent the trafficking to the nucleus would define the NLS more specifically.

A comparison of the aminopeptidase O NLS sequence between mouse, rat, human and puffer fish is as follows:

Mouse	RVNRRPRKRKRRGKR
Rat	RVNRRPGKRQRRKR
Human	GVNRRPRKRKRREK
Fugu (puffer fish)	LSRSHGKSKKRKR

There is not a conserved motif, although the sequence does contain basic amino acids which are common in canonical NLS sequences. An experiment with the eGFP-APO fusion proteins from other species with and without the putative NLS could be done to ascertain if this domain is functionally conserved between species.

The eGFP-APO fusions with the NLS sequence localise to the nucleoli. It is possible a nucleoli specific signal has been identified. Several examples where short nucleolar targeting signals (NTS) are able to traffic heterologous proteins to the nucleolus have been demonstrated in viral proteins (Kalland et

al., 1994; Nosaka et al., 1989) and in the heat shock chaperone proteins (HSC70/HSP70) which conditionally localise to the nucleolus as part of the cellular response to stress (Dang and Lee, 1989; Hunt and Morimoto, 1985). A common feature of these NTSs is that they contain regions of basic amino acids, for example the human immunodeficiency virus (HIV) tat protein is: (GRKKRRQRRRAP) and HSP 70 is: (KRKHKKDISQNKRAVRR). This is very much in keeping with murine APO signal sequence which is: RVNRRPRKRKRRGKR.

Whilst the experimental evidence for the putative NLS is compelling, one should be aware that over expression of a recombinant protein from a strong promoter can have possible misleading results. Proteins expressed at high concentrations can lead to misfolding and incorrect localisation (Garcia-Mata et al., 1999), although it has been demonstrated that a nucleolar protein fibrillarin expressed in the same system localises correctly (Newton et al., 2003). To determine the physiological relevance and show that APO protein truly localises to the nucleoli one should look at the endogenous protein in endothelial cells with an APO specific antibody.

Purification of nucleoli from HeLa cell lines and mass spectrophotometry analysis of the constituent proteins has not identified aminopeptidase O in this sample (www.lamondlab.com/Nopdb (Andersen et al., 2002; Scherl et al., 2002). This may reflect, however, the cell line used, as it has been found in this study that *ApO* is expressed in the vasculature and, therefore, may not be represented in HeLA cells.

The aminopeptidase O has an armadillo repeat in the C-terminus and has been described as a member of the armadillo repeat aminopeptidases (Chapter 3 section and Diaz–Pereles 2005). Many proteins involved with the transfer of proteins to the nucleus have armadillo/HEAT repeats such as the importin family of nuclear transport proteins (Malik et al., 1997). The armadillo repeats form a grooved structure in which basic peptides of the NLS are able to bind (Conti et al., 1998). Interestingly, APO has a putative NLS sequence immediately upstream of the armadillo repeat motif. The protein was able to

translocate to the nucleoli when the armadillo repeat was absent (Figure 5.4 b and f) indicating that the ARM repeat is not essential for nucleolar localisation. Perhaps the APO ARM motif binds to other proteins (as yet undefined) and by virtue of the NLS guides the new protein complex to the nucleolus acting as nucleolar cargo transport protein. Performing immunoprecipitation on endothelial or transfected cells, followed by two dimensional gel electrophoresis might allow their identification (Delahunty and Yates, 2005).

Another important consideration is that *ApO* alternative splicing can generate isoforms without the NLS (Chapter 3 Figure 3.2). Alternative splicing is a mechanism of gene control that can yield proteins with different functions from the same gene such as VEGF where splicing gives rise to functionally distinct isoforms with different signalling properties (Robinson and Stringer, 2001). The *ApO* gene has a number of alternative splice variants which remove regions of the gene including those that encode the catalytic domain and the NLS sequence. These will presumably have important functional consequences for *ApO* and its associated proteins. For example transcripts which do not have the putative NLS would not localise to the nucleoli and it would be expected that they perform a different function to transcripts that are in the nucleoli. The nucleoli are sites for a number of different cellular processes ranging from ribosome biogenesis, cell cycle regulation, DNA damage repair, pre-mRNA processing (Desterro et al., 2003; Olson et al., 2002; Pederson, 1998; Tsai and McKay, 2002). It is possible to speculate that *ApO* may play a critical role in one or more of these processes. Furthermore, because *ApO* is expressed in the vasculature and the relationship between angiogenesis and diseases such as cancer, one could contemplate that *ApO* may offer a potential therapeutic target.

Chapter 6:

The expression and purification of fusion proteins to raise an antibody against aminopeptidase O

6.1 Introduction

The experiments in chapter 5 demonstrate there is a putative NLS signal in the APO protein. To determine whether this has a physiological function *in vivo* and does in fact localise to the nucleolus, the endogenous protein would have to be studied. One approach to facilitate this study would be to generate an antibody against the APO protein. An antibody could also be used to purify the aminopeptidase O protease by immunoprecipitation. The purified protein could then be tested in enzyme assays to determine substrate specificity.

There are different methods available to generate the protein used to raise an antibody. One approach is to analyse the protein sequence computationally and use software that will predict antigenic epitopes. There are a number of programmes in existence, for example (http://bio.dfci.harvard.edu/Tools/antigenic.html). The desired peptide/s is then synthetically manufactured before injecting into the host animal. Another approach would be to clone the gene into an expression vector in- frame with a tag that allows easy purification such as a GST or 6XHistidine which bind sepharose and charged nickel respectively. The fusion protein is expressed and purified from contaminating proteins prior to injection into the recipient animal.

The approach taken in this thesis was to express a GST-Fl-APO protein. The APO protein used did not contain the catalytic domain which was predicted less likely to cross-react with other aminopeptidases. The raising of an antibody against the APO protein and subsequent testing by Western blot on COS 7 cell lysates transfected with eGFP tagged murine APO proteins is described in this chapter.

6.2 Results

6.2.1 Expression and purification of GST-Fl-APO (minus catalytic domain)

The Fl-*ApO* 2.3 kb (Figure 3.1g) was cloned into the bacterial expression vector pGEX 4T-1 and expressed in protease deficient bacteria (BL21). Upon IPTG induction an N-terminal GST tagged Fl-APO fusion protein (116 kDa) was generated. The expression of the fusion protein was optimised using different concentrations of IPTG, temperature, and time of induction (hours). The conditions that yielded the highest fraction of solubilised full length fusion protein (0.4mM IPTG induction at 37°C for 3 hours) were identified and adhered to in further experiments (Figure 6.1a).

The bacterial pellet was mechanically lysed by sonication. The soluble fraction was passed over a glutathione column to bind the GST tagged protein. Unbound proteins were removed by washing. The retained GST bound proteins were eluted from the column with reduced glutathione and collected as fractions (Figure 6.2a). There were a number of smaller fragments copurifying with the GST-Fl-APO fusion protein (Figure 6.2a). To test if these were bacterial contaminating proteins or degraded breakdown products of the fusion protein, a Western blot with antibody against GST was carried out (Figure 6.2b). The result indicated that many of the additional bands were degraded fragments of the GST-Fl-APO fusion protein.

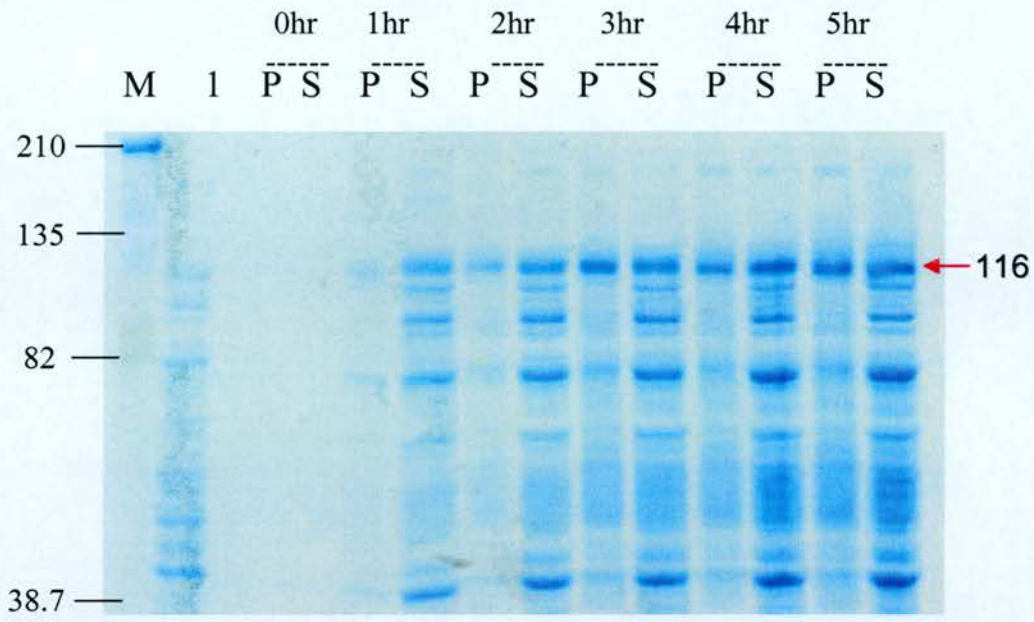


Figure 6.1: Optimisation of GST-Fl-*ApO* (1-stop) expression at 37°C using 0.4mM IPTG.

a). Untransformed wild type BL21 bacteria (1) transformed BL21 pellet (p) and supernatant (s) at different time of induction (hours).

red arrow indicates purified GST fusion protein (GST-Fl-APO) at the expected size of 116 kDa.

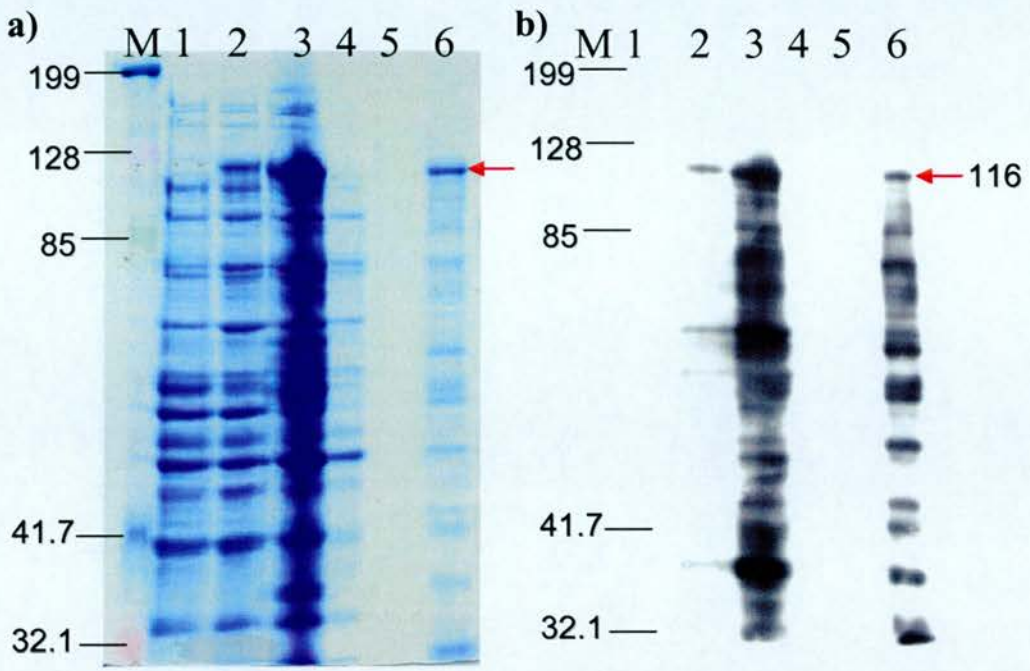


Figure 6.2: Purification of recombinant GST tagged APO protein
a) Purification of GST-Fl-APO and b) Western blot with anti-GST antibody. Negative control, BL21 bacteria without vector (1), BL21 transformed GST-Fl-APO(1-stop) no induction (2), BL21 GST-Fl-*ApO* (1-stop) induced with IPTG (3), first wash after binding supernatant to GST-sepharose column (4), final wash (5), elution from column (6). Red arrow indicates purified GST fusion protein at the expected size of 116 kDa.

6.2.2 Raising and testing of anti-APO antibody

Approximately 1mg of the purified full length GST-APO fusion protein was injected into a rabbit (Scottish Cell Diagnostics). Preimmune sera, first, second and third bleeds (at 8, 12 and 16 weeks, respectively) were collected. To demonstrate if the raised antibody could detect APO protein, a Western blot of COS 7 cells that had been transiently transfected with the cDNA encoding eGFP-Fl-APO protein was performed. Identical blots were hybridised with an antibody against eGFP and the preimmune sera, and the first, second and third bleeds of the rabbit immunised against APO. The preimmune sera did not detect any bands on the Western blot (Figure 6.3a). The antibody against eGFP detected proteins of the predicted sizes (Figure 6.3b) and the third bleed from the rabbit detected a similar band in the lane loaded with lysate from cells transfected with a cDNA encoding the eGFP-APO fusion protein.

The diffuse background on the Western blots resulted in a loss of specific signal with long exposure. To reduce the signal to noise ratio, different dilutions of the antibody were tried (Figure 6.3 c, d and e), but this did not resolve the problem. Similar experiments were conducted with the first and second bleeds but the results were the same (data not shown).

Many different conditions and additives were tried to reduce the background interference, including using protein membranes from different suppliers and a variety of blocking agents (5 and 10% milk, BSA, gelatin all with Tween 20 at a maximum concentration of 2%) with increased blocking and washing times. None of the conditions mentioned dramatically improved the signal to noise ratio (data not shown). The antibody was also immuno-affinity purified on protein 'A' column but this experiment was not successful (data not shown).

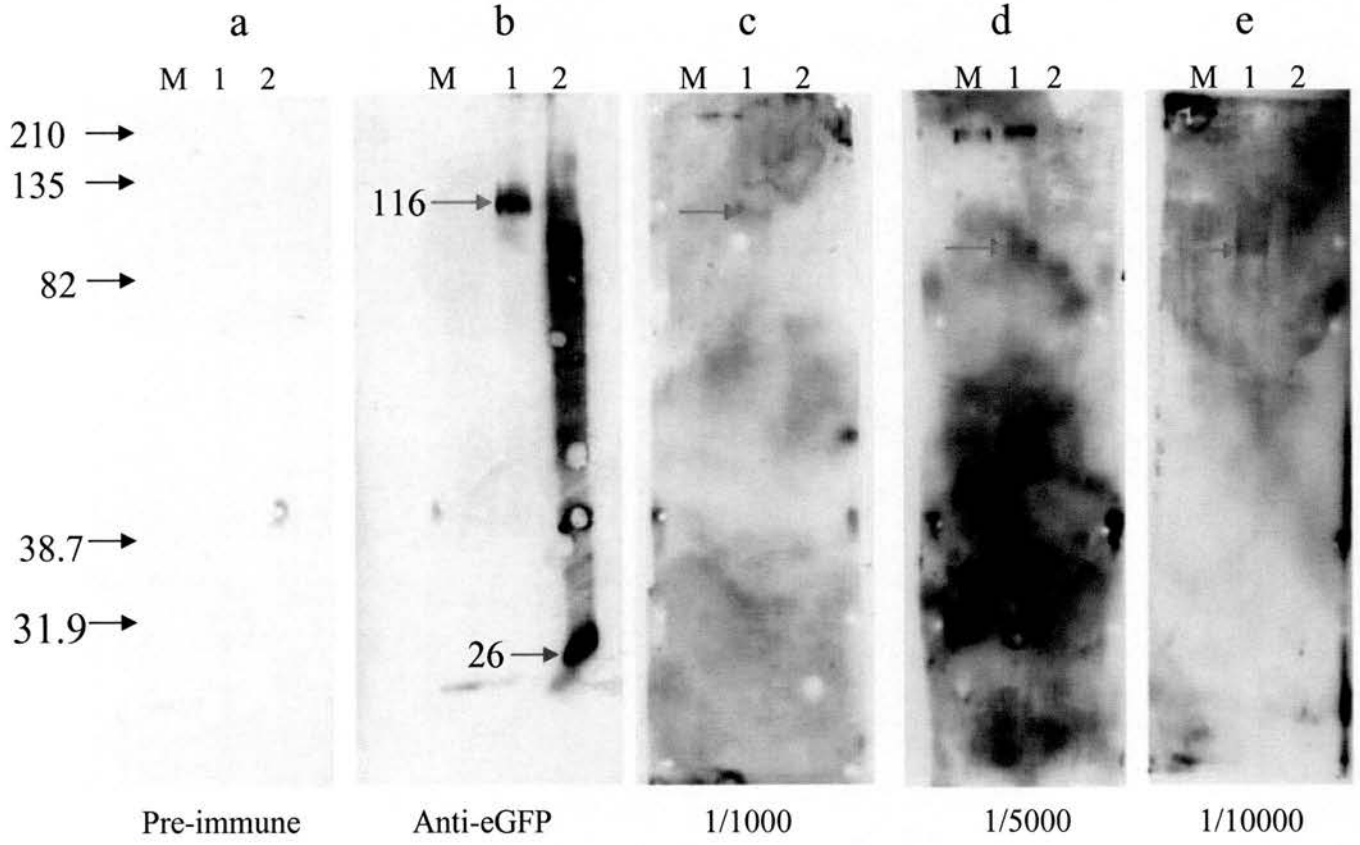


Figure 6.3: Western blot analysis of protein lysates from COS 7 cell transient transfections.
eGFP-Fl-*ApO* (1-stop) (lane 1) or eGFP alone (lane 2) using pre-immune serum (a), an anti eGFP antibody (b) and serum from the rabbit immunised against GST-Fl-*ApO* (1-stop) (3rd bleed) at different dilutions (c,d,e). Red arrow indicates eGFP-*ApO* fusion (116kDa), blue arrow indicates eGFP (26 kDa).

6.2.3 Immuno-affinity purification using a His tagged Fl-APO

The anti APO was purified from serum using a different immuno-affinity purification procedure (Harlow E, 1999). This involved the generation of an N-terminal His tagged Fl-APO (without catalytic domain) fusion protein which was then purified on a charged nickel column. The purified fusion protein was irreversibly bound by cyanogens bromide linkage to sepharose beads and the rabbit serum was then passed over the beads to capture the aminopeptidase antibody (Figure 6.4).

The expression of the fusion protein was optimised using different conditions (Section 6.2.1). In this expression system 1mM IPTG induction at 37°C for 3 hours yielded the highest levels of fusion protein. A Western blot with an antibody to the His tag was used to verify that the induced band was the desired fusion protein, although there were some additional breakdown protein products (Figure 6.5).

The fusion protein yield was low with the majority of the protein in the insoluble aggregate fraction (data not shown). It was decided to solubilise the His tagged fusion protein from the aggregates with 6M urea.

Initial attempts to bind the His-Fl-APO protein using the charged nickel colums for 1 hour failed with no protein detected in the eluted fraction (Figure 6.6 a). However, when this was increased to overnight incubation a significant amount of fusion protein was detected in the eluted fraction (Figure 6.6 b).

This purified fusion protein was then bound to the activated cyanogen bromide sepharose beads. These beads were then used to purify the anti-APO antibody from serum and tested on a Western blot against various eGFP tagged APO fusions.

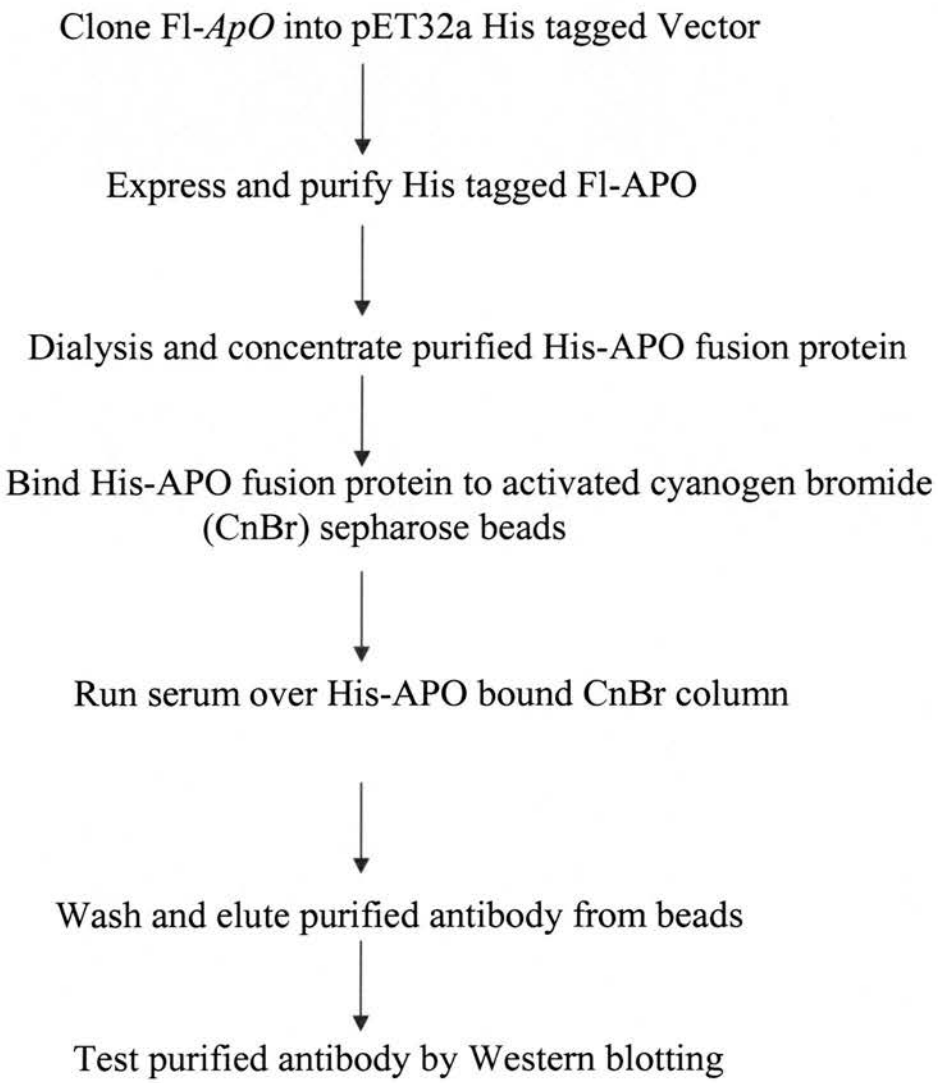


Figure 6.4: Schematic diagram illustrating immuno affinity purification of anti APO from sera.

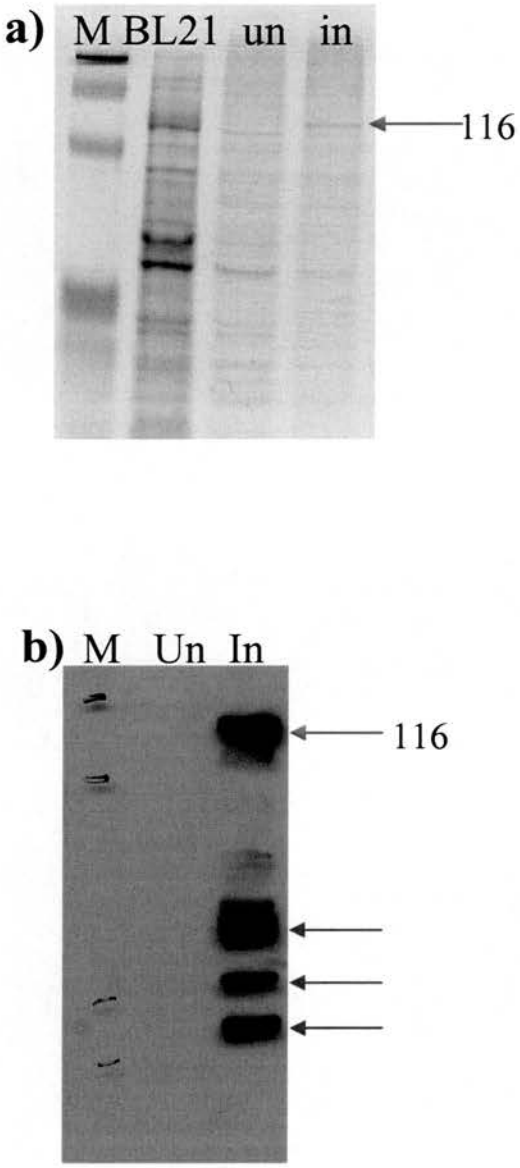


Figure 6.5: Expression of HisTagged Fl-APO (1-stop).

a) Coomassie stained gel of protein lysates from untransformed BL21 bacteria or transformed His-Fl-APO (1-stop) BL21 cells uninduced (un) or induced (in) with IPTG.

b) Western blot analysis of BL21 cell lysates either uninduced (un) or induced (in) with 1 mM IPTG for 3 hours using an anti His antibody.

Red arrow marks full length fusion (116kDa).
Blue arrows indicate breakdown products.

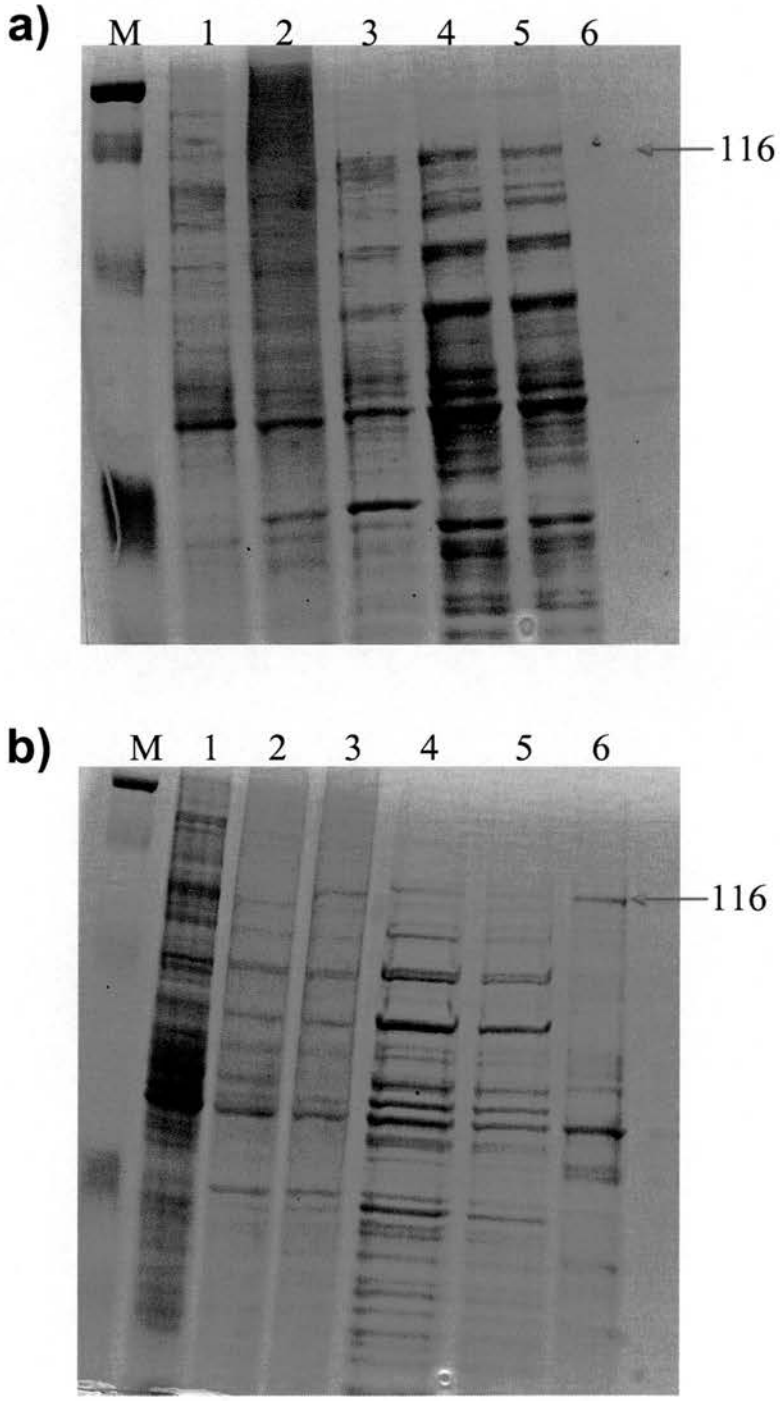


Figure 6.6: Expression and purification of His tagged Fl-*ApO* (1-stop) with out catalytic domain under two different purification conditions.

(a) 2 hours and (b) overnight.

Marker (M), BL21 untransformed (1), uninduced (2), induced (3), solubilised fraction (4), first wash (5) and elution from column (6).

6.2.4 Testing of immuno purified ANTI APO on recombinant APO protein

Western blot analysis using the commercial anti eGFP antibody detected the eGFP protein and tagged eGFP-Fl-APO (1-stop) eGFP-CT- APO (552-stop), and eGFP-NT-APO (1-281) (Figure 6.7 a). The purified anti APO antibody detected similarly sized bands in the eGFP-Fl-APO and eGFP-NT but not the eGFP-CT-APO protein product (Figure 6.7 b). A possible reason might be that the antigenic epitopes were not present on the C-terminus of the protein.

Antibody dilution at 1 in 5000 resulted in an intense specific signal with short exposure times typically less than 1 minute (Figure 6.7 c). The signal to noise ratio was very low with long exposures and the diffuse high background observed with the unpurified sera was eliminated (data not shown). This result indicated that the antibody purification had been successful. The additional band present with the eGFP-NT–APO on both the Westerns was assumed to be degraded APO protein (Figure 6.7 a and b, lane 4).

6.2.5 Testing immuno purified antibody on mouse tissues by Western blot

To assess whether this immunopurified antibody could be used to detect APO in mouse tissue lysates, total protein was extracted from wild type mice and mice carrying the heterozygous and homozygous GT411 gene trap. One would have expected APO-βgeo fusion proteins associated with the gene trap integration to shift in the heterozygote and homozygote tissues when compared to wild type APO protein. Attempts to use the immuno-purified anti APO antibody on mouse tissues were unsuccessful (data not shown).

Given that the anti APO did not work on tissues we tried a control experiment was attempted with a commercial antibody against β-gal. APO-βgeo fusion proteins (60+156kDa) were detected in the heterozygote $ApO^{Gt411For/+}$ and homozygote $ApO^{Gt411For/Gt411For}$ hearts (Figure 6.8).

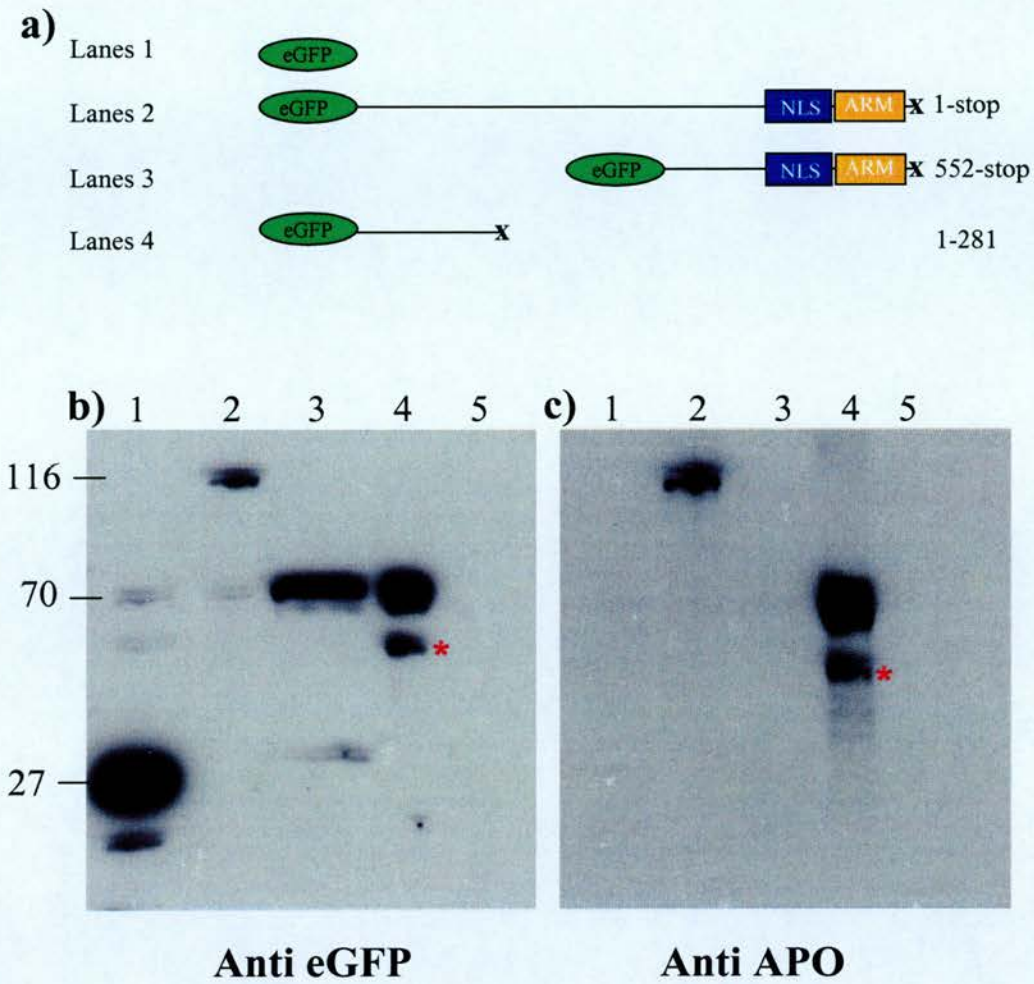


Figure 6.7: Testing of the immunopurified APO antibody using COS 7 cells transfected with a range of eGFP- APO fusion proteins.

a). Schematic illustration of eGFP clones transfected into COS 7 cells.

Western blot analysis of protein lysates isolated from COS 7 cells transiently transfected with eGFP alone (1), eGFP-Fl-*ApO* (2), eGFP-C-Term *ApO* (3), eGFP-N-Term-*ApO* (4) and mock transfected cells (5). Probed with either the anti eGFP (b), or the immunopurified anti APO antibody (c).

Asterisks indicate putative breakdown products.

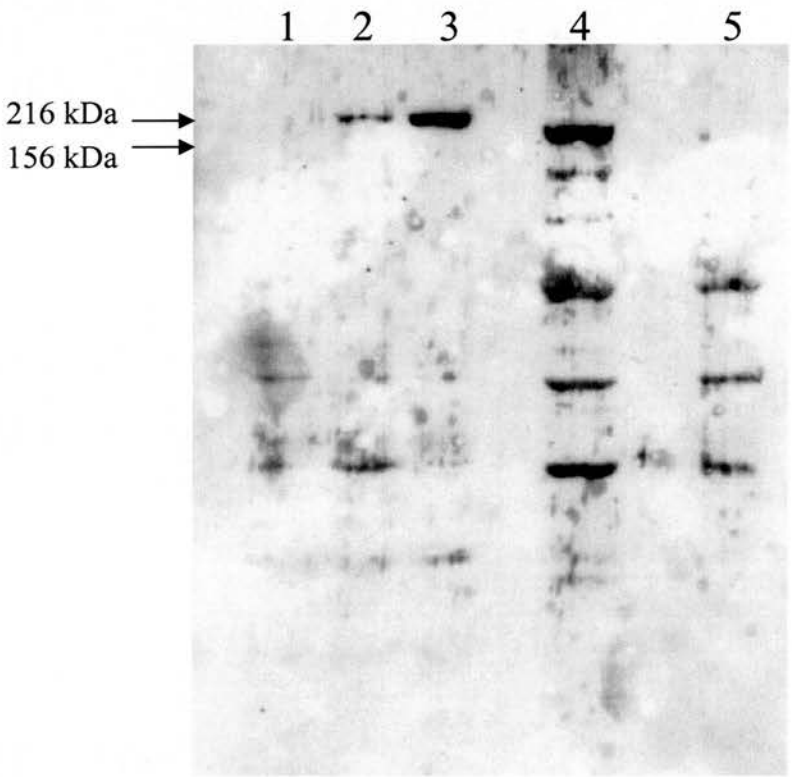


Figure 6.8: Western blot analysis of protein lysates isolated from adult mouse hearts with an anti βgal antibody to detect fusion proteins.

Wild type (1), heterozygous aminopeptidase O $^{Gt411For/+}$ (2), homozygous aminopeptidase O $^{Gt411For/Gt411For}$ (3), and ES cell expressing OCT4-IRES- βgeo (4), and wild type ES cell (5).

The heterozygote $ApO^{Gt411For/+}$ signal was approximately half the intensity of the homozygote $ApO^{Gt411For/Gt411For}$. The positive control (gifted by Dr. Tom Burdon) was a protein lysate from ES cells expressing OCT4-IRES-βgeo (156kDa) (Figure 6.8, lane 4). The negative control was a protein lysate from wild type ES cells (Figure 6.8, lane 5). The bands seen in the heterozygote $ApO^{Gt411For/+}$ and homozygote $ApO^{Gt411For/Gt411For}$ hearts with the commercial β-gal antibody were larger than the OCT4-IRES-βgeo control indicating that they were authentic and not artefactual bands.

6.2.6 Immunohisochemistry with APO antibody on paraffin embedded kidney sections.

To assess the purified aminopeptidase antibody applicability for immunohistochemistry it was tested on paraffin embedded kidney sections. The lacZ staining of $ApO^{Gt411For/+}$ and homozygote $ApO^{Gt411For/Gt411For}$ adult kidney sections had demonstrated staining of the glomeruli and afferent vessels and was considered a suitable tissue to optimise the immunohistochemistry protocol. The *ApO*-LACZ reporter staining within the glomeruli when compared to the afferent vessels appears to be stronger indicating higher expression in this region of the tissue. The purified APO antibody indicated specific staining of the glomeruli which is comparable to the βgal staining in kidney cryosections (Figure 6.9 a, b and c). However, there was no apparent staining of the vessels leading into the glomerulus, which may be due to the fact that APO expression is lower and beyond the sensitivity of the antibody. The immunohistochemistry procedure would require additional trouble shooting to maximise the protocol, for example a nickel staining method could be tried (Paraffin embedded sections cut by Eli Mohammed and staining performed by Julie Buchannan).

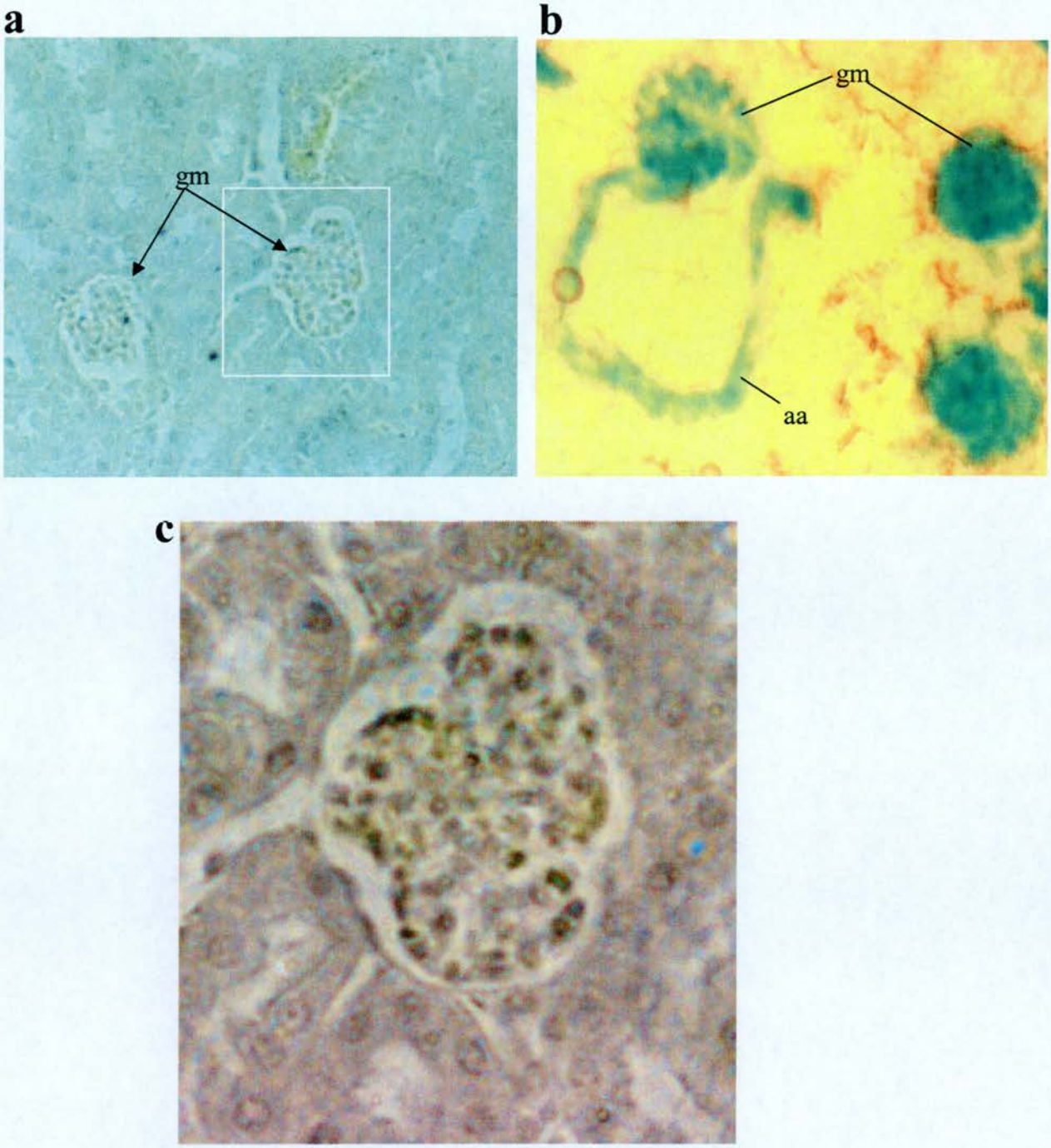


Figure 6.9: Testing of immunoprurified APO antibody on parafin embedded kidney section.

a) Immunohistochemistry of paraffin embedded adult kidney with GT411 purified antibody X10.

b) βgal stained adult kidney homozygote $ApO^{Gt411For/Gt411For}$ cryosection.

c) High power of boxed region showing APO specific staining of cells within the glomerulus X40 (b) gm=glomeruli, aa=afferent arteries.

6.2.7 Immuno Precipitations (IP) of recombinant APO protein

One strategy to isolate purified catalytically active enzyme is to IP the endogenous protein from tissues. This approach has been taken with matrix metallo-proteases where they are then assayed with known substrates in techniques such as zymography. It was decided to test whether antibody APO could IP eGFP tagged APO fusions from transiently transfected COS 7 cells.

The control IP with anti eGFP antibody captured the eGFP (26 kDa), eGFP-Fl-APO (1-stop) and eGFP-CT-APO (552-stop) proteins (Figure 6.10 b). The eGFP-Fl-ApO protein (lane 2) had a ladder of faint bands (largest 116 kDa) and these were assumed to be breakdown products. Note the strong band at 40 kDa. The eGFP-CT-APO protein was also captured (Figure 6.10 b, lane 3), although there are a number of breakdown products, including a strong band at 40 kDa and one at the same position as the eGFP protein alone (26 kDa).

The APO antibody immunoprecipitated bands shown in lane 2 (Figure 6.10 c), however, these were too small to be the full length fusion (116 kDa) and were again assumed to be breakdown products. It was very similar to the bands observed in the eGFP-Fl-APO (Figure 6.10 b lane 2).The APO antibody did not IP the C-Terminal fusion successfully. It is likely that the C-terminus does not contain epitopes recognised by the APO antibody and this concurs with the Western data (see section 6.2.4). The mock transfections served as a negative control (Figure 6.10 b and c, lane 4). In both Western blots there is a constant band in all the lanes which is probably the heavy IGG chain (58kDa).

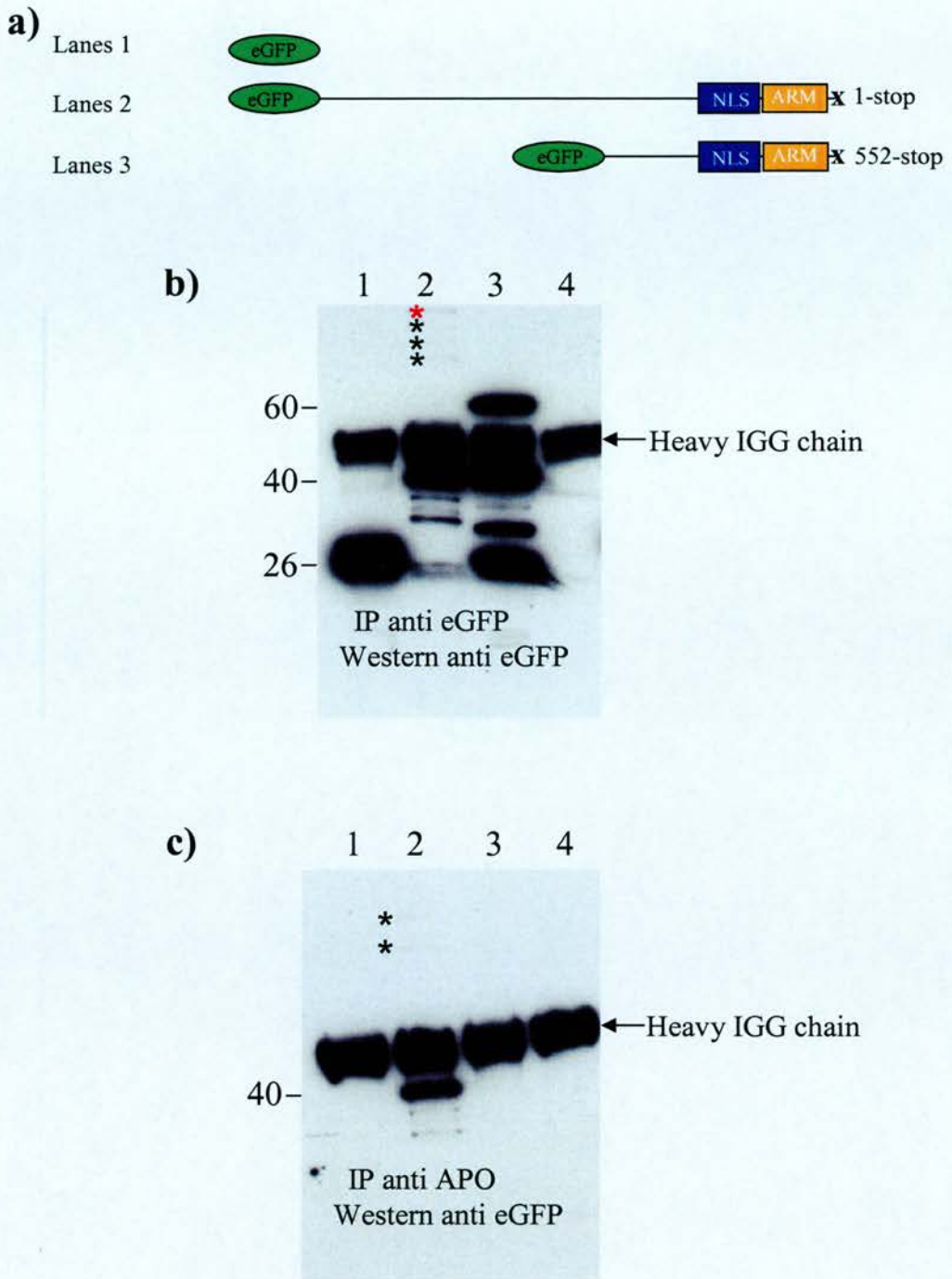


Figure 6.10: Immunoprecipitation of protein lysates from COS 7 cells transiently transfected with eGFP-fusions.
a) Schematic diagram of eGFP-*ApO* fusions used in transfections. b) and c) Western blots of with eGFP antibody eGFP lane 1, eGFP-Fl-APO (1-stop) lane 2, eGFP-C-terminal-APO (552-stop) lane 3, or mock transfected lane 4. Using anti–eGFP (b) or anti-APO (c) antibodies and then subsequently analysed by Western blot using anti eGFP antibody. Asterisks indicate faint full length eGFP-APO fusion (red) and breakdown products (black).

6.3 Discussion

Antibodies provide a powerful approach to characterise gene expression more fully. There are several advantages to this strategy. Firstly, the endogenous gene product can be directly analysed at the cellular and tissue level. Secondly, immunoprecipitation of endogenous protein enables functional activity to be assayed. Thirdly, interacting proteins can be discovered by co-immunoprecipitation. Taken together antibodies provide a valuable resource to evaluate and determine gene function.

In this study an antibody was raised in a rabbit against the full length murine aminopeptidase O (without catalytic domain) and its efficacy tested by Western blotting, immunoprecipitation and immunohistochemistry. It was demonstrated that the APO antibody can detect transiently expressed aminopeptidase O protein from COS 7 cells but not the endogenous protein from whole tissue lysates. It is possible to speculate that failure of the antibody to detect endogenous APO proteins may be due to protein instability or that the protein lysate preparation caused solubility difficulties. These hypotheses seem less credible since the APO-βgeo fusion protein was detected with a commercial β-gal antibody. Perhaps the purified aminopeptidase O antibody developed in-house was inferior to commercial preparations.

The APO antibody was successfully used to IP recombinant fusion protein from COS 7 cell lysates. The C-terminal (552-stop) APO fusion protein was not captured by the APO antibody. Similarly the antibody did not detect this C-terminal fusion protein by Western blot either. A plausible explanation may be that the C-terminal portion of the protein did not contain antigenic epitopes.

Other experiments which should now be undertaken would be to IP endogenous APO from tissues or a mouse endothelial cell line such as the murine cell endothelial cells (MCEC) (Lidington et al., 2002). This would allow purification of the protease which could then be assayed with fluorogenic substrates to confirm its aminopeptidase activity (Grant et al.,

2002). Additionally, interacting partners could be resolved by 2-D electrophoresis and protein sequencing (Delahunty and Yates, 2005).

Chapter 7:

Summary and prospective

7.0 General discussion

7.1 Summary and prospective

During a gene trap screen designed to identify novel regulated genes in cardiovascular development, the ES clone GT411 was isolated. Originally the EST database contained only a short contig for this gene. The full length open reading frame was determined using 5'RACE. The gene was identified to be a novel aminopeptidase but during the course of the research its human orthologue was cloned and named Aminopeptidase O (*APO*) (Diaz-Perales et al., 2005).

Aminopeptidase O contains several functional domains. First, a protease domain has been identified indicating that it is a member of the M1 family of aminopeptidases. These enzymes are involved in a number of important biological processes including angiogenesis, maintenance of blood pressure and inflammatory pathways (Haeggstrom, 2004; Mitsui et al., 2004; Sato, 2003). Secondly, prediction of an armadillo repeat structure suggests that *ApO* may be implicated in the transport of proteins to the nucleus (Malik et al., 1997). In addition, characterisation of a putative nuclear localisation signal sequence (NLS) was undertaken using eGFP-*ApO* fusions with and without the NLS sequence. The eGFP-*ApO* fusion proteins with the NLS were directed to the nucleoli while those without were retained in the cytoplasm. A comparison of the NLS protein sequence between mouse, rat, human and puffer fish did not show an absolutely conserved NLS, although the sequence did contain a stretch of basic amino acids which are common in canonical NLS sequences. An experiment with the eGFP-APO fusion proteins from other species with and without the putative NLS could be done to ascertain if this domain is functionally conserved between species.

Analysis from Northern blot and RT-PCR data has shown that *ApO* has alternatively spliced isoforms which in some instances can remove the

catalytic domain and the NLS. It would be predicted that transcripts that do not contain these domains would have different functional activity.

A novel regulatory mechanism of enzyme function has been described for the aminopeptidase thytotropin releasing hormone degredation enzyme. This gene expresses an isoform which does not contain the catalytic domain and can bind as a heterodimer to the isoform that contains the catalytic domain preventing enzymatic activity (Chavez-Gutierrez et al., 2005). It is possible that the *ApO* isoform without the catalytic domain could act in a similar manner as a negative regulator of *ApO* enzyme function.

Splicing enables expression of functionally diverse proteins from a single gene. There are alternative splicing events for the *ApO* gene involving the catalytic domain and the NLS. The function of these domains could be investigated by a gene targeting knock-in strategy whereby only one particular isoform is expressed. A similar approach was taken with the zinc finger transcription factor ZBP-89 where a splice isoform lacking the amino terminal residues 1-127 of the full-length protein was knocked in to define its function. The ZBP 89ΔN homozygous mice expressing ZBP 89ΔN protein had a phenotype and it was concluded that the amino terminus of the protein was required for gastrointestinal homeostasis (Law et al., 2006).

The gene trap mice did not appear to have a phenotype. Specifically, analysis of 8.5, 9.5 and 10.5 d.p.c. embryos and sections of adult tissues did not show a gross morphological abnormality in the vasculature. Moreover, using the aortic ring assay, no discernable phenotype was observed.

The recent publication of the human aminopeptidase O orthologue and the demonstration that it can enzymatically cleave Ang III into Ang IV clearly implicates APO's involvement in the renin angiotensin system in the maintenance of blood pressure. Comparison of heart and kidney weights dissected from 1 month old wild type, heterozygous $ApO^{Gt411For/+}$ and homozygous $ApO^{Gt411For/Gt411For}$ gene trap mice did not demonstrate any statistically significant differences using T-tests. A limitation of this experiment was that the mice were young and a high blood pressure

phenotype might take longer to develop. This experiment could be performed on older mice and one could also compare and look for thickening of the aorta by paraffin embedded sectioning. Although it has been recently demonstrated that APO can cleave Ang III into Ang IV in an *in vitro* enzyme assay (Diaz-Perales et al., 2005), there is at present no evidence to show that this occurs *in vivo*. Indeed the vasoactive peptides of the renin angiotensin system are systemically transported in the blood and are either enzymatically cleaved by secreted enzymes or cleaved by membrane bound enzymes. APO does not contain transmembrane domains or signal peptides to direct it to the cell membrane. It is, therefore, unlikely that APO is physiologically able to participate directly in the catabolism of the vasoactive peptide Ang III.

The absence of a phenotype in the gene trap mice might be explained by the fact that the integration has occurred in the last intron of the *ApO* gene. This removes 46 amino acids from the protein and it is possible that it may retain all or part of its normal function despite being fused to the βgeo gene.

Future research in this area could involve trying to determine the function of aminopepetidase O. One strategy might be to ablate the *ApO* gene by homologous recombination strategies either involving the generation of a complete null or a conditional knockout. Phenotypic analysis of mutant mice may allow functionality to be ascribed to the gene. Many gene knockout studies have not resulted in a phenotype and a possible cause is that there may be genetic redundancy. The closest relative to aminopeptidase O is aminopeptidase B and this might be a candidate for redundancy. One way to increase the severity of a phenotype is to cross the mutant lines to generate a double knock out (Heber et al., 2000).

Other strategies that could be used to determine gene function include morpholino antisense technology against the frog or zebra fish *ApO* orthologues (Summerton et al., 1997) and RNA interference (RNAi) where a short double stranded RNA specific to the APO gene is introduced into an endothelial cell line (HUVECS) and analysed with an endothelial cell assay (Fukasawa et al., 2006). The effect of these strategies might be to down-

regulate the endogenous gene expression which in turn may cause a phenotype.

In conclusion the murine *ApO* gene was isolated using a gene trapping screen designed to discover novel genes involved in cardiovascular development. Although this approach was successful in isolating this gene and demonstrated that the APO was expressed in the vasculature, it was a labour intensive strategy. A desirable consequence of gene trapping is that it may be mutagenic when transmitted through the germ line although in this instance no apparent phenotype was evident. With the benefit of hindsight and with recent technological advances, one could take different approaches to identify genes involved with vascular development. For example, a subtractive screen comparing the gene expression of an endothelial cell line with a different cell type might identify genes that are required solely for vascular function.

Germline transmission of the GT411 ES cells demonstrated *LacZ* reporter expression in developing embryonic and adult vasculature. Understanding the function of *ApO* therefore might provide insight into vascular biology in both normal developmental and pathological conditions. These could include processes such as angiogenesis, maintenance of blood pressure as well as its deregulation in diseases such as cancer and hypertension and may offer a potential therapeutic target for treatment of these pathologies.

References

Adams, J. M. (2003). Ways of dying: multiple pathways to apoptosis. *Genes Dev* 17, 2481-95.

Ahmad, S. and Ward, P. E. (1990). Role of aminopeptidase activity in the regulation of the pressor activity of circulating angiotensins. *J Pharmacol Exp Ther* 252, 643-50.

Andersen, J. S., Lyon, C. E., Fox, A. H., Leung, A. K., Lam, Y. W., Steen, H., Mann, M. and Lamond, A. I. (2002). Directed proteomic analysis of the human nucleolus. *Curr Biol* 12, 1-11.

Andrade, M. A., Petosa, C., O'Donoghue, S. I., Muller, C. W. and Bork, P. (2001). Comparison of ARM and HEAT protein repeats. *J Mol Biol* 309, 1-18.

Antoine, N., Greimers, R., De Roanne, C., Kusaka, M., Heinen, E., Simar, L. J. and Castronovo, V. (1994). AGM-1470, a potent angiogenesis inhibitor, prevents the entry of normal but not transformed endothelial cells into the G1 phase of the cell cycle. *Cancer Res* 54, 2073-6.

Bailey,J.A. Yavor,A.M., Viggiano,I., Misceo,D., Horvath,J.E., Archidiacono,N., Schwarz,S., Rocchi,M and Eichler,E.E. (2002). Human specific duplication and mosaic trnscripts: the recent paralogous structure of chromosome 22. *Am J Hum Genet* 70,83-100

Bain, G., Ray, W. J., Yao, M. and Gottlieb, D. I. (1996). Retinoic acid promotes neural and represses mesodermal gene expression in mouse embryonic stem cells in culture. *Biochem Biophys Res Commun* 223, 691-4.

Baker, D., Silen, J. L. and Agard, D. A. (1992). Protease pro region required for folding is a potent inhibitor of the mature enzyme. *Proteins* 12, 339-44.

Benita, Y., Oosting, R. S., Lok, M. C., Wise, M. J. and Humphery-Smith, I. (2003). Regionalized GC content of template DNA as a predictor of PCR success. *Nucleic Acids Res* 31, e99.

Bergers, G., Brekken, R., McMahon, G., Vu, T. H., Itoh, T., Tamaki, K., Tanzawa, K., Thorpe, P., Itohara, S., Werb, Z. et al. (2000). Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nat Cell Biol* 2, 737-44.

Bjornsson, J. M., Larsson, N., Brun, A. C., Magnusson, M., Andersson, E., Lundstrom, P., Larsson, J., Repetowska, E., Ehinger, M., Humphries, R. K. et al. (2003). Reduced proliferative capacity of hematopoietic stem cells deficient in Hoxb3 and Hoxb4. *Mol Cell Biol* 23, 3872-83.

Blacher, S., Devy, L., Burbridge, M. F., Roland, G., Tucker, G., Noel, A. and Foidart, J. M. (2001). Improved quantification of angiogenesis in the rat aortic ring assay. *Angiogenesis* 4, 133-42.

Black, R. A., Rauch, C. T., Kozlosky, C. J., Peschon, J. J., Slack, J. L., Wolfson, M. F., Castner, B. J., Stocking, K. L., Reddy, P., Srinivasan, S. et al. (1997). A metalloproteinase disintegrin that releases tumour-necrosis factor-alpha from cells. *Nature* 385, 729-33.

Bogerd, H. P., Benson, R. E., Truant, R., Herold, A., Phingbodhipakkiya, M. and Cullen, B. R. (1999). Definition of a consensus transportin-specific nucleocytoplasmic transport signal. *J Biol Chem* 274, 9771-7.

Brock, T. G., Lee, Y. J., Maydanski, E., Marburger, T. L., Luo, M., Paine, R., 3rd and Peters-Golden, M. (2005). Nuclear localization of leukotriene A4 hydrolase in type II alveolar epithelial cells in normal and fibrotic lung. *Am J Physiol Lung Cell Mol Physiol* 289, L224-32.

Brooks, P. C., Silletti, S., von Schalscha, T. L., Friedlander, M. and Cheresh, D. A. (1998). Disruption of angiogenesis by PEX, a noncatalytic metalloproteinase fragment with integrin binding activity. *Cell* 92, 391-400.

Brun, A. C., Bjornsson, J. M., Magnusson, M., Larsson, N., Leveen, P., Ehinger, M., Nilsson, E. and Karlsson, S. (2004). Hoxb4-deficient mice undergo normal hematopoietic development but exhibit a mild proliferation defect in hematopoietic stem cells. *Blood* 103, 4126-33.

Byrom, J., Mudaliar, V., Redman, C. W., Jones, P., Strange, R. C. and Hoban, P. R. (2004). Loss of heterozygosity at chromosome 9q22-31 is a frequent and early event in ovarian tumors. *Int J Oncol* 24, 1271-7.

Cadel, S., Foulon, T., Viron, A., Balogh, A., Midol-Monnet, S., Noel, N. and Cohen, P. (1997). Aminopeptidase B from the rat testis is a bifunctional enzyme structurally related to leukotriene-A4 hydrolase. *Proc Natl Acad Sci U S A* 94, 2963-8.

Cao, Y., Ji, R. W., Davidson, D., Schaller, J., Marti, D., Sohndel, S., McCance, S. G., O'Reilly, M. S., Llinas, M. and Folkman, J. (1996). Kringle domains of human angiostatin. Characterization of the anti-proliferative activity on endothelial cells. *J Biol Chem* 271, 29461-7.

Carlson, C. M. and Largaespada, D. A. (2005). Insertional mutagenesis in mice: new perspectives and tools. ***Nat Rev Genet*** **6, 568-80.**

Carmeliet, P. and Jain, R. K. (2000). Angiogenesis in cancer and other diseases. ***Nature*** **407, 249-57.**

Cebe-Suarez, S., Zehnder-Fjallman, A. and Ballmer-Hofer, K. (2006). The role of VEGF receptors in angiogenesis; complex partnerships. ***Cell Mol Life Sci*** **63, 601-15.**

Chavez-Gutierrez, L., Bourdais, J., Aranda, G., Vargas, M. A., Matta-Camacho, E., Ducancel, F., Segovia, L., Joseph-Bravo, P. and Charli, J. L. (2005). A truncated isoform of pyroglutamyl aminopeptidase II produced by exon extension has dominant-negative activity. ***J Neurochem*** **92, 807-17.**

Chubinskaya, S., Mikhail, R., Deutsch, A. and Tindal, M. H. (2001). ADAM-10 protein is present in human articular cartilage primarily in the membrane-bound form and is upregulated in osteoarthritis and in response to IL-1alpha in bovine nasal cartilage. ***J Histochem Cytochem*** **49, 1165-76.**

Coates, J. C. (2003). Armadillo repeat proteins: beyond the animal kingdom. ***Trends Cell Biol*** **13, 463-71.**

Conlon, R. A. and Rossant, J. (1992). Exogenous retinoic acid rapidly induces anterior ectopic expression of murine Hox-2 genes in vivo. ***Development*** **116, 357-68.**

Conti, E. and Izaurralde, E. (2001). Nucleocytoplasmic transport enters the atomic age. ***Curr Opin Cell Biol*** **13, 310-9.**

Conti, E., Uy, M., Leighton, L., Blobel, G. and Kuriyan, J. (1998). Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. ***Cell*** **94, 193-204.**

Dang, C. V. and Lee, W. M. (1989). Nuclear and nucleolar targeting sequences of c-erb-A, c-myb, N-myc, p53, HSP70, and HIV tat proteins. ***J Biol Chem*** **264, 18019-23.**

Das, S., Mandal, M., Chakraborti, T., Mandal, A. and Chakraborti, S. (2003). Structure and evolutionary aspects of matrix metalloproteinases: a brief overview. ***Mol Cell Biochem*** **253, 31-40.**

Davis, S., Aldrich, T. H., Jones, P. F., Acheson, A., Compton, D. L., Jain, V., Ryan, T. E., Bruno, J., Radziejewski, C., Maisonpierre, P. C. et al. (1996). Isolation of angiopoietin-1, a ligand for the TIE2 receptor, by secretion-trap expression cloning. ***Cell*** **87, 1161-9.**

de Vries, C., Escobedo, J. A., Ueno, H., Houck, K., Ferrara, N. and Williams, L. T. (1992). The fms-like tyrosine kinase, a receptor for vascular endothelial growth factor. ***Science*** **255, 989-91.**

Decatur, W. A. and Fournier, M. J. (2003). RNA-guided nucleotide modification of ribosomal and other RNAs. ***J Biol Chem*** **278, 695-8.**

Delahunty, C. and Yates, J. R., 3rd. (2005). Protein identification using 2D-LC-MS/MS. ***Methods*** **35, 248-55.**

Deng, J. M. and Behringer, R. R. (1995). An insertional mutation in the BTF3 transcription factor gene leads to an early postimplantation lethality in mice. ***Transgenic Res*** **4, 264-9.**

Deshayes, F. and Nahmias, C. (2005). Angiotensin receptors: a new role in cancer? ***Trends Endocrinol Metab*** **16, 293-9.**

Desterro, J. M., Keegan, L. P., Lafarga, M., Berciano, M. T., O'Connell, M. and Carmo-Fonseca, M. (2003). Dynamic association of RNA-editing enzymes with the nucleolus. ***J Cell Sci*** **116, 1805-18.**

Diaz-Perales, A., Quesada, V., Sanchez, L. M., Ugalde, A. P., Suarez, M. F., Fueyo, A. and Lopez-Otin, C. (2005). Identification of human aminopeptidase O, a novel metalloprotease with structural similarity to aminopeptidase B and leukotriene A4 hydrolase. ***J Biol Chem*** **280, 14310-7.**

Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S. and Smithies, O. (1987). Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. ***Nature*** **330, 576-8.**

Dumont, D. J., Gradwohl, G., Fong, G. H., Puri, M. C., Gertsenstein, M., Auerbach, A. and Breitman, M. L. (1994). Dominant-negative and targeted null mutations in the endothelial receptor tyrosine kinase, tek, reveal a critical role in vasculogenesis of the embryo. ***Genes Dev*** **8, 1897-909.**

Early, P., Huang, H., Davis, M., Calame, K. and Hood, L. (1980a). An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. ***Cell*** **19, 981-92.**

Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. and Hood, L. (1980b). Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. ***Cell*** **20, 313-9.**

Egeblad, M. and Werb, Z. (2002). New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2, 161-74.

Ema, M., Faloon, P., Zhang, W. J., Hirashima, M., Reid, T., Stanford, W. L., Orkin, S., Choi, K. and Rossant, J. (2003). Combinatorial effects of Flk1 and Tal1 on vascular and hematopoietic development in the mouse. *Genes Dev* 17, 380-93.

Estreicher, A., Muhlhauser, J., Carpentier, J. L., Orci, L. and Vassalli, J. D. (1990). The receptor for urokinase type plasminogen activator polarizes expression of the protease to the leading edge of migrating monocytes and promotes degradation of enzyme inhibitor complexes. *J Cell Biol* 111, 783-92.

Eter, N. and Spitznas, M. (2002). DMSO mimics inhibitory effect of thalidomide on choriocapillary endothelial cell proliferation in culture. *Br J Ophthalmol* 86, 1303-5.

Evans, M. J., Carlton, M. B. and Russ, A. P. (1997). Gene trapping and functional genomics. *Trends Genet* 13, 370-4.

Fang, J., Shing, Y., Wiederschain, D., Yan, L., Butterfield, C., Jackson, G., Harper, J., Tamvakopoulos, G. and Moses, M. A. (2000). Matrix metalloproteinase-2 is required for the switch to the angiogenic phenotype in a tumor model. *Proc Natl Acad Sci U S A* 97, 3884-9.

Favor, J. and Neuhauser-Klaus, A. (2000). Saturation mutagenesis for dominant eye morphological defects in the mouse Mus musculus. *Mamm Genome* 11, 520-5.

Folkman, J. (2002). Role of angiogenesis in tumor growth and metastasis. *Semin Oncol* 29, 15-8.

Folkman, J. and Klagsbrun, M. (1987). Angiogenic factors. *Science* 235, 442-7.

Forrester, L. M., Nagy, A., Sam, M., Watt, A., Stevenson, L., Bernstein, A., Joyner, A. L. and Wurst, W. (1996). An induction gene trap screen in embryonic stem cells: Identification of genes that respond to retinoic acid in vitro. *Proc Natl Acad Sci U S A* 93, 1677-82.

Friedrich, G. and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* 5, 1513-23.

Fujimoto, Y., Sasaki, T., Tsuchida, A. and Chayama, K. (2001). Angiotensin II type 1 receptor expression in human pancreatic cancer and growth inhibition by angiotensin II type 1 receptor antagonist. *FEBS Lett* 495, 197-200.

Fukasawa, K., Fujii, H., Saitoh, Y., Koizumi, K., Aozuka, Y., Sekine, K., Yamada, M., Saiki, I. and Nishikawa, K. (2006). Aminopeptidase N (APN/CD13) is selectively expressed in vascular endothelial cells and plays multiple roles in angiogenesis. *Cancer Lett.*

Garcia-Mata, R., Bebok, Z., Sorscher, E. J. and Sztul, E. S. (1999). Characterization and dynamics of aggresome formation by a cytosolic GFP-chimera. *J Cell Biol* 146, 1239-54.

Geballe, A. P. and Morris, D. R. (1994). Initiation codons within 5'-leaders of mRNAs as regulators of translation. *Trends Biochem Sci* 19, 159-64.

Gerety, S. S., Wang, H. U., Chen, Z. F. and Anderson, D. J. (1999). Symmetrical mutant phenotypes of the receptor EphB4 and its specific transmembrane ligand ephrin-B2 in cardiovascular development. *Mol Cell* 4, 403-14.

Good, D. J., Polverini, P. J., Rastinejad, F., Le Beau, M. M., Lemons, R. S., Frazier, W. A. and Bouck, N. P. (1990). A tumor suppressor-dependent inhibitor of angiogenesis is immunologically and functionally indistinguishable from a fragment of thrombospondin. *Proc Natl Acad Sci U S A* 87, 6624-8.

Gorlich, D. (1997). Nuclear protein import. *Curr Opin Cell Biol* 9, 412-9.

Gossler, A., Joyner, A. L., Rossant, J. and Skarnes, W. C. (1989). Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science* 244, 463-5.

Grant, S. K., Sklar, J. G. and Cummings, R. T. (2002). Development of novel assays for proteolytic enzymes using rhodamine-based fluorogenic substrates. *J Biomol Screen* 7, 531-40.

Graveley, B. R. (2001). Alternative splicing: increasing diversity in the proteomic world. *Trends Genet* 17, 100-7.

Greco, S., Muscella, A., Elia, M. G., Salvatore, P., Storelli, C., Mazzotta, A., Manca, C. and Marsigliante, S. (2003). Angiotensin II activates extracellular signal regulated kinases via protein kinase C and epidermal growth factor receptor in breast cancer cells. *J Cell Physiol* 196, 370-7.

Griffith, E. C., Su, Z., Niwayama, S., Ramsay, C. A., Chang, Y. H. and Liu, J. O. (1998). Molecular recognition of angiogenesis inhibitors fumagillin and ovalicin by methionine aminopeptidase 2. *Proc Natl Acad Sci U S A* 95, 15183-8.

Griffith, E. C., Su, Z., Turk, B. E., Chen, S., Chang, Y. H., Wu, Z., Biemann, K. and Liu, J. O. (1997). Methionine aminopeptidase (type 2) is the common target for angiogenesis inhibitors AGM-1470 and ovalicin. *Chem Biol* 4, 461-71.

Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-Czarny, N., Spillett, D., Muselet, D., Prud'homme, J. F., Dib, C., Auffray, C. et al. (1996). A radiation hybrid map of the human genome. *Hum Mol Genet* 5, 339-46.

Haeggstrom, J. Z. (2004). Leukotriene A4 hydrolase/aminopeptidase, the gatekeeper of chemotactic leukotriene B4 biosynthesis. *J Biol Chem* 279, 50639-42.

Hajitou, A., Sounni, N. E., Devy, L., Grignet-Debrus, C., Lewalle, J. M., Li, H., Deroanne, C. F., Lu, H., Colige, A., Nusgens, B. V. et al. (2001). Down-regulation of vascular endothelial growth factor by tissue inhibitor of metalloproteinase-2: effect on in vivo mammary tumor growth and angiogenesis. *Cancer Res* 61, 3450-7.

Hanahan, D. and Folkman, J. (1996). Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell* 86, 353-64.

Handwerger, K. E. and Gall, J. G. (2006). Subnuclear organelles: new insights into form and function. *Trends Cell Biol* 16, 19-26.

Harlow E, L. D. (1999). Using Antibodies a laboratory Manual: Cold Spring Harbor Laboratory Press.

Heber, S., Herms, J., Gajic, V., Hainfellner, J., Aguzzi, A., Rulicke, T., von Kretzschmar, H., von Koch, C., Sisodia, S., Tremml, P. et al. (2000). Mice with combined gene knock-outs reveal essential and partially redundant functions of amyloid precursor protein family members. *J Neurosci* 20, 7951-63.

Herrick, T. M. and Cooper, J. A. (2002). A hypomorphic allele of dab1 reveals regional differences in reelin-Dab1 signaling during brain development. *Development* 129, 787-96.

Hesp, J. R. and Hooper, N. M. (1997). Proteolytic fragmentation reveals the oligomeric and domain structure of porcine aminopeptidase A. *Biochemistry* 36, 3000-7.

Hill, D. P. and Wurst, W. (1993). Gene and enhancer trapping: mutagenic strategies for developmental studies. *Curr Top Dev Biol* 28, 181-206.

Hobson, B. and Denekamp, J. (1984). Endothelial proliferation in tumours and normal tissues: continuous labelling studies. *Br J Cancer* 49, 405-13.

Hol, E. M., van Leeuwen, F. W. and Fischer, D. F. (2005). The proteasome in Alzheimer's disease and Parkinson's disease: lessons from ubiquitin B+1. *Trends Mol Med* 11, 488-95.

Howard, L., Maciewicz, R. A. and Blobel, C. P. (2000). Cloning and characterization of ADAM28: evidence for autocatalytic pro-domain removal and for cell surface localization of mature ADAM28. *Biochem J* 348 Pt 1, 21-7.

Huber, A. H., Nelson, W. J. and Weis, W. I. (1997). Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* 90, 871-82.

Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J. and Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature* 432, 625-30.

Hughes, T. A. and Brady, H. J. (2005a). E2F1 up-regulates the expression of the tumour suppressor axin2 both by activation of transcription and by mRNA stabilisation. *Biochem Biophys Res Commun* 329, 1267-74.

Hughes, T. A. and Brady, H. J. (2005b). Expression of axin2 is regulated by the alternative 5'-untranslated regions of its mRNA. *J Biol Chem* 280, 8581-8.

Hunt, C. and Morimoto, R. I. (1985). Conserved features of eukaryotic hsp70 genes revealed by comparison with the nucleotide sequence of human hsp70. *Proc Natl Acad Sci U S A* 82, 6455-9.

Ikemura, H., Takagi, H. and Inouye, M. (1987). Requirement of pro-sequence for the production of active subtilisin E in Escherichia coli. *J Biol Chem* 262, 7859-64.

Imanishi, T., Hano, T. and Nishio, I. (2004). Angiotensin II potentiates vascular endothelial growth factor-induced proliferation and network formation of endothelial progenitor cells. *Hypertens Res* 27, 101-8.

Itoh, T., Ikeda, T., Gomi, H., Nakao, S., Suzuki, T. and Itohara, S. (1997). Unaltered secretion of beta-amyloid precursor protein in gelatinase A (matrix metalloproteinase 2)-deficient mice. *J Biol Chem* 272, 22389-92.

Iwase, A., Nomura, S. and Mizutani, S. (2001). Characterization of a secretase activity for placental leucine aminopeptidase. *Arch Biochem Biophys* 393, 163-9.

Jacob, R., Peters, K. and Naim, H. Y. (2002). The prosequence of human lactase-phlorizin hydrolase modulates the folding of the mature enzyme. *J Biol Chem* 277, 8217-25.

Joseph, R., Dou, D and Tsang, W. (1995). Neuronatin mRNA: alternatively spliced forms of a novel brain-specific mammalian developmental gene *Brain Research*, 92-98

Kalderon, D., Roberts, B. L., Richardson, W. D. and Smith, A. E. (1984). A short amino acid sequence able to specify nuclear location. *Cell* 39, 499-509.

Kalland, K. H., Szilvay, A. M., Brokstad, K. A., Saetrevik, W. and Haukenes, G. (1994). The human immunodeficiency virus type 1 Rev protein shuttles between the cytoplasm and nuclear compartments. *Mol Cell Biol* 14, 7436-44.

Kato, T., Kure, T., Chang, J. H., Gabison, E. E., Itoh, T., Itohara, S. and Azar, D. T. (2001). Diminished corneal angiogenesis in gelatinase A-deficient mice. *FEBS Lett* 508, 187-90.

Kessel, M. and Gruss, P. (1991). Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* 67, 89-104.

Kester, W. R. and Matthews, B. W. (1977). Crystallographic study of the binding of dipeptide inhibitors to thermolysin: implications for the mechanism of catalysis. *Biochemistry* 16, 2506-16.

Kim, S. and Iwao, H. (2000). Molecular and cellular mechanisms of angiotensin II-mediated cardiovascular and renal diseases. *Pharmacol Rev* 52, 11-34.

Klymkowsky, M. W. (1999). Plakophilin, armadillo repeats, and nuclear localization. *Microsc Res Tech* 45, 43-54.

Kondo, T., Takahashi, N. and Muramatsu, M. (1992). The regulation of the murine Hox-2.5 gene expression during cell differentiation. *Nucleic Acids Res* 20, 5729-35.

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299, 1-34.

Kudelka, A. P., Verschraegen, C. F. and Loyer, E. (1998). Complete remission of metastatic cervical cancer with the angiogenesis inhibitor TNP-470. *N Engl J Med* 338, 991-2.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-5.

Lakka, S. S., Gondi, C. S., Dinh, D. H., Olivero, W. C., Gujrati, M., Rao, V. H., Sioka, C. and Rao, J. S. (2005). Specific interference of urokinase-type plasminogen activator receptor and matrix metalloproteinase-9 gene expression induced by double-stranded RNA results in decreased invasion, tumor growth, and angiogenesis in gliomas. *J Biol Chem* 280, 21882-92.

Lamkanfi, M., Declercq, W., Vanden Berghe, T. and Vandenabeele, P. (2006). Caspases leave the beaten track: caspase-mediated activation of NF-kappaB. *J Cell Biol* 173, 165-71.

Lander, E. S. Linton, L. M. Birren, B. Nusbaum, C. Zody, M. C. Baldwin, J. Devon, K. Dewar, K. Doyle, M. FitzHugh, W. et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.

Law, D. J., Labut, E. M., Adams, R. D. and Merchant, J. L. (2006). An isoform of ZBP-89 predisposes the colon to colitis. *Nucleic Acids Res* 34, 1342-50.

Lee, H. J., Tomioka, M., Takaki, Y., Masumoto, H. and Saido, T. C. (2000). Molecular cloning and expression of aminopeptidase A isoforms from rat hippocampus. *Biochim Biophys Acta* 1493, 273-8.

Leung, A. K., Andersen, J. S., Mann, M. and Lamond, A. I. (2003). Bioinformatic analysis of the nucleolus. *Biochem J* 376, 553-69.

Lichun, Y., Ching Tang, C. M., Wai Lau, K. and Lung, M. L. (2004). Frequent loss of heterozygosity on chromosome 9 in Chinese esophageal squamous cell carcinomas. *Cancer Lett* 203, 71-7.

Lidington, E. A., Rao, R. M., Marelli-Berg, F. M., Jat, P. S., Haskard, D. O. and Mason, J. C. (2002). Conditional immortalization of growth factor-responsive cardiac endothelial cells from H-2K(b)-tsA58 mice. *Am J Physiol Cell Physiol* 282, C67-74.

Lijnen, H. R. (2001). Elements of the fibrinolytic system. *Ann N Y Acad Sci* 936, 226-36.

Limongi, P., Resnati, M., Hernandez-Marrero, L., Cremona, O., Blasi, F. and Fazioli, F. (1995). Biosynthesis and apical localization of the urokinase receptor in polarized MDCK epithelial cells. *FEBS Lett* 369, 207-11.

Lorenz, I. C., Marcotrigiano, J., Dentzer, T. G. and Rice, C. M. (2006). Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature* 442, 831-5.

Malik, H. S., Eickbush, T. H. and Goldfarb, D. S. (1997). Evolutionary specialization of the nuclear targeting apparatus. *Proc Natl Acad Sci U S A* 94, 13738-42.

Marchio, S., Lahdenranta, J., Schlingemann, R. O., Valdembri, D., Wesseling, P., Arap, M. A., Hajitou, A., Ozawa, M. G., Trepel, M., Giordano, R. J. et al. (2004). Aminopeptidase A is a functional target in angiogenic blood vessels. *Cancer Cell* 5, 151-62.

Matsumoto, H., Rogi, T., Yamashiro, K., Kodama, S., Tsuruoka, N., Hattori, A., Takio, K., Mizutani, S. and Tsujimoto, M. (2000). Characterization of a recombinant soluble form of human placental leucine aminopeptidase/oxytocinase expressed in Chinese hamster ovary cells. *Eur J Biochem* 267, 46-52.

Mazar, A. P., Henkin, J. and Goldfarb, R. H. (1999). The urokinase plasminogen activator system in cancer: implications for tumor angiogenesis and metastasis. *Angiogenesis* 3, 15-32.

Mazzieri, R., Masiero, L., Zanetta, L., Monea, S., Onisto, M., Garbisa, S. and Mignatti, P. (1997). Control of type IV collagenase activity by components of the urokinase-plasmin system: a regulatory mechanism with cell-bound reactants. *Embo J* 16, 2319-32.

Meyer, M., Clauss, M., Lepple-Wienhues, A., Waltenberger, J., Augustin, H. G., Ziche, M., Lanz, C., Buttner, M., Rziha, H. J. and Dehio, C. (1999). A novel vascular endothelial growth factor encoded by Orf virus, VEGF-E, mediates angiogenesis via signalling through VEGFR-2 (KDR) but not VEGFR-1 (Flt-1) receptor tyrosine kinases. *Embo J* 18, 363-74.

Middleton, J., Americh, L., Gayon, R., Julien, D., Aguilar, L., Amalric, F. and Girard, J. P. (2004). Endothelial cell phenotypes in the rheumatoid synovium: activated, angiogenic, apoptotic and leaky. *Arthritis Res Ther* 6, 60-72.

Millauer, B., Wizigmann-Voos, S., Schnurch, H., Martinez, R., Moller, N. P., Risau, W. and Ullrich, A. (1993). High affinity VEGF binding and developmental expression suggest Flk-1 as a major regulator of vasculogenesis and angiogenesis. *Cell* 72, 835-46.

Millichip, M. I., Dallas, D. J., Wu, E., Dale, S. and McKie, N. (1998). The metallo-disintegrin ADAM10 (MADM) from bovine kidney has type IV collagenase activity in vitro. *Biochem Biophys Res Commun* 245, 594-8.

Mitsui, T., Nomura, S., Itakura, A. and Mizutani, S. (2004). Role of aminopeptidases in the blood pressure regulation. *Biol Pharm Bull* 27, 768-71.

Mitsui, T., Nomura, S., Okada, M., Ohno, Y., Kobayashi, H., Nakashima, Y., Murata, Y., Takeuchi, M., Kuno, N., Nagasaka, T. et al. (2003). Hypertension and angiotensin II hypersensitivity in aminopeptidase A-deficient mice. *Mol Med* 9, 57-62.

Miyajima, A., Kosaka, T., Asano, T., Asano, T., Seta, K., Kawai, T. and Hayakawa, M. (2002). Angiotensin II type I antagonist prevents pulmonary metastasis of murine renal cancer by inhibiting tumor angiogenesis. *Cancer Res* 62, 4176-9.

Miyashita, H., Yamazaki, T., Akada, T., Niizeki, O., Ogawa, M., Nishikawa, S. and Sato, Y. (2002). A mouse orthologue of puromycin-insensitive leucyl-specific aminopeptidase is expressed in endothelial cells and plays an important role in angiogenesis. *Blood* 99, 3241-9.

Mizutani, S., Goto, K., Mizuno, K., Itakura, A., Kurauchi, O., Kikkawa, F. and Tomoda, Y. (1995). Interaction between pregnancy-induced bioactive peptides and the placental proteases. *Reprod Fertil Dev* 7, 1431-6.

Mizutani, S. and Tomoda, Y. (1996). Effects of placental proteases on maternal and fetal blood pressure in normal pregnancy and preeclampsia. *Am J Hypertens* 9, 591-7.

Mizutani, S., Yamada, R., Kurauchi, O., Ito, Y., Narita, O. and Tomoda, Y. (1987). Serum aminopeptidase A (AAP) in normal pregnancy and pregnancy complicated by pre-eclampsia. *Arch Gynecol* 240, 27-31.

Modrek, B., Lee,C. (2002). A genomic view of alternative splicing *Nature Genetics* 30, 13-19

Morriss-Kay, G. M., Murphy, P., Hill, R. E. and Davidson, D. R. (1991). Effects of retinoic acid excess on expression of Hox-2.9 and Krox-20 and on morphological segmentation in the hindbrain of mouse embryos. *Embo J* 10, 2985-95.

Moses, M. A. (1997). The regulation of neovascularization of matrix metalloproteinases and their inhibitors. *Stem Cells* 15, 180-9.

Moss, T. and Stefanovsky, V. Y. (2002). At the center of eukaryotic life. *Cell* 109, 545-8.

Neurath, H. (1999). Proteolytic enzymes, past and future. *Proc Natl Acad Sci U S A* 96, 10962-3.

Newton, K., Petfalski, E., Tollervey, D. and Caceres, J. F. (2003). Fibrillarin is essential for early development and required for accumulation of an intron-encoded small nucleolar RNA in the mouse. *Mol Cell Biol* 23, 8519-27.

Nicosia, R. F. (1998). What is the role of vascular endothelial growth factor-related molecules in tumor angiogenesis? *Am J Pathol* 153, 11-6.

Nicosia, R. F. and Ottinetti, A. (1990a). Growth of microvessels in serum-free matrix culture of rat aorta. A quantitative assay of angiogenesis in vitro. *Lab Invest* 63, 115-22.

Nicosia, R. F. and Ottinetti, A. (1990b). Modulation of microvascular growth and morphogenesis by reconstituted basement membrane gel in three-dimensional cultures of rat aorta: a comparative study of angiogenesis in matrigel, collagen, fibrin, and plasma clot. *In Vitro Cell Dev Biol* 26, 119-28.

Nosaka, T., Siomi, H., Adachi, Y., Ishibashi, M., Kubota, S., Maki, M. and Hatanaka, M. (1989). Nucleolar targeting signal of human T-cell leukemia virus type I rex-encoded protein is essential for cytoplasmic accumulation of unspliced viral mRNA. *Proc Natl Acad Sci U S A* 86, 9798-802.

Nouet, S. and Nahmias, C. (2000). Signal transduction from the angiotensin II AT2 receptor. *Trends Endocrinol Metab* 11, 1-6.

Nowak, M. A., Boerlijst, M. C., Cooke, J. and Smith, J. M. (1997). Evolution of genetic redundancy. *Nature* 388, 167-71.

O'Reilly, M. S., Holmgren, L., Shing, Y., Chen, C., Rosenthal, R. A., Cao, Y., Moses, M., Lane, W. S., Sage, E. H. and Folkman, J. (1994a). Angiostatin: a circulating endothelial cell inhibitor that suppresses angiogenesis and tumor growth. *Cold Spring Harb Symp Quant Biol* 59, 471-82.

O'Reilly, M. S., Holmgren, L., Shing, Y., Chen, C., Rosenthal, R. A., Moses, M., Lane, W. S., Cao, Y., Sage, E. H. and Folkman, J. (1994b). Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell* 79, 315-28.

Obermann, E. C., Meyer, S., Hellge, D., Zaak, D., Filbeck, T., Stoehr, R., Hofstaedter, F., Hartmann, A. and Knuechel, R. (2004). Fluorescence in situ hybridization detects frequent chromosome 9 deletions and aneuploidy in histologically normal urothelium of bladder cancer patients. *Oncol Rep* 11, 745-51.

Ofner, L. D. and Hooper, N. M. (2002a). The C-terminal domain, but not the interchain disulphide, is required for the activity and intracellular trafficking of aminopeptidase A. *Biochem J* 362, 191-7.

Ofner, L. D. and Hooper, N. M. (2002b). Ectodomain shedding of cystinyl aminopeptidase from human placental membranes. *Placenta* 23, 65-70.

Olson, M. O., Hingorani, K. and Szebeni, A. (2002). Conventional and nonconventional roles of the nucleolus. *Int Rev Cytol* 219, 199-266.

Orning, L., Gierse, J. K. and Fitzpatrick, F. A. (1994). The bifunctional enzyme leukotriene-A4 hydrolase is an arginine aminopeptidase of high efficiency and specificity. *J Biol Chem* 269, 11269-73.

Otani, A., Takagi, H., Oh, H., Koyama, S. and Honda, Y. (2001). Angiotensin II induces expression of the Tie2 receptor ligand, angiopoietin-2, in bovine retinal endothelial cells. *Diabetes* 50, 867-75.

Oudit, G. Y., Crackower, M. A., Backx, P. H. and Penninger, J. M. (2003). The role of ACE2 in cardiovascular physiology. *Trends Cardiovasc Med* 13, 93-101.

Page, N. M. (2002). The endocrinology of pre-eclampsia. *Clin Endocrinol (Oxf)* 57, 413-23.

Pall, G. S., Wallis, J., Axton, R., Brownstein, D. G., Gautier, P., Buerger, K., Mulford, C., Mullins, J. J. and Forrester, L. M. (2004). A novel transmembrane MSP-containing protein that plays a role in right ventricle development. *Genomics* 84, 1051-9.

Pasqualini, R., Koivunen, E., Kain, R., Lahdenranta, J., Sakamoto, M., Stryhn, A., Ashmun, R. A., Shapiro, L. H., Arap, W. and Ruoslahti, E. (2000). Aminopeptidase N is a receptor for tumor-homing peptides and a target for inhibiting angiogenesis. *Cancer Res* 60, 722-7.

Pawlikowski, M. and Kunert-Radek, J. (1997). Angiotensin IV stimulates the proliferation of rat anterior pituitary cells in vitro. *Biochem Biophys Res Commun* 232, 292-3.

Pederson, T. (1998). The plurifunctional nucleolus. *Nucleic Acids Res* 26, 3871-6.

Peifer, M., Orsulic, S., Sweeton, D. and Wieschaus, E. (1993). A role for the Drosophila segment polarity gene armadillo in cell adhesion and cytoskeletal integrity during oogenesis. *Development* 118, 1191-207.

Pickart, C. M. (2004). Back to the future with ubiquitin. *Cell* 116, 181-90.

Prince, V. E. and Pickett, F. B. (2002). Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet* 3, 827-37.

Puente, X. S., Sanchez, L. M., Overall, C. M. and Lopez-Otin, C. (2003). Human and mouse proteases: a comparative genomic approach. *Nat Rev Genet* 4, 544-58.

Quwailid, M. M., Hugill, A., Dear, N., Vizor, L., Wells, S., Horner, E., Fuller, S., Weedon, J., McMath, H., Woodman, P. et al. (2004). A gene-driven ENU-based approach to generating an allelic series in any gene. *Mamm Genome* 15, 585-91.

Ragno, P. (2006). The urokinase receptor: a ligand or a receptor? Story of a sociable molecule. *Cell Mol Life Sci* 63, 1028-37.

Rapaport, D. (2003). Finding the right organelle. Targeting signals in mitochondrial outer-membrane proteins. *EMBO Rep* 4, 948-52.

Rawlings, N. D., O'Brien, E. and Barrett, A. J. (2002). MEROPS: the protease database. *Nucleic Acids Res* 30, 343-6.

Reaux, A., Fournie-Zaluski, M. C., David, C., Zini, S., Roques, B. P., Corvol, P. and Llorens-Cortes, C. (1999). Aminopeptidase A inhibitors as potential central antihypertensive agents. *Proc Natl Acad Sci U S A* 96, 13415-20.

Reinartz, J., Schafer, B., Batrla, R., Klein, C. E. and Kramer, M. D. (1995). Plasmin abrogates alpha v beta 5-mediated adhesion of a human keratinocyte cell line (HaCaT) to vitronectin. *Exp Cell Res* 220, 274-82.

Roberts, C. M., Tani, P. H., Bridges, L. C., Laszik, Z. and Bowditch, R. D. (1999). MDC-L, a novel metalloprotease disintegrin cysteine-rich protein family member expressed by human lymphocytes. *J Biol Chem* 274, 29251-9.

Robinson, C. J. and Stringer, S. E. (2001). The splice variants of vascular endothelial growth factor (VEGF) and their receptors. *J Cell Sci* 114, 853-65.

Rogi, T., Tsujimoto, M., Nakazato, H., Mizutani, S. and Tomoda, Y. (1996). Human placental leucine aminopeptidase/oxytocinase. A new member of type II membrane-spanning zinc metallopeptidase family. *J Biol Chem* 271, 56-61.

Rogozin, I. B., Kochetov, A. V., Kondrashov, F. A., Koonin, E. V. and Milanesi, L. (2001). Presence of ATG triplets in 5' untranslated regions of eukaryotic cDNAs correlates with a 'weak' context of the start codon. *Bioinformatics* 17, 890-900.

Rosenthal, R. P. H. N. (1999). Heart Development: Academic Press.

Roy, R., Zhang, B. and Moses, M. A. (2006). Making the cut: protease-mediated regulation of angiogenesis. *Exp Cell Res* 312, 608-22.

Rozenfeld, R., Muller, L., Messari, S. E. and Llorens-Cortes, C. (2004). The C-terminal domain of aminopeptidase A is an intramolecular chaperone required for the correct folding, cell surface expression, and activity of this monozinc aminopeptidase. *J Biol Chem* 279, 43285-95.

Rudberg, P. C., Tholander, F., Andberg, M., Thunnissen, M. M. and Haeggstrom, J. Z. (2004). Leukotriene A4 hydrolase: identification of a common carboxylate recognition site for the epoxide hydrolase and aminopeptidase substrates. *J Biol Chem* 279, 27376-82.

Russell, W. L., Kelly, E. M., Hunsicker, P. R., Bangham, J. W., Maddux, S. C. and Phipps, E. L. (1979). Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. ***Proc Natl Acad Sci U S A*** **76, 5818-9.**

Sam, M., Wurst, W., Kluppel, M., Jin, O., Heng, H. and Bernstein, A. (1998). Aquarius, a novel gene isolated by gene trapping with an RNA-dependent RNA polymerase motif. ***Dev Dyn*** **212, 304-17.**

Sambrook , J., Fritsch, E.F., Maniatis.T. (1989). Molecular Cloning A Laboratory Manual

Satchi-Fainaro, R., Puder, M., Davies, J. W., Tran, H. T., Sampson, D. A., Greene, A. K., Corfas, G. and Folkman, J. (2004). Targeting angiogenesis with a conjugate of HPMA copolymer and TNP-470. ***Nat Med*** **10, 255-61.**

Sato, Y. (2003). Aminopeptidases and angiogenesis. ***Endothelium*** **10, 287-90.**

Sato, Y. (2004). Role of aminopeptidase in angiogenesis. ***Biol Pharm Bull*** **27, 772-6.**

Scherl, A., Coute, Y., Deon, C., Calle, A., Kindbeiter, K., Sanchez, J. C., Greco, A., Hochstrasser, D. and Diaz, J. J. (2002). Functional proteomic analysis of human nucleolus. ***Mol Biol Cell*** **13, 4100-9.**

Schomburg, L., Kollmus, H., Friedrichsen, S. and Bauer, K. (2000). Molecular characterization of a puromycin-insensitive leucyl-specific aminopeptidase, PILS-AP. ***Eur J Biochem*** **267, 3198-207.**

Seals, D. F. and Courtneidge, S. A. (2003). The ADAMs family of metalloproteases: multidomain proteins with multiple functions. ***Genes Dev*** **17, 7-30.**

Shastry, B. S. (1995). Genetic knockouts in mice: an update. ***Experientia*** **51, 1028-39.**

Shi, X. and Jarvis, D. L. (2006). A new rapid amplification of cDNA ends method for extremely guanine plus cytosine-rich genes. ***Anal Biochem.***

Shibuya, M. (2006). Vascular endothelial growth factor (VEGF)-Receptor2: its biological functions, major signaling pathway, and specific ligand VEGF-E. ***Endothelium*** **13, 63-9.**

Silen, J. L., Frank, D., Fujishige, A., Bone, R. and Agard, D. A. (1989). Analysis of prepro-alpha-lytic protease expression in Escherichia coli reveals that the pro region is required for activity. ***J Bacteriol*** **171, 1320-5.**

Simoneau, M., LaRue, H., Aboulkassim, T. O., Meyer, F., Moore, L. and Fradet, Y. (2000). Chromosome 9 deletions and recurrence of superficial bladder cancer: identification of four regions of prognostic interest. *Oncogene* 19, 6317-23.

Skarnes, W. C. (2005). Two ways to trap a gene in mice. *Proc Natl Acad Sci U S A* 102, 13001-2.

Skarnes, W. C., Auerbach, B. A. and Joyner, A. L. (1992). A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. *Genes Dev* 6, 903-18.

Skarnes, W. C., von Melchner, H., Wurst, W., Hicks, G., Nord, A. S., Cox, T., Young, S. G., Ruiz, P., Soriano, P., Tessier-Lavigne, M. et al. (2004). A public gene trap resource for mouse functional genomics. *Nat Genet* 36, 543-4.

Sorek, R., Shamir, R. and Ast, G. (2004). How prevalent is functional alternative splicing in the human genome? *Trends Genet* 20, 68-71.

Springman, E. B., Angleton, E. L., Birkedal-Hansen, H. and Van Wart, H. E. (1990). Multiple modes of activation of latent human fibroblast collagenase: evidence for the role of a Cys73 active-site zinc complex in latency and a "cysteine switch" mechanism for activation. *Proc Natl Acad Sci U S A* 87, 364-8.

Stanford, W. L., Caruana, G., Vallis, K. A., Inamdar, M., Hidaka, M., Bautch, V. L. and Bernstein, A. (1998). Expression trapping: identification of novel genes expressed in hematopoietic and endothelial lineages by gene trapping in ES cells. *Blood* 92, 4622-31.

Stanford, W. L., Cohn, J. B. and Cordes, S. P. (2001). Gene-trap mutagenesis: past, present and beyond. *Nat Rev Genet* 2, 756-68.

Summerton, J., Stein, D., Huang, S. B., Matthews, P., Weller, D. and Partridge, M. (1997). Morpholino and phosphorothioate antisense oligomers compared in cell-free and in-cell systems. *Antisense Nucleic Acid Drug Dev* 7, 63-70.

Suri, C., Jones, P. F., Patan, S., Bartunkova, S., Maisonpierre, P. C., Davis, S., Sato, T. N. and Yancopoulos, G. D. (1996). Requisite role of angiopoietin-1, a ligand for the TIE2 receptor, during embryonic angiogenesis. *Cell* 87, 1171-80.

Suto, K., Yamazaki, Y., Morita, T. and Mizuno, H. (2005). Crystal structures of novel vascular endothelial growth factors (VEGF) from snake venoms: insight into selective VEGF binding to kinase insert domain-containing receptor but not to fms-like tyrosine kinase-1. *J Biol Chem* 280, 2126-31.

Takahashi, C., Sheng, Z., Horan, T. P., Kitayama, H., Maki, M., Hitomi, K., Kitaura, Y., Takai, S., Sasahara, R. M., Horimoto, A. et al. (1998). Regulation of matrix metalloproteinase-9 and inhibition of tumor invasion by the membrane-anchored glycoprotein RECK. ***Proc Natl Acad Sci U S A*** **95, 13221-6.**

Tarui, T., Miles, L. A. and Takada, Y. (2001). Specific interaction of angiostatin with integrin alpha(v)beta(3) in endothelial cells. ***J Biol Chem*** **276, 39562-8.**

Taylor, A. (1993). Aminopeptidases: structure and function. ***Faseb J*** **7, 290-8.**

Testa, G., Zhang, Y., Vintersten, K., Benes, V., Pijnappel, W. W., Chambers, I., Smith, A. J., Smith, A. G. and Stewart, A. F. (2003). Engineering the mouse genome with bacterial artificial chromosomes to create multipurpose alleles. ***Nat Biotechnol*** **21, 443-7.**

Thomas, K. R. and Capecchi, M. R. (1986). Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene. ***Nature*** **324, 34-8.**

Thomas, K. R. and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. ***Cell*** **51, 503-12.**

Thunnissen, M. M., Nordlund, P. and Haeggstrom, J. Z. (2001). Crystal structure of human leukotriene A(4) hydrolase, a bifunctional enzyme in inflammation. ***Nat Struct Biol*** **8, 131-5.**

Torres, M., Stoykova, A., Huber, O., Chowdhury, K., Bonaldo, P., Mansouri, A., Butz, S., Kemler, R. and Gruss, P. (1997). An alpha-E-catenin gene trap mutation defines its function in preimplantation development. ***Proc Natl Acad Sci U S A*** **94, 901-6.**

Townley, D. J., Avery, B. J., Rosen, B. and Skarnes, W. C. (1997). Rapid sequence analysis of gene trap integrations to generate a resource of insertional mutations in mice. ***Genome Res*** **7, 293-8.**

Tsai, R. Y. and McKay, R. D. (2002). A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells. ***Genes Dev*** **16, 2991-3003.**

Tsujimoto, M., Mizutani, S., Adachi, H., Kimura, M., Nakazato, H. and Tomoda, Y. (1992). Identification of human placental leucine aminopeptidase as oxytocinase. ***Arch Biochem Biophys*** **292, 388-92.**

Tuteja, R. (2005). Type I signal peptidase: an overview. ***Arch Biochem Biophys*** **441, 107-11.**

Uemura, H., Ishiguro, H., Nakaigawa, N., Nagashima, Y., Miyoshi, Y., Fujinami, K., Sakaguchi, A. and Kubota, Y. (2003). Angiotensin II receptor blocker shows antiproliferative activity in prostate cancer cells: a possibility of tyrosine kinase inhibitor of growth factor. *Mol Cancer Ther* 2, 1139-47.

Venter, J. C. Adams, M. D. Myers, E. W. Li, P. W. Mural, R. J. Sutton, G. G. Smith, H. O. Yandell, M. Evans, C. A. Holt, R. A. et al. (2001). The sequence of the human genome. *Science* 291, 1304-51.

von Melchner, H., DeGregori, J. V., Rayburn, H., Reddy, S., Friedel, C. and Ruley, H. E. (1992). Selective disruption of genes expressed in totipotent embryonal stem cells. *Genes Dev* 6, 919-27.

Voss, A. K., Thomas, T. and Gruss, P. (1998). Compensation for a gene trap mutation in the murine microtubule-associated protein 4 locus by alternative polyadenylation and alternative splicing. *Dev Dyn* 212, 258-66.

Vu, T. H., Shipley, J. M., Bergers, G., Berger, J. E., Helms, J. A., Hanahan, D., Shapiro, S. D., Senior, R. M. and Werb, Z. (1998). MMP-9/gelatinase B is a key regulator of growth plate angiogenesis and apoptosis of hypertrophic chondrocytes. *Cell* 93, 411-22.

Wei, Y., Lukashev, M., Simon, D. I., Bodary, S. C., Rosenberg, S., Doyle, M. V. and Chapman, H. A. (1996). Regulation of integrin function by the urokinase receptor. *Science* 273, 1551-5.

Wen, W., Meinkoth, J. L., Tsien, R. Y. and Taylor, S. S. (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell* 82, 463-73.

Wente, S. R. (2000). Gatekeepers of the nucleus. *Science* 288, 1374-7.

Williams, C. L. (2003). The polybasic region of Ras and Rho family small GTPases: a regulator of protein interactions and membrane association and a site of nuclear localization signal sequences. *Cell Signal* 15, 1071-80.

Wobus, A. M., Kaomei, G., Shan, J., Wellner, M. C., Rohwedel, J., Ji, G., Fleischmann, B., Katus, H. A., Hescheler, J. and Franz, W. M. (1997). Retinoic acid accelerates embryonic stem cell-derived cardiac differentiation and enhances development of ventricular cardiomyocytes. *J Mol Cell Cardiol* 29, 1525-39.

Wong, E. A. and Capecchi, M. R. (1987). Homologous recombination between coinjected DNA sequences peaks in early to mid-S phase. *Mol Cell Biol* 7, 2294-5.

Wright, J. W., Mizutani, S., Murray, C. E., Amir, H. Z. and Harding, J. W. (1990). Aminopeptidase-induced elevations and reductions in blood pressure in the spontaneously hypertensive rat. *J Hypertens* 8, 969-74.

Xu, Q., Modrek, B. and Lee, C. (2002). Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic Acids Res* 30, 3754-66.

Yamazaki, T., Akada, T., Niizeki, O., Suzuki, T., Miyashita, H. and Sato, Y. (2004). Puromycin-insensitive leucyl-specific aminopeptidase (PILSAP) binds and catalyzes PDK1, allowing VEGF-stimulated activation of S6K for endothelial cell proliferation and angiogenesis. *Blood* 104, 2345-52.

Yan, D., Swain, P. K., Breuer, D., Tucker, R. M., Wu, W., Fujita, R., Rehemtulla, A., Burke, D. and Swaroop, A. (1998). Biochemical characterization and subcellular localization of the mouse retinitis pigmentosa GTPase regulator (mRpgr). *J Biol Chem* 273, 19656-63.

Yancopoulos, G. D., Davis, S., Gale, N. W., Rudge, J. S., Wiegand, S. J. and Holash, J. (2000). Vascular-specific growth factors and blood vessel formation. *Nature* 407, 242-8.

Yu, Y. and Bradley, A. (2001). Engineering chromosomal rearrangements in mice. *Nat Rev Genet* 2, 780-90.

Zhou, H. M., Weskamp, G., Chesneau, V., Sahin, U., Vortkamp, A., Horiuchi, K., Chiusaroli, R., Hahn, R., Wilkes, D., Fisher, P. et al. (2004). Essential role for ADAM19 in cardiovascular morphogenesis. *Mol Cell Biol* 24, 96-104.

Zhu, X. L., Ohta, Y., Jordan, F. and Inouye, M. (1989). Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. *Nature* 339, 483-4.

Development of a PCR genotyping assay

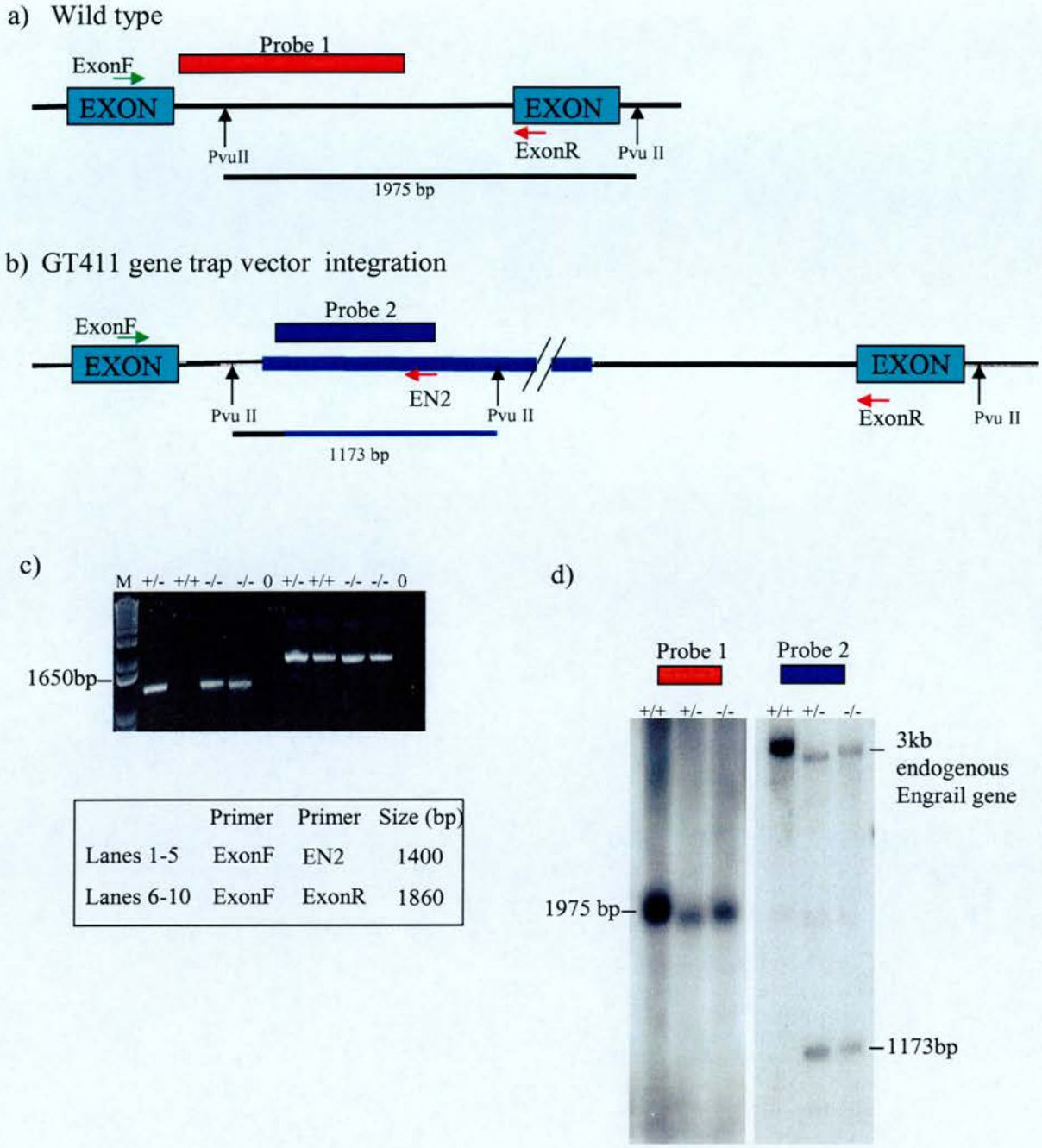
Genomic DNA was extracted from adult tissues of mice that had previously been genotyped with the β-galactosidase staining protocol (see methods). The wild type allele was amplified by PCR using primers Exon F and Exon R (Figure Appendix 1.1a). The gene trap (mutant) allele was amplified by PCR with the primers Exon F and EN2 (Figure Appendix 1.1b). As expected the mutant PCR yielded products for the heterozygote $ApO^{Gt411For/+}$ and homozygous $ApO^{Gt411For/Gt411For}$ DNA, but not for the wild type DNA (Figure Appendix 1.1c). Surprisingly, however, the wild type PCR primers (Exon F-Exon R) yielded product for all the different genotypes (Figure Appendix 1.1c). The expected result would have been that homozygous $ApO^{Gt411For/Gt411For}$ mice would not have yielded a PCR product. The experiment was repeated several times on other homozygous βgal genotyped mice and on the same homozygous $ApO^{Gt411For/Gt411For}$ mouse used in the knockout Northern analysis (Chapter 4). The result was the same despite extreme care being taken to eliminate the possibility of cross contamination during the purification of the DNA.

A genomic Southern was performed on PVU II restriction digested genomic DNA from wild type, heterozygous and homozygous mice. Using probe 1(specific to wild type sequence) a band should only be seen in the wild type and heterozygous lanes at 1975 bp. However, there is clearly a band in the homozygote the same size (Figure Appendix 1.1d, probe1). Using probe 2 (specific to the gene trap vector) bands should be seen in lanes 2 and 3 at 1173 bp. This was the case with this probe (Figure 1.1d, probe 2).

The experiment to design a PCR based genotyping assay was unsuccessful. The PCR primers ExonF- ExonR should have enabled one to determine which samples were homozygous $ApO^{Gt411For/Gt411For}$ and wild type. There was always a PCR fragment in the homozygous $ApO^{Gt411For/Gt411For}$ sample. To test if there was a de novo genomic duplication of this gene region else where in the genome FISH analysis was carried out on wild type metaphase chromosomes. This result indicated that there was only a single

copy on chromosome 13 (Figure 3.7). However, there is still the possibility that there is duplication within the region itself. Indeed thorough sequence analysis of chromosome 22 has reported that as much as 10% of this chromosome is duplicated (Bailey et al., 2002).

Another explanation might be that the wild type genomic sequence has undergone as a consequence of the gene trap integration a duplication event. It certainly is not uncommon for the gene trap vector to integrate multiple times (Pall et al., 2004). The Southern blot analysis with probe 1 supports this hypothesis because a fragment of 1975 bp is observed in the homozygous sample. Another experiment that could have been carried out was to perform FISH analysis on metaphase spreads of both wild type and homozygous gene trap GT411. A comparison of the wild type and homozygous metaphase chromosomes might have allowed a duplication event in the homozygous DNA to be seen. The data suggests the possibility that the endogenous gene locus may have become duplicated during the integration event.



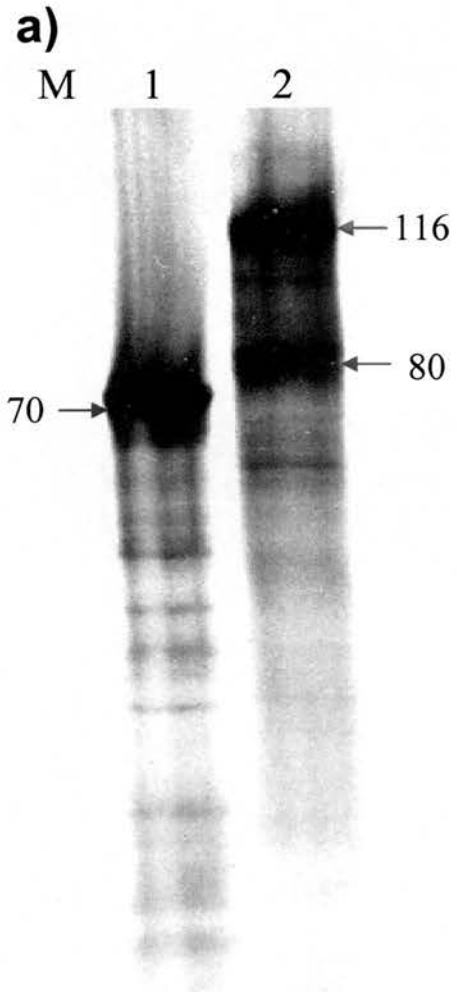
	Primer	Primer	Size (bp)
Lanes 1-5	ExonF	EN2	1400
Lanes 6-10	ExonF	ExonR	1860

Figure Appendix 1.1: Development of genotyping assay for the GT411 gene trap mice.

a) Schematic diagram of the wild type genomic region wild type,
b) gene trap allele. Gene trap vector shown in blue.
c) Genotyping PCR on genomic DNA extracted from wild type, heterozygous *ApO* $^{Gt411For/+}$ and homozygous*Apo* $^{Gt411For/Gt411For}$ samples.
d) Southern blot of wild type, heterozygous *ApO* $^{Gt411For/+}$ and homozygous *ApO* $^{Gt411For/GT411For}$ genomic DNA with restriction enzyme Pvu II and hybridised with probe 1 and probe 2. Note endogenous engrail at 3kb (Pall 1998)

LZ1	ANTISENSE	TAATGGGATAGGTTACG
LZ2	ANTISENSE	AGTATCGGCCTCAGGAAGATCG
EN2	ANTISENSE	TGCTCTGTCAGGTACCTGTTG
ANCHOR PRIMER 1		GGTTGTGAGCTCTTCTAGATGGTTTTTTTTTTTTTTTT
ANCHOR PRIMER 2		GGTTGTGAGCTCTTCTAGATGG
P2	ANTISENSE	CATTTCGCCACCTCGGAGCCTACTTG
P3	ANTISENSE	CTGAGGGCTCAGCGTCTTCTGCTCTA
P9	SENSEG	CAAGTAGGCTCCGAGGTGGCG
P11	ANTISENSE	TGTCCATGTGTTCCTTCACCAGC
P14	SENSE	CAGGTGTTAAGACCCAACAAAGAA
P19	ANTISENSE	TGTGACTGCACTGATGGCTCAGTC
P20	ANTISENSE	CTTCTTTGTTGGGTCTTAACACCTGCATTC
P21	ANTISENSE	GCAGCTGCTGTGCCTCAGCCCAAAATA
P23	SENSE	GGCAGCTGATTCTTTCCCAGGAT
P25	SENSE	AACCCGGAGTCTGCGGAAGTGTAGC
P26	SENSE	TAACCTGCAGCTCCGGTGTCAAGCAG
P28	ANTISENSE	TGTCCAGCCGGGAGAAAGGGAGAGTTC
P36	ANTISENSE	CAGCGGACGCACTCTGGAACCTGCA
P38	SENSE	TGCCTCTCAGCTGCACACTCTGTCCTG
P54	SENSE	GGACAGAGGAGTGGCTCAGTGAAGGA
P55	SENSE	TATTTTGGGCTGAGGCACAGCAGCTG
ECR	ANTISENSE	TTCCTGCAGTCTCAGAGCATGG
PECORI	SENSE	TTCCCCGAATTCCCACCTATGGACATAAAGCTGGACCCTTC
PNOT1	ANTISENSE	GGCCAATTGCGGCCGCTTAGAACAGCATTTCAGTCACCACCTG
HUGT411/5	SENSE	GACAATGTAGGCCCTGAGGAATTAC
HUGT411/6	ANTISENSE	ATACTCTGATTAGAAGCATAAGC
Fib 1	EcoRI	GAATTCATGAAGCCAGGTTTCAGCCCCCGTG
Fib 2	Asp718I	GGTACCTCAGTTCTTCACCTTGGGAGGTG

Appendix 1.2 Table 1: Primers used in this study.



b)

List of vectors used In study	
pGEX4T-1	GST tag
pCDNA3.1+	
pET32A	His tag
eGFP C2	GFP tag

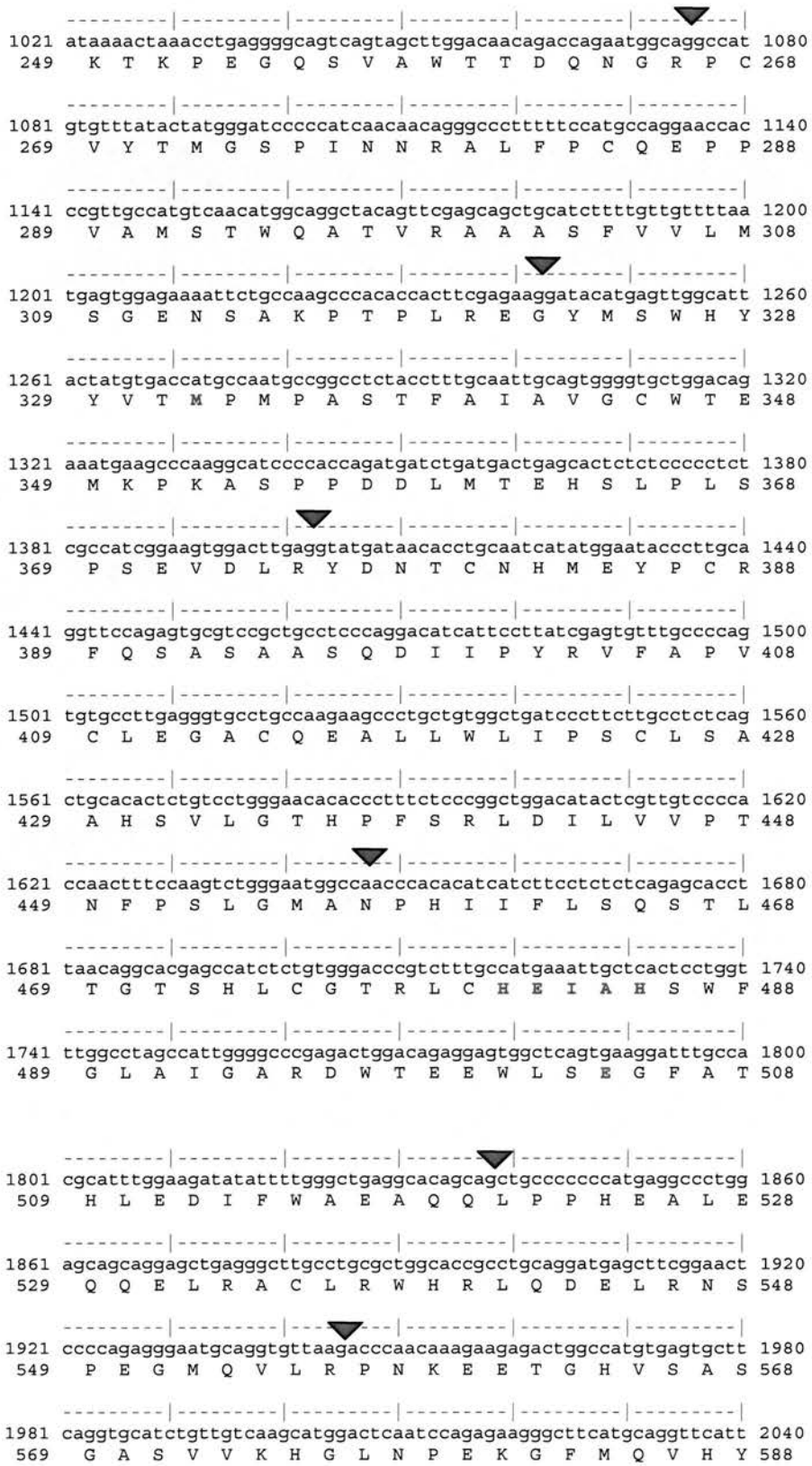
Appendix 1.3 *In vitro* translation of Full length *ApO* and vectors used for cloning

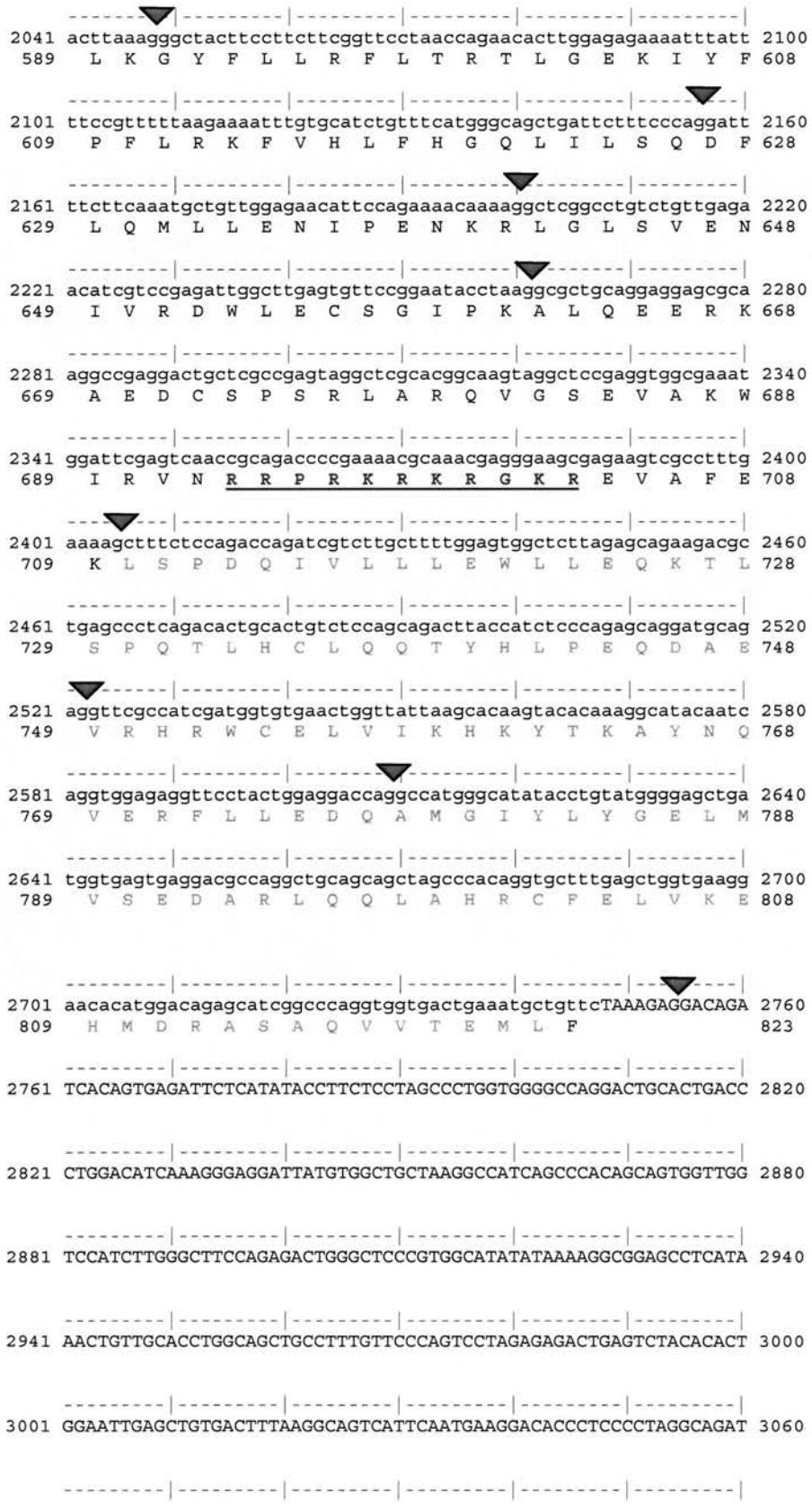
a) Lane 1, Luciferase control (70kDa), Lane 2, Fl-*ApO* (1-stop) cloned into pcDNA3.1+ (116kDa)
and breakdown product (80kDa).
b) List of vectors used in study

APPENDIX 1.4

PRETTYSEQ of cutpaste-9746 from 1 to 3856. BUILD 16/06/05

```
    ---------|---------|---------|---------|---------|---------|
  1 GAATTCGCCCTTAACCCGGAGTCTGCGGAAGTGTAGCTTGCAGCAAGCACCGCAGCTGAG 60

    ---------|---------|---------|---------|---------|---------|
 61 ACAGGTAACCTGCAGCTCCGGTGTCAAGCAGCCCGCTCCGGGTTATGCGGACCCACGACC 120

    ---------|---------|---------|---------|---------|---------|
121 TCTGGTGGGGTGAGCAGGGAATTGAAAGTTGTGGAAAGATAACGAGTGCCAGTGGGTGCT 180

    ---------|---------|---------|---------|---------|---------|
181 TCTTATTTGGTAAACCAGCCATTTGTGATTCTGAGCTACTACAGGAGATTAGAAGCTTCA 240

    ---------|---------|---------|---------|---------|---------|
241 GCTGAGATTCAGCAAAGCTACCAACTGCAGACTACACCatggacataaagctggaccctt 300
  1                                       M  D  I  K  L  D  P  S 8

    ---------|---------|---------|---------|---------|---------|
301 cccgggatgatctgcctctcatggccaacaccagccacatgcttgtgaagcactacatac 360
  9   R  D  D  L  P  L  M  A  N  T  S  H  M  L  V  K  H  Y  I  L 28

    ---------|---------|---------|---------|---------|---------|
361 tggatttggatgtggactttggaaatcaagtcattgagggcaacatagtgcttttctttg 420
 29   D  L  D  V  D  F  G  N  Q  V  I  E  G  N  I  V  L  F  F  G 48

    ---------|---------|---------|---------|---------|---------|
421 gagatggaaacagatttaaaaaccagtcgcgttctacccaggaaaccttccagatggagt 480
 49   D  G  N  R  F  K  N  Q  S  R  S  T  Q  E  T  F  Q  M  E  S 68

    ---------|---------|---------|---------|---------|---------|
481 cagaggaggctgacatatttaggacagctgaaccctgccatgttcctgagatggattcga 540
 69   E  E  A  D  I  F  R  T  A  E  P  C  H  V  P  E  M  D  S  S 88

    ---------|---------|---------|---------|---------|---------|
541 gtaccttctcacttaagatgggacatagagaatgtgcagtctgtggtaaaggtgatcaag 600
 89   T  F  S  L  K  M  G  H  R  E  C  A  V  C  G  K  G  D  Q  D 108

    ---------|---------|---------|---------|---------|---------|
601 atgcctttgataacgatggtacccatgacaaccaggaacgtgattctgagatctctagct 660
109   A  F  D  N  D  G  T  H  D  N  Q  E  R  D  S  E  I  S  S  S 128

    ---------|---------|---------|---------|---------|---------|
661 caaagtactgctgtgacacagggaatcatgggaaaagggatttcttgctagtgttggact 720
129   K  Y  C  C  D  T  G  N  H  G  K  R  D  F  L  L  V  L  D  C 148

    ---------|---------|---------|---------|---------|---------|
721 gctgtgatttatctgtgctaaaggtagaggaggtggatgtggctgctgtgccaggccttg 780
149   C  D  L  S  V  L  K  V  E  E  V  D  V  A  A  V  P  G  L  E 168

    ---------|---------|---------|---------|---------|---------|
781 agaaatttactaaagctccaaagctcttggctactcctgagaagctcaggtgtgagattg 840
169   K  F  T  K  A  P  K  L  L  A  T  P  E  K  L  R  C  E  I  V 188

    ---------|---------|---------|---------|---------|---------|
841 tccgtgaccttgtggctctgcctgcagatgcttggagggagcagttagactgctcacactc 900
189   R  D  L  V  A  L  P  A  D  A  W  R  E  Q  L  D  C  Y  T  R 208

    ---------|---------|---------|---------|---------|---------|
901 gttgcagccaggctcctggctgtggggagctcttgtttgactctgacaattggagcttac 960
209   C  S  Q  A  P  G  C  G  E  L  L  F  D  S  D  N  W  S  L  R 228

    ---------|---------|---------|---------|---------|---------|
961 ggatcaggaagacggggacttcgacacctgctgactttcctcgtgccatcagaatatggt 1020
229   I  R  K  T  G  T  S  T  P  A  D  F  P  R  A  I  R  I  W  Y 248
```





```
3061 GGAGAGCCGGTTGCCTGCTTCCTGTGATTGCAAAAACTAGCTGCCATGCTCTGAGACTGC 3120

     ---------|---------|---------|---------|---------|---------|
3121 AGGAATCATGTGACACCTCCTGTGGCTGGTCCCCAAGGATTCTGCAAAGCCCTTTTGATG 3180

     ---------|---------|---------|---------|---------|---------|
3181 TGTTCAAGTCAACAGAGCATTCCTGTTTTTGAAAATGCAGGAAAGCCTAAAATGTCTTAA 3240

     ---------|---------|---------|---------|---------|---------|
3241 TAAAGGTACCAGATTTGCTAGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAAAAGC 3300

     ---------|---------|---------|---------|---------|---------|
3301 AAAGAAGAACGTGTGTCAGAACAGGGCACTTTCCACACACTTCCTACCTCCAGGTAATGT 3360

     ---------|---------|---------|---------|---------|---------|
3361 TAATTCTAAAAAGGTTGGAAGTAAAGGGAAATTAGTGTGGCTTGGGACTCTGCTCTCTTC 3420

     ---------|---------|---------|---------|---------|---------|
3421 CAAATGTGGAGTTTCAGCCAGCCTGGTACTAGCGGAGTCTCAGAACAAGCTCTAAGCTCT 3480

     ---------|---------|---------|---------|---------|---------|
3481 TTCAGGTCACTGTCTAGAGAATGGTGCTTGAGACTGAGCCATCAGTGCAGTCACATGAGC 3540

     ---------|---------|---------|---------|---------|---------|
3541 CCAGCCAAGGCAGAGCGGTGGACAGAGGAACCAGAAGGAGCATCCCACTGGCTGGCCAGG 3600

     ---------|---------|---------|---------|---------|---------|
3601 CTTCTGCACCCAGACCTGCTGGGTGACAAAGGCTCTAGGACAGACTGAGAGGAGCGTCCG 3660

     ---------|---------|---------|---------|---------|---------|
3661 CACAGAACCATTGGTTTTCAGAGCCTGGGCTGTAGCAGTGACTGCATTTGGATGGTTGCA 3720

     ---------|---------|---------|---------|---------|---------|
3721 GGGGACTTGCAAATGGTTATCTGCTTTGTATACCACAATGGTACAAGTATCAAAGTGTGT 3780

     ---------|---------|---------|---------|---------|---------|
3781 TAATTCTACTTAATGTTACCCTGATTTCTGTGAAGACTTGAAATAAATGGCTGAGTCTAT 3840

     ---------|------
3841 GCCTTCCTGCTCTCAG 3856
```

Aminopeptidase catalytic HEIAH (E18) amino acid position 481

Nuclear localisation signal (NLS) RRPRKRKRGKR amino acid position 694

Armadillo repeat amino acid position 710-822

Putative start codons at amino acid position 1 and 332.

▼ = Exon Boundaries

AATAAA Polyadenylation signals