# Characterisation of human H-ATPase a4 subunit

## Thesis

UNRESTRICTED.

# CHARACTERISATION OF THE HUMAN $H^+$-ATPase a4 SUBUNIT

by

Ya Su B.Sc. (Hons.), M.Sc.

Submitted to the Open University (Life Sciences) for the Degree of Doctor of Philosophy- August 2004

Addenbrooke's NHS Trust,
Hills Road, Cambridge.
U.K.

Department of Medical Genetics
University of Cambridge
U.K.

Submission date: 5 August 2004
Award date: 22 October 2004



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

# ABSTRACT

The vacuolar $H^+$-ATPase (or V-ATPases) are a family of ATP-dependent proton pumps that move protons across the plasma membrane at specialised sites such as kidney epithelial cells and osteoclasts, as well as acidifying intracellular compartments. The 100 kDa polytopic a subunit of this group of ATPases is suggested to play important roles in proton translocation, assembly, and targeting as well as coupling of ATP hydrolysis and proton transport of the V-ATPase. In man, different a subunit paralogues are encoded by four genes. *ATP6V0A4* encodes a4, which is dominantly expressed apically in α-intercalated cells in both human and mouse kidney.

I sought binding partners for a4 in order to address its potential role in the V-ATPase complex. Random peptide phage display analysis using a4's C terminus as a target protein revealed a consensus motif (WLELRP) with almost complete homology to part of the enzyme phosphofructokinase 1 (PFK-1). Activity of this enzyme is the rate-limiting step in glycolysis. Specificity of a4 binding to this peptide was confirmed by phage ELISA. Protein-protein interaction was further demonstrated by co-immunoprecipitation of a4 with PFK-1 from human kidney membrane proteins. An *in vitro* PFK-1 pull-down assay showed that this interaction is also true for the ubiquitously expressed a1 subunit. Finally, PFK-1 co-immunolocalised with a4 in α-IC in the collecting ducts of human kidney.

These findings indicate a direct link between V-ATPase and glycolysis, via the C-terminus of the pump's a subunit, and suggest a novel regulatory mechanism between V-ATPase function and energy supply. This interaction between the a subunit and PFK-1 also provides new evidence that the C-terminus of this subunit lies cytoplasmically *in vivo.*

Finally, SPR analysis suggests a possible alteration of the a4/PFK-1 interaction by the mutation (G820R) within the a4(C) region identified from a patient with rdRTA, providing a potential mechanism for disease.

**TO MY PARENTS, HUSBAND, AND DAUGHTER**

**FOR THEIR ASSISTANCE, ENCOURAGEMENT AND LOVE.**

# DECLARATION

I declare that the work contained in this thesis, submitted by me for the degree of Doctor of Philosophy, is my own original work, except where explicit reference is made to the work of other authors. It has not been previously submitted for a higher degree or any other qualification.

Signed

Date

# TABLE OF CONTENTS

# LIST OF FIGURES

## LIST OF TABLES

# ACKNOWLEDGEMENTS

Firstly, my thanks and gratitude must go to my supervisor Dr Fiona Karet for giving me the opportunity to undertake the research described in this thesis and for her invaluable ideas, continual encouragement, and support throughout the project. I also thank my second supervisor Dr David Lomas for monitoring and supporting my project each year.

I am deeply grateful to Drs Aiwu Zhou, Hui Hong, Rafia Al-Lamki, Timothy Dafforn and Babak Javid for their help on HPLC, MS, Immunohistology, CD and SPR analysis. I would also like to thank Drs Ruth Case and Catherine Boucher for their advice on the phage display and yeast two-hybrid works.

Many thanks go to my colleagues past and present in our group: Drs Annabel Smith, Katherine Borthwick, Mark Devonald, Nanywan Rungroj and Miss Elizabeth Stover for their useful day-to-day scientific discussions and general help.

Special thanks to Dr Katherine Borthwick for the proof-reading of my thesis.

Finally, I would like to thank my parents Ping Su and Shixin Jiang, husband Erdan Gu and daughter Yisu Gu for their love, sharing my happiness and sadness, as well as their continued support during this study.

# RELATED PUBLICATIONS

***Journal papers***

**Su Y, Zhou A, Al-Lamki RS, Karet FE.** The a-subunit of the V-type $H^+$-ATPase interacts with phosphofructokinase-1 in humans, *J Biol. Chem.* 2003 May. 278(22):20013-20018.

**Smith AN, Finberg KE, Wagner CA, Lifton RP, Devonald MAJ, Su Y and Karet FE.** Molecular cloning and characterisation of *Atp6n1b*: a novel fourth murine vacuolar $H^+$-ATPase a-subunit gene, *J. Biol. Chem.* 2001 Nov. 276(45):42382-42388.

***Conference presentations***

**Ya Su, Aiwu Zhou, Rafia Al-Lamki, Fiona E Karet.** The a subunit of H+-ATPase interacts with Phosphofructokinase 1 in the human kidney. Talk at ASN 35$^{th}$ Annual Meeting and Scientific Exposition, Philadelphia, PA, USA, October, 2002,

# ABBREVIATIONS

| | |
|---|---|
| ABC | ATP-binding cassette |
| ABTS | 2,2'-azino-bis[3-ethylbenzthiazoline 6-sulphonic acid] diammonium |
| a4(C) | the C-terminus of human a4 protein |
| α-IC | α-intercalated cell |
| a4(Loop2) | the second loop of human a4 protein |
| amp | ampicillin |
| a4(N) | the N-terminus of human a4 protein |
| APS | ammonium persulphate |
| ATP | adenosine 5'-triphosphate |
| BDT | BigDye$^{TM}$ Terminator |
| Biotin Cadaverine | *N*-(5-aminopentyl) biotinamide |
| BLAST | Basic Local Alignment Search Tool |
| bp | basepair |
| BSA | bovine serum albumin |
| CA II | carbonic anhydrase II |
| CAPS | 3-[cyclohexylamino]-1-propanesulfonic acid |
| CCD | cortical collecting duct |
| CD | circular dichroism |
| CFTR | cystic fibrosis transmembrane regulator |
| cfu | colony-forming units |
| CHAPS | (3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulphonate |
| 2D | two dimensional |
| 3D | three dimensional |
| DCCD | dicyclohexylcarbodiimide |
| ddRTA | dominant distal renal tubular acidosis |
| DES | *Drosophila* expression system |
| $dH_2O$ | deionised water |
| $H_2O$ | double deionised water |
| dRTA | distal renal tubular acidosis |
| DMSO | dimethyl sulphoxide |
| DSG | disuccinimidyl glutarate |
| DTSSP | 3,3'-dithiobis(sulfosuccinimidylpropionate) |
| DTT | dithiothreitol |
| ECL | enhanced chemiluminescence |
| EDAC | 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide, hydrochloride |
| EDC | N-ethyl-N'-(3-dimethylaminopropyl) carbodiimide |

| | |
|---|---|
| EDTA | ethylendiaminetetra-acetic acid |
| EGTA | ethyleneglycol-bis[ß-aminoethyl etherl |
| ELISA | enzyme linked immunosorbent assay |
| EM | electron microscopy |
| ER | endoplasmic reticulum |
| FCS | fetal calf serum |
| FBS | fetal bovine serum |
| FITC | fluorescein isothiocyanate |
| Gal | galactose |
| GEF | guanyl nucleotide exchange factor |
| GFP | green fluorescent protein |
| Glc | glucose |
| GST | glutathione S-transferase |
| HEPES | N-(2-hydroxyethyl)piperazine-N´-(2-ethanesulphonic acid) |
| His | hexahistidine |
| HIV | human immunodeficiency virus |
| hkCp | human kidney cytosolic protein fraction |
| hkMp | human kidney membrane protein fraction |
| HPLC | high performance liquid chromotography |
| HRP | horse radish peroxidase |
| Hyg | Hygromycin |
| IAA | isoamyl alcohol |
| imf | ionic-motive force |
| IMCD | inner medullary collecting duct |
| IPTG | isopropyl ß-D-thiogalactopyranoside |
| ITC | isothermal titration calorimetry |
| kb | kilobase |
| kDa | kilodalton |
| -L | absence of leucine amino acids |
| LB | Luria-Bertani |
| LC-MS | liquid chromatography-mass spectrometry |
| LR-PCR | long range polymerase chain reaction |
| MBP | 4-(N-maleimido)benzophenone |
| MCS | multiple cloning sites |
| MDR | multidrug resistance |
| MeCN | acetonitrile |
| Mr | relative molecular mass |
| MT | metallothionein |

| | |
|---|---|
| NEB | New England Biolabs |
| NEM | nethylmaleimide |
| NHE-RF | $Na^+/H^+$ exchanger regulatory factor |
| NHS | N-hydroxysuccinimide |
| n-NDG | n-Nonyl-β-D-glucopyranoside |
| OD | Optical densities |
| OMCD | outer medullary collecting duct |
| PBS | phosphate-buffered saline |
| PCR | polymerase chain reaction |
| PDGF | platelet-derived growth factor receptor |
| PEG | polyethylene glycol |
| PFK-1 | phosphofructokinase 1 |
| pfu | plaque-forming units |
| PKA | protein kinase A |
| pmf | proton-motive force |
| PMSF | phenylmethylsulfonyl fluoride |
| rdRTA | recessive distal renal tubular acidosis |
| RT | room temperature |
| RTA | renal tubular acidosis |
| RT-PCR | reverse transcription-PCR |
| SAP | shrimp alkaline phosphatase |
| S2 cells | Schneider 2 cells |
| SDS-PAGE | sodium dodecyl sulphate-Polyacrylamide gel electrophoresis |
| SNHL | sensorineural hearing loss |
| SPR | Surface Plasma Resonance |
| TBS | tris buffered saline |
| TE | tris EDTA |
| TEMED | N,N,N',N'-tetramethylethylenediamine |
| tet | tetracycline |
| TFA | trifluoroacetic acid |
| TM | transmembrane |
| trxB | thioredoxin reductase |
| TFA | trifluoroacetic acid |
| TRITC | tetramethyl rhodamine isothiocyanate |
| Triton X-100 | iso-octylphenoxypolyethoxyethanol |
| Tween 20 | polyoxyethylenesorbitan monolaurate |
| -U | absence of uracil amino acid |
| -UL | absence of uracil and leucine amino acids |

| | |
|---|---|
| v/v | volume per volume |
| w/v | weight per volume |
| Y2H | yeast two-hybrid |

# CHAPTER 1
# INTRODUCTION

## 1.1 The Vacuolar $H^+$-ATPase (or V-ATPase)

The vacuolar $H^+$-ATPase (or V-ATPase) is one of the most fundamental enzymes in nature. Null mutations in genes encoding V-ATPase subunits are likely to be lethal for most eukaryotic cells, underscoring the vital role of this enzyme in living cells (Nelson, 1992a; Nelson and Klionsky, 1996). The V-ATPase is an evolutionarily ancient enzyme that is closely related to the F-ATPase (or F-ATP synthase) of mitochondria, chloroplasts and bacteria (Cross and Duncan, 1996; Fillingame, 1997; Weber and Senior, 1997). The V-ATPase was first identified in the 1960s (Kirshner et al., 1966) in the adrenal medulla and characterised in chromaffin granules (Njus and Radda, 1978). Soon thereafter, it was found in a number of intracellular compartments, including coated vesicles, endosomes, lysosomes, and the central vacuole of *Neurospora*, plants, and yeast. In addition, it has also been found in the plasma membrane of some specialised cells, including osteoclasts, certain epithelial cells in the kidney and male genital tract, macrophages, certain tumour cells and goblet cells of the insect midgut (Brown et al., 1987; Chatterjee et al., 1992; Martinez-Zaguilan et al., 1993; Stevens and Forgac, 1997; Wieczorek et al., 1991). V-ATPases are a family of Adenosine 5'-triphosphate (ATP)-dependent proton pumps and have been particularly well characterised in yeast. Over decades of studies, much effort has been devoted to uncovering the functions, regulation, assembly and targeting of this enzyme in different membranes of a wide variety of organelles. In addition, many genes that encode the subunits of V-ATPases have been isolated in an attempt to determine the functions and regulation of these subunits, and to identify interactions among

the subunits which allows modelling of the overall structure of V-ATPase (Forgac, 1998, 1999; Stevens and Forgac, 1997).

### 1.1.1 Function of the V-ATPase

In eukaryotic cells, the pH of the cytoplasm and of intracellular compartments is one of the most crucial parameters, which has to be carefully controlled by certain transporters. Cytoplasmic pH is regulated by transporters such as the $Na^+/H^+$ exchanger, whereas the pH within many intracellular compartments is maintained principally by V-ATPases (Nishi and Forgac, 2002).

Using ATP as an energy source, V-ATPases perform the vital function of acidifying a variety of intracellular compartments in eukaryotic organisms, and also pumping protons across the plasma membrane of some specialised cells of higher organisms. In all cases, it carries out proton transport from the cytoplasm to either the lumen of intracellular compartments or to the extracellular spaces. This ability of proton pumping not only contributes, along with other ion channels and transporters, to maintaining luminal pH (ranging from 4.5 to 6.5) but it also generates a transmembrane (TM) electrochemical proton gradient [or proton motive force (pmf)] across the intracellular compartmental membranes in organelles and the plasma membrane. Both the acidic luminal pH, and the pmf generated, play important roles in a variety of process in eukaryotic cells, as listed below.

*Functions in Intracellular Organelles*

Firstly, luminal acidification generated by the V-ATPase plays an important role in the activation of ligand-receptor dissociation that is essential for receptor recycling in the processes of receptor-mediated endocytosis and targeting of the newly synthesised lysosomal enzymes. During endocytosis, the internalised

ligand-receptor complexes are delivered to the early endosome compartment, where the low pH triggers release of ligands from their receptors (Geuze et al., 1983). This ligand-receptor dissociation allows the uncoupled receptors to be recycled back to the cell surface where they can be reused. This receptor recycling process not only controls the rate of internalisation of ligands, but also controls the density of cell-surface receptors (Nishi and Forgac, 2002). The receptor density can, in turn, control the sensitivity of cells to ligands. A similar mechanism is also used for the intracellular targeting of newly synthesised lysosomal enzymes from the Golgi complex to lysosomes. In this pathway, the newly synthesised lysosomal enzyme first binds to their receptor, mannose-6-phosphate (Man-6-P) in the trans-Golgi, to form lysosome/Man-6-P complexes which are then delivered to the late endosomes with an acidic luminal environment (Kornfeld, 1992). The low pH in the late endosomes activates release of the lysosomal enzymes from Man-6-P receptors, again allowing the uncoupled receptors to be recycled back to the trans-Golgi, where they can be reused.

Secondly, luminal acidification of endosomes is required for budding of endosomal carrier vesicles (Aniento et al., 1996). Formation of carrier vesicles mediates the transport of the released ligands (during endocytosis) from early to late endosomes and then to lysosomes (Clague et al., 1994; van Weert et al., 1995). These movements can be blocked by bafilomycin, a specific V-ATPase inhibitor, indicating the involvement of V-ATPase activity.

Thirdly, luminal acidification of endosomes also activates the fusion of internalised envelope viruses (such as influenza virus) with the endosomal membrane, a step that is essential for viral infection (White, 1992). In addition, vacuolar acidification in yeast is required for protein degradation and transport of small molecules and ions across the vacuolar membrane (Anraku et al., 1992).

Finally, during exocytosis, the pmf, which consists of voltage and pH differences across the plasma membrane, generated by the V-ATPase, provides a driving force for the transport of hormones or transmitters into secretory or synaptic vesicles derived from endosomes or from the Golgi network (Moriyama et al., 1992).

*Functions in the Plasma Membrane*

In osteoclasts, the V-ATPase, as the main proton pump, is present at a high density at a specialised domain of the apical plasma membrane, known as the 'ruffled border'. In the ruffled border, the V-ATPases pump protons into a sealed extracellular space known as the resorption lacuna between the osteoclasts and bone to create acidic pH required for bone digestion (bone resorption) by cathepsin K (Blair et al., 1989). Loss of function of the V-ATPase causes an increased pH in this space, which inhibits the activity of cathepsin K, resulting in osteopetrosis. For example, defects in the a3 subunit of the osteoclast plasma membrane V-ATPase are associated with infantile malignant osteopetrosis (Frattini et al., 2000; Kornak et al., 2000).

In the distal nephron of the mammalian kidney, α-intercalated cells (α-ICs) express large numbers of V-ATPases on their apical plasma membrane. The apical V-ATPases of α-ICs are responsible for pumping protons across the plasma membrane into the urine. This proton translocation plays a central role in maintaining acid-base homeostasis of the organism. Failure or inadequate $H^+$ excretion by the V-ATPase results in recessive distal renal tubular acidosis (rdRTA) as described in Section 1.3.5.3.

The mammalian epididymis and the vas deferens contain significant numbers of specialised proton-secreting epithelial cells that express many V-ATPases on their apical plasma membrane (Brown et al., 1992a). In these cells,

the V-ATPase pumps protons across the plasma membrane into the lumen and is the major source of acidification compared to other acid-base transporters, which include a $Na^+/H^+$ exchanger and a $Na^+/HCO_3^-$ cotransporter (Au and Wong, 1980; Jensen et al., 1999). The low luminal pH is essential for sperm maturation as well as maintenance of sperm in an immotile state during their passage through the epididymis and vas deferens. Failure of luminal acidification is a potential explanation for some cases of male infertility (Hinton and Palladino, 1995).

Certain tumour cells also express V-ATPases at the plasma membrane where they transport protons into the extracellular space which is required for secretion of lysosomal enzymes. Here these enzymes participate in degradation of extracellular matrix, which is required for tumour metastasis (Martinez-Zaguilan et al., 1993).

Finally, the pmf generated by V-ATPases across the membranes of eukaryotic cells is utilised as a driving force for numerous secondary transport processes. For example, V-ATPases occur at the apical surface of goblet cells in the insect midgut where they pump $H^+$ into the lumen of the midgut to generate pmf. This, in turn, provides the driving force to transport $K^+$ into the midgut in exchange for $H^+$ through the $K^+/2H^+$ antiporter (Wieczorek et al., 1991). In sum, the luminal pH of the midgut is in fact alkaline, which is essential for signal sensory generation (Wieczorek, 1992). In addition, V-ATPases are found at the apical surface of frog skin epithelial cells, where they pump protons out of cells to generate pmf. This generated pmf is also used as a driving force to energise $Cl^-$ uptake into the cells through a $Cl^-/HCO_3^-$ exchanger (Jensen et al., 1997).

### 1.1.2 Structure of the V-ATPase

V-ATPases from fungi, plants and animals are structurally very similar to each other (Stevens and Forgac, 1997). They are large multi-subunit heteromeric enzymes with a structure and mechanism similar to the F-ATPase (Adachi et al., 1990b; Forgac, 2000; Nelson and Nelson, 1989; Nelson, 1989, 1992a). Yeast genetics, using null mutations of V-ATPase subunits, has helped to identify the genuine V-ATPase subunits and led to the discovery of some novel subunits as well as proteins that function exclusively in the assembly of the enzyme but which are not the elemental subunits (Nelson and Harvey, 1999). The majority of genes encoding V-ATPase subunits were first cloned from bovine tissues (Nelson and Harvey, 1999). Using techniques such as electron microscopy (EM) and crosslinking, considerable progress has been made in understanding the structure of V-ATPase.

The holoenzyme has a molecular weight of approximately 830 Kilodalton (kDa). At least 13 genes encode the different V-ATPase subunits in the most elaborately explored case: the yeast *Saccharomyces cerevisiae*. However, the total number of subunits may differ across species, probably due to specific regulatory constraints. Like the F-ATPase, the V-ATPase has a ball-and-stalk structure and membrane extrinsic and intrinsic domains, termed $V_1$ and $V_0$, respectively (Forgac, 1989, 1998; Moriyama et al., 1991; Stevens and Forgac, 1997). The $V_1$ domain is a catalytic peripheral complex of approximately 570 kDa that is composed of 8 different kinds of subunits (subunits A-H) with a probable stoichiometry of $A_3B_3C_1D_1F_1G_2H_{1-2}$ (Arai et al., 1988; Kawasaki-Nishi et al., 2003b; Xu et al., 1999). The stoichiometry of majority $V_1$ subunits has been determined by quantitative amino acid analysis (Arai et al., 1988). The main function of the $V_1$ domain is to bind ATP and to catalyse ATP hydrolysis. The $V_0$ domain is an

integral membrane complex of approximately 260 kDa that is composed of five different kinds of subunits (a, c, c', c" and d) with a stoichiometry of $a_1d_1c"_1(c'c)_{4-6}$ (Arai et al., 1988; Powell et al., 2000). Both subunits c' and c" are homologous to the c subunit and they are all highly hydrophobic proteolipid subunits. Genetic studies in yeast have demonstrated that a fully functioned V-ATPase must contain at least one copy of each proteolipid subunit (Hirata et al., 1997). It is not clear why the V-ATPase requires three different proteolipid subunits compared to its evolutionarily related F-ATPase. However, to date, mammalian orthologues of yeast c' have not been identified. In addition, a novel, extremely hydrophobic 9.2-kDa integral membrane protein, named M9.7 or M9.2, was previously identified in *Manduca sexta*, *Arabidopsis thaliana*, and bovine chromaffin granules (Ludwig et al., 1998; Sze et al., 2002; Wieczorek et al., 2000). Very recently, the yeast orthologue of this protein, named Vma9p, was identified and its null mutation showed a typical $Vma^-$ growth phenotype (i.e. cells that grow only in acidic conditions: pH 5.5-6.5) (Sambade and Kane, 2004). The finding of Vma9p has finally positioned this protein as a component element of $V_0$ domain of V-ATPase and it is now referred to as subunit e. The main function of the $V_0$ domain is to transport protons across the membrane. The subunit molecular weights, distributions and proposed functions of the yeast and mammalian V-ATPase are shown in Table 1. Furthermore, several studies on V-ATPases of yeast and insects showed that a reversible dissociation of the $V_1$ and $V_0$ domains occurs, suggesting a dynamic equilibration between assembled and dissociated V-ATPase complex, as described later in Section 1.1.4.

| Domain | Code | Mammalian | | Yeast | | ***F-ATPase Homologous | ***Proposed Function (in V-ATPase) |
|---|---|---|---|---|---|---|---|
| | | Mr (kDa) | *Genes | Mr (kDa) | Genes | | |
| $V_1$ | A | 70-73 | *ATP6V1A* | 69 | *VMA1* | β | Catalytic site, regulation |
| | B | 56-58 | *ATP6V1B* | 60 | *VMA2* | α | Non-catalytic site, targeting? |
| | C | 40-42 | *ATP6V1C* | 42 | *VMA5* | ? | Assembly, peripheral stalk? |
| | D | 33-34 | *ATP6V1D* | 32 | *VMA8* | γ | Central stalk, assembly |
| | E | 31-33 | *ATP6V1E* | 27 | *VMA4* | ? | Assembly, peripheral stalk? |
| | F | 10-14 | *ATP6V1F* | 14 | *VMA7* | ε | Assembly, central stalk? |
| | G | 14-15 | *ATP6V1G* | 13 | *VMA10* | b | Assembly, peripheral stalk? |
| | H | 50-57 | *ATP6V1F* | 54 | *VMA13* | | Not assembly, peripheral stalk? |
| $V_0$ | a | 100-116 | *ATP6V0A* | 100 | *VPH1/STV1* | a | $H^+$ transport, assembly, targeting, bafilomycin site |
| | c | 14-17 | *ATP6V0C* | 17 | *VMA3* | c | $H^+$ transport, **DCCD site |
| | c' | ? | *ATP6V0C'* | 17 | *VMA11* | c | $H^+$ transport, |
| | c" | 19-20 | *ATP6V0C"* | 23 | *VMA16* | c | $H^+$ transport |
| | d | 38-39 | *ATP6V0D* | 36 | *VMA6* | ? | Assembly |
| | e | 9.2 | *ATP6V0E* | 8.4 | *VMA9* | ? | unknown |

**Table 1 Subunits of V-ATPase.**

*The Human Gene Nomenclature Committee (HGNC) recently redesignated all genes in the V-ATPase (ATP6) family (Smith et al., 2003). All symbols used in this study are the new official designations unless otherwise indicated. **DCCD (dicyclohexylcarbodiimide) is an inhibitor of proton transport. ***For references, see (Forgac, 1999; Hunt and Bowman, 1997; Kawasaki-Nishi et al., 2003b; Ludwig et al., 1998; Sambade and Kane, 2004).

Topographical analysis using membrane impermeant reagents and proteolysis indicates that the $V_1$ complex is orientated towards the cytoplasmic side of the membrane and can be dissociated from the membrane using chaotropic agents (such as KI and KNO3) in the absence of detergents (Adachi et al., 1990a; Adachi et al., 1990b; Arai et al., 1988). The $V_1$ domain possesses six potential nucleotide binding sites located on subunits A and B (Forgac, 1999; Nishi and Forgac, 2002). Only those on the A subunit are catalytic binding sites (Feng and Forgac, 1992; Zhang et al., 1995), whereas the remaining three located on the B subunit are non-catalytic, but essential sites, presumed to be regulatory (Liu et al., 1996; Vasilyeva and Forgac, 1996; Zhang et al., 1995). The $V_0$ domain is a membrane-spanning complex. The three dimensional (3D) structure of the $V_0$ domain purified from bovine brain has been determined by EM at a 21 Å resolution (Wilkens and Forgac, 2001). In 1994, Zhang *et al.* had observed a DCCD inhibitable passive proton conductance for the reassembled $V_0$ domain, suggesting that this domain is responsible for proton translocation across the membrane in which it is anchored (Zhang et al., 1994). The DCCD binding site is located on the c subunit of the $V_0$ domain (Arai et al., 1987; Finbow et al., 1992). Although subunit d in the $V_0$ domain does not possess any integral membrane spanning regions (Wang et al., 1988), it remains tightly bound to $V_0$ domain upon dissociation of $V_1$ from $V_0$ (Zhang et al., 1992). Connection of subunit d with the $V_0$ domain may be through interaction with the a subunit cytoplasmic domain (Graham et al., 2000). Apart from the d subunit, the a, c, c' and c" subunits have significant luminal domains (Arai et al., 1988).

To date the yeast and bovine V-ATPases are the best characterised proton pumps in eukaryotes. EM has revealed that, like the F-ATPase, the peripheral and

integral domains of the bovine V-ATPase are connected by two stalks (Wilkens et al., 1999). In addition, crosslinking studies of the bovine V-ATPase revealed the arrangement of certain subunits in the complex (Adachi et al., 1990b; Xu et al., 1999). As a result, a more detailed structural model (Figure 1*A*) can be proposed that is defined by several regions: a catalytic core composed of A and B subunits; a central stalk, suggested to be composed of D and F subunits; a peripheral stalk likely to be composed of the N-terminus of the a subunit, together with C, E, G and H, and a proton-translocating domain composed of the C-terminus of a, with c, c', c'' and d subunits (Forgac, 2000; Landolt-Marticorena et al., 2000; Nishi and Forgac, 2002; Stevens and Forgac, 1997; Xu et al., 1999).

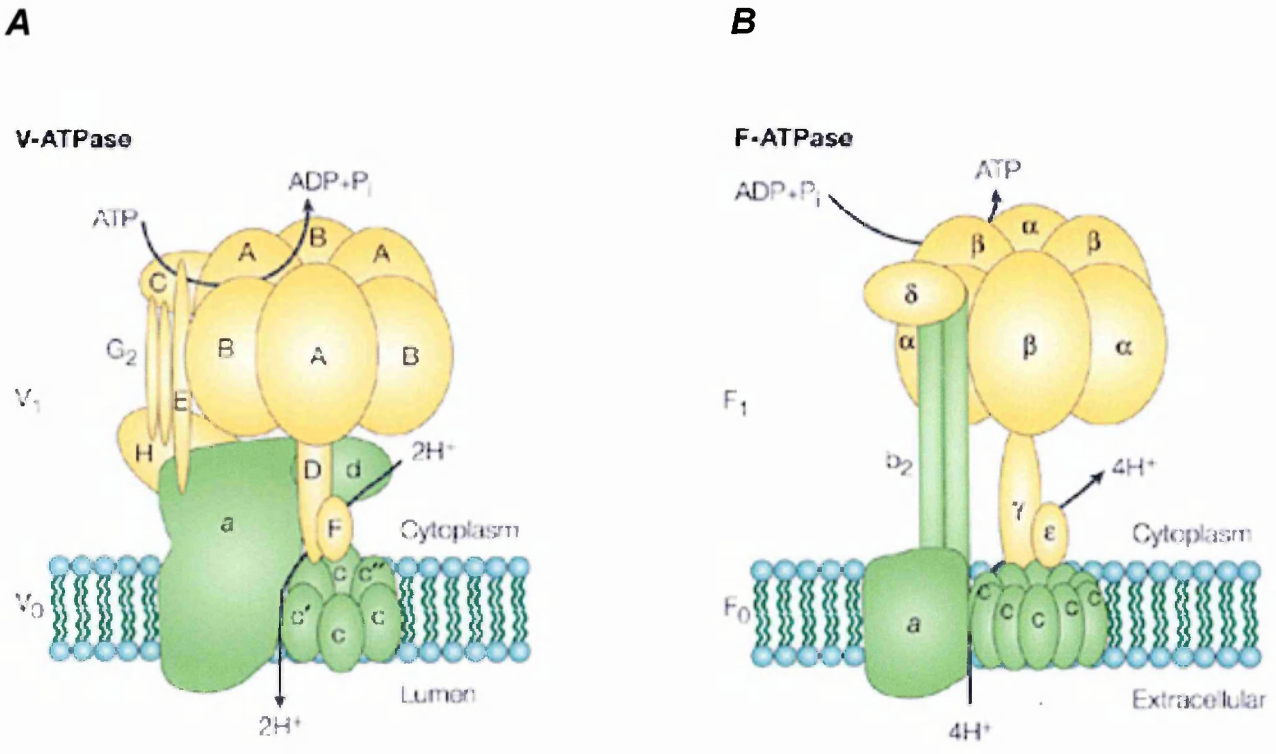



**Figure 1 Schematic models of V-ATPase and F-ATPase.**

Schematic models of V-ATPase (*A*) and F-ATPase (*B*). The peripheral $V_1/F_1$ domain subunits are shown in yellow, and the integral $V_0/F_0$ domain subunits in green. This figure is taken from reference (Nishi and Forgac, 2002).

### 1.1.3 Regulation of the V-ATPase

A critical question is what is the mechanism by which luminal acidification generated by the V-ATPase is controlled in eukaryotic cells. Although the cellular proteins that regulate V-ATPases remain largely unknown (Forgac, 2000), several studies have provided clues into the assembly, targeting, and organisation of V-ATPase, changes in the tightness of coupling between proton transport and ATP hydrolysis, as well as for the regulation of V-ATPase activity.

In yeast, the $V_1$ domain cannot be attached to the $V_0$ in the absence of any of the $V_0$ subunits, although the absence of $V_1$ does not prevent the $V_0$ domain from being assembled and targeted to the central vacuole (Doherty and Kane, 1993; Kane et al., 1992; Nelson and Nelson, 1990; Noumi et al., 1991). However, the $V_0$ domain cannot be assembled/or stabilised in the absence of the c and a subunits (Bauerle et al., 1993; Graham et al., 2000; Kane et al., 1992). So where does assembly of V-ATPase occur? In bovine cells, assembly of V-ATPase complexes from newly synthesized subunits is blocked when transport from the endoplasmic reticulum (ER) to the Golgi complex is inhibited (Graham et al., 1998; Myers and Forgac, 1993). Also, three proteins, Vma12p, Vma21p and Vma22p, required for $V_0$ domain assembly have been localised to the ER of yeast (Graham et al., 1998; Hill and Stevens, 1994; Hirata et al., 1993; Jackson and Stevens, 1997; Tomashek et al., 1997). These results suggest that the assembly of $V_0$ and $V_1$ into the holoenzyme occurs in the ER. These assembled V-ATPases are then targeted to their required destinations, such as vesicular compartments and specialised cell membranes.

Several vertebrate urinary epithelia are able to shuttle V-ATPases between vesicular compartments and the plasma membrane upon stimulation. For example, in the kidney, systemic acidosis causes vesicle exocytosis which in turn

delivers more V-ATPases to the apical plasma membrane, resulting in an increase in $H^+$ secretion (Brown, 2000; Brown and Breton, 2000; Gluck, 1992; Schwartz et al., 1985). In contrast, when the stimulus is reversed, the proton pumps are removed from the apical membrane by endocytosis. These results suggest that a dynamic exchange of V-ATPases may occur between endomembranes and the plasma membrane (Harvey and Wieczorek, 1997).

In addition to the reversible membrane shuttle of the holoenzyme, reversible dissociation of the $V_1$ and $V_0$ domains has been shown to occur in both yeast and insects. In 1995, Kane demonstrated that in yeast, glucose deprivation results in a rapid dissociation of $V_1$ and $V_0$ domains, and this effect is reversed upon addition of glucose (Kane, 1995). This study suggests that there is a dynamic equilibration between the assembled and dissociated V-ATPase complex. A similar conclusion has been drawn from studies of changes in V-ATPase assembly during moulting and starvation in insects (Graf et al., 1996; Sumner et al., 1995; Wieczorek et al., 2000). These studies demonstrated that under these conditions, the amount of V-ATPase holoenzyme in the goblet cell apical membrane of the tobacco hornworm midgut decreases whilst the concentration of cytosolic $V_1$ complexes increases. This suggests a reversible shuttling of $V_1$ complexes that dissociate and reassociate with the membrane $V_0$ complexes, rather than by a reversible membrane shuttle. Because the separated $V_1$ and $V_0$ domains do not function as a ATPase or as a proton pump, this disassembly/reassembly of $V_1$ and $V_0$ domains may provide a potential mechanism for regulating V-ATPase activity *in vivo* (Puopolo et al., 1992b; Zhang et al., 1992).

In addition, certain V-ATPase subunits that have multiple paralogues have been suggested to contain information necessary for differential targeting of V-ATPase. These include subunits B, C, G, E, a and d. Furthermore, a 29-kDa

glycoprotein named Ac45 (also called glycoprotein IV) was identified in association with the $V_0$ domain of V-ATPase from the membrane of chromaffin granules (Getlawi et al., 1996; Supek et al., 1994). This protein is present in many mammalian V-ATPases. As there is no yeast counterpart of Ac45, it is quite difficult, without yeast genetics as a tool, to make a convincing assignment of this protein as a genuine V-ATPase subunit. However, a possible function for Ac45 as an organelle and membrane specific targeting protein was proposed (Supek et al., 1994).

Besides the assembly and targeting described above, V-ATPase activities and proton transport can also be affected by several other mechanisms. Firstly, disulphide bond formation between conserved cysteine residues located at the catalytic site on the A subunit (Cys254 and Cys532, bovine numbering) leads to the reversible inactivation of V-ATPase, suggesting the cysteine residues are involved in redox regulation of the enzyme activity (Feng and Forgac, 1992, 1994; Nishi and Forgac, 2002). Formation of the disulphide bond may be induced by nitric oxide and the mechanism involved is possibly through a conformational change of the A subunit according to X-ray crystal structure analysis of $F_1$ β subunit which is the counterpart of the $V_1$ A subunit (Abrahams et al., 1994).

Secondly, a variety of extrinsic conditions and intrinsic residues of V-ATPase elements cause changes in the efficiency of coupling between the proton transport and the ATP hydrolysis activity of V-ATPases. These factors are listed as follows:

A) A high concentration of ATP causes partial uncoupling of V-ATPase. In this case, ATP hydrolysis continues to increase along with increasing concentration of ATP, but proton transport plateaus and then decreases (Arai et al., 1989).

B) Sodium azide can completely inhibit proton transport by the V-ATPase without affecting ATP hydrolysis (Vasilyeva and Forgac, 1998).

C) Arg-735 of the yeast orthologue Vph1p was identified to be essential for proton transport. A change of this residue to lysine results in a complete inactivity of proton transport, but about 25% of wild-type ATPase activity is retained, suggesting a partial uncoupling of the proton transport and ATPase activity has occurred (Kawasaki-Nishi et al., 2001b).

Thirdly, the luminal pH generated by V-ATPases can be regulated by other transporters such as the $Cl^-$ channel or the $(Na^+/K^+)$-ATPase. For example, in coated vesicles, continued proton transport by V-ATPases creates an electrical potential difference across the membrane. This membrane potential needs to be dissipated by moving other charged species across the membrane (Arai et al., 1989). The dissipation occurs primarily through the action of a coated vesicle chloride channel (Arai et al., 1989; Glickman et al., 1983). Changes in activation and inactivation states of the $Cl^-$ channel through phosphorylation and dephosphorylation by cyclic-AMP dependent protein kinase A (PKA) affects $Cl^-$ conductance, which in turn changes the activity of V-ATPase (Mulberg et al., 1991).

Finally, interactions between the V-ATPase and other chemicals or proteins that are not components of the proton pump are, or might be, involved in the regulation of proton pumping or enzyme activity. For example, the c subunit in $V_0$ domain contains DCCD binding sites (E137, yeast numbering). Binding of DCCD to this subunit inactivates proton pumping as well as ATPase activity of V-ATPase (Nelson and Nelson, 1989, 1990). In addition, associations of subunits E and a with glycolytic enzymes, may be involved in the regulation of V-ATPase activities (Lu et al., 2001; Lu et al., 2003).

### 1.1.4 Comparison of the V-ATPase and F-ATPase

#### 1.1.4.1 Structure and Enzyme Activities

The exchange of energy between ATP and ionic eletrochemical gradients [(or ionic motive force (imf)] is a fundamental property of life. In eukaryotic cells, three broad classes of ATPases [F-ATPases (or F-ATP synthases), V-ATPases, and P-ATPases] mediate the exchange. Among these three ATPases, V-ATPases are structurally and evolutionarily closely related to the F-ATP synthases. In other words, the two enzymes appear to have descended from an ancestral ATPase. This relationship is evident both in their overall structure and in the sequence homology or functional similarity of certain component subunits. However, differences between these two enzymes have also been demonstrated in many aspects, particularly in their function and regulation.

In eukaryotic cells F-ATP synthases are confined to chloroplasts and mitochondria, whereas in all known eubacteria, they are present in the plasma membrane. The primary function of this enzyme in eukaryotic cells is to form ATP at the expense of pmf. In other words, it acts as an ATP synthase. However, this enzyme also functions as an ATPase, i.e. generation of pmf by catalysing ATP-dependent proton pumping. However, the ATPase role of F-ATP synthases can only be performed in exceptional cases, such as under anaerobic conditions in bacteria (Futai et al., 1989; Grabe et al., 2000). By contrast, V-ATPases can only perform the latter function, i.e. operate as an ATP hydrolase, pumping protons away from the cytoplasmic compartment. The loss of the ATP synthase ability in the V-ATPase is probably due to alterations in the membrane domains of this enzyme (Nelson, 1992a).

The two enzymes are compared below, with both their similarities and differences highlighted.

The F-ATPase is a multisubunit heteromeric protein (Figure 1*B*) and it has a relatively small molecular mass of approximately 500 kDa, compared to the V-ATPase with a molecular weight of ~830 kDa (see Section 1.1.2). It consists of at least nine different subunits arranged in a catalytic/regulatory $F_1$ domain and an $H^+$-translocating $F_0$ domain. The X-ray crystal structure of the $F_1$ domain from bovine heart mitochondria has been determined at the 2.8 Å resolution (Abrahams et al., 1994). This domain is a peripheral complex that is composed of five different subunits ($\alpha\beta\gamma\delta\varepsilon$) with a stoichiometry of $\alpha_3\beta_3\gamma_1\delta_1\varepsilon_1$. Of these subunits, $\alpha_3\beta_3$ forms a catalytic core, $\gamma$, $\varepsilon$, and $\delta$ form a central stalk and a peripheral stalk. The $F_0$ domain is an integral membrane complex that is composed of four different types of subunits (abcd) with a stoichiometry of $a_1b_1c_{12}d_1$. The c subunits are arranged in a ring structure which forms the main part of the proton-translocating machinery together with the a subunit. Similar to F-ATPases, V-ATPases are composed of a catalytic $V_1$ domain and a proton translocating $V_0$ domain as described in Section 1.1.2.

Apart from the overall structural similarities, the evolutionary link between these two families of enzymes has also been demonstrated by the sequence homology or functional similarity of certain subunits. The F-ATPase catalytic subunit $\beta$ has roughly 25% amino acid sequence identity with the A subunit of V-ATPase. Catalytic residues identified in the P-loop (GXXXXGKT) and the GERXXE sequence of the $\beta$ subunit are conserved in the A subunit (Futai et al., 1989; Omote and Futai, 1998). Yeast mutagenesis experiments confirmed that the lysine in the P-loop and the glutamate in the GERXXE sequence are catalytically essential (Liu et al., 1997). Conservation of the two residues in the V-ATPase suggests that the V-ATPase may have a similar catalytic mechanism (Futai et al., 2000). In addition to the catalytic subunits, approximately 25% amino acid

sequence identity is shared between two non-catalytic nucleotide binding subunits, α (in the F-ATPase) and B (in the V-ATPase) as well as between the proteolipid subunits (c subunit) of the two enzymes (Bowman et al., 1992; Inatomi et al., 1989). Although no other subunits in the two enzymes display such sequence homology, similar structural motifs, in a coiled-coil arrangement, were identified in the D subunit of $V_1$ and the γ subunit of $F_1$ (Nelson et al., 1995). This suggests that the D subunit may function in the V-ATPase as the counterpart of the $F_1$ γ subunit, which has been implicated in coupling of ATP hydrolysis to proton pumping and in coupling pmf to ATP production (Nelson and Harvey, 1999). Moreover, the F subunit of $V_1$ has been tentatively assigned to be an analogue of $F_1$ δ subunit based on their similarity in affecting proton pumping activity and solubility properties (Graf et al., 1994; Nelson et al., 1994). Finally, subunits a in both $F_0$ and $V_0$ domains function similarly, i.e. as aids in proton translocation across the membrane, although no amino acid sequence homology has been shown between the two subunits.

However, despite the overall structure similarity of the two holoenzymes, differences were revealed from the EM study of V-ATPase from *Neurospora* and the X-ray crystal structure of the $F_1$ domain of F-ATPase. From these studies it was observed that although both enzymes contained a globular head attached to the membrane by a central stalk, the V-ATPase had additional projections emanating from the base of the stalk (Dschida and Bowman, 1992). In addition, subunit types and numbers in each domain of the two enzymes have certain differences. For example, unlike the $F_1$ domain, the $V_1$ domain contains eight different kinds of subunits rather than five and unlike the $F_0$ domain, the $V_0$ domain contains additional proteolipid subunits c' and c'' rather than just the c subunit.

With regard to the enzyme activities, unlike the analogous $F_1$ and $F_0$ domains, the dissociated $V_1$ and $V_0$ do not function as an ATPase and a proton channel, respectively (Puopolo et al., 1992b; Zhang et al., 1992). Also, enzyme activities of the two ATPases are affected differently under certain conditions. For example, activity of the F-ATPase can be specifically inhibited by oligomycin, whereas the V-ATPase shows no sensitivity on to this chemical (Forgac, 1989). Instead, activity of the V-ATPase is specifically inhibited by bafilomycin and concanamycin (Bowman et al., 1988; Drose et al., 1993). In addition, a cysteine residue with a regulatory role through disulphide bonding is conserved in the P-loop (GXXXCGKT) of $V_1$ A subunit but not in $F_1$ β subunit. Modification of this cysteine residue using nethylmaleimide (NEM) inhibits V-ATPase activity. However, the F-ATPase is insensitive to NEM which can probably be attributed to the fact that the $F_1$ β subunit lacks this conserved cysteine in the P-loop (Iwamoto et al., 1994).

### 1.1.4.2 Mechanism of Proton Transport

Overall structural similarities suggest that both of the V- and F-ATPases utilise the same basic mechanism for ATP hydrolysis, proton transport and energy coupling. In other words, the mechanism by which V-ATPases carry out ATP-dependent proton transport can be drawn from the mechanism proposed for F-ATPase hydrolytic function (Cross and Duncan, 1996; Vik and Antonio, 1994). In the latter, energy produced from ATP hydrolysis by the catalytic core is thought to drive the rotation of the central stalk formed mainly by the γ subunit which is connected both to the catalytic core in $F_1$ and the ring of c subunits within $F_0$ (Cross and Duncan, 1996; Noji et al., 1997; Sabbert et al., 1996). This results in the rotation of the c ring, which lies adjacent to the a subunit. The a subunit is held

fixed relative to the catalytic core through the peripheral stalk and thought to bring the protons to a glutamate residue within the carboxyl group of the c subunit. The c subunit then rotates the protons through the membrane bilayer and releases them on the opposite side of the membrane as it rotates. Because rotation occurs in only one direction: from the $F_1$ to $F_0$ induced by ATP hydrolysis, this mechanism serves to transfer energy stored in ATP through the form of rotatory energy to a linear pmf across the membrane (Forgac, 2000).

This rotatory mechanism for the V-ATPase is also supported by a number of studies, carried out very recently. Firstly, a counter-clockwise continuous rotation of the G subunit (possible role in peripheral stalk) relative to the fixed c subunit in yeast V-ATPase was observed from the membrane side (Hirata et al., 2003). Secondly, a counter-clockwise rotation of the D or F subunit (possible role in the central stalk) of *Thermus thermophilus* V-ATPase was viewed from the membrane side (Imamura et al., 2003). Finally, a counter-clockwise rotation of the c subunit ring (role in the proton translocation pore) of *Thermus thermophilus* V-ATPases was also observed from the membrane side (Yokoyama et al., 2003b). All rotations observed above are ATP dependent.

### 1.1.5 Organelle and Tissue Specificity of the V-ATPase Subunits

The subunits forming the functional V-ATPase molecule may vary from tissue to tissue. To date, several subunits in both the $V_1$ and $V_0$ domains of the V-ATPase have been reported as having multiple paralogues which are encoded by different genes and show some organelle and tissue specificity as follows. Two genes, *ATP6V1B1* and *ATP6V1B2*, encoding homologous but distinct 57-kDa B subunits (B1 and B2) have been cloned from many higher eukaryotes (Chatterjee et al., 1992; Nelson, 1992b; Puopolo et al., 1992a). Their amino acid sequences

share approximately 80% identity. The B2 subunit is known to be ubiquitously expressed, whereas B1 is expressed most abundantly in the kidney, but also the cochlea and endolymphatic sac of the inner era (Karet et al., 1999b). In the kidney, the B1 subunit is found at the apical surface of acid-secreting cells (Nelson et al., 1992). Deficiency of B1 causes rdRTA with impaired hearing (Karet et al., 1999b). Also, two 42-kDa C (C1 and C2), three 14-kDa G (G1, G2 and G3), and two 40-kDa d (d1 and d2) paralogues have been identified in both human and mouse (Murata et al., 2002; Smith et al., 2002; Sun-Wada et al., 2003). There are approximately 62%, 53% and 68% identity in amino acid sequences between the two C, three G and two d paralogues, respectively. Among the different human C, G and d transcripts, C1, G1, and d1 are ubiquitously expressed, C2 is seen only in kidney and placenta, d2 is expressed in kidney, osteoclast and lung, and G2 and G3 are seen in the brain and the kidney in human, respectively (Smith et al., 2002). Furthermore, two paralogues of the E subunit, E1 and E2, have been reported, of which the E1 is a testis-specific paralogue in mouse (Sun-Wada et al., 2002). Heterogeneity of the a subunit across species has been widely reported and is described in Section 1.3.4.

The existence of multiple homologues of some subunits and their limited tissue localisation may contribute to the differential targeting and transport properties, as well as the physiological regulation of the V-ATPase in different cell types and membrane domains.

## 1.2 Interactions Involving Subunits of the V-ATPase

Various physical interactions between V-ATPase subunits, or between subunits and other proteins, have been reported. Inter-molecular interactions among different subunits are more likely to provide a structural support for the

proton pump, whereas interactions with other proteins might provide insights into V-ATPase assembly, transport, targeting, or regulation. Here, interactions relating to the a subunit are omitted. Instead, they are listed separately in Section 1.3.6.

### 1.2.1 Interactions between Subunits within the V-ATPase

As mentioned earlier, the overall structure of the V-ATPase has been uncovered mainly by using EM and membrane topology analysis. However, much more information regarding interactions between subunits is required in order to determine the arrangement of the subunits within the V-ATPase complex. This will aid in the exploration of further mechanisms of this enzyme. Using covalent cross-linking techniques followed by Western blotting, Xu *et al.* (1999) have reported interactions between certain subunits mainly within the $V_1$ domain of V-ATPase from calf brain clathrin-coated vesicles (Xu et al., 1999). In this study, using the cross-linking reagent disuccinimidyl glutarate (DSG), they have identified interactions between the D and F, H and E, as well as H and F subunits. Also using the cross-linking reagent 4-(N-maleimido)benzophenone (MBP), evidence for subunit interactions between G and E has been provided. Furthermore, subunits between C and E, and D and E have been cross-linked by another cross-linking reagent, 1-ethyl-3-(dimethylaminopropyl)carbodiimide hydrochloride (EDC). Moreover, crosslinking studies using 3,3'-dithiobis(sulfosuccinimidylpropionate) (DTSSP) followed by 2D gel electrophoresis revealed interactions between the A and B subunits as well as between the c and C, D and E subunits (Adachi et al., 1990b). Finally, a yeast two-hybrid library screen identified interaction between the H and E subunits, which has been further confirmed by GST pull-down assay (Lu et al., 2002).

### 1.2.2 Interactions between Subunits of the V-ATPase and other Proteins

Interactions between V-ATPase subunits and proteins that are not the elements of the enzyme have now been widely observed. Some of these associations have led to the proposal of certain novel regulatory functions.

Firstly, the V-ATPase was found to be associated with the cytoskeleton in osteoclasts (Nakamura et al., 1997). Mice lacking this association showed severe bone deformities due to abnormal osteoclast bone resorption, suggesting that this interaction is essential for recruitment of the V-ATPase to the ruffled border and for normal bone resorption. A potential mechanism for the association of V-ATPase with the cytoskeleton was later identified: a direct binding between the V-ATPase and F-actin (Lee et al., 1999). This V-ATPase/F-actin interaction appeared to be mediated by a high affinity F-actin binding site in the N-terminal domain of the B subunit of V-ATPase (Holliday et al., 2000). This interaction suggests a new mechanism for controlling bone absorption. Secondly, the C-terminal domain of B1 (but not B2) contains a PDZ-binding motif which mediates interaction with the PDZ-binding protein, $Na^+/H^+$ exchanger regulatory factor (NHE-RF) in a cell type-specific manner (Breton et al., 2000). Thirdly, an interaction between the E subunit of V-ATPases and mSos1 was reported (Miura et al., 2001). mSos1 is a guanine nucleotide exchange factor which regulates cell growth, transformation and differentiation (Bowtell et al., 1992). Association of the E subunit and mSos1 suggests that V-ATPase E subunit participates in the regulation of the mSos1-dependent Rac1 signalling pathway. Fourthly, the H subunit of the V-ATPase has been shown to interact with Nef protein, which is the accessory protein of human immunodeficiency virus (HIV). As Nef protein controls expression of CD4 (the principal receptor for HIV) on the surface of infected cells, the Nef/H subunit association may be involved in the regulation of CD4 levels on the surface of T

cells (Lu et al., 1998). Fifthly, the c subunit has been shown to interact with several proteins, including β1 integrin (Skinner and Wildeman, 1999), E5 oncoprotein of papilloma viruses (Finbow et al., 1991), platelet-derived growth factor receptor (PDGF) (Goldstein et al., 1992) and synaptophysin/synaptobrevin (Galli et al., 1996). These associations have indicated a role for this subunit in cell growth or transformation. Finally, a direct link between the V-ATPase and glycolysis has been proposed for the first time through the observation of an interaction between the E subunit and aldolase (Lu et al., 2001).

## 1.3 The a Subunit of the V-ATPase

### 1.3.1 Structure of the a Subunit

The cDNA encoding the a subunit was first cloned from rat brain in 1991 (Perin et al., 1991) and this protein was initially described as the 'large accessory' subunit of V-ATPase. Its presence in kidney, however, was once disputed (Gillespie et al., 1991; Gluck and Caldwell, 1987). This subunit is the largest protein within the V-ATPase complex and is glycosylated in mammalian V-ATPases, but the site of glycosylation has not yet been identified (Apps et al., 1989). The a subunit has a bipartite structure containing an amino-terminal hydrophilic domain of about 45 kDa and a carboxyl-terminal hydrophobic domain of about 55 kDa, which contains multiple putative TM helices and a small soluble C-terminal tail (Perin et al., 1991).

So far, different opinions on membrane topology of the a subunit, mainly focused on the numbers of TM regions, have been reported. The overall topology of the yeast Vph1p is that the N-terminal domain is located at the cytoplasmic side of the membrane, whereas a significant portion of its C-terminal domain is on the luminal side of the membrane (Arai et al., 1988; Leng et al., 1999). Evidence for

the intracellular location of the N-terminal domain is not only from the above studies, but also from the interactions identified between this domain and several other $V_1$ subunits of V-ATPase and a cytosolic protein (Landolt-Marticorena et al., 2000; Lu et al., 2003; Xu et al., 1999). However, the question is how many TM helices are within the C-terminal hydrophobic domain. A commonly accepted description is that there are 6-9 TM helices regardless of species. In other words, the C-terminal soluble tail, after the last TM domain, is either located intracellularly (i.e. with 6 or 8 TM helices) or extracellularly/or luminally (i.e. with 7 or 9 TM helices). A model for yeast Vph1p revealed, using cysteine-scanning mutagenesis and chemical labelling (by membrane permeant and impermeant sulfhydryl reagents) techniques, has 9 TM regions, i.e. the C-terminal tail faces the luminal side of the vacuole (Leng et al., 1999). However, such a model has not been verified in any higher eukaryotes. Nevertheless, data suggesting 6 or 8 TM helices, with cytoplasmic orientation of the C-terminal tail, have also been reported in yeast and *Dictyostelium discoiderm* (Clarke et al., 2002; Urbanowski and Piper, 1999). Both of the studies employed a green fluorescent protein (GFP) tag which was attached to the C-terminus of the a subunit. The results showed a cytoplasmic orientation of the GFP signal. Nevertheless, as the crystal structure of the a subunit has yet to be solved, it is hard to determine definitively whether the N- and C-termini lie on the same or opposite sides of the membrane. Thus, more evidence is required to determine the orientation of the C-terminal tail of the a subunit. The 6-TM model based on computerised analysis (Smith et al., 2000) was used in this project for protein engineering. The predicted two-dimensional (2D) model of the a4 contains a large N-terminal cytoplasmic domain, six TM regions, five loops (three luminal and two cytoplasmic), and one small cytoplasmic C-terminal tail (Table 2, Figure 2).

| Symbol | Region | *Mutations |
|---|---|---|
| a4 | 1-840 (840 **aa) | |
| N-terminal domain [a4(N)] | 1-393 (393 aa) | V35fsX, E82fsX87, intron 6 acceptor splice AG→AA, L103fsX139, N113fsX117, G175D, R194X, K237del, Q276fsX, Q358X, |
| 1st TM | 394-410 (17 aa) | P395fsX407 |
| 1st Loop | 411-449 (39 aa) | R449H |
| 2nd TM | 450-466 (17 aa) | |
| 2nd Loop [a4(Loop2)] | 467-544 (78 aa) | Y502X, P524L |
| 3rd TM | 545-561 (17 aa) | |
| 3rd Loop | 562-579 (18 aa) | Intron17donor splice GT→AT(X565) |
| 4th TM | 580-596 (17 aa) | |
| 4th Loop | 597-640 (44 aa) | S611fsX648 |
| 5th TM | 641-657 (17 aa) | |
| 5th Loop | 658-778 (121aa) | Q753X, V778fsX788 |
| 6th TM | 779-795 (17 aa) | |
| C-terminal tail [a4(C)] | 796-840 (45 aa) | R807Q, G820R, X841Q |

**Table 2 Individual regions of human a4 subunit.**

*The mutations identified in the human a4 gene from patients with rdRTA (Smith et al., 2000; Stover et al., 2002). fs = framshift; del = deleted. **aa = amino acid.

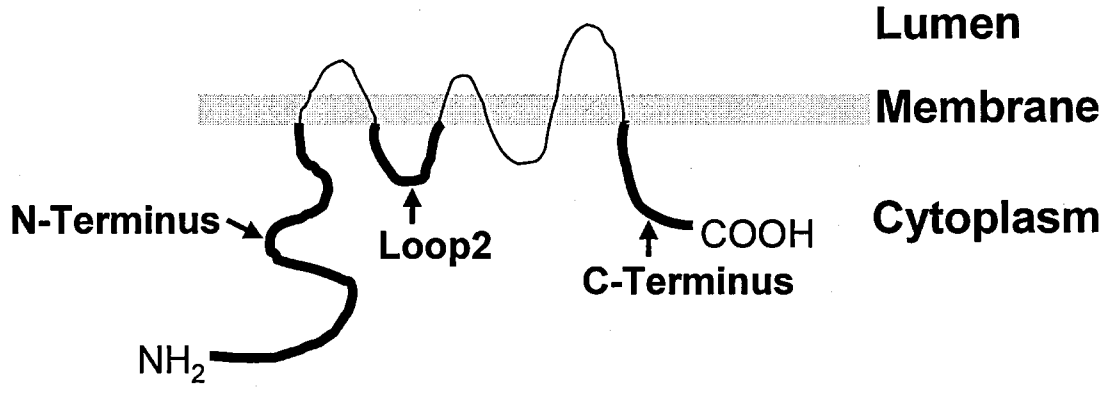



**Figure 2 A Schematic 2D diagram of the human a4 subunit (not to scale).**

The three highlighted regions (N-terminus, Loop2 and C-terminus) of human a4, were chosen for characterisation in this project.

### 1.3.2 Functions of the a Subunit

In contrast to all the other V-ATPase subunits in yeast where null mutations cause growth dependence on acidic media (approximately pH 5.5-6.5) (Munn and Riezman, 1994; Nelson and Nelson, 1990), the resulting phenotype from null mutations in each of the *VPH1 and STV1* encoding yeast a subunits (*VPH1*Δ or *STV1*Δ) is different. These can grow in medium buffered to pH 7.5, similar to wild-type strains (Manolson et al., 1992). This incomplete Vma$^-$ growth phenotype implies compensation from the second a subunit. Interestingly, overexpression of *STV1* in a double-deletion strain (*VPH1*Δ*STV1*Δ) showed significant localisation of Stv1p in the vacuoles rather than only in its normal locations (Golgi and endosomes) (Kawasaki-Nishi et al., 2001c; Manolson et al., 1994). This observation may also provide some evidence that Stv1p could possibly compensate the defect of Vph1p in the *VPH1*Δ strain. However, disruption of both genes (double mutant) revealed a typical Vma$^-$ growth phenotype similar to all the other V-ATPase null mutations. The specific null phenotype of Vph1p made it possible to study functions of the a subunit by amino acid replacement using strategies such as site-directed and random mutagenesis. From the mutagenesis studies, it was found that several charged residues in the C-terminal hydrophobic domain of Vph1p, in particular E789Q, led to a significant loss of ATP-dependent proton transport without affecting stability or assembly of V-ATPase complex. Also mutations of either Leu739 or Leu46 to serine led to almost complete loss of proton-transport activity of V-ATPase (Leng et al., 1998; Leng et al., 1996). In addition residue Arg735 was found to be essential for proton translocation (Kawasaki-Nishi et al., 2001b). Changing this residue to asparagine, glutamic acid or glutamine results in a fully assembled holoenzyme with a total absence of both proton transport and ATPase activities. Furthermore, a monoclonal antibody

directed against the extracellular domain of the a subunit inhibits endosomal acidification, supporting an essential role of this subunit in proton translocation activity of V-ATPase (Sato and Toyama, 1994).

Besides proton translocation, the a subunit was also found to be involved in assembly and targeting as well as coupling of ATP hydrolysis and proton transport of V-ATPase. Manolson *et al.* reported that disruption of the a subunit in yeast affects assembly of $V_1$ onto the vacuolar membrane (Manolson et al., 1992). Also, Kawasaki-Nishi *et al.* have investigated functions of the N- and C-terminal domains of yeast a subunit (Kawasaki-Nishi et al., 2001a). Using chimeras constructed with either domain, they have found that the N-terminal domain of the a subunit appeared to control targeting of V-ATPase, whereas the C-terminal domain affects coupling of ATP hydrolysis and proton transport. In addition, as mentioned earlier, an important regulatory mechanism for controlling V-ATPase activity identified in yeast is the reversible dissociation of the $V_1$ and $V_0$ domains in response to glucose depletion (Kane, 1995). This *in vivo* dissociation process is also likely to be controlled by signals located in the N-terminal domain of a subunit (Kawasaki-Nishi et al., 2001c). Furthermore, a cluster of five mutations was identified between residues 800 and 814 (yeast numbering), in the C-terminal soluble segment, which affected either assembly or stability of V-ATPase complex (Leng et al., 1998). Two of these mutations may also affect targeting of the a subunit (Leng et al., 1998; Leng et al., 1996). Interestingly, many, if not all, of the residues identified in the yeast homologue, Vph1p, important for activity, assembly, targeting or coupling of V-ATPase are conserved among the a subunit orthologues and paralogues in mouse and humans (Smith et al., 2001). Taken together, all these studies suggest that in yeast at least, the a subunit of the proton

pump plays crucial roles not only in proton translocation, but also in assembly, targeting and coupling efficiency of V-ATPase complex.

In addition, associations of the a subunit with certain $V_1$ subunits are likely to contribute to the formation of the peripheral stator, providing a structural support for the V-ATPase as described in Section 1.3.6. Moreover, evidence has been obtained suggesting that the inhibition of V-ATPase activity by bafilomycin and concanamycin are through binding of these chemicals to the a subunit (Bowman et al., 1988; Zhang et al., 1994), although binding sites of these two inhibitors are also reported on the c subunit (Bowman and Bowman, 2002; Huss et al., 2002). Finally, in recent years, it has become evident from the studies of human diseases that the presence of the a subunit is essential for normal pump function at the cell surface of renal ICs (a4) and osteoclasts (a3) (Frattini et al., 2000; Kornak et al., 2000; Smith et al., 2000). Defects in the genes encoding these proteins are associated with the recessively inherited human disease rdRTA and infantile malignant osteopetrosis, respectively, underscoring the functional importance of this subunit in kidney and bone.

### 1.3.3 Comparison of the a Subunits between the V-ATPase and F-ATPase

The a subunit in $F_0$ is a very hydrophobic protein containing 271 amino acid residues (*E. coli* numbering, approximately 25-30 kDa in bacteria and eukaryotes) of which the membrane topology is still unclear. However, a model based on the site-dependent accessibility of cysteine introduced into extramembrane loops of this protein was proposed to contain a N-terminal domain facing extracellularly, five TM α-helices and a C-terminus facing the cytoplasmic surface (Hatch et al., 1995; Valiyaveetil and Fillingame, 1998). The a subunit contains a critical basic residue, Arg210 (*E. coli* numbering) for proton translocation which is coupled to

ATP synthesis or hydrolysis (Valiyaveetil and Fillingame, 1997). The a subunit in $V_0$ is significantly larger in size (approximately 70 kDa in bacteria and 100 kDa in eukaryotes) than that of $F_0$ and no sequence homology is found between the $V_0$ and $F_0$ a subunits (Yokoyama et al., 2003a). The $V_0$ a subunit is likely to be composed of a large intracellular N-terminal domain, 6-9 TM helices, and a small soluble C-terminal tail as described in Section 1.3.1. It also contains some critical residues which are important for proton pumping activity without affecting assembly and ATP hydrolysis of the enzyme as described in Section 1.3.2. The latter is similar to those changes observed in *E. coli* $F_0$ a subunit (Cain and Simoni, 1988). This has prompted researchers to suggest that these residues play a fundamental role in the mechanism of proton translocation, and that the a subunit forms part of the proton-conducting pathway through the membrane (Finbow and Harrison, 1997). Hence, the functional studies mentioned above have suggested that they do in fact behave analogously, although there is no obvious sequence or structural homology between the a subunits of V- and F-ATPases (Bowman et al., 1988; Leng et al., 1996; Zhang et al., 1994).

Finally, direct interactions between subunits a and c as well as a and b within the $F_0$ domain have been observed (Jiang and Fillingame, 1998; Stalz et al., 2003). Although no evidence for such protein-protein interactions has been reported between the counterparts of these subunits within the $V_0$ domain, a very recent study revealed direct interacting helical surfaces of the TM segments of the subunits a and c' (Kawasaki-Nishi et al., 2003a). As mentioned earlier, the c' subunit, together with the subunits c and c", form the c ring which contributes to a main part of the proton transporting machinery.

### 1.3.4 The a Subunit Orthologues and Paralogues

#### 1.3.4.1 Tissue Specificity

The heterogeneity of the a subunit has been widely reported. Various a subunit paralogues show tissue specificity in different species from yeast to human. In yeast, two distinct genes, *VPH1* and *STV1*, have been identified to encode the 100 kDa a subunit and their amino acid sequences share approximately 54% identity (Manolson et al., 1992; Manolson et al., 1994). The Vph1p-containing V-ATPases are located on the vacuolar membrane, while the Stv1p-containing complexes are localised in endosomes and/or Golgi (Kawasaki-Nishi et al., 2001c; Manolson et al., 1994). In *C. elegans*, four genes (*vha-5, vha-6, vha-7, and unc-32*) coding for the a subunit paralogues have been identified (Oka et al., 2001b; Pujol et al., 2001). Transgenic and immunofluorescence analysis revealed that these genes were strongly expressed in distinct cells: *vha-5* in an H-shaped excretory cell, *vha-6* in intestine, *vha-7* in hypodermis, whereas *unc-32* was expressed in nerve cells. Furthermore, the *vha-7* and *unc-32* genes were also expressed in the uteri of hermaphrodites.

In addition, multiple paralogues of the a subunit have been identified in chicken (a1, a2, a3) and cow (a1, a2) (Mattsson et al., 2000; Peng et al., 1999). Among these paralogues, a1 and a2 have the strongest mRNA expression in brain, whilst the a3 is expressed predominantly in bone and liver. In both mouse and man, different a subunit paralogues (a1-a4) are encoded by at least four genes (Oka et al., 2001a; Smith et al., 2001; Smith et al., 2000). Among these a subunit paralogues, *ATP6V0A1* (previously known *ATP6N1A*) encodes a1, a paralogue that is ubiquitously expressed; *ATP6V0A2* encodes a2 (previously known as TJ6), a paralogue that was originally identified in mouse and thought to have a potential immune regulatory role (Lee et al., 1990); *TCIRG1* encodes a3

(previously known as OC116), a paralogue that is expressed specifically in osteoclast, a shorter transcript of a3 (known as TIRC7), which is T-cell specific was also identified (Heinemann et al., 1999; Li et al., 1996) and *ATP6V0A4* (previously known *ATP6N1B*) encodes a4, a paralogue that is expressed predominantly apically in acid secreting cells in the distal nephron of both human and mouse kidney, as well as in male genital tract and the cochlea within the human inner ear (Oka et al., 2001a; Smith et al., 2001; Smith et al., 2000; Stover et al., 2002). A more detailed localisation of the a4 subunit in the nephron is described in Section 1.3.5.2.3.

### 1.3.4.2 Sequence Divergence

Interestingly, sequence divergence between the paralogues of a subunit is more than would be expected due to species differences in vertebrates (Smith et al., 2001). In other words, the difference in sequence is greater between the paralogues than those of the orthologues. The sequence comparison between chicken, mouse, cow and human a subunit paralogues and orthologues are listed in Table 3. It can be seen from this table that approximately 50% amino acid sequence identity has been observed between the human a subunit paralogues (highlighted in blue), which is a similar level to the two yeast a subunit paralogues. However, over 83% sequence identity is found between the a subunit orthologues of chicken, mouse, cow compared to human (highlighted in red), with exception of a2 and a3 in chicken and human (highlighted in green), respectively.

The sequence divergence between the a subunit paralogues of V-ATPase may contribute to the differential targeting as well as other regulatory properties of V-ATPase (Kawasaki-Nishi et al., 2001c; Nishi and Forgac, 2000; Toyomura et al., 2000).

| | a1 (human) | a2 (human) | a3 (human) | a4 (human) |
|---|---|---|---|---|
| a1 (chicken) | 91 | 52 | 47 | 60 |
| a1 (mouse) | 95 | 53 | 47 | 61 |
| a1 (bovine) | 96 | 52 | 48 | 61 |
| a1 (human) | 100 | 53 | 47 | 61 |
| a2 (chicken) | 53 | 71 | 50 | 50 |
| a2 (mouse) | 53 | 91 | 51 | 51 |
| a2 (bovine) | 52 | 92 | 50 | 51 |
| a2 (human) | 53 | 100 | 49 | 52 |
| a3 (chicken) | 51 | 54 | 64 | 51 |
| a3 (mouse) | 47 | 51 | 83 | 47 |
| a3 (bovine) | - | - | - | - |
| a3 (human) | 47 | 49 | 100 | 49 |
| a4 (mouse) | 61 | 53 | 47 | 85 |
| a4 (human) | 61 | 52 | 49 | 100 |

**Table 3 Sequence comparison.**

Comparison of amino acid sequence identities among the a subunit paralogues and orthologues of the V-ATPase in certain vertebrates (chicken, mouse, bovine and human) (Smith et al., 2001).

### 1.3.5 The Human a4 Subunit

#### 1.3.5.1 Discovery of the *ATP6V0A4* Gene

In 1999, Karet *et al.* analysed 31 kindreds with autosomal rdRTA using a genome-wide linkage analysis method to identify genes whose mutations are responsible for disease in these cases. From this, mutations were identified in *ATP6V1B1* (encoding the V-ATPase B1 subunit) in approximately 2/3 of the analysed kindreds with rdRTA and sensorineural hearing loss (SNHL). Interestingly, the B1-encoding gene was observed to be expressed in the cochlea and endolymphatic sac of mouse, which may explain why these patients are deaf. (Karet et al., 1999b).

However, 13 of the initial kindreds with rdRTA and normal hearing were not accounted for by mutations in the B1 gene (Karet et al., 1999a). In order to identify loci linked to this trait, a further genome-wide linkage screen of these kindreds was

performed and from this a new locus, named *rdRTA2*, was identified for this trait. The *rdRTA2* locus was defined to be in a 14-cM (the LOD-3 support interval) or 11-cM (LOD-1 support interval) region of 7q33-34. However, 4 of the 13 kindreds were not linked to the *rdRTA2* locus (Karet et al., 1999a).

What is the actual disease-causing gene in the *rdRTA2* locus? In 2000, Smith *et al.* performed further analysis of those 9 kindreds linked to the *rdRTA2* locus, by performing homozygosity mapping using additional closely linked polymorphic markers within the 11-cM region. This analysis resulted in the identification of *ATP6N1B* (now known as *ATP6V0A4*) (Smith et al., 2000). *ATP6V0A4* is composed of 23 exons with the initiation codon occurring in exon 4. Altogether 8 homozygous mutations (premature termination, frameshift, missense substitution and splice site mutations) (Table 2) were identified in *ATP6V0A4* from these patients, all of which are predicted to disrupt the encoded protein and consequently affect the function of the protein (Smith et al., 2000). Recently Stover *et al.* (2002) screened more cases with rdRTA, from which an additional 9 novel mutations (Table 2) were identified in *ATP6V0A4*. Mutations identified in the a4-encoding gene from the above studies are spread throughout the gene. Interestingly, it was also found that several *ATP6V0A4*-linked patients have developed hearing loss. However, in contrast to those cases associated with *ATP6V1B1* with severe hearing loss in childhood, the identified hearing loss in *ATP6V0A4*-linked patients is usually in young adulthood. The loss of hearing can probably be attributed to the expression of *ATP6V0A4,* at the mRNA level at least, within the cochlea of human inner ear (Stover et al., 2002). Nevertheless, it is not clear why the hearing loss appears at an earlier age for those patients with *ATP6V0B1* mutations than those with *ATP6V0A4* mutations.

### 1.3.5.2 a4 in the Kidney

#### 1.3.5.2.1 The Kidney Nephrons and Renal Acid-base Homeostasis

The kidney contributes to maintaining a stable extracellular environment for all body cells. In order to control water and ionic balance, the kidney performs many important functions, including regulating the excretion of many substances ($H_2O$, $Na^+$, $K^+$, $Cl^-$, $HCO_3^-$, $Ca^{2+}$, $Mg^{2+}$ and phosphate etc.) and acid-base homeostasis.

The kidney can be divided into an outer zone (cortex) and an inner zone (medulla). The basic functional unit in the kidney is the nephron (Figure 3) and each kidney contains approximately 400,000 to 800,000 nephrons. A nephron consists of the glomerulus connected to a tubule system that is composed of the proximal tubule, loop of Henle, distal convoluted tubule and collecting duct. The collecting duct has three sections, named according to their depth in the kidney: the cortical collecting duct (CCD), the outer medullary collecting duct (OMCD), and the inner medullary collecting duct (IMCD). The glomerulus functions as a filtration barrier for the formation of urinary filtrate, which is then modified mainly with respect to its volume and concentration in the tubules by reabsorption and excretion of many substances such as those mentioned above. The majority of reabsorption occurs in the proximal tubule. For example, approximately 65% of filtered $Na^+$ is reabsorbed from the urinary filtrate in the proximal tubule. The second highest reabsorption (25-30%) takes place in the loop of Henle, whereas the distal nephron, including distal convoluted tubule and collecting ducts, contributes to less than 10% of the reabsorption.

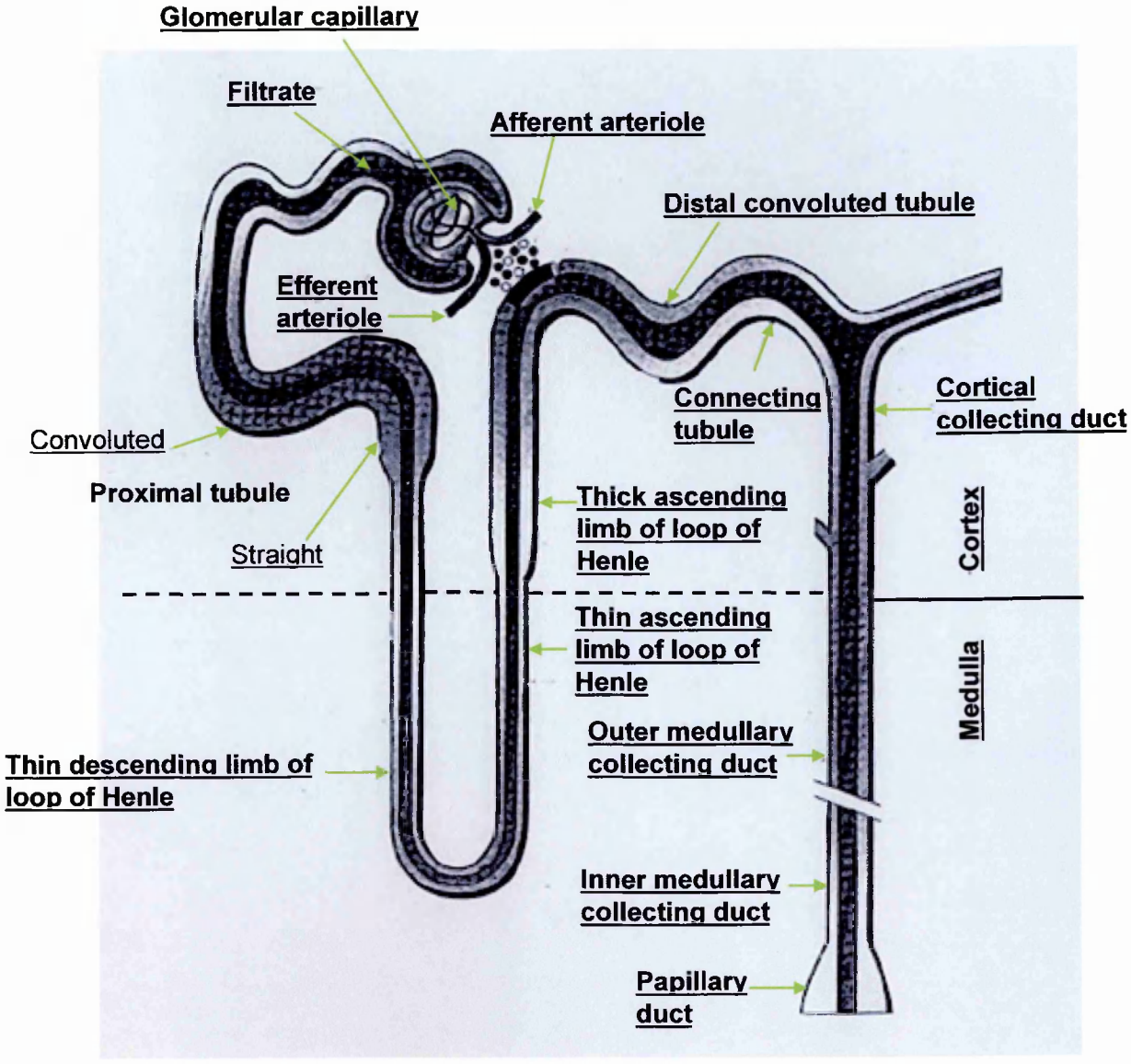
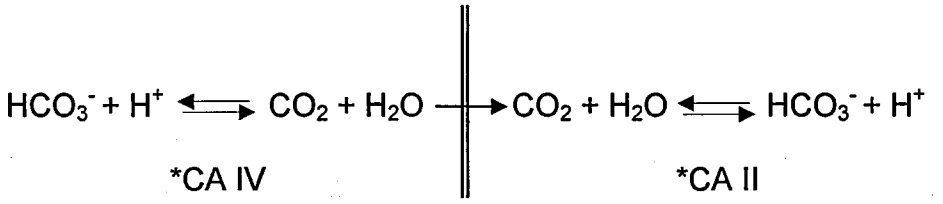



**Figure 3 Structure of a nephron.**

A schematic diagram of a Nephron (not to scale) and collecting system of the kidney. This figure is taken from reference (O'Callaghan and Brenner, 2000).

In adult humans, catabolism of an omnivorous diet generates approximately 70 mmol $H^+$/per day (Penney and Oleesky, 1999). The kidney must excrete this $H^+$ load into the urine to preserve acid-base homeostasis of the body. Buffers in the urine can help to regulate the concentration of $H^+$. Different buffer pairs, buffer acid and buffer base, within the body are in equilibrium with each other. In other words, a change of one system will change body pH, which in turn will alter the ratio of buffer acid and buffer base in other systems. The major independent urinary buffer is sodium phosphate ($Na_2HPO_4/NaH_2PO_4$) and the main extracellular buffer is the bicarbonate system ($H^+/HCO_3^-$), which can be altered by changing the concentration of $HCO_3^-$ for example.

Through altering the concentration of $HCO_3^-$ and excretion of $H^+$, the kidney plays a vital role in maintaining acid-base homeostasis. Of the filtered $HCO_3^-$, approximately 95% is reabsorbed along the nephron to maintain normal plasma $HCO_3^-$ concentration and therefore plasma pH. Approximately 80% of the $HCO_3^-$ is reabsorbed in the proximal tubule, and this reabsorption depends on the secretion of $H^+$ into the filtrate in the lumen. In this region, $H^+$ secretion is mainly through the $Na^+/H^+$ exchanger, but also the V-ATPase (Chan and Giebisch, 1981). In the filtrate, the secreted $H^+$ interacts with $HCO_3^-$ (under catalysis of CA IV) (Bastani and Gluck, 1996) to form $CO_2$ and $H_2O$, which then diffuse back into the plasma where they reform as $H^+$ and $HCO_3^-$ (under catalysis of CA II). The reformed $HCO_3^-$ is secreted into the blood, via a $Cl^-/HCO_3^-$ exchanger and a $Na^+/HCO_3^-$ co-transporter, and the $H^+$ enters the next run of this cycle (Penney and Oleesky, 1999). The intact interaction equation is shown below:




Where, CA IV *and* CA II are carbonic anhydrase enzymes which catalyse this reaction in lumen and cytoplasm, respectively.

Therefore, $H^+$ secretion here mainly contributes to reabsorption of $HCO_3^-$, but not to net acid excretion. However, a small amount of net acid excretion does occur in this part of the nephron, in the form of $NH_4^+$ produced from the interaction of $H^+$ and $NH_3$.

However, in the distal parts of the nephron, especially the connecting tubule and collecting ducts, the secretion of $H^+$ functions not only in the reabsorption of any remaining $HCO_3^-$, but also as net acid excretion. For net acid excretion, $H^+$ is pumped out from the cell into the lumen to enter the urinary filtrate by V-ATPases residing at the apical surface of α-ICs (see next section for details). In the filtrate, pumped $H^+$ is either associated with $HPO_4^{2-}$ to form $H_2PO_4^-$ in the urinary buffer system, or with $NH_3$ to form $NH_4^+$; both are excreted in the urine (Kurtzman, 1990; Penney and Oleesky, 1999). In reality, this urinary acidification is tightly regulated in order to maintain acid-base homeostasis.

#### 1.3.5.2.2 α-Intercalated Cells

The transporting epithelia of the distal nephron contain a group of cells which share certain common features, including large numbers of mitochondria, high levels of cytosolic CA II and membrane associated V-ATPases (Brown and

Breton, 1996; Kim et al., 1990). These cells are known as intercalated cells (ICs). At least two types of IC, named α and β, are found in the mammalian nephron distal tubule and CCD where they together make up approximately 40% of the total number of cells (Al-Awqati, 1996; Brown and Breton, 1996). However, only α-ICs are found in the kidney medulla, including both OMCD (accounting for ~40% of the epithelial cell population) and IMCD (accounting for ~10% of the epithelial cell population) (Schuster, 1993).

The α-IC (Figure 4) has a high density of V-ATPases on its apical plasma membrane as well as stored in specialised intracellular tubulovesicular compartment located near the cell surface. In response to a drop in arterial pH, the apical V-ATPase is activated and pumps protons into the urine. This drop in blood pH also triggers rapid recruitment and insertion of stored V-ATPases to the apical plasma membrane of α-IC to increase $H^+$ secretion (Bastani et al., 1991; Tisher et al., 1991). On the same side, a P-type ATPase ($H^+/K^+$) is also present which might contribute to a lesser extent to luminal $H^+$ secretion under certain circumstances, such as dietary $K^+$ depletion (Silver and Soleimani, 1999). In addition to the apical proton pump, this cell has a basolateral $HCO_3^-/Cl^-$ exchanger (AE1), which secretes bicarbonate ions into the blood (for reabsorption and buffering) in exchange for $Cl^-$ (Alper et al., 1989; Bastani and Gluck, 1996). Proper α-IC function is important not only for the maintenance of the body acid-base homeostasis but also for the solubility of calcium in urine as well as for the stability of calcium in bone. Functional failure of α-IC function results in metabolic acidosis, a condition characterised by dRTA as described later in Section 1.3.5.3.

In contrast to the α-IC, the β-IC (Figure 4) has a $HCO_3^-/Cl^-$ exchanger located at its apical surface which secretes $HCO_3^-$ into the lumen and a V-ATPase at its basolateral side to secrete $H^+$ into the blood (Bastani and Gluck, 1996).

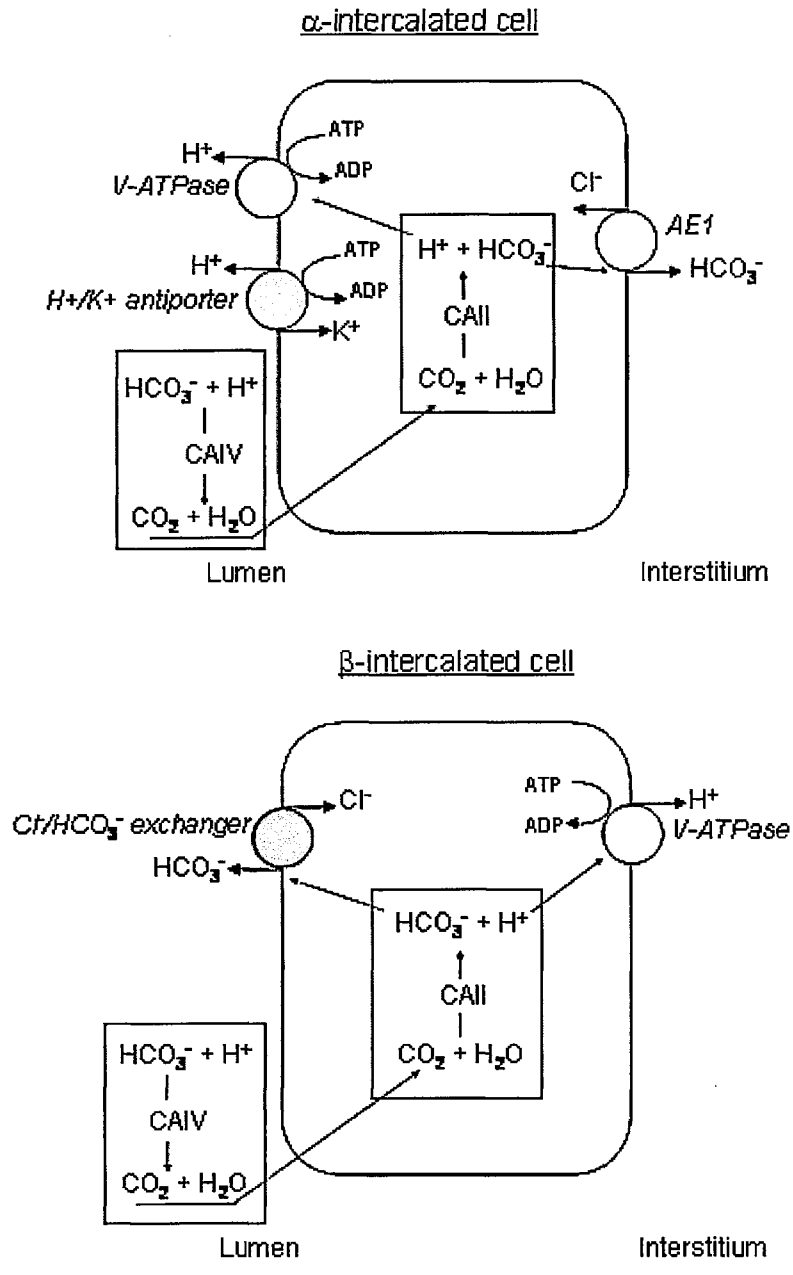



**Figure 4 α- and β-intercalated cell.**

V-ATPases pump protons from the cytoplasm into the lumen (α-IC) or blood (β-IC). In the cytoplasm, protons are produced together with $HCO_3^-$ from interaction between $CO_2$ and $H_2O$, catalysed by CA II. The resulting $HCO_3^-$ is reabsorbed into the blood through the AE1 (α-IC) or secreted into lumen by the anion exchanger ($Cl^-/HCO_3^-$) (β-IC).

#### 1.3.5.2.3 Localisation of V-ATPase and a4 within the Nephron

The plasma membrane V-ATPase makes a major contribution to $H^+$ secretion in several nephron segments (Bastani et al., 1994; Bastani et al., 1997; Brown et al., 1992b; Chan and Giebisch, 1981). The expression of plasma membrane V-ATPases has been found in the proximal tubule, the loop of Henle, the connecting tubule, CCD, OMCD and IMCD (Bastani et al., 1991; Brown et al., 1988). As mentioned in the above section, the α-IC has a high density of V-ATPases at its apical membrane, whereas the β-IC is able to insert V-ATPases into its basolateral membrane.

Different observations of the localisation of the a4 subunit in nephron segments have been reported. In 2000, Smith *et al.* observed, for the first time, that localisation of a4 was restricted to the apical surface of α-ICs in the collecting ducts of the human nephron. This experiment was performed using an antibody raised against the last 14 amino acids of the human a4 subunit (Smith et al., 2000). Later, in 2001, Oka *et al.* observed that the a4 subunit is localised to both α-ICs and β-ICs in the distal portion of mouse nephron (Oka et al., 2001a), which the Karet group has reported in the same year (Smith et al., 2001). In addition, a4 staining was observed, to a lesser extent, at the base of the brush border of proximal cells in the initial part of the proximal tubule of mouse nephron (Smith et al., 2001). Very recently, Stehberger *et al.* observed immunostaining for a4 in various nephron segments in both human and mouse kidney (Stehberger et al., 2003). These differences seen in experimental data could possibly have resulted from using different techniques for antigen retrieval or tissue fixation.

### 1.3.5.3 Recessive Distal Renal Tubular Acidosis

As discussed above, the kidney must excrete the $H^+$ generated from the catabolism of an omnivorous diet into the urine to maintain acid-base homeostasis. The pH range of urine has between 4.5 and 8.5. In humans, this urinary acidification is mainly carried out by the apical proton pump of the α-IC in the distal portion of the nephron as described in Sections 1.3.5.2.1 and 1.3.5.2.2. Failure or inadequate acid excretion results in type 1 distal renal tubular acidosis (dRTA). In RTA, the metabolic acidosis is usually accompanied by hypokalemia and hypercalciuria. In untreated cases, the metabolic acidosis leads to other conditions such as nephrocalcinosis and osteomalacia (Barzel and Hart, 1973; Bastani and Gluck, 1996; Courey and Pfister, 1972). Early diagnosis and adequate alkali treatment can reverse the metabolic acidosis, which results in improved growth and bone cell function (Domrongkitchaiporn et al., 2002).

dRTA can be inherited in two different forms: autosomal dominant or autosomal recessive. In the case of autosomal dominant dRTA (ddRTA) (MIM 179800), nephrocalcinosis is a common feature whereas osteomalacia is usually less prominent. Mutations in the AE1-encoding gene (*SLC4A1*) are responsible for ddRTA in all kindreds reported so far (Bruce et al., 1997; Jarolim et al., 1998; Karet et al., 1998). However, recessive dRTA (rdRTA) is rarely the result of mutations in this gene (Karet et al., 1998), although there have been a few reported cases of rdRTA caused by AE1 mutations in South Asia (Tanphaichitr et al., 1998). ddRTA normally presents in adulthood and often has milder symptoms compared to rdRTA. rdRTA has historically distinguished into two subtypes: rdRTA with progressive bilateral sensorineural deafness (OMIM 267300) and rdRTA with preserved hearing (OMIM 602722). rdRTA normally presents at a very young age (infancy or early childhood).

Both subtypes of rdRTA are associated with mutations in tissue specific subunit-encoding genes of the V-ATPase in the collecting duct of the nephron. As mentioned earlier in Section 1.3.5.1, mutations in the B1-encoding gene are associated with rdRTA with deafness, whereas mutations in the a4-encoding gene are associated with rdRTA with milder hearing loss.

### 1.3.6 Interactions of the a Subunit

Despite the importance of the a subunit to the proton pump, not much is known about its protein interactions. However, some evidence regarding this aspect has been provided very recently as listed below. Using crosslinking techniques (with crosslinking reagent MBP), Xu *et al.* have provided evidence for possible interaction between the E and a subunits, although it has not been confirmed by alternative assays (Xu et al., 1999). In addition, Landolt-Marticorena *et al.* demonstrated through a yeast two-hybrid assay that the N-terminal domain of Vph1p interacts directly with the A subunit and this interaction was further confirmed by a co-immunoprecipitation analysis (Landolt-Marticorena et al., 2000). In the same study, they also demonstrated an *in vitro* interaction between the N-terminal domain of Vph1p and the H subunit (Landolt-Marticorena et al., 2000). Because the a subunit forms part of the proton translocation machinery, the presence of interactions between this subunit and certain $V_1$ subunits supports the a subunit involvement in coupling ATP hydrolysis to proton translocation. On the other hand, these associations are likely to contribute to structural support within the enzyme, making the a subunit likely candidate for the formation of part of the peripheral stalk of V-ATPase. In addition to interacting with some $V_1$ subunits, a direct contact between the a and c' subunits was reported through structural analysis (Kawasaki-Nishi et al., 2003a). Also, a possible interaction between the a

and d subunits was suggested from the observation that the d subunit was retained in cytosol in yeast cells lacking Vph1p (Graham et al., 2000).

In addition to interacting with the component elements of V-ATPase, the a subunit was also found to interact with other chemicals or proteins. As mentioned earlier, binding of bafilomycin to the a subunit inhibits V-ATPase activity (Bowman et al., 1988; Zhang et al., 1994). Very recently, Lu *et al.* demonstrated interaction between the N-terminal domain of human a4 and aldolase, a glycolytic enzyme, through yeast two-hybrid library screening (Lu et al., 2003).

## 1.4 The Aims of the Present Study

I am particularly interested in the a4 homologue of the a subunit of V-ATPase because of its essential contribution to renal acid-base homeostasis. However, information concerning potential functional or regulatory contributions of this tissue-specific paralogue of the mammalian V-ATPase a subunit is scarce. The main aim of this project was to characterise the human a4 subunit focusing on protein-protein interactions which can help us to determine the biological functions of this protein, by situating it relative to other proteins in cellular pathways or functional classes. The plan was first to identify potential binding partner(s) of the human a4 through library screening using such techniques as Phage display and/or yeast two-hybrid systems. Possible interactions identified from the screens would be further characterised using other strategies such as pull-down, co-immunoprecipitation, and immunochemistry.

# CHAPTER 2

# MATERIALS AND METHODS

## 2.1 Chemicals, Enzymes and Antibodies

Chemicals of analytical grade or higher were purchased from Sigma, BDH/Merck or BioRad. Restriction endonucleases and modifying enzymes were purchased from New England Biolabs (NEB) and Taq DNA polymerase and PfuTurbo DNA polymerase from Roche and Stratagene, respectively. DNA purification kits were from Qiagen. All antibodies used in this study are listed in Appendix Table A.1.

## 2.2 General Buffers, Bacterial Growth Media and Antibiotic Solutions

General buffers (*PBS, PBST, TBE, TBS, TAE, TE*), bacterial growth media (*LB, 2x YT, SOC*), and antibiotic solutions were mainly obtained from the media room in Cambridge Institute for Medical Research building. Recipes are listed in Appendix Section A.1.

## 2.3 Cell Strains and Storage

All cell strains used in this project are listed in Appendix Table A.2.

Bacterial and yeast strains were stored as glycerol stocks. Overnight cultures were diluted by the addition of an equal volume (bacteria) or 1/3 volume (yeast) of sterile 50% glycerol, vortexed, chilled in dry ice and then stored at -80°C.

## 2.4 Plasmids

All plasmids used in this study are listed in Appendix Table A.4. Maps of bacterial expression vectors (pGEX-4T1 and mini-prseta-max), the *Drosophila*

expression vector (pMT/Bip/V5-His A) and yeast expression vectors (pSOS and pMyr) used in this project are presented in Appendix Figures A.1, A.2, and A.3, respectively.

## 2.5 DNA Analysis

### 2.5.1 Quantification of DNA Concentration and DNA Storage

DNA concentrations were determined by spectrophotometry using the GeneQuant II system (Pharmacia Biotech). Samples were diluted with double deionised water ($H_2O$) and transferred into a 0.5 mm quartz cuvette. Optical densities (OD) were taken at 260 nm and 280 nm. Plasmid DNA was kept in Tris-Cl (10 mM, pH 8.0) or TE buffer and stored at -20°C.

### 2.5.2 Agarose Gel Electrophoresis of DNA

Depending on the size of the DNA fragments (or plasmids) of interest, agarose electrophoresis gels typically contained 1-2% ultra pure electrophoresis grade agarose in 1x TAE buffer. Prior to casting, ethidium bromide was added to the agarose gel solution to give a final concentration of 0.5 μg/ml. DNA samples were mixed with 0.2x volume of agarose gel loading buffer [40% (w/v) sucrose, 0.25% xylene cyanol FF, and 1 mM ethylendiaminetetra-acetic acid (EDTA)] before eletrophoresis, which was carried out at 120 V in 1x TAE buffer. 1 kilobase (kb) DNA ladder and/or 100 basepair (bp) DNA ladder were loaded parallel with the samples in order to estimate the size of the samples. After eletrophoresis, DNA was visualised by exposure to UV light on a transilluminator.

## 2.6 Purification of DNA

### 2.6.1 Using Commercial Kits

Purification of plasmid DNA from bacterial cultures, Polymerase Chain Reaction (PCR) products from their reaction mixtures and DNA fragments from agarose gel slices were carried out using appropriate Qiagen Kits, unless stated otherwise, following manufacturer's guidelines.

### 2.6.2 Phenol/Chloroform Extraction

200-300 μl DNA solution was vortexed with an equal volume of phenol/chloroform/Isoamyl alcohol (IAA) (25:24:1 in volume) and centrifuged [11,000 x *g*, 2 min, room temperature (RT)] in a microfuge (Eppendorf). The supernatant was recovered and re-extracted at least two more times. This supernatant was then ethanol precipitated in order to recover the DNA.

### 2.6.3 Ethanol Precipitation

DNA in solution was precipitated by either the addition of 1/10 volume of NaAc (3 M, pH 5.2) or $NH_4Ac$ (4 M, pH 5.3), and 2-3 volumes of cold (-20°C) ethanol followed by incubation at -80°C for 30 min. DNA was recovered in the form of a pellet by *centrifugation (15,115 x *g*, 30 min 4°C). The DNA pellet was washed once with 70% ethanol, air-dried, resuspended in the appropriate buffer and stored at -20°C.

*All centrifugation steps were carried out with Beckman centrifuges, unless otherwise indicated.

## 2.7 Design, Synthesis and Storage of Oligonucleotides (Primers)

Primers for gene cloning and sequencing were designed to contain a GC content of higher than 50% if possible. Primer pairs used for gene cloning were designed to have similar melting temperatures. In addition, cloning primers were also designed to have further sequences at the 5' end, upstream of restriction sites in order to gain better digestion of DNA fragments produced using the primers. The melting temperature (Tm) value of primers was calculated using the formula: 4°C x CG content + 2°C x AT content, unless otherwise indicated. All primers were synthesised by Invitrogen Life Technologies and were dissolved in $H_2O$ to give a final concentration of 0.1 mM and then stored at -20°C. All primers used in this study are listed in Appendix Table A.3.

## 2.8 Polymerase Chain Reaction (PCR) Methods

PCR (Saiki et al., 1985) was carried out using a programmable MJ PTC-225 thermal cycle (Techne Ltd.). All PCRs included a negative control where DNA template was replaced by $H_2O$.

### 2.8.1 Standard PCR

A standard PCR was used to produce DNA fragments for cloning and for sequencing of positive clones obtained from yeast two hybrid assays. A reaction mixture was prepared as described in Table 4. PfuTurbo DNA polymerase was included in the reaction mixture, in addition to DNA polymerase, due to its proof-reading function. The cycling parameters used were as follows, unless otherwise stated: 1 cycle for 3 min at 95°C; 35 cycles of [95°C, 0.5 min; *X°C, 0.5 min; 72°C, **Y min] and 1 cycle for 10 min at 72°C, (*X = calculated Tm - 4°C and **Y depends on the length of the amplified region).

| | STANDARD PCR | COLONY PCR | TEMPERATURE CYCLING |
|---|---|---|---|
| **Reagent** | **Volume or amount** | **Volume** | **Volume or amount** |
| Template DNA | ~20 ng plasmid DNA | 3 μl colony DNA | 12, 24, or 60 ng plasmid DNA |
| *10x conc. PCR buffer | 2 μl | 2 μl | 5 μl |
| dNTPs (2 mM) | 2 μl | 2 μl | 2.5 μl (10 mM dNTPs) |
| Primer 1 (10 μM) | 1 μl | 1 μl | 1.5 μl |
| Primer 2 (10 μM) | 1 μl | 1 μl | 1.5 μl |
| Taq DNA polymerase (5 u/μl) | 0.1 μl | 0.1 μl | - |
| PfuTurbo DNA polymerase (2.5 u/μl) | 0.1 u | - | 2.5 u |
| $H_2O$ | **X μl | 10.9 μl | ***Y μl |
| **Total Volume** | **20 μl** | **20 μl** | **50 μl** |

**Table 4 Reaction mixtures for standard PCR, colony PCR and temperature cycling of site-directed mutagenesis.**

*The buffer contains 15 mM $Mg^{2+}$. **X and ***Y depend on the volume of template DNA.

### 2.8.2 Long-Range PCR (LR-PCR)

The Expand™ Long Template PCR System (Roche) was used to produce DNA fragments longer than 1 kb for cloning, following manufacturer's guidelines. Cycling parameters are shown in Table 5.

| Number of Cycles | Temperature | | Time |
|---|---|---|---|
| 1 | 92°C | | 2 min |
| 10 | 92°C | | 10 sec |
| | X°C | | 30 sec |
| | 68°C | | 3 min |
| 15 | 92°C | | 10 sec |
| | X°C | | 30 sec |
| | 68°C | (+20 sec/cycle for enlongation only) | 3 min |
| 1 | 68°C | | |

**Table 5 Cycle conditions for LR-PCR.**

*X depends on Tm of individual primers

### 2.8.3 Colony PCR

Colony PCR for screening of recombinant DNA clones was carried out as follows: a single colony from a bacterial agar growth plate was resuspended in 20 μl $H_2O$ and the resuspension was heated at 95°C for 10 min to release bacterial colony DNA. 3 μl of the denatured cell sample was used in a 20 μl PCR reaction mixture (Table 4). Cycle conditions were the same as described in Section 2.8.1.

## 2.9 Site-Directed Mutagenesis

The QuikChange™ Site-Directed Mutagenesis Kit (Stratagene) was used to carry out single amino acid changes.

### 2.9.1 Plasmid Preparation

Plasmids were purified using Qiagen Qiaprep® Midiprep Kit and used as a template vector. Mutagenic primers were designed according to manufacturer's instructions, to create a specific single amino acid change upon expression,.

### 2.9.2 Temperature Cycling

Three reaction mixtures, using 12, 24 or 60 ng of the template plasmid were prepared (Table 4). Each of the reaction mixture was overlaid with 30 µl of mineral oil and then submitted to the reaction cycle. The cycling parameters used to create a specific single amino acid change were as follows: 16 cycles each of, 95°C for 0.5 min, 55°C for 1 min and 68°C for 12 min.

### 2.9.3 Digestion

Following Temperature Cycling, the products were incubated on ice for 2 min before being submitted to parental methylated DNA template removal by *Dpn*I restriction enzyme digestion. Digestion was carried out by the addition of 10 units of *Dpn*I directly to the reaction mixture under the mineral oil overlay followed by incubation at 37°C for 1 hr.

### 2.9.4 Transformation

Following digestion, 1 µl of the *Dpn*I-treated DNA was transformed into 20 µl XL-10 Gold chemical competent cells (Stratagene) using the same procedures as described in Section 2.10.5.2. Transformants were plated onto LA plates containing the appropriate antibiotic and grown overnight at 37°C.

## 2.10 Gene Cloning

### 2.10.1 Digestion of DNA by Restriction Endonucleases

Restriction endonuclease digestion of vector and insert DNA was performed according to the supplier's instructions. DNA was cut for 2-16 hrs at 37°C in 1x enzyme reaction buffer. After restriction digestion, enzymes were removed using either the Qiagen PCR Purification Kit or the Qiagen Gel Extraction Kit. Where possible, multiple digestions were performed, if the buffer conditions were compatible for both enzymes. Otherwise, serial digestions were performed.

### 2.10.2 Removal of 5' Phosphate Group (SAP Treatment)

Following the restriction enzyme digestion, 5' phosphate groups of vector DNA were removed using Shrimp Alkaline Phosphatase (SAP) (USB). 1 unit SAP per 5 μg DNA was added into the digestion mixture and incubated for 1 hr at 37°C, after which the enzyme was inactivated by heating at 65°C for 20 min.

### 2.10.3 Ligation of insert DNA Fragment into Vector DNA

Prior to ligation reaction, both vector and insert were digested with the same restriction enzyme(s) in order to join them at designed positions. All the ligation reaction mixtures were carried out in volumes of 20 μl and prepared as shown in Table 6. The molar ratio of DNA insert : vector was 3-8:1. The following equation was used to calculate the amount of insert needed. For example, if the molar ratio of insert : vector is 3:1, the amount of insert X (ng) required for 100 ng of vector:

$$\mathrm{X\,(ng)} = 3 \times \frac{100\,\mathrm{ng}}{\mathrm{size\ of\ vector\,(bp)}} \times \mathrm{size\ of\ insert\,(bp)}$$

Reaction mixture containing no insert DNA provided a negative control. The ligation reaction was carried out at 16°C for approximately 16-18 hrs.

| Reagent | Volume |
|---|---|
| Vector | 100 ng |
| Insert | X ng |
| 1X T4 DNA ligase buffer | 2 µl |
| T4 DNA ligase (20 u/µl) | 1 µl |
| $H_2O$ | *Y µl |
| **Total Volume** | **20 µl** |

**Table 6 Ligation reaction mixture.**

*Y depends on the volume of vector and insert used.

Where electroporation was employed for transformation after ligation, the ligation mixture was ethanol precipitated. The final DNA pellet was resuspended in 15 µl $H_2O$; 5 µl of this resuspension was then used for each electroporation.

### 2.10.4 Preparation of Bacterial Electro-competent Cells

LA plates being plated with bacterial cells contained, if necessary, appropriate antibiotic/s. The plates were incubated at 37°C for 16-18 hrs. A single colony from such a plate was used to inoculate 5 ml of LB medium containing the same antibiotic/s. This culture was incubated for 16-18 hrs at 37°C with vigorous shaking and used to inoculate 400 ml of the same medium, but omitting antibiotic/s. The culture was grown at 37°C until the $A_{600}$ reached 0.4 and then cooled on ice for 15 min before being centrifugated (4648 x *g*, 10 min, 4°C). The cell pellet was washed 3 times with firstly 400 ml, then 200 ml and finally 10 ml of ice-cooled $H_2O$. The final cell pellet was resuspended in 0.5 ml 10% glycerol and stored as small aliquots at -80°C after freezing on dry ice.

### 2.10.5 Transformation of Bacterial Cells with Bacteriai Plasmid DNA

#### 2.10.5.1 Transformatlon by Electroporation

A sample of frozen electro-competent cells was allowed to thaw slowly on ice. About 1 ng plasmid DNA was mixed with 40 μl of cells. This mixture was transferred to a chilled electroporation cuvette (2 mm) and subjected to electroporation using Gene Pulser (BioRad) under the following conditions: Resistance: 200 Ohms, Capacitance: 25 μF, and Voltage: 2.3 kV. Immediately, 500 μl of SOC medium was added to the mixture followed by incubation, without shaking, at 37°C for 1 hr. After incubation, the transformants were plated onto LA plates containing an appropriate antibiotic/s and then grown at 37°C for 16-18 hrs. This method typically yielded transformation efficiencies of 1 x $10^{9-10}$ transformants per μg plasmid DNA.

#### 2.10.5.2 Heat Shock Transformation

A sample of frozen chemical competent cells was allowed to thaw slowly on ice. About 1 ng plasmid DNA was mixed with 20 μl of cells and the mixture was incubated on ice for 2 min. Cells were then shocked by heating at 42°C for 30 sec followed by incubation in 600 μl of SOC medium for 1 hr at 37°C whilst shaking. After incubation, the transformants were plated onto LA plates containing the appropriate antibiotic/s and incubated at 37°C for 16-18 hrs. This method typically yielded transformation efficiencies of 1 x $10^{6-7}$ transformants per μg plasmid DNA.

### 2.11 DNA Sequencing

Automated DNA sequencing was carried out using an ABI 377 sequencing machine according to the standard ABI Prism® BigDye™ Terminator (BDT) protocol.

### 2.11.1 Sequencing Reactions

Sequencing reaction mixtures (Table 7) were prepared in microtitre plates, The cycling parameters used were: 25 cycles each of [96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min]. On completion, excess dye terminators in the sequencing product mixture were disabled by SAP treatment, which was performed in 1x SAP reaction buffer as described in Section 2.10.2.

| Reagent | Volume | Volume |
|---|---|---|
| DNA + $H_2O$ | 5.5 | 11 µl |
| BDT | 2 µl | 2 µl |
| 5X Sequencing buffer | 2 µl | 6 µl |
| Primer (10 µM) | 0.5 µl | 1 µl |
| **Total Volume** | **10 µl** | **20 µl** |

**Table 7 Sequencing reaction mixtures.**

### 2.11.2 Ethanol/Isopropanol Precipitation

Following SAP treatment, sequencing reaction products were precipitated using either ethanol or isopropanol, as follows: the reaction mixtures (Table 8) were prepared and incubated at RT for 15 min. They were then centrifuged (2413 x *g*, 30-40 min, 4°C) before being inverted onto tissues and centrifuged briefly to remove the supernatant. 70% ethanol was added to each pellet and centrifuged for 20 minutes under the same centrifuge conditions. After the removal of the supernatant, pellets were stored at -20°C for up to one month in the dried-state.

| Reagent | Volume | Volume |
|---|---|---|
| Sequencing reaction products | 10 µl | 20 µl |
| *Ethanol/NaAc mixture | 80 µl | - |
| Isopropanol | - | 60 µl |
| $H_2O$ | 20 µl | 20 µl |
| **Total Volume** | **110 µl** | **100 µl** |

**Table 8 Precipitation reaction mixtures.**

*Ethanol/NaAc mixture contains 89% (v/v) ethanol + 6.25% (w/v) NaAc (3 M, pH 5.2).

### 2.11.3 Sequencing Gel Electrophoresis and Result Analysis

A sequencing gel mixture (Table 9) was prepared and assembled onto the ABI 377 sequencing machine. Immediately prior to loading, 3 µl of loading dye (1 in 5 dilution, with formamide, of the ABI sequencing gel-loading dye) was added to each sample and the resuspension was heated at 95°C for 3 min followed by immediate chilling on ice. 1.7 µl of each sample was loaded onto each well of the gel in 1x TBE buffer.

Sequencing results were analysed using either *Sequence Navigator* or *SeQuencher* software.

| Reagent | Volume |
|---|---|
| 10x TBE | 3 ml |
| *Seqtagel AutoMatrix 4.5 | 27 ml |
| Ammonium persulphate (APS) | 150 µl |
| N,N,N',N'-tetramethylethylenediamine (TEMED) | 21 µl |
| **Total Volume** | **30 ml** |

**Table 9 Sequencing gel mixture.**

*Seqtagel AutoMatrix 4.5 was from National Diagnostics Ltd.

## 2.12 Protein Quantification and Detection

### 2.12.1 Quantification of Protein Concentration

Protein concentration was determined either by using the BioRad protein assay reagent according to the manufacturer's protocol with bovine serum albumin (BSA) as a standard, or by using spectrophotometry.

#### 2.12.1.1 Using BioRad Protein Assay Kit

When the BioRad Protein Assay Kit was used, 200 µl of 1 in 5 diluted dye provided was added to a protein sample of 10 µl and the $A_{595}$ was measured, using a microplate reader (Anthos HTII , Anthos Labtec HT2), after holding at RT for 5 min. Protein concentration was determined by comparison with a standard curve constructed from BSA (concentration ranging between 0.25 mg/ml and 0.0313 mg/ml) as a protein standard, which was co-assayed with the test protein.

#### 2.12.1.2 Using Spectrophotometry

When the spectrophotometry technique was used, absorbance at 280 nm was measured using GeneQuant II system. Protein concentration was calculated according to the Lambert-Beer equation:

$$A = KC$$

Where, $A$ is the absorbance measured, $K$ (mg/ml) is the extinction coefficient of a protein, which can be calculated according to the protein sequence using a programme under *ExPASy* (http://www.up.univ-mrs.fr/~wabim/d_abim/compo-p.html) and $C$ is the resulting concentration in mg/ml.

### 2.12.2 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Proteins or peptides were analysed by means of SDS-PAGE using a Tris-glycine buffer system (Laemmli, 1970) with MiniProtean III™ Electrophoresis System (BioRad).

#### 2.12.2.1 Gel Preparation

Glass plates were cleaned with methanol and clamped together, separated by fixed spacers. A resolving gel of appropriate acrylamide concentration was prepared (Table 10), and poured between the two plates. The gel was overlaid with methanol or $H_2O$ and allowed to polymerise for 30 min. When polymerised, the methanol was decanted and replaced by a stacking gel solution (Table 10) and a comb inserted to produce the loading wells. The use of a stacking gel, which is of low acrylamide concentration, allows relatively large samples to be concentrated before reaching the resolving gel, increasing the resolution of bands. Once polymerisation was complete (at least 30 min), the comb was gently removed and the plates were assembled onto the electrophoresis tank of the electrophoresis system. 1x electrophoresis buffer [10x electrophoresis stock solution: 3.03% (w/v) Tris-base, 14.4 (w/v) glycine and 1% (w/v) SDS] was poured into both the bottom and top tanks.

| | RESOLVING GEL | | STACKING GEL |
|---|---|---|---|
| **Reagent** | **12% Gel** | **16% Gel** | **3% Gel** |
| Acrylamide 40% | 6 ml | 8 ml | 0.47 ml |
| Tris-Cl (1 M, pH 8.8) | 7.5 ml | 7.5 ml | - |
| Tris-Cl (1 M, pH 6.8) | - | - | 0.78 ml |
| SDS (20%) | 100 µl | 100 µl | 31 µl |
| $H_2O$ | 6.12 ml | 4.12 ml | 4.89 ml |
| APS (10%) | 266.7 µl | 266.7 µl | 75 µl |
| TEMED | 13.3 µl | 13.3 µl | 5.2 µl |
| **Total Volume** | **20 ml** | **20 ml** | **6.25 ml** |

**Table 10 SDS-PAGE solutions.**

### 2.12.2.2 Sample Preparation

Immediately prior to electrophoresis, 0.2 volumes of 3x SDS sample buffer [0.175 M Tris-Cl (pH 6.8), 5.14% (w/v) SDS, 18% (v/v) glycerol, 0.3 M dithiothreitol (DTT), 0.006% (w/v) bromophenol blue] was mixed with each sample. The mixtures were heated at 90-95°C for 3-5 min and cooled to RT before loading.

### 2.12.2.3 Sample Separation

15-20 µl of samples were loaded into each well of the gels immediately prior to separation. Appropriate Rainbow Molecular Weight Markers (Amersham) were loaded in parallel with all of the samples in order to estimate the size of the samples. The electrophoresis was carried out at 120 V-160 V in Laemmli buffer (25 mM Tris-base, 250 mM glycine, 0.1% (w/v) SDS) . The gels were then subjected to either staining or Western blotting.

### 2.12.2.4 Gel Staining

When staining with coomassie brilliant blue R250, the gel was incubated in an aqueous solution containing 0.1% (w/v) coomassie brilliant blue R250, 10% (v/v) acetic acid and 40% (v/v) methanol for 30 min. It was then destained with a

solution made of methanol : acetic acid : water (4:1:5 by volume) until the background staining disappeared. Both stain and destain procedures were carried out at RT.

When SimplyBlue Safe Stain reagent (Invitrogen) was used, the manufacturer's instructions were followed.

#### 2.12.2.5 Gel Drying

Following gel destaining, gels were first incubated in drying solution [30% (v/v) ethanol, 10% (v/v) acetic acid, 3% (v/v) glycerol] and then packed in a sandwich arrangement with Gel Wrap (BDH). The packed gel was fixed into a set of frames, and air-dried. All steps were carried out at RT.

### 2.12.3 Western Blot Analysis

Western blotting was carried out by the method of Towbin (Towbin et al., 1979). A Trans-Blot Semi-Dry Transfer Cell (BioRad) was used to transfer proteins separated by SDS-PAGE onto nitrocellulose membranes (Schleicher & Schuell).

#### 2.12.3.1 Transfer of Proteins to Nitrocellulose Membrane

Following SDS-PAGE, the gel was placed in contact with a protein nitrocellulose transfer membrane in a sandwich arrangement surrounded by pads of Whatman 3 mm chromatography paper, which had been soaked in transfer buffer [0.048 M Tris, 0.039 M glycine, 0.032% (w/v) SDS and 20% (v/v) methanol, pH 8.3). Air bubbles were carefully removed. Transformation was carried out at 0.35 A per blot for 50 min.

### 2.12.3.2 Immunodetection of Blotted Proteins

Following blotting, nitrocellulose filters (blots) were blocked in a blocking buffer [PBST, 5% dried milk powder (Marvel)] at RT for 1 hr, or overnight at 4°C. After blocking, blots were first rinsed briefly in PBST, and then incubated in primary antibody diluted with PBST containing dried milk or BSA for 1hr at RT, or overnight at 4°C. Secondary antibody conjugated with horse radish peroxidase (HRP) was diluted in PBST containing the same milk and applied onto the blots for 30-60 min at RT. Following each of the antibody reaction stages, blots were washed (for 3x 10 min each time) in PBST to remove excess antibodies. The blots were incubated in enhanced chemiluminescence (ECL) reagents (KPL) according to manufacturer's instructions, drawn to remove excess liquid, covered in Saran wrap and the results were visualised on Hyperfilm ECL (Amersham).

### 2.12.3.3 Stripping of Membranes

If required, blots could be stripped of primary and secondary antibodies after immunodetection, and re-used. No signal remaining from the previous analysis was detectable after stripping.

After previous detection, filters were wet with PBST and incubated in 62.5 mM Tris-Cl (pH 6.8), 2% (w/v) SDS, and 100 mM ß-mercaptoethanol at 50°C for 30 min, with occasional agitation. The filters were then washed twice in PBST, for 10 min each time, at RT, blocked and reprobed.

## 2.13 Protein Expression

### 2.13.1 Expression of Proteins using Bacterial System

#### 2.13.1.1 Expression Vectors

Two *E. coli* expression vectors, pGEX-4T1 (Pharmacia Biotech) and miniprseta-mac (a gift from Dr Jane Clark, Department of Chemistry, University of Cambridge) were used to express glutathione S-transferase (GST)- and hexahistidine (His)-tagged fusion proteins, respectively. Both vectors are constructed to direct the synthesis of foreign polypeptides in *E. coli* upon induction with Isopropyl ß-D-thiogalactopyranoside (IPTG). cDNA encoding the protein of interest was cloned in frame into either vector and subsequently expressed in *E. coli* bacterial cells.

#### 2.13.1.2 Screening Colonies for Fusion Protein Expression

Screening colonies for fusion protein expression was always performed on a small-scale in 10 ml culture volume in order to select the best-expressed colony. Cells from the glycerol stock of a clone to be expressed were streaked onto a LA plate containing an appropriate antibiotic and incubated overnight at 37°C. Several colonies from the plate were randomly picked, each of which was used to inoculate 10 ml LB or 2x TY medium containing the same antibiotic. The culture was incubated at 37°C with vigorous shaking until the $A_{600}$ reached 0.8. The cells were induced with 0.1 mM IPTG, and grown for a further 2 hrs. Bacterial cells were pelleted by centrifugation (4,648 x *g*, 10 min, 4°C), resuspended in an appropriate lysis buffer and lysed. Following centrifugation (48,384 x *g*, 30 min, 4°C), both the supernatant and insoluble materials were analysed by SDS-PAGE and Western blotting to assess protein expression. The best-expressed colony was chosen for optimisation of expression conditions.

#### 2.13.1.3 Optimisation of Expression Conditions and Large-scaie Expression

Prior to large-scale expression, optimisation of expression conditions was carried out on a small-scale in 10 ml culture volume. Investigations were focused on the following aspects: using different *E. coli* strains or expression vectors, varying concentration of IPTG as well as temperature and length of time for induction.

Once optimum conditions were established, large-scale culture volume expression was carried out as follows. 2 ml (or 10 ml) culture containing appropriate antibiotic was grown at 37°C for 8-10 hrs and used to inoculate 100 ml (or 600 ml) fresh LB (or 2x TY) at 37°C for 16-18 hrs with vigorous shaking. This culture was diluted, 1 in 50, with fresh LB (or 2x TY) to 1 litre (or 30 litres, for fermentation) and grown at 37°C until the $A_{600}$ reached 0.8. Temperature was adjusted, if necessary, just before IPTG was added and induction was carried out according to the optimum conditions. Cells were pelleted and stored at -20°C for purification. The same bacterial cells containing empty vector provided the negative control.

### 2.13.2 Expression of Protein using *Drosophila* Cell Line

Expression of a protein of interest in *Drosophila* Schneider 2 cells (S2 cells) was carried out using the *Drosophila* Expression System (DES) Kit (Invitrogen) and the manufacture's instructions were essentially followed. A *Drosophila* expression vector, pMT/BiP/V5-His A, was used to produce a secretion fusion protein with a hexahistidine tag at the C-terminus.

### 2.13.2.1 Solutions and Media

*Complete DES expression medium*

DES Expression Medium plus 10% (w/v) fetal bovine serum (FBS), 1% (w/v) penicillin and 1% L-glutamine (2 mM)

*Selective medium*

Complete DES Expression Medium, 300 μg/ml Hygromycin B (hyg. B)

*Conditioned medium*

Supernatant of Cell Culture

*Freezing medium*

Conditioned medium/complete DES Expression Medium (1:1 in volume), 10% (v/v) FBS, 10% (v/v) dimethyl sulphoxide (DMSO)

*Solution A (300 μl)*

36 μl of 2 M $CaCl_2$, 10 μl (1 μg) pCoHygro vector, 42 μl (19 μg) plasmid, 212 μl of $H_2O$

*Solution B (2x HEPES solution)*

50 mM N-(2-hydroxyethyl)piperazine-N′-(2-ethanesulphonic acid) (HEPES), 1.5 mM $Na_2HPO_4$, 280 mM NaCl, pH 7.1

*Transfection mix*

Solution A was added dropwise into solution B with continuous mixing until solution A had depleted. The resulting solution was incubated at RT for 30-40 min until a fine precipitate had formed.

### 2.13.2.2 Cell Culture

Cell lines were cultured in complete DES expression medium at 25°C without $CO_2$. At a density of 6-20 x $10^6$ cell/ml, cells were subcultured into fresh medium, which was made by 1 in 2 dilution of the conditioned medium with

complete DES expression medium. Penicillin was included in the complete DES expression medium to avoid contamination.

### 2.13.2.3 Storage and Recovery of Cells in Liquid Nitrogen

Aliquots of S2 cells were stored in liquid nitrogen at a density of 1.1 x $10^7$ cells/ml (in freezing medium) following overnight freezing at -80°C.

To recover cells from liquid nitrogen, an aliquot was thawed quickly at 37°C, transferred to a flask containing complete DES expression medium, and then incubated overnight. Following centrifugation, cells were collected, resuspended in complete DES expression medium and cultured normally as described in Section 2.13.2.2.

### 2.13.2.4 Selection of Stable Cell Lines

Stably transfected cells were selected using the selection vector pCoHYGRO, which contains the hyg resistance gene.

Freshly made transfection mix was added dropwise to cells with a density of between 1.0-2.0 x $10^7$ cells/ml in a 35 mm petri dish and incubated for 24 hrs. Following centrifugation, cells were collected, washed with complete DES expression medium, resuspended in selective medium, replated back into the used petri dish and then incubated. The selective medium was changed every 4-5 days up to 4 weeks before expansion.

### 2.13.2.5 Expression

For induction, $CuSO_4$ was added to a final concentration of 500 μM and the culture was incubated under the same conditions as before. Both supernatant and

cells were collected every 24 hrs immediately after incubation up to the 5$^{th}$ day and analysed by Western blotting.

#### 2.13.2.6 Preparation of Genomic DNA from the S2 Cells

S2 cells from approximately 10 ml culture were pelleted by centrifugation and resuspended in a 200 µl low salt buffer containing 10 mM Tris-Cl (pH 7.4), 10 mM KCl, 10 mM $MgCl_2$, and 2 mM EDTA (pH 8.0). Following a brief vortex, 20 µi of 10% SDS was added to the mixture, which was then incubated for 2 hrs at 50°C. Phenol/chloroform extraction was performed, followed by ethanol precipitation, as described in Sections 2.6.2 and 2.6.3. The final pellet was resuspended in 20-50 µl of TE buffer, followed by the addition of 1 µl of RNase A (1 mg/ml). This genomic DNA preparation was stored at -20°C.

### 2.14 Protein Purification

#### 2.14.1 Purification of GST Fusion Proteins

Expressed cells were re-suspended (about 5 ml buffer per gram of cells) in an ice-cold lysis buffer containing PBS, 2 mM EDTA, 0.1% β-mercaptoethanol, 1 µM phenylmethylsulfonyl fluoride (PMSF) and protease inhibitor tablet (EDTA free, Roche). Cell lysis was achieved by sonication for 18-36 x 10 sec depending on the volume of cell suspension. In the majority of cases, GST fusion proteins are soluble under non-denaturing conditions so that they can be purified from crude bacterial lysates, under native conditions, by affinity chromatography on immobilised glutathione. Therefore, cell lysates were centrifuged (48,384 x *g*, 30 min, 4°C), and supernatants were collected followed by purification using either Glutathione Sepharose 4B slurry or GSTrap pre-packed column (Amersham Bioscience) according to manufacturer's guidelines. Bound protein was eluted with

freshly made glutathione solution [20 mM of glutathione (reducing form), 50 mM Tris-Cl (pH 8.0)] at RT. The eluates were dialysed to remove glutathione and concentrated, if necessary, using Vivaspin 20 columns (5000 MWCO) (Vivascience), followed by analysis on both SDS-PAGE and Western blotting.

### 2.14.2 Dialysis

In most cases, dialysis was carried out using dialysis membrane tubing (3500 MWCO) (Spectrum Laboratories) with at least two buffer changes of greater than 50 times the volume of the sample. However, for dialysis of small sample volumes, (< 3 ml), dialysis cassettes (Pierce) were used. When a dialysis technique was used for protein refolding, the concentration of a solution containing the protein that is to be refolded was adjusted into the 10-50 μg/ml range prior to the dialysis.

### 2.14.3 Thrombin Digestion

To remove GST from GST fusion proteins, thrombin was added (10 u thrombin per 1 mg protein) and incubated for 16 -18 hrs at RT. In order to stabilise thrombin, $Ca^{2+}$ was included, at a final concentration of 1-5 mM. The digestion was stopped by addition of PMSF to a final concentration of 0.2 mM followed by incubation for 15 min at 4°C. The digestion mixture was then loaded onto high performance liquid chromatography (HPLC) column for further purification as described below.

### 2.14.4 Purification using a MS Linked HPLC Technique

HPLC purification was kindly performed by Dr Aiwu Zhou in the Department of Haematology, University of Cambridge.

A C18 HPLC column (Phenomenex, PRODIGY 5u ODS3, 21.2mm x 250mm, 100Å) was equilibrated with 5% (HPLC grade) acetonitrile (MeCN) + 0.1% trifluoroacetic acid (TFA). A thrombin digestion mixture of a GST fusion protein was then injected onto the column followed by a gradient wash: starting at 5% MeCN + 0.1% TFA and ending at 100% MeCN + 0.1% TFA. The protein was eluted across a gradient of 20-60% MeCN in 0.1% TFA and then lyophilised and stored at -80°C.

The purity and molecular weight of the eluted protein was analysed by liquid chromatography-mass spectrometry (LC-MS) (HP1100 coupled with LCQ, Finnigan MAT) which was kindly performed by Dr Hui Hong in the Department of Chemistry, University of Cambridge.

### 2.14.5 Purification of His-tagged Fusion Proteins

Expressed bacterial cells were resuspended in ice-cold lysis buffer containing PBS, 10 mM imidazole, 1 μM PMSF and protease inhibitor tablet, pH 8.0. Lysozyme was then added, to a final concentration of 0.35 mg/ml, and incubated at RT for 30 min. Cell lysis was achieved by sonication. If the expressed fusion protein was in the form of inclusion bodies, the insoluble materials separated by centrifugation were collected and solubilised with a buffer containing PBS, 6 M urea and 10 mM imidazole pH 8.0 for 30 min at 4°C. Following centrifugation (48384 x *g*, 30 min, 4°C), the supernatant was collected. Ni-NTA agarose resin (Qiagen) was added into the supernatant and incubated for 1 hr at 4°C. The resin containing bound protein was washed 4 times with a solution containing PBS, 30 mM imidazole and 6 M urea (pH 8.0). Bound proteins were eluted with PBS, 300 mM NaCl, 500 mM imidazole and 6 M urea, pH 8.0. The eluates were first analysed on SDS-PAGE and Western blotting for assessment of

the fusion protein expression and purification before being submitted to protein refolding using either a dialysis technique or commercially available protein refolding kit (see the following Section for details).

If an expressed His-tagged protein was soluble, cleared cell lysate was collected and purified using the same resin, but under native conditions. In this case, all buffers used were the same as above but in the absence of urea.

## 2.15 Protein Refolding

Refolding of protein from inclusion bodies was carried out using a Protein Refolding Kit (Novagen), with some modifications to the manufacturer's guidelines as follows.

### 2.15.1 Buffers and Solutions

*Resuspension buffer*

20 mM Tris-Cl (pH7.5), 10 mM EDTA, 1% Triton X-100, 5 mM DTT, 1 mM PMSF

*Solubilisation buffer*

50 mM 3-[cyclohexylamino]-1-propanesulfonic acid (CAPS) (pH 11.0), 0.3% (w/v) N-lauroylsarcosine, 1 mM DTT

*Dialysis buffer*

20 mM Tris-Cl (pH 8.5), 0.1 mM DTT

### 2.15.2 Purification and Solubilisation of Inclusion Bodies

Purification of inclusion bodies contained in insoluble materials was carried out as follows. After cell lysis (Section 2.14.5), DNase I was added, to a final concentration of 1%, and incubated at 37°C for 1 hr followed by centrifugation

(48,384 x *g*, 30 min, 4°C). The insoluble materials were collected and purified by repeatedly resuspending (with the resuspension buffer), sonicating and centrifugating the materials until the inclusion bodies, as analysed by SDS-PAGE, were clean.

Solubilisation of the inclusion bodies was achieved by mixing them with solubilisation buffer, followed by incubation at RT for 3 hrs. Insoluble materials were removed by centrifugation (48,384 x *g*, 30 min, RT) and the supernatant was carefully collected and submitted for two dialysis at 4°C overnight and then for 3 hrs, both of which were against 4 litres of dialysis buffer.

## 2.16 Human Kidney Protein Fractionation

### 2.16.1 Preparation of Crude Cytosolic and Membrane Protein Fractions

Samples of normal human kidney were obtained from nephrectomy specimens resected due to renal tumours. Informed, written consent was obtained, with the approval of the Histopathology Department at Addenbrooke's Hospital, Cambridge's Tissue Bank Committee. 1 g of snap frozen kidney tissue was chopped in liquid nitrogen and transferred into 3.5 ml of ice-cold homogenisation buffer [10 mM Tris-Cl (pH 7.4), 150 mM NaCl, 2 mM PMSF, 1 mM ethyleneglycol-bis[ß-aminoethyl etherl (EGTA), 1 mM EDTA, 1 mM DTT, protease inhibitor cocktail tablet] followed by homogenisation at 15,000 rpm with a Polytron® homogeniser (Kinematica). The homogenate was centrifuged (1000 x *g*, 10 min, 4°C) to remove nuclei and cell debris. Following collection of the supernatant, a further centrifugation (100,000 x *g*, 1 hr, 4°C) was performed to separate cytosolic and crude membrane fractions. The pelleted membrane fraction was resuspended in the same homogenisation buffer and both fractions were stored in small aliquots at -80°C.

### 2.16.2 Solubilisation of Human Kidney Membrane Proteins

Based on earlier work (Gluck and Caldwell, 1987), human kidney membrane proteins prepared as described above, were solubilised in a solubilisation buffer containing 10 mM Tris-Cl (pH 7.4), 1 mM EDTA, 1 mM DTT, 10% glycerol, either 1% n-Nonyl-β-D-glucopyranoside (n-NDG) or 0.6% (3-[(3-cholamidopropyl)dimethylammonio]-1-propane-sulphonate (CHAPS) or both, and protease inhibitor cocktail tablet. Following gentle rotation, solubilisation (for 30 min) and centrifugation (100,000 x *g*, 1 hr) were applied in order to recover the supernatant. All steps were carried out at 4°C.

## 2.17 Circular Dichroism (CD) Spectroscopy

The secondary structures of expressed proteins were analysed with far-UV CD spectroscopy. After a protein solution was filtered through a polyethersulfone membrane, the CD spectra were measured by Jasco J-810 spectropolarimeter at 20°C (or 22°C). The concentration of the protein was normally adjusted to a range of 0.20-1 mg/ml before measurement. The cell length used was 1 mm. Data were collected every 0.5 nm with a 1 nm bandwidth and a 1 sec time constant. Scan speed was 50 nm/min. Ten scanning spectra were averaged for each sample.

## 2.18 N-terminal Sequence Analysis

N-terminal sequence analysis was kindly performed by Mr M. Weldon in the Protein and Nucleic Acid Chemistry Facility in the Department of Biochemistry, University of Cambridge.

## 2.19 Identification of Protein Ligands

### 2.19.1 Phage Display Random Peptide Library Screening

The identification of short peptide sequences that are able to bind to a protein of interest was performed with the Ph.D.-7™ and Ph.D.-12™ Phage Display Random Peptide Libraries (NEB). Experimental procedures were carried out according to manufacturer's guidelines but with some alterations.

#### 2.19.1.1 Buffers, Solutions and Media

In this section, all concentrations expressed as "%" are in w/v and all media and solutions were prepared in $H_2O$ and then sterilised by autoclaving, unless otherwise indicated.

*Agarose top*

10 g/l bacto-tryptone, 5 g/l yeast extract, 5 g/l NaCl, 1 g/l $MgCl_2 \cdot 6H_2O$, 7 g/l agarose

*Blocking buffer*

0.1 M $NaHCO_3$ (pH 8.6), 0.5% BSA, 0.02% $NaN_3$. Filter to sterilise.

*Elution buffer (pH 2.2)*

0.2 M Glycine-HCl, 1 mg/ml BSA

*Iodide buffer*

10 mM Tris-Cl (pH 8.0), 1 mM EDTA, 4 M NaI

*IPTG/Xgal plates*

0.5% IPTG, 0.4% Xgal contained in LA plates

*Polyethylene glycol/NaCl solution*

20% Polyethylene Glycol-8000 (PEG-8000), 2.5 M NaCl

*ABTS solution (pH 4.0)*

0.022% (w/v) 2,2'-azino-bis[3-ethylbenzthiazoline 6-sulphonic acid] diammonium (ABST), 50 mM sodium citrate, 0.05% $H_2O_2$

### 2.19.1.2 Biopanning

Target protein was diluted in 0.1 M $NaHCO_3$ (pH 8.6) to a final concentration of 0.1 mg/ml. 10-15 μg of the protein was coated onto a micro-plate well at 4°C overnight in a humidifier. Unbound protein was removed and the well was blocked with blocking buffer for 2 hrs at 4°C. Following blocking, the well was washed 6 times with TBS containing 0.1% Tween 20 (TBST). 10 μl of the original phage peptide library containing about 2 x $10^{11}$ phage virions was diluted with 100 μl TBST and added into the blocked well for the first round of panning. After incubation of 1 hr at RT, unbound phage was removed and the well was washed 10 times with TBST. Specifically bound phage was eluted and submitted for amplification in *E. coli* ER2738 strain (see the following Section for details of phage amplification). About $10^{11}$ amplified phage virions were used for the next round. The whole procedure was repeated three to four times, individual phage clones were isolated, as described in Section 2.19.1.4, and characterised by DNA sequencing. However, the Tween 20 concentration of the TBST was increased from 0.1% to 0.5% from the 2$^{nd}$ panning onwards to reduce non-specific phage binding.

### 2.19.1.3 Phage Amplification and Titration

*Phage Amplification*

A single ER2738 colony was used to inoculate 20 ml LB medium containing tet (LB-tet) at 37°C until $A_{600}$ reached ~0.2. Phage eluate from a panning was added into the culture and grown for a further 4.5 hrs before centrifugation (12,000

x *g*, 10 min, 4°C). The supernatant was collected and re-spun under the same conditions. To the upper 80% of the supernatant, 1/6 volume of PEG/NaCl solution was added and the mixture was incubated overnight at 4°C to precipitate phage particles. Following centrifugation (12,000 x *g*, 15 min, 4°C), the resulting pellet was resuspended in 1 ml of TBS. Further precipitation of the phage particles from the 1 ml resuspension was carried out by the addition of 1/6 volume of PEG/NaCl followed by incubation of the mixture on ice for 1 hr. After centrifugation (16,000 x *g*, 10 min, 4°C) phage particles were collected and re-suspended in 200 µl of TBS containing 0.02% $NaN_3$. This amplified phage eluate was stored at 4°C.

*Phage Titration*

A single ER2738 colony was used to inoculate LB-tet medium at 37°C until the $A_{600}$ reached ~0.5. 200 µl of the culture was separately mixed with 5 µl, 10 µl, 25 µl, 50 µl, 100 µl and 250 µl of a 1 in $10^9$ dilution of the amplified phage eluate. Each mixture was first incubated at RT for 5 min, then mixed with 3 ml pre-melted agarose top (45°C). The resulting mixture was immediately poured onto pre-warmed (37°C) LB/IPTG/Xgal plates followed by incubation overnight at 37°C, before counting the plaques the following day. The phage titer was expressed as plaque-forming units (pfu) per 10 µl of phage resuspension.

### 2.19.1.4 Isolation Phage DNA for Sequencing

Individual phage clones were amplified as described above, and pellets were resuspended thoroughly in 100 µl iodide buffer containing 250 µl ethanol, followed by incubation for 10 min at RT to precipitate single-stranded phage DNA. After centrifugation, pellets were first washed with 70% ethanol, air-dried and resuspended in 10 µl Tris buffer (10 mM, pH 8.5). Standard sequencing reactions were carried out using 5 µl of the phage DNA with primer 19 (Table A.3).

### 2.19.1.5 Phage ELISA

The specificity of binding of bound phage to target protein was confirmed by enzyme linked immunosorbent assay (ELISA), which was carried out essentially under the same conditions as those used in the panning procedures. 96-well plates were coated with 10-15 μg of the target protein in 0.1 M $NaHCO_3$ (pH 8.6), at 4°C overnight in a humid box. Excess protein was removed and wells were first blocked with the blocking buffer for 1.5 hrs at 4°C, and then washed 6 times with TBST. Serial dilutions of the consensus sequence displaying phage were added into the wells and incubated at RT for 1.5 hrs, followed by 6 washes with TBST. Non-specific binding was evaluated in parallel wells lacking the target protein.

For detection, HRP conjugated anti-M13 antibody diluted 1:5000 in PBST containing 2% Marvel was added into each well and incubated for 2 hrs at RT. Following 6 washes in PBST, ABTS developing solution was added to each well, and incubated at RT until a suitable green colour had formed. The $A_{405}$ value was then measured.

## 2.19.2 Yeast Two-hybrid (Y2H) Assay

Y2H assay was carried out using the CytoTrap two-hybrid system (Stratagene). cDNA encoding the protein of interest (bait protein) was cloned into the pSOS vector, and cDNA libraries containing inserts encoding target proteins were cloned into the pMyr vector.

### 2.19.2.1 Buffers, Solutions and Media

In this section, all media and buffers/solutions were prepared with $dH_2O$ and $H_2O$, respectively, and sterilised by autoclaving unless stated otherwise. All concentrations expressed as "%" are in w/v.

*10x dropout solution [without uracil and leucine (-UL])*

0.03% L-isoleucine, 0.15% L-valine, 0.02% L-adenine hemisulphate salt, 0.05% arginine HCl, 0.02% L-histidine HCl monohydrate, 0.05% L-lysine HCl, 0.02% L-methionine, 0.05% L-phenylalanine, 0.2% L-threonine, 0.05% L-tryptophan, 0.05% L-tyrosine, 0.1% L-glutamic acid, 0.1% L-Aspartic acid, 0.04% L-serine. Filter to sterilise.

*LiSORB solution*

100 mM LiAc, 10 mM Tris-Cl (pH 8.0), 1 mM EDTA, 1 M sorbitol.

*PEG/LiAc solution*

10 mM Tris-Cl (pH 8.0), 1 mM EDTA (pH 8.0), 100 mM LiAc (pH 7.5), 40% PEG 3350.

*YPAD broth*

10 g/l yeast extract, 20 g/l bacto peptone, 20 g/l dextrose (D-glucose), 0.04 g/l adenine sulphate.

*SD/Glucose medium (-UL)* [SD/Glc(-UL)]

0.17 g/l yeast nitrogen base without amino acid, 5 g/l ammonium sulphate, 20 g/l glucose, 1x dropout solution.

*SD/Glc medium [without leucine (-L) / or uracil (-U)]*

SD glucose medium (-UL) plus 0.2 g/l uracil / or 1 g/l leucine

*SD/Galactose medium (-UL)* [SD/Gal(-UL)]

0.017% yeast nitrogen base without amino acid, 0.5% ammonium sulphate, 2% galactose, 1% raffinose, 1x dropout solution.

*Solid agar plate*

Solid agar plates were of identical composition to that of liquid medium but with the addition of Bacto agar. YPAD plates were composed of YPAD broth plus

2% Bacto agar and all other solid SD glucose/galactose plates were composed of relative liquid medium with the addition of 1.7% Bacto agar.

### 2.19.2.2 Preparing Yeast Competent Cells

The following method yields 6.8 ml competent cells with one preparation. 4 to 5 colonies of yeast cdc25H (strain a or α) from YPAD plates were used to inoculate 50 ml of fresh YPAD broth at 25°C for about 18 hrs ($A_{600}$ > 1). The appropriate volume of culture was diluted to 300 ml with YPAD until the $A_{600}$ reached 0.2, before being followed by a 3 hr-incubation at 25°C to achieve $OD_{600}$ > 0.7. After centrifugation (1000 x *g*, 5 min, 25°C), the cell pellet was collected, washed once in $H_2O$, resuspended in LiSORB solution, and incubated for 30 min at RT before being subjected to centrifugation. The pellet was resuspended in 300 μl of LiSORB, to which 600 μl of LiSORB solution containing 4% (v/v) denatured salmon sperm DNA was added, followed by the addition of 5.4 ml PEG/LiAc solution and 530 μl of DMSO. The mixture containing freshly made yeast competent cells was used immediately for transformation.

### 2.19.2.3 Transformation of Yeast Competent Cells

Mixtures of 100 μl yeast competent cells, and either 100 ng of a plasmid (single transformation) or 300 ng of each plasmid (co-transformation), were prepared followed by the addition of 2 μl of 1.4 M β-mercaptoethanol. After incubation at RT for 30 min, cells were heat shocked at 42°C for 20 min, followed by incubation for 3 min on ice. Cells were pelleted by centrifugation, resuspended in 0.5 ml 1 M sorbitol solution and plated out onto the appropriate plates, which were then incubated at RT until the colonies became visible. Transformation

efficiency was expressed as colony-forming units (cfu) per μg of plasmid DNA transformed.

### 2.19.2.4 Assessment of Bait Constructs

All bait plasmids were subjected to the recommended control tests to verify suitability of the bait plasmids for CytoTrap interaction assay. Bait plasmid was co-transformed with empty pMyr vector into cdc25H competent cells before being plated onto a SD/Glc(-UL) agar plate. After 4-5 day incubation at RT, the colonies that appeared were picked up and each of them was patched onto two SD/Gal(-UL) agar plates and two SD/Glc(-UL) agar plates. One plate each was incubated at 37°C and the others were at RT. If no colonies appeared on SD/Gal(-UL) agar plates after 5-day incubation at 37°C, the bait construct was deemed suitable to be used in this system, otherwise not.

### 2.19.2.5 Library Screening and Selection of Positive Colonies

40 μg of bait plasmid were co-transformed with the same amount of a cDNA library into the cdc25H (strain α) competent cells, plated onto SD/Glc(-UL) plates and incubated at RT for 48 hrs. Transformants were replicated onto SD/Gal(-UL) plates and incubated at 37°C for up to ten days. Colonies, termed *interactor candidates*, were picked from the plates incubated between 6-10 days. pSos vector replaced the bait plasmid as a negative control.

The *interactor candidates* were patched on SD/Glc(-UL) plates followed by incubation for 48 hrs at RT to repress *GAL1* promoter-driven expression from the pMyr library. After incubation, cells were patched onto two SD/Glc(-UL) and one SD/Gal(-UL) plate. One plate of each type was incubated at 37°C and the 2nd SD/Glc(-UL) plate was incubated at RT for 48 hrs. This patching test was

proceeded one more time from cells grown on the 2nd SD/Glc(-UL) plate at RT. Only patches, which were grown twice on SD/Gal(-UL) plates, but not on SD/Glc(-UL) at 37°C, were selected as *putative positives*.

### 2.19.2.6 Mating Test to Verify the Putative Positives

To prepare cells for mating, the bait vector in the *putative positives* must be removed. This procedure is known as *curing*. To achieve this, a *putative positive* clone was used to inoculate SD/Glc(-U) liquid medium for 2 days at 25°C, then plated onto SD/Glc(-U) plate and incubated at RT for a further 4 days. The single colony obtained was then plated onto both SD/Glc(-U) and SD/Glc(-L) plates. Following incubation at RT for about 4 days, clones that were grown on SD/Glc(-U) plates only [SD/Glc(-U)$^+$], but not on SD/Glc (-L) plates [SD/Glc(-L)$^-$ ] were selected for mating tests.

For mating tests, each selected cured *putative positive* (in stain α), i.e. SD/Glc(-L)$^-$(-U)$^+$ clones, was co-patched with either the yeast cells/$H_2O$ suspension containing bait construct or pSOS Col I, which was provided by the system as a negative control, (both were in Strain a) on PYAD plates. Following incubation at RT for 24 hrs, the mated cells were patched onto two SD/Glc(-UL) and two SD/Gal(-UL) plates. One plate of each type was incubated at 37°C and the 2nd plate of each type incubated at RT for at least 5 days. The cured *putative positives* which grew at 37°C on SD/Gal(-UL) plates mated with bait containing cells, but not pSOS Col I containing cells, termed *Positive* clones. These were then submitted to plasmid isolation followed by sequencing analysis.

### 2.19.2.7 Plasmid Isolation and Sequencing Analysis

*Positive* clones containing only pMyr cDNA plasmid were resuspended thoroughly in 100 µl TE buffer followed by addition of 25 µl of 5 M NaCl, 5 µl of 20% SDS, and 130 µl of acid-washed glass beads (0.5 mm) (Sigma). Phenol/chloroform extraction was carried out, followed by precipitation using a solution containing 25 µl of 5 M NaCl and 340 µl of 100% ethanol. The final DNA pellet was resuspended in 20 µl $H_2O$; 1 µl of this resuspension was then used in a 20 µl-standard PCR reaction (Table 4); 5 µl of the PCR product was subsequently used in a sequencing reaction.

## 2.20 Protein Labelling: Biotinylation of Protein

Lyophilised protein was dissolved at 0.5 mg/ml in $H_2O$. *N*-(5-aminopentyl) biotinamide (biotin cadaverine) (Molecular probes), was dissolved in 0.1 M 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide, hydrochloride (EDAC) to a final concentration of 2.4 mg/ml. Biotin cadaverine/EDAC solution was added to 50 µl of the protein solution, containing 25 µg of the protein, and incubated at RT for 1 hr. Excess biotin cadaverine was quenched by the addition of 1 M NaAc (pH 5.0), followed by incubation at RT for 2 hrs. Water replaced the protein as a control. The labelled samples were stored at -20°C.

## 2.21 Characterisation of Protein-Protein Interactions

### 2.21.1 Affinity Pull-down Assay of Protein-Protein Interaction

Affinity pull-down assay between a bead-bound protein and either another purified protein or a tissue extract was achieved as follows. Samples were mixed and incubated for 1-18 hrs at 4°C with gentle rotation. The beads were collected by centrifugation (200 x *g*, 2 min, 4°C) and washed (4x 5 volumes) with ice-cold

PBST or PBS plus 0.1% Tween 20. Bound proteins were eluted from the beads by boiling in SDS sample buffer for 5-10 min at 95°C and the supernatants were either spotted onto nitrocellulose membrane or subjected to SDS-PAGE. Appropriate primary antibody was used to probe the spotted membrane or Western blots according to standard methods.

### 2.21.2 Assay for Protein Immunoreactivity in Human Kidney

Protein immunoreactivity in human kidney was tested as follows. 10-40 μg of each human kidney cytosolic and membrane protein fractions, prepared as described in Section 2.16.1, were subjected to 12% SDS-PAGE. Western blots were probed using appropriate antibodies according to the standard methods.

### 2.21.3 Co-Immunoprecipitation Assay in Human Kidney

#### 2.21.3.1 Preparation of Affinity Agarose Gel

Immobilisation of antibody onto agarose beads using the Seize Primary Mammalian Immunoprecipitation Kit (PIERCE Biotechnology) was essentially carried out according to the manufacturer's guidelines.

Briefly, 200 μl of 50% agarose beads was mixed with 60 μg of the affinity purified antibody in coupling buffer (0.1 M $Na_3PO_4$, 0.15 M NaCl, pH 7.2) followed by the addition of sodium cyanoborohydride, a chemical used for covalently coupling the amine group of the antibody to the beads. After overnight incubation at 4°C, excess antibody was removed by washing 6 times with 1 M NaCl. The antibody-coupled beads were washed 3 times with binding buffer (8 mM $Na_3PO_4$, 2 mM $K_3PO_4$, 10 mM KCl, pH 7.4) and then stored in 200 μl of the binding buffer containing 0.02% $NaN_3$ at 4°C.

### 2.21.3.2 Co-Immunoprecipitation

30 μl of unconjugated beads (50%) from the same kit were added to solubilised human kidney membrane fraction (see Section 2.16.2 for details), and incubated for 1 hr. These beads were replaced by 50 μl of the antibody-coupled beads prior to overnight incubation at 4°C. The beads were then washed three times with 1 ml of buffer containing 20 mM Tris-Cl (pH 7.4), 5 mM $NaN_3$ and either 0.3% n-NDG or 0.3% CHAPS or both; then a further 3 times with the same buffers containing 500 mM NaCl; and then 3 final times with buffers without NaCl. Bound proteins were eluted from the agarose beads by boiling in SDS sample buffer at 95°C and supernatants were subjected to SDS-PAGE. Western blots were probed using the appropriate antibodies according to standard methods.

## 2.22 Binding Affinity Study by Surface Plasmon Resonance (SPR)

Binding affinity study of a protein-protein interaction by SPR was carried out with great help from Dr Babak Javid in the Department of Medicine, University of Cambridge.

Binding analysis was performed on a BIAcore™ 2000 biosensor system (Pharmacia Biosensor AB) using SPR measurements. A carboxymethylated sensor chip (type CM5) was activated with 1:1 mixture of 0.4 M N-ethyl-N'-(3-dimethylaminopropyl) carbodiimide (EDC) and 0.1 M N-hydroxysuccinimide (NHS) in $H_2O$. Protein X (ligand) was then immobilised on the sensor chip by amine-coupling according to the manufacturer's instructions. Unreacted sites were blocked with 1 M ethanolamine/HCl (pH 8.0). Control flow cells were activated and blocked with absence of the ligand. The flow cells were routinely equilibrated with running buffer (PBS, 0.005% surfactant P20). Protein Y (analyte) was diluted in the running buffer and allowed to interact with the sensor surface by a 180-second

injection. Seven different concentrations (7.5-250 μM) of the analyte were injected at a flow rate of 10 μl per min at 25°C. The data were fitted to a 1:1 binding equilibrium model using Grafit (Erithacus software).

## 2.23 Immunohistochemistry

### 2.23.1 Tissue Section Preparation

Kidney tissue sections were kindly prepared by Dr R Al-Lamki in the Department of Medicine, University of Cambridge.

Human kidney Cortico-medullary sections < 1 mm thick were embedded in OCT embedding compound and snap-frozen in liquid nitrogen-cooled isopentane. 5 μm sections were cryostat-cut and thaw-mounted on APES-coated glass slides. Sections were stored at -80°C until use.

### 2.23.2 Immunodetection

Sections were fixed in 100% methanol at -20°C for 5 min and rehydrated by a 5 min immersion in TBS (pH 7.5) containing 0.01% Tween 20 (TBST). All subsequent incubations and rinses were in TBST. Non-specific antibody binding was blocked by incubation in blocking buffer [TBST, 10% fetal calf serum (FCS)] for 15 min. Primary antibodies were applied at 1:100-1:5000 dilutions overnight at 4°C in the blocking buffer. Appropriate fluorescent conjugated secondary antibodies were applied at 1:100 dilutions for 40 min at RT. Following mounting in Citifluor mounting medium, bound antibodies were visualised using a Leica TCS-NT Confocal Laser Scanning Microscope (work with help from Dr R Al-Lamki). Replacement of primary antibody with the appropriate pre-immune serum provided negative controls.

# CHAPTER 3
# EXPRESSION AND PURIFICATION OF a4(N), a4(Loop2) AND a4(C)

## 3.1 Introduction

The a subunit of the proton pump is deemed crucial for the coupling of ATP hydrolysis ($V_1$) and proton transport ($V_0$). As noted from the yeast orthologue Vph1p, evidence obtained from several studies has revealed an essential role of this subunit in proton translocation activity (Leng et al., 1996; Sato and Toyama, 1994). In addition to this, the N-terminal domain of the a subunit is likely to contribute to the formation of the peripheral stator, and also to the control of targeting and *in vivo* dissociation of the V-ATPase (Kawasaki-Nishi et al., 2001c). Furthermore, the soluble C-terminal tail of the a subunit has been shown to play a potential role in the assembly, stabilisation and targeting of the a subunit (Leng et al., 1998). However, information concerning potential functional or regulatory contributions of the N- or C-terminal domains of the different tissue-specific homologues of the mammalian V-ATPase a subunit is scarce.

I have been particularly interested in the a4 paralogue because of its essential contribution to renal acid-base homeostasis. Noting firstly the suggested importance of both domains (as outlined above), and secondly the C-terminal region being the most homologous among the four human a subunit paralogues, I therefore sought further to explore the roles of these two domains of human a4 [here designated a4(N) and a4(C), respectively]. In addition to these two domains, the 2$^{nd}$ loop is most homologous compared to the other loops among the four human a subunit paralogues. However, to date, no information concerning potential functional or regulatory contributions of this domain has been reported. I

was therefore also interested in investigating this domain of human a4 [here designated a4(Loop2)].

In order to understand the biological functions of this subunit, I sought further to explore the roles of this protein through identification of its potential-binding partners. A powerful approach to this is to first express the coding sequence of a4 in order to produce the protein, which is then used to identify its binding partners. I have already mentioned before that the a4 is an integral transmembrane protein. Generally, expression levels of recombinant membrane proteins have been rather low and the purification procedure inefficient. Attempts to express recombinant a4 in bacterial system were not successful (data not shown), which might be due to the toxicity of the protein inserted into the bacterial membrane. Therefore, it was decided to express the coding sequences for a4(N) (residues 1-393), a4(Loop2) (residues 467-544, 78 amino acids) and a4(C) (residues 796-840, 45 amino acids) instead of the whole protein. The molecular weight of a4(N), a4(loop2) and a4(C) were predicted to be 45.30 kDa, 10.77 kDa and 5.264 kDa respectively, using the program Compute pI/Mw tool (http://us.expasy.org/tools/pi_tool.html). These recombinant domains were then used to identify their binding partners, and subsequently characterise their binding.

This chapter describes the construction, expression, purification and structural analysis of the a4(N), a4(loop2) and a4(C) domains. Two bacterial cell expression vectors, pGEX-4T1 and miniprseta-mac, and one *Drosophila* cell expression vector, pMT/Bip/V5-His A, were used to produce the proteins. The pGEX-4T1 vector (Figure A.1*a*) used to express GST fusion protein is constructed to direct the synthesis of foreign polypeptides in *E. coli* as fusions with the C terminus of a 26-kDa GST protein. This vector contains the strong IPTG-inducible *tac* promoter, which controls expression of inserts cloned into the vector. This

vector is also engineered with an internal *LacI*$^{q}$ gene. The *LacI*$^{q}$ gene product is a repressor protein that binds to the operator region of the *tac* promoter, preventing expression until induction by IPTG. In this way, a control over the expression of the insert is maintained. The vector also contains a thrombin cleavage site just before the multiple cloning sites (MCS), so that the GST carrier can be cleaved from fusion proteins by thrombin. The miniprseta-mac vector (Figure A.1*b*) contains the bacteriophage T7 promoter induced by T7 RNA polymerase, which was induced by addition of IPTG to bacterial cells. This vector produces a fusion protein with a His tag at the N-terminus. The pMT/Bip/V5-His A vector (Figure A.2) contains the *Drosophila* metallothionein (MT) promoter which is tightly regulated and is easily induced by the addition of $CuSO_4$. This vector also carries the *Drosophila* BiP signal sequence for efficiently targeting BiP and its fusion protein into the secretary pathway of S2 cells. Frequently, secretion of recombinant proteins into the culture medium improves their yield and quality. This vector produces a fusion protein, with a His tag at the C-terminus, in culture medium.

## 3.2 Results

### 3.2.1 Expression and Purification of a4(N)

#### 3.2.1.1 Using *E. coli* Bacterial Expression System

In order to express a4(N), an *E. coli* bacterial system was first chosen due to its simplicity and potentially high levels of protein expression.

*Construction of a4(N) into pGEX-4T1 Expression Vector*

The cDNA sequence encoding the intact a4 protein was previously cloned in pGEM-T Easy vector by Dr Annabel Smith (Department of Medical Genetics, University of Cambridge). The cDNA sequence spanning the a4(N) region was amplified from this vector by high fidelity LR-PCR using primer 1 (forward) and

primer 2 (reverse) (Table A.3). The resulting 1.193 kb *Eco*RI-*Not*I PCR fragment was digested using the two restriction endonucleases, and subsequently ligated into pGEX-4T1, which had been cleaved with the same restriction enzymes. The ligation product was used to transform two *E. coli* strains, C41(DE3) [which was derived from BL21(DE3)] and AD494(DE3), by electroporation. Positive recombinant clones were identified by both colony PCR (using the same primers as in PCR), and restriction endonuclease digestion (using the same enzymes as in cloning). The insert sequence in the positive clone was verified by DNA sequencing, in both directions, using primers 17, 25, 26, 27 (forward) and 18, 28, 29, 30 (reverse) (Table A.3). This is to ensure the absence of PCR-induced mutations and that the fusion to GST was in-frame. A recombinant plasmid containing the correct *a4(N)* sequence, termed pGEX-a4(N), was subsequently subjected to small-scale expression trials.

*Expression and Purification*

Prior to large-scale expression, small-scale expression trials - as listed below, were performed to screen the best-expressed colonies and to optimise expression conditions.

Firstly, three randomly chosen BL21 colonies, harbouring the pGEX-a4(N) plasmid, were screened for GST fused a4(N) protein [GST-a4(N)] expression at 37°C with 2 hr induction. After cell lysis, samples of insoluble and soluble protein fractions were analysed by SDS-PAGE and Western blotting. The Western blot was probed with anti-GST antibody (Pharmacia), and then followed by HRP-conjugated secondary antibody. Both assays confirmed the production of GST-a4(N) in all samples collected and there were no visible differences in expression levels observed between the three colonies (data not shown).

Secondly, small-scale cultures, in 10 ml volume, were grown at 37°C until the $OD_{600}$ reached 0.8. Cultures were allowed to equilibrate at 30°C for 30 min prior to induction with 0.1 mM IPTG. Samples taken at 2, 4 and 5 hr post induction were lysed, and the GST-a4(N) in the soluble sonicate was purified, under native conditions, using glutathione sepharose. All eluates were submitted to SDS-PAGE to assess levels of expression. Detection of GST-a4(N) (with expected size of 71.3 kDa) purified from soluble fraction (*lanes 4-6* of Figure 5) indicates that the fusion protein remained soluble and the highest level of the expression appeared with the 4 hr-induction (*lane 5*). However, the overall yields were relatively low for all samples collected.

Thirdly, expression trials using a different *E. coli* strain, AD494(DE3), were attempted under 4 hr-induction at 30°C. As demonstrated by Figure 6, the yield of soluble GST-a4(N) produced from using AD494 was even lower than that from BL21. There was obviously insufficient protein produced for further structural study.

Based on the above trials, large-scale expression was performed as follows: 6 litres of BL21 host was grown to $OD_{600}$ of 0.8 at 37°C, the incubation temperature was lowered to 30°C for 30 min, then IPTG was added (final concentration 0.1 mM), and incubation was continued for a further 4 hrs. After cell lysis, samples of soluble and insoluble fractions, flow-through before wash and eluate were analysed by SDS-PAGE and Western blotting (Figure 7). From Figure 7*B*, it can be seen that there was roughly equivalent amounts of insoluble and soluble

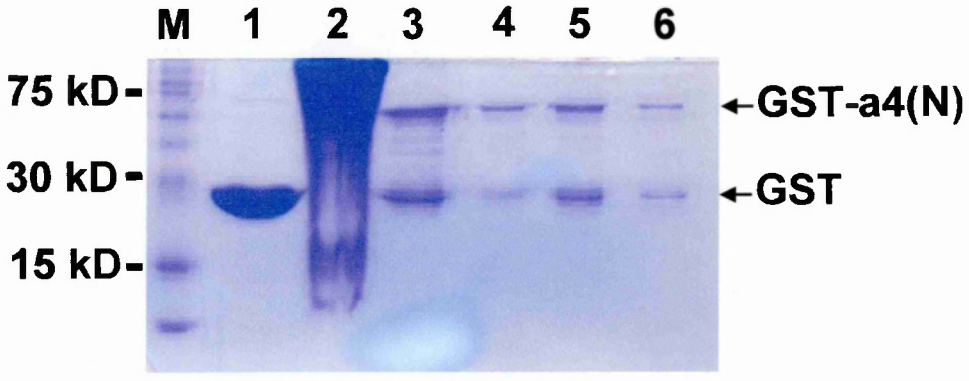



**Figure 5 SDS-PAGE analysis.**

SDS-PAGE (12%, stained with coomassie brilliant blue R250) analysis of the expression and purification of GST-a4(N). This figure demonstrates the effect of different lengths of induction time on yields of the GST fusion protein. The culture was induced with 0.1 mM IPTG at 30°C for various lengths of time. Soluble sonicates of each sample were purified and eluted. *Lanes M*: full-range rainbow Marker, *1*: GST alone, *2*: insoluble sonicate, *3*: beads after elution, *4-6*: eluates from samples with 2, 4 and 5 hr-induction.

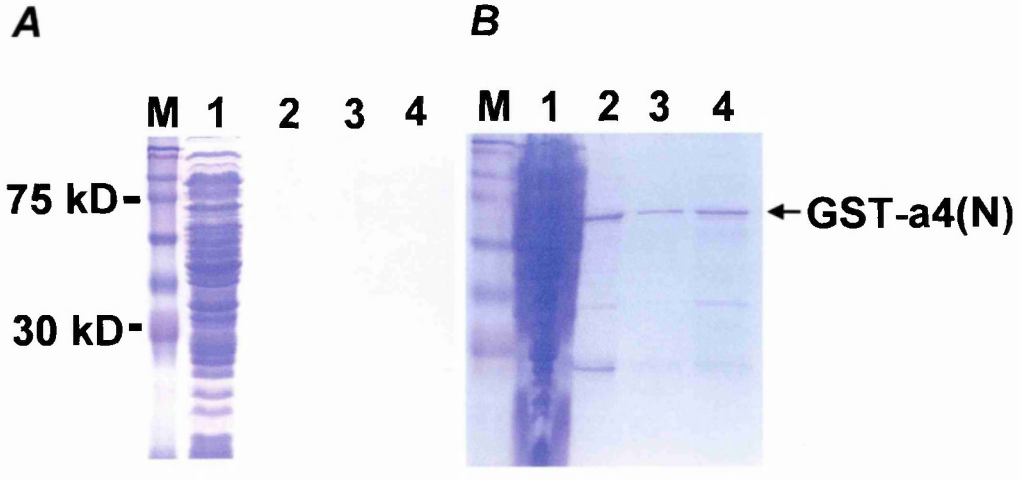



**Figure 6 SDS-PAGE analysis.**

SDS-PAGE (12%, stained with coomassie brilliant blue R250) analysis of the expression trials of GST-a4(N). This figure demonstrates the effect that different bacterial strain on yields of the protein. The culture was induced with 0.1 mM IPTG for 4 hrs at 30°C and soluble sonicates of each sample were purified. *Panels* ***A*** and ***B*** show the expression with AD494 and BL21, respectively. *Lanes M*: full-range rainbow marker, *1*: insoluble sonicate, *2* and *3*: the 1st and 2nd eluates, *4*: beads after elution.

GST-a4(N) produced (*lanes 1* and 2). Production of insoluble protein indicates the existence of unfolded or misfolded forms of the protein. The yield of soluble GST-a4(N) was low, approximately 50 μg per litre of bacterial culture expressed, as measured by Bio-Rad Protein Assay.

*CD Spectroscopy Analysis of GST-a4(N)*

CD spectroscopy was used to probe the secondary structure of the purified GST-a4(N) (work undertaken by Dr Timothy Dafforn, Department of Haematology, University of Cambridge). The analysis was carried out with 100 μl of 0.22 mg/ml GST-a4(N) in 0.1 M phosphate buffer (pH 8.0). GST alone, which was produced in parallel, replaced GST-a4(N) as the control. Figure 8*A* shows the CD spectra of GST and GST-a4(N). The spectrum of GST shows a large negative band around 208 nm and a small negative band around 222 nm, indicating considerable α-helical content as well as some β-sheet structure compared with the standard CD spectra (Figure 8*B*) (Perczel et al., 1991; Perczel et al., 1992a, b). This result is consistent with the published structure of GST (Masino et al., 2002). However, the CD spectrum of GST-a4(N) exhibits a very weak negative band around 205 nm and the curve gradually flattens out and approaches zero mdeg of CD absorbance. This spectrum is analogous to the standard CD spectrum of irregular structure (Figure 8*C*), suggesting that no clear secondary elements were detected in this sample. In other words, the GST-a4(N) might not have been folded properly in solution, although it was soluble.

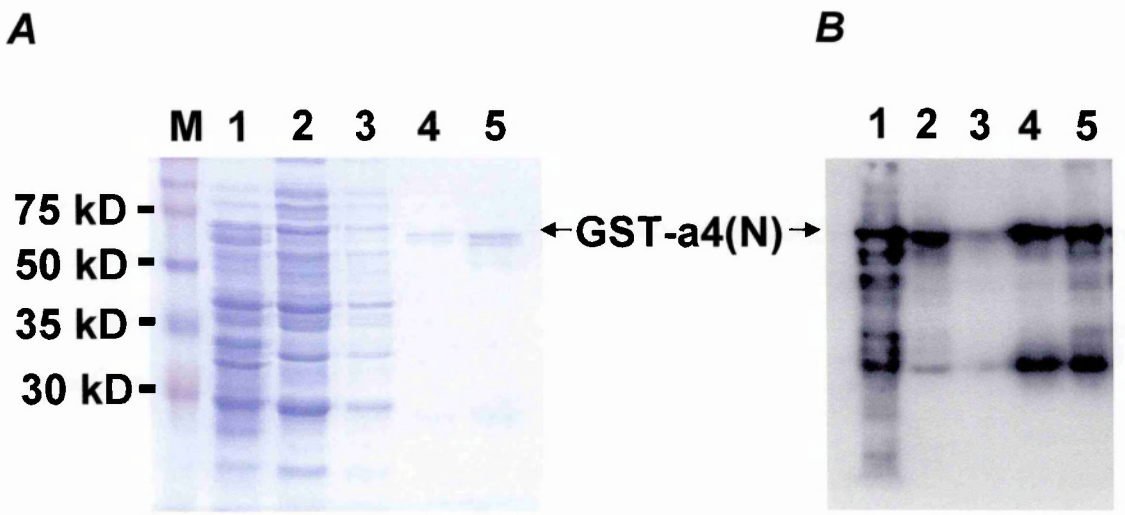



**Figure 7 SDS-PAGE and Western blot analysis.**

SDS-PAGE (12%, stained with coomassie brilliant blue R250, *panel **A***) and Western blot (*panel **B***) analysis of large-scale expression, purification and detection of GST-a4(N). The culture was induced with 0.1 mM IPTG for 4 hrs at 30°C. *Lanes M*: full-range marker, *1* and *2*: soluble and insoluble sonicates, *3*: flow-throw before washes, *4*: eluate, *5*: beads after elution. Western blot was probed using anti-GST antibody (1 in 1000 dilution).

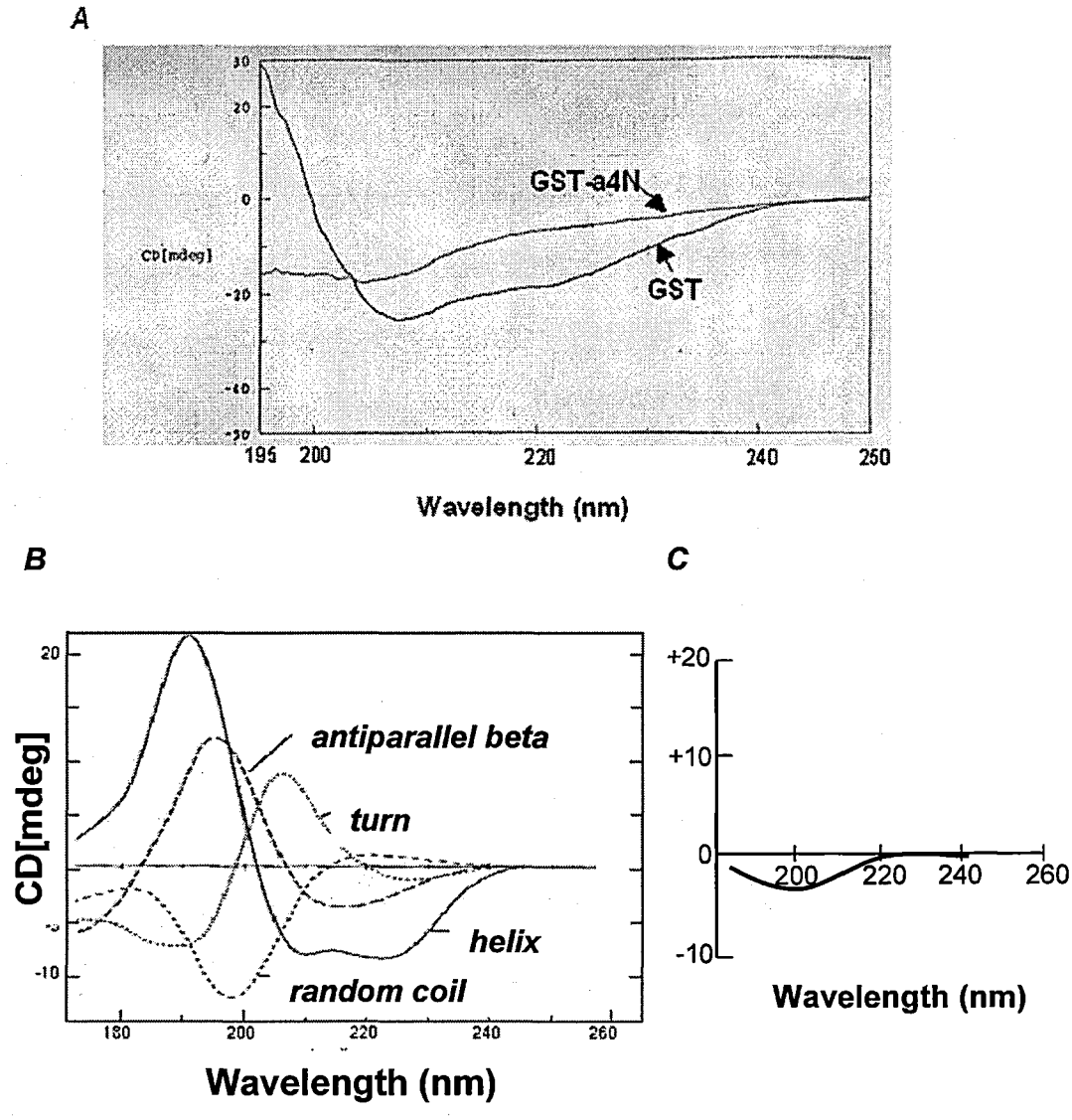



**Figure 8 Far-UV CD analysis.**

*Panel **A**,* spectra of GST-a4(N) and GST proteins. Data points were collected at protein concentration of 0.22 mg/ml, suspended in 0.1 M phosphate buffer pH 8.0, 22°C. Each spectrum is the averaged result of ten spectra. Standard curves of known secondary structure elements and irregular structure of proteins are shown in *panels **B*** and ***C**,* respectively [for references, see (Perczel et al., 1991; Perczel et al., 1992a, b)].

### 3.2.1.2 Using *Drosophila* S2 Cell Line

Further attempts were made to express a4(N) by using *Drosophila* S2 Cells as follows.

*Construction of a4(N) into pMT/Bip/V5-His A Expression Vector*

The cDNA sequence encoding a4(N) was amplified by high fidelity LR-PCR using primer 5 (forward) and primer 6 (reverse) (Table A.3). The resulting 1.191kb *Spe*I-*Age*I PCR fragment was cloned into the *Spe*I and *Age*I sites of the pMT/BiP/V5-His A vector. The subsequent identification of positive clones and verification of inserts were carried out as described in Section 3.2.1.1. The recombinant plasmid containing correct *a4(N)* sequence, termed pMT-a4(N), was used to transfect S2 cells as follows.

*Transfection and Selection of Stable Cells*

Purified pMT-a4(N) plasmid was co-transfected into the S2 cells together with pCoHygro vector which was used for selection, in a mass ratio of 19:1. The empty pMT/BiP/V5-His A vector replaced pMT-a4(N) as the control. After a few weeks of selection with hyg.B, stably transfected cells were obtained. To check whether the transfected plasmid DNA had integrated into the chromosomes of the host, genomic DNA was made from either the stable S2 cell line transfected with pMT-a4(N) or from the control. Using the genomic DNA as a template, a standard PCR was carried out using primers 5 and 6 (Table A.3) and the results were analysed on agarose gel (Figure 9). A band with the expected size of *a4(N)* (*lanes 4* and *5*) is seen from the sample of S2 cells transfected with pMT-a4(N), demonstrating that pMT-a4(N) plasmid not only transfected into the cells, but also integrated into the genomic DNA of these cells. This band is not present in the control (*lanes 1* and *2*), indicating the specificity of the PCR.

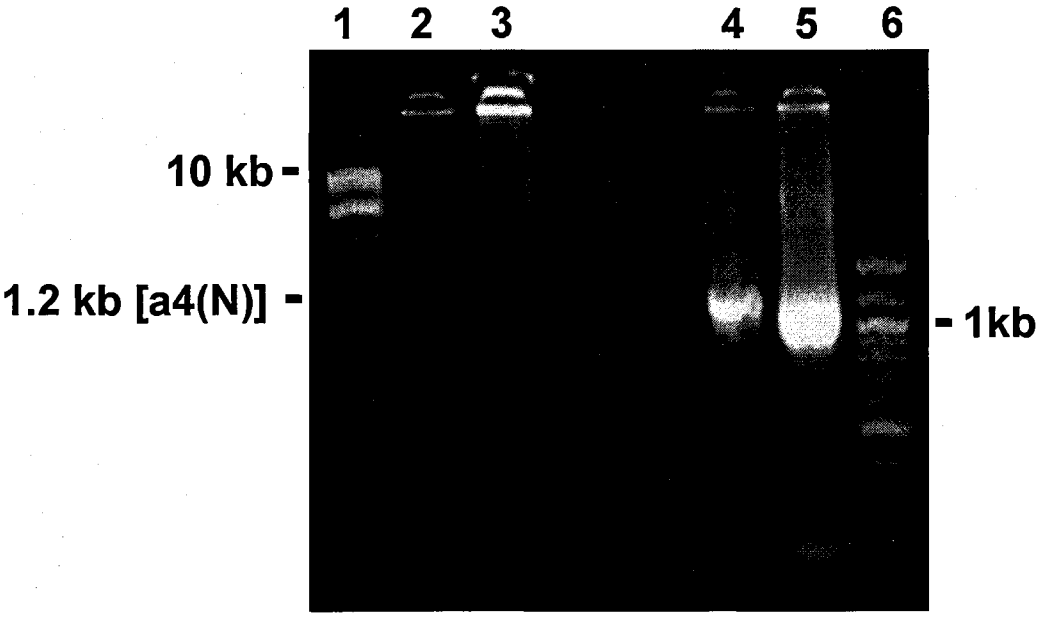



**Figure 9 Agarose gel analysis.**

Agarose gel (2%) analysis of PCR products from genomic DNA of pMT-N transfected S2 cells and the control. *Lanes 1* and *6*: 1 kb and 100 bp ladders, respectively, *4* and *5*: 2 μl and 8 μl of PCR products from pMT-N transfected S2 cells, *2* and *3*: 2 μl and 8 μl of PCR products from the control.

*Expression and Purification*

4x 15 ml of the stably transfected S2 cells were induced with $CuSO_4$ and both culture media and cell pellets were collected every 24 hrs immediately after induction up to the $5^{th}$ day. No protein bands with the expected size (~ 46 kDa) of His-a4(N) were seen on SDS-PAGE for all the collected samples (data not shown). The media collected were then mixed, purified (with Ni-NTA agarose resin under native conditions) and concentrated (with Vivaspin 20 column). All samples, including the cell extracts and eluates, were analysed by SDS-PAGE and Western blotting. No visible bands of this protein could be seen on SDS-PAGE (data not shown). When Western blotting was carried out using anti-histidine (anti-His) antibody (Figure 10), a 46 kDa band was seen in the concentrated sample purified from the mixed media (*lanes 1*). However, a band with a similar size also appeared in the control (*lane 2*), suggesting that this band is not specific. Similarly, no a4(N) was detected in the cell extract containing transfected pMT-a4(N) (*lanes 3*) compared to the control (*lane 4*).

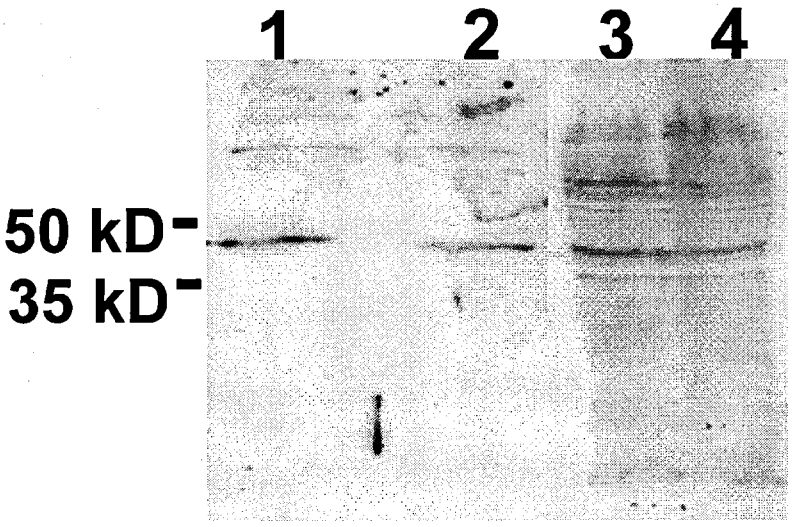



**Figure 10 Western blot analysis.**

Western blot analysis using anti-His antibody (1 in 1000 dilution) for expression of His-tagged a4(N) in *Drosophila* S2 cells. *Lanes 1*: mixture of media collected from samples taken at 1-5 days post induction from pMT-a4(N) transfected S2 cells, *2*: mixture of media collected from the control after induction, *3* and *4*: cell extracts from the pMT-a4(N) transfected and control cells.

### 3.2.2 Expression and Purification of a4(Loop2)

*Construction of a4(Loop2) into miniprseta-mac Expression Vector*

The cDNA sequence encoding a4(Loop2) was amplified by high fidelity PCR using primer 7 (forward) and primer 8 (reverse) (Table A.3). The resulting 234 bp *Bam*HI-*Eco*RI PCR fragment was cloned into the *Bam*HI and *Eco*RI sites of miniprseta-mac vector. The subsequent transformation (into BL21), identification of positive clones and verification of inserts were carried out as described in Section 3.2.1.1. The recombinant plasmid containing the correct *a4(Loop)* sequence, termed prseta-a4(Loop2), was subjected to expression trials.

*Expression and Purification*

Small-scale expression trials to screen for the best-expressed colonies and to optimise expression conditions were performed prior to large-scale expression. For the screening, six randomly chosen colonies, each harbouring a prseta-a4(Loop2) recombinant, were cultured and expressed at 37°C for 2 hrs. After cell lysis, both insoluble and soluble sonicates from each individual sample were analysed on SDS-PAGE (Figure 11). As demonstrated by the SDS-PAGE, a high yield of His-tagged a4(Loop2) fusion protein [(His-a4(Loop2)] was produced from all the colonies, although one of them expressed at a slightly lower level (*lane 12*). However, almost all of the produced protein remained in the insoluble fraction (*lanes* with even numbers) and this type of insoluble product is normally termed an inclusion body. Several other attempts were made to obtain at least a significant portion in the soluble fraction by changing expression vectors as well as using a variety of induction conditions. These included using pGEX-4T1 instead of miniprseta-mac, variations in IPTG concentrations (0.05, 0.1, or 0.2 mM) and induction times (0.5, 1, or 2 hrs) as well as different expression temperatures (25,

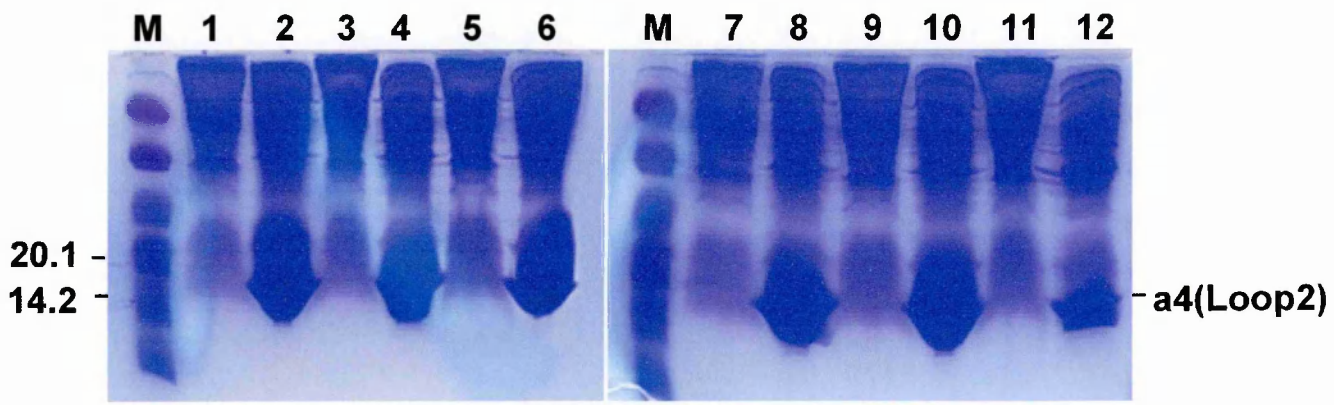



**Figure 11 SDS-PAGE analysis.**

Analysis of colony screen for expression of His-a4(Loop2). The culture inoculated from each colony (6 colonies in total) was induced with 0.1 mM IPTG for 2 hrs at 37°C. Soluble and insoluble sonicates of each sample were analysed on SDS-PAGE (16%) stained with coomassie brilliant blue R250. Lane M is low-range Marker, all lanes with odd numbers are soluble sonicates, and those with even numbers are insoluble sonicates.

30 or 37°C). However, none of these conditions tested could increase the solubility of the protein (data not shown). It was therefore decided to prepare His-a4(Loop2) from the inclusion bodies by a refolding strategy.

Insoluble His-a4(Loop2) was produced in a 500 ml volume culture induced with 0.2 mM IPTG at 37°C for 2 hrs. After cell lysis, the insoluble sonicate containing His-a4(Loop2) was solubilised in 6 M urea and after centrifugation, the supernatant was purified using Ni-NTA column under denaturing conditions. Samples were collected at several stages during purification and analysed by SDS-PAGE (Figure 12*A*) and Western blotting (Figure 12*B*). The highly expressed and purified His-a4(Loop2) was demonstrated by SDS-PAGE (*lanes 2* and *3/4*, before and after purification). An extra band approximately double the size of His-a4(Loop2) was detected (*panel B*), suggesting the existence of a dimer. The yield of the purified denatured His-a4(Loop2) was approximately 50 mg per litre of bacterial culture.

To refold the purified protein, the sample was first diluted to a concentration of less than 50 μg/ml and dialysed using 3,500 MWCO dialysis tubing against various buffers: 1) PBS (pH 7.4); 2) 20 mM Tris (pH 8.0); and 3) PBS containing 300 mM NaCl, 10% glycerol, 0.005% Tween 20, 0.1% β-mercaptoethanol, 5 M urea (pH 7.4). The concentration of glycerol and urea in buffer 3 was then gradually reduced, e.g. 1% and 0.5 M lower in the successive dialysis buffers changed every 12 hrs. However, precipitation occurred immediately when the protein was dialysed against both the 1$^{st}$ and 2$^{nd}$ buffers. Although no precipitate was seen until the urea concentration in buffer 3 had fallen to 2 M, precipitation occurred very rapidly when the urea concentration was further reduced to 1 M. As a urea-containing buffer was not suitable for use in this experiment, guanidine-HCl was used instead as the solubilisation reagent followed by dialysis against non

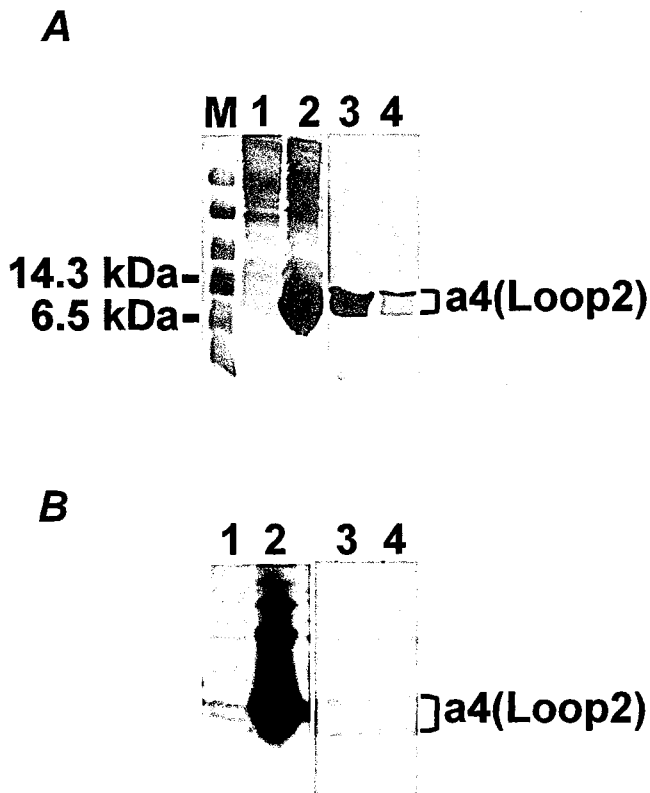



**Figure 12 SDS-PAGE and Western blot analysis.**

SDS-PAGE (16%, *panel **A***, stained with coomassie brilliant blue R250) and Western blot analysis (*panel **B***) of the expression and purification of His-a4(Loop2). The culture was induced with 0.2 mM IPTG for 2 hrs at 37°C. The insoluble fraction was solubilised with 6 M urea followed by purification under denaturing conditions. *Lanes M*: low-range rainbow marker, *1*: cleared lysate, *2*: insoluble lysate, *3* and *4*: the 1st and 2nd eluates. Western blots were probed using anti-His antibody (1 in 1000 dilution).

detergent sulfobetaines (NDSB) containing buffer as described by Vuillard *et. al* (Goldberg et al., 1996; Vuillard et al., 1995). However, protein precipitation also appeared heavily during the following dialysis.

Progress in making soluble His-a4(Loop2) was achieved through further attempts in which a Protein Refolding Kit was employed as described in Section 2.15. His-tagged a4(Loop2) inclusion bodies were produced in 500 ml volume culture (Figure 13*A*) and highly purified from insoluble materials (Section 2.15.2) as shown in Figure 13*B* (*lane 3*). To investigate refolding of the inclusion bodies, they were first solubilised in a buffer containing N-lauroylsarcosine, and then dialysed against 20 mM Tris-Cl (pH 7.4) and 0.1 mM DTT in order to remove the detergent. No visible precipitation occurred during dialysis and a high yield, about 20 mg per litre of culture, of soluble His-a4(Loop2) was obtained. Figure 13*C* demonstrates the final confirmation of refolded a4(Loop2) by Western blotting. This sample was subjected to structural analysis as follows.

*CD Spectroscopy Analysis*

CD spectroscopy was used to determine whether the soluble His-a4(Loop2) protein had secondary structure (work undertaken by Dr Timothy Dafforn, Department of Haematology, University of Cambridge). The analysis was carried out with 100 μl of the 0.8 mg/ml His-a4(Loop2) and the spectrum is shown in Figure 14. The CD spectrum of His-a4(loop2) exhibits a strong negative absorbance at about 215 nm, which is analogous to the standard CD spectrum of antiparallel β-sheets (Figure 8*B*). It indicates that this protein is folded with a significant portion of β-sheets in solution. Therefore, the refolded His-a4(loop2) was used in further studies.

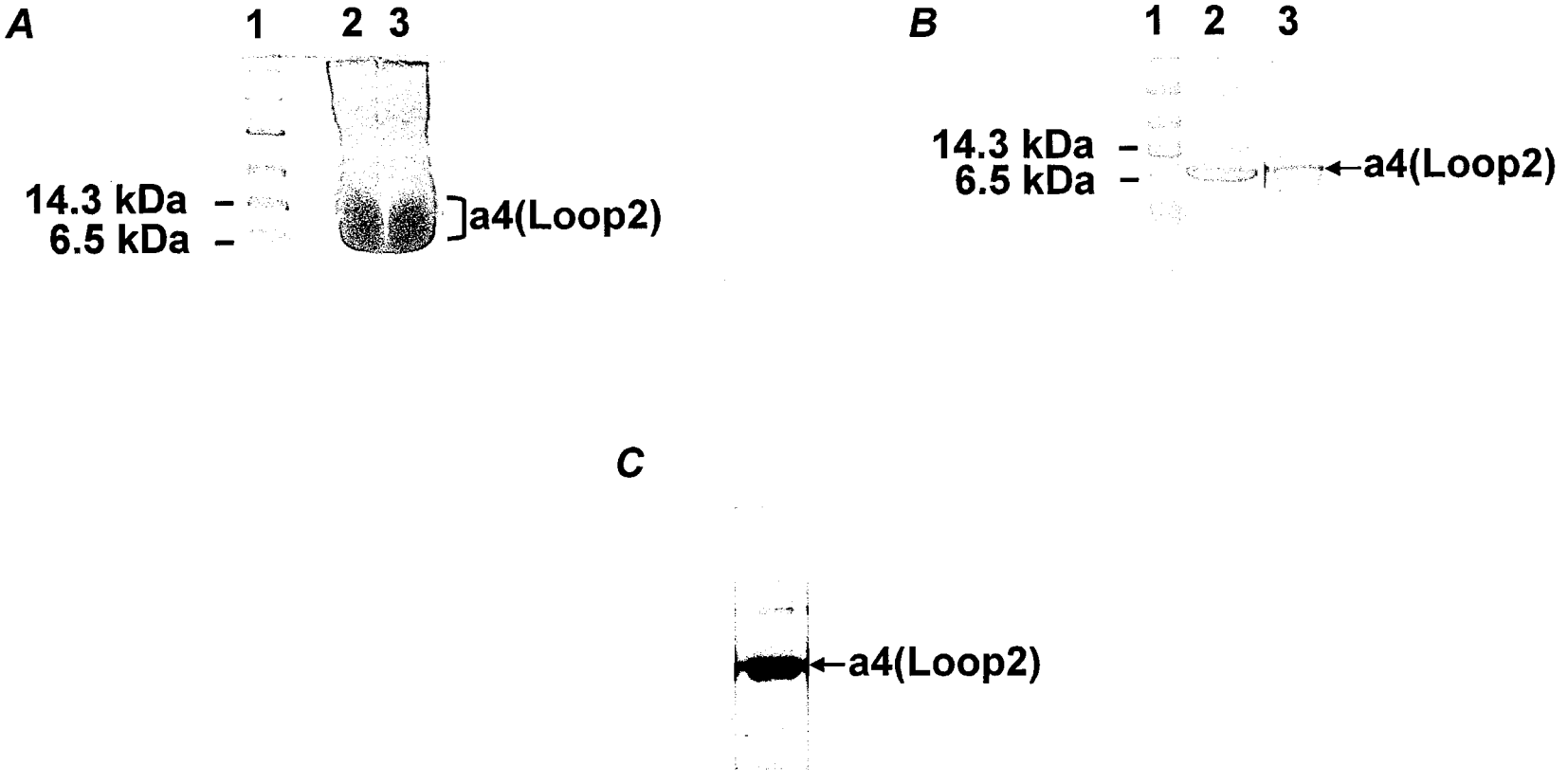



**Figure 13 SDS-PAGE and Western blot analysis.**

SDS-PAGE (16%, *panels* ***A*** and ***B***, stained with SimplyBlue Safe Stain reagent) and Western blotting (*panel* ***C***) of expression, purification, refolding and detection of His-tagged a4(Loop)2. *Lanes 1* are low range Marker. ***A***, *lanes 2* and *3* are two preparations of initial inclusion bodies. ***B***, *lanes 2* and *3* are different stages of purification of the inclusion bodies, and *lane 3* is the final purified and also the refolded one. ***C***, Western blot analysis of the refolded a4(Loop2) using anti-His antibody (1 in 1000 dilution).

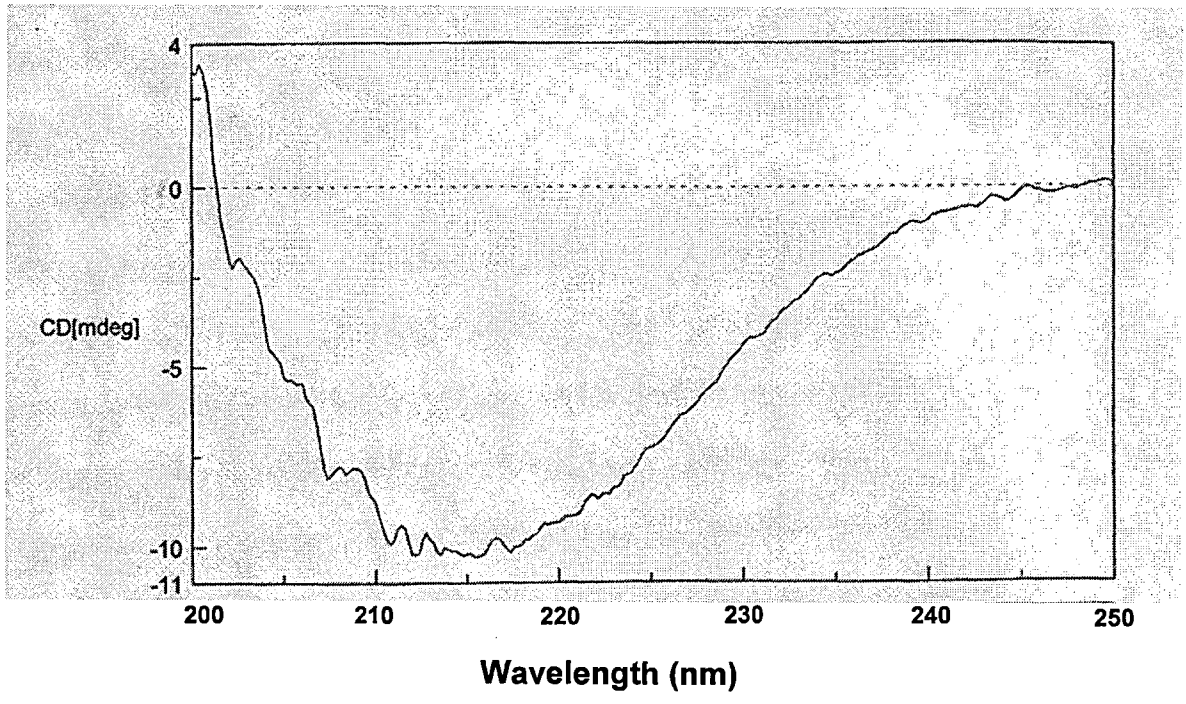



**Figure 14 CD spectroscopy.**

Far-UV CD spectrum of the refolded His-tagged a4(loop2) protein. Data points were collected at protein concentration of 0.8 mg/ml, suspended in 20 mM Tris-Cl, 0.1 mM EDTA buffer (pH 7.4), 22°C. This spectrum is the averaged result of ten spectra.

### 3.2.3 Expression and Purification of a4(C)

*Construction of a4(C) into pGEX-4T1 Expression Vector*

cDNA encoding the last 45 amino acids of a4 [a4(C)] was amplified by high fidelity PCR using primer 11 (forward) and primer 12 (reverse) (Table A.3). The resulting 141 bp *Bam*HI-*Eco*RI PCR fragment was cloned into the *Bam*HI and *Eco*RI sites of the pGEX-4T1 vector. The subsequent transformation (into BL21), identification of positive clones and verification of inserts were carried out as described in Section 3.2.1.1. The recombinant containing correct *a4(C)* sequence, termed pGEX-a4(C), was subjected to expression trials.

*Expression, Purification, and Detection*

Prior to large-scale expression, the small-scale expression trials listed below were performed to screen for the best-expressed colonies and to optimise expression conditions.

Firstly, three randomly chosen BL21 colonies, each harbouring the pGEX-a4(C) plasmid, were screened for the GST fused a4(C) protein [GST-a4(C)] as described for GST-a4(N). A Western blot was probed with the polyclonal antibody RA2922, directed against the last 14 amino acids. This confirmed the production of GST-a4(C) in all samples collected and there were no visible differences in expression levels between the three colonies (data not shown).

Secondly, the effect of different induction times (2, 3, 5 hrs) on the yield of GST fusion protein was investigated as described for GST-a4(N). The empty vector, i.e. without *a4(C)* insert, provided a control. The same amount of each eluate was analysed on both SDS-PAGE and Western blotting using RA2922 antibody. The Western blotting (Figure 15A) demonstrated the production of a protein with a molecular weight of approximately 31 kDa, corresponding to the recombinant GST-a4(C) with a predicted molecular weight of 31.54 kDa. Also the

highest level of expression was from the 3 hr-induction (*lane 3*). The detection of GST-a4(C) purified from the supernatant of cell lysate indicated that the protein remained soluble. However, bands with the similar size (31 kDa) could not be observed on SDS-PAGE (Figure 15*B*), indicating that the overall expression level of the fusion protein was very low. In addition, there were bands (*lanes 1-3* Figure 15*B*) of the same size (~26 kDa) as the control GST (*lane 4*) which were not recognised by RA2922 antibody (Figure 15*A*), suggesting they were GST alone. This result indicates that the purified protein contained mainly GST and only small portion (approximately less than 5% estimated from SDS-PAGE analysis, data not shown) of GST-a4(C) fusion protein.

Finally, other attempts were made to increase the expression level of a4(C) by cloning its coding sequence into miniprseta-mac vector followed by expression in BL21. However, no obvious improvements were obtained (data not shown).

As the GST-a4(C) fusion protein was expressed in this system only at very low levels, a large-scale fermentation technique was employed to obtain sufficient a4(C) protein (work was carried out by Ms. Galina Dimitrova, The MRC Dunn Human Nutrition Unit, Cambridge). Based on the above trials, fermentation was carried out in a 30 litre volume, with induction with 0.1 mM IPTG for 3 hrs at 30°C using BL21.

Similarly, the two missense mutations previously identified in this *a4(C)* region from two different patients diagnosed with rdRTA (R807Q or G820R) (Table 2) (Smith et al., 2000; Stover et al., 2002), were also expressed in order to compare with a4(C) in further studies carried out later (detailed in Chapter 5). In order to do this, site-directed mutagenesis using pGEX-a4(C) as a template

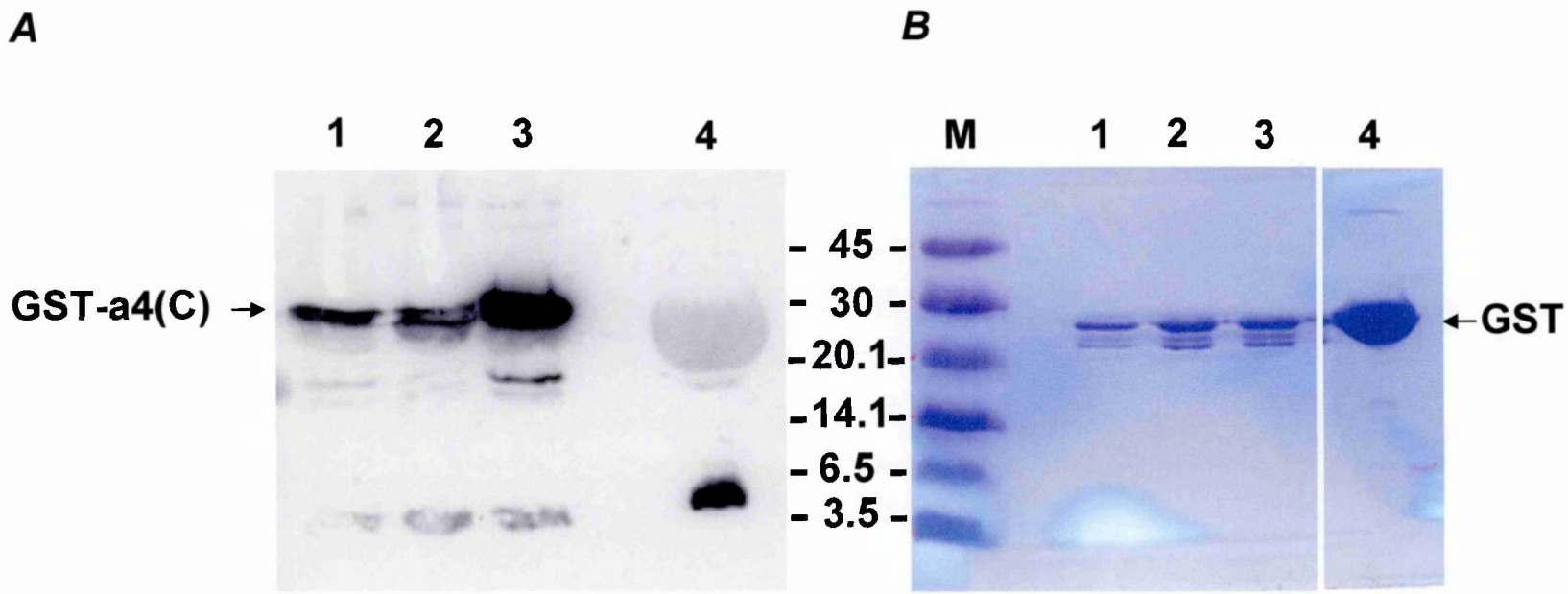



**Figure 15 SDS-PAGE and Western blot analysis.**

SDS-PAGE (16%, stained with coomassie brilliant blue R250) (*panel **B***) and Western blotting (*panel **A***) of expression trials of GST-a4(C). These results demonstrate the effect of different lengths of induction time on yields of the GST fusion protein. The culture was induced with 0.1 mM IPTG at 30°C for various lengths of time. Soluble fractions of each sample were purified. *Lanes M*: Marker, *1-3*: eluates from 2, 5 and 3 hr-induction, *4*: GST alone. Western blot was probed with RA2922 antibody (1 in 1000 dilution).

was carried out, as described in Section 2.9, to create a single base substitution in codon 807 (CGA → CAA) or in codon 820 (GGG → AGG) which then resulted in the single amino acid change R → Q or G → R, respectively. The obtained constructs were verified by DNA sequencing (Figure 16) and used to transform BL21 cells followed by fermentation under the same conditions as those used for pGEX-a4(C). The proteins produced are designated as pGEX-a4(C)-G (for mutation G820R) and pGEX-a4(C)-Q (for mutation R807Q).

During fermentation, bacterial cell growth was monitored (Figure 17). As demonstrated from Figure 17, before the addition of IPTG (*arrowed*), a normal exponential bacterial cell growth was observed for all three proteins. However, a while after the addition of IPTG, the increase of cell growth was slowing down for all samples. This change in cell growth is very likely due to production of the proteins. Interestingly, both mutant samples, especially pGEX-a4(C)-G, showed more effects on cell growth than that of wild type.

Following cell lysis, fusion proteins were purified from the soluble fraction using a GSTrap column, under native conditions. Eluates were concentrated and then thrombin-digested to remove GST before being submitted to HPLC purification. Samples of fusion protein eluates (concentrated), thrombin digests and the a4(C) [or a4(C)-G or a4(C)-Q] proteins from HPLC purification were analysed by SDS-PAGE (Figure 18*A*) and Western blot analysis (Figure 18*B*). The analysis of GST-a4(C), demonstrated the presence of the fusion protein (Figure 18*A*, *lanes 4*, *5*, *6* and *1, 2, 3*, and Figure 18*B*, *lanes 1* and *2*, before and after thrombin digestion) in both intact and degraded forms. A band corresponding to a4(C) is evident in the post-thrombin-digested sample, as shown by

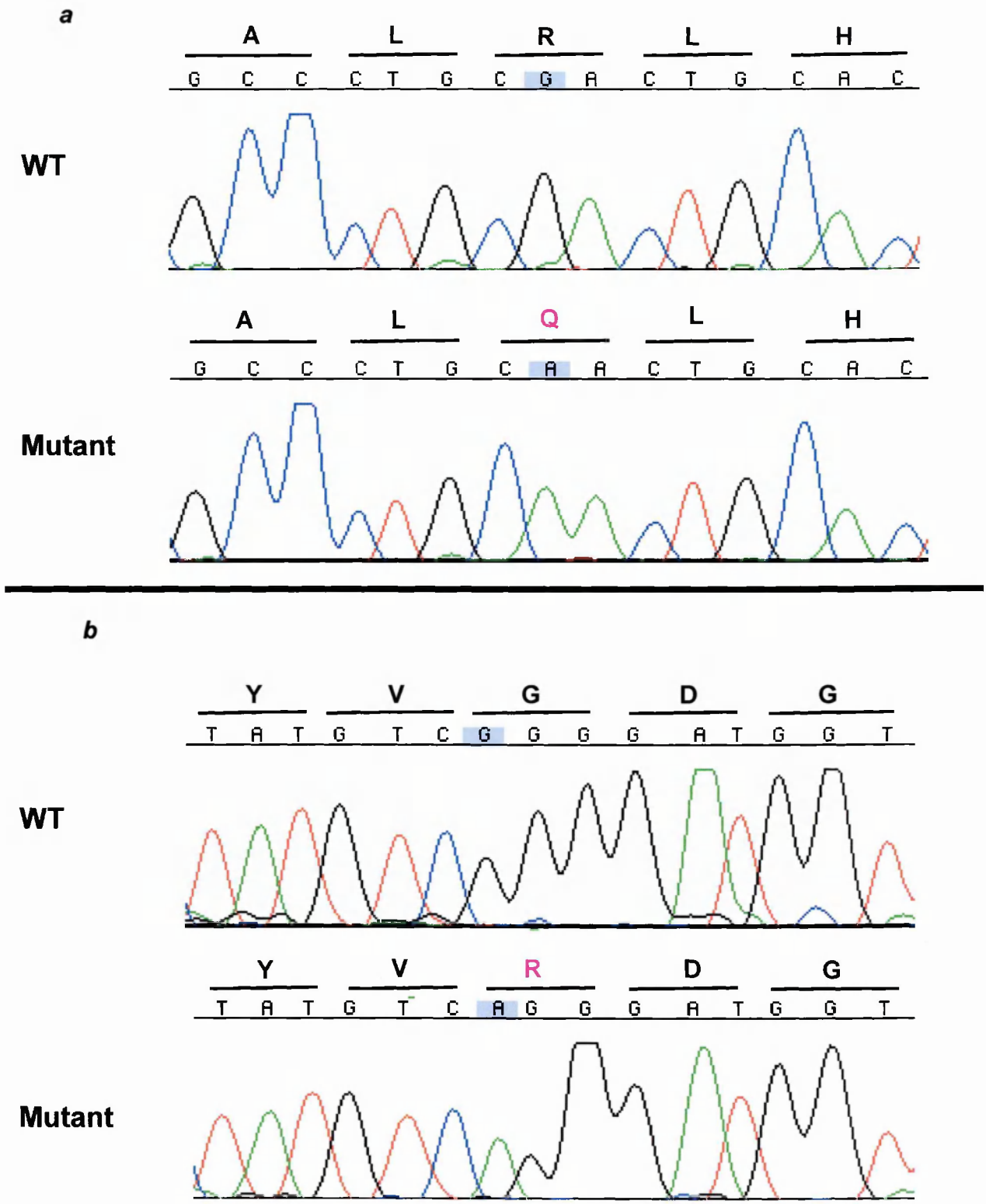



**Figure 16 Site-directed mutagenesis.**

Mutations created in a4(C) using site-directed mutagenesis method. *Panel* ***a*** shows a single base substitution in codon CGA → CAA results in changing of arginine to glutamine [a4(C)-Q]. *Panel* ***b*** shows a single base substitution in codon GGG → AGG results in changing of glycine to arginine [a4(C)-G].

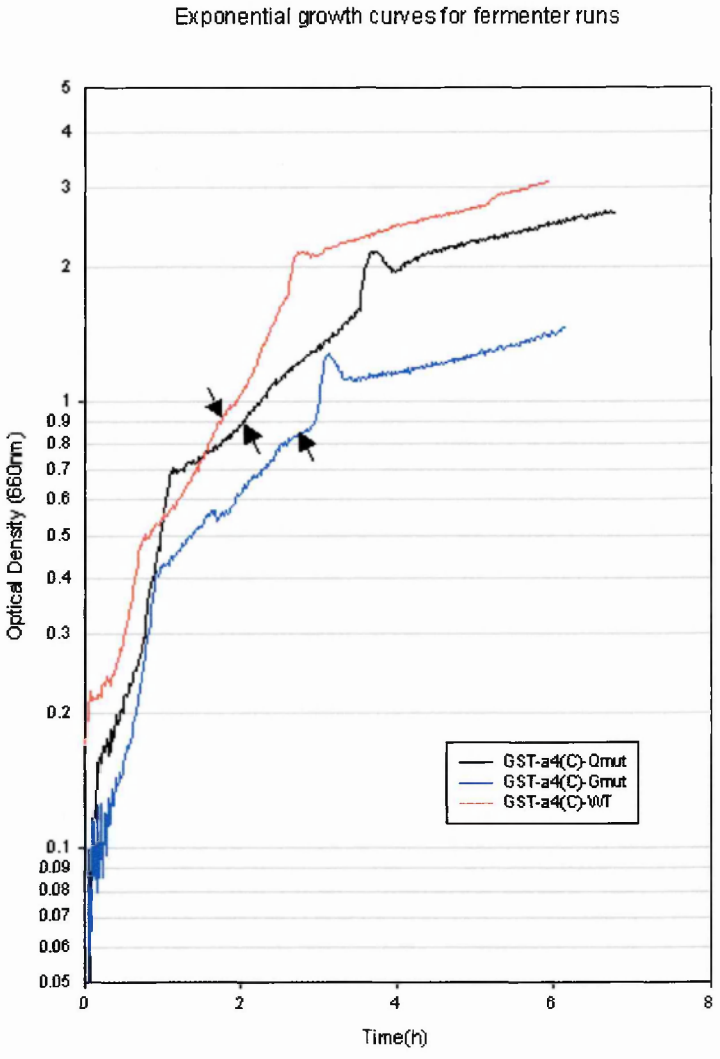



**Figure 17 Comparison of exponential growth.**

Comparison of exponential growth curves during fermentation for expression of GST-a4(C), GST-a4(C)-G and GST-a4(C)-Q. The cultures were induced at $OD_{600}$ ~0.8, with 0.1 mM IPTG for 3 hrs at 30°C. *Arrows*, the points of addition of IPTG.

Figure 18*B* (*lane 2*). This thrombin-digested product was subjected to HPLC purification. Results from HPLC (Figure 19) demonstrated that of the three main peaks observed, the second peak (*arrowed*) corresponded to recombinant a4(C). The first peak represented a fraction of smaller mass than expected, but N-terminal sequencing confirmed that this was indeed part of a4(C). This short form indicates the presence of an alternative thrombin cleavage site within a4(C). The third peak corresponded to non-digested GST-a4(C) fusion protein. As can be seen in Figure 20, further passage of the second fraction through an HPLC column linked to a mass spectrophotometer confirmed a single peak representing the a4(C) fragment, with the expected mass of 5.41 kDa. This corresponded well to the predicted size of 5.264 kDa, the difference being accounted for by two additional amino acids, glycine and serine, which were introduced by the *Bam*HI restriction site. As expected, this fragment was recognised by RA2922 (Figure 18*B*, *lane 3*), further confirming its identity. The two mutant proteins showed similar purification and thrombin digestion profiles (data not shown). Moreover, the HPLC-purified a4(C) and the two mutants were all subjected to N-terminal sequencing which confirmed their identity. The final yield of HPLC-purified proteins varied between 0.2-0.4 mg from a 30-litre volume culture and these purified proteins were used in further studies described in the following chapters.

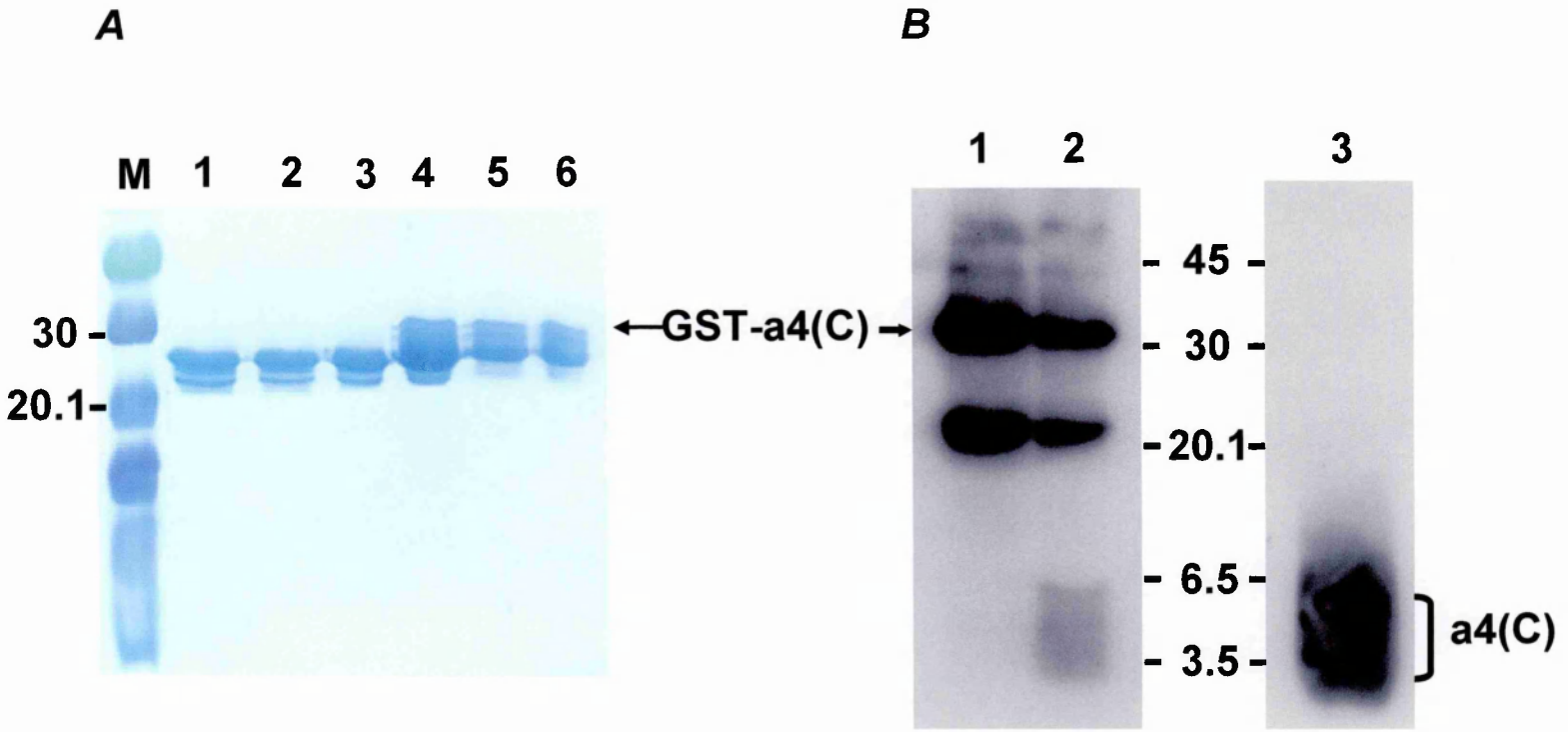



**Figure 18 SDS-PAGE and Western blot analysis.**

Expression, purification and detection of constructed GST-a4(C) fusion protein. ***A***, SDS-PAGE (16%) analysis stained with SimplyBlue Safe Stain reagent. *Lanes M*: low range Marker, *1-3*: thrombin digested GST fusion protein, *4-6*: non-thrombin digested GST fusion protein. ***B***, Western blot analysis of non-thrombin digested GST fusion protein (*lane 1*) and thrombin digested GST fusion protein (*lane 2*) and HPLC purified a4(C) (*lane 3*). Blot was probed with RA2922 antibody (1 in 1000 dilution). Data for a4(C)-G and a4(C)-Q mutant proteins are of similar quality, but not shown.

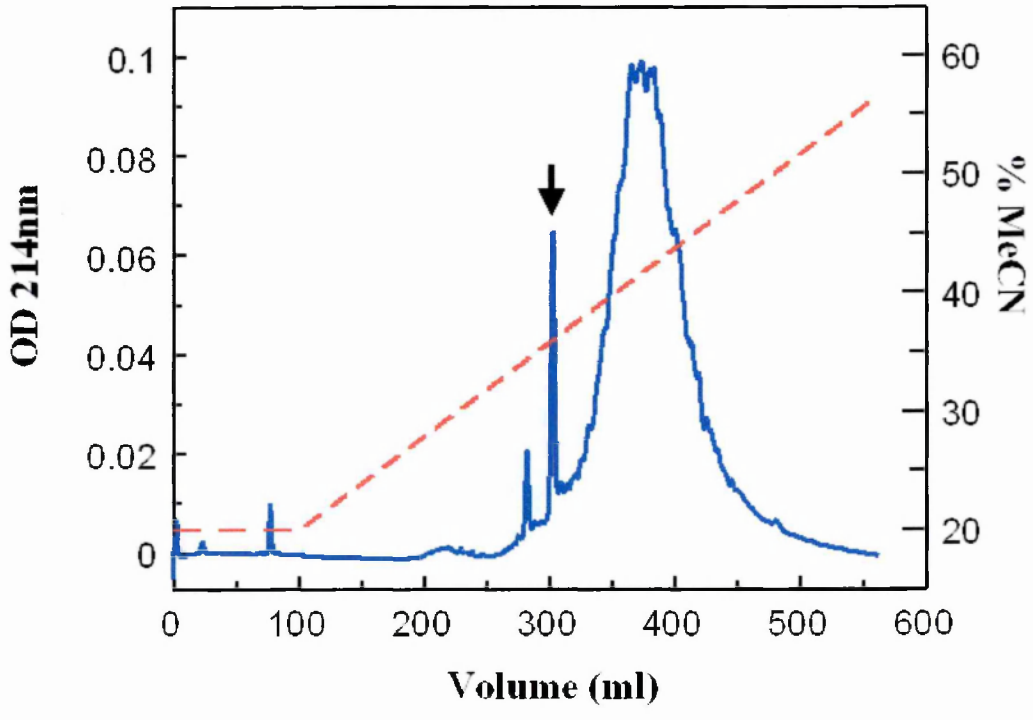



**Figure 19 HPLC purification.**

Following thrombin digestion, the mixture containing a4(C) was subjected to HPLC purification. The *arrowed* peak represents the fraction containing intact a4(C) protein. Data for a4(C)-G and a4(C)-Q mutant proteins are of similar quality, but not shown.

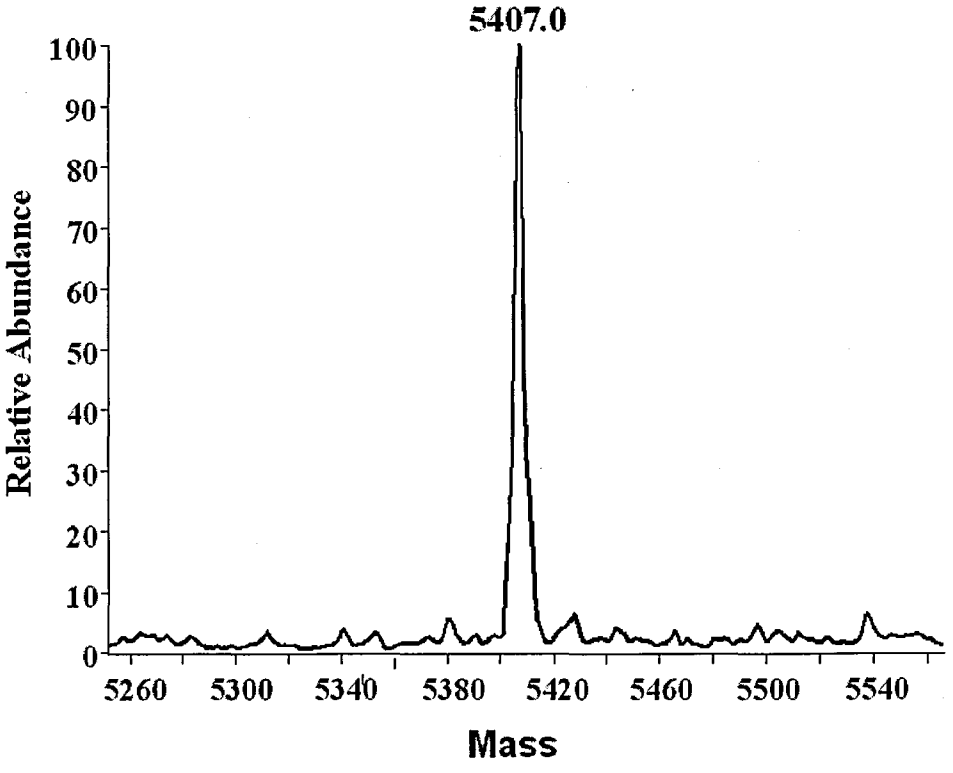



**Figure 20 Mass spectroscopy.**

Following thrombin digestion, the mixture containing a4(C) was subjected to HPLC purification. The fraction containing intact a4(C) protein was further analysed by LC-MS. Its molecular mass, 5.41 kDa, confirms it as a4(C). Data for a4(C)-G and a4(C)-Q mutant proteins are of similar quality, but not shown.

*CD Spectroscopy Analysis*

As such low yields of the a4(C), a4(C)-G and a4(C)-Q proteins were obtained, it was decided later to synthesise them for certain further studies described in Chapter 5. The polypeptides were synthesised by CovalAB and analysed by CD spectroscopy to probe the secondary elements of the proteins. All products were N-acetylated, HPLC-purified, and were all completely water-soluble. Protein concentrations were determined using UV absorption measurement, with calculated extinction coefficients at 280 nm for all three proteins. The analysis was carried out with 200 μl of 100 μM of each protein in $H_2O$ plus 5% (v/v) PBS (pH 7.4) and results are shown in Figure 21. The CD spectra of all three proteins exhibit two large negative bands around 208 nm and 222 nm. These spectra are analogous to the standard CD spectrum of α-helix (Figure 8*B*), suggesting these proteins mainly have α-helix structure in solution. It was noticed that both mutants showed much stronger CD signals than that of wild type. This indicates a higher level of α-helix in the mutants compared to the wild type. In addition, the CD signal around 208 nm is relatively stronger than that of 222 nm in the spectrum of a4(C)-G compared to the other two proteins, suggesting a higher content of coils in this protein.

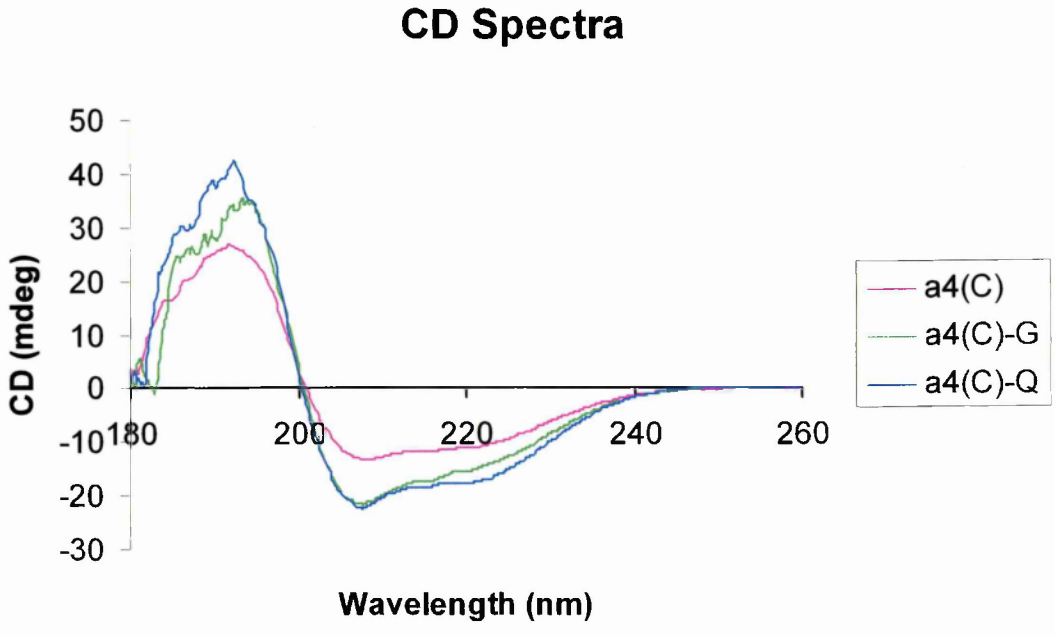



**Figure 21 CD spectroscopy.**

Far-UV CD spectra of a4(C), a4(C)-G and a4(C)-Q. Data points were collected at protein concentration of 100 μM, suspended in $H_2O$ containing 5% (v/v) PBS (pH 7.4), 20°C. Each spectrum is the averaged result of ten spectra.

## 3.3 Discussion

Modern molecular biological techniques make it possible to produce individual domains of multidomain proteins as independent molecules, allowing the investigator to study the structures and functions of individual folded domains. Due to the advantage of simplicity and a relatively high level of expression, a bacterial *E. coli* cell expression system was chosen as the first choice, for the expression of a4(N), a4(Loop2) and a4(C) domains of the human a4 protein. The *E. coli* BL21 strain was initially chosen because it is defective in OmpT and Lon protease production, meaning it is able to aid in the expression of fusion proteins by minimising the effects of proteolytic degradation by the host (Miroux and Walker, 1996). All three proteins were produced in the above system, but each of the initial products differed in yield, solubility and folding aspects.

Although the production of recombinant fusion protein in *E. coli* is well established, there are numerous factors which may present obstacles for successful expression and purification of foreign proteins (Smith and Johnson, 1988). In other words, successful expression of foreign proteins in *E. coli* depends upon many events, which collectively result in obtaining high yields of a soluble protein whose functions have been maintained throughout the expression and purification processes. As individual proteins vary widely in their toxicity, solubility, and susceptibility to degradation in the *E. coli* environment, the conditions required to obtain optimal yields of a soluble and active protein need to be empirically determined for each construct. A variety of growth and expression parameters also need to be investigated.

Host strains may in some cases help to increase the yield of intact soluble fusion proteins. As seen in the expression of GST-a4(N), expression level of soluble GST-a4(N) was higher in BL21 than that of AD494. In 1998, Saluta and

Bell had observed the highest yield of GST-luciferase fusion protein in BL21 compared to other strains tested (Saluta, 1998). The higher level of expression in BL21 cells might be due to its ability to minimise the effects of proteolytic degradation, which in turn increases protein stability. In addition, the yields of both soluble GST-a4(N) and GST-a4(C) varied with different periods of induction period. However, overall yields of these two proteins, including their soluble and insoluble forms, were both very low.

The possible reasons for extremely low yields of a4(C) could be because: firstly, it was seen that the cell growth slowed down after addition of IPTG during fermentation. This change is presumably caused by production of the protein, suggesting a4(C) might be toxic to the cells. Another interesting aspect was that the GST alone always appeared in a much larger fraction along with the production of GST-a4(C). Whether or not this phenomenon also implies that a4(C) might be toxic to the cells remains unclear. However, if we assume a4(C) is toxic, then in order to protect themselves, the cells may digest the a4(C) part from its initial product of GST fusion protein. This effect reduces the levels of GST fusion protein. Secondly, relatively large amounts of GST alone in cell lysates influenced purification efficiency of GST-a4(C) after cell lysis. This would also reduce the level of purified fusion protein. Thirdly, it is difficult for thrombin digestion to reach completion. In addition, the existence of an internal digestion site of thrombin made it more difficult to control the amount of enzyme used. The amount of thrombin added resulted in either incomplete digestion or over digestion with the creation of short forms of a4(C), and both decrease the final yield of a4(C).

However, a high-level production of foreign fusion proteins in *E. coli* does not always indicate a successful outcome, because it often results in the formation of inclusion bodies, as demonstrated by expression of His-a4(Loop2). Although the

formation of inclusion bodies can be used as a means to purify an expressed fusion protein, which may otherwise be unstable in the soluble fraction, the conditions needed to refold the inclusion bodies afterwards can be highly variable.

Apart from the yield and solubility aspects, another important aspect is protein folding. When proteins are being synthesised, they must adopt the correct conformation for their functions. Proteins may either fold spontaneously or they may need the assistance of folding factors, such as chaperons, to gain the correct final conformations.

When folded, proteins are normally soluble in aqueous solutions. However, this is not always the case and *vice versa*. As seen in the case of GST-a4(N), although large-scale expression in BL21 obtained reasonable amounts of the soluble protein for structural analysis, the CD spectrum obtained suggested no clear natural secondary structure elements in the sample. In other words, this protein might be unfolded or misfolded. CD is an excellent method for analysing the conformation of proteins and peptides in solution. Each of the three basic secondary structures of a polypeptide chain (helix, sheet, and turn) shows a characteristic CD spectrum and the standard curves of CD spectra were first published in 1969 by Greenfield and Fasman (Greenfield and Fasman, 1969). Although those are actually for poly-lysine in different conformations, only little improvement in the accuracy of fits has been achieved by attempting to generate other standard data sets from protein spectra of known structure (Johnson, 1990). The standard spectra (Figure 8*B*) are theoretical curves of the six pure components derived by the convex constraint analysis method (Perczel et al., 1991; Perczel et al., 1992a, b). The analysis of CD spectra can yield valuable information about the secondary structure of biological macromolecules. GST is a well-characterised protein of known structure and its CD spectrum is typical of an

α/β protein, with predominance of the α signal (Masino et al., 2002). My results regarding the structure of GST are consistent with the known ones. Due to possible unfolding or misfolding of GST-a4(N), the solubility of this protein might benefit from a GST tag. However, whether or not the produced a4(N) itself would still remain soluble after removal of the GST tag is unknown.

Attempts were also made to express the protein using another *E. coli* strain: AD494. This strain contains thioredoxin reductase (trxB) mutants which enable disulphide bond formation in the cytoplasm, providing the potential to produce properly folded, active proteins. This advantage is lacking in many other commonly used *E. coli* strains, including BL21, where the reducing environment of the cytoplasm prevents the generation of protein disulphide bonds (Derman et al., 1993). Although, to date, no 3D structural information is available for the a subunit of the proton pump in any species, having nine cysteine residues in the a4(N) region predicts a high likelihood of intrachain disulphide bridges. In many proteins, disulphide bonds (intra- or interchain or both) contribute to stabilising their native structures. If this is also the case for a4(N) its unfolding or misfolding, after production in BL21 cells, might partly be the consequence of lacking properly formed disulphide bonds. Therefore, using AD494 instead of BL21 might have helped reduce misfolded GST-a4(N) by formation of proper disulphide links. This in turn would increase the level of soluble GST-a4(N) by reducing the levels of insoluble ones. However this is not the case as it was found that the amount of soluble and insoluble proteins produced still remained similar. In addition, the yield of the soluble GST-a4(N) was even lower in AD494 than in BL21. It might be too perfunctory to say that this result suggests no intradomain disulphide bonds within a4(N). However, seeing the results above perhaps implies other post-translational modifications are necessary for the proper folding of this protein. It was

disappointing to not obtain a soluble protein with structural elements and it reinforces the importance of checking the folding situation of an overexpressed protein before using it in further studies.

In addition to the folding problem of a4(N), a4(Loop2) also generated difficulties. In contrast to a4(N), a4(loop2) was produced at a significant level, but mainly entered an insoluble inclusion body fraction after cell lysis. Despite using a variety of procedures (different expression vectors, growth temperatures, induction conditions) that have previously been shown to enhance the amount of soluble protein that would otherwise be inclusion bodies, a4(Loop2) still remained insoluble after initial expression. I therefore resorted to solubilisation of the protein from inclusion bodies. Inclusion bodies are insoluble aggregates of misfolded protein. Although they can be easily purified, solubilisation can usually only be obtained by using strong denaturing conditions, which normally causes a major problem later in achieving efficient folding *in vitro*. The renaturation process may not always be achieved, as demonstrated by solubilisation of His-a4(Loop2) inclusion bodies with urea followed by dialysis against different buffers. In the mid 90s, Vuillard *et al* reported successful refolding of protein from inclusion bodies by using a non-detergent reagent, NDSB (Goldberg et al., 1996; Vuillard et al., 1995). Following this technique, some other proteins are also reported to have been successfully refolded (Benetti et al., 1998; Maiorano et al., 2000; Ochem et al., 1997). However, application of this method to His-a4(Loop2) inclusion bodies also failed. Nevertheless, attempts to refold a4(Loop2) from inclusion bodies was finally successful using a Protein Refolding Kit (Novagen), where a mild detergent, N-lauroylsarcosine, that aids solubilisation of membrane proteins was used during the solubilisation stage.

Although the folding situations of the expressed GST-a4(N) and His-a4(Loop2) were investigated, I did not check the relative aspect of the expressed a4(C). This is mainly due to, a) insufficient final protein product [a4(C)] obtained from an initial fermentation run, and b) a4(C) is completely water-soluble. For a polypeptide that is small, not tagged, and water soluble, it is more likely to be folded. This was confirmed by CD analysis using synthetic a4(C). All three types of a4(C) measured by CD spectroscopy were folded and mainly existed as α-helices in solution. However, the substitution of R to Q (in mutant a4(C)-Q) or G to R (in mutant a4(C)-G) increased the α-helix content. This is consistent with published data showing the Conformational Preferences of Amino Acids (Williams et al., 1987; Wilmot and Thornton, 1988).

Attempts to express a4(N) using a different system, *Drosophila* S2 cells, also failed to produce detectable a4(N). The possible reasons could be: firstly, low levels of expression, which could be caused by low transfection efficiency. Secondly, it may be due to this protein being produced in an unfolded or misfolded form, which would subsequently be degraded by the ER quality control in S2 cells. Thirdly, failure to produce a4(N) on both bacterial and *Drosophila* cell systems might imply that certain specific mechanisms necessary for proper folding of this protein are scarce in the two systems used. Fourthly, as a4(N) is only a part of a4 protein, its folding might also rely on some other parts of a4, such as the existence of interdomain disulphide links. However, no conclusion can be made for this unfolding or misfolding property until further structural information becomes available.

In conclusion, a4(C) and a4(Loop2), but not a4(N), were successfully produced and purified using bacterial expression systems, although the yield of a4(C) was very low. As a result, both purified a4(C) and a4(Loop2) were used to

screen phage display random peptide libraries for potential binding partner(s). Attempts were made to produce a4(N) in both bacterial and insectile expression systems. However, no clear structure was detected from the GST-tagged a4(N) produced in bacterial system (although it is soluble) and also no a4(N) was detected from the insectile expression system. The inability to obtain folded a4(N) prevented further investigation at the *in vitro* level with the purified protein, such as usage of the phage display method to identify ligand(s). However, I was able to use a yeast two-hybrid system to search for potential binding partner(s) of a4(N), as this method does not rely on the purified protein. Investigations using phage display and yeast two-hybrid systems are described in the following chapter.

# CHAPTER 4

# SCREENING FOR BINDING PARTNERS

## 4.1 Introduction

Despite the importance of the a subunit, very little is known about the protein-protein interactions involving this subunit. In order to understand the biological functions of this subunit, I sought further to explore the roles of the protein through identification of its potential-binding partners.

Protein-protein interactions are critical to all cellular processes, and understanding them is crucial to understanding any biological system. Therefore, the identification of protein-protein interactions can help to determine the biological functions of proteins, by situating them relative to other proteins in cellular pathways or functional classes (Walhout and Vidal, 2001). Two groups of strategies, physical and library-based, are normally employed for the identification of protein-protein interactions. The physical strategy is a group of methods for the *in vitro* identification of interactions and it involves the use of proteins with no linkage to their encoding DNA. These methods include affinity binding, co-immunoprecipitation and chemical crosslinking. This strategy is generally applied in the validation of protein interactions determined by library-based strategy. The library-based strategy includes *in vitro* and *in vivo* methods, and both depend on the ability to link polypeptides to their encoding DNA. The *in vitro* methods use libraries in which polypeptides are on the outside of the cell where they are accessible for binding to exogenous ligands. The most widely used *in vitro* method is phage display (Sidhu et al., 2000). The *in vivo* methods can be defined as those in which the protein-protein interaction under study is detected inside of living

cells. The most widely used *in vivo* method is the yeast two-hybrid (Y2H) system (Fashena et al., 2000).

In recent years, the phage display approach has become an increasingly popular *in vitro* selection technique in protein research (Sidhu et al., 2000). This method was first demonstrated with the *E. coli*-specific M13 bacteriophage (Smith, 1985). Based on the idea, several other *E. coli* phage have been adapted for phage display and eukaryotic systems have also been developed (Possee, 1997; Ren and Black, 1998; Santini et al., 1998). In the case of phage display, a foreign peptide (or protein) fused to a bacteriophage coat protein is displayed on the surfaces of phage particles that also contain the cognate DNA packed inside of the phage. These particles can be selected for according to the affinity of the displayed peptide (or protein) to certain ligands (target proteins), and then the gene encoding the peptide (or protein) can be identified by sequencing. The greatest advantage of this method is that large numbers of clones (greater than $10^9$ different displayed sequences) can be easily screened. Furthermore, the direct linkage between an observed phenotype and encapsulated genotype allows fast determination of selected sequences. This powerful tool has been successfully used in a number of applications, including epitope mapping, determining enzyme specificity, and exploring protein-protein (or protein-DNA) interactions (Baltrusch et al., 2001; Hong and Boulanger, 1995; Scott and Smith, 1990). In this study, M13 phage display libraries (NEB) containing up to $10^{11}$ members were screened by a simple *in vitro* selection procedure called "biopanning" (Figure 22). $10^{10-11}$ phage particles were used for each panning, and after extensive washing with a mild detergent, bound phage were eluted for propagation in *E. coli*. A single round of selection can enrich for binding phage by 20-1000 fold (Marks et al., 1991). The multiple rounds of selection provides a powerful strategy of enriching for specific

binding phage. The specificity of the binding between the selected phage and target protein can be verified by ELISA using anti-M13 phage antibody.

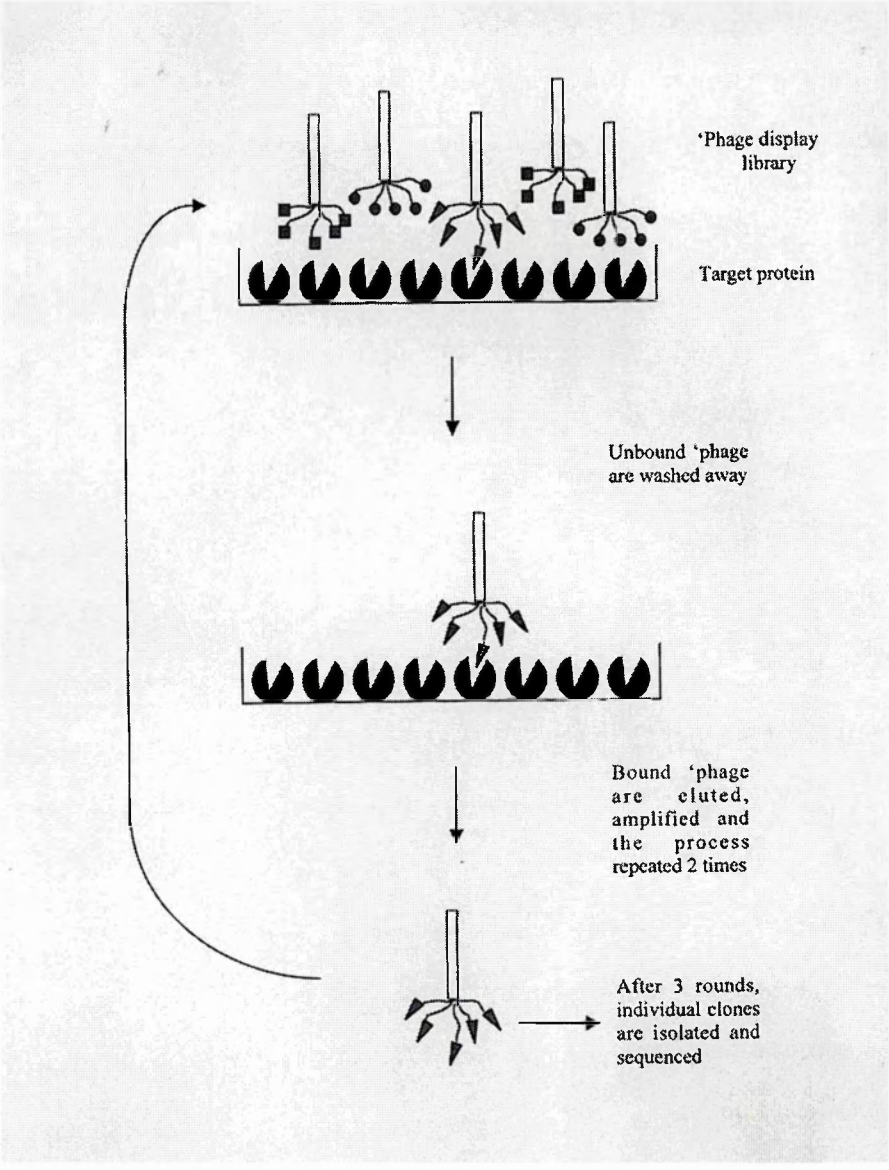



**Figure 22 Outline of phage display procedures.**

The library containing phage-displayed peptides is incubated in a plate coated with target protein. Unbound phage particles are washed away and the specifically bound phage is eluted, amplified and submitted for the next run. Details are from NEB.

In addition to phage display, another powerful approach for identification of protein interacting ligands is an *in vivo* selection using the Y2H system. Conventional Y2H systems relying on transcriptional activation of reporter genes in the nucleus to detect protein-protein interactions were first established by Field *et*

*al* in the late 1980s (Fields and Song, 1989). Since then, several different Y2H systems have been developed (Drees, 1999; Fashena et al., 2000; Vidal and Legrain, 1999). One of these systems is the recently developed protein recruitment strategy (such as Sos recruitment) (Aronheim, 2000; Huang et al., 2001). The Sos recruitment system is commercialised as CytoTrap™ (Stratagene) and was used in this study. This system greatly increases the opportunities for finding unique protein-protein interactions by taking the search for these interactions to the cytoplasm, rather than to the nucleus. In this system, proteins are expressed in the cytoplasm where, unlike in the nucleus, they may undergo post-translational modifications. Figure 23 illustrates the principle of the CytoTrap two-hybrid system.

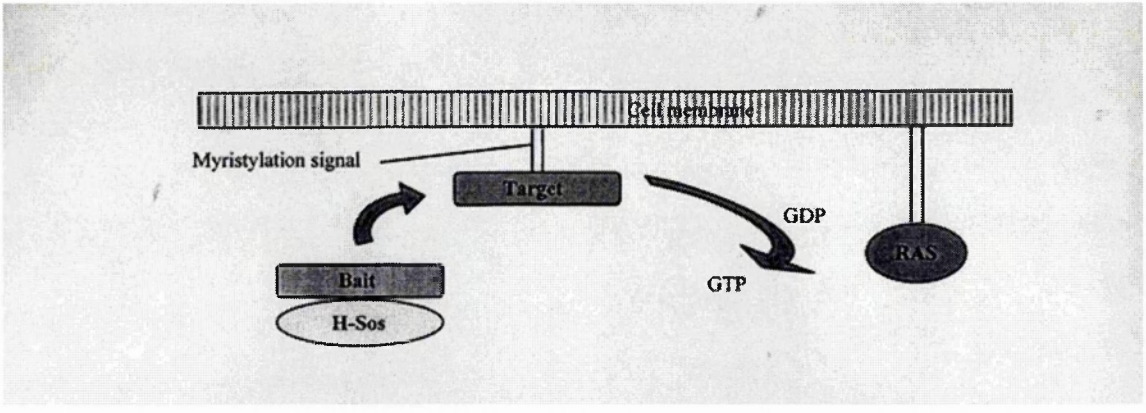



**Figure 23 Principle of the CytoTrap two-hybrid system.**

Target protein is anchored on the yeast cell membrane through Myr protein. hSos protein is localised onto the membrane via bait-target interaction, where it activates the Ras pathway which permits mutant yeast cdc25H strain to grow at 37°C. Details are from Stratagene.

In the CytoTrap two-hybrid system, two plasmids (Table A.4) are constructed, one (pSos) encoding protein of interest (bait) fused to the human Sos [hSos, a Ras guanyl nucleotide exchange factor (Ras GEF)], and the other one (pMyr) encoding protein (target) fused to a v-Src myristylation membrane-

localisation signal (Myr). The plasmids are co-transformed into a temperature-sensitive yeast mutant strain, cdc25H, which contains a mutation in the *Cdc25* gene, which encodes yeast Ras GEF (Petitjean et al., 1990). Because of the mutation, the cells can only grow at 25°C but not at 37°C. The Myr-target fusion is targeted to the yeast plasma membrane. Interaction between the bait and the target protein recruits hSos to the yeast plasma membrane where it complements the cdc25 mutation by activating the Ras signalling cascade. The interaction is detected through growth of the yeast cells at the restricted temperature (37°C). The pSos vector contains *ADH1* promoter driving the expression of the hSos-bait which is constitutively active. The pMyr vector contains a *GAL1* promoter driving the expression of the Myr-target fusion which is only induced by galactose. In addition, the pSos and pMyr vectors also carry, respectively, yeast biosynthetic genes *LEU2* and *URA3* for selection of yeast transformants based on nutritional requirements.

This chapter describes identification of potential binding partners of a4(N), a4(Loop2) and a4(C) using phage display and CytoTrap two-hybrid systems. In the case of phage display analysis, two random peptide (7-mer and 12-mer) M13 phage libraries were screened using either a4(Loop2) or a4(C) as a target protein. Both target proteins were produced as described in the previous chapter. In the case of Y2H analysis, either a4(Loop2) or a4(N) was used as a bait protein to screen a human testis cDNA library using the CytoTrap two-hybrid system.

## 4.2 Results

### 4.2.1 Phage Display Assays

#### 4.2.1.1 Screening Phage Display Library with a4(C)

*Identification of Potential Binding Phage-Displayed Peptide*

A 7-mer random peptide M13 phage display library was used to screen immobilised a4(C) (as described in Section 2.19.1.2) for potential interaction peptides displayed on the phage. For this purpose, HPLC purified recombinant a4(C) protein was fixed to microtitre plates through hydrophobic interactions, and subjected to three rounds of 'panning' (recovery of bound phage followed by their reapplication for the next round). Following the 3$^{rd}$ panning, 17 enriched phage plaques were sequenced. The sequence analysis yielded the peptide sequence SWLELRP, which was found in 7 out of 17 (approximately 40%) phage plaques sequenced (Table 11*A*).

*Verification of the a4(C)/Peptide Interaction*

Phage ELISA (Section 2.19.1.5) was used to verify the interaction of a4(C) with the SWLELRP-displaying phage. Phage concentration ranging from 1 x $10^{10}$ to 2.26 x $10^{11}$ pfu, employing two-fold dilutions was applied to immobilised a4(C). Results obtained (Figure 24) show a significant binding affinity of this phage displayed peptide to a4(C) protein compared to the control where no pre-coated a4(C) was present. Maximum binding was observed in the range of 1-2 x $10^{11}$ phage virions/well, which is similar to the concentration of phage virions used in each of the panning procedures in the phage display procedure.

*Identification of Candidate Ligands*

The program Basic Local Alignment Search Tool (BLAST) (http://www.ncbi.nlm.nih.gov/BLAST/) was used to search for homology between the identified amino acid sequence (SWLELRP) and protein database sequences. Using the 'short nearly exact matches' option, comparative BLAST analysis revealed at least four proteins that contain either almost the entire SWLELRP or very similar peptide sequences (Table 12). Of the four proteins identified, phosphofructokinase 1 (PFK-1), a key participant in the glycolytic pathway,

contains a sequence (at the C-terminus of PFK-1) with almost complete homology to SWLELRP. Aligning this sequence with other selected phage clone sequences revealed a longer consensus motif EQWWLKLRP (Table 11*B*). This matched region within PFK-1 is highly conserved among mammalian PFK-1 orthologues and paralogues (Table 11*C*). In addition to PFK-1, PKC and cytochrome P450 also contain sequences with relatively high homology to the peptide SWLELRP. PKC plays a key role in regulating the differentiation and growth of diverse cell types and cytochrome P450 catalyses reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids. The involvement of PKC in the dissociation of V-ATPase induced by glucose removal in yeast has been ruled out (Parra and Kane, 1998). However, a protein-protein interaction between aldolase (a glycolytic enzyme) and the E subunit of the V-ATPase has previously been reported (Lu et al., 2001). Therefore, PFK-1 was considered to be a very good candidate for an a4(C) ligand.

*A*

| Clone numbers | Selected Sequences | Peptide Sequences |
|---|---|---|
| 1 | TCT TGG CTT GAG TTG CGT CCT | SWLELRP |
| 2 | TCT TGG CTT GAG TTG CGT CCT | SWLELRP |
| 3 | TCT TGG CTT GAG TTG CGT CCT | SWLELRP |
| 4 | TCT TGG CTT GAG TTG CGT CCT | SWLELRP |
| 5 | TCT TGG CTT GAG TTG CGT CCT | SWLELRP |
| 6 | TCT TGG CTT GAG TTG CGT CCT | SWLELRP |
| 7 | TCT TGG CTT GAG TTG CGT CCT | SWLELRP |
| 8 | GAG GGT TGG CAT GCT CAT ACG | EGWHAHT |
| 9 | GAG GGT TGG CAT GCT CAT ACG | EGWHAHT |
| 10 | GAG GGT TGG CAT GCT CAT ACG | EGWHAHT |
| 11 | GAG GGT TGG CAT GCT CAT ACG | EGWHAHT |
| 12 | GAG GGT TGG CAT GCT CAT ACG | EGWHAHT |
| 13 | AAG CTG TGG ACT ATT AAG CCG | KLWTIKP |
| 14 | AAG CTG TGG ACT ATT AAG CCG | KLWTIKP |
| 15 | GCG CAT ACT CTT CCT GGG CGT | AHTLPGR |
| 16 | GCT CAT CCT CTG ATG CTT TAT | AHPLMLY |
| 17 | AAT CAG AAG GAG TAT ACG CTT | NQKEYTL |

*B*

| Peptide sequences | | | | | | | | | | | Number of clones |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | S | W | L | E | L | R | P | | | 7 |
| | | | | | K | L | W | T | I | K | P | 2 |
| | A | H | P | L | M | L | Y | | | | 1 |
| | A | H | T | L | P | G | R | | | | 1 |
| N | Q | K | E | Y | T | L | | | | | 1 |
| E | G | W | H | A | H | T | | | | | 5 |
| **E** | **Q** | **W** | **W** | **L** | **K** | **L** | **R** | **P** | | | |

*C*

| | SWLELRP<br>EQWWLKLRP |
|---|---|
| Human muscle type PFK-1 | 739 EQWWLKLRP 747 |
| Mouse muscle type PFK-1 | 739 EQWWLKLRP 747 |
| Rabbit muscle type PFK-1 | 739 EQWWLKLRP 747 |
| Human platelet type PFK-1 | 748 EQWWLKLRP 756 |
| Mouse platelet type PFK-1 | 747 EQWWLKLRP 755 |
| Rabbit platelet type PFK-1 | 748 EQWWLKLRP 756 |
| Human liver type PFK-1 | 738 EQWWLSLRL 746 |
| Mouse liver type PFK-1 | 738 EQWWLNLRL 746 |

**Table 11 Peptide sequences selected against a4(C) from the 3rd panning of the 7-mer random peptide M13 phage display library.**

***A***, Sequences of a4(C)-binding peptides selected from M13 phage display library. Phage clones that bound a4(C) were isolated after three rounds of high stringency panning from a 7-mer random peptide M13 phage display library. ***B***, *Red coloured* residues indicate contributors from both the directly selected peptide sequence SWLELRP and other related clones to a 12-mer consensus peptide, sequence EQWWLKLRP. ***C***, Conservation of the directly identified and the consensus peptide sequence within human, mouse and rabbit muscle, platelet and liver PFK-1 isozymes.

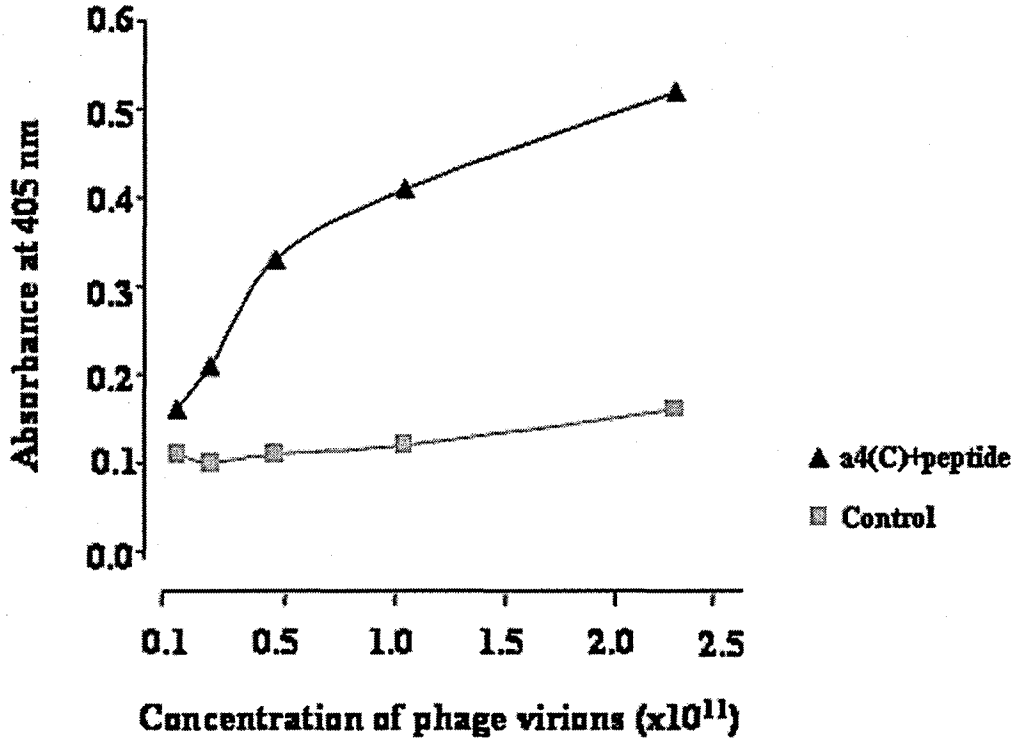



**Figure 24 Phage ELISA of output phage (SWLELRP) from the 3rd panning of the 7-mer random peptide library.**

Binding affinity of SWLELRP-displaying phage to a4(C) was determined with 2-fold serial dilutions of phage virions from 2.26 x $10^{11}$ phage/well. Specific binding of a4(C) to the peptide SWLELRP is shown.

| Protein Name* | Accession No. | Known Functions | Interacting Regions** |
|---|---|---|---|
| phosphofructokinase 1 (M-PFK-1 ) | NP_000280 | a key rate-limiting enzyme in glycolysis and represents a major control point in the metabolism of glucose | 742 **WL**K**LRP** 747 |
| phosphofructokinase 1 (P-PFK-1) | NP_002618 | The same as above | 751 **WL**K**LRP** 756 |
| Protein Kinase C-theta | NP_006248 | Plays a key role in regulating the differentiation and growth of diverse cell types | 106 **WLEL**K**P** 111 |
| Cytochrome P450 | NP_009184 | Catalyses reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids. | 7 **SWL**G**LRP** 13 |

**Table 12 Database search for proteins containing an SWLELRP-like sequence.**

Proteins were identified using programme BLAST ('short nearly exact matches' option). *Protein name was obtained from the Human Protein Database (web site). **Directly interacting regions within the ligand are numbered at the side of the sequence, and the interacting amino acids in bold.

#### 4.2.1.2 Screening Phage Display Library with a4(Loop2)

*Identification of Potential Binding Phage Displayed Peptide*

Refolded a4(Loop2) was used as a target protein to screen both the Ph.D 7- and 12-mer random peptide M13 phage display libraries (NEB) as described in Section 2.19.1.2. Following the 3$^{rd}$ panning, 19 and 21 enriched phage clones from each screen were sequenced. Of the 19 phage clones isolated from the 7-mer library screen, no specific residue pattern which was predominant over others was observed (Table 13). However, 5 out of 19 clones (approximately 26%) were found to display the sequence, K(V/L)WVIP(Q/R) (clones 1-5), although others were almost completely random (clones 6-19). This K(V/L)WVIP(Q/R) residue pattern was not clearly shown in the clones selected from the 12-mer library screen, although a few clones containing K(V/L)WVIP(Q/R)-like patterns were present (data not shown). Again, no overwhelming residue pattern emerged from this screen.

To gain further information, an additional round (4$^{th}$ panning) was carried out for both screens. 36 and 21 phage clones from each screen were sequenced. The sequence analysis from the 7-mer library screen yielded the peptide sequence KLWVIPQ, which was found in 14 out of 36 phage plaques sequenced (clones 1-14, Table 14*A*). Aligning this predominant residue pattern, KLWVIPQ, with 13 other peptide sequences (clones 15-27) revealed the sequence motif K(L/V)WVIPQ (Table 14*B*), which also appeared after the 3$^{rd}$ panning as mentioned above. This sequence motif represented about 75% (27 out of 36) of the total phage clones sequenced, indicating a large tendency of this sequence motif to react with a4(Loop2). However, the rest of the nine sequences (clones 28-36) showed no homology to the predominant residue pattern and appeared to be relatively random. Again, no residue pattern, which is predominant over others,

emerged with the 12-mer library screen after the 4th panning (Table 15). Nevertheless, the K(L/V)WVIPQ-like pattern was seen within several bound peptide sequences (bold residues).

*Verification of the a4(Loop2)/Peptide Interactions*

The interaction of a4(Loop2) with the KLWVIPQ-displaying phage was confirmed by ELISA binding assay (Figure 25). Phage concentration ranging from 0.4 to 10 x $10^{11}$ pfu was used which are similar to the concentration of phage virions used in each of the panning in the phage display procedure. Results obtained showed significant binding affinities of KLWVIPQ phage displayed peptides to immobilised a4(Loop2) protein compared to the control where no pre-coated a4(Loop2) was present. This peptide exhibited a strong interaction with a4(Loop2) that was significantly above the background level across the whole range of phage concentration used. In addition, the interaction of a4(Loop2) with the KVWTLPAHVTPR-displaying phage (Table 15, clones 15 and 16) was also confirmed by phage ELISA (data not shown).

| Clone numbers | Selected Sequences | Peptide Sequences |
|---|---|---|
| 1 | TTC CAC ACC CAA GAC GGA TCC | **KVWVLPR** |
| 2 | TTC CAA ACC TGC TAA GGA GCC | **KVWTIPR** |
| 3 | TTC GAC ACC CAC TAA GGA GTC | **KLWVIPQ** |
| 4 | TTC GAC ACC CAC TAA GGA GTC | **KLWVIPQ** |
| 5 | TTC CAC ACC TTC TAA CTA CTC | **KVW**KIDE |
| 6 | AGC GAA CTA AGC GGC CAA AGA | SLDSPVS |
| 7 | TTC TTC TAC TTC TCC GCC TTC | KKMKRRK |
| 8 | AAC GGA GGA TTA AGC GCC TGC | LPPNSRT |
| 9 | GTC GTA ATA GGA ATA TGC GCA | QHYPYTR |
| 10 | CTA GGA CAA GAA TAA GAC GAA | DPVLILL |
| 11 | GTA TGA GGA TTC GTC TTA GGA | HTPKQNP |
| 12 | AGA GTA ACC AGC GAC CGC GTA | SHWSLAH |
| 13 | CAC AAC GTC CTC TAA CTC GCC | VLQEIER |
| 14 | AAA AGC GTC GGC GTC TAA AGC | FSQPQIS |
| 15 | TGC TAC CGC GTC GTA TAC TGA | TMAQHMT |
| 16 | AAA GGC GGA GTC GTC GGC GGC | LPPQQPP |
| 17 | ATA TGA CGA ATA GGC GTC GGA | YTAYPQP |
| 18 | CGC TGA TTA AAA TAC TGA CCA | ATNFMTG |
| 19 | CCA AGC GGA CCC GAC GAA GGC | GSPGLLP |

**Table 13 Peptide sequences selected against a4(Loop2) from the 3rd panning of the 7-mer random peptide M13 phage display library.**

***A***

| Clone numbers | Sequences correspond to codon strand (5' → 3') | Peptide Sequences |
|---|---|---|
| 1 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 2 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 3 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 4 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 5 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 6 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 7 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 8 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 9 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 10 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 11 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 12 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 13 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 14 | AAG CTG TGG GTG ATT CCT CAG | **KLWVIPQ** |
| 15 | AAG GTT TGG ACG ATT CCT CGG | KVWTIPR |
| 16 | AAG GTT TGG ACG ATT CCT CGG | KVWTIPR |
| 17 | AAG GTT TGG ACG ATT CCT CGG | KVWTIPR |
| 18 | AAG GTT TGG ACG ATT CCT CGG | KVWTIPR |
| 19 | AAG GTT TGG ACG ATT CCT CGG | KVWTIPR |
| 20 | AAG GTT TGG ACG ATT CCT CGG | KVWTIPR |
| 21 | AAG CTT TGG AAG ATT CCT ACT | KLWKIPT |
| 22 | AAG GTG TGG TAG ATG ACT TAT | KVWQMTY |
| 23 | AAG GTG TGG CAG CTG CAT TCT | KVWQLHS |
| 24 | AAG GTG TGG CAG CTG CAT TCT | KVWQLHS |
| 25 | AAG GTG TGG TAG ATT AAT TCG | KVWQINS |
| 26 | AAG GTT TGG ATT ATT AAT TCT | KVWIINS |
| 27 | AAG GTT TGG TAT ATT ACG CCT | KVWYITP |
| 28 | CAG GGG CAG ACT CCG AGT ACG | QGQTPST |
| 29 | AAG AAG ATG AAG AGG CGG AAG | KKMKRRK |
| 30 | AAG AAG ATG AAG AGG CGG AAG | KKMKRRK |
| 31 | AAG AAG ATG AAG AGG CGG AAG | KKMKRRK |
| 32 | ATG TAT TCT GGG CCG ACT AGG | MYSGPTR |
| 33 | TCT ATT CTG CCG TAT CCT TAT | SILPYPY |
| 34 | TCT TCT AGT GTT GTG ACT CAT | SSSVVTH |
| 35 | TCT TCT AGT GTT GTG ACT CAT | SSSVVTH |
| 36 | TCG CTT ACT AGT ACG CAT ATG | SLTSTHM |

***B***

| Peptide Sequences | | | | | | | No. of Clones |
|---|---|---|---|---|---|---|---|
| K | L | W | V | I | P | Q | 14 |
| K | V | W | T | I | P | R | 6 |
| K | L | W | K | L | P | T | 1 |
| K | V | W | Q | M | T | Y | 1 |
| K | V | W | Q | L | H | S | 2 |
| K | V | W | Q | I | N | S | 1 |
| K | V | W | Y | I | T | P | 1 |
| K | V | W | I | I | N | S | 1 |
| **K** | **L/V** | **W** | **V** | **I** | **P** | **Q** | 27 |

**Table 14 Peptide sequences selected against a4(Loop2) from the 4th panning of the 7-mer random peptide M13 phage display library.**

***A***, Peptide sequences selected from the a4(Loop2) protein screen of the random 7-mer peptide M13 phage display library after the 4th panning. ***B***, *Red coloured* residues indicate contributors from both the directly selected peptide sequence KLWVIPQ and other related clones to the sequence motif K(L/V)WVIPQ.

| Clones | Segences for codon strand (5' → 3') | Peptides |
|---|---|---|
| 1 | TAT AGT CTG AGG GCT GAT TCT AGG TGG | YSLRADSRWMPS |
| 2 | AAG TGT TGT TAT TAT GAT CAT TCG CAT | KCCYYDHSHALS |
| 3 | AAG TGT TGT TAT TAT GAT CAT TCG CAT | KCCYYDHSHALS |
| 4 | AAG GTT TGG CCT CCG CAT CCT ATT CCT | **KVW**PPHPIPTRT |
| 5 | AAG GTT TGG CCT CCG CAT CCT ATT CCT | **KVW**PPHPIPTRT |
| 6 | ACT ATG AAG TGT TGT TAT TCG AAT ACT | TMKCCYSNTSPP |
| 7 | ACT ATG AAG TGT TGT TAT TCG AAT ACT | TMKCCYSNTSPP |
| 8 | ACT ATG AAG TGT TGT TAT TCG AAT ACT | TMKCCYSNTSPP |
| 9 | AAG GTT TGG TAT CAT ACT TGG CCG TCG | **KVW**YHTWPSKTP |
| 10 | CAT AGT CTG CGT ACG GAT TGG TCT TCG | HSLRTD**W**SS**P**SR |
| 11 | CAT AGT CTG CGT ACG GAT TGG TCT TCG | HSLRTD**W**SS**P**SR |
| 12 | AAG GTT TGG GAT TGG CAG CCG TCT CAG | **KVW**DWQPSQATV |
| 13 | ACG CTT TCG CGG AAG AAG GAT CGG TTT | TLSRKKDRFKNK |
| 14 | AAG ATT GTT CCT ACG GAT TGG GTT TCG | KIVPTDWVSARS |
| 15 | AAG GTT TGG ACG CTG CCT GCT CAT GTT | **KVW**TL**P**AHVTRP |
| 16 | AAG GTT TGG ACG CTG CCT GCT CAT GTT | **KVW**TL**P**AHVTRP |
| 17 | AAG GTT TGG GAT TTT AAG CCG CAT AAT | **KVW**DFKPHNNMY |
| 18 | GCG TCT CTG AGG CAT GAT CAT TCT CAT | ASLRHDHSHVLP |
| 19 | AAG CTG TGG ACG ATT CCT AGT AAT GAT | **KLW**T**IP**SNDYPP |
| 20 | TTG GAG GCG AAG ATT TGG GTG GTG CCT | LEA**K**I**WV**V**P**APS |
| 21 | CAT CCT CAT AAT AAG TGG CCT CCT GCG | HPHNKWPPATPT |

**Table 15 Peptide sequences selected against a4(Loop2) fron the 4th panning of the 12-mer random peptide M13 phage display library.**

Bold residues indicate contributors to the sequence motif K(L/V)WVIPQ.

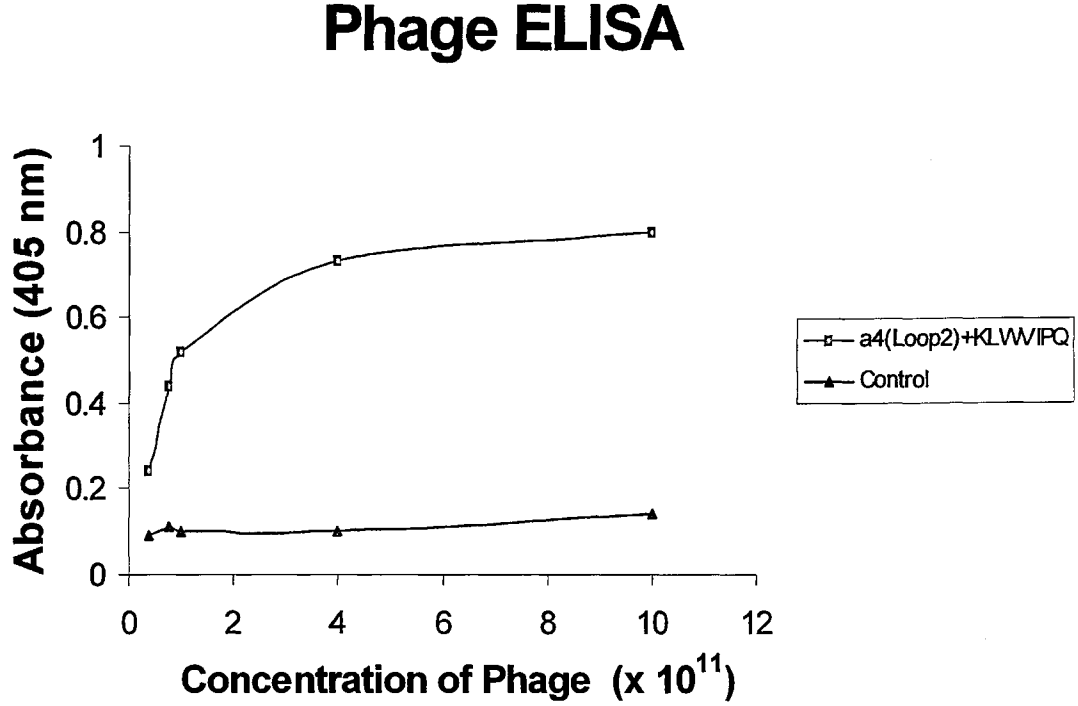



**Figure 25 Phage ELISA of output phage (KLWVIPQ) from the 4$^{th}$ panning of the 7-mer random peptide library.**

Binding affinity of KLWVIPQ-displaying phage to a4(Loop2) was determined with serial dilutions (~2.5-fold dilutions) of phage virions from 10 x $10^{11}$ phage/well. Specific binding of a4(Loop2) to the peptide KLWVIPQ is shown.

*Identification of Candidate Ligands*

The database searches using the same program as described in Section 4.2.1.1 revealed at least four proteins that contain either almost the entire KLWVIPQ or very similar peptide sequences (Table 16). Of the identified proteins, the ATP-binding cassette (ABC) protein C11 (ABC C11), also called MRP8, has the highest homology to the peptide identified. This protein is a newly identified member of ABC transporter superfamily with an integral membrane location and relatively high-level expression in kidney tissues (Bera et al., 2001; Tammur et al., 2001; Yabuuchi et al., 2001). Very recently, ABC C11 has been characterised as an amphipathic anion transporter that is able to efflux cAMP and cGMP and to function as a resistance factor for commonly employed purine and pyrimidine nucleotide analogues (Guo et al., 2003). However, the physiological relevance of this protein to the functions of proton pump is not apparent. The functions of other proteins identified have not ever been reported.

| Protein Name* | Accession No. | Known Functions | Interacting Regions** |
|---|---|---|---|
| Hypothetical protein | NP_116256 | Protein of unknown function | 65 **LWVFPQ** 70 |
| Unnamed protein | BAC87557 | Protein of unknown function | 60 **KLWV-PQ** 65 |
| ATP-binding cassette protein C11 (MRP8) | NP_149163 | Amphipathic anion transport that is able to efflux cAMP and cGMP and to function as a resistance factor for commonly employed purine and pyrimidine nucleotide analogues | 1217 **KLSVIPQ** 1223 |
| Transducin β-like 3 | NP_006444 | Protein of unclear function | 504 **KLWALPQ** 510 |

**Table 16 Database search for proteins containing a KLWVIPQ-like sequence.**

Proteins were identified using programme BLAST ('short nearly exact matches' option). *Protein name was obtained from the Human Protein Database (web site). **Directly interacting regions within the ligand are numbered at the side of the sequence, and the interacting amino acids in bold.

### 4.2.2 Yeast Two-hybrid Assays

To search for protein binding partners of the a4(Loop2) [or a4(N)], Y2H assays were performed using the CytoTrap two-hybrid system, with a4(Loop2) [or a4(N)] as the bait, selecting from a human testis cDNA library (Stratagene).

#### 4.2.2.1 Construction of Bait Plasmids

The cDNA sequences encoding a4(Loop2) [or a4(N)] were amplified by high fidelity PCR [or LR-PCR] using primer 9 [or 3] (forward) and primer 10 [or 4] (reverse) (Table A.3). The resulting 235 bp [or 1.194 kb] *SalI-NotI* PCR fragment was cloned into *SalI* and *NotI* sites of pSos vector. The subsequent identification of positive clones and verification of inserts were carried out as described in Section 3.2.1.1.

#### 4.2.2.2 Assessment of Bait Plasmid

Prior to proceeding with library screens, both bait plasmids, pSos-a4(Loop2) and pSos-a4(N) constructed as above were subjected to the recommended bait verification test as described in Section 2.19.2.4. Each bait construct was co-transformed with empty pMyr vector into cdc25H and assessed for growth at 37°C on both SD/Gal (-UL) and SD/Glc (-UL) plates. Yeast colonies were observed at 25°C, but not at 37°C after 5-day incubation with both constructs (Figure 26), indicating that the Sos-a4(Loop2) or Sos-a4(N) fusion does not interact with the myristylation signal in the absence of an interaction partner. This result also indicates that neither a4(loop2) nor a4(N) contain sequences that target the Sos-bait fusions to the cell membrane. In other words, the two constructs did not show any auto-activation and therefore are both suitable for use with the Y2H system.

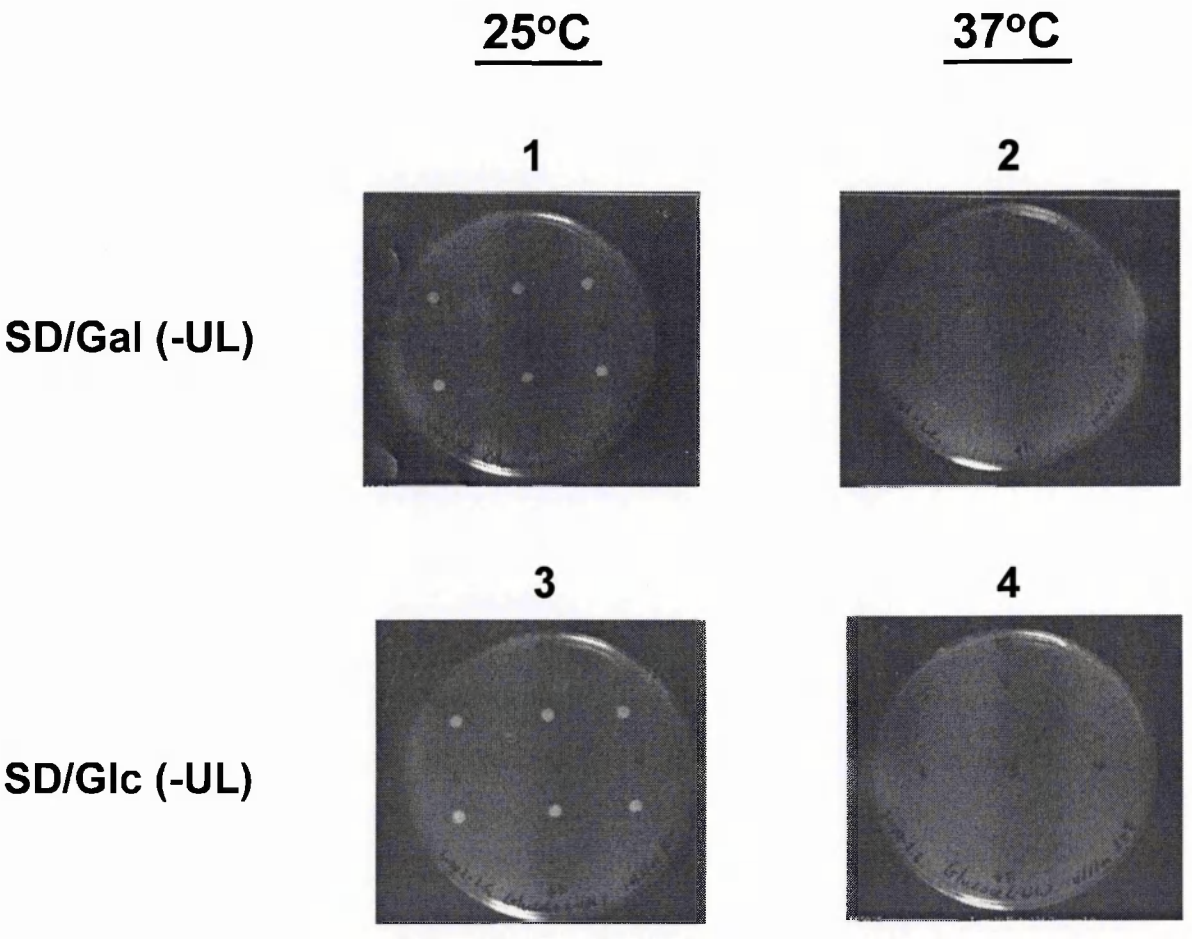



**Figure 26 Assessment of the bait plasmids.**

The pSos-a4(Loop2) bait plasmid was co-transformed with empty pMyr vector into cdc25H. Growth was observed at 25°C (*Panel 1 and 3)*, but not at 37°C (*Panel 2 and 4*) on both SD/Gal (-UL) and SD/Glc (-UL) plates after 5-day incubation. Data for construct tests of pSos-a4(N) were of similar quality, but not shown.

### 4.2.2.3 Yeast Two-hybrid Screen with a4(Loop2)

*Library Screening*

Library screens and substantial patching tests were carried out as described in Section 2.19.2.5. pSos-a4(Loop2) bait plasmid was co-transformed with human testis cDNA library into the cdc25H (strain α) competent cells. Empty pSos vector replaced the bait plasmid as a negative control to estimate the number of false positive colonies from the library and of temperature sensitive revertants. After screening approximately 1.5 x $10^4$ yeast clones, 138 *interactor candidate* colonies were obtained on SD/Gal(-UL) plates incubated at 37°C. Altogether 8 colonies appeared on the control plate. Patching tests revealed that all of the 8 colonies from the control plate were temperature sensitive revertants of cdc25H strain, as they all showed growth phenotype on SD/Glc(-UL) plate incubated at 37°C. This indicates that no false positive colonies from the library were obtained. To exclude temperature sensitive revertants from the 138 *interactor candidates*, the same patching test was performed as with the control sample. From this, 4 colonies grew on SD/Gal(-UL) plates, rather than SD/Glc(-UL), at 37°C, suggesting that interactions occurred between the bait and target proteins. These identified colonies were known as *putative positives.*

*Verification of Putative Positive*

The mating test described in Section 2.19.2.6 was used to verify the identified *putative positive* clones. Prior to the assay, the 4 *putative positives* were cured to remove pSos-a4(Loop2) bait plasmid. Following curing, the *putative positives* (in strain α) containing only the potential target constructs were co-patched with pSos-a4(Loop2) which was previously transformed into strain a of cdc25H cells. pSos Col I replaced pSos-a4(Loop2) as a negative control. Of the 4 *putative positives,* 1 grew, at 37°C, on SD/Gal(-UL) plates mated with bait plasmid

containing cells, but not with the control, demonstrating that an interaction had occurred between a4(Loop2) and the target protein. This result also indicates that the interaction between the bait and target is specific in the Y2H system, as no growth was observed from the controls (Figure 27). This clone was named *positive*. However, the other 3 *putative positives* showed mating with both bait containing cells and control, suggesting that non-specific interaction had occurred.

This *positive* clone was subsequently subjected to the target plasmid isolation, and the plasmid isolated was then used as template in a PCR reaction using primers 22 (forward) and 23 (reverse) (Table A.3). The resulting PCR fragment (~2.5 kb) was submitted to DNA sequencing using primer 22 (Table A.3). Database searches (http://www.ncbi.nlm.nih.gov/BLAST/) revealed that the cDNA clone had 60% identity to the coding sequence for a human hypothetical protein (NP_078891). However, no known functions of this protein have been reported.

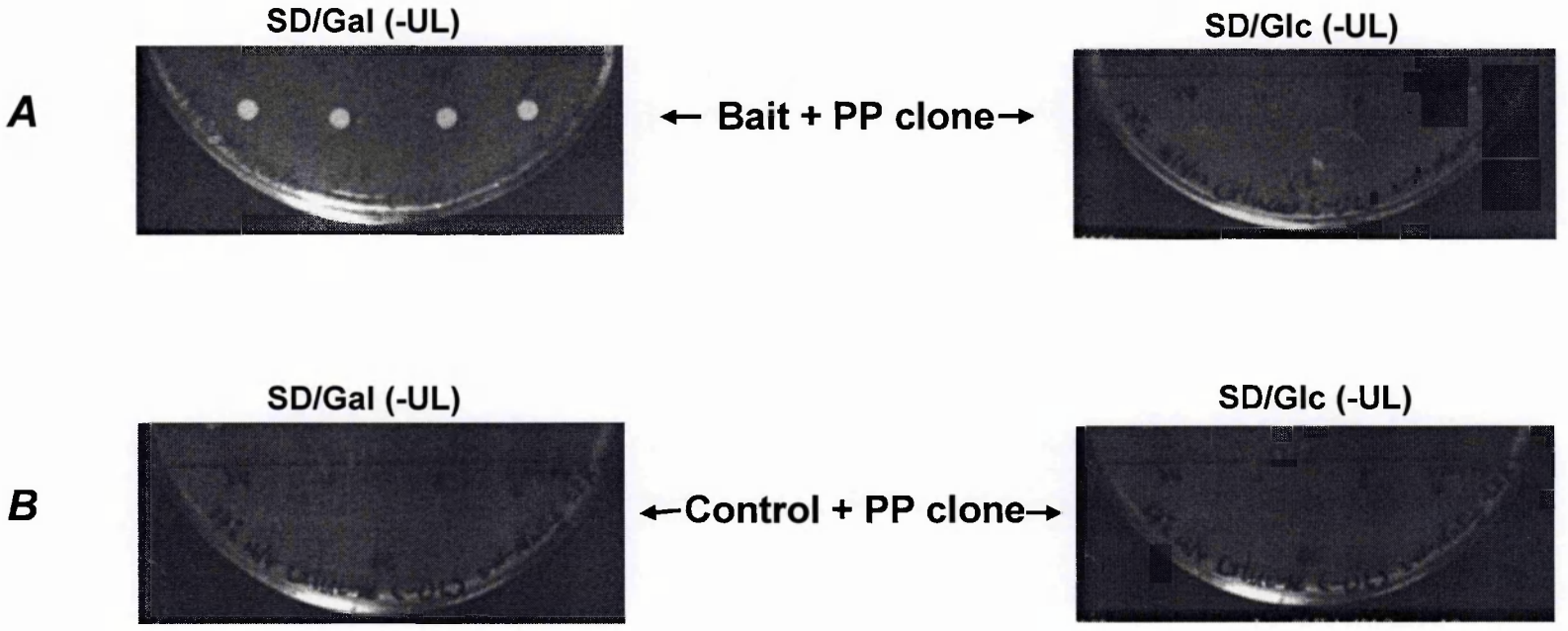



**Figure 27 Mating tests between the target protein and a4(Loop2).**

The cdc25 cells (strain α) containing the *putative positive* (PP) was patched (four patches were made from the same clone) with cdc25 cells (strain a) containing the bait plasmid followed by incubation at 37°C. Growth was observed on SD/Gal (-UL), but not on SD/Glc (-UL) plates (*Panel* ***A***). pSos Col I replaced pSos-a4(Loop2) as a control. No colonies were observed on either of the plates (*Panel* ***B***).

#### 4.2.2.4 Yeast Two-Hybrid Screen with a4(N)

*Library Screening*

The pSos-a4(N) bait plasmid was co-transformed with human testis cDNA library into the cdc25H (strain α) and empty pSos vector replaced the bait plasmid as a negative control, as described in Section 4.2.2.3. After screening approximately 1.75 x $10^4$ yeast clones, 205 *interactor candidate* colonies were obtained. Altogether 14 colonies appeared on the control plate, which were all identified as temperature sensitive revertants after the patching tests. The same patching tests were performed to all *interactor candidate* colonies selected, and this resulted in 10 *putative positives.*

*Verification of Putative Positives*

The 10 *putative positives* were first cured and then subjected to mating tests. Of these 10 *putative positives,* 5 clones grew, at 37°C, on SD/Gal(-UL) plates mated with bait plasmid containing cells, but not with the control, demonstrating interaction had occurred between a4(N) and these five target proteins within the colonies identified. This result also indicates that the interactions between the bait and targets are specific, as no growth was observed from the controls. Data for the mating tests were of similar quality as shown in Figure 27 (but are not actually shown). However, the other five *putative positives* showed mating with both bait containing cells and control, suggesting that non specific interactions had occurred.

Four of the five *positive clones* were successfully sequenced and database searches revealed that the four cDNA clones had almost complete homology to the coding sequences of three human proteins. These proteins are shown in Table 17, together with the region of the protein responsible for the interaction and their putative assigned functions. Two cDNA clones had approximately 90% identity to

the coding sequence for an unknown protein (AAH19030), the third clone had 90% identity to a hypothetical protein (NP_116256) and the fourth clone had 100% identity to the coding sequence for human proteasome beta 1 subunit (NP_002784). The proteasomes are protein degradation machines, which play a number of important roles in cell life. One of their major roles is to remove abnormal and misfolded proteins from the cell (Coux et al., 1996). The proteasome is a large complex which is present in the nucleus, cytosol, ER and lysosomes of all eukaryotic cells examined (Rivett et al., 1992; Scherrer and Bey, 1994).

| Protein Name* | Acoession No. | Identities | Known Functions | Interacting Regions** |
|---|---|---|---|---|
| Unknown protein | AAH19030 | 90% | Protein of unknown function | Around aa 457- 800 |
| Hypothetical protein | NP_116256 | 90% | Protein of unknown function | Around aa 1-945 |
| Proteasome beta 1 subunit | NP_002784 | 100% | Essential component of the ATP-dependent proteolytic pathway in eukaryotic cells and is responsible for the degradation of most cellular proteins. | Around aa 1- 280 |

**Table 17 Database search for proteins containing sequenoes homology to the target proteins.**

BLAST was used to search for proteins interacting with a4(N) using CytoTrap two-hybrid system. *Protein name was obtained from the Human Protein Database (web site). **Interacting regions were estimated according to sizes of PCR products used for sequencing.

## 4.3 Discussion

The purpose of this study was to identify ligands binding to a4(N), a4(Loop2) and a4(C). Two library-based methods, phage display and the Y2H system, were used for this investigation. Successful expression of a4(Loop2) and a4(C) enabled them to be used as target proteins to screen random peptide phage display libraries. However, the identification of ligands for a4(N) could only be attempted with the Y2H system due to unavailability of folded a4(N) protein.

Phage display analysis using a4(C) as a target to screen the 7-mer random peptide library selected a single specific displayed peptide in approximately 40% of cases. The interaction of a4(C) with this peptide, SWLELRP, was confirmed by a phage ELISA. Database searching revealed that this residue pattern is almost completely homologous to part of the enzyme PFK-1. This matched region is at the C-terminus of PFK-1, and this region within PFK-1 is highly conserved among mammalian PFK-1 orthologues and paralogues. Activity of this enzyme is the rate-limiting step in glycolysis. Linkage of the glycolytic enzyme aldolase to the proton pump through protein-protein interaction has been previously reported (Lu et al., 2001). This resulted in PFK-1 being in a more advantageous position to be a potential binding partner compared to the other proteins identified. It was therefore decided to examine this interaction in greater detail as described in the next chapter.

The screening of a random 7-mer peptide phage display library with immobilised a4(Loop2) protein selected a single specific displayed peptide, KLWVIPQ, which overwhelmed others. Alignment of this sequence with other bound phage peptide sequences revealed a sequence motif K(L/V)WVIPQ in approximately 75% of cases. Although this residue pattern had not clearly emerged through panning with random 12-mer peptide phage display library,

some K(L/V)WVIPQ-like patterns were seen from this screen. The interaction of a4(Loop2) with the peptide, KLWVIPQ, was confirmed by a phage ELISA. However, database searches did not reveal any proteins which have apparent physiological relevance to V-ATPases. ABC C11 (MRP8) seems, however, to be interesting for two reasons. Firstly, it is an integral membrane protein and is expressed at relatively high levels in human kidney (Yabuuchi et al., 2001), although absence of expression of this gene in the kidney was also reported (Tammur et al., 2001). Secondly, some multidrug resistance (MDR) cell lines have been found that overexpress subunit C of the V-ATPase while they overexpress MRPs (Ma and Center, 1992). Nevertheless, evidence that directly links the activities of V-ATPase and MRPs is scarce. Therefore it would have been hard, at this stage, to select this protein as a potential binding candidate for further characterisation. In addition, for those unknown and hypothetical proteins, it would also have been difficult to perform further investigations due to lack of any functional information and of availability of antibodies etc.

Identification of a binding sequence, but not an actual protein of interest might be due to the following disadvantages of the phage display libraries and database searches used. Firstly, in both libraries, the peptides are very short linear sequences. In reality, protein-protein interactions can be roughly divided into two basic types: those that involve the recognition of a continuous, linear stretch of amino acids and those that involve binding contacts spread over large, non-contiguous surfaces. In the latter type, formation of multiple residues involved for interaction requires a correctly folded protein, and these residues may be far apart in the primary amino acid sequence. However, the methods used for searching for binding ligands from databases rely only on the primary structure of proteins. In other words, the searches can only fish for those proteins that have interacting

sequences of a continuous and linear stretch of amino acids to the identified peptide, whereas proteins having interacting sequences formed by non-contiguous residues similar to the identified peptide would be missed. Secondly, the peptide sequences in the libraries were constructed artificially according to the possible recombination of the 20 amino acids, and it is very likely that some of them do not occur in proteins from natural sources. Also, it is possible that some artificial peptides might somehow have a greater affinity for binding to a target protein than those naturally occurring sequences in an *in vitro* assay. Both of the above cases can sometimes cause confusion during investigation. However, strong binding of a4(Loop2) to the peptide, KLWVIPQ, shown by ELISA indicated that the interactions were specific and real, at least confirmed in the phage display assays.

To gain further insight into potential partners for a4(Loop2), an alternative library-based method, Y2H assay was performed. One positive clone whose sequence matched (~60%) with a hypothetical protein in the protein databases was identified. Regarding the hypothetical protein, little information is available, so further investigations on this interaction were hard to be envisaged. The Y2H system was also used to identify proteins capable of binding to a4(N), which revealed three proteins from the database. Interestingly, the proteasome appeared as a potential binding partner in one of the positive clone identified. The proteasome is an essential component of the ATP-dependent proteolytic pathway. This protein catalyses the rapid degradation of most cellular proteins, such as those highly abnormal or misfolded proteins which may arise by mutation or postsynthetic damage (Coux et al., 1996). In addition, proteasomes were also found to be involved, together with other ER quality control components, in the degradation of proteins that appeared to have a normal structure but failed to be assembled in the ER (Hill and Cooper, 2000). Whether or not the binding of

proteasome to a4(N) seen in the Y2H analysis implies that a4(N) has not properly folded or is no longer necessary in that particular clone remains unclear. However, results obtained from a series of tests during the Y2H analysis indicated that the a4(N)/target interactions were specific and real, at least in this Y2H system. However, this needs to be further confirmed by other physical strategies in order to exclude the possibility of 'false positives'.

There are two important considerations that must be taken into account when using the Y2H system. One possibility is that the proteins may adopt a different tertiary structure when expressed as fusions with either Sos or Myr proteins, which could potentially inhibit the binding of true protein partners. Another possibility is that the efficiency of the yeast cell transformation could also affect the outcome of the assays. In both library screens, either using a4(Loop2) or a4(N) as a bait, the number of clones screened were at a magnitude of $10^4$, whereas the number of primary clones in the library used are at an magnitude of $10^6$ . Therefore, even though each of the $10^4$ clones screened contained different target cDNA plasmid, there were still approximately $10^2$ target cDNAs missing from the screen. Both of these problems could possibly give rise to a 'false negative' result.

In conclusion, through the systematic technique of phage display, I identified a potential a4(C) interacting partner, PFK-1. The possible interaction between these two proteins suggests a direct link between proton pump activity and the glycolytic pathway. Therefore, further investigations, in greater detail of this interaction, were carried out as described in the next chapter. In addition, I was interested to see whether or not the ubiquitously expressed human a1 subunit of V-ATPase also interacts with PFK-1. If this is the case, coupling between the V-

ATPase and glycolysis through the PFK-1/a subunit interaction could be a common mechanism.

# CHAPTER 5
# CHARACTERISATION OF INTERACTION BETWEEN a4(C) AND ITS POTENTIAL BINDING PARTNER PFK-1

## 5.1 Introduction

As mentioned earlier, all studies to date have demonstrated that the a subunit of the proton pump plays crucial roles in the coupling of ATP hydrolysis ($V_1$) and proton transport ($V_0$). Since its soluble C-terminal tail has already been shown to play potential roles in the assembly, stabilising or targeting of the a subunit as well as proton translocation (Leng et al., 1998), I chose to further investigate the possible binding partners for this part of the molecule to further explore its possible functions. As described in Chapter 4, to identify proteins that interact with a4(C), a random 7-mer peptide phage display library was used to select epitopes with high binding affinity. Through this library-based strategy, one a4(C)-binding peptide sequence, SWLELRP, was identified that showed high homology to the glycolytic enzyme PFK-1. This pattern is highly conserved among mammalian PFK-1 orthologues and paralogues. Activity of this enzyme is the rate-limiting step in glycolysis. The specificity of binding of a4(C) to this peptide was confirmed by phage ELISA. This finding suggests that PFK-1 is a potential binding partner of the a4 C-terminal soluble tail.

However, 'false-positives' can occur in any library-based screen. Thus, it is imperative that any interactions that have been identified in a screen are confirmed by alternative assays. A common strategy is to perform an *in vitro* pull-down assay using recombinant protein(s), which can provide evidence of physical association between the tested proteins. However, this type of assay is usually not enough to extend the systematic assay, since the recombinant proteins used are usually at relatively high concentration, which may not accurately reflect the situations that

occur within cells. In addition to *in vitro* pull-down analysis, co-localisation of the putative interacting proteins in cells in culture or tissue samples by immunofluorescence microscopy would be supportive of a 'genuine' interaction. Nevertheless, the optimal method for confirming the validity of a physiological *in vivo* interaction between two proteins is the co-immunoprecipitation of the interacting endogenous proteins from a tissue sample.

This chapter presents several lines of evidence to demonstrate the a4/PFK-1 protein-protein interaction, via the C-terminus of a4. Assays, including *in vitro* pull-down analysis, co-immunoprecipitation and immunofluorescence microscopy were performed and the results obtained indicate binding of PFK-1 to a4. The association between these two proteins indicates a direct link between the V-ATPase and the ATP-generating glycolytic pathway in the kidney, suggesting the possibility of a regulation mechanism between energy supply and V-ATPase function.

Having identified the specific interaction between a4(C) and PFK-1, I wanted to investigate whether the interaction was disrupted or influenced by the mutations in a4(C) identified from patients. In an attempt to address this question, additional *in vitro* pull-down assays with recombinant a4(C) mutants were carried out. Also, studies using the Surface Plasma Resonance (SPR) technique were subsequently performed in order to compare the binding affinities of a4(C)/PFK-1 and a4(C) mutants/PFK-1 interactions.

SPR provides a powerful tool for analysis of protein-protein interactions, which allows macromolecular interactions to be measured in real time. BIAcore is, so far, the most widely used SPR-based system (Jonsson et al., 1991). In recent years, use of BIAcore biosensor technology has increased steadily in the investigation of protein-protein interactions. The use of BIAcore biosensors

provides excellent instrumentation for a label-free, real-time investigation of protein-protein interactions. This system offers particular advantages for analysing weak macromolecular interactions, allowing measurements that are not possible using some other techniques (van der Merwe and Barclay, 1994, 1996). Another major advantage of this system over other techniques, such as stopped-flow, analytic centrifugation and isothermal titration calorimetry (ITC), is that much smaller amounts of protein samples are required. The principle underlying this technology relies on the SPR phenomenon, which transforms the specific incident angle of the light reflected from a metallic surface in response to the substance bound to the surface. In other words, signals are gained from the changes of refractive index due to the increased mass obtained from a bound protein. Based on these data a broad range of applications, including the comparison of a specific protein-protein interaction between wild-type and mutants, can be achieved.

## 5.2 Results

### 5.2.1 PFK-1 and a4 Immunoreactivities in Human Kidney

Prior to examining the interaction between PFK-1 and the a4 subunit in human kidney, a commercially available goat polyclonal antibody directed against rabbit muscle-type PFK-1 (Chemicon International Ltd) was tested for immunoreactivity against human kidney cytosolic and membrane protein fractions prepared as described in Section 2.16.1. 10 μg of each sample was subjected to 12% SDS-PAGE and Western blotted using α-PFK-1 antibody at 1:2000 dilution. As displayed in Figure 28 (*panel A*), this antibody recognised a band of the correct size for human PFK-1 (approximately 85 kDa) in both fractions, which as expected was more prominent in the cytosolic portion. The RA2922 antiserum directed against the last 14 amino acid residues of human a4 was also tested for

immunoreactivity against human kidney cytosolic and membrane portions in the same way and the results obtained are shown in Figure 28 (*panel B*). As previously documented (Smith et al., 2000), the RA2922 antibody recognised a band of the correct size for human a4 (approximately 116 kDa) in membrane protein fraction, but not in cytosolic protein fraction. What the lower band, (approximately 55 kDa) seen in both fractions, represents is not clear.

### 5.2.2 Co-immunoprecipitation

To examine whether a4 and PFK-1 interact *in vivo*, a co-immunoprecipitation assay was carried out using the α-PFK-1 antibody for precipitation and RA2922 antibody for detection. The experimental procedures were carried out as described in Section 2.21.3. The α-PFK-1 antibody was covalently coupled on agarose beads followed by incubation with solubilised human kidney membrane protein fraction. After extensive washes, bound proteins were eluted and subjected to 12% SDS-PAGE electrophoresis and Western blotted using both RA2922 antibody (1 in 1000 dilution) and α-PFK-1 antibody (1:1000 dilution). The resulting blot, showing a single major band at approximately 116 kDa recognised by the α-a4 antiserum (Figure 29*A, +lane*), conforms to previous analysis (Smith et al., 2000) and the activity assay shown in Figure 28*B* using this antibody. An identical blot probed with the α-PFK-1 antibody revealed a band of the correct size, confirming the presence of the enzyme (Figure 29*B, +lane).* This demonstration of the co-precipitation of these two proteins indicates their interaction. Interestingly, a4 was present only in the protein sample prepared with n-NDG as the detergent in both solubilisation and washing buffers (*panel C, lane 2*), and not when CHAPS was used [*panel C, lanes 1*(CHAPS only) and *3* (CHAPS and n-NDG)]. Specificity of the assay was confirmed by absence of a4 when the precipitating antibody was omitted (*-lanes*).

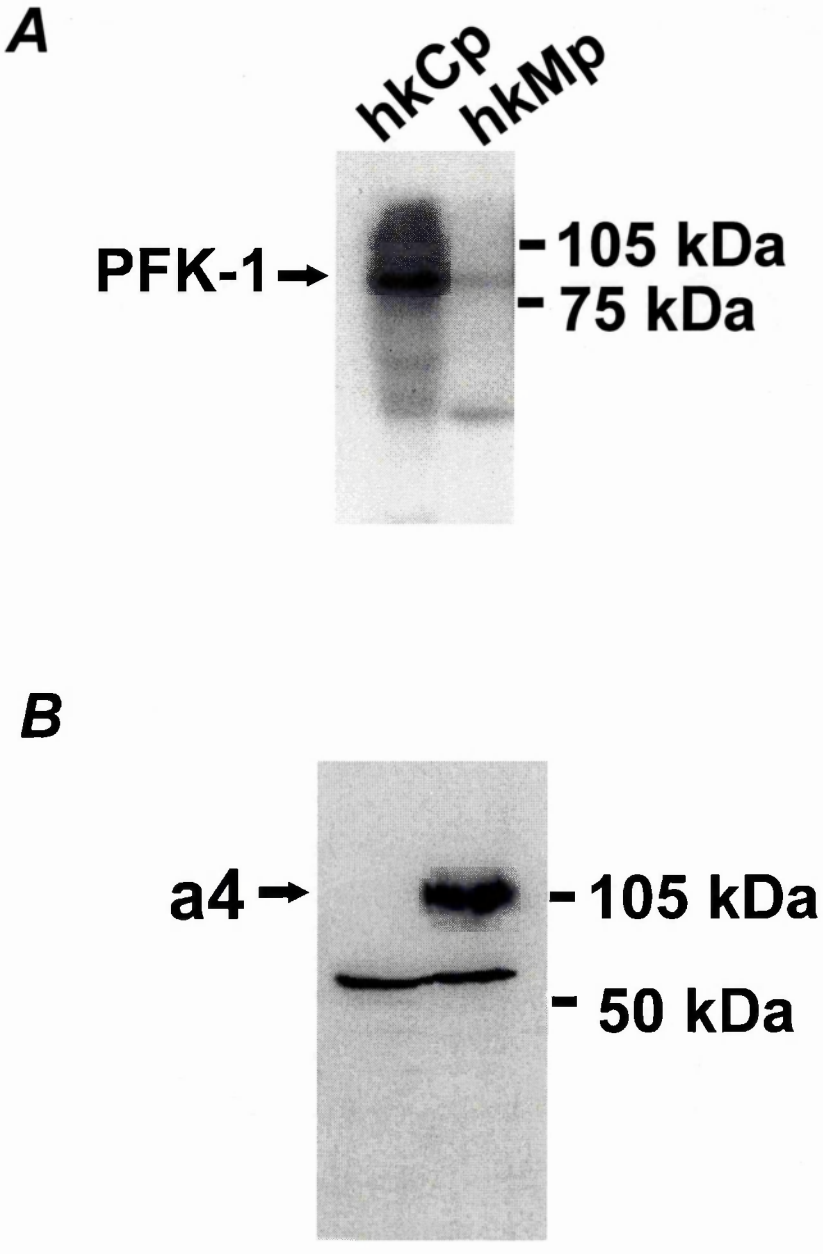



**Figure 28 Immunoreactivity assays**

10 μg of either human kidney cytosolic protein fraction (hkCp) or membrane protein fraction (hkMp) were subjected to SDS-PAGE. Western blot analysis using α-PFK-1 antibody demonstrated an 85 kDa band corresponding to the PFK-1 protein in both fractions (*panel **A***). Western blot analysis using RA2922 antibody demonstrated a 116 kDa band corresponding to the a4 protein in membrane protein fractions, but not in cytosolic protein fraction (*panel **B***).

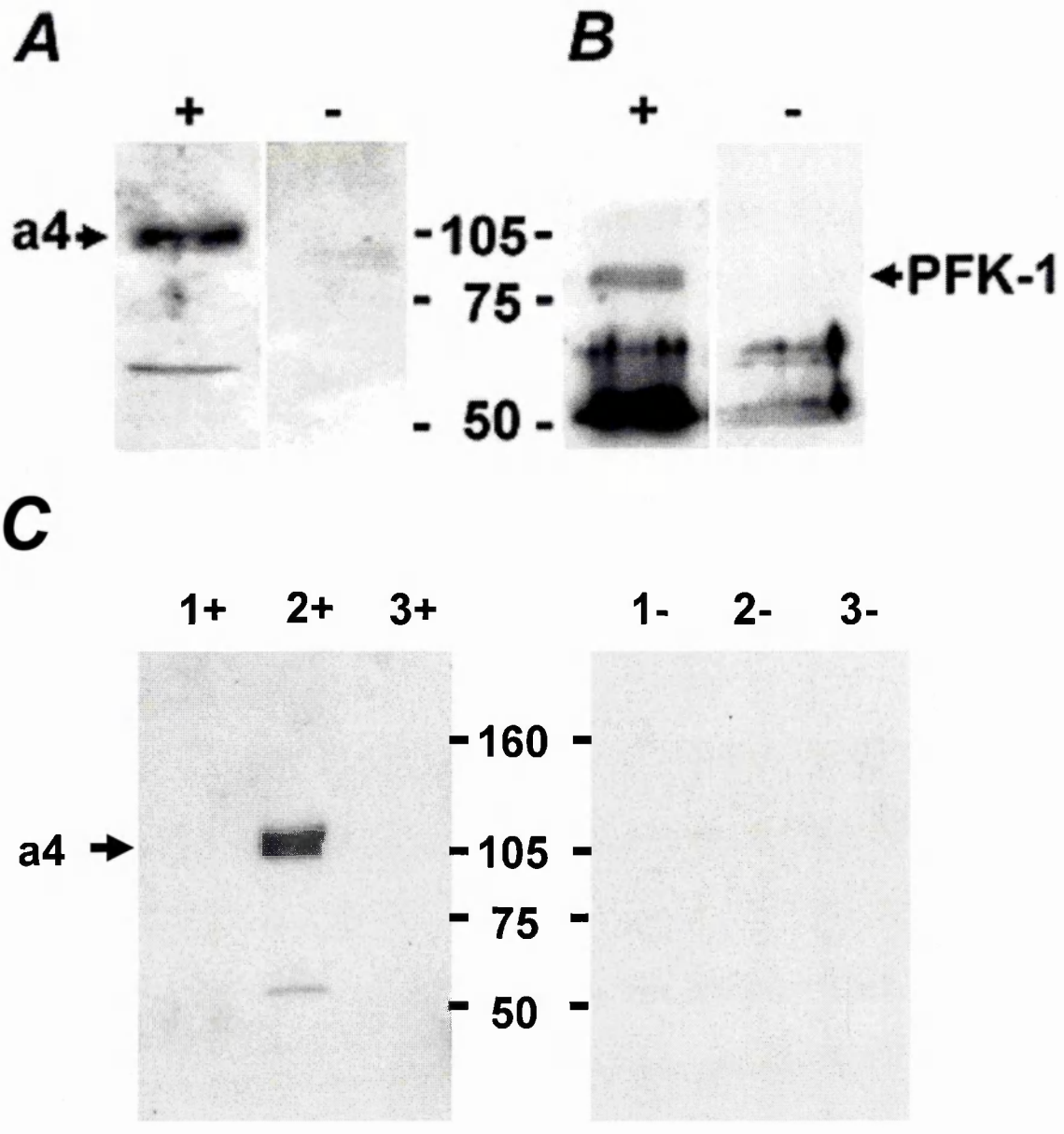



**Figure 29 Co-immunoprecipitation.**

The membrane protein fraction of fresh frozen human kidney was solubilised and immunoprecipitated using rabbit muscle type PFK-1 antibody (+ *lanes*). ***A***, Detection of a4 (approximate 116 kDa, *arrowed*) was performed using RA2922 [anti-a4(C)] antiserum, and indicates co-precipitation of these two proteins. ***B***, Probing with the precipitating antibody confirmed the presence of PFK-1 (approximately 85 kDa, *arrowed*) in the precipitated complex. ***C***, Effects of detergents. - indicates negative control where α-PFK-1 antibody was omitted.

### 5.2.3 Affinity Pull-down Assays of the Protein-protein Interaction

I was interested to see whether the interaction between a4(C) and PFK-1 was also true of the C-terminus of the ubiquitously expressed a1 subunit [a1(C)]. If this was the case, coupling between the V-ATPase and glycolysis through the PFK-1/a subunit interaction could be a common mechanism. To address this question, I wished to perform similar immunoprecipitation studies, but found that the available a1 antibody (kind gift from Dr M. Futai, Osak University, Japan) appeared to cross-react with a4 (Figure 30). I therefore designed a PFK-1 pull-down assay instead. HPLC-purified a4(C) or synthesised a1(C) were first labelled with biotin (Section 2.20), and following confirmation of successful biotinylation (Figure 31, *panel A*), were then incubated with agarose beads to which rabbit muscle type PFK-1 was bound (Sigma). After extensive washing with PBST, bound proteins were eluted and spotted on nitro-cellulose membrane and analysed using avidin α-biotin conjugate. Binding of PFK-1 to both a4(C) and a1(C) was evident from this assay (*panels C* and *E*). As can be seen from *panels D* and *F*, this binding was specific, since when an unrelated protein (protein A) conjugated to identical agarose beads was used, no significant binding of a4(C) or a1(C) was observed. Finally, *panel G* confirmed there was almost no binding between PFK-1 and the avidin-HRP conjugate alone.

### 5.2.4 Immunolocalisation of a4 and PFK-1 in Human Kidney

The distributions of a4 and PFK-1 in human kidney were compared by double label immunohistochemistry, as described in Section 2.23, using the antibodies to human a4 (RA2922) (1 in 5000 dilution) and to rabbit muscle type PFK-1 (1 in 100 dilution) employed earlier. Following the primary antibody reactions, tetramethyl rhodamine isothiocyanate (TRITC)-labelled anti-rabbit or

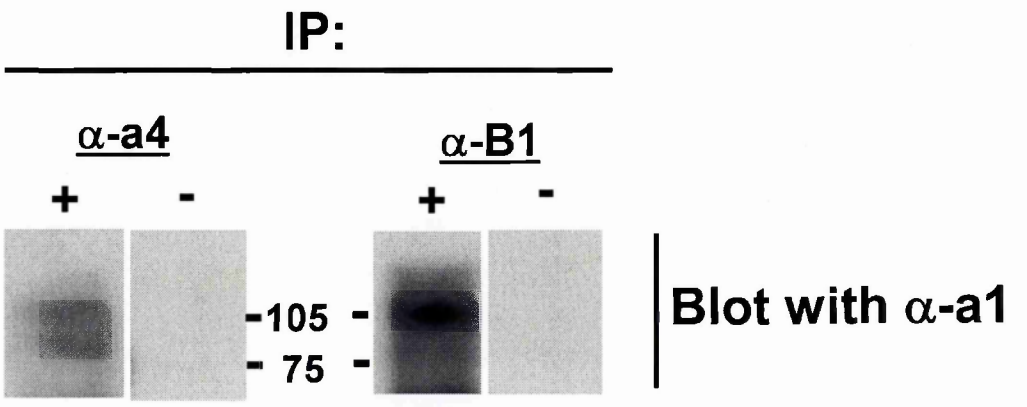



**Figure 30 Immunoprecipitation**

The membrane protein fraction of fresh frozen human kidney was solubilised and immunoprecipitated using RA2922 [anti-a4(C)] antiserum or anti-human B1 antibody (+ *lanes*). Western blots were probed with anti-mouse a1 antibody (1 in 500 dilution). - indicates negative control where the precipitating antibodies were omitted.

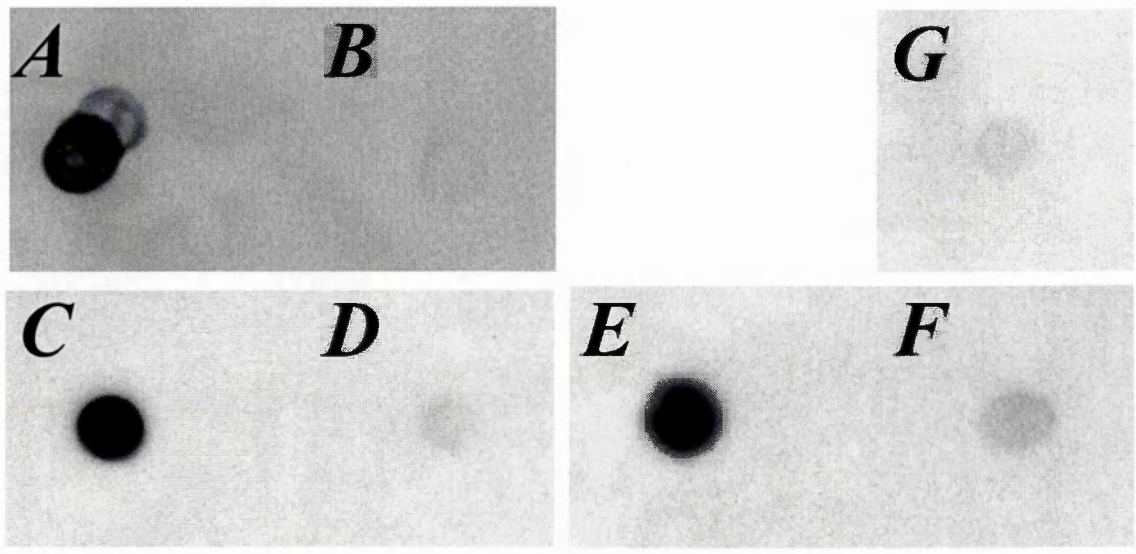



**Figure 31 PFK-1 pull-down analysis.**

Initially, purified a4(C) peptide (***A***) or a sample where a4(C) was omitted (as negative control, ***B***) were biotinylated and detected with HRP-conjugated avidin, confirming successful biotinylation and quenching. Subsequently, agarose bead-bound PFK-1 was able to pull down both biotinylated a4(C) (***C***) and a1(C) (***E***), indicating binding. Specificity of this interaction was confirmed by replacing PFK-1 beads with protein A beads (***D*** and ***F***) or applying avidin to PFK-1 alone (***G***).

fluorescein isothiocyanate (FITC)-labelled anti-goat secondary antibodies (Vector Laboratories) were then applied at 1:100 dilution. As previously documented (Smith et al., 2000), a4 localised to the apical surface of α-IC in the collecting duct (Figure 32, *the left panel*). In contrast, it was found that PFK-1 was extensively distributed in all nephron segments, as predicted from its ubiquitous function. However, some enrichment was observed in glomeruli and proximal tubules (not shown), and collecting duct (*middle panel*). Merged confocal optical sections (*right panel*) suggested co-localisation of a4 and PFK-1 in α-IC in the cortical collecting duct.

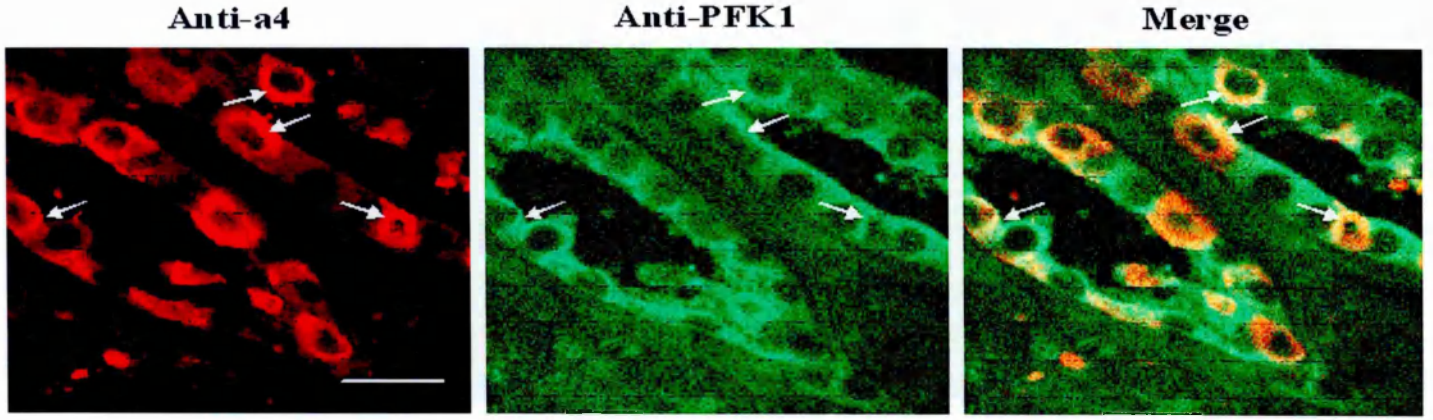



**Figure 32 Immunohistochemistry.**

5 µm cryostat-cut sections of human kidney were co-stained with goat anti-rabbit muscle PFK-1 and rabbit RA2922 antibodies. TRITC-labelled anti-rabbit or FITC-labelled anti-goat secondary antibodies were applied for a4 and PFK-1, respectively. At *left panel*, the typical appearance of a4 is shown, localised to the apical surfaces of ICs in the collecting duct, as previously observed. As expected, PFK-1 distribution was widespread throughout all nephron segments, with some enrichment apically in the collecting duct (*centre panel*). The merged image (*right panel*) demonstrates co-localisation of a4 and PFK-1 (*arrows*) in α-IC. *Scale bar* = 10 µm.

### 5.2.5 Comparison of the Interaction of PFK-1 with a4(C) and the a4(C) Mutants

As mentioned earlier in Chapter 1 (Section 1.2.5.1 and Table 2), two missense mutations (R807Q or G820R) were previously identified in this *a4(C)* region from two different patients diagnosed with rdRTA (Smith et al., 2000; Stover et al., 2002). In order to investigate whether the specific interaction between a4(C) and PFK-1 was disrupted or influenced by these a4(C) mutant proteins [a4(C)-G or a4(C)-Q], the following assays were carried out.

#### 5.2.5.1 Pull-down Assays

*PFK-1 Pull-down Assay*

PFK-1 pull-down assay, as described above (Section 5.2.3), was first performed using recombined a4(C), a4(C)-G or a4(C)-Q incubated with agarose beads bound to rabbit muscle type PFK-1 (Figure 33, *+panels*). Protein A replaced PFK-1 as a control (*-panels*). Specific binding of PFK-1 to a4(C) was again evident from this assay (*panels A*). However, it was not clear whether or not binding of a4(C)-G or a4(C)-Q to PFK-1 is specific. As demonstrated from *panels B-* and *C-*, both mutant proteins bind to either the agarose beads, or to protein A, or both.

*GST Pull-down Assay*

Prior to performing the assay, each GST fusion protein [GST-a4(C), GST-a4(C)-G, and GST-a4(C)-Q] was expressed (in 1 litre volume) and purified using glutathione sepharose beads as previously described (Section 3.2.3). Prior to the pull-down assay, small quantities (10 μl) of the purified fusion proteins (each contained a large amount of GST alone as described in Section 3.2.3) were analysed by Western blotting using RA2922 antibody (Figure 34, *panel A*) followed

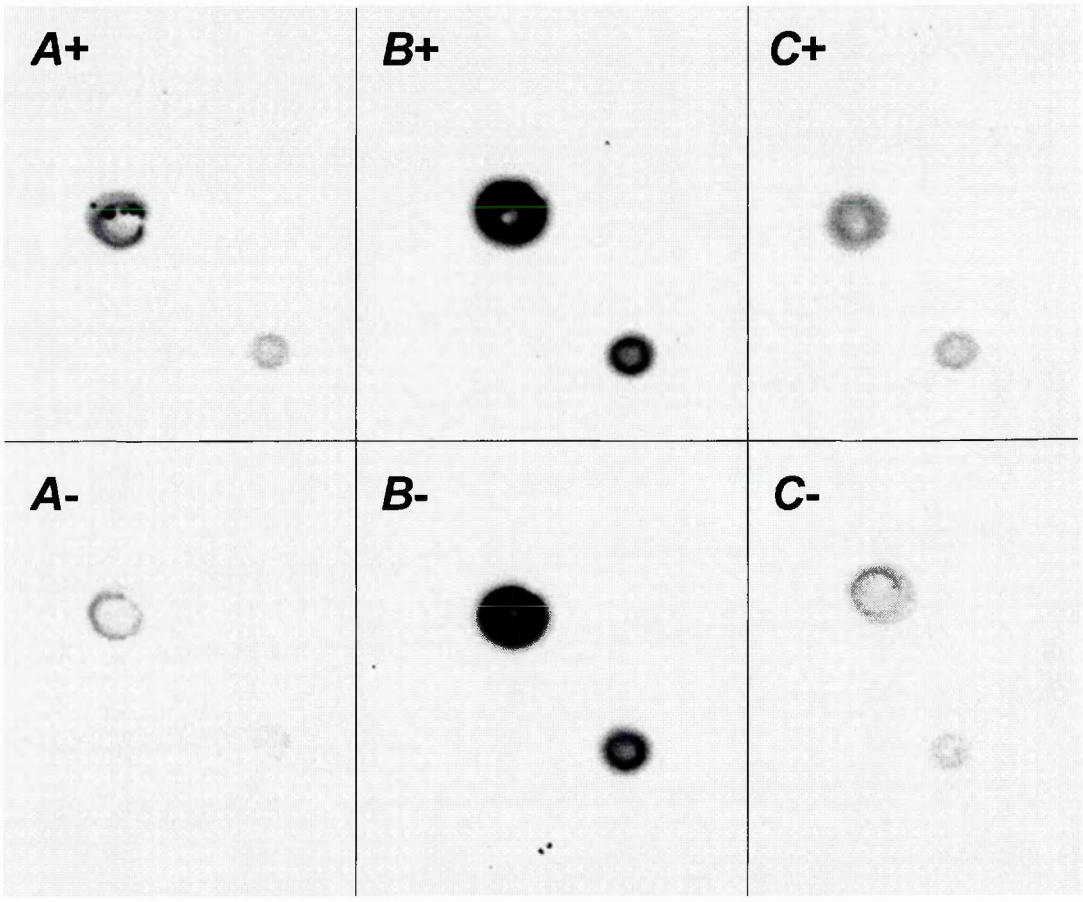



**Figure 33 PFK-1 pull-down analysis.**

Initially, recombinant a4(C), a4(C)-G and a4(C)-Q peptides were successfully biotinylated. Subsequently, biotinylated a4(C) was able to be pulled down by agarose bead-bound PFK-1 (*panel A*+), but not by protein A beads (*panel A*-) as observed previously. However, both biotinylated mutants showed positive to both incubating with the bead-bound PFK-1 (*panel **B** and **C**,* +), and bead-bound protein A (*panel **B*** and ***C***, -), indicating existence of nonspecific binding of the mutants, to either protein A or to the agarose beads, or to both.

by densitometry analysis (Figure 34, *panel B*) to quantify the level of fusion proteins. The percentage densities obtained from the assay for GST-a4(C), GST-a4(C)-G and GST-a4(C)-Q, are 17.25, 11.8 and 11.8, respectively. From these data, volumes equivalent to a designated density (3500) of each fusion protein were calculated (Table 18) and used in the pull-down assay. Briefly, human kidney cytosolic protein fraction was precleaned with GST-bound sepharose beads followed by incubation with each of the bead-bound fusion protein. GST, which was produced in parallel, replaced the GST fusion proteins as a control. After extensive washing with PBS plus 0.15% Tween 20, bound proteins were eluted and the supernatant was subjected to 12% SDS-PAGE and Western blotted using α-PFK-1 antibody (1:1000 dilution). Binding of PFK-1 to a4(C), a4(C)-G, and a4(C)-Q was evident from this assay (Figure 34, *lanes 1-3*). In contrast, no binding of PFK-1 to GST alone was detected (*lane 4*), confirming the specificity of the interaction.

| Samples (10 μl) | Density (%) | Volume of sample |
|---|---|---|
| GST-a4(C) | 17.25 | 202 μl (= 3500 density) |
| GST-a4(C)-G | 11.8 | 296 μl (= 3500 density) |
| GST-a4(C)-Q | 11.8 | 296 μl (= 3500 density) |

**Table 18 Densitometry analysis**

### 5.2.5.2 Binding Affinity Studies

Having identified that both mutants also interact with PFK-1 from the pull-down assay, I then asked whether there are any differences, in terms of binding affinity, between the binding of wild type and mutant a4(C) to PFK-1. To address

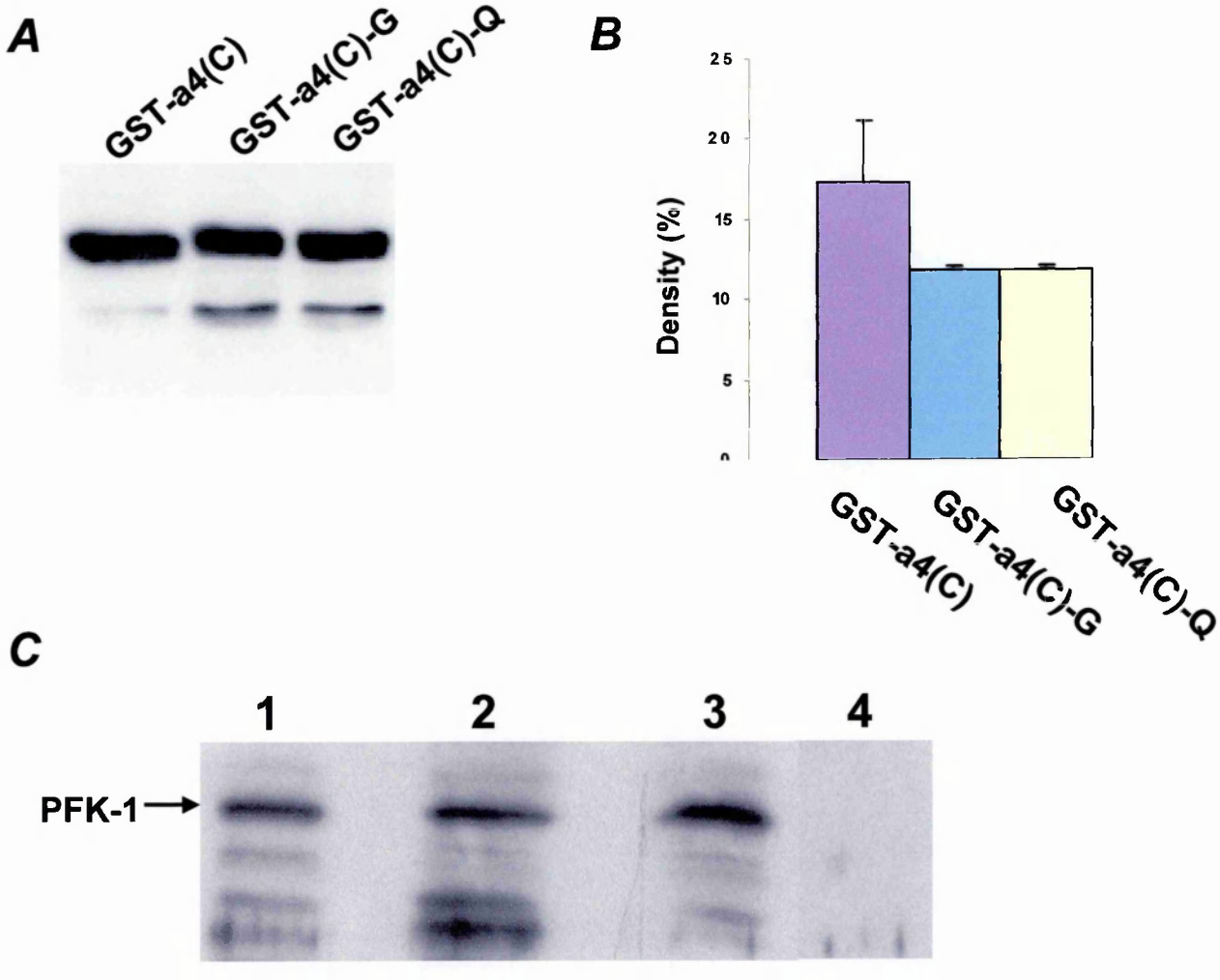



**Figure 34 GST pull-down analysis.**

GST-a4(C), GST-a4(C)-G and GST-a4(C)-Q were expressed, purified, analysed by Western blotting using RA2922 antibody (*panel **A***) and quantified by densitometry (*panel **B***) prior to be used in the pull-down assay. Subsequently, all glutathione sepharose bead-bound GST fusion proteins were able to pull down PFK-1 from kidney cytosol (*panel **C,** lanes 1-3*), indicating binding. Specificity of the interactions was confirmed by replacing the GST fusion proteins with GST (*panel C, lane 4*).

this question, SPR analysis was carried out using the BIAcore system as described in Section 2.22. In this assay, the rabbit liver-type PFK-1 protein (Sigma) (20 μg/ml in 10 mM NaAc, pH 5.0) was used as a ligand to be immobilised to a sensor chip. The synthetic peptides [a4(C), a4(C)-G and a4(C)-Q] (CovalAB) were used as analytes to be run over the surface of the chip.

BIAcore measurements are expressed in resonance units (RU) proportional to the concentration of a protein bound to the chip/or the ligand. The SPR signal generated from the immobilised PFK-1 was approximately 330 RU. Control flow cells were activated and blocked in the absence of PFK-1. Binding of the peptides to PFK-1 was evaluated over a range of peptide concentrations (7.5-250 μM) in the running buffer under a continuous flow of 10 μl/min at 25°C. Figure 35-37 displays sensorgrams of the wide type or mutant a4(C) binding to PFK-1 (*panels A*) as well as the relationships between the relative binding response (signals from the control were subtracted) and the peptide's concentration (C) (*panels B*). From these preliminary results, it can be seen that all of the three peptides showed binding with PFK-1, but the sensorgrams revealed a markedly lower response from the binding of a4(C)-Q to PFK-1 compared with the other two a4(C) peptides. In addition, the intensity of the RU signals increased with the peptide concentration within the tested range for both a4(C) and a4(C)-G. However, this behaviour has not been shown for a4(C)-Q. Furthermore, The $K_D$ values obtained using the programme Grafit for a4(C)/PFK-1 and a4(C)-G/PFK-1 are 24.4 ± 4.5 μM and 49.9 ± 9.6 μM ($p < 0.05$ obtained from the T-test), respectively. The data revealed that the affinity of PFK-1 to a4(C)-G mutant is ~2-fold lower than that of the a4(C) wild type peptide.

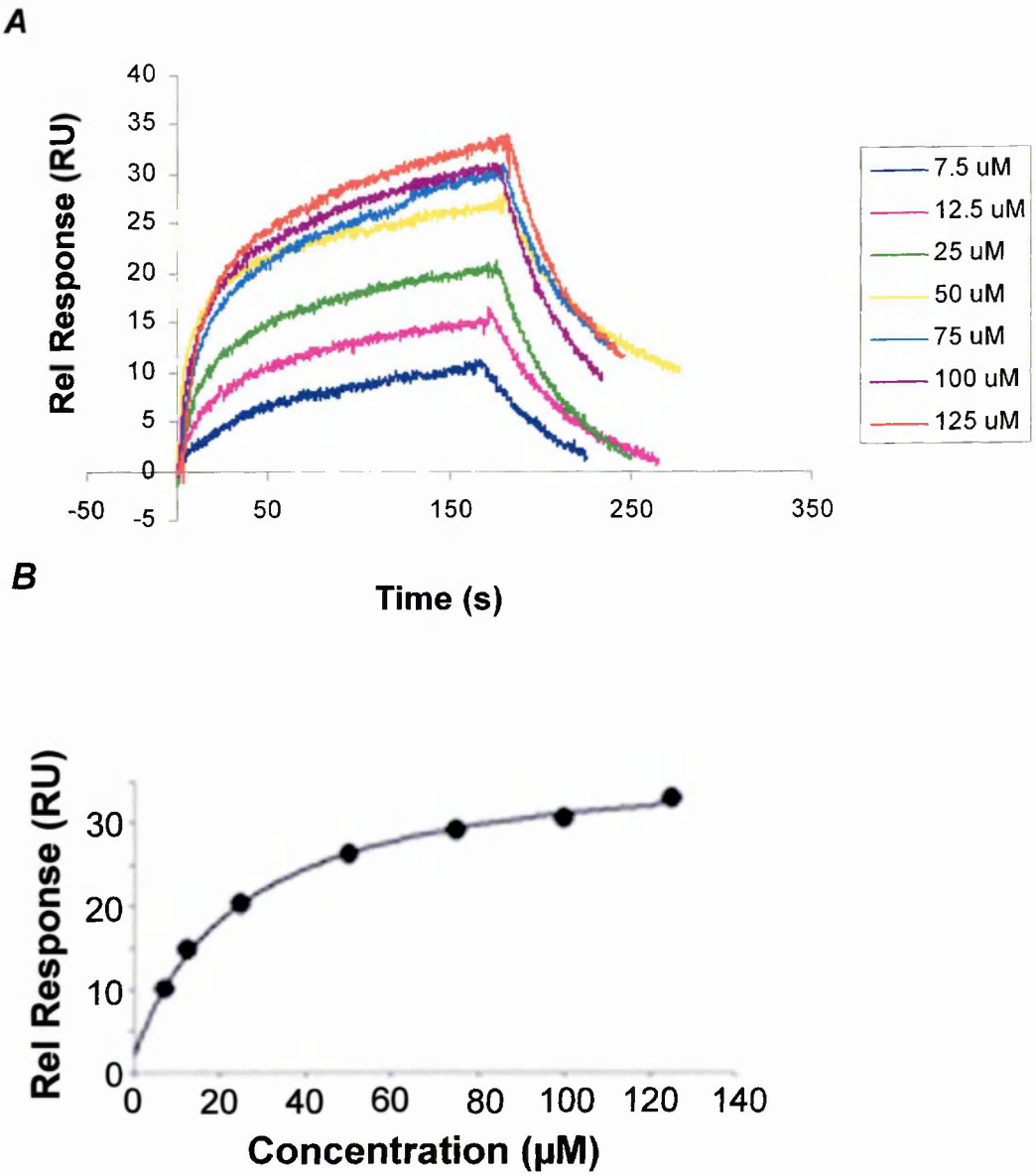



**Figure 35 Surface plasmon resonance assay.**

Binding of a4(C) to PFK-1 indicated by analysis of SPR. Synthetic a4(C) at concentrations of 7.5, 12.5, 25, 50, 75, 100 or 125 µM in PBS containing 0.005% surfactant P20 were injected (10 µl/min) over immobilised PFK-1 on the sensorchip CM-5. ***A***, Sensorgrams of binding of a4(C) and PFK-1. ***B***, Relationship between the relative response (RU) and the peptide concentration.

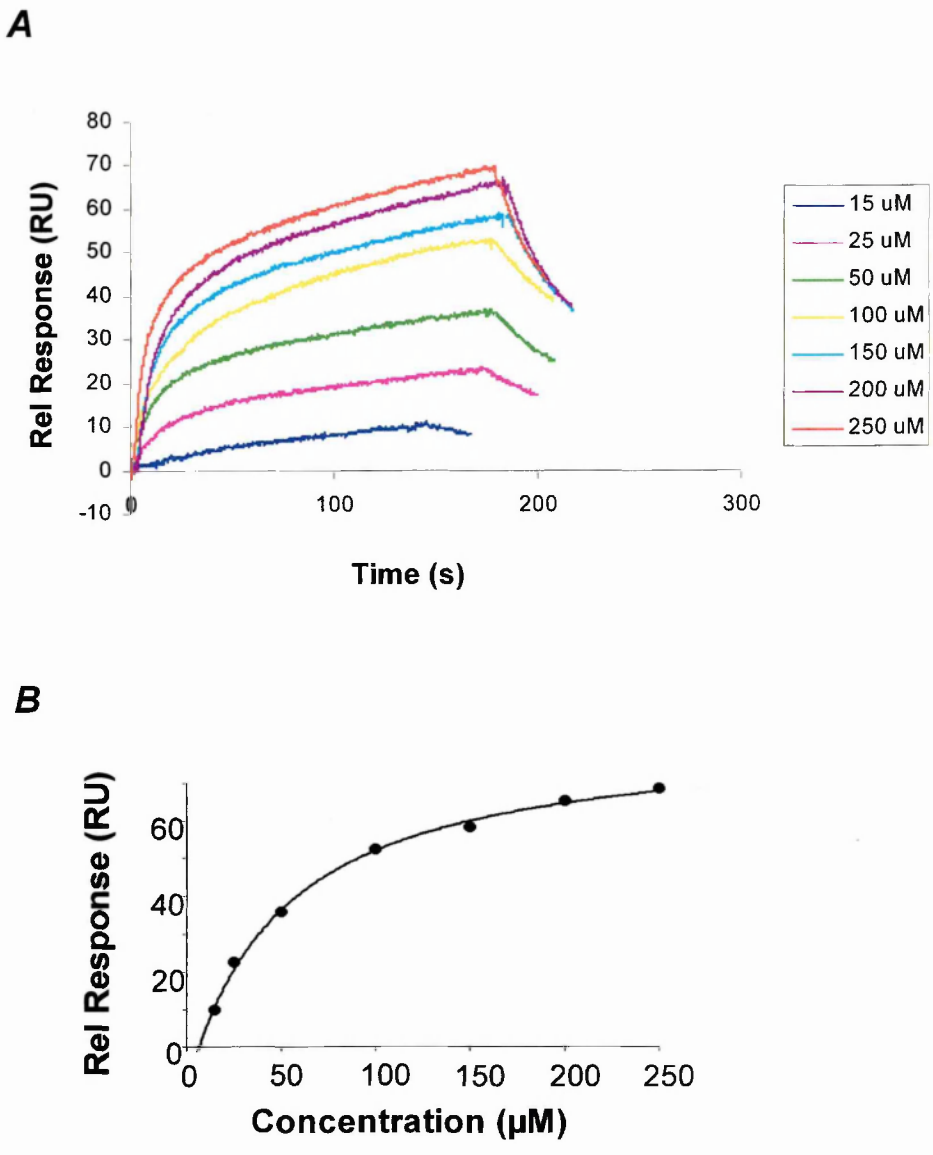



**Figure 36 Surface plasmon resonance assay.**

Binding of a4(C)-G to PFK-1 indicated by analysis of SPR. Synthetic a4(C)-G at concentrations of 15, 25, 50, 100, 150, 200 or 250 µM in PBS containing 0.005% surfactant P20 were injected (10 µl/min) over immobilised PFK-1 on the sensorchip CM-5. ***A***, Sensorgrams of binding of a4(C)-G and PFK-1. ***B***, Relationship between the relative response (RU) and the peptide concentration.

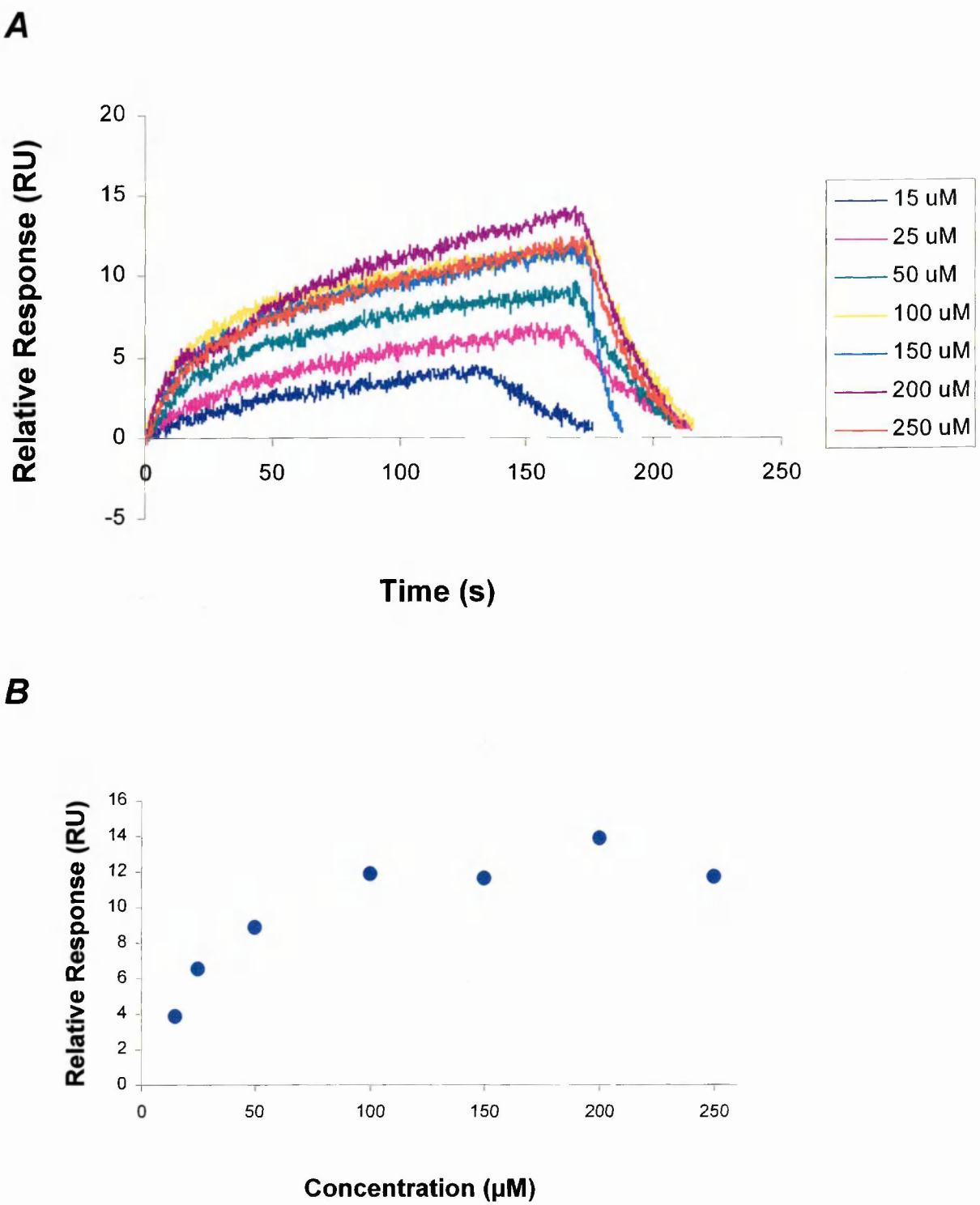



**Figure 37 Surface plasmon resonance assay.**

Binding of a4(C)-Q to PFK-1 indicated by analysis of SPR. Synthetic a4(C)-Q at concentrations of 15, 25, 50, 100, 150, 200 or 250 μM in PBS containing 0.005% surfactant P20 were injected (10 μl/min) over immobilised PFK-1 on the sensorchip CM-5. ***A***, Sensorgrams of binding of a4(C)-Q and PFK-1. ***B***, Relationship between the relative response (RU) and the peptide concentration.

## 5.3 Discussion

The a subunit of the proton pump is deemed crucial for the coupling of ATP hydrolysis and proton transport. Physical interactions between the V-ATPase subunits, or between subunits and other proteins, have been variously reported as discussed in Sections 1.2 and 1.3.6. Inter-molecular interactions among different subunits are more likely to provide a structural support for the proton pump, whereas interactions with other proteins might provide insights into V-ATPase assembly, transport, targeting, or regulation.

This study presents multiple lines of evidence that identify the glycolytic enzyme PFK-1 as a novel binding partner for the C-terminus of the pump's a subunit. I initially focused on the a4 subunit because in contrast to its intracellular counterpart a1, it has a differently targeted distribution to the apical surface of polarised cells in the kidney and an essential role in renal acid-base homeostasis. I confirmed the initial phage peptide interaction by proceeding to *in vitro* binding studies of a4(C) with intact PFK-1, and also demonstrated that a4 and PFK-1 can be co-precipitated from human kidney membranes. Lastly, immunolocalisation in human kidney suggests some enrichment of PFK-1 at the same sites as high intensity a4 immunoreactivity in acid secreting cells of the distal nephron.

It was subsequently shown that PFK-1 also binds to the ubiquitous a1 subunit's C-terminus. However, the available polyclonal anti-a1 antibody (Toyomura et al., 2000) appeared to recognise both the a1 and a4 subunits, meaning it could not be used to make a meaningful interpretation of immunoblots. Instead I used a specific *in vitro* assay to demonstrate the interaction. Of the 45 amino acids in the a subunit's C-terminus, the first 23 are identical in the a1 and a4 paralogues, with only 59% identity thereafter, and it is therefore likely that this region contains the binding domain.

The above results strongly suggest that there is a direct link between proton transport and glycolysis, and implies that the energy source for pump function may be glycolytic rather than mitochondrial. Lu *et al.* (Lu et al., 2001; Lu et al., 2003) have reported interactions at the protein level between aldolase and the ubiquitously expressed E subunit as well as with the N-terminal domain of a4 of the V-ATPase, again supporting a direct coupling between the proton pump and glycolysis. Aldolase is the next enzyme in the glycolytic pathway which, following the action of PFK-1, cleaves fructose 1,6-bisphosphate. Taken together with results obtained from this study, it is probable that the many pump and glycolytic components form a 'metabolon' to maximize the efficiency of energy provision. Whether other glycolytic enzymes, such as hexokinase, also physically interact with V-ATPase components, is unknown.

In the glycolytic pathway, PFK-1 catalyses the phosphorylation of fructose 6-phosphate by Mg-ATP to form fructose 1,6-bisphosphate and Mg-ADP. This reaction is the rate-limiting step in glycolysis, which is therefore critically dependent on the level of activity of PFK-1. This in turn is allosterically controlled both by the ratio of ATP to AMP, and by several other metabolites including citrate and fructose-2,6-bisphosphate (Dunaway, 1983). Mammalian PFKs have yet to be crystallised, but comparison with the bacterial form suggests that the binding region I have identified for a4(C) is probably distant from the described active sites of substrate binding and allosteric regulation, which involve residues R48, R97, R433, R481 and S541 among others (Kemp et al., 1987; Kemp and Gunasekera, 2002; Li et al., 1999; Zheng and Kemp, 1994). This might imply that the binding of a4(C) to PFK-1 should not directly affect its kinase function. Future studies will be required to address this issue.

The enzymatically active form of PFK-1 may be a tetramer or high order oligomer with a subunit molecular mass of approximately 85 kDa in eukaryotes (Uyeda, 1979). In humans, isozymes of PFK-1 have been grouped as muscle type (M-PFK-1), liver type (L-PFK-1), and platelet type (P-PFK-1) (Vora, 1982). All three types have been detected in human kidney (Dunaway, 1983; Nakajima et al., 1990). Rabbit M-PFK-1, to which the only commercially available antibody was raised, shares a sequence identity of about 96% with human M-PFK-1. The sequence identity between rabbit M-PFK-1 and human L-PFK-1 is not as high, about 69%, but comparative gene structural analysis has revealed a high degree of similarity, from which it is implied that this will also be true at the functional level (Elson et al., 1990). However, the potential for differences between the PFK-1 homologues means that the relative power of the available antibody to interact with the V-ATPase-PFK-1 complex may be less than optimal.

The involvement of glycolytic enzymes with the proton pump is of especial interest in the IC, which has been labelled the 'mitochondrion-rich' cell type in the nephron (Brown and Breton, 1996). Other energy-dependent transport functions of this cell are likely to depend on the mitochondrial supply of ATP, but in the case of the apical V-ATPase that is responsible for urinary acidification, I propose the alternative.

Several other lines of evidence support this hypothesis. Firstly, functional assessment of proton transport in the isolated turtle urinary bladder showed that it could be driven by the energy from both aerobic and anaerobic glycolysis (Beauwens and Al-Awqati, 1976; Steinmetz et al., 1981). Secondly, an investigation of the effect of glucose on the reversible assembly of the $V_1$ and $V_0$ domains of the pump complex in yeast (Parra and Kane, 1998) suggested coupling between V-ATPase activity and glycolysis. Interestingly, accumulation of

glucose-6-phosphate was insufficient to maintain or induce this assembly, suggesting that further glucose metabolism is required. In addition, the signalling involved in V-ATPase assembly did not appear to involve the Ras-cyclic AMP pathway, Snf1p, protein kinase C, or the stress response protein Rts1p, which are major proteins for glucose signalling in yeast (Parra and Kane, 1998). This study suggested that the transient cytosolic pH drop resulting from the initiation of glycolysis could provide a signal for activation of V-ATPase by triggering the glucose-induced assembly of the $V_1$ domain and $V_0$ domains. Thus it can be speculated that the activity of PFK-1 may have a role as an indirect regulator of the proton pump.

The a subunit was initially described as the 'large accessory' subunit of the V-ATPase, and its presence in the kidney was once disputed (Gillespie et al., 1991; Gluck and Caldwell, 1987). In recent years it has become evident from the study of human diseases that its presence is essential for normal pump function at the cell surface of renal ICs and osteoclasts at least (Frattini et al., 2000; Kornak et al., 2000; Smith et al., 2000). Through the use of different solutions in the co-immunoprecipitation experiments, I have shown that the detection of this subunit is critically dependent on the detergent employed in disrupting the cell. In particular, unsuccessful detection following a4 co-immunoprecipitation with PFK-1 when using buffers containing CHAPS suggests that this interaction, at least, is sensitive to this detergent. In a similar way, the differences in subunit composition reported for V-ATPases in the kidney may in fact have related to methodological differences in tissue preparation.

Having identified the specific interaction between a4(C) and PFK-1, I then investigated whether or not the interaction is abolished or interrupted by either of the mutations (R807Q or G820R) that are located in this domain in order to reveal

potential disease causing mechanisms. The GST pull-down assay showed that both of the mutants can interact with PFK-1. However, it is clear that this type of assay is not able to probe the roles of the mutations in terms of affinity. To address this question, BIAcore analysis was performed. Ideally, a4(C) should be used as the ligand (i.e. to be immobilised onto the chip) rather than the analyte, due to its low molecular weight (approximately 5 kDa). However, PFK-1 had to be used as the ligand due to its limited available quantity. Thus, a4(C) had to be used as the analyte rather than the ligand, although the sensitivity of the assay would be reduced if done in this way. The results showed a 2-fold lower in binding affinity for a4(C)-G/PFK-1 interaction compared to the a4(C)/PFK-1 interaction, suggesting a possible effect on the stability of the 'metabolon' formed by the V-ATPase subunits and glycolytic components. However, this is only a preliminary result which needs to be further confirmed by repeated experiments and/or using alternative techniques, once PFK-1 is available, in order to draw a conclusion.

Finally, topological studies of Vph1p, employing cysteine mutagenesis and chemical labeling, have led to a model for the a subunit in the yeast vacuole that contains 9 transmembrane helices. In this model, the amino-terminal domain lies on the cytoplasmic side of the membrane and the carboxyl-terminus on the lumenal side of the vacuole (Leng et al., 1999). However, data suggesting 6 transmembrane domains, with cytoplasmic orientation of both N-and C-termini, have also been reported in yeast and Dictyostelium (Clarke et al., 2002; Urbanowski and Piper, 1999). In the absence of crystallographic or other structural information, it has not been clear whether in the case of a4 the C-terminal tail would be found below the apical cell membrane in ICs, or protruding into the urinary space. The finding of the interaction between PFK-1 and a4(C) provides new evidence for an intracellular location of a4(C).

# CHAPTER 6
# CONCLUSION AND GENERAL DISCUSSION

## 6.1 Conclusion

In this dissertation, I have used protein engineering in combination with biochemical and biophysical methods to characterise the human a4 subunit of the V-ATPase, focusing on its potential protein-protein interactions. Previous work from our laboratory has focused on linkage of the human a4 subunit locus to rdRTA, and the identification of the coding region of the gene as well as the localisation of the protein in kidney (Karet et al., 1999a; Smith et al., 2000). Information concerning interactions involving the a4 protein is scarce. The aim of this project was to identify binding partner(s) of the human a4 subunit which may provide further understanding of the biological functions of this subunit and regulation of V-ATPase activity. In the present work, a4(N), a4(Loop2) and a4(C) were chosen as the engineering targets for *in vitro* expression attempts. Using a bacterial expression system, I have successfully expressed both the a4(Loop2) and a4(C) domains as His- and GST-tagged fusion proteins respectively, although the yield of the GST-a4(C) was at a rather low level. The GST tag in the GST-a4(C) fusion protein was subsequently removed followed by further purification of a4(C) using HPLC. CD spectrometry revealed that a4(Loop2) and a4(C) contain mainly β-sheets and α-helix secondary structural elements, respectively. However, a4(N) failed to be expressed in both bacterial and *Drosophila* S2 cell systems.

Through screening a random peptide (7-mer) M13 phage display library using a4(C) as a target protein, I identified a potential binding peptide, SWLELRP. Phage ELISA was then carried out to analyse specificity of the a4(C)/peptide interaction. Database searches revealed PFK-1 as a candidate protein for

interacting with a4(C). Assays, including *in vitro* pull-down, co-immunoprecipitation, immunohistochemistry and SPR, were subsequently performed to characterise the a4(C)/PFK-1 protein-protein interaction.

From the results presented, five conclusions can be drawn for the a4(C)/SWLELRP and a4(C)/PFK-1 interactions:

1) The binding of SWLELRP peptide to a4(C) is specific, a more than 3 times higher signal than background was observed in the ELISA assay.

   Comparative BLAST analysis revealed an almost complete match of the peptide sequence, SWLELRP, to the enzyme PFK-1, which is a key participant in the glycolytic pathway. This matched region within PFK-1 is highly conserved among mammalian PFK-1 orthologues and paralogues.

2) Several lines of evidence confirmed interaction between a4(C) and the intact PFK-1 protein. Firstly, biotin labelled a4(C) can be pulled down by bead-bound PFK-1. Secondly, bead-bound GST-a4(C) is also able to pull down PFK-1 from human kidney cytosol. Thirdly, intact a4 can be co-immunoprecipitated with PFK-1 from human kidney membrane. Finally, a4 and PFK-1 are colocalised at the apical surface of the $\alpha$-ICs in human kidney.

3) Evidence was also presented to confirm an interaction between PFK-1 and human a1 (through the C-terminus of the a1), a ubiquitously expressed paralogue of the human a subunit of the V-ATPase. This finding suggests a general property of protein-protein interaction between the human a subunit and PFK-1.

   Finding of the a subunit/PFK-1 interaction indicates a direct link between V-ATPases and glycolysis, via the C-terminal region of the pump's a subunit. This link through the a/PFK-1 interaction could be triggered by a systematic pH drop and subsequently leads to the activation of the V-ATPase for proton

pumping in expenditure of ATP generated from the glycolysis. Thus provides a potential novel regulatory mechanism between V-ATPase functions and energy supply. However, further work will be needed to confirm this hypothesis.

4) Further characterisation of this protein-protein interaction was carried out to investigate whether the two a4(C) mutants [a4(C)-G and a4(C)-Q] disrupt the a4(C)/PFK-1 interaction. As mentioned earlier, the two mutations were identified in patients diagnosed with rdRTA (Smith et al., 2000; Stover et al., 2002). Through a GST pull-down assay, it was seen that both mutants, like the wild type a4(C), were also able to interact with PFK-1.

   To examine this in more detail, SPR analysis was performed using the BIAcore system. The preliminary data obtained from this assay showed an approximately 2-fold difference, in terms of $K_D$, between the wild-type a4(C) and the mutant a4(C)-G. However, this result needs to be further confirmed.

5) The cellular location of the C-terminal tail of the a subunit has been disputed. As mentioned earlier, data suggesting an extracellular location (Leng et al., 1999) and a cytoplasmic orientation of the C-termini (Clarke et al., 2002; Urbanowski and Piper, 1999) have been reported. In the absence of crystallographic or other structural information, it is not clear whether in the case of a4 the C-terminal tail would be found below the apical cell membrane in ICs, or found protruding into the urinary space. The finding of an interaction between PFK-1 and a4(C) provides new evidence for an intracellular location of a4(C).

## 6.2 General Discussion

Protein-protein interactions are operative at almost every level of cell function: in the structure of intracellular organelles; the transport machinery across

the various membranes; packaging of chromatin; muscle contraction; signal transduction; and regulation of gene expression etc. Over decades of study, much effort has been devoted to uncover the regulation of V-ATPase activity in different membranes of a wide variety of organelles, tissues and organisms. Increasing evidence indicates that binding of V-ATPase subunits to other proteins, other than the component elements of this enzyme, is more than likely to be involved in regulatory events as mentioned earlier in Section 1.2.2. Therefore, in order to understand better how the activity of the V-ATPase is regulated, it is necessary to identify and study those proteins that bind to the V-ATPase subunits but are not elements of the core structure of this enzyme.

However, the majority of studies of the mammalian a4 subunit of the V-ATPase, to date, have focused on the gene structure and cellular localisation, so little is known about its protein-protein interactions. As noted from the yeast orthologue Vph1p, the N- and C-terminal domains are likely to contain signals in controlling proton translocation, targeting, assembly and regulation of V-ATPase activity (Kawasaki-Nishi et al., 2001a; Kawasaki-Nishi et al., 2001c; Leng et al., 1998; Leng et al., 1996; Manolson et al., 1992). However, molecular mechanisms of these events have not yet been unravelled. In this study, I have chosen to seek binding partners for a4 in order to address its potential role, on a molecular basis, in the V-ATPase complex.

### 6.2.1 Expression of a4(N), a4(Loop2) and a4(C)

I chose a4(N), a4(Loop2) and a4(C) domains, rather than the intact a4 protein, as the engineering targets for the following reasons. Firstly, intact a4 is a relative large integrated membrane protein containing multiple putative TM helices. Although overexpression of full-length integral membrane proteins is highly

desirable since their TM regions often contain important structural information directing their folding, oligomerisation or subcellular sorting etc, high-level expression of them in *E. coli* remains a difficult task due to toxic effects exerted by the hydrophobic protein domains on host cells (Cosson and Bonifacino, 1992; Lemmon and Engelman, 1994; Shaw and Miroux, 2003). Therefore, bacterial expression of membrane proteins has frequently been restricted to their soluble domains. Indeed, my attempts to express and purify the intact human a4 protein in BL21 cells were not successful in the present work. Secondly, both N- and C-terminal domains of the yeast a subunit orthologue have been implicated in the regulation of V-ATPase activity. Thirdly, the N-terminal domain of the yeast a subunit was previously found to interact with subunits A and H of V-ATPase, which is likely to provide structural support as well as coupling the ATP hydrolysis and proton translocation of the V-ATPase (Kawasaki-Nishi et al., 2001a). Finally, the 2$^{nd}$ loop is the most conserved region, in terms of amino acid sequence, among the loops of human a subunit paralogues.

A bacterial *E. coli* expression system was chosen as the first choice to produce the three domains for several reasons. Firstly, the cells are easy and quick to grow to high cell densities which could possibly allow for large-scale production of proteins. a4(Loop2) was expressed at a rather high level in this manner. Secondly, there are many commercial and non-commercial expression vectors available with different N- and C-terminal tags. Also many different strains exist which are optimised for special applications as well as a number of well developed transformation protocols. Finally, the very well understood molecular genetics of bacteria permit the expression of recombinant proteins to be tightly controlled. This is important because it could possibly allow for *E. coli* to be used as a host for expressing proteins that are toxic, or inhibit cell growth, by using

promoters whose expression is inducible. However, many eukaryotic proteins still cannot be successfully expressed to a high level or with a properly folded form in bacterial cells even under controlled expression. In the case of a4(C), although it was successfully expressed as a soluble and functional form, the expression level was rather low. This might be the result of either the tocxicity of this protein to BL21 cells and/or the sensitivity of this protein to proteinase (Gillespie et al., 1991; Gluck and Caldwell, 1987). Nevertheless, the *E. coli* expression system has several disadvantages, the main one being the lack of machinery for post-translational modification. This makes *E. coli* cells an inadequate expression system for many eukaryotic proteins; especially for those that require post-translational modification for folding and functionality. In the case of a4(Loop2) expression, although the yield of the a4(Loop2) was very high, it was initially formed as inclusion bodies which can often be very difficult, if not impossible, to refold. On the other hand, although a4(N) was expressed in a soluble form with a GST tag, it was not properly folded. Despite it being disappointing not to obtain a soluble protein with structural elements, it reinforces the importance of checking the folding situation of an overexpressed protein before using it in further studies.

In addition to a bacterial expression system, other systems including yeast, insect, and mammalian, are also widely used for protein expression. These share the property of post-translational modification machinery, which is necessary for the expression of many eukaryotic proteins, especially those that have failed to be expressed in bacterial systems. Both yeast and insect cell expression systems can perform similar post-translational modifications to those occurring in mammalian cells. Also, both systems are suitable for the relatively large-scale production of proteins. However, they can be challenging from the point of view of the transformation and transfection processes. For example, attempts to express

a4(N) using *Drosophila* S2 cells failed to produce detectable a4(N). One possible reason could be due to low levels of expression caused by low transfection efficiency. If this is the case, the efficiency could possibly be increased by using an alternative selection vector, which would in turn increase the level of protein expression. Alternatively, the a4(N) protein may have been produced in an unfolded (or misfolded) form, which would subsequently be degraded by the ER quality control in S2 cells. If this was the case, using an alternative expression system(s), such as mammalian cells, might have to be the choice in future studies, although the expression level in such systems is generally quite low.

### 6.2.2 Identification of Interacting Ligands

Both the expressed and purified a4(C) and a4(Loop2) were used as target proteins to screen random peptide M13 phage display libraries for potential binding partners. Through screening of a 7-mer random peptide phage display library with a4(C), a potential binding peptide sequence, SWLELRP, was identified and the specificity of the a4(C)/peptide interaction was further confirmed by a phage ELISA analysis. This peptide sequence almost completely matched a linear stretch of amino acids at C-terminus of PFK-1, and the matched sequence is highly conserved among mammalian PFK-1 orthologues and paralogues. This finding suggests that PFK-1 is a potential binding partner for a4(C). However, not all phage display library screening produced such a successful output - as seen in the assays using a4(Loop2) as a target protein to screen both 7- and 12-mer random peptide phage display libraries. The library screens identified a consensus interacting motif K(L/V)WVIPQ in approximately 75% of cases and this a4(Loop2)/peptide interaction was further confirmed by phage ELISA analysis. However, database searches did not reveal any protein containing this motif that

had apparent physiological relevance to the V-ATPase. There could be a number of reasons for identification of a binding sequence, but not an actual protein of interest as described in Section 4.3. The main possible reason is that the identified peptide sequence is only formed in a binding ligand with a 3D structure. If this is the case, an alternative approach, in future work, is either to use alternative phage display libraries containing structured peptides, or to use alternative strategies, such as Y2H system. The latter usually combines libraries with larger insert size and is especially good for searching binding ligands that require post-translational modification for interaction.

The inability to obtain folded a4(N) prevented the use of a phage display strategy. Instead, a Y2H based CytoTrap two-hybrid system was employed to screen a human testis cDNA library (contains insert size greater than 0.4 kb) for binding ligand(s) of a4(N). In addition, as an alternative, a4(Loop2) was also used as a bait to screen the same library. Through the screens, several hypothetical/unknown proteins were identified as potential binding ligands, one for a4(Loop2) and two for a4(N), as described in Sections 4.2.2.3 and 4.2.2.4. In February 2001, the sequence of the human genome was reported with over 90% coverage, revealing an estimated 31,000 proteins, many of which, however, are of unknown function (Lander et al., 2001; Venter et al., 2001). In order to determine the biological functions of these proteins, much effort has now been devoted to identifying protein-protein interactions using various combinational biological methods. Therefore, more information regarding these identified unknown proteins may become available in the future. If this is the case, decisions for further characterisations of the interactions between these potential ligands and a4(N) [or a4(Loop2)] can be made accordingly. Nevertheless, in addition to the unknown proteins a potential binding candidate of a4(N), the proteasome (β subunit), was

also identified through the Y2H library screen. The proteasome is a protein degradation machine which is a large protein complex (also called 26S proteasome) formed by a 20S core chamber and two 19S caps. The 20S complex, which is the heart of the proteasome, does the job of protein degradation and is composed of 4 rings. Each of the 4 rings is composed of $\alpha$ and $\beta$ subunits. When a protein is misfolded, it is guided into the 20S core chamber where the peptide bonds of the protein are cleaved every 8-9 amino acids. In addition to degrading misfolded proteins, proteasomes are also to be involved in breaking down proteins which have normal structures but are retained as unassembled. An example of this came from studies of the degradative pathway of non-assembled Vph1p, the yeast orthologue of the a subunit of V-ATPase (Hill and Cooper, 2000). Whether or not the binding of a4(N) to proteasome is due to its misfolding or other reasons remained unclear.

Finally, screening different libraries associated with the Y2H system could possibly identify additional protein(s) of interest. Initially, there were several cDNA libraries available for the CytoTrap two-hybrid system. The human testis cDNA library was chosen in this study mainly due to three reasons. Firstly, there was no mammalian kidney cDNA library suitable for this system. Secondly, the human a4 protein is also expressed in testis. Finally, testis cDNA library is supposed to contain most of transcripts compared to those from other tissues. However, a mouse kidney cDNA library, which is suitable to be used with the CytoTrap two-hybrid system, has very recently become available and this provides an alternative material for identification additional binding partners for both a4(N) and a4(Loop2) in the future.

### 6.2.3 Characterisation of a4/PFK-1 Protein-protein Interaction

Through the phage display assay, PFK-1 was identified as a potential binding partner for a4, via its C-terminal soluble tail. To test whether a4(C) indeed interacts with PFK-1 at a protein level, multiple assays including *in vitro* pull-down, co-immunoprecipitation and immunofluorescence microscopy were carried out and from this several lines of evidence were presented to demonstrate the a4(C)/PFK-1 protein-protein interaction. Association between these two proteins indicates a direct link between the V-ATPase and the ATP-generating glycolytic pathway, via the a subunit, and suggests a novel regulation mechanism between energy supply and V-ATPase function.

Having identified a specific interaction between a4(C) and PFK-1, the next question was what are the influences of the mutations (R807Q and G820R) within the a4(C) region identified from patients with rdRTA, on the binding of the a4(C) to PFK-1. In other words, do these mutations cause rdRTA through abolishing or altering the a4/PFK-1 interaction? To address this question, I next carried out a GST pull-down assay. Like the wild-type a4(C), a4(C) domains containing either R807Q or G820R mutations fused with GST bound sepharose beads did pull down PFK-1 from human kidney cytosol. This suggests, at least, that both mutants do not disrupt the a4(C)/PFK-1 interaction completely. Therefore, I further asked if binding affinity was affected by the mutations. There are several methods, including stopped-flow, analytic centrifugation, ITC, fluorescence titration and SPR, which are commonly used for binding affinity analysis of a protein-protein interaction. However, the number of applications which could be used in the present study was restricted by either the limited amount of PFK-1 protein or the existence of tryptophan residues in both PFK-1 and a4(C). The restricted quantity of PFK-1 prevented the use of techniques such as stopped-flow, analytic

centrifugation and ITC, whereas the latter limits the use of fluorescent titration. The SPR based BIAcore system was, therefore, the only method which could be used at present.

To perform the assay, PFK-1 had to be used as a ligand (to be immobilised on the BIAcore chip), rather then an analyte (to be run over the chip surface), again due to insufficient quantity. This is not the favoured way to run the assay, as a4(C) is a relatively small molecule (~ 5 kDa). Performing the assay in this way would normally reduce its sensitivity. The preliminary results obtained through the BIAcore analysis showed that mutation G820R in a4(C) [a4(C)-G] can cause approximately 2-fold decrease of the PFK-1 binding affinity compared to the wild type a4(C). However, if there was a sufficient quantity of PFK-1, BIAcore analysis could be performed more optimally, i.e. the a4(C) was immobilised to the chip as the ligand and PFK-1 was run over the a4(C) as an analyte. In this way, the sensitivity of the analysis might be much increased, which in turn might reduce potential errors.

### 6.2.4 Future Studies

*Further Characterisation of the a4/PFK-1 Interaction*

First of all, efforts will need to be made to produce sufficient quantities of PFK-1 protein for example by purification from certain tissue extracts. This would allow me to carry out several further studies regarding the a4/PFK-1 interaction as follows.

Firstly, studies of regulation of V-ATPase activity, in yeast, have shown that extracellular glucose concentrations regulate V-ATPase activity *in vivo* by regulating the extent of the association between $V_1$ and $V_0$ domains (Graf et al., 1996; Kane, 1995; Sumner et al., 1995; Wieczorek et al., 2000). Therefore, further

characterisation of the direct coupling between glucose metabolism and V-ATPase activity through a/PFK-1 interaction could be performed if a PFK-1 deficient cell line was available. If this was the case, the a/PFK-1 interaction might regulate V-ATPase activity by controlling disassembly/reassembly of $V_1$ and $V_0$ domains - a potential mechanism which has been suggested for regulating V-ATPase activity *in vivo* (Nakamura et al., 1997; Puopolo et al., 1992b; Zhang et al., 1992).

Secondly, as described in Section 5.3, it was assumed that binding of a4 to PFK-1 would not affect its enzyme catalytic activities. However, in the absence of 3D structural information for eukaryotic PFK-1, it is hard to draw a precise conclusion. One way to elucidate this aspect would be to perform some enzyme activity assays to see whether binding of a4(C) to PFK-1 affects binding of fructose 6-phosphate which is the substrate of PFK-1, or vice versa. This would provide more information to understand better the mechanism of the coupling of glycolysis directly to V-ATPase activity through the a4/PFK-1 interaction.

Thirdly, more measurements of binding affinity and kinetic analysis of the interaction between the wild-type or mutant a4(C) and PFK-1 will be performed using SPR and/or other methods mentioned above.

Finally, attempts would be made to grow crystals of the a4(C)/PFK-1 complex in order to gain 3D structural information of this complex. Knowledge of this can provide information on how they interact and why a particular mutation causes disease.

*Further Characterisation of the a4(N) and a4(Loop2) Domains*

As mentioned above, additional screening of the mouse kidney cDNA library with a4(Loop2) will be carried out for potential binding ligands. In addition, a technique combining *in vitro* pull-down assay and mass spectrometry has recently been established locally, which has opened an alternative method for

identification of binding partners of a4(Loop2). In this technique, the His-a4(Loop2) expressed can be captured onto Ni-NTA agarose resin and incubated with human kidney crude protein fraction, followed by SDS-PAGE analysis. Unique visible band(s) on the gel compared to a control could possibly represent potential binding ligand(s) which can be characterised by mass spectrometry.

Further characterisation of any a4(N) protein-protein interactions will be carried out if a potential a4(N)-binding partner is identified in an additional Y2H screen. In addition to performing immunoprecipitation and immunohistochemistry assays, this will involve attempts to express a4(N) in systems other than bacteria to produce a4(N) suitable for *in vitro* pull-down analysis. If the expression level is relatively high, i.e. if there is an adequate quantity of a4(N), structural analysis will be carried out to gain more information about functional domains and motifs contained in this part of the a subunit protein. This kind of information will not just provide evidence to support any identified interactions, but may also be useful in predicting other potential binding partners of a4(N).

*Cell Biology and Model Systems*

It was assumed that the sequence divergence between the a subunit paralogues of V-ATPase may contribute to the differential targeting as well as to the other regulatory properties of V-ATPase (Kawasaki-Nishi et al., 2001c; Nishi and Forgac, 2000; Toyomura et al., 2000). Studies using chimeras constructed with either N- and C-terminal domains of the yeast a subunits (Vph1p and Stv1p) have indeed demonstrated that the differential targeting and proton translocating signals are likely to be embedded in the N- and C- terminus, respectively (Kawasaki-Nishi et al., 2001c; Manolson et al., 1994). It is interesting to see whether this is true in the acid-secreting cells in mammalian kidney. Unfortunately, no cell lines derived from kidney epithelial cells have, so far, been identified to be

able to investigate targeting of the a4 subunit onto the cell surface. However, if a suitable cell line becomes available in the future, similar studies could be carried out between a4 and its ubiquitously expressed counterpart a1 subunit.

In addition to cell lines, model systems based on organisms would be good for the functional studies of a4. The mouse is closely related to humans: most human genes have functional mouse counterparts and the genome is organised in a very similar manner. Indeed, both mouse and human contain four paralogues of the a subunit of the V-ATPase. So mouse models can be established to see whether the a4-encoding gene (*Atp6v0a4*) knockout mice show similar phenotypes as seen in those *ATP6V0A4*-linked patients and also to investigate the molecular mechanisms underlying this disease development as well as the identification of targeting and regulatory pathways.

Finally, a coated vesicle $Cl^-$ channel was proposed to function in dissipating the membrane potential generated by V-ATPase in coated vesicles (Mulberg et al., 1991). When the $Cl^-$ channel activity decreased by treatment with a phosphatase, a parallel decrease of vesicular acidification by the V-ATPase was also observed. The decrease in $Cl^-$ conductance and ATP-dependent acidification can be reversed by treatment with PKA and MgATP. These results indicate that modulation of the $Cl^-$ channel activity by PKA affects vesicular acidification by V-ATPase. If this regulation mechanism does exist for the a4-containing V-ATPase in the kidney, the possibility of the existence of a chloride channel at the apical surface of α-ICs can be proposed. If a suitable cell line or organism model becomes available, it would be interesting to test this hypothesis. If this is the case, it will provide new insight to the structure of the α-IC in human kidney.

# APPENDIX

## A.1 General Buffers, Bacterial Growth Media and Antibiotic Solutions

### A.1.1 General Buffers

The following buffers were made up in deionised water ($dH_2O$), sterilised by autoclaving and stored at RT.

*10x Phosphate-buffered saline (PBS) (pH 7.4)*

20 mM $KH_2PO4$, 100 mM $Na_2HPO_4$, 1.37 M NaCl, 27 mM KCl

*PBST*

1x PBS, 0.05% polyoxyethylenesorbitan monolaurate (Tween 20), unless otherwise stated.

*5x TBE*

0.45 M Tris-borate, 10 mM EDTA, pH 8.0

*10x Tris-buffered saline (TBS) (pH 7.4)*

8 g/l NaCl. 0.2 g/l KCl, 3 g/l Tris base, 0.015 g/l phenol red

*50x TAE*

2 M Tris-acetate, 50 mM EDTA (pH 8.0)

*Tris EDTA (TE) (pH 8.0)*

10 mM Tris-Cl (pH 8.0), 0.1 mM EDTA (pH 8.0)

### A.1.2 Bacterial Growth Media and Antibiotic Solutions

The following growth media were prepared in $dH_2O$, sterilised in the absence of antibiotics by autoclaving and stored at RT. Antibiotic stock solutions were made with $H_2O$, sterilised using 0.22 μm filters and stored at -20°C.

*Luria-Bertani (LB) medium (pH 7.0)*

10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl

*LB agar (LA)*

LB medium, 15 g/l Bacto agar added before autoclaving

*2x YT medium (pH 7.0)*

16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl

*SOC medium (pH7.0)*

20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl, 2.03 g/l $MgCl_2$, 20 mM glucose

*Antibiotic solutions*

Sterile-filtered stock solution of ampicillin (amp) (50 mg/ml $H_2O$) and tetracycline (tet) (15 mg/ml 50% ethanol (v/v))] were kept at -20°C and added to sterilised media as required [amp (100 μg/ml), tet (15 μg/ml)], unless otherwise stated.

## A.2 Antibodies, Cell Strains, Plasmids and Primers

| Antibodies | Description | Applications (dilutions) | Reference/Source |
|---|---|---|---|
| RA2922 | Rabbit anti-human a4 (polyclonal, serum) | Wb, 1 : 1000<br>IHC, 1 : 5000 | (Smith et al., 2000) |
| α-PFK-1 | Goat anti-rabbit muscle type PFK-1 (Polyclonal, purified) | Wb, 1 : 1000-2000<br>IHC, 1 : 100 | Amersham |
| α-M13 | Mouse anti-M13 (Monoclonal) | EL, 1 : 5000 | Amersham |
| α-His | Mouse anti-RGS-$His_4$ (Monoclonal) | Wb, 1 : 1000 | QIAGEN |
| Secondary | Goat anti-rabbit IgG (HRP-conjugated) | Wb, 1 : 5000 | DACO Ltd. |
| Secondary | Rabbit anti-mouse IgG (HRP-conjugated) | Wb, 1 : 15,000 | DACO Ltd. |
| Secondary | Swine α-rabbit IgG (TRITC-labelled) | IHL, 1 : 100 | Vector Laboratories |
| Secondary | Rabbit anti-goat IgG (FITC-labelled) | IHL, 1 : 100 | Vector Laboratories |

**Table A.1 Antibodies.**

Wb: Western blotting; EL: ELISA; IHC: Immunohistochemistry

| Name | Description/Genotype | Reference or Source |
|---|---|---|
| *BL21(DE3)C41 (*E. coli*) | Derived from BL21(DE3) which has the advantage of being deficient in both lon and ompT proteases. No antibiotic resistance. | (Miroux and Walker, 1996) |
| AD 494 (DE3)(*E. coli*) | Containing thioredoxin reductase (trxB) mutants enable disulfide bond formation in the cytoplasm, providing the potential to produce properly folded, active proteins. Selectable on kan. | (Derman et al., 1993) |
| ER2738 (*E. coli*) | *Tet$^r$ Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacI$^q$ZΔM15 Tn10 (Tet$^r$) Amy Cam$^r$]$^a$* | NEB |
| XL10-Gold (*E. coli*) | *Tet$^r$ Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte [F' proAB lacI$^q$ZΔM15 Tn10 (Tet$^r$) Amy Cam$^r$]$^a$* | Stratagene |
| Schneider 2 cells (*Drosophila* S2 cells) | A cell line derived from *Drosophila melanogaster*. They are easily maintained in loosely adherent or suspension culture at RT with no requirement of $CO_2$. | (Schneider, 1972) |
| cdc25H strain a (*yeast*) | *MATa ura3-52 his-3-200 ade2-101 lys2-801 trp 1-901 leu2-3 112 cdc25-2 Gal$^+$* | Stratagene |
| cdc25H strain a (*yeast*) | *MaTα ura3-52 his-3-200 ade2-101 lys2-801 trp 1-* | Stratagene |

**Table A. 2 Cell strains.**

*The strain C41(DE3) was derived from BL21(DE3) [E. coli F$^-$ *ompT hsdS*$_B$ ($r_B^-$ $m_B^-$) *gal dcm* (DE3)]

| No. | Name | Sequences (5'→ 3') | Tm (°C) | *RS/(Application |
|---|---|---|---|---|
| 1 | 5'sen(N)N1B EcoRI | CC**GAATTC**ATGGCGTCTGTGTTTC GAAG | 60 | *Eco*RI (Cloning in pGEX-4T1) |
| 2 | 3'anti(N)N1B NotI | TCCACGTC**GCGGCCGC**CTACTCCT CGGCTGTGCC | 58 | *Not*I (Cioning in pGEX-4T1) |
| 3 | 5'sen(N)N1B Sal I/Not I | CC**GAATTCGTCGAC**GATGGCGTCT GTGTTTCGAAG | 64 | *Sal*I (Cloning in pSOS) |
| 4 | 3'anti (N)N1B NotI/sal I | TCCACATC**GCGGCCGC**GTGGGTTT ATCTCCCGGTAGC | 66 | *Sal*I (Cloning in pSOS) |
| 5 | 5'sen(N)N1B Spe I/Age I | CG**ACTAGT**ATGGCGTCTGTGTTTC GAAGC | 64 | *Spe*I (Cloning in pMT/BiP/V5-His |
| 6 | 3'anti (N)N1B Age I/Spe I | CG**ACCGGT**TGGGTTTATCTCCCGG TAGC | 62 | *Age*I (Cloning in pMT/BiP/V5-His |
| 7 | 5'senN1Bloop2B/E | CG**GGATCC**AATGACTGCTTCTCCA AGTCCTTG | 70 | *Bam*HI (Cloning in miniprseta- |
| 8 | 3'antiN1Bloop2E/B | CC**GAATTC**CGACATCTTCATTTTAT A CGAG | 60 | *Eco*RI (Cloning in miniprseta-mac) |
| 9 | 5'sen(Loop2)N1B Sal I/Not I | CC**GTCGAC**GAATGACTGCTTCTCC AAGTCC | 62 | *Sal*I (Cloning in pSOS) |
| 10 | 3'anti (N)N1B NotI/sal I | TCCACATC**GCGGCCGC**CGACATCT TCATTTTATACGAG | 60 | *Not*I (Cloning in pSOS) |
| 11 | 5'sen(C)N1B BamHI | CG**GGATCC**ATGGAGGGCCTCTCTG CTTTC | 66 | *Bam*HI (Cloning in pGEX-4T1) |
| 12 | 3'anti(C)N1B EcoRI | CC**GAATTC**CTACTCCTCGGCTGTG CCATC | 68 | *Eco*RI (Cloning in pGEX-4T1) |
| 13 | 5'-F-a4(C)-R→Q-muta. | CTGCACGCCCTGCAACTGCACTGG GTTGAG | 81.65 | Mutagenesis |
| 14 | 3'-R-a4(C)-R→Q-muta. | CTCAACCCAGTGCAGTTGCAGGGC GTGCAG | 81.65 | Mutagenesis |
| 15 | 5'-F-a4(C)-G→R-muta. | CCAGAACAAGTTCTATGTCAGGGAT GGTTACAAGTTTTCTCC | 80.6 | Mutagenesis |
| 16 | 3'-R-a4(C)-G→R-muta | GGAGAAAACTTGTAACCATCCCTGA CATAGAACTTGTTCTGG | 80.6 | Mutagenesis |
| 17 | 5' pGEX | GGGCTGGCAAGCCACGTTTGGTG | 76 | Sequencing |
| 18 | 3' pGEX | CCGGGAGCTGCATGTGTCAGAGG | 76 | Sequencing |
| 19 | 96 gIII 5' | CCCTCATAGTTAGCGTAACG | 60 | Sequencing |
| 20 | Sos 5' primer | CCAAGACCAGGTACCATG | 56 | Sequencing |
| 21 | Sos 3' primer | GCCAGGGTTTTCCCAGT | 54 | Sequencing |
| 22 | Myr 5' primer | ACTACTAGCAGCTGTAATAC | 56 | PCR; Sequencing |
| 23 | Myr 3' primer | CGTGAATGTAAGCGTGACAT | 58 | PCR; Sequencing |
| 24 | pMT Forward | CATCTCAGTGCAACTAAA | 50 | Sequencing |
| 25 | RTA2Pr6 | TATGAATGTGAACAGCTTTCAAAGG | 68 | Sequencing |
| 26 | RTA2Pr10 | GCTGATGATTTCTTTACTGAGGAC | 68 | Sequencing |
| 27 | RTA2Pr14 | GAGCCTGCGGTGGAGCGCAG | 70 | Sequencing |
| 28 | RTA2Pr17 | CTTCTGCACCTTGATGAGCCAG | 68 | Sequencing |
| 29 | RTA2Pr13 | CTCGGCAGATTCGCCACAGTAAC | 72 | Sequencing |
| 30 | RTA2Pr7 | AGGATTCTCTCCAGTGATTCACAC | 70 | Sequencing |

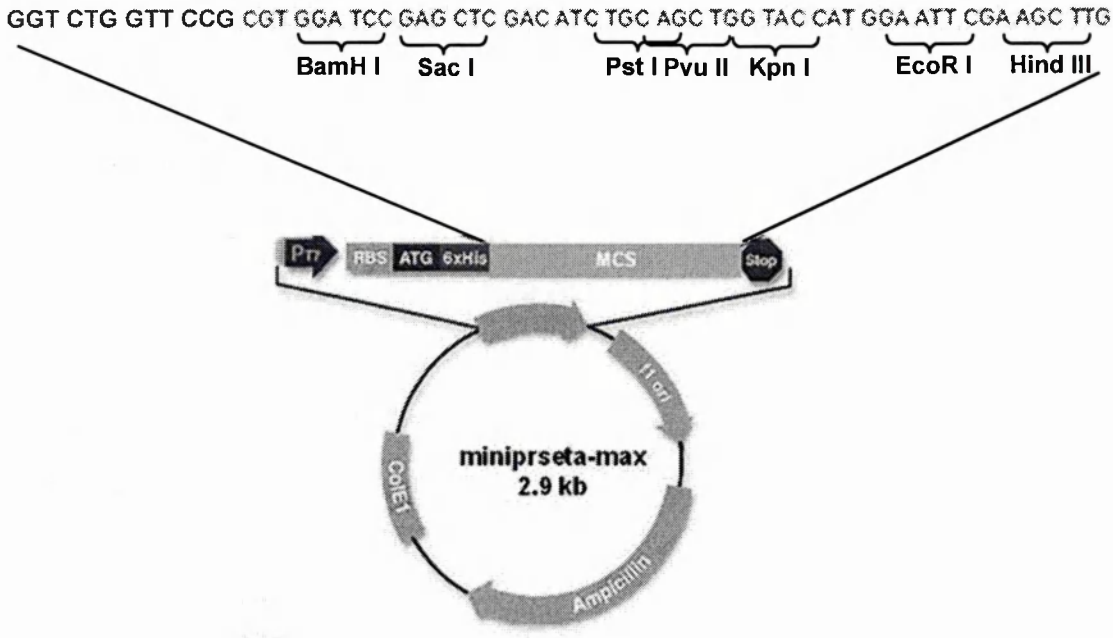
## Table A.3 Primers

*Tm value For primers No.13-16, was calculated according to a formula: Tm = 81.5 + 0.41(%GC) – 675 / N - % mismatch, which was provided by the manufacturer's instruction. Where, N is the primer length in base pairs. The highlighted bases in primers indicate restriction sites (*RS), and, the underlined bases (in primers No.13-16) indicate where codons have been changed.

| Plasmid | Characteristics | Source |
|---|---|---|
| pGEM-T Easy –a4 | Cloning vector containing human *a4* cDNA fragment, *amp*$^r$, *lacZ* | Dr Smith A.N (Department of Medical Genetics, University of Cambridge) |
| pGEX-4T-1 | Bacterial expression vector, *amp*$^r$ | Pharmacia Biotech, |
| pGEX-a4(C) | pGEX-4T-1 containing *a4(C)-WT* cDNA fragment | This study |
| pGEX-a4(C)-G | pGEX-4T-1 containing *a4(C)-G* cDNA fragment | This study |
| pGEX-a4(C)-Q | pGEX-4T-1 containing *a4(C)-Q* cDNA fragment | This study |
| pGEX-a4(N) | pGEX-4T-1 containing *a4(N)* cDNA fragment | This study |
| pSOS | Yeast expression vector, *LEU2*, *amp*$^r$ | Stratagene |
| pSOS-a4(N) | pSOS containing *a4(N)* cDNA fragment | This study |
| pSOS-a4(Loop2) | pSOS containing *a4(Loop2)* cDNA fragment | This study |
| pSOS-Col I | pSOS containing *Col I* cDNA fragment | Stratagene |
| pMyr | Yeast expression vector, Myristylation signal, *URA3*, *cam*$^r$ | Stratagene |
| miniprseta-mac | Bacterial expression vector, *amp*$^r$, His-tagged | Dr Jane Clark (department of Chemistry, University of Cambridge) |
| prseta- a4(Loop2) | miniprseta-mac containing *a4(Loop2)* cDNA fragment | This Work |
| pMT/Bip/V5-His A | *Drosophila* cell expression vector, His-tagged, *amp*$^r$, | Invitrogen |
| pMT-a4(N) | pMT/Bip/V5-His A containing *a4(N)* cDNA fragment | This study |
| pCoHYGRO | Selection vector, *amp*$^r$, *hyg*$^r$, | Invitrogen |

**Table A.4 Plasmids**

**a**




**b**

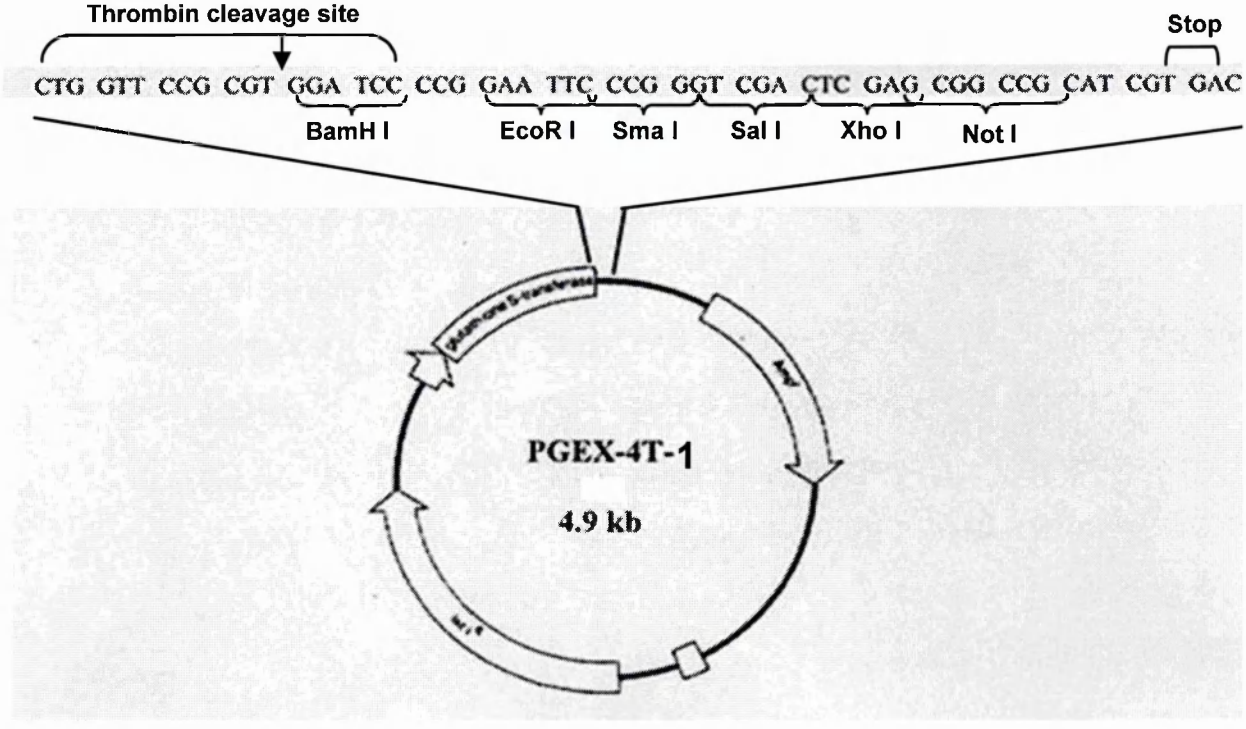



**Figure A.1 Bacterial expression vectors for expression in *E. coli* cells.**

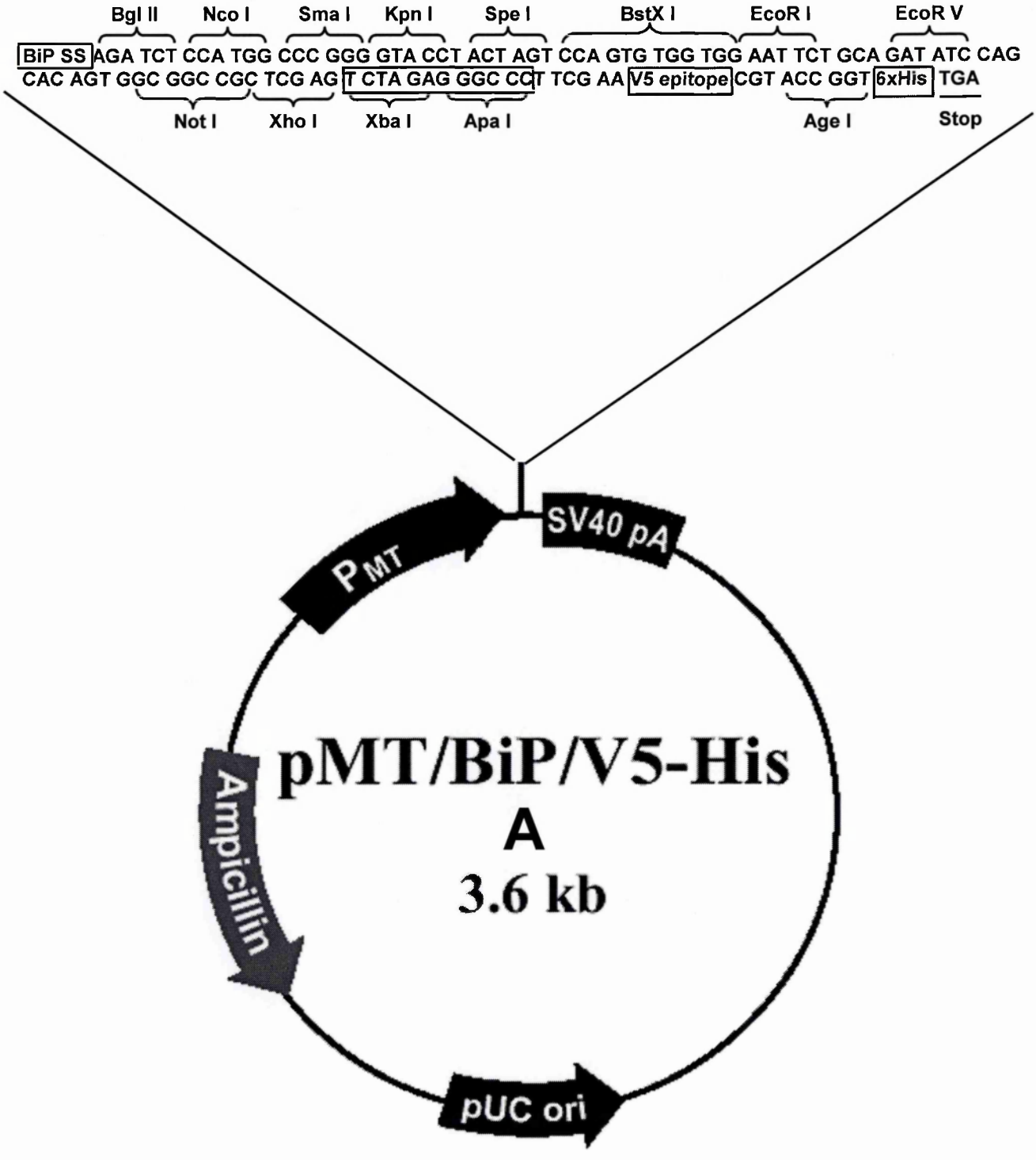



**Figure A.2 Insectile expression vector for expression in S2 cells.**
Boxed nucleotides indicate the variable region for pMT/BiP/V5-His A, B and C vectors.

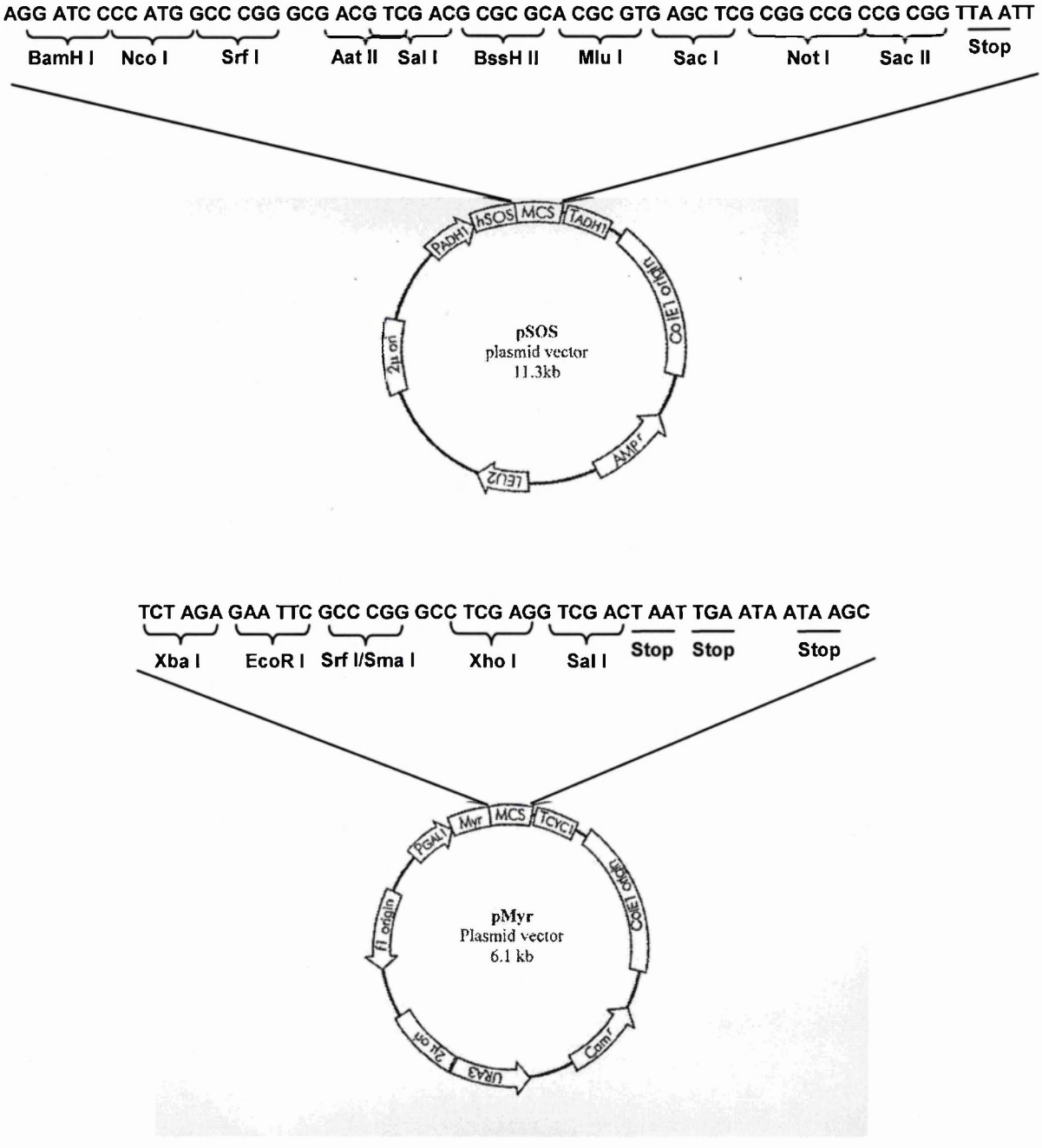



**Figure A.3 Yeast expression vectors.**

# REFERENCE


Abrahams, J. P., Leslie, A. G., Lutter, R., and Walker, J. E., 1994, Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria, *Nature* **370**(6491):621-8.

Adachi, I., Arai, H., Pimental, R., and Forgac, M., 1990a, Proteolysis and orientation on reconstitution of the coated vesicle proton pump, *J Biol Chem* **265**(2):960-6.

Adachi, I., Puopolo, K., Marquez-Sterling, N., Arai, H., and Forgac, M., 1990b, Dissociation, cross-linking, and glycosylation of the coated vesicle proton pump, *J Biol Chem* **265**(2):967-73.

Al-Awqati, Q., 1996, Plasticity in epithelial polarity of renal intercalated cells: targeting of the H(+)-ATPase and band 3, *Am J Physiol* **270**(6 Pt 1):C1571-80.

Alper, S. L., Natale, J., Gluck, S., Lodish, H. F., and Brown, D., 1989, Subtypes of intercalated cells in rat kidney collecting duct defined by antibodies against erythroid band 3 and renal vacuolar H+-ATPase, *Proc Natl Acad Sci U S A* **86**(14):5429-33.

Aniento, F., Gu, F., Parton, R. G., and Gruenberg, J., 1996, An endosomal beta COP is involved in the pH-dependent formation of transport vesicles destined for late endosomes, *J Cell Biol* **133**(1):29-41.

Anraku, Y., Umemoto, N., Hirata, R., and Ohya, Y., 1992, Genetic and cell biological aspects of the yeast vacuolar H(+)-ATPase, *J Bioenerg Biomembr* **24**(4):395-405.

Apps, D. K., Percy, J. M., and Perez-Castineira, J. R., 1989, Topography of a vacuolar-type H+-translocating ATPase: chromaffin-granule membrane ATPase I, *Biochem J* **263**(1):81-8.

Arai, H., Berne, M., and Forgac, M., 1987, Inhibition of the coated vesicle proton pump and labeling of a 17,000-dalton polypeptide by N,N'-dicyclohexylcarbodiimide, *J Biol Chem* **262**(23):11006-11.

Arai, H., Pink, S., and Forgac, M., 1989, Interaction of anions and ATP with the coated vesicle proton pump, *Biochemistry* **28**(7):3075-82.

Arai, H., Terres, G., Pink, S., and Forgac, M., 1988, Topography and subunit stoichiometry of the coated vesicle proton pump, *J Biol Chem* **263**(18):8796-802.

Aronheim, A., 2000, Protein recruitment systems for the analysis of protein-protein interactions, *Biochem Pharmacol* **60**(8):1009-13.

Au, C. L., and Wong, P. Y., 1980, Luminal acidification by the perfused rat cauda epididymidis, *J Physiol* **309**:419-27.

Bailleul-Winslett, P. A., Newnam, G. P., Wegrzyn, R. D., and Chernoff, Y. O., 2000, An antiprion effect of the anticytoskeletal drug latrunculin A in yeast, *Gene Expr* **9**(3):145-56.

Baltrusch, S., Lenzen, S., Okar, D. A., Lange, A. J., and Tiedge, M., 2001, Characterization of glucokinase-binding protein epitopes by a phage-displayed peptide library. Identification of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase as a novel interaction partner, *J Biol Chem* **276**(47):43915-23.

Barzel, U. S., and Hart, H., 1973, Studies in calcium absorption: initial entry calcium into the gastrointestinal tract in hyperparathyroidism and in a case of renal tubular acidosis, *Nephron* **10**(2):714-87.

Bastani, B., and Gluck, S. L., 1996, New insights into the pathogenesis of distal renal tubular acidosis, *Miner Electrolyte Metab* **22**(5-6):396-409.

Bastani, B., McEnaney, S., Yang, L., and Gluck, S., 1994, Adaptation of inner medullary collecting duct vacuolar H-adenosine triphosphatase to chronic acid or alkali loads in the rat, *Exp Nephrol* **2**(3):171-5.

Bastani, B., Purcell, H., Hemken, P., Trigg, D., and Gluck, S., 1991, Expression and distribution of renal vacuolar proton-translocating adenosine triphosphatase in response to chronic acid and alkali loads in the rat, *J Clin Invest* **88**(1):126-36.

Bastani, B., Underhill, D., Chu, N., Nelson, R. D., Haragsim, L., and Gluck, S., 1997, Preservation of intercalated cell H(+)-ATPase in two patients with lupus nephritis and hyperkalemic distal renal tubular acidosis, *J Am Soc Nephrol* **8**(7):1109-17.

Bauerle, C., Ho, M. N., Lindorfer, M. A., and Stevens, T. H., 1993, The Saccharomyces cerevisiae VMA6 gene encodes the 36-kDa subunit of the vacuolar H(+)-ATPase membrane sector, *J Biol Chem* **268**(17):12749-57.

Beauwens, R., and Al-Awqati, Q., 1976, Active H+ transport in the turtle urinary bladder. Coupling of transport to glucose oxidation, *J Gen Physiol* **68**(4):421-39.

Benetti, P. H., Kim, S. I., Chaillot, D., Canonge, M., Chardot, T., and Meunier, J. C., 1998, Expression and characterization of the recombinant catalytic subunit of casein kinase II from the yeast Yarrowia lipolytica in Escherichia coli, *Protein Expr Purif* **13**(3):283-90.

Bera, T. K., Lee, S., Salvatore, G., Lee, B., and Pastan, I., 2001, MRP8, a new member of ABC transporter superfamily, identified by EST database mining and gene prediction program, is highly expressed in breast cancer, *Mol Med* **7**(8):509-16.

Blair, H. C., Teitelbaum, S. L., Ghiselli, R., and Gluck, S., 1989, Osteoclastic bone resorption by a polarized vacuolar proton pump, *Science* **245**(4920):855-7.

Bowman, B. J., and Bowman, E. J., 2002, Mutations in subunit C of the vacuolar ATPase confer resistance to bafilomycin and identify a conserved antibiotic binding site, *J Biol Chem* **277**(6):3965-72.

Bowman, B. J., Vazquez-Laslop, N., and Bowman, E. J., 1992, The vacuolar ATPase of Neurospora crassa, *J Bioenerg Biomembr* **24**(4):361-70.

Bowman, E. J., Siebers, A., and Altendorf, K., 1988, Bafilomycins: a class of inhibitors of membrane ATPases from microorganisms, animal cells, and plant cells, *Proc Natl Acad Sci U S A* **85**(21):7972-6.

Bowtell, D., Fu, P., Simon, M., and Senior, P., 1992, Identification of murine homologues of the Drosophila son of sevenless gene: potential activators of ras, *Proc Natl Acad Sci U S A* **89**(14):6511-5.

Breton, S., Wiederhold, T., Marshansky, V., Nsumu, N. N., Ramesh, V., and Brown, D., 2000, The B1 subunit of the H+ATPase is a PDZ domain-binding protein. Colocalization with NHE-RF in renal B-intercalated cells, *J Biol Chem* **275**(24):18219-24.

Brown, D., 2000, Targeting of membrane transporters in renal epithelia: when cell biology meets physiology, *Am J Physiol Renal Physiol* **278**(2):F192-201.

Brown, D., and Breton, S., 1996, Mitochondria-rich, proton-secreting epithelial cells, *J Exp Biol* **199**(Pt 11):2345-58.

Brown, D., and Breton, S., 2000, H(+)V-ATPase-dependent luminal acidification in the kidney collecting duct and the epididymis/vas deferens: vesicle recycling and transcytotic pathways, *J Exp Biol* **203 Pt 1**:137-45.

Brown, D., Gluck, S., and Hartwig, J., 1987, Structure of the novel membrane-coating material in proton-secreting epithelial cells and identification as an H+ATPase, *J Cell Biol* **105**(4):1637-48.

Brown, D., Hirsch, S., and Gluck, S., 1988, Localization of a proton-pumping ATPase in rat kidney, *J Clin Invest* **82**(6):2114-26.

Brown, D., Lui, B., Gluck, S., and Sabolic, I., 1992a, A plasma membrane proton ATPase in specialized cells of rat epididymis, *Am J Physiol* **263**(4 Pt 1):C913-6.

Brown, D., Sabolic, I., and Gluck, S., 1992b, Polarized targeting of V-ATPase in kidney epithelial cells, *J Exp Biol* **172**:231-43.

Bruce, L. J., Cope, D. L., Jones, G. K., Schofield, A. E., Burley, M., Povey, S., Unwin, R. J., Wrong, O., and Tanner, M. J., 1997, Familial distal renal tubular acidosis is associated with mutations in the red cell anion exchanger (Band 3, AE1) gene, *J Clin Invest* **100**(7):1693-707.

Cain, B. D., and Simoni, R. D., 1988, Interaction between Glu-219 and His-245 within the a subunit of F1F0-ATPase in Escherichia coli, *J Biol Chem* **263**(14):6606-12.

Chan, Y. L., and Giebisch, G., 1981, Relationship between sodium and bicarbonate transport in the rat proximal convoluted tubule, *Am J Physiol* **240**(3):F222-30.

Chatterjee, D., Chakraborty, M., Leit, M., Neff, L., Jamsa-Kellokumpu, S., Fuchs, R., and Baron, R., 1992, Sensitivity to vanadate and isoforms of subunits A and B distinguish the osteoclast proton pump from other vacuolar H+ ATPases, *Proc Natl Acad Sci U S A* **89**(14):6257-61.

Clague, M. J., Urbe, S., Aniento, F., and Gruenberg, J., 1994, Vacuolar ATPase activity is required for endosomal carrier vesicle formation, *J Biol Chem* **269**(1):21-4.

Clarke, M., Kohler, J., Arana, Q., Liu, T., Heuser, J., and Gerisch, G., 2002, Dynamics of the vacuolar H(+)-ATPase in the contractile vacuole complex and the endosomal pathway of Dictyostelium cells, *J Cell Sci* **115**(Pt 14):2893-905.

Cosson, P., and Bonifacino, J. S., 1992, Role of transmembrane domain interactions in the assembly of class II MHC molecules, *Science* **258**(5082):659-62.

Courey, W. R., and Pfister, R. C., 1972, The radiographic findings in renal tubular acidosis: analysis of 21 cases, *Radiology* **105**(3):497-503.

Coux, O., Tanaka, K., and Goldberg, A. L., 1996, Structure and functions of the 20S and 26S proteasomes, *Annu Rev Biochem* **65**:801-47.

Cross, R. L., and Duncan, T. M., 1996, Subunit rotation in F0F1-ATP synthases as a means of coupling proton transport through F0 to the binding changes in F1, *J Bioenerg Biomembr* **28**(5):403-8.

Derman, A. I., Prinz, W. A., Belin, D., and Beckwith, J., 1993, Mutations that allow disulfide bond formation in the cytoplasm of Escherichia coli, *Science* **262**(5140):1744-7.

Doherty, R. D., and Kane, P. M., 1993, Partial assembly of the yeast vacuolar H(+)-ATPase in mutants lacking one subunit of the enzyme, *J Biol Chem* **268**(22):16845-51.

Domrongkitchaiporn, S., Pongskul, C., Sirikulchayanonta, V., Stitchantrakul, W., Leeprasert, V., Ongphiphadhanakul, B., Radinahamed, P., and Rajatanavin, R., 2002, Bone histology and bone mineral density after correction of acidosis in distal renal tubular acidosis, *Kidney Int* **62**(6):2160-6.

Drees, B. L., 1999, Progress and variations in two-hybrid and three-hybrid technologies, *Curr Opin Chem Biol* **3**(1):64-70.

Drose, S., Bindseil, K. U., Bowman, E. J., Siebers, A., Zeeck, A., and Altendorf, K., 1993, Inhibitory effect of modified bafilomycins and concanamycins on P- and V-type adenosinetriphosphatases, *Biochemistry* **32**(15):3902-6.

Dschida, W. J., and Bowman, B. J., 1992, Structure of the vacuolar ATPase from Neurospora crassa as determined by electron microscopy, *J Biol Chem* **267**(26):18783-9.

Dunaway, G. A., 1983, A review of animal phosphofructokinase isozymes with an emphasis on their physiological role, *Mol Cell Biochem* **52**(1):75-91.

Elson, A., Levanon, D., Brandeis, M., Dafni, N., Bernstein, Y., Danciger, E., and Groner, Y., 1990, The structure of the human liver-type phosphofructokinase gene, *Genomics* **7**(1):47-56.

Fashena, S. J., Serebriiskii, I., and Golemis, E. A., 2000, The continued evolution of two-hybrid screening approaches in yeast: how to outwit different preys with different baits, *Gene* **250**(1-2):1-14.

Feng, Y., and Forgac, M., 1992, A novel mechanism for regulation of vacuolar acidification, *J Biol Chem* **267**(28):19769-72.

Feng, Y., and Forgac, M., 1994, Inhibition of vacuolar H(+)-ATPase by disulfide bond formation between cysteine 254 and cysteine 532 in subunit A, *J Biol Chem* **269**(18):13224-30.

Fields, S., and Song, O., 1989, A novel genetic system to detect protein-protein interactions, *Nature* **340**(6230):245-6.

Fillingame, R. H., 1997, Coupling H+ transport and ATP synthesis in F1F0-ATP synthases: glimpses of interacting parts in a dynamic molecular machine, *J Exp Biol* **200**(Pt 2):217-24.

Finbow, M. E., Eliopoulos, E. E., Jackson, P. J., Keen, J. N., Meagher, L., Thompson, P., Jones, P., and Findlay, J. B., 1992, Structure of a 16 kDa integral membrane protein that has identity to the putative proton channel of the vacuolar H(+)-ATPase, *Protein Eng* **5**(1):7-15.

Finbow, M. E., and Harrison, M. A., 1997, The vacuolar H+-ATPase: a universal proton pump of eukaryotes, *Biochem J* **324**(Pt 3):697-712.

Finbow, M. E., Pitts, J. D., Goldstein, D. J., Schlegel, R., and Findlay, J. B., 1991, The E5 oncoprotein target: a 16-kDa channel-forming protein with diverse functions, *Mol Carcinog* **4**(6):441-4.

Forgac, M., 1989, Structure and function of vacuolar class of ATP-driven proton pumps, *Physiol Rev* **69**(3):765-96.

Forgac, M., 1998, Structure, function and regulation of the vacuolar (H+)-ATPases, *FEBS Lett* **440**(3):258-63.

Forgac, M., 1999, Structure and properties of the vacuolar (H+)-ATPases, *J Biol Chem* **274**(19):12951-4.

Forgac, M., 2000, Structure, mechanism and regulation of the clathrin-coated vesicle and yeast vacuolar H(+)-ATPases, *J Exp Biol* **203 Pt 1**:71-80.

Frattini, A., Orchard, P. J., Sobacchi, C., Giliani, S., Abinun, M., Mattsson, J. P., Keeling, D. J., Andersson, A. K., Wallbrandt, P., Zecca, L., Notarangelo, L. D., Vezzoni, P., and Villa, A., 2000, Defects in TCIRG1 subunit of the vacuolar proton pump are responsible for a subset of human autosomal recessive osteopetrosis, *Nat Genet* **25**(3):343-6.

Futai, M., Noumi, T., and Maeda, M., 1989, ATP synthase (H+-ATPase): results by combined biochemical and molecular biological approaches, *Annu Rev Biochem* **58**:111-36.

Futai, M., Oka, T., Sun-Wada, G., Moriyama, Y., Kanazawa, H., and Wada, Y., 2000, Luminal acidification of diverse organelles by V-ATPase in animal cells, *J Exp Biol* **203 Pt 1**:107-16.

Galli, T., McPherson, P. S., and De Camilli, P., 1996, The V0 sector of the V-ATPase, synaptobrevin, and synaptophysin are associated on synaptic vesicles in a Triton X-100-resistant, freeze-thawing sensitive, complex, *J Biol Chem* **271**(4):2193-8.

Getlawi, F., Laslop, A., Schagger, H., Ludwig, J., Haywood, J., and Apps, D., 1996, Chromaffin granule membrane glycoprotein IV is identical with Ac45, a membrane-integral subunit of the granule's H(+)-ATPase, *Neurosci Lett* **219**(1):13-6.

Geuze, H. J., Slot, J. W., Strous, G. J., Lodish, H. F., and Schwartz, A. L., 1983, Intracellular site of asialoglycoprotein receptor-ligand uncoupling: double-label immunoelectron microscopy during receptor-mediated endocytosis, *Cell* **32**(1):277-87.

Gillespie, J., Ozanne, S., Tugal, B., Percy, J., Warren, M., Haywood, J., and Apps, D., 1991, The vacuolar H(+)-translocating ATPase of renal tubules contains a 115-kDa glycosylated subunit, *FEBS Lett* **282**(1):69-72.

Glickman, J., Croen, K., Kelly, S., and Al-Awqati, Q., 1983, Golgi membranes contain an electrogenic H+ pump in parallel to a chloride conductance, *J Cell Biol* **97**(4):1303-8.

Gluck, S., and Caldwell, J., 1987, Immunoaffinity purification and characterization of vacuolar H+ATPase from bovine kidney, *J Biol Chem* **262**(32):15780-9.

Gluck, S. L., 1992, The structure and biochemistry of the vacuolar H+ ATPase in proximal and distal urinary acidification, *J Bioenerg Biomembr* **24**(4):351-9.

Goldberg, M. E., Expert-Bezancon, N., Vuillard, L., and Rabilloud, T., 1996, Non-detergent sulphobetaines: a new class of molecules that facilitate in vitro protein renaturation, *Fold Des* **1**(1):21-7.

Goldstein, D. J., Andresson, T., Sparkowski, J. J., and Schlegel, R., 1992, The BPV-1 E5 protein, the 16 kDa membrane pore-forming protein and the PDGF receptor exist in a complex that is dependent on hydrophobic transmembrane interactions, *Embo J* **11**(13):4851-9.

Grabe, M., Wang, H., and Oster, G., 2000, The mechanochemistry of V-ATPase proton pumps, *Biophys J* **78**(6):2798-813.

Graf, R., Harvey, W. R., and Wieczorek, H., 1996, Purification and properties of a cytosolic V1-ATPase, *J Biol Chem* **271**(34):20908-13.

Graf, R., Lepier, A., Harvey, W. R., and Wieczorek, H., 1994, A novel 14-kDa V-ATPase subunit in the tobacco hornworm midgut, *J Biol Chem* **269**(5):3767-74.

Graham, L. A., Hill, K. J., and Stevens, T. H., 1998, Assembly of the yeast vacuolar H+-ATPase occurs in the endoplasmic reticulum and requires a Vma12p/Vma22p assembly complex, *J Cell Biol* **142**(1):39-49.

Graham, L. A., Powell, B., and Stevens, T. H., 2000, Composition and assembly of the yeast vacuolar H(+)-ATPase complex, *J Exp Biol* **203 Pt 1**:61-70.

Greenfield, N., and Fasman, G. D., 1969, Computed circular dichroism spectra for the evaluation of protein conformation, *Biochemistry* **8**(10):4108-16.

Guo, Y., Kotova, E., Chen, Z. S., Lee, K., Hopper-Borge, E., Belinsky, M. G., and Kruh, G. D., 2003, MRP8, ATP-binding cassette C11 (ABCC11), is a cyclic nucleotide efflux pump and a resistance factor for fluoropyrimidines 2',3'-dideoxycytidine and 9'-(2'-phosphonylmethoxyethyl)adenine, *J Biol Chem* **278**(32):29509-14.

Harvey, W. R., and Wieczorek, H., 1997, Animal plasma membrane energization by chemiosmotic H+ V-ATPases, *J Exp Biol* **200 ( Pt 2)**:203-16.

Hatch, L. P., Cox, G. B., and Howitt, S. M., 1995, The essential arginine residue at position 210 in the alpha subunit of the Escherichia coli ATP synthase can be transferred to position 252 with partial retention of activity, *J Biol Chem* **270**(49):29407-12.

Heinemann, T., Bulwin, G. C., Randall, J., Schnieders, B., Sandhoff, K., Volk, H. D., Milford, E., Gullans, S. R., and Utku, N., 1999, Genomic organization of the gene coding for TIRC7, a novel membrane protein essential for T cell activation, *Genomics* **57**(3):398-406.

Hill, K., and Cooper, A. A., 2000, Degradation of unassembled Vph1p reveals novel aspects of the yeast ER quality control system, *Embo J* **19**(4):550-61.

Hill, K. J., and Stevens, T. H., 1994, Vma21p is a yeast membrane protein retained in the endoplasmic reticulum by a di-lysine motif and is required for the assembly of the vacuolar H(+)-ATPase complex, *Mol Biol Cell* **5**(9):1039-50.

Hinton, B. T., and Palladino, M. A., 1995, Epididymal epithelium: its contribution to the formation of a luminal fluid microenvironment, *Microsc Res Tech* **30**(1):67-81.

Hirata, R., Graham, L. A., Takatsuki, A., Stevens, T. H., and Anraku, Y., 1997, VMA11 and VMA16 encode second and third proteolipid subunits of the Saccharomyces cerevisiae vacuolar membrane H+-ATPase, *J Biol Chem* **272**(8):4795-803.

Hirata, R., Umemoto, N., Ho, M. N., Ohya, Y., Stevens, T. H., and Anraku, Y., 1993, VMA12 is essential for assembly of the vacuolar H(+)-ATPase subunits onto the vacuolar membrane in Saccharomyces cerevisiae, *J Biol Chem* **268**(2):961-7.

Hirata, T., Iwamoto-Kihara, A., Sun-Wada, G. H., Okajima, T., Wada, Y., and Futai, M., 2003, Subunit rotation of vacuolar-type proton pumping ATPase: relative rotation of the G and C subunits, *J Biol Chem* **278**(26):23714-9.

Holliday, L. S., Lu, M., Lee, B. S., Nelson, R. D., Solivan, S., Zhang, L., and Gluck, S. L., 2000, The amino-terminal domain of the B subunit of vacuolar H+-ATPase contains a filamentous actin binding site, *J Biol Chem* **275**(41):32331-7.

Hong, S. S., and Boulanger, P., 1995, Protein ligands of the human adenovirus type 2 outer capsid identified by biopanning of a phage-displayed peptide library on separate domains of wild-type and mutant penton capsomers, *Embo J* **14**(19):4714-27.

Huang, W., Wang, S. L., Lozano, G., and de Crombrugghe, B., 2001, cDNA library screening using the SOS recruitment system, *Biotechniques* **30**(1):94-8, 100.

Hunt, I. E., and Bowman, B. J., 1997, The intriguing evolution of the "b" and "G" subunits in F-type and V-type ATPases: isolation of the vma-10 gene from Neurospora crassa, *J Bioenerg Biomembr* **29**(6):533-40.

Huss, M., Ingenhorst, G., Konig, S., Gassel, M., Drose, S., Zeeck, A., Altendorf, K., and Wieczorek, H., 2002, Concanamycin A, the specific inhibitor of V-ATPases, binds to the V(o) subunit c, *J Biol Chem* **277**(43):40544-8.

Imamura, H., Nakano, M., Noji, H., Muneyuki, E., Ohkuma, S., Yoshida, M., and Yokoyama, K., 2003, Evidence for rotation of V1-ATPase, *Proc Natl Acad Sci U S A* **100**(5):2312-5.

Inatomi, K., Eya, S., Maeda, M., and Futai, M., 1989, Amino acid sequence of the alpha and beta subunits of Methanosarcina barkeri ATPase deduced from cloned genes. Similarity to subunits of eukaryotic vacuolar and F0F1-ATPases, *J Biol Chem* **264**(19):10954-9.

Iwamoto, A., Orita-Saita, Y., Maeda, M., and Futai, M., 1994, N-ethylmaleimide-sensitive mutant (beta Val-153-->Cys) Escherichia coli F1-ATPase: cross-linking of the mutant beta subunit with the alpha subunit, *FEBS Lett* **352**(2):243-6.

Jackson, D. D., and Stevens, T. H., 1997, VMA12 encodes a yeast endoplasmic reticulum protein required for vacuolar H+-ATPase assembly, *J Biol Chem* **272**(41):25928-34.

Jarolim, P., Shayakul, C., Prabakaran, D., Jiang, L., Stuart-Tilley, A., Rubin, H. L., Simova, S., Zavadil, J., Herrin, J. T., Brouillette, J., Somers, M. J., Seemanova, E., Brugnara, C., Guay-Woodford, L. M., and Alper, S. L., 1998, Autosomal dominant distal renal tubular acidosis is associated in three families with heterozygosity for the R589H mutation in the AE1 (band 3) Cl-/HCO3- exchanger, *J Biol Chem* **273**(11):6380-8.

Jensen, L. J., Schmitt, B. M., Berger, U. V., Nsumu, N. N., Boron, W. F., Hediger, M. A., Brown, D., and Breton, S., 1999, Localization of sodium bicarbonate cotransporter (NBC) protein and messenger ribonucleic acid in rat epididymis, *Biol Reprod* **60**(3):573-9.

Jensen, L. J., Sorensen, J. N., Larsen, E. H., and Willumsen, N. J., 1997, Proton pump activity of mitochondria-rich cells. The interpretation of external proton-concentration gradients, *J Gen Physiol* **109**(1):73-91.

Jiang, W., and Fillingame, R. H., 1998, Interacting helical faces of subunits a and c in the F1Fo ATP synthase of Escherichia coli defined by disulfide cross-linking, *Proc Natl Acad Sci U S A* **95**(12):6607-12.

Johnson, W. C., Jr., 1990, Protein secondary structure and circular dichroism: a practical guide, *Proteins* **7**(3):205-14.

Jonsson, U., Fagerstam, L., Ivarsson, B., Johnsson, B., Karlsson, R., Lundh, K., Lofas, S., Persson, B., Roos, H., Ronnberg, I., and et al., 1991, Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology, *Biotechniques* **11**(5):620-7.

Kane, P. M., 1995, Disassembly and reassembly of the yeast vacuolar H(+)-ATPase in vivo, *J Biol Chem* **270**(28):17025-32.

Kane, P. M., Kuehn, M. C., Howald-Stevenson, I., and Stevens, T. H., 1992, Assembly and targeting of peripheral and integral membrane subunits of the yeast vacuolar H(+)-ATPase, *J Biol Chem* **267**(1):447-54.

Karet, F. E., Finberg, K. E., Nayir, A., Bakkaloglu, A., Ozen, S., Hulton, S. A., Sanjad, S. A., Al-Sabban, E. A., Medina, J. F., and Lifton, R. P., 1999a, Localization of a gene for autosomal recessive distal renal tubular acidosis with normal hearing (rdRTA2) to 7q33-34, *Am J Hum Genet* **65**(6):1656-65.

Karet, F. E., Finberg, K. E., Nelson, R. D., Nayir, A., Mocan, H., Sanjad, S. A., Rodriguez-Soriano, J., Santos, F., Cremers, C. W., Di Pietro, A., Hoffbrand, B. I., Winiarski, J., Bakkalogiu, A., Ozen, S., Dusunsel, R., Goodyer, P., Hulton, S. A., Wu, D. K., Skvorak, A. B., Morton, C. C., Cunningham, M. J., Jha, V., and Lifton, R. P., 1999b, Mutations in the gene encoding B1 subunit of H+-ATPase cause renal tubular acidosis with sensorineural deafness, *Nat Genet* **21**(1):84-90.

Karet, F. E., Gainza, F. J., Gyory, A. Z., Unwin, R. J., Wrong, O., Tanner, M. J., Nayir, A., Alpay, H., Santos, F., Hulton, S. A., Bakkaloglu, A., Ozen, S., Cunningham, M. J., di Pietro, A., Walker, W. G., and Lifton, R. P., 1998, Mutations in the chloride-bicarbonate exchanger gene AE1 cause autosomal dominant but not autosomal recessive distal renal tubular acidosis, *Proc Natl Acad Sci U S A* **95**(11):6337-42.

Kawasaki-Nishi, S., Bowers, K., Nishi, T., Forgac, M., and Stevens, T. H., 2001a, The amino-terminal domain of the vacuolar proton-translocating ATPase a subunit controls targeting and in vivo dissociation, and the carboxyl- terminal domain affects coupling of proton transport and ATP hydrolysis, *J Biol Chem* **276**(50):47411-20.

Kawasaki-Nishi, S., Nishi, T., and Forgac, M., 2001b, Arg-735 of the 100-kDa subunit a of the yeast V-ATPase is essential for proton translocation, *Proc Natl Acad Sci U S A* **98**(22):12397-402.

Kawasaki-Nishi, S., Nishi, T., and Forgac, M., 2001c, Yeast V-ATPase complexes containing different isoforms of the 100-kDa a- subunit differ in coupling efficiency and in vivo dissociation, *J Biol Chem* **276**(21):17941-8.

Kawasaki-Nishi, S., Nishi, T., and Forgac, M., 2003a, Interacting helical surfaces of the transmembrane segments of subunits a and c' of the yeast V-ATPase defined by disulfide-mediated cross-linking, *J Biol Chem* **278**(43):41908-13.

Kawasaki-Nishi, S., Nishi, T., and Forgac, M., 2003b, Proton translocation driven by ATP hydrolysis in V-ATPases, *FEBS Lett* **545**(1):76-85.

Kemp, R. G., Fox, R. W., and Latshaw, S. P., 1987, Amino acid sequence at the citrate allosteric site of rabbit muscle phosphofructokinase, *Biochemistry* **26**(12):3443-6.

Kemp, R. G., and Gunasekera, D., 2002, Evolution of the allosteric ligand sites of mammalian phosphofructo-1-kinase, *Biochemistry* **41**(30):9426-30.

Kim, J., Tisher, C. C., Linser, P. J., and Madsen, K. M., 1990, Ultrastructural localization of carbonic anhydrase II in subpopulations of intercalated cells of the rat kidney, *J Am Soc Nephrol* **1**(3):245-56.

Kirshner, N., Kirshner, A. G., and Kamin, D. L., 1966, Adenosine triphosphatase activity of adrenal medulla catecholamine granules, *Biochim Biophys Acta* **113**(2):332-5.

Kornak, U., Schulz, A., Friedrich, W., Uhlhaas, S., Kremens, B., Voit, T., Hasan, C., Bode, U., Jentsch, T. J., and Kubisch, C., 2000, Mutations in the a3 subunit of the vacuolar H(+)-ATPase cause infantile malignant osteopetrosis, *Hum Mol Genet* **9**(13):2059-63.

Kornfeld, S., 1992, Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors, *Annu Rev Biochem* **61**:307-30.

Kurtzman, N. A., 1990, Disorders of distal acidification, *Kidney Int* **38**(4):720-7.

Laemmli, U. K., 1970, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* **227**(259):680-5.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray,

A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al., 2001, Initial sequencing and analysis of the human genome, *Nature* **409**(6822):860-921.

Landolt-Marticorena, C., Williams, K. M., Correa, J., Chen, W., and Manolson, M. F., 2000, Evidence that the NH2 terminus of vph1p, an integral subunit of the V0 sector of the yeast V-ATPase, interacts directly with the Vma1p and Vma13p subunits of the V1 sector, *J Biol Chem* **275**(20):15449-57.

Lee, B. S., Gluck, S. L., and Holliday, L. S., 1999, Interaction between vacuolar H(+)-ATPase and microfilaments during osteoclast activation, *J Biol Chem* **274**(41):29164-71.

Lee, C., Ghoshal, K., and Beaman, K. D., 1990, Cloning of a cDNA for a T cell produced molecule with a putative immune regulatory role, *Mol Immunol* **27**(11):1137-44.

Lemmon, M. A., and Engelman, D. M., 1994, Specificity and promiscuity in membrane helix interactions, *FEBS Lett* **346**(1):17-20.

Leng, X. H., Manolson, M. F., and Forgac, M., 1998, Function of the COOH-terminal domain of Vph1p in activity and assembly of the yeast V-ATPase, *J Biol Chem* **273**(12):6717-23.

Leng, X. H., Manolson, M. F., Liu, Q., and Forgac, M., 1996, Site-directed mutagenesis of the 100-kDa subunit (Vph1p) of the yeast vacuolar (H+)-ATPase, *J Biol Chem* **271**(37):22487-93.

Leng, X. H., Nishi, T., and Forgac, M., 1999, Transmembrane topography of the 100-kDa a subunit (Vph1p) of the yeast vacuolar proton-translocating ATPase, *J Biol Chem* **274**(21):14655-61.

Li, Y., Rivera, D., Ru, W., Gunasekera, D., and Kemp, R. G., 1999, Identification of allosteric sites in rabbit phosphofructo-1-kinase, *Biochemistry* **38**(49):16407-12.

Li, Y. P., Chen, W., and Stashenko, P., 1996, Molecular cloning and characterization of a putative novel human osteoclast-specific 116-kDa vacuolar proton pump subunit, *Biochem Biophys Res Commun* **218**(3):813-21.

Liu, Q., Kane, P. M., Newman, P. R., and Forgac, M., 1996, Site-directed mutagenesis of the yeast V-ATPase B subunit (Vma2p), *J Biol Chem* **271**(4):2018-22.

Liu, Q., Leng, X. H., Newman, P. R., Vasilyeva, E., Kane, P. M., and Forgac, M., 1997, Site-directed mutagenesis of the yeast V-ATPase A subunit, *J Biol Chem* **272**(18):11750-6.

Lu, M., Holliday, L. S., Zhang, L., Dunn, W. A., Jr., and Gluck, S. L., 2001, Interaction between aldolase and vacuolar H+-ATPase: evidence for direct coupling of glycolysis to the ATP-hydrolyzing proton pump, *J Biol Chem* **276**(32):30407-13.

Lu, M., Sautin, Y. Y., Holliday, L. S., and Gluck, S. L., 2003, The glycolytic enzyme aldolase mediates assembly, expression and activity of V-ATPase, *J Biol Chem.*

Lu, M., Vergara, S., Zhang, L., Holliday, L. S., Aris, J., and Gluck, S. L., 2002, The amino-terminal domain of the E subunit of V-ATPase interacts with the H subunit and is required for V-ATPase function, *J Biol Chem* **5**:5.

Lu, X., Yu, H., Liu, S. H., Brodsky, F. M., and Peterlin, B. M., 1998, Interactions between HIV1 Nef and vacuolar ATPase facilitate the internalization of CD4, *Immunity* **8**(5):647-56.

Ludwig, J., Kerscher, S., Brandt, U., Pfeiffer, K., Getlawi, F., Apps, D. K., and Schagger, H., 1998, Identification and characterization of a novel 9.2-kDa membrane sector-associated protein of vacuolar proton-ATPase from chromaffin granules, *J Biol Chem* **273**(18):10939-47.

Ma, L., and Center, M. S., 1992, The gene encoding vacuolar H(+)-ATPase subunit C is overexpressed in multidrug-resistant HL60 cells, *Biochem Biophys Res Commun* **182**(2):675-81.

Maiorano, D., Moreau, J., and Mechali, M., 2000, XCDT1 is required for the assembly of pre-replicative complexes in Xenopus laevis, *Nature* **404**(6778):622-5.

Manolson, M. F., Proteau, D., Preston, R. A., Stenbit, A., Roberts, B. T., Hoyt, M. A., Preuss, D., Mulholland, J., Botstein, D., and Jones, E. W., 1992, The VPH1 gene encodes a 95-kDa integral membrane polypeptide required for in vivo assembly and activity of the yeast vacuolar H(+)-ATPase, *J Biol Chem* **267**(20):14294-303.

Manolson, M. F., Wu, B., Proteau, D., Taillon, B. E., Roberts, B. T., Hoyt, M. A., and Jones, E. W., 1994, STV1 gene encodes functional homologue of 95-kDa yeast vacuolar H(+)- ATPase subunit Vph1p, *J Biol Chem* **269**(19):14064-74.

Martinez-Zaguilan, R., Lynch, R. M., Martinez, G. M., and Gillies, R. J., 1993, Vacuolar-type H(+)-ATPases are functionally expressed in plasma membranes of human tumor cells, *Am J Physiol* **265**(4 Pt 1):C1015-29.

Masino, L., Kelly, G., Leonard, K., Trottier, Y., and Pastore, A., 2002, Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins, *FEBS Lett* **513**(2-3):267-72.

Mattsson, J. P., Li, X., Peng, S. B., Nilsson, F., Andersen, P., Lundberg, L. G., Stone, D. K., and Keeling, D. J., 2000, Properties of three isoforms of the 116-kDa subunit of

vacuolar H+-ATPase from a single vertebrate species. Cloning, gene expression and protein characterization of functionally distinct isoforms in Gallus gallus, *Eur J Biochem* **267**(13):4115-26.

Miroux, B., and Walker, J. E., 1996, Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels, *J Mol Biol* **260**(3):289-98.

Miura, K., Miyazawa, S., Furuta, S., Mitsushita, J., Kamijo, K., Ishida, H., Miki, T., Suzukawa, K., Resau, J., Copeland, T. D., and Kamata, T., 2001, The Sos1-Rac1 signaling. Possible involvement of a vacuolar H(+)-ATPase E subunit, *J Biol Chem* **276**(49):46276-83.

Moriyama, Y., Maeda, M., and Futai, M., 1992, The role of V-ATPase in neuronal and endocrine systems, *J Exp Biol* **172**:171-8.

Moriyama, Y., Yamamoto, A., Tashiro, Y., and Futai, M., 1991, Chromaffin granule H(+)-ATPase has F1-like structure, *FEBS Lett* **291**(1):92-6.

Mulberg, A. E., Tulk, B. M., and Forgac, M., 1991, Modulation of coated vesicle chloride channel activity and acidification by reversible protein kinase A-dependent phosphorylation, *J Biol Chem* **266**(31):20590-3.

Munn, A. L., and Riezman, H., 1994, Endocytosis is required for the growth of vacuolar H(+)-ATPase-defective yeast: identification of six new END genes, *J Cell Biol* **127**(2):373-86.

Murata, Y., Sun-Wada, G. H., Yoshimizu, T., Yamamoto, A., Wada, Y., and Futai, M., 2002, Differential localization of the vacuolar H+ pump with G subunit isoforms (G1 and G2) in mouse neurons, *J Biol Chem* **277**(39):36296-303.

Myers, M., and Forgac, M., 1993, Assembly of the peripheral domain of the bovine vacuolar H(+)-adenosine triphosphatase, *J Cell Physiol* **156**(1):35-42.

Nakajima, H., Kono, N., Yamasaki, T., Hamaguchi, T., Hotta, K., Kuwajima, M., Noguchi, T., Tanaka, T., and Tarui, S., 1990, Tissue specificity in expression and alternative RNA splicing of human phosphofructokinase-M and -L genes, *Biochem Biophys Res Commun* **173**(3):1317-21.

Nakamura, I., Takahashi, N., Udagawa, N., Moriyama, Y., Kurokawa, T., Jimi, E., Sasaki, T., and Suda, T., 1997, Lack of vacuolar proton ATPase association with the cytoskeleton in osteoclasts of osteosclerotic (oc/oc) mice, *FEBS Lett* **401**(2-3):207-12.

Nelson, H., Mandiyan, S., and Nelson, N., 1994, The Saccharomyces cerevisiae VMA7 gene encodes a 14-kDa subunit of the vacuolar H(+)-ATPase catalytic sector, *J Biol Chem* **269**(39):24150-5.

Nelson, H., Mandiyan, S., and Nelson, N., 1995, A bovine cDNA and a yeast gene (VMA8) encoding the subunit D of the vacuolar H(+)-ATPase, *Proc Natl Acad Sci U S A* **92**(2):497-501.

Nelson, H., and Nelson, N., 1989, The progenitor of ATP synthases was closely related to the current vacuolar H+-ATPase, *FEBS Lett* **247**(1):147-53.

Nelson, H., and Nelson, N., 1990, Disruption of genes encoding subunits of yeast vacuolar H(+)-ATPase causes conditional lethality, *Proc Natl Acad Sci U S A* **87**(9):3503-7.

Nelson, N., 1989, Structure, molecular genetics, and evolution of vacuolar H+-ATPases, *J Bioenerg Biomembr* **21**(5):553-71.

Nelson, N., 1992a, Evolution of organellar proton-ATPases, *Biochim Biophys Acta* **1100**(2):109-24.

Nelson, N., 1992b, Structural conservation and functional diversity of V-ATPases, *J Bioenerg Biomembr* **24**(4):407-14.

Nelson, N., and Harvey, W. R., 1999, Vacuolar and plasma membrane proton-adenosinetriphosphatases, *Physiol Rev* **79**(2):361-85.

Nelson, N., and Klionsky, D. J., 1996, Vacuolar H(+)-ATPase: from mammals to yeast and back, *Experientia* **52**(12):1101-10.

Nelson, R. D., Guo, X. L., Masood, K., Brown, D., Kalkbrenner, M., and Gluck, S., 1992, Selectively amplified expression of an isoform of the vacuolar H(+)- ATPase 56-kilodalton subunit in renal intercalated cells, *Proc Natl Acad Sci U S A* **89**(8):3541-5.

Nishi, T., and Forgac, M., 2000, Molecular cloning and expression of three isoforms of the 100-kDa a subunit of the mouse vacuolar proton-translocating ATPase, *J Biol Chem* **275**(10):6824-30.

Nishi, T., and Forgac, M., 2002, The vacuolar (H+)-ATPases--nature's most versatile proton pumps, *Nat Rev Mol Cell Biol* **3**(2):94-103.

Njus, D., and Radda, G. K., 1978, Bioenergetic processes in chromaffin granules a new perspective on some old problems, *Biochim Biophys Acta* **463**(3-4):219-44.

Noji, H., Yasuda, R., Yoshida, M., and Kinosita, K., Jr., 1997, Direct observation of the rotation of F1-ATPase, *Nature* **386**(6622):299-302.

Noumi, T., Beltran, C., Nelson, H., and Nelson, N., 1991, Mutational analysis of yeast vacuolar H(+)-ATPase, *Proc Natl Acad Sci U S A* **88**(5):1938-42.

O'Callaghan, C., and Brenner, B., 2000, The kidney at a glance, Blackwell Science.

Ochem, A. E., Skopac, D., Costa, M., Rabilloud, T., Vuillard, L., Simoncsits, A., Giacca, M., and Falaschi, A., 1997, Functional properties of the separate subunits of human DNA helicase II/Ku autoantigen, *J Biol Chem* **272**(47):29919-26.

Oka, T., Murata, Y., Namba, M., Yoshimizu, T., Toyomura, T., Yamamoto, A., Sun-Wada, G. H., Hamasaki, N., Wada, Y., and Futai, M., 2001a, a4, a unique kidney-specific isoform of mouse vacuolar H+-ATPase subunit a, *J Biol Chem* **276**(43):40050-4.

Oka, T., Toyomura, T., Honjo, K., Wada, Y., and Futai, M., 2001b, Four subunit a isoforms of Caenorhabditis elegans Vacuolar H+-ATPase: Cell-specific expression during development, *J Biol Chem* **276**:33079-33085.

Omote, H., and Futai, M., 1998, Mutational analysis of F1F0 ATPase: catalysis and energy coupling, *Acta Physiol Scand Suppl* **643**:177-83.

Parra, K. J., and Kane, P. M., 1998, Reversible association between the V1 and V0 domains of yeast vacuolar H+-ATPase is an unconventional glucose-induced effect, *Mol Cell Biol* **18**(12):7064-74.

Peng, S. B., Li, X., Crider, B. P., Zhou, Z., Andersen, P., Tsai, S. J., Xie, X. S., and Stone, D. K., 1999, Identification and reconstitution of an isoform of the 116-kDa subunit of the vacuolar proton translocating ATPase, *J Biol Chem* **274**(4):2549-55.

Penney, M. D., and Oleesky, D. A., 1999, Renal tubular acidosis, *Ann Clin Biochem* **36 ( Pt 4)**:408-22.

Perczel, A., Hollosi, M., Tusnady, G., and Fasman, G. D., 1991, Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins, *Protein Eng* **4**(6):669-79.

Perczel, A., Park, K., and Fasman, G. D., 1992a, Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: a practical guide, *Anal Biochem* **203**(1):83-93.

Perczel, A., Park, K., and Fasman, G. D., 1992b, Deconvolution of the circular dichroism spectra of proteins: the circular dichroism spectra of the antiparallel beta-sheet in proteins, *Proteins* **13**(1):57-69.

Perin, M. S., Fried, V. A., Stone, D. K., Xie, X. S., and Sudhof, T. C., 1991, Structure of the 116-kDa polypeptide of the clathrin-coated vesicle/synaptic vesicle proton pump, *J Biol Chem* **266**(6):3877-81.

Petitjean, A., Hilger, F., and Tatchell, K., 1990, Comparison of thermosensitive alleles of the CDC25 gene involved in the cAMP metabolism of Saccharomyces cerevisiae, *Genetics* **124**(4):797-806.

Possee, R. D., 1997, Baculoviruses as expression vectors, *Curr Opin Biotechnol* **8**(5):569-72.

Powell, B., Graham, L. A., and Stevens, T. H., 2000, Molecular characterization of the yeast vacuolar H+-ATPase proton pore, *J Biol Chem* **275**(31):23654-60.

Pujol, N., Bonnerot, C., Ewbank, J. J., Kohara, Y., and Thierry-Mieg, D., 2001, The Caenorhabditis elegans unc-32 gene encodes alternative forms of a vacuolar ATPase a subunit, *J Biol Chem* **276**(15):11913-21.

Puopolo, K., Kumamoto, C., Adachi, I., Magner, R., and Forgac, M., 1992a, Differential expression of the "B" subunit of the vacuolar H(+)-ATPase in bovine tissues, *J Biol Chem* **267**(6):3696-706.

Puopolo, K., Sczekan, M., Magner, R., and Forgac, M., 1992b, The 40-kDa subunit enhances but is not required for activity of the coated vesicle proton pump, *J Biol Chem* **267**(8):5171-6.

Ren, Z., and Black, L. W., 1998, Phage T4 SOC and HOC display of biologically active, full-length proteins on the viral capsid, *Gene* **215**(2):439-44.

Rivett, A. J., Palmer, A., and Knecht, E., 1992, Electron microscopic localization of the multicatalytic proteinase complex in rat liver and in cultured cells, *J Histochem Cytochem* **40**(8):1165-72.

Sabbert, D., Engelbrecht, S., and Junge, W., 1996, Intersubunit rotation in active F-ATPase, *Nature* **381**(6583):623-5.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., and Arnheim, N., 1985, Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia, *Science* **230**(4732):1350-4.

Sambade, M., and Kane, P. M., 2004, The yeast V-ATPase contains a subunit homologous to the M. sexta and bovine e subunits that is essential for function, *J Biol Chem*.

Santini, C., Brennan, D., Mennuni, C., Hoess, R. H., Nicosia, A., Cortese, R., and Luzzago, A., 1998, Efficient display of an HCV cDNA expression library as C-terminal fusion to the capsid protein D of bacteriophage lambda, *J Mol Biol* **282**(1):125-35.

Sato, S. B., and Toyama, S., 1994, Interference with the endosomal acidification by a monoclonal antibody directed toward the 116 (100)-kD subunit of the vacuolar type proton pump, *J Cell Biol* **127**(1):39-53.

Scherrer, K., and Bey, F., 1994, The prosomes (multicatalytic proteinases; proteasomes) and their relationship to the untranslated messenger ribonucleoproteins, the cytoskeleton, and cell differentiation, *Prog Nucleic Acid Res Mol Biol* **49**:1-64.

Schneider, I., 1972, Cell lines derived from late embryonic stages of Drosophila melanogaster, *J Embryol Exp Morphol* **27**(2):353-65.

Schuster, V. L., 1993, Function and regulation of collecting duct intercalated cells, *Annu Rev Physiol* **55**:267-88.

Schwartz, G. J., Barasch, J., and Al-Awqati, Q., 1985, Plasticity of functional epithelial polarity, *Nature* **318**(6044):368-71.

Scott, J. K., and Smith, G. P., 1990, Searching for peptide ligands with an epitope library, *Science* **249**(4967):386-90.

Shaw, A. Z., and Miroux, B., 2003, A general approach for heterologous membrane protein expression in Escherichia coli: the uncoupling protein, UCP1, as an example, *Methods Mol Biol* **228**:23-35.

Sidhu, S. S., Lowman, H. B., Cunningham, B. C., and Wells, J. A., 2000, Phage display for selection of novel binding peptides, *Methods Enzymol* **328**:333-63.

Silver, R. B., and Soleimani, M., 1999, H+-K+-ATPases: regulation and role in pathophysiological states, *Am J Physiol* **276**(6 Pt 2):F799-811.

Skinner, M. A., and Wildeman, A. G., 1999, beta(1) integrin binds the 16-kDa subunit of vacuolar H(+)-ATPase at a site important for human papillomavirus E5 and platelet-derived growth factor signaling, *J Biol Chem* **274**(33):23119-27.

Smith, A. N., Borthwick, K. J., and Karet, F. E., 2002, Molecular cloning and characterization of novel tissue-specific isoforms of the human vacuolar H(+)-ATPase C, G and d subunits, and their evaluation in autosomal recessive distal renal tubular acidosis, *Gene* **297**(1-2):169-77.

Smith, A. N., Finberg, K. E., Wagner, C. A., Lifton, R. P., Devonald, M. A., Su, Y., and Karet, F. E., 2001, Molecular cloning and characterization of Atp6n1b: a novel fourth murine vacuolar H+-ATPase a-subunit gene, *J Biol Chem* **276**(45):42382-8.

Smith, A. N., Lovering, R. C., Futai, M., Takeda, J., Brown, D., and Karet, F. E., 2003, Revised nomenclature for mammalian vacuolar-type H+ -ATPase subunit genes, *Mol Cell* **12**(4):801-3.

Smith, A. N., Skaug, J., Choate, K. A., Nayir, A., Bakkaloglu, A., Ozen, S., Hulton, S. A., Sanjad, S. A., Al-Sabban, E. A., Lifton, R. P., Scherer, S. W., and Karet, F. E., 2000, Mutations in ATP6N1B, encoding a new kidney vacuolar proton pump 116-kD subunit, cause recessive distal renal tubular acidosis with preserved hearing, *Nat Genet* **26**(1):71-5.

Smith, D. B., and Johnson, K. S., 1988, Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase, *Gene* **67**(1):31-40.

Smith, G. P., 1985, Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface, *Science* **228**(4705):1315-7.

Stalz, W. D., Greie, J. C., Deckers-Hebestreit, G., and Altendorf, K., 2003, Direct interaction of subunits a and b of the F0 complex of Escherichia coli ATP synthase by forming an ab2 subcomplex, *J Biol Chem* **278**(29):27068-71.

Stehberger, P. A., Schulz, N., Finberg, K. E., Karet, F. E., Giebisch, G., Lifton, R. P., Geibel, J. P., and Wagner, C. A., 2003, Localization and regulation of the ATP6V0A4 (a4) vacuolar H+-ATPase subunit defective in an inherited form of distal renal tubular acidosis, *J Am Soc Nephrol* **14**(12):3027-38.

Steinmetz, P. R., Husted, R. F., Mueller, A., and Beauwens, R., 1981, Coupling between H+ transport and anaerobic glycolysis in turtle urinary bladder: effect of inhibitors of H+ ATPase, *J Membr Biol* **59**(1):27-34.

Stevens, T. H., and Forgac, M., 1997, Structure, function and regulation of the vacuolar (H+)-ATPase, *Annu Rev Cell Dev Biol* **13**:779-808.

Stover, E. H., Borthwick, K. J., Bavalia, C., Eady, N., Fritz, D. M., Rungroj, N., Giersch, A. B., Morton, C. C., Axon, P. R., Akil, I., Al-Sabban, E. A., Baguley, D. M., Bianca, S., Bakkaloglu, A., Bircan, Z., Chauveau, D., Clermont, M. J., Guala, A., Hulton, S. A., Kroes, H., Li Volti, G., Mir, S., Mocan, H., Nayir, A., Ozen, S., Rodriguez Soriano, J., Sanjad, S. A., Tasic, V., Taylor, C. M., Topaloglu, R., Smith, A. N., and Karet, F. E., 2002, Novel ATP6V1B1 and ATP6V0A4 mutations in autosomal recessive distal renal tubular acidosis with new evidence for hearing loss, *J Med Genet* **39**(11):796-803.

Sumner, J. P., Dow, J. A., Earley, F. G., Klein, U., Jager, D., and Wieczorek, H., 1995, Regulation of plasma membrane V-ATPase activity by dissociation of peripheral subunits, *J Biol Chem* **270**(10):5649-53.

Sun-Wada, G. H., Imai-Senga, Y., Yamamoto, A., Murata, Y., Hirata, T., Wada, Y., and Futai, M., 2002, A proton pump ATPase with testis-specific E1 subunit isoform required for acrosome acidification, *J Biol Chem* **in press**.

Sun-Wada, G. H., Yoshimizu, T., Imai-Senga, Y., Wada, Y., and Futai, M., 2003, Diversity of mouse proton-translocating ATPase: presence of multiple isoforms of the C, d and G subunits, *Gene* **302**(1-2):147-53.

Supek, F., Supekova, L., Mandiyan, S., Pan, Y. C., Nelson, H., and Nelson, N., 1994, A novel accessory subunit for vacuolar H(+)-ATPase from chromaffin granules, *J Biol Chem* **269**(39):24102-6.

Sze, H., Schumacher, K., Muller, M. L., Padmanaban, S., and Taiz, L., 2002, A simple nomenclature for a complex proton pump: VHA genes encode the vacuolar H(+)-ATPase, *Trends Plant Sci* **7**(4):157-61.

Tammur, J., Prades, C., Arnould, I., Rzhetsky, A., Hutchinson, A., Adachi, M., Schuetz, J. D., Swoboda, K. J., Ptacek, L. J., Rosier, M., Dean, M., and Allikmets, R., 2001, Two new genes from the human ATP-binding cassette transporter superfamily, ABCC11 and ABCC12, tandemly duplicated on chromosome 16q12, *Gene* **273**(1):89-96.

Tanphaichitr, V. S., Sumboonnanonda, A., Ideguchi, H., Shayakul, C., Brugnara, C., Takao, M., Veerakul, G., and Alper, S. L., 1998, Novel AE1 mutations in recessive distal renal tubular acidosis. Loss-of- function is rescued by glycophorin A, *J Clin Invest* **102**(12):2173-9.

Tisher, C. C., Madsen, K. M., and Verlander, J. W., 1991, Structural adaptation of the collecting duct to acid-base disturbances, *Contrib Nephrol* **95**:168-77.

Tomashek, J. J., Graham, L. A., Hutchins, M. U., Stevens, T. H., and Klionsky, D. J., 1997, V1-situated stalk subunits of the yeast vacuolar proton-translocating ATPase, *J Biol Chem* **272**(42):26787-93.

Towbin, H., Staehelin, T., and Gordon, J., 1979, Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications, *Proc Natl Acad Sci U S A* **76**(9):4350-4.

Toyomura, T., Oka, T., Yamaguchi, C., Wada, Y., and Futai, M., 2000, Three subunit a isoforms of mouse vacuolar H(+)-ATPase. Preferential expression of the a3 isoform during osteoclast differentiation, *J Biol Chem* **275**(12):8760-5.

Urbanowski, J. L., and Piper, R. C., 1999, The iron transporter Fth1p forms a complex with the Fet5 iron oxidase and resides on the vacuolar membrane, *J Biol Chem* **274**(53):38061-70.

Uyeda, K., 1979, Phosphofructokinase, *Adv Enzymol Relat Areas Mol Biol* **48**:193-244.

Valiyaveetil, F. I., and Fillingame, R. H., 1997, On the role of Arg-210 and Glu-219 of subunit a in proton translocation by the Escherichia coli F0F1-ATP synthase, *J Biol Chem* **272**(51):32635-41.

Valiyaveetil, F. I., and Fillingame, R. H., 1998, Transmembrane topography of subunit a in the Escherichia coli F1F0 ATP synthase, *J Biol Chem* **273**(26):16241-7.

van der Merwe, P. A., and Barclay, A. N., 1994, Transient intercellular adhesion: the importance of weak protein-protein interactions, *Trends Biochem Sci* **19**(9):354-8.

van der Merwe, P. A., and Barclay, A. N., 1996, Analysis of cell-adhesion molecule interactions using surface plasmon resonance, *Curr Opin Immunol* **8**(2):257-61.

van Weert, A. W., Dunn, K. W., Gueze, H. J., Maxfield, F. R., and Stoorvogel, W., 1995, Transport from late endosomes to lysosomes, but not sorting of integral membrane proteins in endosomes, depends on the vacuolar proton pump, *J Cell Biol* **130**(4):821-34.

Vasilyeva, E., and Forgac, M., 1996, 3'-O-(4-Benzoyl)benzoyladenosine 5'-triphosphate inhibits activity of the vacuolar (H+)-ATPase from bovine brain clathrin-coated vesicles by modification of a rapidly exchangeable, noncatalytic nucleotide binding site on the B subunit, *J Biol Chem* **271**(22):12775-82.

Vasilyeva, E., and Forgac, M., 1998, Interaction of the clathrin-coated vesicle V-ATPase with ADP and sodium azide, *J Biol Chem* **273**(37):23823-9.

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., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L.,

Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., et al., 2001, The sequence of the human genome, *Science* **291**(5507):1304-51.

Vidal, M., and Legrain, P., 1999, Yeast forward and reverse 'n'-hybrid systems, *Nucleic Acids Res* **27**(4):919-29.

Vik, S. B., and Antonio, B. J., 1994, A mechanism of proton translocation by F1F0 ATP synthases suggested by double mutants of the a subunit, *J Biol Chem* **269**(48):30364-9.

Vora, S., 1982, Isozymes of phosphofructokinase, *Isozymes Curr Top Biol Med Res* **6**:119-67.

Vuillard, L., Braun-Breton, C., and Rabilloud, T., 1995, Non-detergent sulphobetaines: a new class of mild solubilization agents for protein purification, *Biochem J* **305 ( Pt 1)**:337-43.

Walhout, A. J., and Vidal, M., 2001, Protein interaction maps for model organisms, *Nat Rev Mol Cell Biol* **2**(1):55-62.

Wang, S. Y., Moriyama, Y., Mandel, M., Hulmes, J. D., Pan, Y. C., Danho, W., Nelson, H., and Nelson, N., 1988, Cloning of cDNA encoding a 32-kDa protein. An accessory polypeptide of the H+-ATPase from chromaffin granules, *J Biol Chem* **263**(33):17638-42.

Weber, J., and Senior, A. E., 1997, Catalytic mechanism of F1-ATPase, *Biochim Biophys Acta* **1319**(1):19-58.

White, J. M., 1992, Membrane fusion, *Science* **258**(5084):917-24.

Wieczorek, H., 1992, The insect V-ATPase, a plasma membrane proton pump energizing secondary active transport: molecular analysis of electrogenic potassium transport in the tobacco hornworm midgut, *J Exp Biol* **172**:335-43.

Wieczorek, H., Grber, G., Harvey, W. R., Huss, M., Merzendorfer, H., and Zeiske, W., 2000, Structure and regulation of insect plasma membrane H(+)V-ATPase, *J Exp Biol* **203 Pt 1**:127-35.

Wieczorek, H., Putzenlechner, M., Zeiske, W., and Klein, U., 1991, A vacuolar-type proton pump energizes K+/H+ antiport in an animal plasma membrane, *J Biol Chem* **266**(23):15340-7.

Wilkens, S., and Forgac, M., 2001, Three-dimensional structure of the vacuolar ATPase proton channel by electron microscopy, *J Biol Chem* **276**(47):44064-8.

Wilkens, S., Vasilyeva, E., and Forgac, M., 1999, Structure of the vacuolar ATPase by electron microscopy, *J Biol Chem* **274**(45):31804-10.

Williams, R. W., Chang, A., Juretic, D., and Loughran, S., 1987, Secondary structure predictions and medium range interactions, *Biochim Biophys Acta* **916**(2):200-4.

Wilmot, C. M., and Thornton, J. M., 1988, Analysis and prediction of the different types of beta-turn in proteins, *J Mol Biol* **203**(1):221-32.

Xu, T., Vasilyeva, E., and Forgac, M., 1999, Subunit interactions in the clathrin-coated vesicle vacuolar (H(+))- ATPase complex, *J Biol Chem* **274**(41):28909-15.

Yabuuchi, H., Shimizu, H., Takayanagi, S., and Ishikawa, T., 2001, Multiple splicing variants of two new human ATP-binding cassette transporters, ABCC11 and ABCC12, *Biochem Biophys Res Commun* **288**(4):933-9.

Yokoyama, K., Nagata, K., Imamura, H., Ohkuma, S., Yoshida, M., and Tamakoshi, M., 2003a, Subunit arrangement in V-ATPase from Thermus thermophilus, *J Biol Chem* **278**(43):42686-91.

Yokoyama, K., Nakano, M., Imamura, H., Yoshida, M., and Tamakoshi, M., 2003b, Rotation of the proteolipid ring in the V-ATPase, *J Biol Chem* **278**(27):24255-8.

Zhang, J., Feng, Y., and Forgac, M., 1994, Proton conduction and bafilomycin binding by the V0 domain of the coated vesicle V-ATPase, *J Biol Chem* **269**(38):23518-23.

Zhang, J., Myers, M., and Forgac, M., 1992, Characterization of the V0 domain of the coated vesicle (H+)-ATPase, *J Biol Chem* **267**(14):9773-8.

Zhang, J., Vasilyeva, E., Feng, Y., and Forgac, M., 1995, Inhibition and labeling of the coated vesicle V-ATPase by 2-azido-[32P]ATP, *J Biol Chem* **270**(26):15494-500.

Zheng, R. L., and Kemp, R. G., 1994, Identification of interactions that stabilize the transition state in Escherichia coli phosphofructo-1-kinase, *J Biol Chem* **269**(28):18475-9.