Mammalian ADP-dependent glucokinase

A thesis presented in partial fulfilment of the requirement for the degree of
Master of Science in Biochemistry at Massey University,
Palmerston North, New Zealand.

Rebecca Hole

2009

Acknowledgements

I would like to express the absolute appreciation I have for my supervisors, Dr. Kathryn Stowell and Dr. Ron Ronimus. You have supported and encouraged me throughout this journey, and without your enthusiasm and approachability, I would not have grown or achieved nearly as much as I did.

I would also like to acknowledge and thank my family for their unending encouragement, especially when things got tough. I could not have finished this without you.
Thank you to Jeebs for being with me through trying times and for the hours you had to spend waiting for me to finish.

Thank you especially to everyone in the Twilight zone. I could not have asked for a better research lab experience. A special thanks to Natisha for her invaluable advice, suggestions, and tolerance of my chatting and questions. I couldn't have done this without you. A special thanks also to Miranda for her support, encouragement, advice and friendship.

Thank you also to Carlene and Trevor for their help, advice and patience, and to Massey University for the financial support.

Abstract

The mammalian ADP-dependent glucokinase is the most recent mammalian glucokinase to have been discovered, and is unique in its ability to catalyse the phosphorylation of glucose to glucose-6-phosphate using ADP as the phosphoryl donor. Up until the discovery of this enzyme, the traditional biochemical view was that the first step of glycolysis was solely catalysed by ATP-dependent hexokinases, types I-IV.

The particular role played by ADP-GK in the mammalian cell and the significance of this role has not yet been determined, although it is hypothesised that the ADP-dependent glucokinase could be potentially significant in contributing to the survival of cells under low energy and hypoxic or ischemic conditions. By using ADP as the energy investment in phase one of the glycolytic cycle instead of ATP, it is predicted that glycolysis could be sustained for longer during lower energy conditions (conditions of high ADP:ATP ratios). Since the phosphorylation of glucose by ADP-GK results in the production of AMP, it may also be possible that this has a direct effect on the energy charge of the cell. The AMP produced could lead to the regulation of cellular metabolism during hypoxia and/or ischemia via the activation of the cell-energy regulator AMPK.

The study of mammalian ADP-dependent glucokinase is a very new area, and prior to this no investigation of the human ADP-GK enzyme had been undertaken. The main objective of this project was to clone, express and purify the recombinant ADP-GK so it could be kinetically characterised and directly compared with the recombinant mouse kinetic characteristics, the only other mammalian ADP-GK to have been studied. Unfortunately, due to complications in the expression and purification of soluble recombinant human ADP-GK, the project did not incorporate the kinetic characterisation of the enzyme. Acquiring data on the kinetic characteristics of the human ADP-GK will, in the long term, assist in the elucidation of the metabolic role of this enzyme, so the continuation of this project would be worthwhile.

Abbreviations

ADP-GK	ADP-dependent glucokinase
Amp	Ampicillin
APS	Ammonium persulfate
ATP	Adenosine triphosphate
bp	Base pairs (DNA)
BSA	Bovine serum albumin
CD	Circular dichroism
cDNA	Complimentary DNA
CHAPS	3[(3-Cholamidopropyl)dimethylammonio]-propanesulfonic acid
DMSO	Dimethyl sulfoxide
DNA	Deoxyribose nucleic acid
DNase I	Deoxyribonuclease I
dNTP	Deoxynucleoside triphosphate (dATP, dTTP, dGTP, dCTP)
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylene diamine tetra-acetic acid
FPLC	Fast protein liquid chromatography
hADPGK	Human ADP-dependent glucokinase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horse radish peroxidase
IEF	Isoelectric focusing
IEX	Ion exchange chromatography
IPTG	Isopropyl β-D-thiogalactoside
KCl	Potassium chloride
kDa	Kilodaltons
LB	Luria Bertani bacteriological media
mAPDGK	Mouse ADP-dependent glucokinase
MGC	Mammalian gene collection
MOPS	3-(N-morpholino)propanesulfonic acid

mRNA	Messenger RNA
NaCl	Sodium chloride
NADH	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pI	Isoelectric point
PMSF	Phenylsulfonylmethyl fluoride
RNA	Ribonucleic acid
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
TAE	Tris acetate EDTA buffer
TBST	Tris-buffered saline-Tween 20
TEMED	N,N,N’,N’-Tetramethylethylenediamine
Tris	tris (hydroxymethyl)-aminomethane
UV	Ultra violet

List of Figures

Page number

List of Tables

Page number

Table of Contents

Chapter 1: Introduction

After more than 70 years of research into glycolysis and glucose utilization, the discovery of a new mammalian enzyme (ADP-GK) with the ability to catalyse the first step in the catabolism of glucose is highly significant. Up until the discovery of this enzyme, the traditional biochemical view was that the first step of glycolysis was solely catalysed by ATP-dependent hexokinases, types I-IV. However, it is now known that the ADP-dependent glucokinase has the ability to catalyse the phosphorylation of glucose via the following reaction (Ronimus and Morgan, 2004):

Glucose + ADP => glucose-6-phosphate + AMP

While this thesis focuses on ADP-dependent glucokinase, the regulation of glycolysis and the ATP-dependent enzymes that direct glucose into this fundamental pathway will be discussed. This will be followed by a brief overview of the ADP sugar kinases and the potential metabolic significance of mammalian ADP-GK.

1.1 Glycolysis

Glycolysis is a catabolic process that is carried out in almost every living cell. It is a process that has been studied in much detail, as it was the first metabolic pathway to be elucidated. The journey of its discovery began in 1860 when Louis Pasteur realised that micro-organisms were responsible for fermentation, and was concluded in 1940 with the completed metabolic pathway. Glycolysis is known alternatively as the Embden-Meherhof-Parnas pathway in commemoration of the three individuals involved foremost in its elucidation.

Glycolysis is the process by which one glucose molecule is converted to two pyruvate molecules, and in so doing, releases enough free energy to synthesise two ATP from ADP and P_i. This is achieved by a pathway of ten enzyme-catalysed reactions that involve chemically coupled phosphoryl-transfer reactions.

Glycolysis can be thought to occur in two stages: the first stage involves the phosphorylation of glucose (a hexose) by a hexokinase, and the cleavage of the resulting glucose-6-phosphate into two molecules of glyceraldehyde-3-phosphate (a triose). In the process, two molecules of ATP are utilised. The second stage converts the two molecules of glyceraldehyde-3-phosphate into two molecules of pyruvate and thereby generates four molecules of ATP, giving a net gain of two ATP per glucose catabolised. The overall reaction of glycolysis is:

$$\text{Glucose} + 2NAD^+ + 2P_i \rightarrow 2NADH + 2 \text{ pyruvate} + 2ATP + 2H_2O + 4H^+$$

In aerobic conditions, the pyruvate generated by glycolysis is transported into the mitochondrial matrix where it is converted to acetyl-CoA and further metabolised by the citric acid cycle. NADH produced by glycolysis is re-oxidised by the transfer of its electrons to the mitochondria. The oxidation of glucose under aerobic conditions is known as oxidative phosphorylation and results in the generation of approximately 32 molecules of ATP (Voet and Voet, 2004).

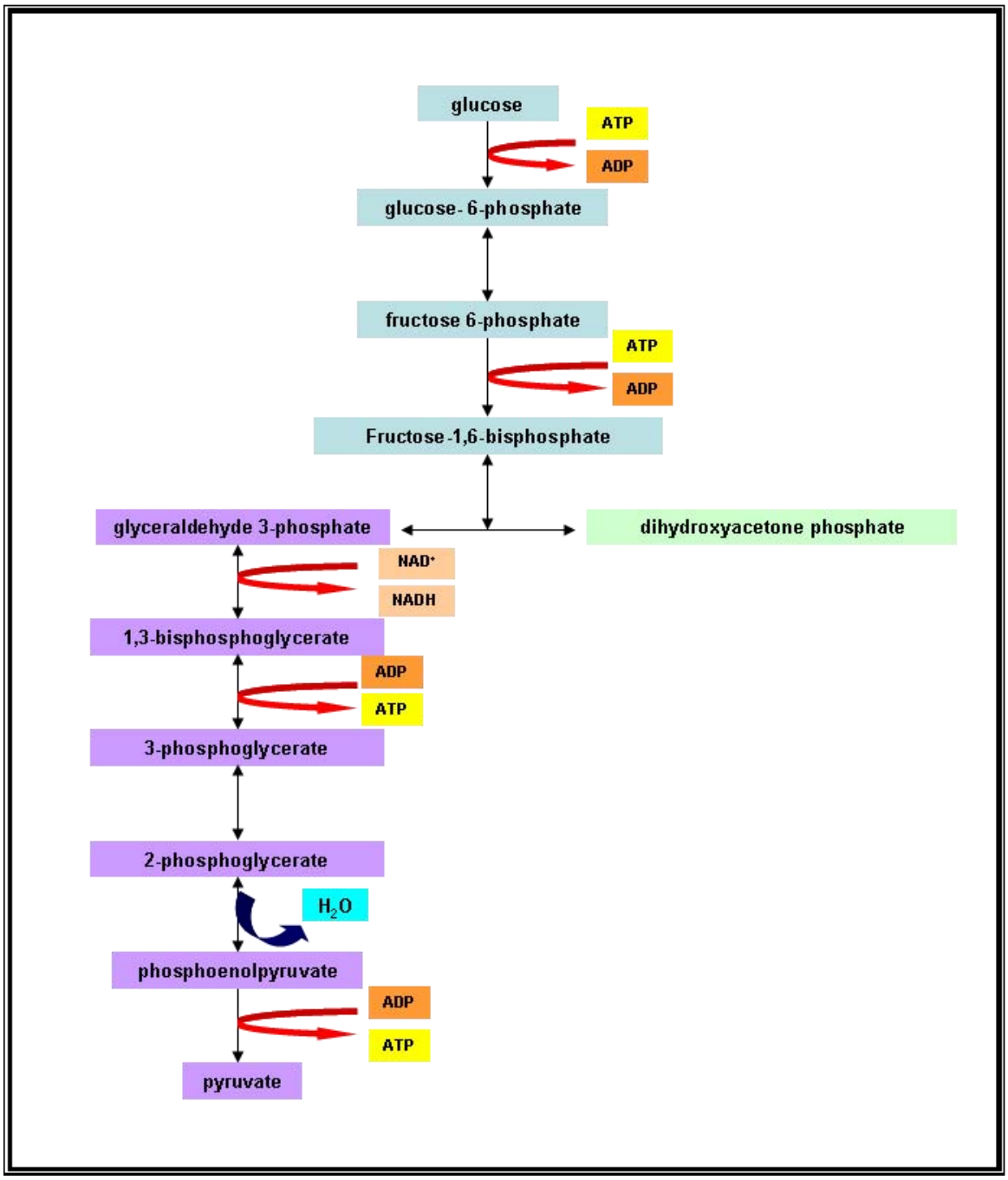


Figure 1.1 Schematic diagram of the glycolytic pathway

Glycolysis is the process by which one glucose molecule is converted to two pyruvate molecules, thereby releasing enough free energy to synthesise two ATP from ADP and P_i.

One of the critical consequences of oxygen deprivation for a cell is that it becomes limited in its ability to carry out oxidative phosphorylation, which is the most efficient way of generating ATP. To compensate for this, stressed cells become greatly dependent on glycolysis. This reliance on glycolysis under anaerobic conditions is known as the Pasteur Effect (Guppy *et al*., 2005). In order to re-oxidise $NADH/H^+$ to NAD^+ under anaerobic conditions, pyruvate is instead converted into lactate by lactate dehydrogenase as follows:

$$\text{Pyruvate} + \text{NADH} + H^+ \rightarrow \text{Lactate} + NAD^+$$

Lactate can subsequently be converted back to pyruvate and then glucose in the liver. Anaerobic glycolysis is common in skeletal muscle under conditions of repeated muscle contraction, where the production of ATP needed for contraction cannot be met by the mitochondria alone (Voet and Voet, 2004).

1.2 Isoenzymes of mammalian hexokinase

Prior to 1964, what were considered to be two types of ATP:hexose 6-phosphotransferases (hexokinases) had been isolated from rat liver using ammonium sulfate fractionation; one with a low K_m for glucose (hexokinase) and one with a high K_m for glucose (glucokinase). However, in 1964 it became clear after chromatography experiments using DEAE-cellulose, that there were four fractions of enzyme that had hexokinase activity (Gonzalez *et al.,* 1964). These four isoenzymes, commonly referred to as hexokinase types I, II III and glucokinase (type IV), were later shown to have unique electrophoretic mobilities and kinetic properties (Katzen and Schimke, 1965).

Hexokinases types I- III and glucokinase catalyse the phosphorylation of glucose, allowing it to enter any of the several alternative pathways of glucose metabolism. However, the hexokinases have distinctive differences in their kinetic and regulatory properties, their abundance and distribution in tissue, and their intracellular localisation. These variances are considered to specify the isoenzymes of hexokinase for particular roles in glucose metabolism (Sebastian *et al*., 2001; Kaselonis *et al*., 1999).

The hexokinase isoenzymes I-III are predicted to have arisen through the duplication and subsequent fusion of an ancestral hexokinase gene encoding a 50 kDa hexokinase. Accordingly, mammalian hexokinases I-III are approximately 100 kDa and have internal sequence repetition. The N- and C- terminal halves have almost fifty per cent sequence identity to each other, as well as extensive similarity to the 50 kDa isoenzyme hexokinase IV (glucokinase), indicating its close evolutionary relationship with the higher molecular weight isoenzymes. Both the N- and C-terminal halves are catalytically active in the type II isoenzyme, however, the type I and type III isoenzymes have a catalytically active C-terminal half only, while the N-terminal halves have gained a regulatory function (Sebastian *et al*., 2001).

In mammalian tissues, hexokinase type I is widely expressed, and is found principally in tissues reliant on glucose as their primary source of energy. One example of this is the brain, which is essentially completely dependent on glycolysis. Hexokinase type I has a predominantly catabolic role, with the function of introducing glucose into glycolytic metabolism in order to generate energy (ATP) (Sebastian *et al*., 2001; Kaselonis *et al*., 1999). Hexokinase I is clearly differentiated from hexokinases II and III by its response to inorganic phosphate (P_i). Hexokinase types I-III are susceptible to potent inhibition by glucose-6-phosphate, but only in hexokinase I can physiological levels of inorganic phosphate relieve the inhibition glucose-6-phosphate causes (Aleshin *et al*., 1998). Inorganic phosphate acts as an inhibitor to hexokinase types II and III. During periods of heightened energy demand, an increase in the $[P_i]$/[glucose-6-phosphate] ratio occurs due to hydrolysis of high energy phosphate compounds and increased glycolytic metabolism. This results in increased hexokinase I activity, fitting with its catabolic profile (Wilson, 2003).

Hexokinase type II has a more restricted expression in mammals, and is found primarily in insulin-sensitive tissues such as the liver, skeletal muscle, cardiac muscle, mammary gland, and adipose tissue. In contrast with hexokinase type I, hexokinase type II has an anabolic function, assisting in glycogen and lipid synthesis. It is the only 100 kDa hexokinase that is regulated by insulin (Sebastian *et al*., 2001; Kaselonis *et al*., 1999; Heikkinen *et al*., 2000). It has been demonstrated that the presence of insulin increases hexokinase type II mRNA

levels, protein content, and activity in skeletal muscle, within a physiologically relevant timescale (Vogt *et al*., 2000). The anabolic role played by hexokinase II is consistent with the findings that selective enhancement of hexokinase type II expression occurs in livers of transgenic mice over-expressing the transcriptionally active nuclear form of SREBP-1a. Sterol regulatory element-binding protein SREBP-1a is considered to be a primary influence on the expression of enzymes involved in fatty acid and triglyceride biosynthesis. The generation of glucose-6-phosphate by hexokinase II and its subsequent entry into metabolism provides carbons and reducing equivalents (NADPH) required for lipid biosynthesis (Sebastian *et al*., 2000). It has also been observed that in the process of mammary tissue undergoing lactation, hexokinase type II mRNA becomes expressed. Consequently, it has been suggested that hexokinase II has a specific role in the mammary gland in terms of the adequate carbohydrate production required for lactation (Kaselonis *et al*. 1999).

Hexokinase type III has proved to be more difficult to analyse in terms of its specific role in tissue metabolism. This isoenzyme has not been found as the predominant form in any tissue, however it is found as a relatively minor amount in most tissues, which makes it difficult to correlate the enzyme with a particular metabolic role (Preller and Wilson, 1992). So far, the primary function of type III hexokinase has not been demonstrated (Kaselonis *et al*., 1999). Hexokinase type III has the least hydrophobic N-terminal segment of the isoenzyme types I-III. This suggests it is unlikely that hexokinase type III can bind to the mitochondria in an equivalent manner to that of types I and II, if at all. It has been shown via immunohistochemical experiments that type III has a predominantly nuclear localisation, although the association formed tends to be weak, and can be readily lost when cellular structure is disrupted, hence the several reports considering type III to be a cytoplasmic enzyme. Although it would be expected that the cells containing nuclear type III hexokinase would share some metabolic characteristics, a clear linkage has not been demonstrated. However, there has been a connection made between ubiquitous endothelial cells and hexokinase type III, which helps to explain the observation that at low levels at least, this isoenzyme is present in all tissues (Preller and Wilson, 1992).

Hexokinase type IV, more commonly referred to as glucokinase, has the highest K_m for glucose of the four mammalian hexokinase isoenzymes. It is primarily located in pancreatic β-cells and liver, but has also been found in pituitary cells, the entero-endocrine K and L cells, and neurons of the hypothalamus and brainstem (neuro-endocrine cells). It is hypothesised that all the glucokinase expressing cells of the mammalian system operate as a kind of network with the objective of maintaining glucose homeostasis. However, the exact role of pituitary, the entero-endocrine, and neuronal forms of glucokinase have yet to be demonstrated (Matschinsky *et al*., 2006; Zelent *et al*., 2006).

Glucokinase has a well-specified dual function: in the liver, its enzymatic function results in the initiation of glucose storage in the form of glycogen, and in the insulin-producing pancreatic β cells, its role is described as that of a glucose sensor (Baltrusch *et al*., 2006).
With regard to glucose, glucokinase shows cooperative kinetics. It is not directly inhibited by metabolites of glucose (such as glucose-6-phosphate or glucose-1,6-bisphosphate), unlike the other three hexokinase isoenzymes, and hence its rate is determined by intra-cellular glucose concentrations only. Of its gene family, glucokinase alone can adjust the rate of glucose phosphorylation in response to changes in physiological glucose concentrations (Jetton *et al*., 1994).

In the liver, glucokinase expression is transcriptionally controlled primarily by the hormone insulin. Therefore, in the presence of insulin, increased uptake of glucose occurs. This is due to an increase in the rate of glucose phosphorylation, which in turn creates a glucose gradient for glucose transport into the liver. The synthesis of glycogen can then follow, as glycogen synthase activation relies solely on glucose-6-phosphate produced by glucokinase (de la Iglesia *et al.*, 2000).

In pancreatic β-cells, glucokinase plays a critical role in the multi-step process of glucose sensing. It is thought that the kinetic characteristics of glucokinase specify many glucose sensing features in the islets as phosphorylation of glucose is the rate-limiting step in glucose utilisation, and therefore glucose-induced insulin secretion, by the β-cell. It has also

been noted that this neuro-endocrine form of glucokinase appears to be expressed constitutively, and is regulated only minimally by hormones (Matschinsky *et al.*, 2006).

The differences in regulation and expression of these enzymatically identical forms of glucokinase can be explained by the two different promoters found in the glucokinase gene. The upstream promoter is used for the β-cell form, and the downstream promoter for the hepatocyte. This leads to a difference in the 15 amino acids at the N-terminus of glucokinase, which allows regulation to be specified to the tissue, while having no apparent functional impact on the enzyme (Magnuson and Shelton, 1989).

In addition to regulation by insulin, hepatocyte glucokinase activity is also regulated on a short term basis by a 68 kDa glucokinase regulatory protein (GKRP). This protein has an inhibitory effect on hepatic glucokinase and is found in the nucleus. Under low glucose conditions, glucose complexes with GKRP and sequesters the enzyme to the nucleus. Translocation of glucokinase to the cytoplasm occurs at raised glucose (or fructose) concentrations. GKRP is expressed during development, preceding the expression of glucokinase. So far, no physiological situation has been encountered where hepatic glucokinase is expressed without the regulatory protein (Matschinsky *et al*., 2006; de la Iglesia *et al*., 2000).

1.3 Warburg Effect

Over 80 years ago it was recognised by Nobel Prize winner Otto H. Warburg that cancerous liver cells have increased glycolytic activity in the presence of oxygen (aerobic glycolysis) when compared to normal liver cells. This observation (now known as the Warburg effect) has had a significant impact on the way in which cancer has subsequently been studied. The common view is that the reliance of cancer cells on glycolysis for the production of ATP, rather than mitochondrial oxidative phosphorylation, is an important metabolic difference distinguishing malignant cells from normal cells. The exact mechanism for the Warburg effect still has not been clearly elucidated.

The shift in metabolism in cancer cells observed was explained by Warburg as being a result of respiratory injury. He suggested that aerobic fermentation (glycolysis) was a means for the cell to compensate for the loss of energy gained by respiration (Warburg, 1956). Although mitochondrial defects and mutations have been associated with a number of cancers (Carew and Huang, 2002), it has also been observed that aerobic glycolysis confers an advantage to cancer cells as it produces metabolic intermediates used for biosynthetic reactions. This is especially important for rapidly dividing neoplastic cells which ultimately require large amounts of building blocks for the biosynthesis of fatty acids, DNA replication, and protein production. Increased levels of glucose-6-phosphate resulting from the high rates of aerobic glycolysis are also used to produce the biosynthetically useful cofactor, NADPH, via the pentose phosphate pathway (Pastorino *et al*., 2002).

The increase in the glycolytic rate of transformed cells has been exploited in both the approach to imaging of tumours as well as anti-cancer drug design. Positron emission tomography (PET) imaging can be used to detect and map many tumours by using a fluorescent glucose analogue (^{18}F-fluorodeoxyglucose). The 'glucose-hungry' tumours can preferentially take up the injected analogue and be detected by a PET scanner (Garber, 2006).

Glycolysis is a favoured target for anti-cancer drugs due to the observed reliance of tumours on glycolysis for energy and metabolic intermediates. Compounds such as 3-bromopyruvate, which has a potent inhibitory effect on glycolysis (a known inhibitor of hexokinase type II), are being tested as possible clinical treatments (Xu *et al*., 2005). 3-bromopyruvate is an alkylating agent which is an analogue of lactate and pyruvate. It can selectively deplete ATP and, as a result, induce cell death. It has been demonstrated in 19 rats with advanced cancerous tumours of 2-3cm in diameter, that the cancer can be treated with 3-bromopyruvate. In all 19 animals, the tumours were eradicated without apparent toxicity or recurrence (Ko *et al*., 2004).

The high glycolytic rate observed in neoplastic cells has been shown to be partly due to the significantly increased expression of hexokinase type II. Up to 70 per cent of hexokinase type II is bound to the mitochondria in cancerous cells and this has been shown to be mediated through an interaction with a Voltage-dependent anion channel (VDAC), an outer membrane channel-forming porin (Nakashima *et al*., 1988; Pastorino *et al*., 2002). In this bound state, hexokinase type II has preferential access to ATP produced by the mitochondria. The availability of ATP from the mitochondria means the cell can maintain a greater rate of glycolysis than normal.

It has also been demonstrated that the binding of hexokinase type II to the mitochondria contributes to the survival of transformed cells by preventing Bax from interacting with the mitochondria. During apoptosis, Bax, a proapoptotic protein, translocates to the mitochondria and binds to VDAC. This mediates the release of intermembrane space proteins, such as cytochrome *c*. Bax and VDAC form a channel through which cytochrome *c* can pass and this helps to trigger the activation of caspases in the cytoplasm. Caspases are integral to the break up of apoptotic cells (Pastorino *et al*., 2002).

1.4 Hypoxia, cellular energetics and cancer

AMP-activated protein kinase (AMPK) is a master sensor of the energy status of the cell, otherwise known as the energy charge of the cell. It is responsible for switching on catabolic pathways and switching off many ATP-consuming processes (Carling *et al*., 1987; Corton *et al*., 1994). Healthy cells, under ideal conditions, maintain an average ADP:ATP ratio of 1:10. This is due to the operation of ATP synthases and adenylate kinase. Under these conditions, adenylate kinase operates as shown in the following equation from right to left keeping AMP very low (an AMP:ATP ratio of 1:100):

$$2ADP \leftrightarrow ATP + AMP$$

If cellular stress causes the rate of ATPases to exceed that of ATP synthases, the ADP:ATP ratio will typically rise by 5-fold and the AMP:ATP ratio will rise 25-fold. Thus the change

in AMP concentration is much more dramatic than the changes that occur in ADP or ATP, making AMP a useful energy charge indicator to which AMPK responds (Hardie and Hawley, 2001).

AMP binds allosterically to AMPK, this being the method in which AMPK senses an increase in the cellular [AMP]/[ATP] ratio. Activated (via phosphorylation) AMPK phosphorylates a range of targets, many that are involved in controlling cellular energy metabolism directly. It is believed that when a cell is exposed to an energy depleting stress, AMPK functions to inhibit ATP-consuming processes and to activate ATP-producing processes. This way, total cellular ATP levels are optimised in order to maintain critical physiological functions. However, it has recently been demonstrated that AMPK activity is also induced soon after the onset of either physiological of pathophysiological hypoxia. This can occur even when total cellular ATP levels are not considerably depleted (Laderoute *et al*., 2006). Hypoxic regions have been observed within the microenvironments of most cancerous tumours, as a result of insufficient blood supply. The extent of hypoxia can be correlated with prognosis in a number of tumour types (Talks *et al*., 2000). It can also be associated with chemo- and radio-therapeutic resistance, and has also been implicated in enhancing metastasis (Guppy *et al*., 2005).

1.5 Hypoxia-inducible Factor 1

Hypoxia-inducible factor 1 (HIF-1), a basic helix-loop-helix-PAS (Per-ARNT-Sim) domain transcription factor, has been described as the master regulator of the transcriptional response to oxygen deprivation (Rapisarda *et al*., 2002), or hypoxia. It has been implicated in the up-regulation of more than 40 genes (Vaupel *et al*., 2004), including genes involved in angiogenesis (*e.g*. vascular endothelial growth factor (VEGF), erythropoietin (EPO) and inducible nitric oxide synthase), as well as genes encoding glycolytic enzymes and glucose transporters (GLUT) (Lu *et al*., 2002; Rapisarda *et al*., 2002).

HIF-1 is a heterodimer made up of HIF-1α and HIF-1β. Both of these subunits are constitutively expressed in mammalian cells, however, levels of HIF-1α are principally

dependent on intracellular oxygen concentration. This largely constitutes the regulation of the HIF-1 complex. Under normoxic conditions, HIF-1α is ubiquitinated and undergoes proteasomal degradation. In order for this to occur, the oxygen-dependent degradation domain on the HIF-1α protein must be bound by the von Hippel-Lindau tumour suppressor protein. This binding is regulated by a family of prolyl hydroxylase enzymes via the hydroxylation of key proline residues on HIF-1α. The prolyl hydroxylases require oxygen and iron for their activity, hence explaining HIF-1α accumulation under hypoxic conditions. Accumulated HIF-1α protein relocates to the nucleus where an active complex is formed with HIF-1β. Activation of target genes occurs by binding to the DNA consensus sequence 5′-RCGTG-3′.

The presence of the active HIF-1 complex has been clearly associated with cancer cell growth and survival, tumour angiogenesis and development, as well as poor clinical prognosis (Rapisarda *et al*., 2002; Lu *et al*., 2002).

1.6 ADP-dependent glucokinase

The first two ADP-dependent kinases were discovered somewhat accidentally in 1994 as enzymes in an Embden-Meyerhof-like pathway used as the major route for glucose oxidation by the hyperthermophilic archaeon *Pyrococcus furiosus* (Kengen *et al*., 1994). Due to its favourable culture conditions, *Pyrococcus furiosus*, a rapidly growing marine heterotroph first isolated in 1986, is one of the most extensively studied hyperthermophiles.

When studying the route used by this archaeon to oxidise glucose (by *in vivo* ^{13}C NMR), researchers (Kengen *et al*., 1994) realised that a modified Embden-Meyerhof pathway was involved. However, the expected hexokinase and phosphofructokinase activity could not be detected using standard assay conditions. It was observed that in the presence of glucose and ATP, some hexokinase activity could be measured, but declined rapidly followed by a recovery in activity when incubation was prolonged. This led to the conclusion that a product formed during the incubation was required for activity, and it turned out that product required was ADP. The researchers found significant hexokinase and

phosphofructokinase activity could be measured in the presence of ADP, and hence discovered the first ADP-dependent glucokinase and phosphofructokinase (Kengen *et al*., 1994).

ADP-dependent glucokinases (ADP-GKs) and ADP-dependent 6-phosphofructokinases (ADP-PFKs) are found in the euryarchaeotal branch of archaea. ADP-dependent sugar kinases and their genes have been characterised from *Pyrococcus furiosus, Thermococcus litoralis,* and *Archaeoglobus fulgidus* as well as the bifunctional ADP-dependent sugar kinase from the glycogen-forming methanogenic archaeaon, hyperthermophilic *Methanococcus jannaschii.* This enzyme has both ADP-dependent glucokinase and ADP-dependent phosphofructokinase activity.

The extremely thermophilic *Pyrococcus furiosus, Thermococcus litoralis,* and *Archaeoglobus fulgidus* ADP-GKs represent, respectively, a homodimeric and two monomeric proteins with subunits of approximately 50 kDa, that are highly specific for ADP and glucose. The *M. jannaschii* bi-functional ADP-GK/PFK is monomeric and also shows high specificity for ADP, however, it accepts both glucose and fructose 6-phosphate as phosphoryl acceptors (Labes and Schonheit, 2003).

The ADP-dependent glucokinases are unrelated to the ATP-dependent hexokinases, instead structural comparisons indicate that they are distantly related to Family B sugar kinases similar to ribokinase (Ronimus and Morgan, 2003).

The first mammalian ADP-GK to be characterised was that of the mouse (*Mus musculus*). The gene encoding the mouse ADP-GK is a single copy gene with limited homology to the archaeal ADP-dependent sugar kinase gene, and is located on chromosome 9, (the corresponding human gene is located on chromosome 15).

In 2004, the mouse ADP-GK gene was cloned into *Escherichia coli,* recombinantly expressed and the soluble protein was subsequently kinetically characterised. Mouse ADP-

GK represents the first new mammalian glucokinase to be described in nearly 40 years (Ronimus and Morgan, 2004).

The estimated molecular mass of mouse ADP-GK is 51.5 kDa and is considered to be monomeric. The key biochemical characteristics of mouse ADP-GK include a high specificity for D-glucose (a diagnostic criterion of glucokinases) and a low apparent K_m (96 μM) for glucose and ADP (0.28 mM). When assayed at physiological temperatures, it has a specific activity of 11.7 U/mg protein. The enzyme has bimodal pH optima, one between 5.75 and 6.5 and one near 8.75–9.0. The bimodal optima are thought to be significant because this mirrors relative rates of glycolysis versus intracellular pH; as acid and alkaline intracellular conditions can both be associated with enhanced glycolytic rates (Ronimus and Morgan, 2004). No inhibition of mouse ADP-GK is detected with either 1.0 mM or 10.0 mM of glucose-6-phosphate, but it is markedly inhibited by high concentrations of glucose. Maximal enzyme activity occurs near 0.35 mM glucose. The mouse ADP-GK is also inhibited by its product, AMP (Ronimus and Morgan, 2004).

In terms of phylogenetics, a tree construction performed by Ronimus and Morgan, (2004) (Figure 1.2) showed the human and mouse ADP-GKs to be grouped together, and a grouping of the *Drosophila* and *Anopheles* enzymes. From the tree, it can be observed that these four proteins combined with the *Caenorhabditis elegans* ADP-GK, were most closely related to the mesophilic methanogen ADP-GKs. It appears that the ADP-GKs are a homogeneous clade, with the next closest relative being the bi-functional *Methanococcus jannaschii* ADP-GK/ ADP-PFK enzyme (Ronimus and Morgan, 2004).

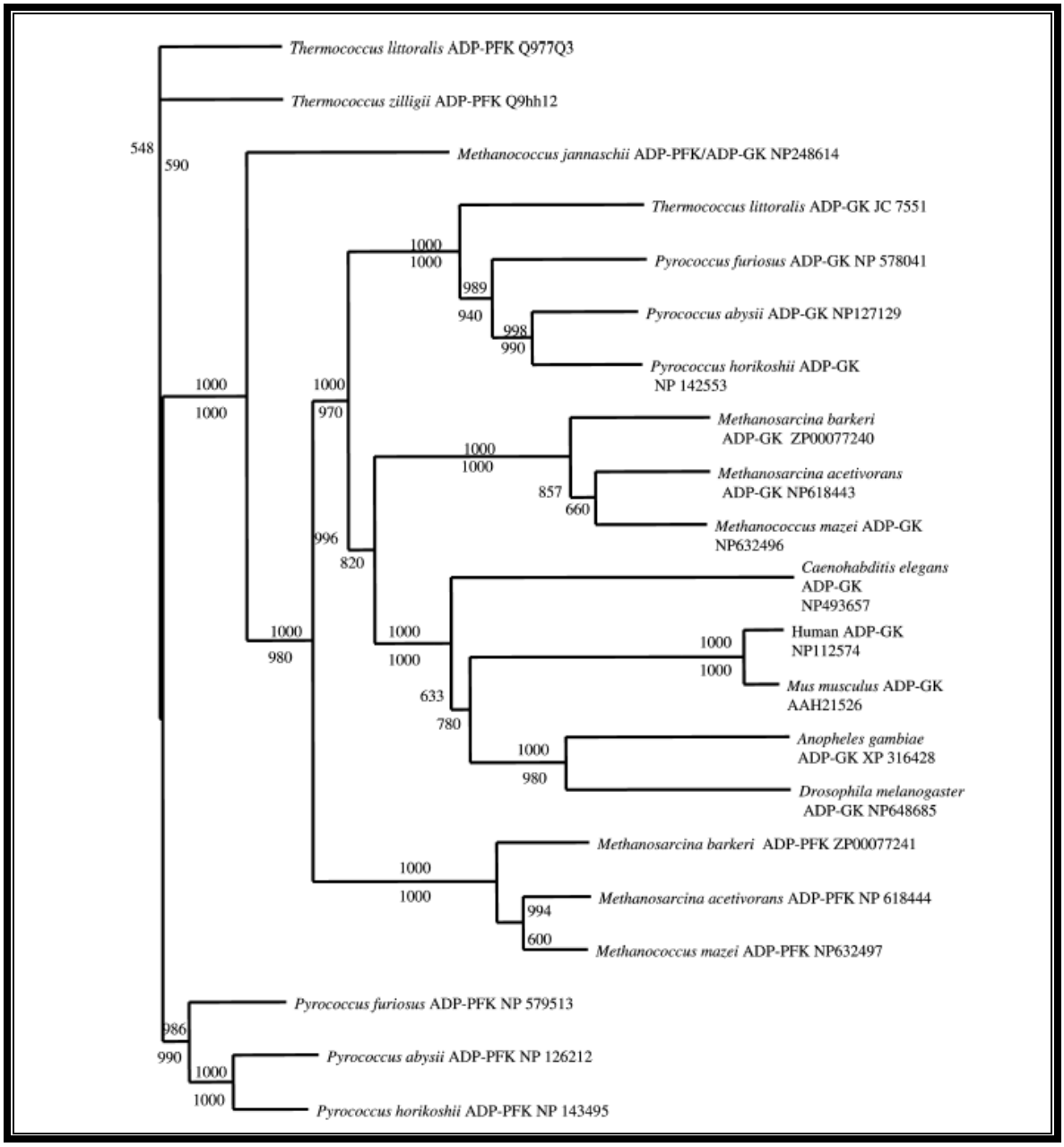


Figure 1.2 Phylogenetic tree of higher eukaryote ADP-GKs and archaeal ADP -GKs and ADP-PFKs (Ronimus and Morgan, 2004)

The GenBank database accession codes are shown after the species names. Numbers above the nodes signify the neighbour-joining bootstrap analysis and numbers below denote maximum parsimony (Ronimus and Morgan, 2004).

1.7 The potential significance of ADP-dependent glucokinase

At this point in time, there is no obvious metabolic role that can be attributed to the enzymatic activity of mammalian ADP-GK. It can only be hypothesised that the ADP-GK may confer some kind of advantage to a cell under energy stressed conditions, in terms of net savings of ATP. By using ADP as the energy investment in phase one of the glycolytic cycle instead of ATP, it is predicted that glycolysis could be sustained for longer during lower energy conditions (conditions of high ADP:ATP ratios). Since the phosphorylation of glucose by ADP-GK results in the production of AMP, it may also be possible that this has a direct effect on the energy charge of the cell. The AMP produced could lead to the regulation of cellular metabolism during hypoxia and/or ischemia via the activation of the cell-energy regulator AMPK.

The ability of ADP-GK to use GDP as an alternative to ADP (Ronimus and Morgan, 2004), may contribute further to the maintenance of glycolysis under low energy conditions. Protein synthesis, which generally requires approximately 25 per cent of cellular energy, is down-regulated during hypoxia (Buttgereit and Brand,1995; Rolfe and Brown, 1997), thus allowing the GTP/GDP pool to be accessed where it would normally be unavailable.

Hypoxia and ischemia are linked with most major diseases including cancer, heart attack, stroke, and chronic lung diseases (Brahimi-Horn and Pouyssegur, 2007; Roffe *et al*., 2003), making the potential role of ADP-GK in the survival of cells under low-oxygen and low energy conditions clinically significant. Importantly, ADP-GK is structurally distinct from the ATP-hexokinases I-IV (Ronimus and Morgan, 2004), and therefore, there are strong prospects for its selective inhibition and/or activation, in new drug-based therapies. Differential expression of ADP-GK could potentially become a marker for hypoxia and thereby aid in the diagnosis of strokes, heart attacks, and hypoxic tumours.

1.8 Research aims

The study of ADP-dependent glucokinases is a very new area, and prior to this no investigation of the human ADP-GK enzyme had been undertaken. The main objective of the project addressed this with the aim to clone, express and purify the recombinant ADP-GK so it could be kinetically characterised and directly compared with the recombinant mouse kinetic characteristics. Acquiring data on the kinetic characteristics of the human ADP-GK will, in the long term, assist in the elucidation of the metabolic role of this enzyme.

The specific objectives of this research were as follows:

- To clone the human ADP-GK coding sequence into a pET151/D-TOPO® vector
- To express and purify the recombinant human ADP-GK from *Escherichia coli*
- To kinetically characterise the recombinant human ADP-GK by biochemical assay
- To screen mammalian tissues for the expression of ADP-GK

Chapter 2: Materials and methods

2.1 Materials

Protein manipulations

Complete™ Mini EDTA-free protease inhibitor cocktail tablets were obtained from Roche Molecular Biochemicals, IN, USA. Bradford protein assay dye concentrate and Dual Colour Precision Plus protein markers were purchased from BioRad Laboratories, CA, USA. N,N,N',N'-Tetramethylethylenediamine (TEMED), lysozyme and rabbit anti-mouse secondary antibody conjugated to horseradish peroxidase were obtained from Sigma Chemical Company, St Louis, MO, USA. Polyacrylamide-bis (29.1:0.9) ready-to-use solution (40%) was acquired from Merck, Darmstadt, Germany. Monoclonal mouse antibody to ADP-dependent glucokinase was purchased from Abnova, Taipei, Taiwan. Positively charged nylon membrane and BM chemiluminescence blotting substrate A and B were obtained from Roche, Mount Wellington, New Zealand. Bovine serum albumin (BSA) was supplied by New England Biolabs, MA, USA. Acrodisc Filter Unit (0.45 µm) filters were supplied by Pall Corporation, MI, USA. Ni-NTA resin was purchased from Invitrogen Corporation, CA, USA. HisTrap™ HP 5 mL affinity columns were acquired from GE Lifesciences Pty. Ltd., Rydalmere, NSW 2116, Australia, catalogue number 17-5247-01. Centrifugal filter units with 50 kDa or 30 kDa membrane cut-offs were supplied by Millipore, North Ryde, NSW 2113, Australia, or Sartorius, Aubagne Cedex, France.
Whole-cell human protein tissue arrays were purchased from Biochain Institute, Inc., CA, USA, catalogue numbers A1234709 and A1235713-1.

DNA manipulations

Dimethyl sulfoxide (DMSO) and primers ADPGK-RH1, -RH2, -RH3, -RH4, Forward-454 and Reverse-612 primers were obtained from Sigma-Aldrich, St Louis, MO, USA. *PfuTurbo*® DNA polymerase and 10× reaction buffer was purchased from Stratagene, La Jolla, CA, USA. *Taq* polymerase, 10× reaction buffer, and 1 kb plus DNA ladder were supplied by Invitrogen Corporation, CA, USA. GC-RICH solution was purchased from Roche Molecular Biochemicals, IN, USA. dNTPs were supplied by GE Healthcare Bio-

Sciences Pty. Ltd., Rydalmere, NSW 2116, Australia. Restriction endonucleases and buffers were acquired from New England Biolabs, MA, USA. The full length human ADP-GK cDNA clone was purchased from Open Biosystems, 601 Genome Way Ste. 2100, Huntsville, AL. The TOPO® cloning kit was supplied by Invitrogen Corporation, CA, USA. Quantum Prep® Plasmid Miniprep Kit was purchased from BioRad Laboratories, CA, USA and the PureLink™ HiPure Plasmid Midiprep Kit was supplied by Invitrogen Corporation, CA, USA. QIAquick PCR purification kit was purchased from QIAGEN Sciences, MD, USA.

Enzyme assays

All chemicals used for inhibitor/activator screens, NAD^+, ADP and glucose-6-phosphate dehydrogenase linker enzyme (*Leuconostoc mesenteroides),* were purchased from Sigma-Aldrich, St Louis, MO, USA. MOPS buffer, potassium chloride and glucose were obtained from Merck, Darmstadt, Germany. Plastic disposable microcuvettes (100 µL reaction volume) were purchased from Eppendorf South Pacific, NSW, Australia, catalogue number 0030 106.300.

General laboratory supplies

1.5 mL micro-centrifuge tubes, thin-walled PCR tubes and all pipette tips were obtained from Axygen, Union City, California USA. Gel loading tips (200 µL) were sourced from Raylab NZ Ltd, Glendene, Auckland, catalogue number BP3700GL. Ammonium persulfate (APS), NaCl, imidazole, sodium dodecyl sulfate (SDS), urea, CHAPS and Triton-X-100 were acquired from Merck, Darmstadt, Germany. 15 mL and 50 mL tubes and 1.0 mL cryotubes were purchased from Nunc Inc, Naperville, IL, USA. Phenylsulfonylmethyl fluoride (PMSF) and dithiothreitol (DTT) were supplied by Sigma Chemical Company, St Louis, MO, USA. LB broth (10 g SELECT peptone 140, 5 g SELECT Yeast Extract, 5 g sodium chloride per litre) was purchased from Invitrogen Corporation, CA, USA. Mini-Protean 3 cell apparatus, glass spacer plates (0.75 mm or 1.5 mm), and short plates were obtained from BioRad Laboratories, CA, USA.

All other chemicals and reagents used were of analytical grade.

2.2 Methods

2.2.1 Agarose gel electrophoresis

Following PCR, 15 µL of each reaction was analysed by agarose gel electrophoresis. Tris-acetate EDTA (50×) (2.0 M Tris, 2.0 M acetic acid, 50 mM EDTA, pH 8.5) was diluted 50-fold in MilliQ water to give 1× TAE. 1% agarose gels (0.6 g agarose melted in 60 mL of 1× TAE) containing 1 µL of ethidium bromide (10 µg/µL) were used to separate PCR products. The liquid agarose was poured into the electrophoresis apparatus followed by the placement of a 10 well comb. The gel was allowed to set for 30 to 60 minutes. Once set, the comb was removed and the gel was covered with 1× TAE, and 2 µL of ethidium bromide (10 µg/µL) was added to the buffer. DNA samples were mixed with 3 µL of DNA loading dye (40% w/v sucrose, 0.25% bromophenol blue) and loaded on the gel. Electrophoresis was carried out at 100 V until the dye front had travelled approximately two thirds down the gel (about 1 hour). DNA can be visualised by the detection of fluorescence when exposed to UV light, due to the complex formed between ethidium bromide and the DNA duplex, which absorbs UV irradiation at 260 nm. Gels were analysed after exposure to UV light using the Gel Doc™ (BioRad, Hercules, California, USA) to visualise the DNA. The sizes of the DNA bands were determined by comparing them to a set of known DNA size standards (1 kb plus ladder).

2.2.2 Polymerase chain reaction

The polymerase chain reaction is a method by which relatively short, specific sequences of DNA are amplified exponentially using a pair of primers and a thermostable polymerase such as *Taq* (*Thermus aquaticus)* polymerase. Polymerase chain reactions amplifying sequences to be used directly for cloning were set up using *PfuTurbo*® as this polymerase has a reported error rate six-fold lower than that of *Taq* polymerase. Co-solvents, such as DMSO and GC-RICH solution were added to PCR reactions because of their ability to improve the denaturation of GC-rich DNA and help overcome the difficulties of polymerase extension through secondary structures.

(i) Reactions using PfuTurbo® polymerase

PCR reactions were set up in 0.2 mL thin-walled PCR tubes in a total reaction volume of 50 µL. To the PCR tube was added 37.5 µL of sterile water, 5 µL 10× reaction buffer (200 mM Tris-HCl, pH 8.8, 20 mM $MgSO_4$, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 1% Triton X-100, 1 mg/ml nuclease-free BSA), 1 µL of 10 mM dNTP mix, 2.5 µL of DMSO, 1 µL of the forward and reverse primer, each at a concentration of 250 ng/µL, 1.0 µL of the template DNA (100 ng/µL) and 1.0 µL of the *PfuTurbo®* polymerase (2.5 U/µL). The tube was gently mixed by vortexing and briefly centrifuged to collect the contents at the bottom of the tube. PCR was carried out using the program outlined below:

95ºC	3.0 min	Initial denaturation
95ºC	30 sec	Amplification 20-30 cycles
61ºC	30 sec	
72ºC	110 sec	
72ºC	10.0 min	Final extension

(ii) Reactions using Taq polymerase

PCR reactions were set up in 0.2 mL thin-walled PCR tubes (Axygen) in a total reaction volume of 50 µL. 36.1 µL of sterile water, 5 µL 10× PCR reaction buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl) (Invitrogen), 0.4 µL of 25 mM dNTP mix, 4 µL GC-RICH solution, 1 µL of the forward and reverse primer, each at a concentration of 250 ng/µL, 1.0 µL of the template DNA (100 ng/µL) and 0.5 µL of the *Taq* polymerase (5.0 U/µL) were added to the PCR tube. The tube was gently mixed by vortexing and briefly centrifuged to collect the tube contents at the bottom. PCR was carried out using the program outlined below:

94ºC	4.0 min	Initial denaturation
94ºC	54 sec	Amplification 20-30 cycles
53ºC	30 sec	
72ºC	110 sec	
72ºC	10.0 min	Final extension

2.2.3 Elution of plasmid DNA from IsoCode® paper

A full-length human ADP-GK cDNA sequence verified by the mammalian gene collection (MGC), which is a database created by the National Institute of Health, was purchased from Open Biosystems as a clone in the non-expression vector pOTB7 (Appendix 1). The clone was provided on IsoCode paper which is designed for the ambient storage of plasmid clones. The following protocol provided by Open Biosystems was used to elute the plasmid clone:

One punch of IsoCode paper was washed by placing it in a micro-centrifuge tube with 100 µL of sterile water and vortexing the tube three times. The wash was discarded and the punch resuspended in 20 µL of sterile water and boiled at 95-100°C for 30 minutes in a water bath. This was followed by pulse vortexing the tube 60 times. The tube was spun briefly in a microcentrifuge to pool the supernatant containing the plasmid DNA. The punch was removed from the tube using sterile forceps and the supernatant cooled on ice. Between 1 and 5 µL of the supernatant was used in transformation reactions as outlined in Section 2.2.7**.**

2.2.4 Purification of PCR products

Positive PCR products were gel purified before use in cloning reactions using the QIAquick PCR purification kit. Double-stranded DNA was purified using a silica-based spin cartridge which removes impurities in the solution, while DNA remains trapped in the column. The DNA was then eluted in 1× TAE or water. Purification was performed following the manufacturer's instructions.

2.2.5 DNA quantification

The absorbance of a sample of double-stranded DNA at a wavelength of 260 nm (A_{260}) can be used to calculate DNA concentration using the formula:

Concentration = $A_{260} \times 50$ µg/mL

At this wavelength and DNA concentration, the absorbance of double-stranded DNA is 1.0. Double stranded DNA concentrations were determined by ultraviolet spectrophotometry using a Nanodrop® ND-1000 spectrophotometer (Nanodrop) using 1 µL of sample. The purity of a DNA sample was analyzed by measuring the ratio of A_{260}/ $A_{280.}$ Pure DNA has a ratio of approximately 1.8; a ratio less than 1.8 indicates protein contamination and a ratio greater than 1.8 indicates RNA contamination.

2.2.6 Cloning PCR products into pET151/D-TOPO®

A Champion™ pET Directional TOPO® Expression Kit was used to clone the human hADP-GK cDNA sequence directly into a pET151/D-TOPO® vector (Appendix 1). To use a PCR product directly for cloning using this system, four bases (CACC) must be added to the 5′ end of the forward primer. The pET vector itself has a complementary GTGG sequence which invades the 5′ end of the PCR product and anneals to the added bases. This stabilizes the PCR product in the correct orientation for cloning. The TOPO technology uses topoisomerase I from *Vaccinia* virus to clone PCR products by charging the vector with topoisomerase at the 3′ phosphates of the cloning site. When a 5′ hydroxyl from the stabilized PCR product attacks the phospho-tyrosyl bond between the vector and the enzyme, this joins the DNA ends and releases the topoisomerase.

The following reaction was set up using 6 ng of purified PCR product (0.375 µL), the supplied salt solution (1.2 M NaCl, 0.06 M $MgCl_2$) and vector. The reaction was mixed gently and incubated for 10 minutes at room temperature (22-23°C).

Reagents	Chemically competent *E. coli*
Fresh PCR products	0.5-4 µL
Salt solution	1 µL
Dilute salt solution (1:4)	--
Sterile water	Added to a final volume of 5 µL
TOPO® vector	1 µL
Total Volume	**6 µL**

Table 2.1 Constituents of the Champion™ pET Directional TOPO® cloning reaction

Each of the reaction components were added, mixed gently and incubated for 10 minutes at room temperature before being used in a transformation reaction.

A 3 µL volume of the cloning reaction was used to transform 200 µL of competent *E. coli* TOP10 cells (used to maintain the plasmid containing the hADP-GK cDNA sequence).

2.2.7 Transformation of Escherichia coli (E. coli) (XL-1, BL21(DE3), and Rosetta strains)

(i) Preparation of competent E. coli cells

A 50 mL volume of LB broth was inoculated with 1 mL of overnight culture of an appropriate strain of *E. coli* cells, and the culture was grown to an optical density (O.D) at A_{550} of between 0.45 and 0.50 at 37ºC. The culture was pelleted for 10 minutes at 1000 *g* in a refrigerated Sorvall RT7 centrifuge in a RTH750 rotor. The pellet was resuspended in 50 mL ice-cold 50 mM $CaCl_2$ and then incubated on ice for 30 minutes. The cells were pelleted again at 1000 *g* for 10 minutes and resuspended in 2 mL of ice-cold 50 mM $CaCl_2$. The cells were incubated on ice in a 4ºC room for 2 hours or overnight before being used in a transformation reaction.

(ii) Transformation of E. coli

Plasmid DNA (4-200 ng) was added to 200 µL of competent cells and the mixture was incubated on ice for 30 minutes. The cells were heat shocked at 42ºC for two minutes and

then the tube was place on ice and incubated for between 15 and 30 minutes. LB broth was added (no antibiotic) to make up the total volume to 0.5 mL, and the cells were incubated at 37 ºC for between 30 minutes and one hour. 100 µL of the transformation reaction was plated on LB plates containing 25-34 µg/mL chloramphenicol and/or 100 µg/mL ampicillin (enabling positive selection of transformed cells) and incubated at 37ºC overnight. The pET151/D-TOPO® vector encodes an ampicillin resistance gene; the pOBT7 vector encodes a chloramphenicol resistance gene, as does the pRARE plasmid in the Rosetta strain.

(iii) Glycerol stocks

Glycerol stocks were made in 1 mL sterile cryotubes from selected transformed Rosetta and BL21 lines by mixing 500 µL of an overnight culture (5 mL LB broth plus 5 µL chloramphenicol (25-34 mg/mL) and/or 5 µL ampicillin (100 mg/mL)) with 500 µL sterile 40% glycerol. Glycerol stocks were stored at -80ºC.

2.2.8 Isolation of plasmid DNA from E. coli

(i) Rapid boil plasmid DNA preparation

A 1.5 mL volume of overnight culture (*E. coli)* was pelleted by centrifugation at 10 000 *g* for 1 minute. The supernatant was removed and the cell pellet resuspended in 350 µL of STET buffer (0.1 M NaC1, 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0, 5% Triton X-100) and 25 µL of freshly prepared lysozyme (10 mg/mL). The tubes were mixed and then placed in a boiling water bath for exactly 40 seconds. Immediately afterwards, centrifugation at 10 000 *g* for ten minutes was carried out. The pellet was removed and discarded. A 400 µL volume of isopropanol was added to the supernatant, mixed well, and the mixture allowed to stand at -20ºC for 30 minutes in order to precipitate the DNA. The sample was centrifuged again for five minutes at 10 000 *g* to pellet the plasmid DNA and the supernatant removed. A 500 µL volume of cold 95% ethanol was added to the pellet in a wash step and centrifugation for one minute at 12 000 rpm was performed. The supernatant was removed and the pellet allowed to air dry until all there was no trace of

ethanol. The pellet was resuspended in 50 µL of sterile water. Samples were stored at -20ºC.

(ii) Small-scale plasmid DNA preparation

Small-scale plasmid DNA preparations were used when a small quantity of high quality plasmid DNA was required for sequencing. In general, 2 mL of a 5 mL overnight culture was pelleted and the plasmid DNA isolated using a Quantum Prep® Plasmid Miniprep Kit according to the manufacturer's instructions. The kit uses a method based on that of alkaline lysis (Birnboim and Doly, 1979) to release plasmid DNA from the cell. Purification of the plasmid DNA takes place using a patented Quantum Prep matrix which is created from silicon dioxide exoskeleton of diatoms.

(iii) Medium-scale plasmid DNA preparation

Medium-scale preparations of plasmid were on occasion used to obtain sufficient pure plasmid DNA for an unambiguous sequencing reaction. The required culture was prepared by inoculating 100 mL of LB plus 100 µL of ampicillin (100 mg/mL) with 1 mL of a fresh 5 mL overnight culture of transformed TOP10 strain of *E. coli* (plasmid maintenance strain), and grown overnight at 37°C with shaking. Plasmid DNA was isolated using PureLink™ HiPure Plasmid Midiprep Kit (Invitrogen) according to the manufacturer's instructions. This system is based on a modified alkaline lysis method (Birnboim and Doly, 1979) and the plasmid DNA is subsequently purified by anion exchange resin. The resin is a patented ion-exchange moiety which provides high efficiency of DNA binding and high efficiency of separation of DNA from cellular contaminants.

2.2.9 Restriction endonuclease digests

Restriction endonucleases were used to characterise pET151/D-TOPO plasmid DNA. Approximately 2 µg of purified plasmid DNA was digested with 0.3 units of restriction endonuclease *Acc*I or *Eco*RV in the presence of the appropriate buffer (NE buffer 4 and NE buffer 2, respectively, as specified by the manufacturer) in a total reaction volume of 100 µL. The digest was incubated at 37°C for longer than one hour to facilitate a complete

digest. Samples were analysed for completeness of digestion by agarose gel electrophoresis.

2.2.10 DNA sequencing

Sequencing was performed by Ms Lorraine Berry at the Allan Wilson Centre genome sequencing service (Massey University). Plasmid DNA (600 ng) and 3.2 pmol of the appropriate primer were supplied in sterile water to make up 15 µL total reaction volume. Sequencing was performed using an ABI3730 Genetic Analyzer and the BigDye^{TM} Terminator Version 3.1 Ready Reaction cycle sequencing kit (Applied Biosystems Inc., Foster City, CA,USA).

2.2.11 Recombinant protein expression

LB broths of 50 mL, 100 mL, 500 mL or 1.0 L were inoculated with 10% v/v overnight culture grown in the presence of selective antibiotic. The cultures were grown at 37°C and shaken at 230 rpm to ensure adequate aeration. Once the optical density at A_{600} of approximately 0.6 had been reached, the temperature for expression was adjusted (expression temperatures ranged from 18°C to 37°C), and the culture allowed to equilibrate, before expression was induced. Expression was induced using isopropyl β-D-thiogalactoside (IPTG) concentrations ranging from 0.005 mM to 5 mM. Expressing cultures were grown for between three and 24 hours before the cells were harvested.

(i) Analysis of recombinant protein expression

Cell culture samples of 1 mL were collected before induction of expression, as a control, and every hour post-induction, for up to four hours, and at 18, 20 or 24 hours as appropriate. The cell pellet was harvested at 16 000 *g* for 10 minutes and the supernatant aspirated. The cells were solubilised in 20 µL of 10% SDS solution and 20 µL of 5× treatment buffer. The whole cell samples were analysed for recombinant protein expression by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

2.2.12 Inclusion body expression of recombinant protein

Protein expression was induced in 250 mL cultures with an A_{600} O.D. of approximately 0.6 at 30°C with 0.1 mM IPTG. The cultures were left to express for between 18 and 20 hours before the cells were harvested by centrifugation at 2000 *g* in a refrigerated Sorvall RT7 centrifuge using a RTH750 rotor.

(i) Purification of inclusion bodies

E. coli Rosetta cells harvested from 500 mL of culture were lysed in 35 mL of lysis buffer (50 mM Tris-HCl, pH 8.5, 150 mM NaCl, 2 mM β-mercaptoethanol, 1 mg/mL lysozyme, 1 Complete™ Mini EDTA-free protease inhibitor cocktail tablet) and left to incubate for 20 minutes at 4°C. The cell solution was lysed completely by using the French press three times under a pressure of 5000 psi. This was followed by incubation in the presence of DNase I and 5 mM $MgCl_2$ for 20 minutes at room temperature. Soluble protein was separated from insoluble protein, inclusion bodies and cell debris by centrifuging at 16 000 *g* for 40 minutes at 4°C. The pellet was then washed in three steps (shown below) in order to remove adsorbed impurities such as lipopolysaccharides, host cell proteins, plasmid, DNA, RNA polymerase and enzymes (Dasari *et al.*, 2008). To wash the collected inclusion bodies, the pellet was resuspended in wash buffer, briefly vortexed and recollected by centrifuging at 16 000 *g* for 10 minutes.

Wash 1 (pH 8.5)	**Wash 2 (pH 8.5)**	**Wash 3 (pH 8.5)**
50 mM Tris-HCl	50 mM Tris-HCl	50 mM Tris-HCl
150 mM NaCl	150 mM NaCl	2 M NaCl
2% Triton-X-100	2 mM CHAPS	-

Table 2.2 Inclusion body wash buffer constituents

At each washing stage, the inclusion body pellet was resuspended in the wash buffer followed by vortexing and a 10 minute 16 000 *g* centrifugation step.

(ii) Solubilisation of inclusion bodies

Subsequent to the washing procedure, the inclusion bodies were solubilised in solubilisation buffer (50 mM Tris-HCl, pH 9.0, 150 mM, NaCl 10 mM β-mercaptoethanol, 8 M urea) with gentle stirring for several hours at room temperature or at 4°C overnight. Post-solubilisation, the denatured protein solution was centrifuged at 16 000 *g* at 4°C for 30 minutes to remove insoluble debris.

(iii) Refolding recombinant human ADP-dependent glucokinase

Freshly solubilised protein was introduced to gently stirring refolding buffer (50 mM Tris-HCl, pH 8.0, 50 mM KCl) by direct dilution to give a final protein concentration of between 0.02 and 0.05 mg/mL.

2.2.13 Concentration of protein samples

Protein solutions of a greater volume than 50 mL were concentrated using a 350 mL Amicon stirred cell concentrator. A 30 kDa cut-off ultra-filtration membrane was used to prevent the loss of recombinant protein. Nitrogen gas was applied at a pressure of approximately 60 psi to concentrate the protein solution, while stirring. Smaller volumes of protein solution were concentrated using centrifugal filter units with 50 kDa or 30 kDa membrane cut-offs.

2.2.14 Recombinant protein purification

Protein purification was performed using the ÄKTA FPLC system. Refolded protein was dialysed into buffer A (50 mM Tris-HCl, pH 8.0, 50 mM KCl, 5 mM imidazole) ready for protein purification using a HisTrap™ HP 5 mL affinity column. The column was first equilibrated with five column volumes of buffer A before the dialysed protein solution was loaded at 1 mL/min, via the sample pump, directly onto the column. Once all protein had been loaded, the column was washed with two column volumes of buffer A (2.5 mL/min) to remove any unbound protein. The purification process followed three imidazole steps using an increasing percentage of buffer B (50 mM Tris-HCl, pH 8.0, 50 mM KCl, 500

mM imidazole) as follows: 8 column volumes of 10% buffer B, 6 column volumes of 40% buffer B and 5 column volumes of 100% buffer B. The recombinant human ADP-GK was eluted at 40% buffer B (200 mM imidazole). The eluted purified protein was immediately dialysed into the original refolding buffer (50 mM Tris-HCl, (pH 8.0), 50 mM KCl).

2.2.15 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is a technique which separates proteins by gel electrophoresis, based on their relative molecular mass. This is achieved by first treating the protein samples to be analysed so that they are fully denatured. Treatment buffer is made up of SDS, an anionic denaturing detergent that disrupts almost all non-covalent interactions, and dithiothreitol (DTT) or β-mercaptoethanol to reduce disulfide bonds. Anions of SDS surround and coat the polypeptides of a protein, causing the protein to unfold. The negative charges carried by the SDS molecules confer a negative charge on the protein proportional to the size of the protein. Therefore the migration of the protein through the gel in the presence of an electrical field is relative to its size and not the charge it carries due to its amino acid sequence.

(i) Gel casting

Glass spacer plates (0.75 mm or 1.5 mm) and short plates were cleaned using isopropanol and lint-free wipes and placed in a casting frame. The frame was put in the gel casting stand ready for pouring an acrylamide gel. Resolving and stacking gel solutions (per two 0.75 mm gels) were prepared as outlined in Table 2.3.

	10.5% Polyacrylamide		8% Polyacrylamide	
Components	**Resolving gel**	**Stacking gel**	**Resolving gel**	**Stacking gel**
1.5 M Tris-HCl, pH 8.8	**3.75 mL**	**-**	**3.75 mL**	**-**
0.5 M Tris-Cl, pH 6.8	**-**	**1.25 mL**	**-**	**1.25 mL**
Acrylamide-bis ready-to-use solution, 40%	**3.54 mL**	**1.023 mL**	**2.7 mL**	**838 µL**
MilliQ water	**6.06 mL**	**2.67 mL**	**6.9 mL**	**2.86 mL**
10% SDS (w/v)	**150 µL**	**50 µL**	**150 µL**	**50 µL**
13% APS*	**75 µL**	**25 µL**	**75 µL**	**25 µL**
TEMED*	**7.5 µL**	**5 µL**	**7.5 µL**	**5 µL**

Table 2.3 Resolving and stacking gel solutions for SDS-PAGE

* APS and TEMED were added to the solution immediately before casting.

Immediately before the resolving gel was poured into the casting apparatus, APS and TEMED were added to the solution to initiate polymerisation. The solution was carefully mixed in order to prevent the introduction of air bubbles, and approximately 3.4 mL was dispensed by pipette between the glass plates. The resolving gel layer was overlaid with isopropanol to exclude air from the gel and to ensure the resolving boundary of the gel set level and smooth. The resolving gel was left to polymerise for a minimum of 20 minutes at room temperature before the isopropanol was washed off using MilliQ water. All water was carefully removed with blotting paper. APS and TEMED were then added to the stacking gel solution, which was mixed carefully, and approximately 1 mL was dispensed by pipette on top of the polymerised resolving gel. A well forming comb (10 or 15 wells) was placed into the unset stacking gel without introducing bubbles, and was left to polymerise for at least 10 minutes. Once set, the comb was carefully removed and any extraneous stacking gel left behind by the comb was washed out with MilliQ water to ensure the wells were clear.

(ii) SDS-PAGE

SDS-PAGE was undertaken using Mini-Protean 3 cell apparatus. Gels were removed from the casting frame and placed into the cells according to the manufacturer's instructions. In order to ensure the apparatus had been assembled properly, the inner cell was filled 1× electrode buffer (25 mM Tris, 192 mM glycine, 0.5% SDS, pH 8.3) and left for a few minutes to check for any leaks. The outer tank was then filled to approximately a third of its volume with 1× electrode buffer. Protein samples to be analysed by SDS-PAGE were denatured by the addition of 5× treatment buffer (60 mM Tris-HCl, pH 6.8, 25% glycerol (v/v), 2% SDS, 0.1% bromophenol blue, 14.1 mM β-mercaptoethanol) diluted five-fold in the sample solution and heat treatment by placement of samples in a boiling water bath for 5 minutes. After boiling, the samples were left to cool to room temperature, vortexed to mix, and briefly centrifuged to collect the sample at the bottom of the tube. Samples were loaded into wells using protein loading tips and 4 µL of Precision Plus Protein™ Dual Colour Standards were loaded into end wells. Electrophoresis was carried out at 120 V for approximately 120 minutes or until the dye front had just migrated off the bottom of the gel.

(iii) Coomassie staining of polyacrylamide gels

In order to detect and analyse separated proteins, gels were stained for 15 minutes at room temperature, with shaking, using Coomassie R-250 stain (50% methanol, 10% glacial acetic acid, 0.25% Coomassie Brilliant Blue R-250). Stain was poured off and excess stain rinsed away with water followed by incubating in de-stain (15% methanol, 8% acetic acid) at room temperature with shaking for several hours or overnight.

(iv) Silver-staining of polyacrylamide gels

At the completion of SDS-PAGE, the gel was removed and soaked in fixing solution (50% ethanol, 10% acetic acid) for between 30 minutes and one hour, at room temperature, with shaking. The gel was subsequently washed, twice, in MilliQ water for five minutes. The gel was then soaked in 0.8 mM sodium thiosulfate solution for two minutes and this was followed by another five minute wash in MilliQ water. The gel was incubated in freshly prepared silver nitrate solution (2 g/L silver nitrate) for between 15 and 20 minutes. Two

strictly 30 second MilliQ washes were performed and the gel was placed in developing solution (6 g/mL sodium carbonate, 0.05% (v/v) formaldehyde, 0.016 mM sodium thiosulfate) for between 2 and 10 minutes, or until sufficient development of protein bands. The developing reaction was stopped by transferring the gel into cold 5% acetic acid for at least 15 minutes.

2.2.16 Immunoblotting

After SDS-PAGE, proteins were transferred from the acrylamide gel to a membrane and an antibody was used to detect the ADP-GK protein.

(i) Transfer of proteins

The Mini-Protean 3 cell system was also utilised for the transfer of proteins from a polyacrylamide gel to a positively charged nylon membrane for immunoblotting. The set-up for transfer was as follows: a sponge soaked in transfer buffer (25 mM Tris, 192 mM glycine, pH 8.3) was placed into the transfer cassette. Two pieces of 3 MM Whatmann® paper completely soaked in transfer buffer were placed on top of the sponge, being careful to eliminate any air trapped between the layers. The gel, having also been equilibrated in transfer buffer, was placed on the Whatmann® paper in the correct orientation, also without introducing air bubbles. The equilibrated nylon membrane was placed on top of the gel, making sure the edges met the edges of the gel and that all air bubbles were excluded. This was followed by two more pieces of Whatmann® paper and another sponge, all soaked in transfer buffer. The transfer cassette was closed and inserted into the transfer apparatus, ensuring it was in the correct orientation so the proteins migrated in the right direction. An ice block was placed into the tank to keep the system cold during the transfer. Electroblotting was carried out at 450 mA for 45 minutes.

(ii) Immunoblotting

After transfer, the set-up was carefully disassembled and the nylon membrane soaked in 5% non-fat skim milk powder in Tris-buffered saline-Tween 20 (TBST) (50 mM Tris, 150 mM NaCl, 0.1% Tween-20) also known as 5% blocking solution, overnight at 4°C, with

shaking. This was followed by incubating the blocked membrane at room temperature for three hours in 0.5% blocking solution containing the primary antibody at the appropriate dilution. The membrane was washed with shaking for 10 minutes, twice with TBST and twice with 0.5% blocking solution. Incubation in the presence of the appropriate horseradish peroxidase-conjugated secondary antibody diluted in 0.5% blocking solution was then performed for 45 minutes to one hour, with shaking, at room temperature. Lastly, the membrane was washed four times for 10 minutes with TBST at room temperature.

Horseradish peroxidase-conjugated secondary antibody was detected by enhanced chemiluminescence. Chemiluminescence blotting substrate A and B were mixed at a ratio of 100:1 and the membrane was incubated in the mixture for three minutes. The membrane was removed from the blotting substrate and excess solution was allowed to drip off before the membrane was placed between two clear plastic sheets in an X-ray cassette. Light produced by the reaction was detected in a dark room by X-ray film. The X-ray films were developed using a 100Plus^{TM} Automatic X-ray film developer.

2.2.17 Protein quantification

The protein concentration of total protein extracts prepared from porcine tissue samples were quantified using a method that is based on the assay described by Bradford (1976). This method is based on the shift in the absorbance of Coomassie Brilliant Blue G-250 from 465 nm to 595 nm on its binding to protein. The increase in absorbance at 595 nm can be measured and used to calculate protein concentrations of unknown samples. A series of standards ranging from 0 µg to 2.0 µg of protein were prepared in triplicate using a 10 mg/mL stock solution of BSA and dispensed into wells of a 96-well microplate. Protein samples to be measured were diluted appropriately in MilliQ water and 5 µL applied to the microplate, also in triplicate. Protein Assay dye reagent concentrate was diluted 5-fold in MilliQ water and 200 µL was added to each well containing protein solution. The plate was left for 10 minutes at room temperature to allow colour to develop. The absorbance at 595 nm was measured using an Anthos LabTec HT2 plate reader. A standard curve was constructed using the absorbance values of the standard samples and their known

concentrations. The concentration of the unknown samples was determined using a standard curve.

2.2.18 Circular dichroism

Circular dichroism (CD) is an important technique for analysing the structure of proteins in solution. Circular dichroism refers to the differential absorption of left-handed circularly polarised light and right-handed circularly polarised light, the two components of plane polarised light. Absorption in the region of 240 nm and below is due principally to protein peptide bonds and different types of regular secondary structure give rise to characteristic CD spectra in this region (known as far UV).

Pure protein to be analysed by circular dichroism was dialysed into a minimal buffer. In this case 10 mM K_2HPO_4 and 20 mM KCl was used, as buffer components can interfere by absorbing at wavelengths of interest. The final protein concentration was required to be between 0.1 and 0.5 mg/mL as a 0.1 cm high-transparency quartz cuvette was used to collect spectral data in the Chirascan™ Circular Dichroism Spectrometer. The wavelength to be scanned was set between 180 nm and 260 nm with 0.5 nm steps, and the bandwidth at 1.0 nm. The time taken to measure absorbance at each point was set at 0.5 seconds and the scan performed in triplicate. The absorbance spectrum of the buffer alone was measured first, in order for background absorbance to be subtracted from the spectra of the protein solutions. Data was visualised, averaged, smoothed and overlaid using the modular Pro-Data instrument software. This is specific for the Chirascan™ Circular Dichroism Spectrometer, which provides overall instrument control and data manipulation post-acquisition (http://www.photophysics.com/pdf/chirascaninfo.pdf).

2.2.19 Enzyme activity assays

ADP-dependent glucokinase catalyses the reaction:

$$\text{ADP + Glucose} \longrightarrow \text{Glucose-6-phosphate + AMP}$$

This reaction is linked with the enzyme glucose-6-phosphate dehydrogenase (*Leuconostoc mesenteroides)* as follows:

$$\text{Glucose-6-phosphate} + NAD^+ \longrightarrow \text{Gluconate-6-phosphate} + NADH$$

The production of NADH can be measured at 340 nm.

The standard assay conditions as described by Ronimus and Morgan (2004), were used to assay both the recombinant mouse ADP-GK and recombinant human ADP-GK in a temperature controlled Pharmacia Biotech Ultraspec® 3000 UV/visible spectrophotometer (37ºC).

Assay Components (100 µL total assay volume)
50 mM MOPS (pH 7.25)
100 mM KCl
0.35 mM Glucose
1 mM ADP
1 mM $MgCl_2$
1.0 mM NAD^+
MilliQ water
0.2 Units Glucose-6-phosphate dehydrogenase

Table 2.4 Standard ADP-GK assay components

A master mix for ten reactions was made up freshly before each set of assays and included all components listed, apart from the glucose-6-phosphate dehydrogenase and ADP-dependent glucokinase enzymes. A tenth of the master mix was dispensed into the 100 µL microcuvette, followed by 0.2 units of glucose-6-phosphate dehydrogenase. The assay was initiated by the addition of the ADP-dependent glucokinase, with rapid mixing.

All enzyme assays were performed in triplicate at 37°C using 100 µL microcuvettes. The standard assays contained 1.0 mM NAD^+, 100 mM KCl, a 1:1 ratio of ADP:$MgCl_2$ (1 mM), and 50 mM MOPS buffer. To each assay was added 0.2 units of glucose-6-phosphate dehydrogenase, and the assay was initiated by addition of the recombinant ADP-GK. The assay was monitored for 120 seconds with an approximate 30 second pre-incubation.

2.2.20 Extraction of total cellular protein from tissue

Tissues to be screened for the expression of the ADP-GK protein were collected from pigs freshly slaughtered at a local abattoir and transported on ice to the laboratory. Approximately 1 g samples of tissue were homogenised in 10 mL of cold sucrose buffer (20 mM HEPES, pH 7.4, 0.25 M sucrose, 1 mM EDTA, Complete mini-protease inhibitor) using an Ultra-Turrax® T50 homogeniser. The homogenate was centrifuged at 16 000 *g* for 30 minutes at 4°C and the supernatant collected. The resulting pellet was resuspended in 0.1 M Tris-HCl, pH 7.8, 0.5 % (v/v) Triton X-100. Protein concentration was quantified by Bradford protein assay and the cytoplasmic and membrane protein extracts were dispensed in 1 mL aliquots in fresh micro-centrifuge tubes and stored at -20°C.

Chapter 3: Cloning and expression of recombinant human ADP-dependent glucokinase

3.1 Introduction

This chapter describes the cloning of the full-length human ADP-GK cDNA sequence into a pET151/D-TOPO expression vector and the expression trials undertaken in order to optimise the soluble expression of the recombinant human ADP-GK enzyme.

The full length coding sequence was used to express the human ADP-GK as the purpose of the project was to kinetically characterise the recombinant human ADP-GK enzyme and compare it to the kinetic data obtained from the recombinant mouse ADP-GK. The recombinant mouse ADP-GK was also expressed from the full-length coding sequence, therefore a direct comparison could be made between the two enzymes.

The mouse enzyme had been expressed with the polyhistidine tag on the C-terminal, however, the pET151/D-TOPO expression vector used in this study meant the polyhistidine tag was present at the N-terminus of the recombinant human ADP-GK. As the globular region of the ADP-GK is predicted to be at the C-terminal of the protein, it was considered the N-terminal polyhistidine tag could be an advantage, making it less likely to cause any interference with protein folding.

3.2 Cloning strategy

Cloning reactions were carried out using Invitrogen's Champion™ pET Directional TOPO® Expression Kit. This system was selected because the cloning process is directional and PCR products can be cloned directly, as Invitrogen's TOPO® Cloning technology utilises topoisomerase I from the virus *Vaccinia.*

Topoisomerase I from *Vaccinia* binds to duplex DNA and forms a cleavable complex at its 5′- CCCTT recognition sequence. The topoisomerase cleaves the phosphodiester backbone of the single-strand containing the 5′-CCCTT sequence and conserves the energy from the broken phosphodiester bond in the formation of a covalent bond between the 3′ phosphate of the cleaved strand and a topoisomerase I tyrosyl residue. The reaction is reversed when the phospho-tyrosyl bond between the DNA strand and the enzyme is subsequently attacked by the 5′ hydroxyl of the original cleaved strand, thereby releasing the topoisomerase (Shuman, 1991).

Invitrogen supplies vector DNA with a four nucleotide single-stranded sequence (GTGG) which is charged with *Vaccinia* topoisomerase I (TOPO®-charged). In order for cloning to take place, a 3′ sequence must be added to the incoming DNA (i.e. PCR product to be cloned) by using a forward primer with four additional bases (CACC). This creates a template for the complementary vector single-stranded sequence to invade. The four nucleotide sequence in the cloning vector invades the 5′ end of the PCR product, annealing to the added bases, and in so doing, stabilising the PCR product in the correct orientation (Figure 3.1).

A pET151/D-TOPO® vector (Appendix 1) was used, and this vector enables the expression of recombinant proteins with an N-terminal tag containing the V5 epitope and a 6× His tag. The V5 epitope is derived from the epitope present on the P and V proteins of the simian virus (SV5). It is a commonly included epitope on expression vectors, and can be used in the immunoprecipitation or immunodetection of the expressed recombinant protein as there are antibodies to this epitope available commercially. The N-terminal leader sequence also includes a TEV (Tobacco Etch Virus) protease cleavage site to enable removal of the tag after protein purification using TEV protease.

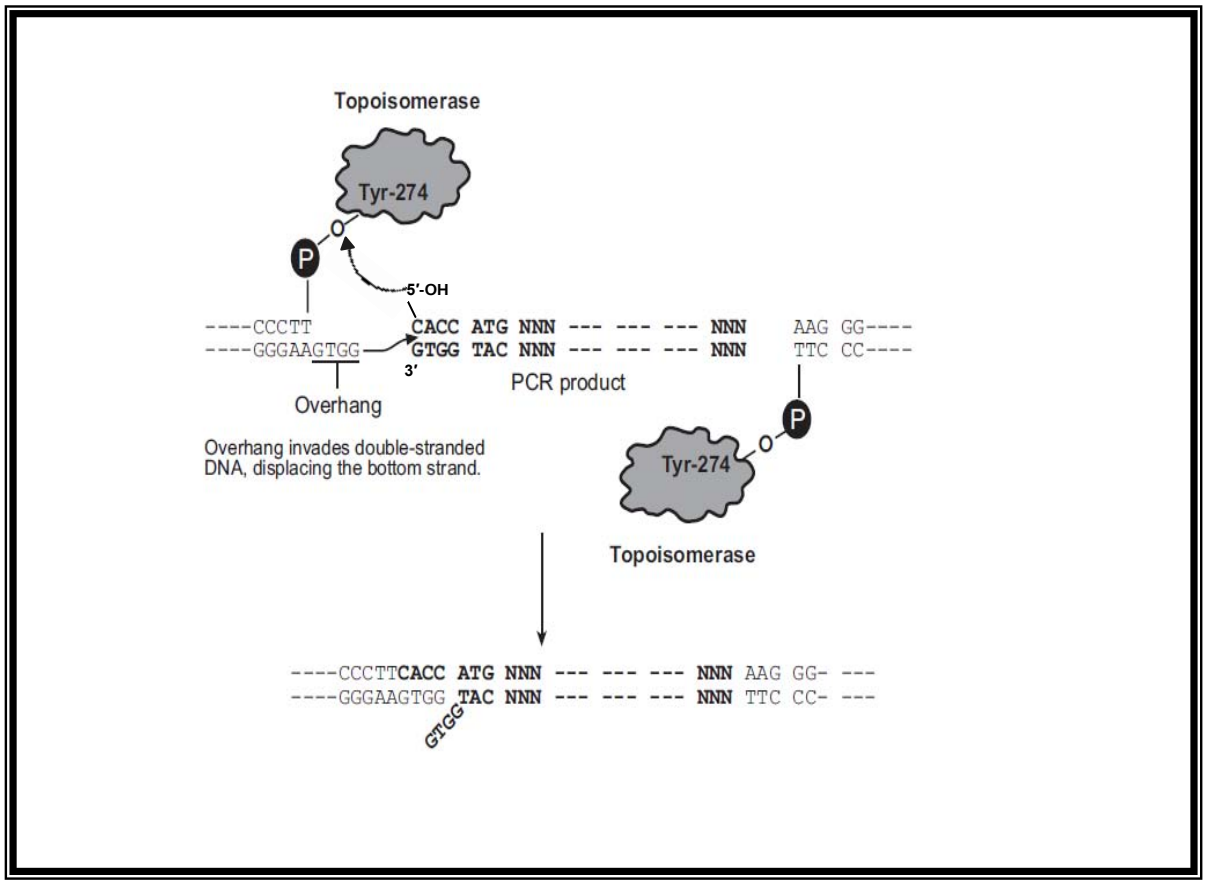


Figure 3.1 Schematic of the Invitrogen TOPO® Cloning technology

Vector DNA is supplied with a four nucleotide single-stranded sequence (GTGG) which is charged with *Vaccinia* topoisomerase via a covalent bond between the 3′ phosphate of the backbone of a single strand of the vector and tyrosine-274 residue of the topoisomerase. The sequence in the cloning vector is complementary to the four base pair sequence added to the PCR product to be cloned (CACC) and invades the 5′ end of the PCR product, annealing to the added bases, and in so doing, stabilising the PCR product in the correct orientation. The reaction is reversed when the phospho-tyrosyl bond between the DNA strand and the enzyme is subsequently attacked by the 5′ hydroxyl of the remaining free strand, thereby releasing the topoisomerase (Shuman, 1991).

3.2.1 Amplification of the human ADP-GK coding sequence by PCR

PCR reactions were carried out using *Pfu Turbo*® polymerase as the product was to be used directly for cloning and *Pfu Turbo* polymerase has a reported error rate six-fold lower than that of *Taq* polymerase (Cline *et al*., 1996).

Forward (RH1) and reverse (RH2) primers were designed to the regions of the start and stop codon of the human ADP-GK coding sequence, to produce an amplicon of 1 499 base pairs. The appropriate addition required for TOPO cloning was added to the forward primer, as indicated in Figure 3.2. The native stop codon was changed from UAG to UAA which is the most commonly used stop codon in *Escherichia coli.*

RH1 forward primer	5′ **CACC**ATGGCGCTGTGGCGCGG 3′
RH2 reverse primer	5′ TTAATAGTGAGGGTGTACTTCC 3′

Figure 3.2 Primer sequences

Forward (RH1) and reverse (RH2) primers were designed to the start and stop codon regions of the coding sequence of the human ADP-GK gene, respectively.

Several annealing temperatures were tested in an attempt to obtain the human ADP-GK coding sequence. The reaction, described in Table 3.1, was eventually successful when an annealing temperature of 61°C was used (Figure 3.3).

Human ADP-GK coding sequence amplification		
	95°C for 3.0 min	
Denaturing	95°C for 30 sec	20× cycles
Annealing	61°C for 30 sec	
Extension	72°C for 90 sec	
Final extension	72°C for 10.0 min	

The PCR mixture contained 1 µL (250 ng/µL) of each primer RH1 and RH2, 1 µL (10 mM) dNTPs, 5 µL 10× PCR buffer (200 mM Tris-HCl, pH 8.8, 20 mM $MgSO_4$, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 1% Triton X-100, 1 mg/mL nuclease-free BSA, Stratagene), 2.5 µL of DMSO, 1.0 µL of *PfuTurbo*® polymerase (2.5 U/µL), 1.0 µL of the template DNA (100 ng/µL) and 37.5 µL of sterile water to make the total volume 50 µL.

Table 3.1 PCR protocol for amplifying human ADP-GK coding sequence

The PCR reaction used to amplify the human ADP-GK coding sequence with a 5′ four base pair addition (CACC).

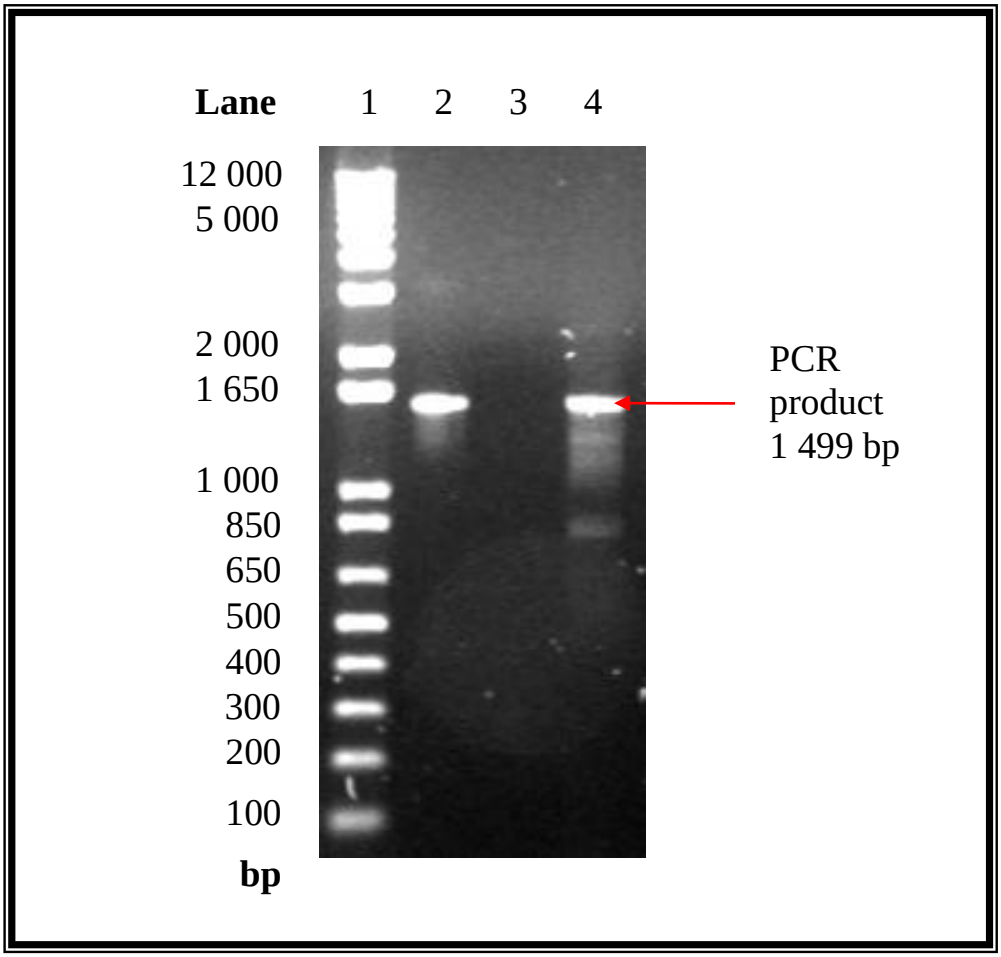


Figure 3.3 PCR reaction to generate fragment for TOPO® cloning

Fifteen microlitres of each PCR reaction was separated on a 1% agarose gel in 1× TAE buffer by electrophoresis at 100 V for approximately one hour. DNA was visualised by the addition of ethidium bromide (0.5 µg/mL) in the gel and the running buffer and exposure to UV light. Ten microlitres of 1 kb plus ladder was loaded in lane one. The sizes of the molecular weight markers are shown on the left in base pairs. The PCR was performed using RH1 and RH2 primers (lane four). Lane two contains the gel-purified PCR product (1 499 bp). A negative control reaction was performed and is shown in lane three.

3.3.2 Cloning and transformation of Escherichia coli strains

Cloning of the human ADP-GK coding sequence with the CACC Topo I site gained by the PCR reaction was performed using a Champion™ pET Directional TOPO® Expression Kit as described in Section 2.2.6. Both *E. coli* BL21(DE3) and Rosetta strains were transformed using chemically competent cells and heat shock treatment as described in Section 2.2.7. Transformed BL21(DE3) or Rosetta cells were selected by plating on LB-ampicillin plates or LB-ampicillin/chloramphenicol plates, respectively. Plasmid DNA was isolated from positive clones by the rapid boil method as outlined in Section 2.2.8.

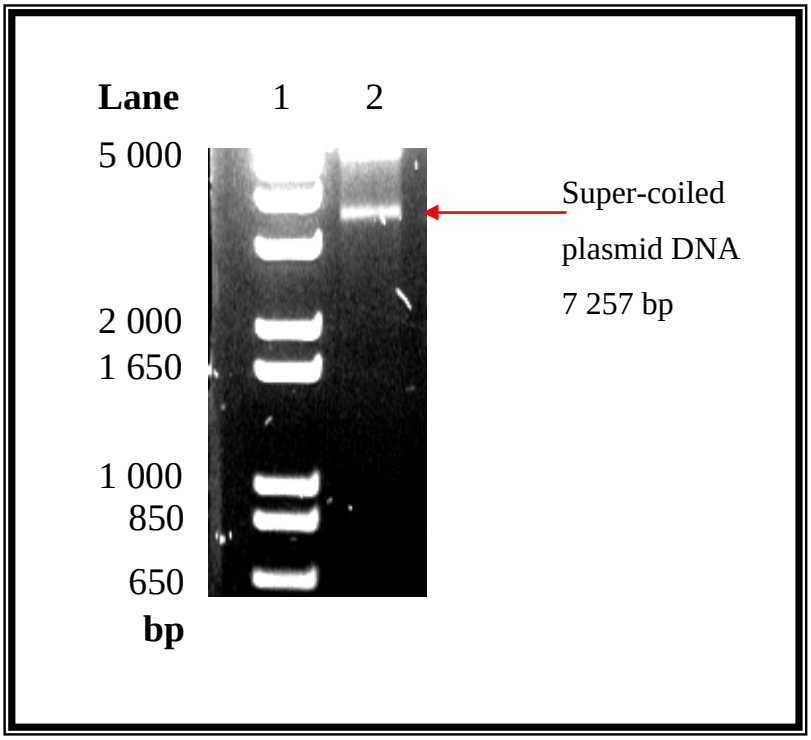


Figure 3.4 Plasmid DNA isolated from a positive clone

Plasmid DNA was isolated from positive clones chosen from antibiotic selection plates and grown up in overnight broths. The rapid boil method was used to isolate the plasmid DNA and the result was visualised by gel electrophoresis performed on a 1% agarose gel.

Lane one: 1 kb plus ladder (10 µL); **Lane two:** 1:2 dilution of plasmid DNA from a positive clone (5 µL).

3.3 Confirmation of cloning strategy

In order to confirm that the positive colonies chosen from the antibiotic selection plates did have the human ADP-GK cDNA sequence insert in the pET151/D-TOPO plasmid, a further PCR reaction was undertaken. This reaction used the T7 forward primer and the RH2 reverse primer in order to include the 5′ T7 promoter region, *lac* operator, ribosome binding site, polyhistidine region, V5 epitpope and TEV recognition site in a product of ~1 680 bp. As this reaction used the T7 forward primer, a new set of conditions had to be optimised.

T7 forward and RH2 reverse clone sequence amplification		
	95°C for 3.0 min	
Denaturing	95°C for 40 sec	20× cycles
Annealing	53°C for 45 sec	
Extension	72°C for 1.68 min	
Final extension	72°C for 10.0 min	
The PCR mixture contained 1 µL (250 ng/µL) of each primer T7 and RH2, 1 µL (10 mM) dNTPs, 5 µL 10× PCR buffer (200 mM Tris-HCl, pH 8.8, 20 mM $MgSO_4$, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 1% Triton X-100, 1 mg/mL nuclease-free BSA, Stratagene), 2.5 µL of DMSO, 1.0 µL of *PfuTurbo*® polymerase (2.5 U/µL), 1.0 µL of the template DNA (100 ng/µL) and 37.5 µL of sterile water to make the total volume 50 µL.		

Table 3.2 PCR protocol for amplifying ~1 680 bp product utilising T7 forward primer and RH2 reverse primer

A range of annealing temperatures were used (53-56°C inclusive) in an attempt to optimise the PCR conditions. It was found that products were successfully obtained at 53°C and 53.6°C, consistently, from three of the four colonies chosen to be screened. One colony was not positive for the predicted product under any of the temperature conditions tested.

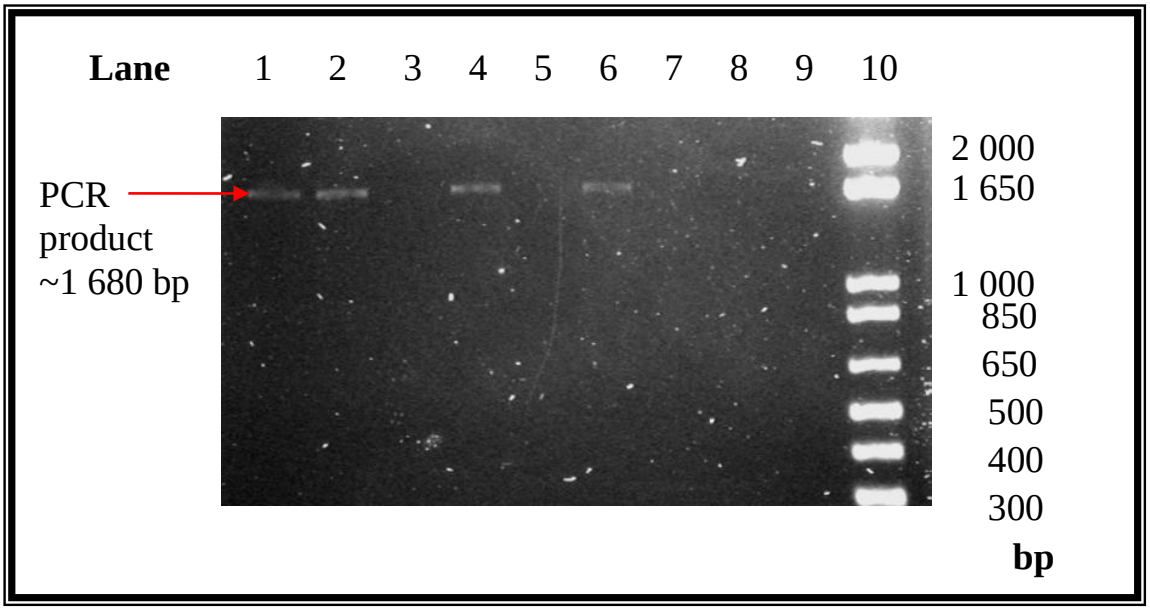


Figure 3.5 PCR product produced from pET151/D-TOPO plasmid isolated from positive clones using T7 forward primer and RH2 reverse primer

Four positive clones were chosen, from the antibiotic selection plates, to be screened for the cloned human ADP-GK cDNA sequence by PCR. Fifteen microlitres of PCR product obtained from each colony, with the stated annealing temperature, was loaded in lanes one through to nine. Separation took place on a 1% agarose gel in 1× TAE buffer by electrophoresis at 100 V for approximately one hour. DNA was visualised by the addition of ethidium bromide (0.5 µg/mL) in the gel and the running buffer, and exposure to UV light.

Lane one: Colony 1 (53°C); **Lane two:** Colony 2 (53°C); **Lane three:** Colony 3 (53°C); **Lane four:** Colony 4 (53°C); **Lane five:** Colony 3 (56°C); **Lane six:** Colony 2 (53.6°C); **Lane seven:** Colony 3 (54.1°C); **Lane eight:** Colony 4 (54.1°C); **Lane nine:** Colony 4 (54.4°C); **Lane ten:** Ten microlitres of 1 kb plus ladder. The sizes of the molecular weight markers are shown on the right in base pairs.

3.3.1 Restriction endonuclease digests

As further confirmation that the sequence of the human ADP-GK cDNA had been cloned into the pET 151/D-TOPO plasmid, a restriction endonuclease digest was performed. Approximately 2 µg of purified plasmid DNA was digested with 0.3 units of restriction endonuclease *Acc*I or *Eco*R V (as described in Section 2.2.9). The size of each visible band was determined by reference to a 1 kb plus ladder. Two bands were expected for each digest: The *Acc*I digest was expected to produce bands of 4 454 bp and 2 803 bp. The *Eco*R V digest was expected to produce bands of 4 237 bp and 3 020 bp. The result of the restriction endonuclease digests are shown in Figure 3.7. The appropriately sized restriction endonuclease fragments were obtained. This supported evidence that the human ADP-GK cDNA had been successfully cloned.

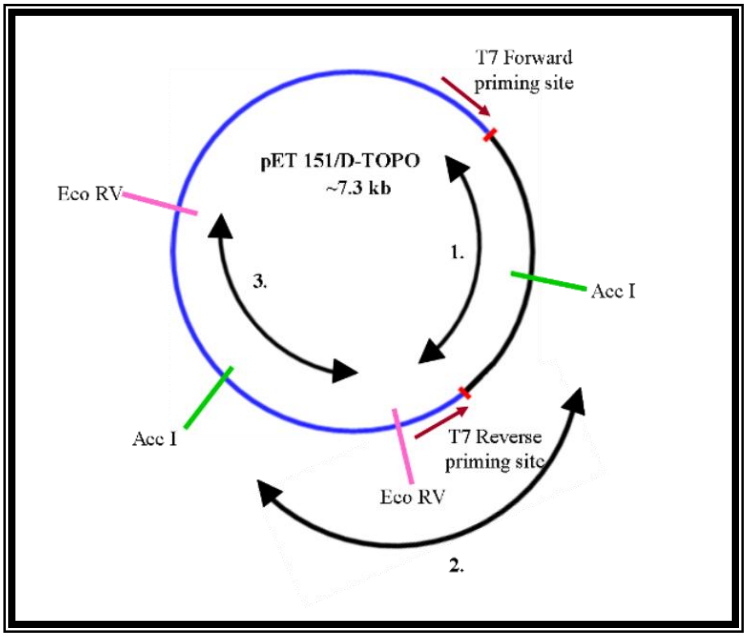


Figure 3.6 Schematic representation of the restriction digest performed on pET151/D-TOPO containing the cloned cDNA sequence of the human ADP-GK gene

1. Cloned cDNA sequence of the human ADP-GK gene (1497 bp)
2. *Acc*I digest (4454 bp, 2803 bp)
3. *Eco*RV digest (4237 bp, 3020 bp)

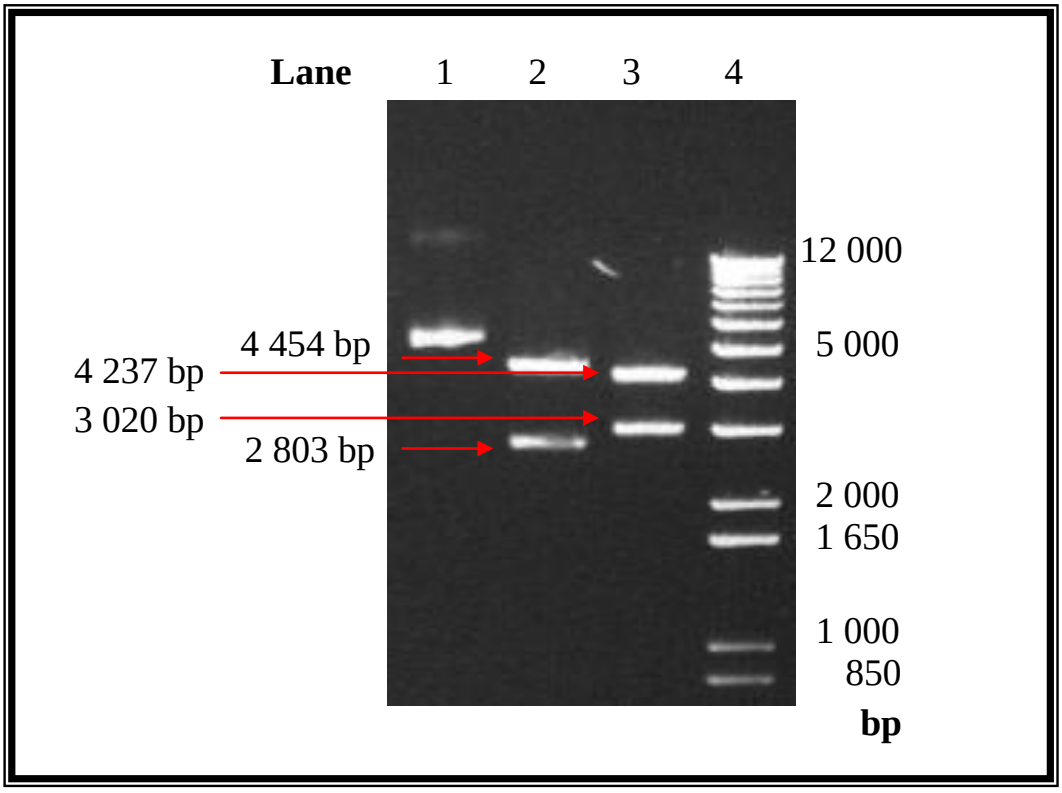


Figure 3.7 Digestion of pET151/D-TOPO plasmid containing the cDNA sequence of the human ADP-GK gene

Plasmid DNA was purified from positively colonies grown in a 100 mL culture overnight at 37°C with shaking. The plasmid was purified using a PureLink™ HiPure Plasmid Midiprep Kit (Invitrogen). The undigested and digested plasmid DNA was separated by electrophoresis on a 1% agarose gel in 1× TAE buffer at 100 V for approximately one hour. The bands were visualised by exposure to UV light due to the incorporation of ethidium bromide in the agarose gel and running buffer. The sizes of the molecular weight markers are shown on the right, in base pairs (bp).

Lane one: 200 ng of undigested, purified plasmid DNA in its supercoiled state; **Lane two:** 15 µL of the *Acc*I digestion reaction; **Lane three:** 15 µL of the *Eco*RV digestion reaction; **Lane four:** 10 µL of a 1 kb plus ladder.

3.3.2 DNA sequencing

Recombinant protein expression can be severely affected by any changes in the DNA sequence of the protein to be expressed, particularly if a change results in the recombinant protein sequence moving out of frame. For this reason, DNA sequencing was performed on the pET151/D-TOPO plasmid sample which had been shown to be positive for the human ADP-GK cDNA sequence by PCR and restriction digest. The sequencing reaction used the T7 forward primer and the RH2 reverse primer so it could be demonstrated that the cDNA sequence was in frame and the polyhistidine tag was intact. In addition, two other pairs of internal primers were used to ensure complete sequencing of the DNA between the T7 and RH2 primers (Figure 3.8)

FWD1 forward primer	5′ TTGCCCAGGTTGCGTCAGAG 3′
REV2 reverse primer	5′ CTGAAACTCCTCCAGGCT 3′
RH3 forward primer	5′ GGCCATGAATATGCTGGAGGT 3′
RH4 reverse primer	5′ TCCCTGTTAGTCATACTGGCC 3′

Figure 3.8 Primer sequences

Forward (FWD1) and reverse (REV2) primers were designed to produce an amplicon of 338 base pairs. Forward (RH3) and reverse (RH4) primers were designed to produce an amplicon of 201 base pairs.

As the plasmid with the cDNA sequence included was quite large, (~7.3 kb) it was found that in order to get unambiguous sequencing results, twice as much plasmid DNA than was recommended needed to be submitted in the sequencing reaction mixture. To obtain enough plasmid DNA for sequencing, it was necessary to use the PureLink^{TM} HiPure Plasmid Midiprep Kit as described in Section 2.2.8.

Sequencing was performed by Lorraine Berry (Allan Wilson Centre Genome Sequencing Service). The sequence of the product obtained from the T7 and RH2 primers was aligned with the sequence of the pET151/D-TOPO plasmid carrying the human ADP-GK cDNA

sequence (NCBI nucleotide database, accession number BC006112), using the unix-1 alignment tool BestFit. There were no mismatches. A section of the positive sequencing chromatogram is presented in Appendix 3.

3.4 Recombinant protein expression

The BL21(DE3) strain of *E. coli* was chosen as the bacterial expression strain for initial recombinant expression trials. BL21(DE3) is a standard bacterial strain used for expression studies and it functions by utilising the T7 expression system. BL21 cells have been engineered to carry a gene for T7 RNA polymerase on a specially designed DE3 bacteriophage lambda lysogen. The T7 RNA polymerase is vital to this expression system as it is specific for a T7 promoter, and this is the promoter that controls the expression of, in this case, the recombinant protein carried on the pET151/D-TOPO expression plasmid. Expression of the T7 RNA polymerase is controlled by the *lac* repressor, and this is encoded by the *lac I* gene, which is found on the bacteriophage lambda lysogen. The *lac* repressor, under normal circumstances, binds to the *lac* operator and this inhibits expression of the T7 RNA polymerase. Since there is no T7 RNA polymerase being expressed, there is, therefore, no expression of recombinant protein. When the inducer of this T7 expression system (IPTG) is added to a BL21 bacterial expression culture, it binds to the *lac* repressor protein and changes its conformation in such a way that is can no longer bind to the *lac* operator. This means expression of the T7 RNA polymerase will take place, which in turn means that expression of the recombinant protein, carried by the expression plasmid (pET151/D-TOPO), will be induced.

Protein expression trials began with basic conditions that were varied only in terms of the final concentration of IPTG used to induce expression of the recombinant protein in culture, and the temperature of the cultures during the period of recombinant protein expression. To begin with, 20 mL cultures were trialled for expression studies. Positively transformed BL21(DE3) cells were used to inoculate 5 mL LB broths containing 100 µg/mL ampicillin in order to maintain the pET151/D-TOPO expression plasmid. The inoculated LB broths were grown overnight at 37°C with shaking. 1 mL of overnight culture was used to

inoculate 20 mL LB broths, also containing antibiotic. Cultures were grown at 37°C, with shaking, until an optical density at A_{600} of approximately 0.6 was reached.

The first trials screened expression conditions of 25°C and 37°C using final concentrations of IPTG of 0.1 mM and 0.5 mM. 1 mL samples of culture were collected before the induction of protein expression and every hour post induction for three hours. Samples were treated with 20 µL of 10% SDS and 20 µL of 5× treatment buffer (60 mM Tris-HCl, pH 6.8, 25% glycerol (v/v), 2% SDS, 0.1% bromophenol blue, 14.1 mM β-mercaptoethanol) to produce whole cell protein lysates. The whole cell protein lysates were used to monitor recombinant protein expression over time by SDS-PAGE. Figure 3.9 is an example of the 37°C temperature trial using 0.1 mM (a) and 0.5 mM (b) IPTG. The pre-induction sample was compared to the post-induction samples on the protein gel and a band representing recombinant protein (with an apparent molecular weight of approximately 51 kDa) was visible from one hour post induction. The calculated molecular weight of the recombinant human ADP-GK based on amino acid composition is 57.75 kDa, but the apparent molecular weight can be considered close to the calculated value, within the resolving power of the gel.

In the trial using 0.5 mM IPTG, it appeared that expression levels increased over the three hours monitored. The 25°C expression trial showed a much lower level of recombinant protein expression over the same time period. The lower temperature slowed down the expression of recombinant hADP-GK noticeably (data not shown). In order to determine whether any recombinant protein had been expressed as soluble protein, samples from these expression cultures (3 hours post-induction) were fractionated into soluble and insoluble protein. Collected culture samples were centrifuged at 16 000 *g* for 10 minutes to pellet cells. The pellet was resuspended and incubated in lysis buffer (1 mg/mL lysozyme, 20% (w/v) sucrose, 30 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0) for 10 minutes on ice. The suspension was treated with three freeze-thaw cycles, followed by centrifugation at 16 000 *g* for five minutes at 4ºC. Insoluble proteins were collected in the pellet and soluble proteins in the supernatant. Figure 3.10 illustrates that at 37°C, recombinant protein was being expressed at a much higher percentage as insoluble inclusion bodies. At 25°C, after three hours, total recombinant protein expression was much lower than at 37°C.

Correspondingly, less recombinant protein was visualised in the fractionated sample, but the majority of it was also insoluble. These results clearly showed that a number of conditions would need to be optimised to achieve a greater percentage of soluble recombinant protein.

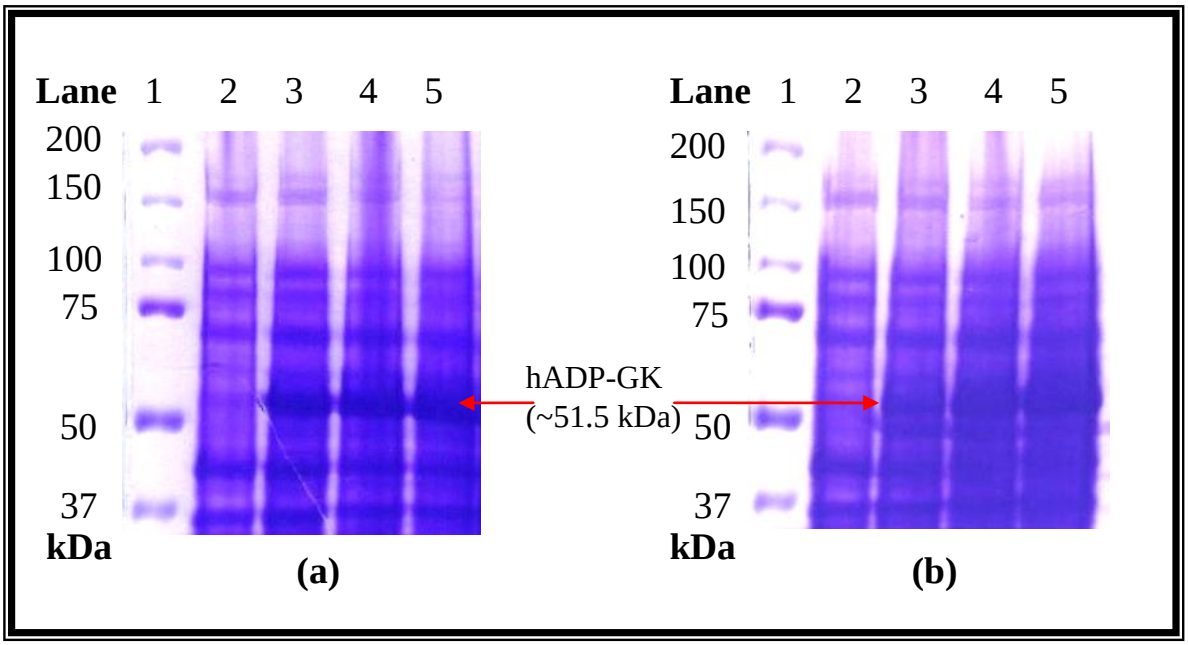


Figure 3.9 SDS-PAGE of whole cell lysates collected from 37ºC BL21(DE3) trial expression cultures

The expression of recombinant human ADP-GK protein by transformed BL21(DE3) cells was monitored by SDS-PAGE of whole cell protein lysates. One millilitre samples of culture were collected directly before the induction of recombinant protein expression at 37ºC by the addition of 0.1 mM (a) or 0.5 mM (b) IPTG, and every hour, post induction, for three hours. Cells were harvested by centrifugation and whole cell lysis was performed by treating cells with 20 µL of 10% SDS and 20 µL of 5× treatment buffer. The 40 µL samples of whole cell protein lysates from each stage of expression were heat treated at 100°C for five minutes before being loaded in total onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie Blue.

Lane one: Precision Plus Protein™ Dual Colour Standards; **Lane two:** Whole cell lysate pre-induction of protein expression; **Lane three:** Whole cell lysate one hour post-induction; **Lane four:** Whole cell lysate two hours post-induction; **Lane five:** Whole cell lysate three hours post-induction.

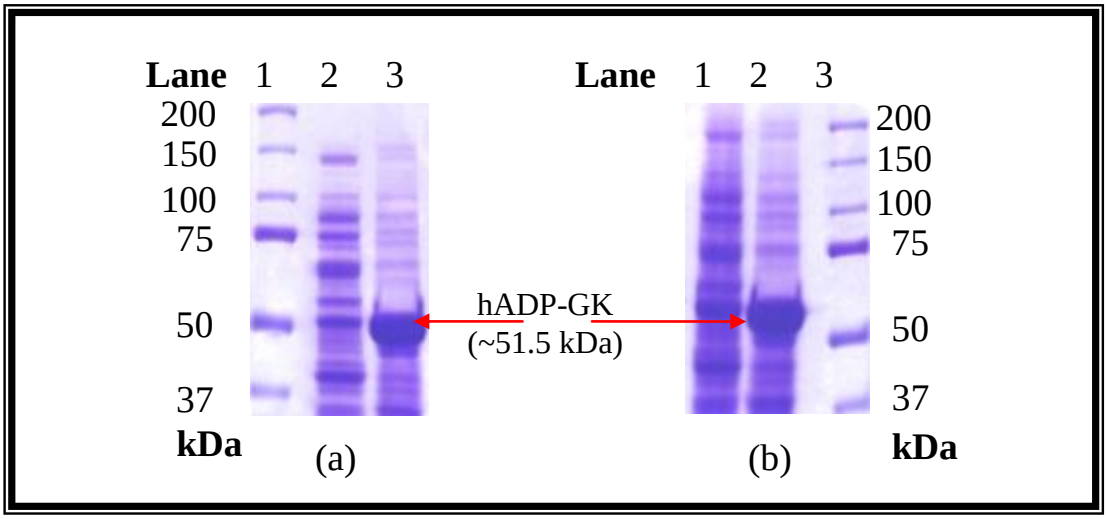


Figure 3.10 SDS-PAGE of soluble and insoluble protein fractions produced from 37°C BL21(DE3) trial expression cultures

The soluble expression of recombinant human ADP-GK protein by transformed BL21(DE3) cells was monitored by SDS-PAGE of fractionated whole cell protein lysates. One millilitre samples of culture were collected three hours post-induction of recombinant protein expression at 37ºC, induced by the addition of 0.1 mM (a) or 0.5 mM (b) IPTG. Cells were harvested by centrifugation and protein fractionation was performed. The cells were lysed by incubation in lysis buffer for ten minutes on ice. The suspension was treated with three freeze-thaw cycles, followed by centrifugation at 16 000 *g* for five minutes at 4ºC. Insoluble proteins were collected in the pellet and soluble proteins were located in the supernatant. Twenty-five microlitres of 5× treatment buffer was added to each 125 µL protein fraction sample and this was followed by heat treatment at 100°C for five minutes. Forty microlitres of soluble and insoluble fractions were loaded onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie Blue.

Lane one (a) and **three** (b): Precision Plus Protein™ Dual Colour Standards; **Lane two** (a) and **one** (b): Soluble protein fraction, **Lane three** (a) and **two** (b): Insoluble protein fraction.

3.5 Improving the solubility of expressed recombinant ADP-dependent glucokinase

Once it was evident that the recombinant protein was being expressed primarily as insoluble protein, trials which changed expression conditions with the aim of achieving a greater percentage of soluble protein were performed.

Since actual recombinant protein expression was high, targeting the rate of protein production was the main objective. The cytosol of an *E. coli* cell is not an ideal environment for eukaryotic protein folding and the compromised conditions can be exacerbated by the incredibly fast rate of recombinant protein expression. Even proteins that have a demonstrated ability to fold into their native conformation under bacterial cytosolic conditions have a greater tendency to move off the correct folding pathway and cause aggregation events to a more frequent extent as protein concentration increases (Anfinsen, 1972). To address this issue, the conditions of recombinant protein expression were first modified in terms of temperature and final inducing concentrations of IPTG.

BL21(DE3) cells carrying the expression plasmid were induced to express recombinant protein at 16°C, 25°C, 30°C and 37°C. IPTG at 0.05, 0.1 or 0.5 mM was used to induce recombinant protein expression at the temperatures of 25°C and 30°C. IPTG at 0.1 mM or 0.5 mM was used to induce protein expression at 16°C and 37°C. Lowering the expression temperature delayed the onset of recombinant protein expression, but once expression commenced, it was generally substantial. Lowering the concentration of IPTG used to induce expression did not influence the solubility of the recombinant protein, as observed by SDS-PAGE (Figure 3.11). This result was not entirely unforeseen as it is known that the T7 polymerase of BL21(DE3) cells is very efficient and only a few molecules of IPTG are required for over-expression of a recombinant protein. Therefore, there is probably less capacity for expression to be gradually stimulated by low levels of IPTG. Instead, it is more like an expression switch, where expression of recombinant protein is either on or off. At the expression temperature of 16°C, the culture was allowed to express for 24.5 hours. No expression of recombinant hADP-GK from this trial could be visualised by SDS-PAGE (data not shown).

3.5.1 Escherichia coli Rosetta host cell strain

Codon usage can be a cause of insoluble protein expression. Some eukaryotic proteins contain codons rarely used in *E. coli* and this can limit translation of non-endogenous protein in *E. coli.* In order to rule this potential problem out, a Rosetta strain of *E. coli* was trialled. The Rosetta host strain is a derivative of the BL21 strain but it carries a chloramphenicol-resistant plasmid, pRARE, which contributes tRNAs for the codons rarely used in *E. coli.*

Initial expression trials were carried out with the Rosetta strain to confirm recombinant protein expression and check its relative solubility. The first three trials were carried out using an expression temperature of 30°C and either 0.05 mM, 0.1 mM or 0.5 mM IPTG to induce protein expression. SDS-PAGE was carried out with whole cell lysates created from samples collected over a four hour period of expression. The Rosetta strain did express the recombinant protein under these conditions (Figure 3.11), however, when the samples collected from the expression cultures were lysed to fractionate soluble protein from insoluble, the strongly visible recombinant human ADP-GK was present in the insoluble fraction. It was clear that methods supplementary to variation in expression temperature and IPTG concentrations were required to aid with soluble protein expression.

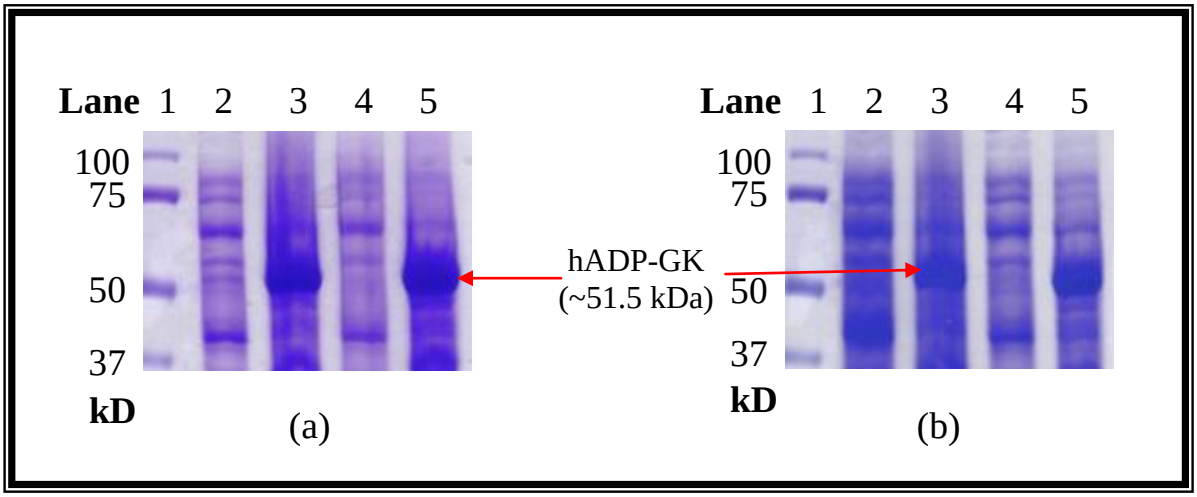


Figure 3.11 SDS-PAGE of soluble and insoluble protein fractions produced from 30ºC Rosetta trial expression cultures

The soluble expression of recombinant human ADP-GK protein by transformed Rosetta cells was monitored by SDS-PAGE of fractionated whole cell protein lysates. One millilitre samples of culture were collected three hours post-induction of recombinant protein expression at 30ºC, induced by the addition of 0.05 mM (a) or 0.1 mM (b) IPTG. Cells were harvested by centrifugation and protein fractionation was performed. The cells were lysed by incubation in lysis buffer for ten minutes on ice. The suspension was treated with three freeze-thaw cycles, followed by centrifugation at 16 000 *g* for five minutes at 4ºC. Insoluble proteins were collected in the pellet and soluble proteins were located in the supernatant. Twenty-five microlitres of 5× treatment buffer was added to each 125 µL protein fraction sample and this was followed by heat treatment at 100°C for five minutes. Forty microlitres of soluble and insoluble fractions were loaded onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie Blue.

Lane one: Precision Plus Protein™ Dual Colour Standards; **Lane two:** Soluble protein fraction (3 hours post-induction of protein expression), **Lane three:** Insoluble protein fraction (3 hours post-induction of protein expression), **Lane four:** Soluble protein fraction (4 hours post-induction of protein expression), **Lane five:** Insoluble protein fraction (4 hours post-induction of protein expression).

3.5.2 Induction of molecular chaperone expression

A technique often used to assist with the expression of soluble recombinant protein is the co-expression of molecular chaperones. Molecular chaperones can prevent protein aggregates forming and actively scavenge and convert aggregates into natively refolded species (de Marco *et al*., 2005). However, coordinating the expression of molecular chaperones from chaperone-harbouring plasmids, so that the combination and ratio expressed resembles that of a normal cell, can be time consuming. Due to time constraints, this option was not explored further. Another experiment based on molecular chaperone assisted expression of soluble recombinant protein was undertaken instead, based on a method reported by de Marco *et al*. (2005). This method suggested the addition of benzyl alcohol to the bacterial culture during expression of recombinant protein. Benzyl alcohol selectively increases motion deep within a cell's lipid bilayer, mimicking the effect of an increase in environmental temperature (Cooper and Meddings, 1991), and hence stimulates the expression of endogenous chaperones that are produced under conditions of heat shock. de Marco *et al*. (2005) demonstrated that the addition of benzyl alcohol to the bacterial medium to promote chaperone expression can be as successful in assisting with the soluble expression of recombinant protein as the more complicated, plasmid-based method of artificially expressing molecular chaperones.

A more recent report (de Marco, 2007) outlined the concept that cessation of the over-expression of recombinant protein can aid in chaperone-assisted refolding of recombinant protein. Recombinant protein expression represents a significant energetic burden for the host cell. Under such conditions, the protein refolding machinery (even if there are artificially increased levels of molecular chaperones available) is likely to be insufficient to attain appropriate protein folding. Under normal conditions, it has been demonstrated that wild-type cells exhibit a dynamic exchange between soluble and aggregate protein fractions; in order to facilitate this exchange under recombinant protein expression conditions, de Marco suggests that complete inhibition of protein expression in the host cell will allow time for the molecular chaperones present to disaggregate pre-existing precipitates, instead of being overwhelmed with newly synthesised protein. This dynamic

exchange is achieved in a basic two-step protocol as follows: bacterial cells are collected after protein expression has taken place, and are resuspended in IPTG-free media, containing 200 µg/mL chloramphenicol to inhibit any further protein expression. The bacteria are then left to incubate for two hours (to disaggregate and refold misfolded protein) before being re-pelleted and analysed for soluble protein expression.

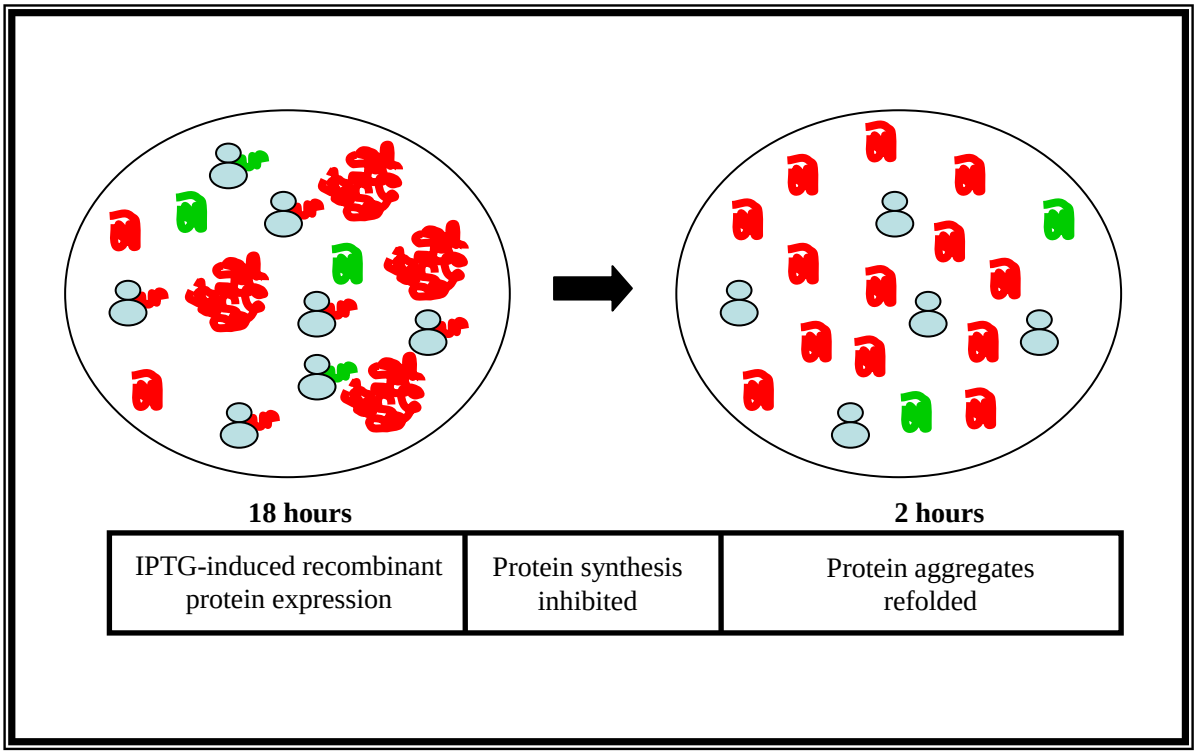


Figure 3.12 A schematic of the two-step expression protocol, adapted from (de Marco, 2007)

Expression of a recombinant protein creates an unbalanced scenario where an excessive amount of recombinant protein is produced at a very fast rate, and in many cases overwhelms the protein folding machinery available within the host cell. This diagram illustrates how protein aggregates can form as a result of rapid protein expression. By inhibiting new protein synthesis with the addition of chloramphenicol, and leaving the cell to recover for two hours, the protein folding machinery has the opportunity to disaggregate precipitated protein and thereby increase the soluble yield of recombinant protein.

An expression trial was designed which utilized benzyl alcohol as the trigger for the expression of molecular chaperones to aid the increase of soluble recombinant ADP-GK (de Marco *et al*., 2005, 2007). Cultures were grown in the presence of selective antibiotics at 37°C until they had reached an optical density at A_{600} of approximately 0.4. The cultures were moved to a 22°C room and left to equilibrate for approximately ten minutes before benzyl alcohol was added to a final concentration of 10 mM to two of the three flasks. The expression temperature was set at 22°C to avoid protein expression being too rapid. The 50 mL cultures were induced to express recombinant protein twenty minutes after the addition of benzyl alcohol with a final IPTG concentration of 0.1 mM. Samples of culture were taken before the induction of protein expression, one hour and two hours post-induction, and twenty hours post-induction. After twenty hours, all cells were collected from the expressing cultures. The cells from two cultures were subsequently resuspended in fresh LB broth, with no IPTG present, and 200 µg/mL chloramphenicol to inhibit any further protein expression. These cultures were incubated at 22°C for a further two hours before the cells were re-collected. Whole cell lysates were created first to check that expression of the recombinant human ADP-GK had taken place. The protein gels were positive for expression (Figure 3.13), although the human ADP-GK band was not as strong as in previous expression trials. This was an encouraging result because the expression temperature had had a perceptible effect on the amount of protein expression. Fractionated protein lysates were then created, to analyse the effect of new experimental conditions on the solubility of the recombinant human ADP-GK. As can be observed in Figure 3.14, all three combinations of the conditions trialled in the expression experiment gave essentially the same result as previous expression trials. The visible band representing the recombinant human ADP-GK was evident in the insoluble fractions only, even in the presence of benzyl alcohol and after the inhibition of protein expression and a further two hour incubation to facilitate the disaggregation of protein.

The same experimental conditions were again trialled with the Rosetta host strain, except the expression temperature was reduced to 18°C. The whole cell lysate samples produced from this experiment were analysed on a denaturing gel and the band representing the human ADP-GK protein looked less over-expressed and therefore solubility seemed more

likely. However, when the samples were fractionated, the solubility of the recombinant protein had not improved.

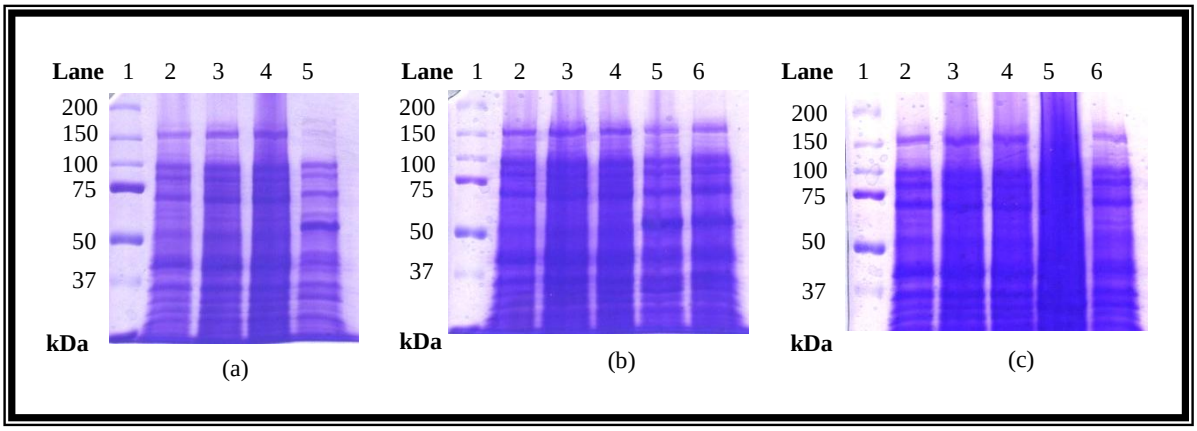


Figure 3.13 SDS-PAGE of whole cell lysates from 22ºC Rosetta trial expression cultures

The expression of recombinant human ADP-GK protein by transformed Rosetta cells was monitored by SDS-PAGE of whole cell protein lysates. Benzyl alcohol was added to a final concentration of 10 mM to cultures (a) and (b) (but not to culture (c)) twenty minutes before protein expression was induced. One millilitre samples of culture were collected directly before the induction of recombinant protein expression at 22ºC by the addition of 0.1 mM IPTG, and for one, two and twenty hours post-induction. Cells were harvested by centrifugation and whole cell lysis was performed by treating cells with 20 µL of 10% SDS and 20 µL of 5× treatment buffer. The 40 µL samples of whole cell protein lysates from each stage of expression were heat treated at 100°C for five minutes before being loaded in total onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie Blue.

Lane one: Precision Plus Protein™ Dual Colour Standards; **Lane two:** Whole cell lysate pre-induction of protein expression; **Lane three:** Whole cell lysate one hour post-induction; **Lane four:** Whole cell lysate two hours post-induction; **Lane five:** Whole cell lysate twenty hours post-induction. **Lane six:** Whole cell lysate twenty-two hours post-induction following the removal of IPTG from the culture medium and the addition of 200 µg/mL of chloramphenicol.

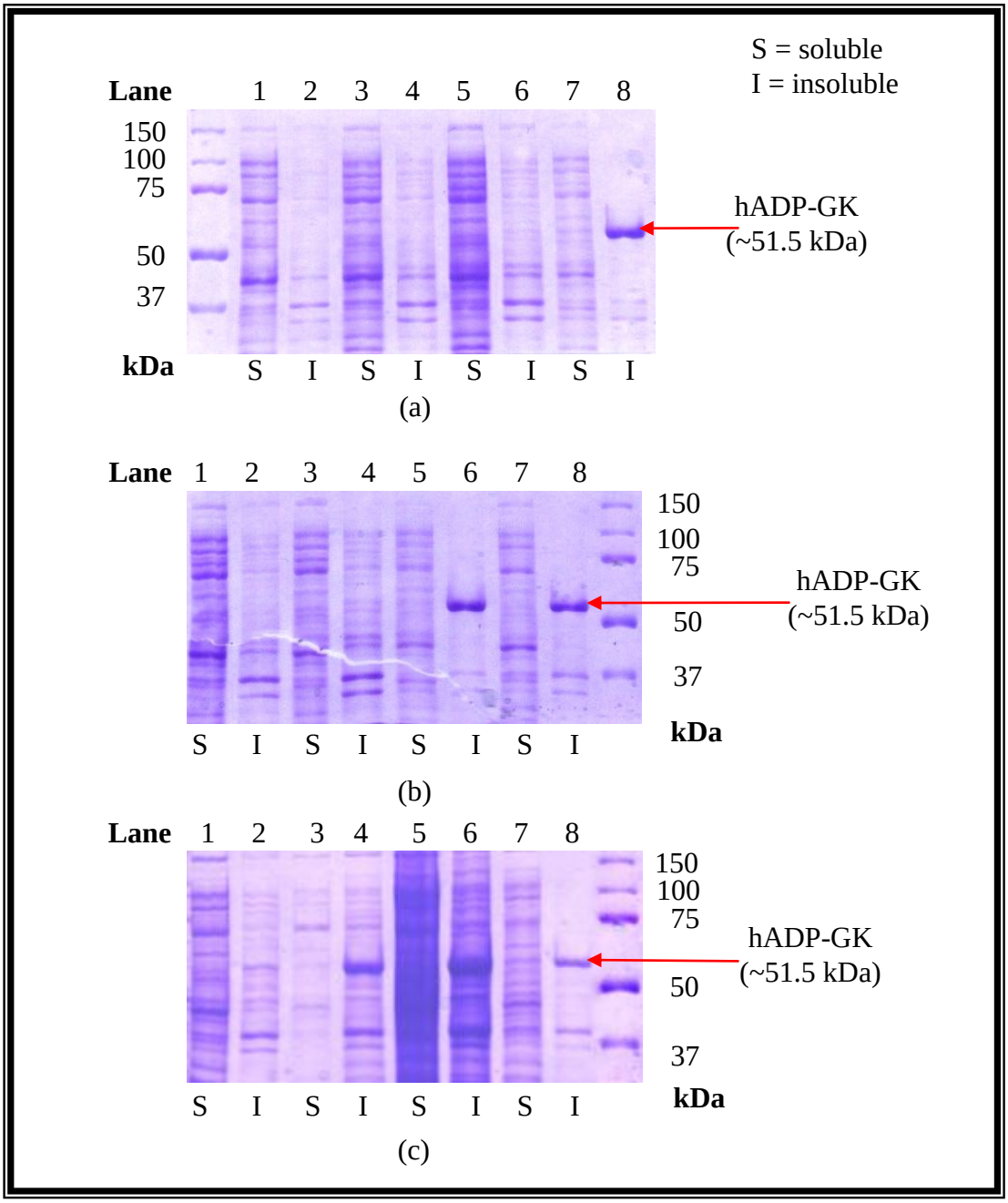


Figure 3.14 SDS-PAGE of soluble and insoluble protein fractions produced from 22ºC Rosetta trial expression cultures

The soluble expression of recombinant human ADP-GK protein by transformed Rosetta cells was monitored by 8% SDS-PAGE of fractionated whole cell protein lysates. Benzyl alcohol was added to a final concentration of 10 mM to cultures (a) and (b) twenty minutes before protein expression was induced. One millilitre samples of culture were collected

directly before the induction of recombinant protein expression at 22ºC by the addition of 0.1 mM IPTG, and for one, two and twenty hours post-induction. Culture (c) was induced to express recombinant protein without the addition of benzyl alcohol. Cells were harvested by centrifugation and protein fractionation was performed. The cells were lysed by incubation in lysis buffer for ten minutes on ice. The suspension was treated with three freeze-thaw cycles, followed by centrifugation at 16 000 *g* for five minutes at 4ºC. Insoluble proteins were collected in the pellet and soluble proteins were located in the supernatant. Twenty-five microlitres of 5× treatment buffer was added to each 125 µL protein sample and this was followed by heat treatment at 100°C for five minutes. Forty microlitres of soluble and insoluble fractions were loaded onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie Blue.

Lane one: Soluble fraction (one hour post-induction of protein expression); **Lane two:** Insoluble fraction (one hour post-induction of protein expression); **Lane three:** Soluble fraction (two hours post-induction of protein expression); **Lane four:** Insoluble fraction (two hours post-induction of protein expression); **Lane five:** Soluble fraction (twenty hours post-induction of protein expression); **Lane six:** Insoluble fraction (twenty hours post-induction of protein expression); **Lane seven:** Soluble fraction (twenty-two hours post-induction following the removal of IPTG from the culture medium and the addition of 200 µg/mL of chloramphenicol); **Lane eight:** Insoluble fraction (twenty-two hours post-induction following the removal of IPTG from the culture medium and the addition of 200 µg/mL of chloramphenicol). The marker lane contains eight microlitres of Precision Plus Protein™ Dual Colour Standards.

3.5.3 A comparison of lysis buffers

Due to the consistent and apparent complete insolubility of the recombinant human ADP-GK under the range of expression conditions trialled, it was considered whether the method and buffers used in the process of cell lysis and protein fractionation was influencing the observed protein solubility. It was possible there could have been some expression of soluble protein, but that the fractionation conditions may have been causing the soluble recombinant protein present to denature. To address this issue, a more thorough cell lysis method was generated and a trial using a range of buffer components was undertaken.

The effect of the cell lysis buffer on the perceived solubility of the recombinant human ADP-GK solubility was evaluated by SDS-PAGE and immunoblotting. The lysis buffers investigated differed in their NaCl concentration (0.6 M or 0.25 M), the presence or absence of 0.1% Triton X-100 and/or the presence or absence of 4 mM 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate, commonly referred to as CHAPS. Salt concentration can influence the solubility of a protein in solution, and high salt concentrations can inhibit non-specific protein-protein interactions. Detergents are a common component of lysis buffers and can aid in the disruption of protein aggregates formed due to hydrophobic interactions. Triton X-100 is a non-ionic detergent, CHAPS is a zwitterionic detergent, both of which are non-denaturing.

Cultures of 50 mL were grown at 18ºC, 25ºC and 30ºC for 23 hours, 18 hours and six hours, respectively. Protein expression was induced with 0.01 mM or 0.1mM IPTG. The cells were collected from the 50 mL cultures by centrifugation and were resuspended in 1 mL of lysis buffer, containing 1 mg/mL lysozyme. The cell solution was lysed completely using the French press three times under a pressure of 5 000 psi. This was followed by incubation in the presence of 200 µg/mL of DNase I and 5 mM $MgCl_2$. Soluble protein was then separated from insoluble protein by centrifugation at 16 000 *g* for 40 minutes at 4ºC.
The samples created from the 30ºC cultures gave the clearest view by SDS-PAGE of the effect that the lysis buffer components were having on the solubility of the recombinant protein. The denaturing gels produced from these samples demonstrated no visible effect

when the concentration of NaCl was changed from 0.6 M to 0.25 M. The presence or absence of Triton X-100 in combination with the changes of salt concentration also did not have a significant effect on the apparent solubility of the recombinant protein. The presence or absence of CHAPS in the lysis buffer, however, did appear to have an influence on the amount of protein visible in the soluble fraction. In Figure 3.15, the most significant soluble band of recombinant human ADP-GK is present in lane eleven and lane thirteen. To confirm this observation, immunoblotting was performed on three fractionated samples from the cultures grown at 18°C (Figure 3.16). These were chosen because there was less recombinant human ADP-GK in the fractionated samples. Immunoblotting performed on samples from cultures grown at 30ºC gave unclear results, due to the abundance of the recombinant protein. The immunoblot result supported the observations made from the Coomassie stained protein gels. When no detergent was added to the lysis buffer, (Figure 3.16, lanes one and two) almost all the recombinant protein was in an insoluble state. When 0.1% Triton X-100 was present in the lysis buffer (Figure 3.16, lanes three and four), the recombinant protein appeared to be distributed fairly evenly between the soluble and insoluble fractions. When 4 mM CHAPS was added to the lysis buffer (Figure 3.16, lanes five and six), there did not appear to be relatively more soluble recombinant protein than when Triton X-100 was present, although no insoluble protein was detected.

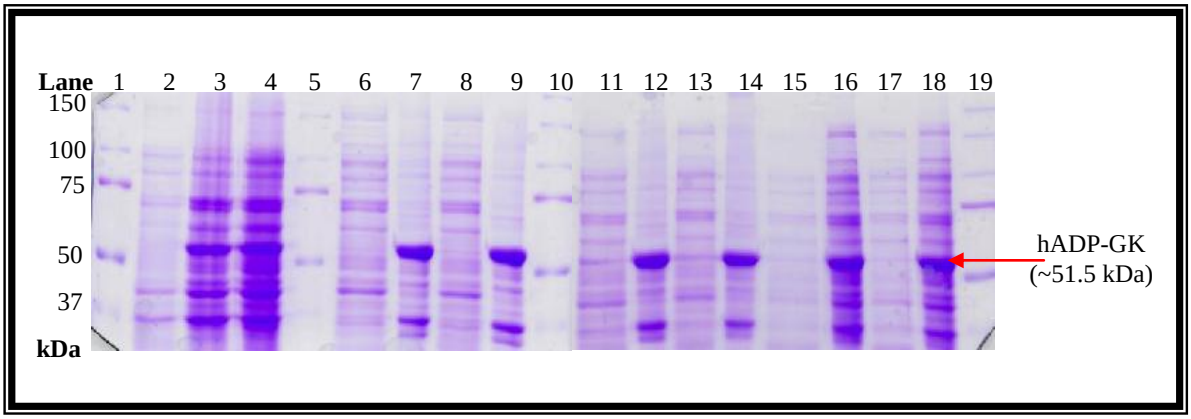


Figure 3.15 SDS-PAGE of protein samples demonstrating the effect of changing lysis buffer components

The soluble expression of recombinant human ADP-GK protein by transformed Rosetta cells was monitored by SDS-PAGE of fractionated whole cell protein lysates. One millilitre samples of culture were collected directly before the induction of recombinant protein expression at 30ºC by the addition of 0.1 mM IPTG, and at four and a half and six hours post-induction. Cells were harvested by centrifugation and protein fractionation was performed. The cells were resuspended in 1 mL of a specific lysis buffer before the cell solution was lysed completely by using the French press three times under a pressure of 5 000 psi. This was followed by incubation in the presence of 200 µg/mL of DNase I and 5 mM $MgCl_2$. The resulting solution was centrifuged at 16 000 *g* for five minutes at 4ºC. Insoluble proteins were collected in the pellet and soluble proteins were located in the supernatant. Two hundred and fifty microlitres of 5× treatment buffer was added to each 1 mL protein fraction sample and this was followed by heat treatment at 100°C for five minutes. Eighty microlitres of soluble and insoluble fractions were loaded onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie Blue.

Lane one: Precision Plus Protein™ Dual Colour Standards; **Lane two:** Whole cell protein (pre-induction of protein expression); **Lane three:** Whole cell protein (4.5 hours post-induction of protein expression); **Lane four:** Whole cell protein (six hours post-induction of protein expression); **Lane five:** Precision Plus Protein™ Dual Colour Standards; **Lane six:** Soluble fraction (lysed in 0.6 M NaCl, 4 mM CHAPS, 0.1% Triton X-100); **Lane seven:** Insoluble fraction from the conditions used in lane six; **Lane eight:** Soluble fraction

(lysed in 0.25 M NaCl, 4 mM CHAPS, 0.1% Triton X-100); **Lane nine:** Insoluble fraction from the conditions used in lane eight; **Lane ten:** Precision Plus Protein™ Dual Colour Standards; **Lane eleven:** Soluble fraction (lysed in 0.6 M NaCl, 4 mM CHAPS); **Lane twelve:** Insoluble fraction from the conditions used in lane eleven; **Lane thirteen:** Soluble fraction (lysed in 0.25 mM NaCl, 4 mM CHAPS); **Lane fourteen:** Insoluble fraction from the conditions used in lane thirteen; **Lane fifteen:** Soluble fraction (lysed in 0.6 M NaCl); **Lane sixteen:** Insoluble fraction from the conditions used in lane fifteen; **Lane seventeen:** Soluble fraction (lysed in 0.25 M NaCl); **Lane eighteen:** Insoluble fraction from the conditions used in lane seventeen; **Lane nineteen:** Precision Plus Protein™ Dual Colour Standards.

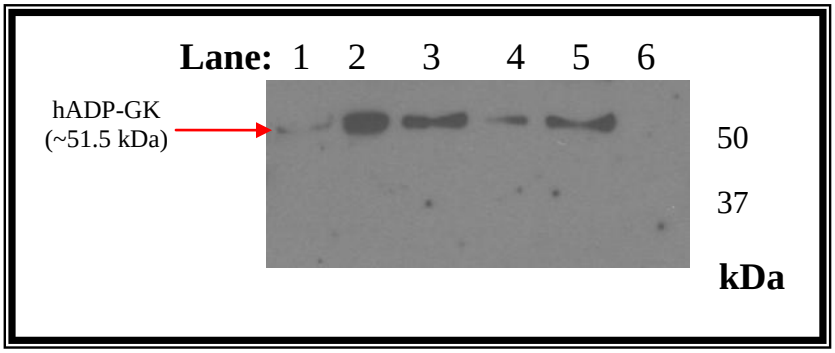


Figure 3.16 Immunoblot of recombinant human ADP-GK fractionated lysis samples

An immunoblot was performed on the soluble and insoluble fractions created from three different lysis buffer combinations. This was carried out to demonstrate the effect of changing the buffer components on the apparent solubility of the recombinant human ADP-GK. Fractionated protein samples, created from cultures grown at 18°C induced to express recombinant protein with 0.01 mM IPTG, were separated by 8% SDS-PAGE at 120 V for approximately 120 minutes. Separated protein was transferred to a positively charged nylon membrane for 45 minutes at 450 mA. The membrane was immunoblotted to detect the recombinant human ADP-GK using a 1:1000 dilution of monoclonal ADP-GK mouse primary antibody and a 1:1000 dilution of rabbit anti-mouse secondary antibody conjugated to horseradish peroxidase. The membrane was exposed to x-ray film for approximately one minute.

Lane one: Soluble fraction (lysed in 0.6 M NaCl); **Lane two:** Insoluble fraction from lane one; **Lane three:** Soluble fraction (lysed in 0.6 M NaCl, 0.1% Triton X-100); **Lane four:** Insoluble fraction from lane three; **Lane five:** Soluble fraction (lysed in 0.6 M NaCl, 4 mM CHAPS); **Lane six:** Insoluble fraction from lane five.

3.6 Chapter summary

The cDNA sequence for the human ADP-GK protein was cloned into the *E. coli* host strains Bl21(DE3) and Rosetta, using a pET151/D-TOPO vector. The cloning strategy was confirmed by endonuclease restriction digest and a positive sequence alignment with the cDNA sequence obtained from the NCBI nucleotide database (accession number BC006112). The recombinant human ADP-GK was expressed at temperatures that ranged from 18°C to 37°C and IPTG concentrations from 0.05 mM to 0.5 mM. However, the percentage of soluble recombinant protein expressed was very low, even when protein expression was slowed right down by low temperature conditions. A trial to induce the expression of molecular chaperones by the addition of benzyl alcohol was undertaken, in the expectation that increased chaperone levels would assist in the disaggregation and refolding of insoluble recombinant protein. This did not appear to significantly change the insoluble state of the recombinant human ADP-GK, and neither did the cessation of new protein synthesis by the addition of chloramphenicol, to give the cell's protein folding machinery time to process protein aggregates.

It was considered that cell lysis and fractionation conditions could be influencing the apparent distribution of soluble and insoluble recombinant protein. Consequently, a trial of lysis buffers was carried out, examining the effect of low and high NaCl concentrations, and the presence and/or absence of the non-denaturing detergents, Triton X-100 and CHAPS. It was observed by SDS-PAGE that the most significant influence on the distribution of recombinant protein between the soluble and insoluble states was the addition of detergent, and in particular CHAPS. Immunoblotting of a fractionated lysed sample created from an expression culture grown at 18°C demonstrated that in the presence of CHAPS, no insoluble recombinant protein could be detected.

Under the expression conditions trialled, and using the pET151/D-TOPO expression vector, it is clear that the recombinant human ADP-GK had a predisposition toward insolubility. However, with the optimisation of the lysis buffer, it was decided that purification of the small percentage of soluble recombinant protein was possible, and that upscaling the volume of culture could aid in the final yield of soluble protein.

Chapter 4: Purification of recombinant human ADP-dependent glucokinase

4.1 Introduction

The expression of recombinant protein with a polyhistidine tag is a method commonly used to assist in protein purification. More than 60% of proteins expressed for the purposes of structural studies make use of polyhistidine tags. The purification of his-tagged proteins is based on the affinity the imidazole side chain of histidine has for metal ions. Polyhistidine tags interact with metal ions such as nickel, and this affinity is exploited by complexing nickel ions with an immobilised chelating agent used as a chromatographic support. The most commonly used chromatographic support is nitrilotriacetic acid (NTA).

Elution of the recombinant protein from the affinity column is accomplished by competing against the interaction between the polyhistidine tag and the nickel ions by adding increasing concentrations of imidazole. The specificity of the binding between the polyhistidine tag and the immobilised nickel ions is so specific that protein purification can be successfully carried out under both non-denaturing and denaturing conditions. The use of affinity tags in the chromatographic purification of recombinant proteins typically results in high yields (over 90%) of protein and purification can be achieved in one step, directly from the crude extract of the host cell (Arnau *et al*., 2005). This makes affinity chromatography an economical method, both in terms of material used and time taken to attain purified protein.

The recombinant ADP-GK protein was expressed with an N-terminal 6× histidine tag to facilitate protein purification, and the aim of the work presented in this chapter was to optimise the purification process to achieve protein homogeneity.

4.2 Initial purification trials and optimisation of Ni-NTA resin purification conditions

The recombinant human ADP-GK was expressed with an N-terminal 6×Histidine tag to aid in its selective purification. A Ni-NTA Purification System from Invitrogen, which incorporated Ni-NTA agarose resin as a slurry and 10 mL plastic purification columns, was used in the protein purification trials. Using this system meant that the amount of resin used in a trial could be adjusted and the purification could be performed on the laboratory bench.

To begin with, a 50 mL culture of BL21 (DE3) cells containing expression plasmid, induced with 0.5 mM IPTG, was grown for five hours at 30°C. The cells were harvested by centrifugation and resuspended in lysis buffer (50 mM NaH_2PO_4, 500 mM NaCl, 1 mg/mL lysozyme, one Complete^{TM} Mini EDTA-free protease inhibitor cocktail tablet, pH 8.0).

The Ni-NTA resin was equilibrated with binding buffer before resin was used to bind protein. The soluble protein fraction was incubated with 0.7 mL of resin (approximately enough to bind between 0.7 mg and 3.5 mg of His-tagged recombinant protein) for 30 minutes at 4°C, with shaking. The Ni-NTA agarose was then settled in the 10 mL plastic purification column by centrifugation for one minute at 800 *g*.

The 0.7 mL column was washed four times, using 8 mL of wash buffer (50 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, pH 8.0). Washing was performed by allowing the wash buffer to drip through the packed Ni-NTA resin.

A stepwise elution followed using five column volumes of elution buffer at each step (50 mM NaH_2PO_4, 500 mM NaCl in combination with 20 mM, 60 mM, 70 mM, 120 mM, 150 mM or 250 mM imidazole). Fractions of 1 mL in volume were collected.

As the elution samples had very low concentrations of protein, the denaturing gel was silver-stained to detect protein (Figure 4.1). Lane two (Figure 4.1 (a)) showed that not all the recombinant protein that was present in the cell lysate sample bound to the nickel resin.

Some recombinant protein was still detected in the unbound protein sample. The first wash contained detectable protein, but it seemed that the wash was unspecific as it contained a range of proteins representative of that which did not bind to the Ni-NTA in the first place, including the recombinant protein. Protein present in the three subsequent washes was not detectable by silver stain.

Lanes two to nine of Figure 4.1 (b) represent the stepwise elution. Recombinant human ADP-GK was visible at each imidazole concentration used in the stepwise elution, as well as one to four other proteins. Lanes two and three appeared to have only one other significant contaminating protein. By the 250 mM imidazole step, most of the recombinant protein seemed to have been eluted. It was clear that since the recombinant protein eluted at essentially every imidazole concentration, including the 20 mM imidazole step, it was necessary to lower the concentration of imidazole in the wash buffer and even the binding buffer, as they contained 20 mM and 10 mM imidazole, respectively. Also, a single step elution was considered, to avoid diluting the eluted recombinant protein.

Once the lysis buffer had been optimised for the maximum yield of soluble protein, as described in Chapter 3, (Section 3.5.3), this was used in conjunction with the purification trials. Cultures of the Rosetta strain of *E. coli* (50 mL) were grown at 18°C for 22 hours before cells were harvested and lysed, using the French press, in 50 mL of lysis buffer (50 mM Tris-HCl, 4 mM CHAPS, 300 mM NaCl, 5.0 mL glycerol, 1 mg/mL lysozyme, one Complete^{TM} Mini EDTA-free protease inhibitor cocktail tablet, pH 8.0). Following the first trial purification results, Triton X-100 (0.01%) was added to the lysate and incubation at 4°C took place for 30 minutes. The addition of Triton X-100 assists with the solubilising of protein, the removal of nucleic acids, and can inhibit non-specific protein-protein interactions (Dasari *et al.*, 2008).

The lysate was then centrifuged at 16 000 *g* for 30 minutes at 4°C to pellet insoluble protein. Approximately 50 mL of the soluble protein fraction was incubated with 9 mL of resin for 30 minutes on ice, with shaking. A washing step was performed with four column volumes of wash buffer (5 mM imidazole, 50 mM NaH_2PO_4, 500 mM NaCl, pH 8.0).

Elution of recombinant protein was attempted using four column volumes of elution buffer (120 mM imidazole, 50 mM NaH_2PO_4, 500 mM NaCl, pH 8.0). The pooled eluate was then concentrated to a volume of 0.5 mL using 30 kDa molecular mass cut-off Centricon™ concentrators and analysed by SDS-PAGE followed by silver staining.

Figure 4.2 shows that in the soluble protein sample, the presence of the recombinant protein was not abundant. The dilution of eluted protein was considerable as no protein was present in lane two. However, the eluted protein that had been concentrated (lane four) contained many protein species, indicating that several different proteins were associating with the nickel and eluting off at a concentration of imidazole of 120 mM. The recombinant human ADP-GK could also be visualised in the concentrated eluate, but it was obvious that further optimisation of the purification conditions was required.

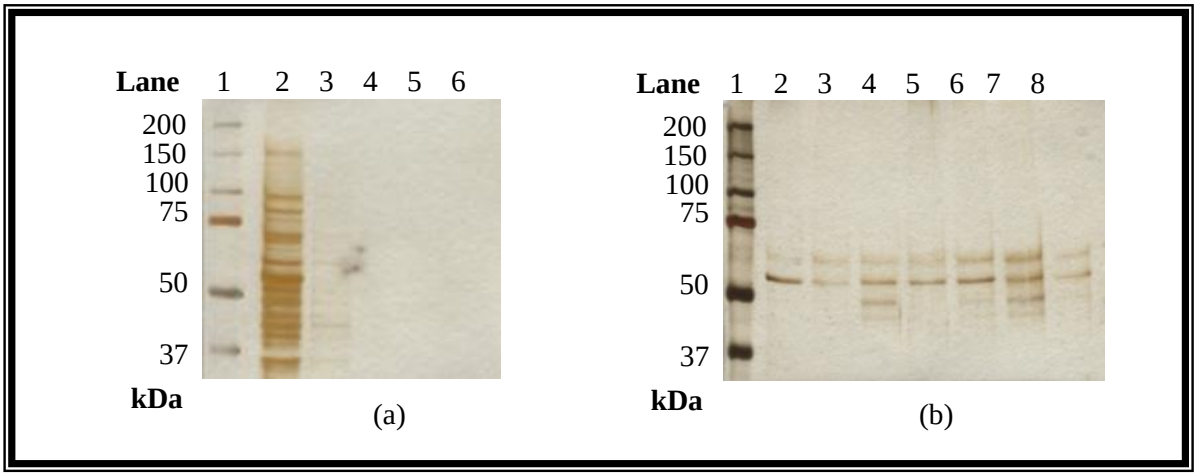


Figure 4.1 SDS-PAGE of a trial purification using Ni-NTA resin

The relative purification of soluble recombinant human ADP-GK protein eluted off the Ni-NTA column was monitored by SDS-PAGE. Five microlitres of 5× treatment buffer was added to each 20 µL protein sample and this was followed by heat treatment at 100°C for five minutes. Twenty microlitres of each protein fraction was loaded onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was silver-stained.

Gel (a): Lane one: Precision Plus Protein™ Dual Colour Standards; **Lane two:** Soluble protein loaded onto Ni-NTA column; **Lane three:** Wash one; **Lane four:** Wash two; **Lane five:** Wash three; **Lane six:** Wash four.

Gel (b): Lane one: Precision Plus Protein™ Dual Colour Standards; **Lane two:** 20 mM imidazole elution; **Lane three:** 50 mM imidazole elution; **Lane four:** 60 mM imidazole elution; **Lane five:** 70 mM imidazole elution; **Lane six:** 120 mM imidazole elution; **Lane seven:** 150 mM imidazole elution; **Lane eight:** 250 mM imidazole elution.

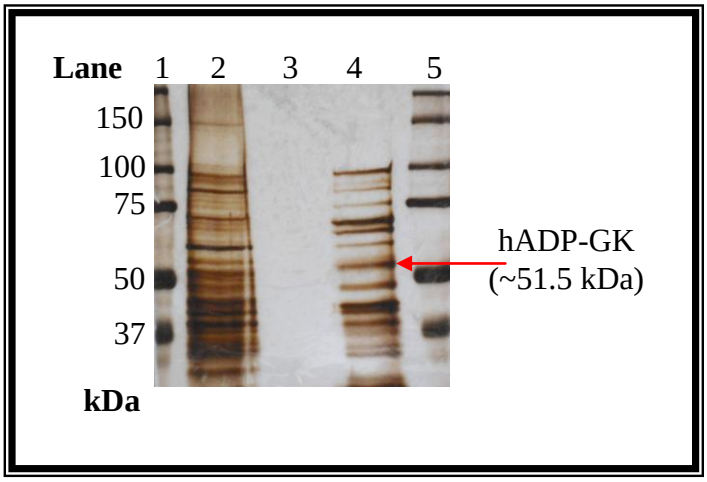


Figure 4.2 SDS-PAGE of a trial purification using Ni-NTA resin

The relative purification of soluble recombinant human ADP-GK protein eluted off the Ni-NTA column was monitored by SDS-PAGE. Five microlitres of 5× treatment buffer was added to each 20 µL protein sample and this was followed by heat treatment at 100°C for five minutes. Twenty microlitres of each protein fraction was loaded onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was silver-stained.

Lanes one and five: Precision Plus Protein™ Dual Colour Standards; **Lane two:** Soluble protein loaded onto Ni-NTA column; **Lane three:** Pooled 120 mM imidazole elution (36 mL); **Lane four:** Concentrated eluted protein.

Due to non-specific protein binding, the wash buffer was modified by dropping the pH from 8.0 to 6.9 (1.1 pH units above the predicted isoelectric point of 5.8). The volume of culture used to make the 40 mL volume of soluble protein lysate was increased from a 50 mL culture to a 500 mL culture. All other conditions remained the same. Figure 4.3 illustrates the resulting protein gel used to analyse this experiment. Recombinant human ADP-GK was visible in the soluble cell lysate, however, the differences in the population of protein in the lysate before and after the resin binding step was negligible. Certainly the amount of recombinant protein still visible in the "post-binding" lysate was significant, which suggested the recombinant protein was not preferentially binding to the Ni-NTA resin. The wash steps appeared somewhat more effective as a protein of approximately 40 kDa was present in each wash step and this had not been observed previously. The 120 mM imidazole elution step still eluted non-recombinant proteins; however, the recombinant protein band was also visible. The ability to visualise the recombinant protein by Coomassie staining was an improvement as previously it was only detectable by silver staining. Once concentrated, the recombinant protein was more clearly visible.

The concentrated eluted sample was used in a standard ADP-GK enzyme assay in an attempt to detect any recombinant ADP-GK activity, but none could be measured. It is unclear whether this was due to incorrect protein folding, interference in the assay by the contaminating *E. coli* proteins or not enough recombinant protein being present in the concentrated eluted sample for enzyme activity to be measured.

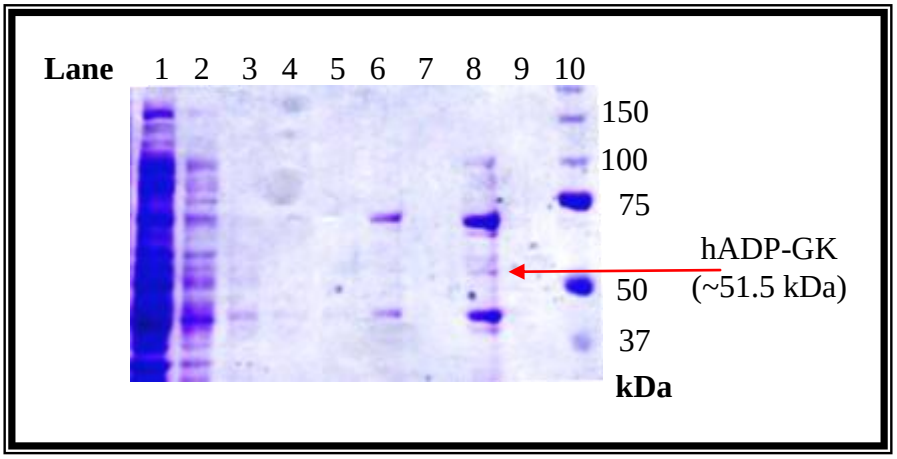


Figure 4.3 SDS-PAGE of a trial purification using Ni-NTA resin

The relative purification of soluble recombinant human ADP-GK protein eluted off the Ni-NTA column was monitored by SDS-PAGE. Five microlitres of 5× treatment buffer was added to each 20 µL protein sample and this was followed by heat treatment at 100°C for five minutes. Twenty-five microlitres of each protein fraction was loaded onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie stain.

Lane one: Soluble protein fraction; **Lane two:** First wash; **Lane three:** Second wash; **Lane four:** Third wash; **Lane five:** Fourth wash; **Lane six:** 120 mM imidazole elution step; **Lane seven:** 250 mM elution step; **Lane eight:** Concentrated 120 mM imidazole eluate; **Lane nine:** Unused; **Lane ten:** Precision Plus Protein™ Dual Colour Standards.

A 500 mL culture of the transformed Rosetta strain of *E. coli* was lysed and a final volume of soluble protein of 20 mL was incubated with 9.6 mL of Ni-NTA resin, and the following purification trial was undertaken: the NaCl concentration was increased to 600 mM in the resin equilibration buffer (50 mM Na_2PO_4, 600 mM NaCl, 5 mM imidazole, pH 8.0) and the binding buffer (50 mM Na_2PO_4, 600 mM NaCl, 5 mM imidazole, 4 mM CHAPS, 20% v/v glycerol, pH 8.0). A sequential decrease in pH was trialled with the wash buffer, with the four wash steps decreasing from pH 8.0 to pH 6.8. The imidazole was raised to 10 mM and the NaCl to 1M, with the aim of increasing the stringency of the washes. A stepwise increase of imidazole was trialled again, with the intention of removing some contaminating proteins before eluting the bulk of the recombinant protein. The elution was six steps (five column volumes each step) as follows: 50 mM, 80 mM, 100 mM, 110 mM, 120 mM and 130 mM imidazole, (pH 8.0).

This experiment resulted in the same conclusion as previous experiments. An insignificant amount of recombinant protein bound to the Ni-NTA resin, relative to other *E. coli* proteins that also bound. Also, the recombinant protein that did bind to the resin appeared to start eluting from the Ni-NTA resin from the first imidazole concentration used, with no single elution step resulting in a more strongly positive recombinant protein band than any other step.

A trial was undertaken which explored whether the amount of imidazole present in the binding buffer influenced how well the recombinant protein bound to the resin. Soluble protein produced from the Rosetta strain of *E. coli* grown at 18°C for 21 hours was incubated with nickel resin at the following imidazole concentrations: 0 mM, 2 mM, 5 mM, 10 mM and 20 mM. The resin and protein was incubated for one hour at 4°C, with shaking. The protein was washed off in one 8 mL wash (50 mM Na_2PO_4, 600 mM NaCl, pH 8.0) with imidazole concentration of the wash step reflecting that of its respective binding buffer. This was followed by a one step elution at 250 mM imidazole. The corresponding denaturing gel showed that imidazole concentration in the binding and wash buffer was not what was influencing the poor yield of recombinant protein. The band representing the recombinant protein appeared essentially equal under all imidazole concentrations.

4.3 Isolating recombinant protein from washed inclusion bodies

As it was clear that most of the recombinant protein expressed was insoluble, it was decided that the best next step, taking time constraints into consideration, was to attempt isolating inclusion bodies and artificially refolding the denatured recombinant human ADP-GK. As had been observed earlier, when protein expression took place at a temperature of 30°C (Figure 3.11), a considerable amount of insoluble recombinant protein was produced. Consequently, this was the expression condition chosen to produce the inclusion bodies to be isolated.

Background research was done to investigate the range of methods used to isolate inclusion bodies and an appropriate trial method was devised. Transformed Rosetta *E. coli* were grown in 500 mL cultures at 30°C for between 18 and 20 hours. The cells were harvested, lysed and the inclusion bodies collected (Section 2.2.12). Once the inclusion bodies had been washed and solubilised in buffer (50 mM Tris-HCl, pH 9.0, 150 mM, NaCl 10 mM β-mercaptoethanol, 8.0 M urea), SDS-PAGE confirmed that the recombinant protein was present. Figure 4.4 shows that the inclusion bodies did not contain purely recombinant human ADP-GK, but a significant proportion of the protein was recombinant human ADP-GK. A preliminary attempt at refolding and purifying the denatured recombinant protein was undertaken by first incubating the denatured protein with Ni-NTA resin and then incubating the resin and any bound protein, in a trial refolding buffer (50 mM HEPES, 300 mM NaCl, 10 mM imidazole, pH 7.4) for one hour at 4°C. This was followed by a native elution using a modified version of the refolding buffer which included imidazole at a concentration of 250 mM. If the recombinant protein had refolded under these particular conditions, the only significant interaction it should have had with the Ni-NTA resin was via its 6×Histidine tag. If this had been the case, the imidazole present in the native elution buffer should have disrupted the Ni-NTA-6×His tag association, thereby causing the protein to elute. Figure 4.4, lane 4 was representative of the native elution, and it is apparent that no protein eluted. To prove that the recombinant protein had indeed bound to the Ni-NTA resin, regardless of its state of folding, a denaturing elution step was undertaken. The elution buffer in this case (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 250

mM imidazole, 8.0 M urea) is based on the principle that hydrophobic interactions between the denatured recombinant protein and the Ni-NTA resin need to be disrupted. This was achieved by the addition of 8.0 M urea, and the Ni-NTA-6×His tag interactions were disrupted by the 250 mM imidazole. Lane five represents the denaturing elution. The recombinant protein was clearly visible and had specifically associated with the Ni-NTA resin as the other *E. coli* proteins present in the solubilised inclusion body sample were no longer apparent.

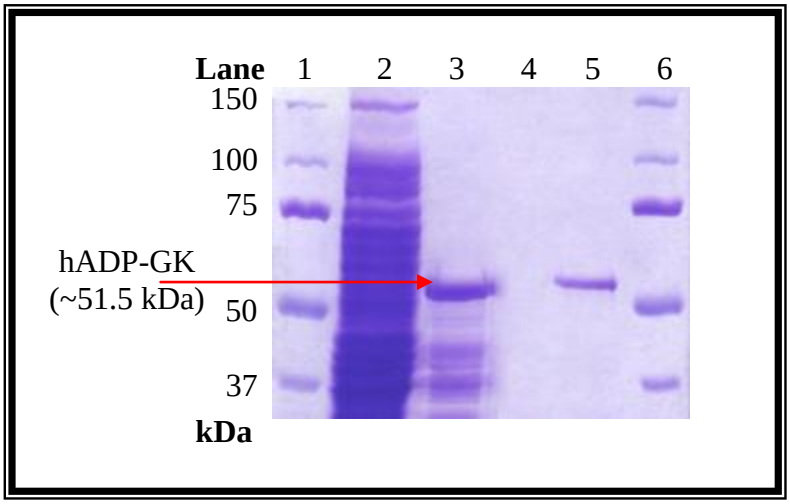


Figure 4.4 Denatured recombinant human ADP-GK isolated from inclusion bodies

Inclusion bodies were isolated from a culture of transformed Rosetta strain of *E. coli* that had been induced to express recombinant protein for between 18 and 20 hours. Protein solubilised from washed inclusion bodies was used in a protein purification trial. Six microlitres of 5× treatment buffer was added to 24 µL protein samples eluted off the Ni-NTA column and this was followed by heat treatment at 100°C for five minutes. Thirty microlitres of each protein fraction was loaded onto an 8% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie stain.

Lane one: Precision Plus Protein™ Dual Colour Standards; **Lane two:** Whole cell lysate; **Lane three:** Washed and solubilised inclusion body protein; **Lane four:** Native elution; **Lane five:** Denaturing elution; **Lane six:** Precision Plus Protein™ Dual Colour Standards.

4.4 Refolding recombinant human ADP-GK protein

The successful refolding of denatured protein often requires very specific conditions. In the interest of using time efficiently, it was decided that the most effective way to approach this task was to screen a range of refolding conditions in a single experiment. Vincentelli *et al*., (2004) reported a refolding screen using a 96-well plate format. The conditions screened followed a logic that exposed the protein to be refolded to a variety of the most common conditions that influence protein solubility. This same logic was used in choosing the conditions to screen refolding of the recombinant ADP-GK. Vincentelli *et al*., (2004) used what they termed a "fractional factorial approach" to their screen, because if every combination of the 20 chemicals they had selected were used, it would have resulted in 2560 combinations. Based on the conditions they chose for the first 80 wells of their 96 well plate (the remainder involved the presence of molecular chaperones) a 57 condition screen was designed.

A wide pH range is useful as proteins are likely to behave differently at a pH below or above their pI. Buffers of pH 4.0 (50 mM NaAc), 5.0 (50 mM MES), 7.0 (50 mM HEPES), 8.0 and 9.0 (50 mM Tris) were selected. Ionic strength is important as solubility can increase or decrease at different ionic strengths. To address this, final concentrations of 100 mM KCl, 100 mM NaCl, or 200 mM NaCl were used. Refolding by direct dilution was the only way this screen could be carried out, so the use of amphiphilic components was important to rule out that the interaction of hydrophobic residues exposed during the intermediate stages of protein refolding was causing protein precipitation. Glycerol and PEG were used for this purpose along with the detergents Triton X-100 and CHAPS. Glucose and arginine are often included in commercially available refolding kits and are known to influence the refolding of proteins in some instances, so these were also added to the screen. β-mercaptoethanol was an additive of the screen to prevent incorrect disulfide bonds being formed, which can also cause precipitation of protein.

A "cocktail" based on that of Vincentelli *et al*., (2004) was created, containing potential co-factors that may be required during the refolding process. In this case it contained divalent

cations and ADP. EDTA was also used as divalent cations can also cause protein precipitation.

An aliqout (90 µL) of the specific refolding buffer being screened was added to each well. The absorbance of the buffers was measured at a wavelength of 350 nm, to measure the background absorbance of the potential refolding buffers. To assist in the analysis of how successful a refolding buffer was, measurement of absorbance at 350 nm was to be undertaken. Absorbance was used to help detect precipitation as it can be difficult to visualise in a plate. Vincentelli *et al*., (2004) demonstrated that the best wavelength to measure the light scattering properties of precipitated protein in refolding studies was 350 nm.

An aliquot (5 µL) of solubilised recombinant protein at a concentration of approximately of 6.0 mg/mL was added to each of the 57 wells. This gave a final protein concentration of approximately 0.3 mg/mL, a 20-fold dilution of the solubilised, denatured protein. Addition of protein to all the wells took approximately 10 minutes, and this was followed directly by measuring the absorbance of the 96-well plate again. Absorbance was measured on the premise that any immediate precipitant formed as a result of incompatibility with the refolding buffer would be detected at 350 nm (Vincentelli *et al*., 2004). As it turned out, precipitate was easily visible and the absorbance readings were not particularly accurate, as some wells that contained precipitate did not have an increased absorbance.

After a period of 24 hours, the plate was analysed and a clear trend was observed. Any set of refolding conditions at a pH of 5.0 produced precipitate; however, no precipitant was formed in any of the refolding trials at pH 8.0 or pH 9.0, with or without additives. Based on these results, it was decided that a refolding trial on a larger basis using Tris buffer at pH 8.0, and no other additives, should be undertaken. It is often better to avoid any additives if possible, and add them only as necessary, as many additives can interfere with protein purification techniques further along in the process.

One millilitre of solubilised protein at a concentration of 0.6 mg/mL was introduced slowly by direct dilution into 19 mL of Tris buffer pH 8.0. No precipitate was visualised so another 1 mL of solubilised protein was added to the buffer to give an approximate protein concentration of 0.06 mg/mL. No precipitate was formed. To assess whether the solubilised protein in the Tris buffer had refolded, an enzyme assay was undertaken under standard assay conditions. An aliquot of 40 µL of the protein solution was added to the assay and visible activity was observed for a period of four minutes.

NaAc (pH 4.0) β-mercaptoethanol 800 mM Arginine	NaAc (pH 4.0) β-mercaptoethanol 35 mM CHAPS	MES (pH 5.0) 1 mM EDTA	HEPES (pH 7.0) 0.05% PEG 3250 β-mercaptoethanol	Tris (pH 8.0) 100 mM KCl β-mercaptoethanol 800 mM Arginine 4 mM CHAPS	Tris (pH 9.0) β-mercaptoethanol 500 mM Glucose 4 mM CHAPS
NaAc (pH 4.0) β-mercaptoethanol 100 mM KCl 35 mM CHAPS	MES (pH 5.0)	MES (pH 5.0) 100 mM NaCl 800 mM Arginine	HEPES (pH 7.0) 200 mM NaCl 4 mM CHAPS	Tris (pH 8.0) β-mercaptoethanol	Tris (pH 9.0) 200 mM NaCl β-mercaptoethanol 35 mM CHAPS
NaAc (pH 4.0) 100 mM NaCl 4 mM CHAPS 500 mM Glucose	MES (pH 5.0) 35 mM CHAPS	MES (pH 5.0) β-mercaptoethanol 35 mM CHAPS	HEPES (pH 7.0) 20% Glycerol β-mercaptoethanol	Tris (pH 8.0) 500 mM Glucose	Tris (pH 9.0) β-mercaptoethanol 800 mM Arginine 1 mM EDTA
NaAc (pH 4.0) 100 mM NaCl β-mercaptoethanol	MES (pH 5.0) Cocktail	MES (pH 5.0) 0.05% PEG 3250 β-mercaptoethanol 4 mM CHAPS	HEPES (pH 7.0) 35 mM CHAPS 800 mM Arginine 0.05% PEG 3250	Tris (pH8.0) 35 mM CHAPS 500 mM Glucose	Tris (pH 9.0)
NaAc (pH 4.0) 0.05% PEG 3250 500 mM Glucose	MES (pH 5.0) 1 mM EDTA β-mercaptoethanol 4 mM CHAPS	HEPES (pH 7.0) β-mercaptoethanol	HEPES (pH 7.0) Cocktail	Tris (pH8.0) 500 mM Glucose 200 mM NaCl β-mercaptoethanol	Tris (pH 9.0) 20% Glycerol 4 mM CHAPS
NaAc (pH 4.0) 20% Glycerol	MES (pH 5.0) β-mercaptoethanol 20% Glycerol	HEPES (pH 7.0)	Tris (pH 8.0) β-mercaptoethanol 500 mM Glucose	Tris (pH8.0) β-mercaptoethanol 20% Glycerol	Tris (pH 9.0) 100 mM NaCl 500 mM Glucose
NaAc (pH 4.0) 1 mM EDTA	MES (pH 5.0) 1 mM EDTA	HEPES (pH 7.0) 4 mM CHAPS 800 mM Arginine	Tris (pH 8.0) 800 mM Arginine	Tris (pH8.0) 100 mM NaCl 1 mM EDTA	Tris (pH 9.0) 0.05% PEG 3250 β-mercaptoethanol 500 mM Glucose
NaAc (pH 4.0) 4 mM CHAPS	MES (pH 5.0) 200 mM NaCl	HEPES (pH 7.0) β-mercaptoethanol	Tris (pH 8.0) 100 mM NaCl β-mercaptoethanol 35 mM CHAPS	Tris (pH 9.0) β-mercaptoethanol	
NaAc (pH 4.0)	MES (pH 5.0) 100 mM KCl 500 mM Glucose	HEPES (pH 7.0) 100 mM NaCl 4 mM CHAPS	Tris (pH 8.0) 4 mM CHAPS	Tris (pH 9.0) Cocktail	
NaAc (pH 4.0) 500 mM Glucose Cocktail	MES (pH 5.0) β-mercaptoethanol	HEPES (pH 7.0) β-mercaptoethanol 800 mM Arginine	Tris (pH 8.0)	Tris (pH 9.0) 100 mM NaCl 35 mM CHAPS	

Table 4.1 Detailed representation of the refolding screen undertaken using 57 wells of a 96 well plate

4.5 Purification of refolded recombinant human ADP-GK

It was decided that purification of recombinant protein using nickel resin would be most effective using a 5 mL HisTrap column on the ÄKTA HPLC system as it is a much more sensitive method in terms of controlling the imidazole concentration and the volume of eluted fractions collected.

The re-folding of recombinant protein, from solubilised inclusion bodies, was scaled up from 20 mL to 2.0 L. The 2.0 L solution was concentrated in batches using an Amicon stirred cell concentrator with a 30 kDa cut-off membrane. Nitrogen gas was applied at a pressure of approximately 60 psi to concentrate the stirred protein solution. A final solution of a volume of 40 mL and approximate protein concentration of 0.6 mg/mL (protein precipitation did occur during the process of concentrating) was then used in purification trials.

To begin with, a gradient elution protocol was used. Approximately 40 mL of soluble protein solution was loaded onto the 5 mL column at a rate of 1 mL/min. A wash of 10 column volumes was performed using 0% buffer B. Buffer B had a total imidazole concentration of 500 mM; buffer A had a total imidazole concentration of 5 mM. A gradient of 0 to 100% buffer B was undertaken over 20 column volumes. As can be seen in Figure 4.8, a small peak appeared at between 10% and 20% buffer B. The SDS-PAGE of this peak shows bands representing proteins of a molecular mass at or below 50 kDa. The second broad peak appeared at between 40% and 60% buffer B. As is observed on the protein gels, a strong band representing the recombinant human ADP-GK is present in each of the fractions collected from the peak. At the end of the gradient elution, the band of human ADP-GK gradually disappears. Interestingly, a band was visible at approximately 103 kDa and followed the same elution pattern as that of the recombinant human ADP-GK. It was considered that the band represented a dimer of the recombinant protein as it was approximately double the molecular mass of the human ADP-GK. This preliminary purification trial proved that the recombinant human ADP-GK did bind efficiently to nickel

resin and eluted at about 200-300 mM imidazole. Optimisation was still required as smaller molecular mass proteins still appeared in the elution fractions of the recombinant ADP-GK.

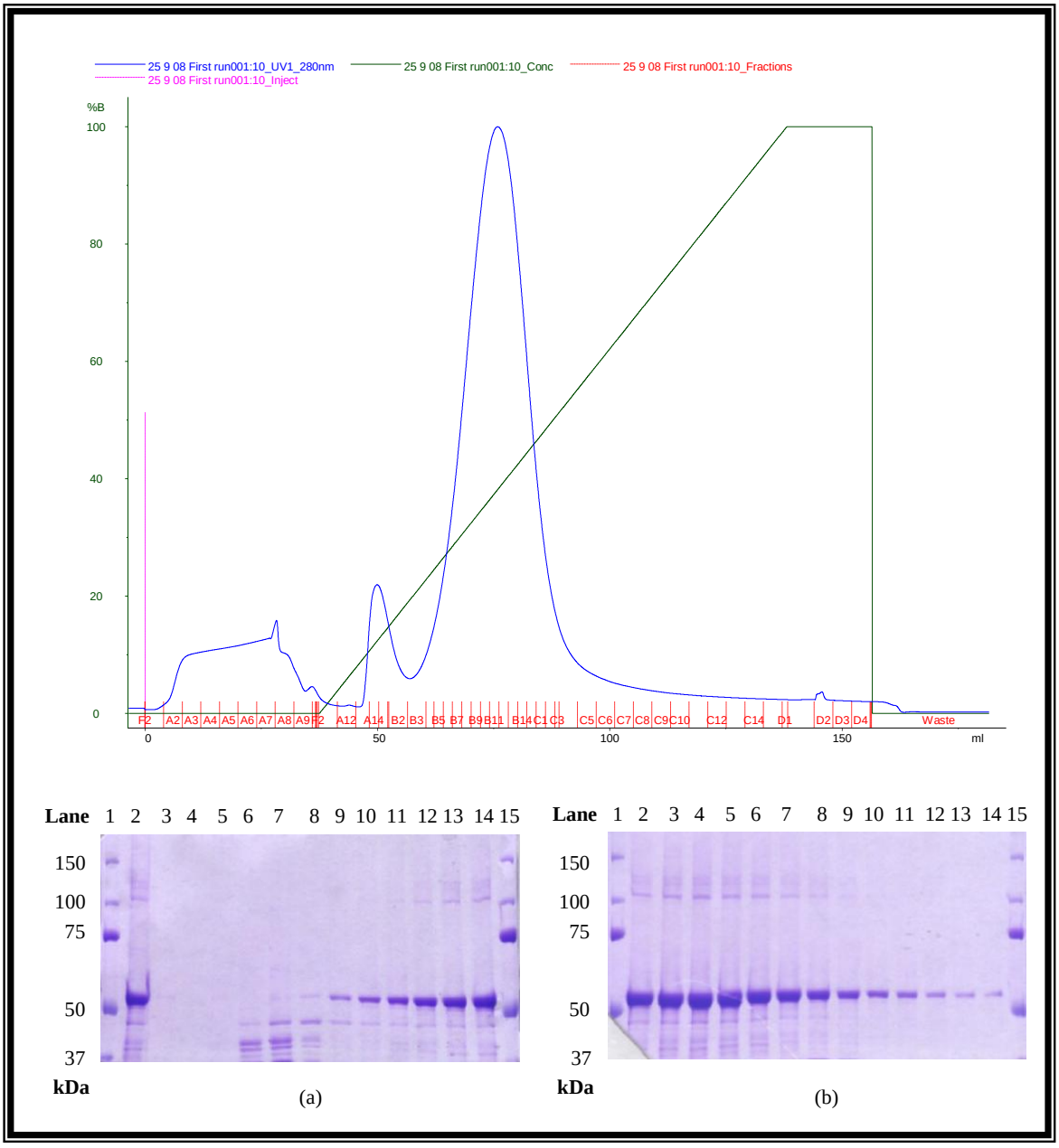


Figure 4.5 Elution schematic and resulting SDS-PAGE for purification trial

The relative purification of soluble recombinant human ADP-GK protein eluted off the HisTrap™ HP 5 mL affinity column was monitored by SDS-PAGE. Three microlitres of 5× treatment buffer was added to 12 µL samples of eluted protein and this was followed by heat treatment at 100°C for five minutes. Twelve microlitres of each protein fraction was loaded onto a 10.5% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie stain.

Gel (a): Lanes one and fifteen: Precision Plus Protein™ Dual Colour Standards; **Lane two:** Soluble protein loaded onto Ni-NTA column; **Lane three:** Eluted protein fraction A7; **Lane four to Lane seven:** Eluted protein fractions A12 to A15; **Lane eight to Lane fourteen:** Eluted protein fractions B2 to B8.

Gel (b): Lane one and fifteen: Precision Plus Protein™ Dual Colour Standards; **Lane two to Lane eight:** Eluted protein fractions B9 to B15; **Lane nine to Lane fourteen:** Eluted protein fraction C1 to C6.

In an endeavour to optimise the purification and eliminate the smaller molecular mass contaminating proteins, a stepwise elution protocol was trialled with seven steps of buffer B as follows; 10%, 20%, 25%, 30%, 35%, 45% and 100%. Each step lasted for 5 column volumes and 1 mL fractions were collected. At the 10% buffer B step (lane (a)), many contaminating proteins were eluted from a range of molecular masses, which was a positive sign. Lane (e) of Figure 4.6 represents the elution of the most concentrated amount of human ADP-GK. Unfortunately, still visible were the small molecular mass contaminating proteins. The recombinant protein band appeared less intense in lanes (g) and (h), and the contaminating proteins were hardly visible. It was thought that the contaminating proteins were still being eluted, but were not as visible due to dilution.

Still present, and mimicking the elution pattern of the recombinant protein, were the protein bands of approximately two and three times the molecular mass of the recombinant human ADP-GK. Although the gel had been run under denaturing conditions, these bands were thought to represent persistent dimers and trimers of the recombinant protein. It has been observed that protein-protein interactions often occur during the protein refolding process, and the aggregates formed often consist of chains with substantial native secondary structure and interact via a number of non-native hydrophobic contacts. This occurs most often when refolding of a protein is taking place at a temperature below the optimal range, where proteins tend to aggregate (Smith and Hall, 2001). Temperature had not been screened as one of the conditions of refolding the recombinant protein, and it is possible that optimising temperature conditions during the refolding process would have influenced the creation of the apparent aggregates.

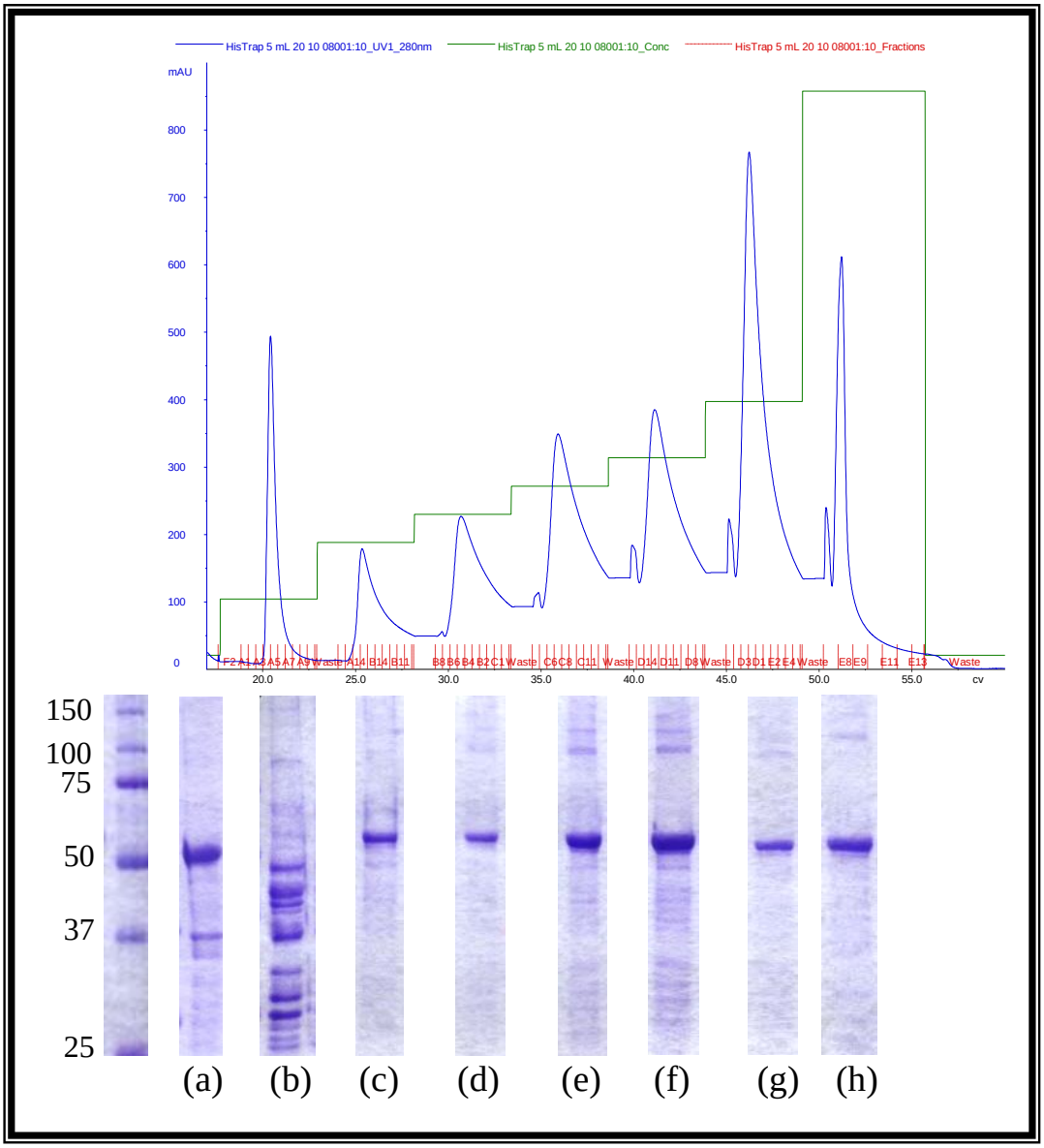


Figure 4.6 Elution schematic and resulting SDS-PAGE for purification trial

The relative purification of soluble recombinant human ADP-GK protein eluted off the HisTrap™ HP 5 mL affinity column was monitored by SDS-PAGE. Three microlitres of 5× treatment buffer was added to 12 µL samples of eluted protein and this was followed by heat treatment at 100°C for five minutes. Twelve microlitres of each protein fraction was loaded onto a 10.5% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie stain.

Lane (a): Soluble protein loaded onto Ni-NTA column; **Lane (b):** Fraction A6; **Lane (c):** Fraction B14; **Lane (d):** Fraction B3; **Lane (e):** Fraction C9; **Lane (f):** D10; **Lane (g):** Fraction E2; **Lane (h):** Fraction E8.

As splitting the elution protocol into multiple steps did not achieve purification to homogeneity, it was decided that using a one step elution may give a better yield of recombinant protein overall, instead of splitting the protein into multiple fractions and thereby diluting it. An elution at 40% buffer B was chosen and this worked effectively as can be observed in Figure 4.7. As the contaminating proteins seemed to follow the elution pattern of the recombinant human ADP-GK, it was decided that an immunoblot should be carried out. It was expected an immunoblot could help to determine if the smaller molecular mass proteins were in fact truncated versions of the recombinant protein caused by incomplete expression or perhaps degradation products, instead of different protein species.

The 10% elution peak was the only imidazole concentration that seemed specific for protein bands not including the 51.5 kDa recombinant protein band. Therefore a sample from the 10% elution peak, the 40% elution peak, the denatured and solubilised protein isolated from washed inclusion bodies, and the refolded protein from the direct dilution protocol were used in an immunoblot. Figure 4.8 illustrates that the mouse monoclonal antibody, specific for ADP-GK, reacted with more than one band that represented proteins of differing molecular masses. This occurred in a repeated immunoblot, confirming the specific interaction. Lane (a) of Figure 4.8 represents approximately 4.8 µg of protein isolated from inclusion bodies followed by solubilisation in 8.0 M urea. Lane (b) represents approximately 0.36 µg of protein refolded by direct dilution. Lane (c) represents 12 µL of a 1 mL fraction from the 10% buffer B wash. Lane (d) represents 1.2 µg of a 1 mL fraction from the 40% buffer B elution.

As lane (c) appears to have a greater amount of proteins smaller than 51.5 kDa than that of the solubilised protein in lane (a), it was considered that the recombinant protein was

degrading over time, hence the build up of smaller molecular mass products that reacted with the ADP-GK monoclonal antibody. In the purified protein lane, apart from higher molecular mass products thought to be aggregates of recombinant protein as dimers and trimers, only two persistent bands that reacted with the ADP-GK antibody were present.

In attempt to remedy this supposed situation, 0.1 mM PMSF (serine protease inhibitor) and 0.1 mM DTT (to prevent the formation of multimers) was added to all the buffers used to isolate and purify the recombinant protein to inhibit proteases. A full purification was undertaken with the addition of PMSF and DTT, however the resulting SDS-PAGE showed that the smaller and higher molecular mass proteins were still present in the same abundance as without the DTT and PMSF.

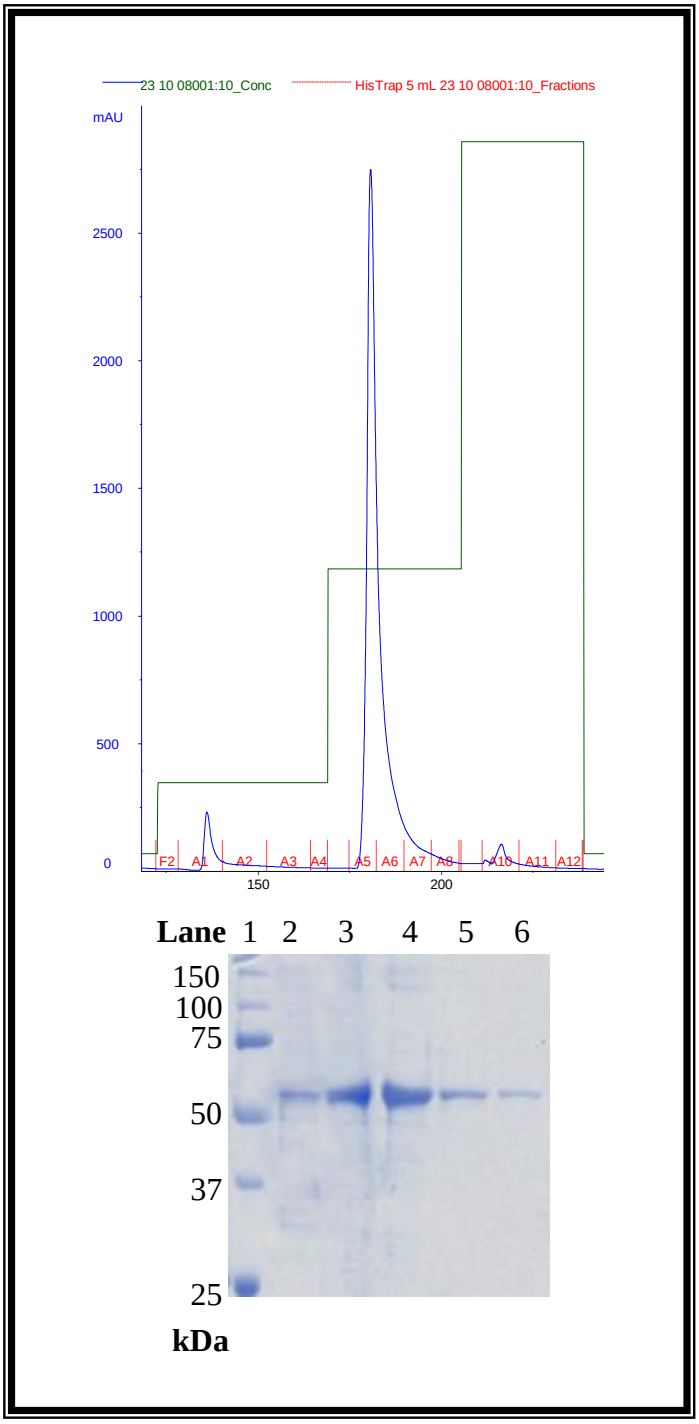


Figure 4.7 Elution schematic and resulting SDS-PAGE for purification trial

The relative purification of soluble recombinant human ADP-GK protein eluted off the HisTrap™ HP 5 mL affinity column was monitored by SDS-PAGE. Three microlitres of 5× treatment buffer was added to 12 µL samples of eluted protein and this was followed by heat treatment at 100°C for five minutes. Twelve microlitres of each protein fraction was

loaded onto a 10.5% acrylamide denaturing protein gel. Electrophoresis was carried out at 120 V for approximately 120 minutes. The resulting gel was stained with Coomassie stain. **Lane one:** Precision Plus Protein™ Dual Colour Standards; **Lane two:** Soluble protein loaded onto Ni-NTA column; **Lane three:** Fraction A5; **Lane four:** Fraction A6; **Lane five:** Fraction A7; **Lane six:** Fraction A8.

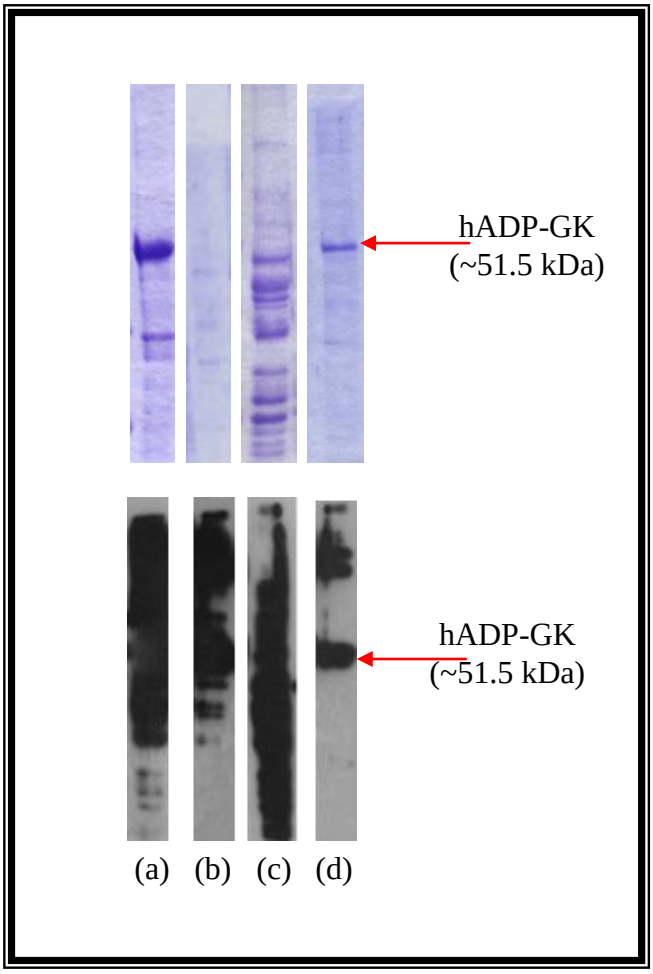


Figure 4.8 Immunoblot of ADP-GK pre- and post- HisTrap column purification

An immunoblot was carried out to help determine the identity of bands visible by SDS-PAGE that were not the 51.5 kDa recombinant human ADP-GK band. A 1:1000 dilution of the mouse primary monoclonal ADP-GK antibody was used, followed by a 1:1000 dilution of horseradish peroxidise-conjugated secondary anti-mouse antibody. The horseradish peroxidase-conjugated secondary antibody was detected by enhanced chemiluminescence. Light produced by the reaction was detected in a dark room by X-ray film.

Upper panel: Coomassie stain; **Lower panel:** Immunoblot.

Lane (a): Solubilised protein isolated from washed inclusion bodies; **Lane (b):** Refolded protein recombinant protein; **Lane (c):** Fraction collected from 10% buffer B wash; **Lane (d):** Purified recombinant human ADP-GK eluted at 40% buffer B.

4.6 Chapter summary

Purification of the human ADP-GK recombinantly expressed in the Rosetta strain of *E. coli* was trialled using Ni-NTA resin and 10 mL plastic columns. Using non-denaturing conditions and SDS-PAGE to monitor the efficiency of the purification process, it became clear that enrichment of the recombinant human ADP-GK protein was not being achieved. A number of conditions were trialled in an attempt to purify the recombinant protein, including changes in imidazole and salt concentrations in the equilibration buffer and the elution buffers, but with no clear success. Due to time constraints, it was decided that approaching the purification of the recombinant human ADP-GK using denaturing conditions was appropriate, as the recombinant protein was expressed at high levels, but was insoluble. A protocol for isolating, washing and solubilising inclusion bodies formed in the Rosetta expression strain of *E. coli* after 18 to 20 hours of expression at 30°C was devised, and this resulted in a considerable yield of partially purified recombinant protein.

A trial refolding process was designed based on a trial used by Vincentelli *et al*., (2004) which utilised a 96 well plate to screen multiple refolding conditions simultaneously. After carrying out the refolding trial, it became clear that the recombinant protein did not precipitate out after direct dilution into refolding buffer that had a pH of 8.0 or 9.0, regardless of additives. This led to a larger scale trial of refolding by direct dilution into Tris buffer at pH 8.0. No precipitate was visualised, so a standard ADP-GK enzyme assay was set up. Enzyme activity was subsequently detected and continued for approximately four minutes, which indicated that at least a portion of the recombinant protein present in solution had refolded.

Purification trials were begun using recombinant protein that had been isolated from inclusion bodies and refolded in Tris buffer at pH 8.0. Using a 5 mL HisTrap column on the ÄKTA HPLC system, an imidazole gradient was used in a protein purification trial. It was apparent that the refolded protein was specifically associating with the nickel resin and it eluted at a total imidazole concentration of between 200 and 300 mM. Contaminating protein bands of a lesser molecular mass than the recombinant protein were visible by SDS-

PAGE. Attempts were made to isolate the recombinant protein without the presence of these contaminating proteins by optimising the concentration of imidazole in the elution step. It was found that the elution of contaminating proteins did occur at 50 mM imidazole, but this did not eliminate them from the recombinant protein elution fractions. Finally 200 mM imidazole was optimised as the concentration for the highest yield of recombinant protein. It was thought that perhaps the smaller molecular mass proteins were possibly degradation productions due to proteases present in buffers used in the purification process so a trial was undertaken which incorporated DTT and PMSF in every buffer. This, however, did not change the range of eluted protein products.

Due to their very similar elution pattern and the detection of these proteins in immunoblots performed using the ADP-GK monoclonal antibody, it was concluded that these proteins could actually be truncated expression products caused by the rapid rate of recombinant expression of the human ADP-GK. Overall, the percentage of protein bands isolated by the purification process visible by SDS-PAGE, other than the 51.5 kDa recombinant human ADP-GK band, was very low.

Chapter 5: Recombinant protein characterisation

5.1 Introduction

The extended period of time taken to produce soluble and pure recombinant human ADP-GK protein meant that the kinetic characterisation of the protein could not be carried out. Instead, only preliminary analysis both the protein's secondary structure and its kinetic behaviour in response to pH was completed, as described in this chapter.

Circular dichroism (CD) is a technique based on the optical absorbance of protein in the wavelength range of 170 -240 nm (far UV). Absorption that occurs below the wavelength of 240 nm is due predominantly to the backbone orientation of peptide bonds, with minor influences due to side-chains. As regular secondary structures of different types give rise to characteristic CD spectra in this region, CD can be used to analyse the secondary structure of proteins in solution.

Circular dichroism refers to the differential absorption of the two components that make up plane polarised light; one of left-handed polarity which rotates counter-clockwise and one of right-handed polarity which rotates clockwise. These two circularly polarised components are of equal magnitude. CD instruments measure the difference in absorbance between the left and right handed circularly polarised components after the light has passed through a protein solution (Kelly *et al*., 2005).

Since recombinant human ADP-GK could not be purified in a soluble form from the Rosetta expression strain of *E. coli,* and had to instead be refolded from denatured protein purified from inclusion bodies, it was deemed necessary to check the refolded protein for detectable secondary structure. As the coding sequences for the mouse ADP-GK (Accession BC021526.1) and the human ADP-GK (Accession BC006112) have 89% identity, (BLAST Local Alignment Search tool), it was estimated that the two proteins should have similar secondary structure, and therefore the spectrum generated by the recombinant mouse ADP-GK protein would be a useful comparison.

5.2 Analysis of refolded recombinant human ADP-GK secondary structure

Analysis by circular dichroism requires protein to be dialysed into a minimal buffer, or, preferably, water. This is due to the fact that many buffers or components of buffers can interfere with analysis of protein secondary structure by absorbing at wavelengths of interest. Time was very limited as this stage of the project, so a buffer trial could not be carried out to optimise the stability of the recombinant proteins in minimal buffer. As a result, the final protein concentration achieved in minimal buffer (10 mM K_2HPO_4 and 20 mM KCl) was 0.1 mg/mL.

A 0.1 cm high-transparency quartz cuvette was used to collect spectral data. The wavelength to be scanned was set between 180 nm and 260 nm with 0.5 nm steps, and the bandwidth at 1.0 nm. The time taken to measure absorbance at each point was set at 0.5 seconds and the scan was performed in triplicate. The absorbance spectrum of the buffer alone was measured first, in order for background absorbance to be subtracted from the spectra of the protein solutions. Data was visualised, averaged, smoothed and overlaid using the modular Pro-Data instrument software, specific for the Chirascan™ Circular Dichroism Spectrometer (http://www.photophysics.com/pdf/chirascaninfo.pdf).

The data collected (in millidegrees) over the wavelengths 190-240 nm for the mouse and the human recombinant proteins were then converted to molar absorptivity by the following equations (Greenfield, 2006):

Ellipticity, $\mathbf{[\theta]}$ = millidegrees / pathlength in millimetres × the molar concentration of protein ×the number of amino acid residues in the protein

Molar absorptivity, $\mathbf{\Delta\varepsilon = [\theta]} / 3\,298$

Once the data (Appendix 4) had been converted to molar absorptivity over the wavelength range of 190-240 nm, it was submitted to the K2D2 web server (http://www.ogic.ca/projects/k2d2) to give a percentage estimation of secondary structure

(Tables 5.1 and 5.2). K2D2 is a method which uses a self-organised map of spectra from proteins with known structure to deduce a map of protein secondary structure. This is used for the prediction of protein secondary structure in proteins with unknown structure. The K2D2 method is trained using CD spectra from 43 soluble proteins with a variety of secondary structure compositions, such as mainly alpha helix, mainly beta sheet, and proteins made up of both types of secondary structure. K2D2 compares well with other published methods for the prediction of protein secondary structure from CD spectra, and it is unique in its ability to warn the user when the prediction between the user's input spectrum and the one computed from the training set is not reliable (Perez-Iratxeta and Andrade-Navarro, 2008).

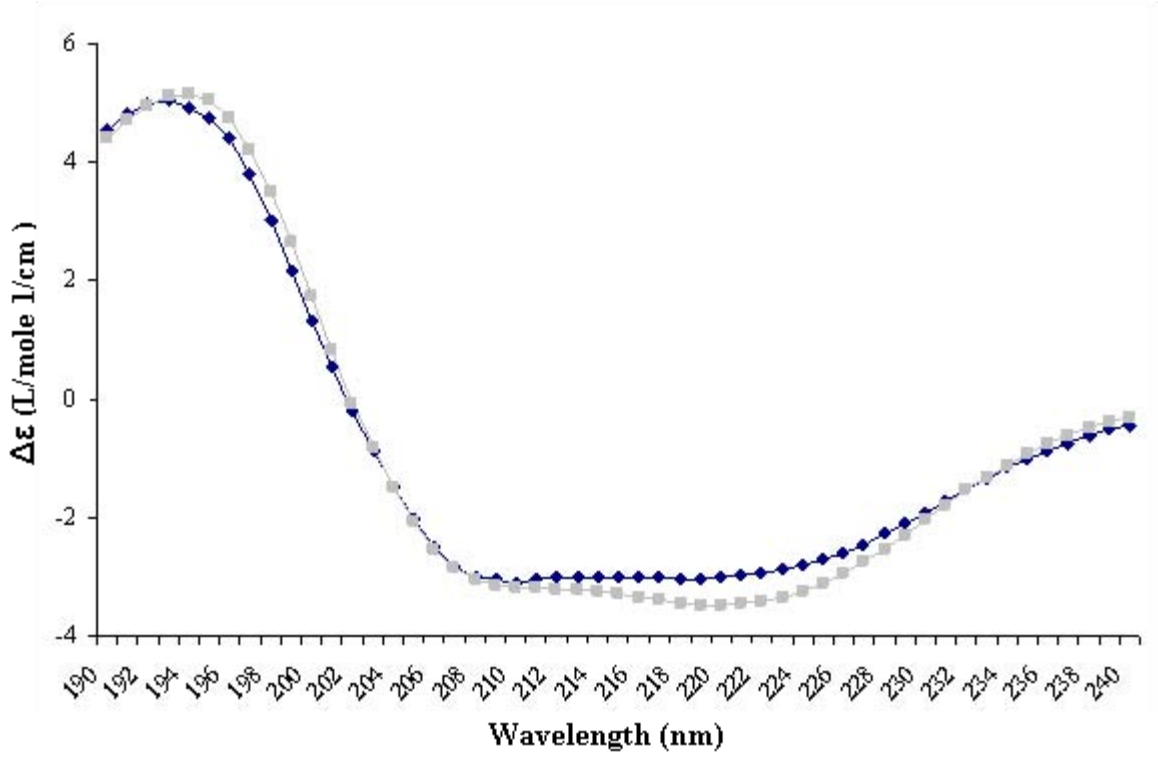


Figure 5.1 Circular dichroism spectrum of recombinant human ADP-GK

The relative distance between circular dichroism query (the molar absorptivity values measured for the recombinant human ADP-GK over the wavelength range of 190-240 nm) and the circular dichroism estimate was 0.96. This is within the limit of reliability, so the estimated secondary structure percentages calculated can be taken into consideration.

Structure	Estimate	Error
Alpha	42.9 %	6.1 %
Beta	19.1 %	14.2 %
Turn	11.6 %	2.3 %
Random	26.5 %	7.5 %

Table 5.1 K2D2 estimated percentages of secondary structure for the recombinant human ADP-GK

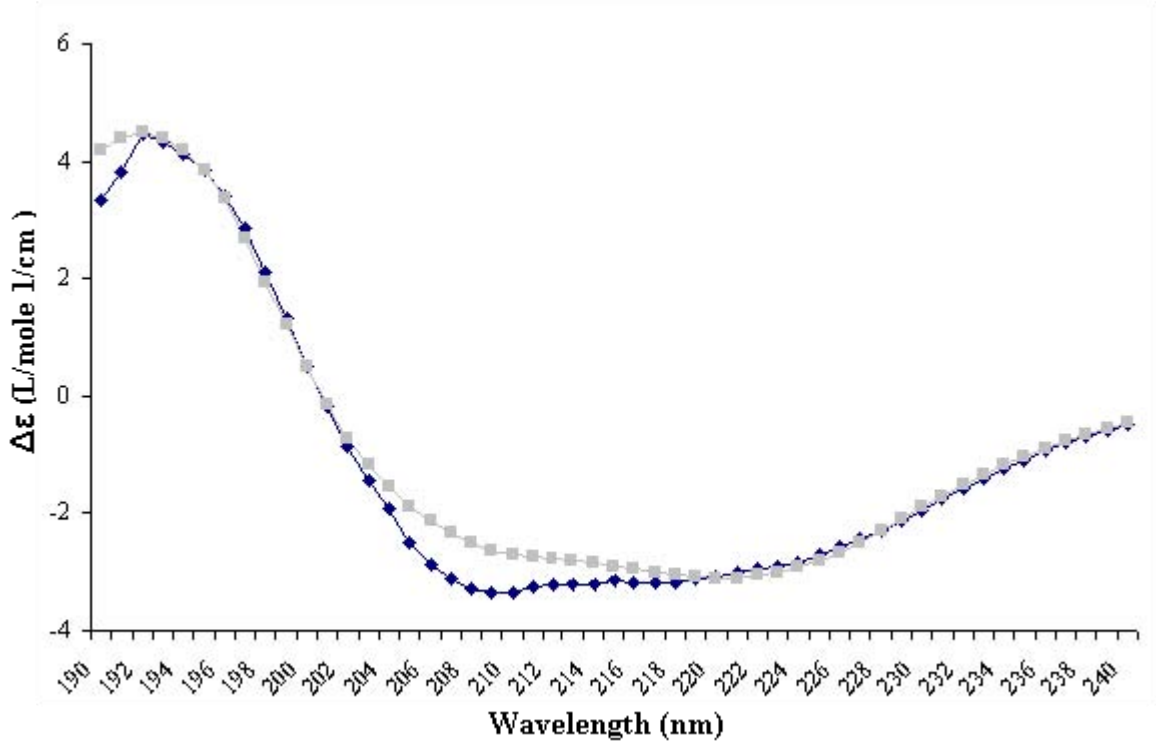


Figure 5.2 Circular dichroism spectrum of recombinant mouse ADP-GK

The relative distance between circular dichroism query (the molar absorptivity values measured for the recombinant mouse ADP-GK over the wavelength range of 190-240 nm) and the circular dichroism estimate was 1.22. This is within the limit of reliability, so the estimated secondary structure percentages calculated can be taken into consideration.

Structure	Estimate	Error
Alpha	41.8 %	3.8 %
Beta	12.3 %	6.7 %
Turn	11.8 %	3.7 %
Random	34.2 %	6.4 %

Table 5.2 K2D2 estimated percentages of secondary structure for the recombinant mouse ADP-GK

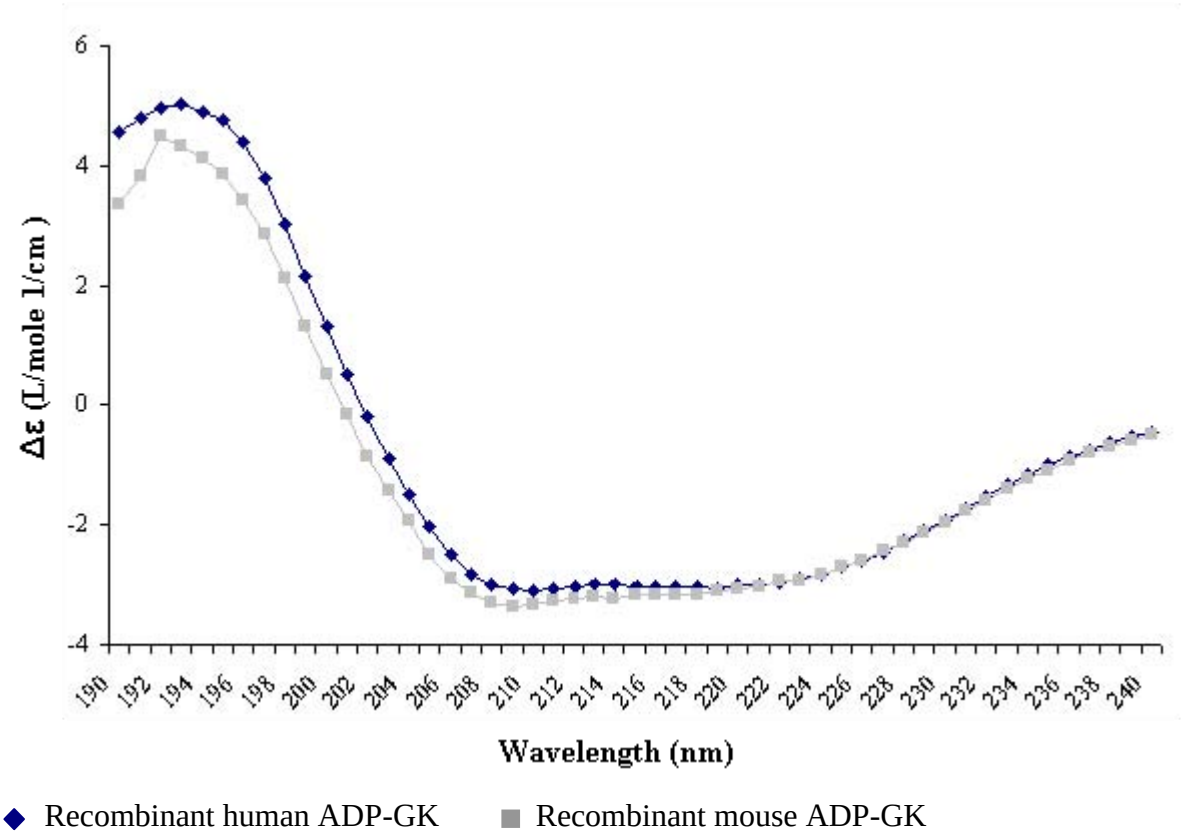


Figure 5.3 Circular dichroism spectrum of recombinant mouse and human ADP -GK

Based on the K2D2 predicted percentages (Tables 5.1 and 5.2) of alpha helix (42.9 % and 41.8 %, respectively) and turns (11.6 % and 11.8 %, respectively), it would appear that the recombinant mouse and human proteins have similar secondary structure. The estimated percentage of beta sheet is not very close between the two proteins, however, it has been noted that the error in the estimation of beta sheet for the recombinant human ADP-GK had a very large margin.

5.3 Enzyme activity relative to pH

Due to time constraints, a complete investigation of the kinetic characteristics of the recombinant human ADP-GK could not be carried out as planned. However, enough recombinant protein was purified to undertake a preliminary screen of the specific activity of the protein over a pH range of 5.0 to 9.0.

It is well known that the pH of an assay can influence the activity of an enzyme. Most enzymes have a narrow or very specific pH range in which they work optimally. The recombinant mouse enzyme appeared to have bi-modal pH optima (between pH 5.75 and 6.5 and one near 8.75–9.0) (Ronimus and Morgan, 2004), so it was of interest to compare the behaviour of the recombinant human ADP-GK over this pH range to see if it too behaved in a similar fashion. The stability of the recombinant protein was very poor in refolding buffer, and although the addition of up to 20% glycerol had been trialled, a concentration of greater than 0.2 mg/mL was not stable. Stability trials will be required if the kinetic characterisation of the protein is to be carried out fully.

The maximum volume of recombinant protein that could be added to the assay if all the stock buffer concentrations were to remain the same was 55.8 µL. The recombinant protein solution had a concentration of 0.15 mg/mL, so each assay contained 8.37 µg of recombinant human ADP-GK. The assays were conducted with the use of Bis-Tris propane, as the pH range being examined could be studied with this one buffer. All enzyme assays were performed in triplicate at 37°C using 100 µL microcuvettes. The standard assays contained 1.0 mM NAD^+, 100 mM KCl, 0.35 mM glucose, a 1:1 ratio of ADP:$MgCl_2$ (1 mM), and 50 mM Bis-Tris propane buffer. To each assay was added 0.2 units of glucose-6-phosphate dehydrogenase, and the assay was initiated by addition of the recombinant ADP-GK. The assay was monitored for 120 seconds with an approximate 30 second pre-incubation (Table 5.3).

Specific Activity*	pH	Approximate lag time	Averaged specific activity
0.127	pH 5.0	No lag	0.134
0.134			
0.140			
0.117	pH 5.5	No lag	0.114
0.115			
0.111			
0.104	pH 6.0	No lag	0.103
0.102			
0.104			
0.065	pH 6.5	50 sec	0.063
0.063			
0.061			
0.033	pH 7.0	100 sec	0.034
0.033			
0.035			
0.015	pH 7.5	100 sec	0.012
0.012			
0.010			
0.012	pH 8.0	150 sec	0.011
0.012			
0.010			
-	pH 8.5	-	-
-			
-			
-	pH 9.0	-	-
-			
-			

Table 5.3 pH dependent specific activity values *(µmol/min/mg) for recombinant human ADP-GK

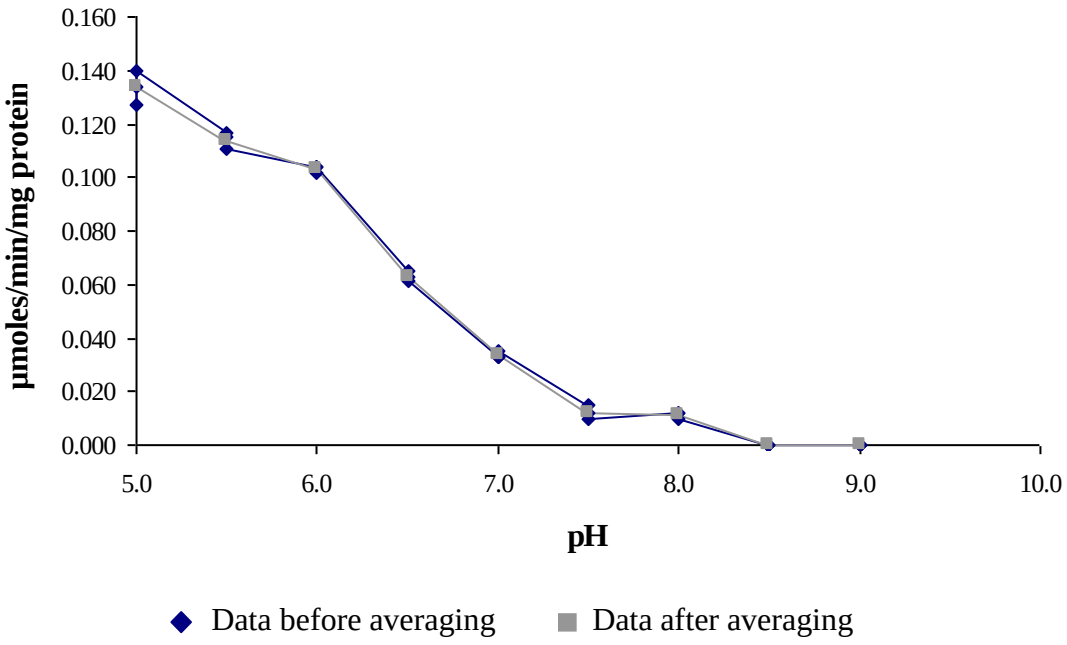


Figure 5.4 Effect of varying pH on recombinant human ADP-GK activity

The preliminary pH screen demonstrated that the behaviour of the recombinant human ADP-GK over the pH range of pH 5.0-9.0 varies greatly to that of the recombinant mouse ADP-GK. Instead of bi-modal optima, the recombinant human ADP-GK had optimum activity at the acidic end of the pH range and activity decreased as pH became more alkaline. By pH 8.5, no activity could be detected (Figure 5.4).

5.4 Chapter summary

Maintaining the purified recombinant protein in a soluble state at a useful concentration proved to be practically impossible under the conditions trialled in this study. In order for the purified protein to be used for in-depth kinetic characterisation studies, the buffer in which the protein is stored will need to be optimised.

Despite this instability, the preliminary circular dichroism study showed the refolded recombinant human ADP-GK did have significant secondary structure (Figure 5.1), and the estimated percentages of alpha helix and turns demonstrate that the mouse and human recombinant ADP-GK proteins probably have similar overall secondary structure.

Although these enzymes may be similar structurally, it seems that their behaviour kinetically could be different in some respects. The preliminary screen of activity versus pH indicated that the human ADP-GK has optimal activity at an acidic pH, and is slowed down as the alkalinity increases, and at pH 8.5, activity can no longer be detected. This contrasts with the behaviour of the recombinant mouse enzyme, which has bi-modal optima between 5.75 and 6.5 and one near 8.75–9.0 (Ronimus and Morgan, 2004).

Chapter 6: Tissue screen for ADP-GK protein expression

6.1 Introduction

No investigation into the expression profile of ADP-GK in mammalian tissues had taken place prior to this project. Although immunoblotting does not divulge information on post-translational modifications or the protein's state of activation, it can give some indication of the protein's importance in the cell. This chapter describes the immunoblotting that was undertaken to examine the expression of ADP-GK in a range of mammalian tissues, and to explore the differential expression of the protein in human foetal tissues, adult tissues, and adult cancerous tissues.

6.2 Immunoblotting for ADP-GK expression in porcine tissue

Pigs have been used as a model organism for a number of years due to their physiological similarity to humans. Pig tissue was readily available from a local abattoir, and utilising this resource, an experiment was undertaken to screen for the presence of the ADP-GK protein in 18 different tissue types.

Tissues to be screened for the expression of the ADP-GK protein were collected from pigs freshly slaughtered at a local abattoir and transported on ice to the laboratory. Approximately 1 g samples of tissue were homogenised in 10 mL of cold sucrose buffer (20 mM HEPES, pH 7.4, 0.25 M sucrose, 1 mM EDTA, Complete mini-protease inhibitor) using an Ultra-Turrax® T50 homogeniser. The homogenates were centrifuged at 16 000 *g* for 30 minutes at 4°C and the supernatants collected. The resulting pellets were resuspended in 0.1 M Tris-HCl, pH 7.8, 0.5 % (v/v) Triton X-100. Protein concentration was quantified by Bradford protein assay (Section 2.2.17). SDS-PAGE was performed using 20 µg of cytoplasmic protein obtained from each of the eighteen tissue types, and proteins were then transferred onto a positively charged nylon membrane (Section 2.2.16).

ADP-GK was detected using the mouse monoclonal anti-ADP-GK antibody at a dilution of 1/500. HRP-conjugated rabbit anti-mouse antibody was used at a dilution of 1/5000.

To begin with, immunoblots were carried out on a few tissues at a time. Both the cytoplasmic fraction and the membrane fraction of each tissue extract were immunoblotted. It became clear that the strongest signal was obtained from the antibody reacting with the expressed ADP-GK protein in the cytoplasmic fraction of the extracts, with a much weaker signal gained from the membrane fraction. It was pointed out at a later date that, in order for the result to conclusively show that the ADP-GK protein is a cytoplasmic protein, the whole cell protein extracts would have had to have been ultra-centrifuged to eliminate all cytoplasmic proteins from the membrane protein fraction. As ultra-centrifugation was not performed, the membrane fraction results will not be examined.

The immunoblot (Figure 6.1) showed that expression of ADP-GK occurred in all 18 tissues screened, and it appeared that expression at the protein level was differential.

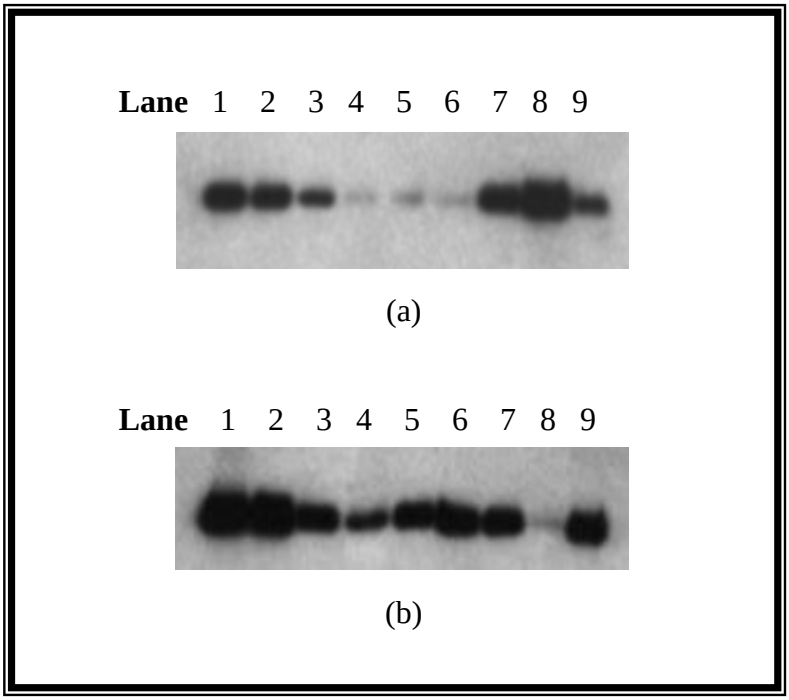


Figure 6.1 Immunoblot of cytoplasmic ADP-GK

SDS-PAGE (8% acrylamide gel) was performed using 20 µg of cytoplasmic protein extracted from 18 different tissue types obtained from freshly culled pigs. The protein was transferred to a positively charged membrane. A 1:500 dilution of the mouse primary

monoclonal ADP-GK antibody was used, followed by a 1:5000 dilution of horse-radish peroxidise-conjugated secondary anti-mouse antibody. The horseradish peroxidase-conjugated secondary antibody was detected by enhanced chemiluminescence. Light produced by the reaction was detected in a dark room by X-ray film.

Film (a): **Lane one:** Heart; **Lane two:** Lung; **Lane three:** Thymus; **Lane four:** Spleen; **Lane five:** Liver; **Lane six:** Pancreas; **Lane seven:** Bladder; **Lane eight:** Cervix; **Lane nine:** Prostate.

Film (b): **Lane one:** Ovary; **Lane two:** Uterus; **Lane three:** Large intestine; **Lane four:** Stomach; **Lane five:** Smooth muscle oesophagus; **Lane six:** Trachea; **Lane seven:** Oesophagus; **Lane eight:** Small intestine; **Lane nine:** Colon.

6.3 Immunoblotting for ADP-GK expression in human tissue

The screening of pig tissues was followed by the screening of two purchased human total protein arrays. The first array (Figure 6.2) compared adult and foetal tissues and the second array (Figure 6.3) compared normal adult tissues to cancerous adult tissues. Total protein had been spotted in duplicate on the nitrocellulose membrane, and each spot contained the same amount of protein (approximately 30 ng). The arrays included a positive (protein extracted from human normal placenta) and negative (water) control, also spotted in duplicate.

Processing of the arrays began at the point of incubation with primary antibody. ADP-GK was detected using the mouse monoclonal anti-ADP-GK antibody at a dilution of 1/500. HRP-conjugated rabbit anti-mouse antibody was used at a dilution of 1/5000.

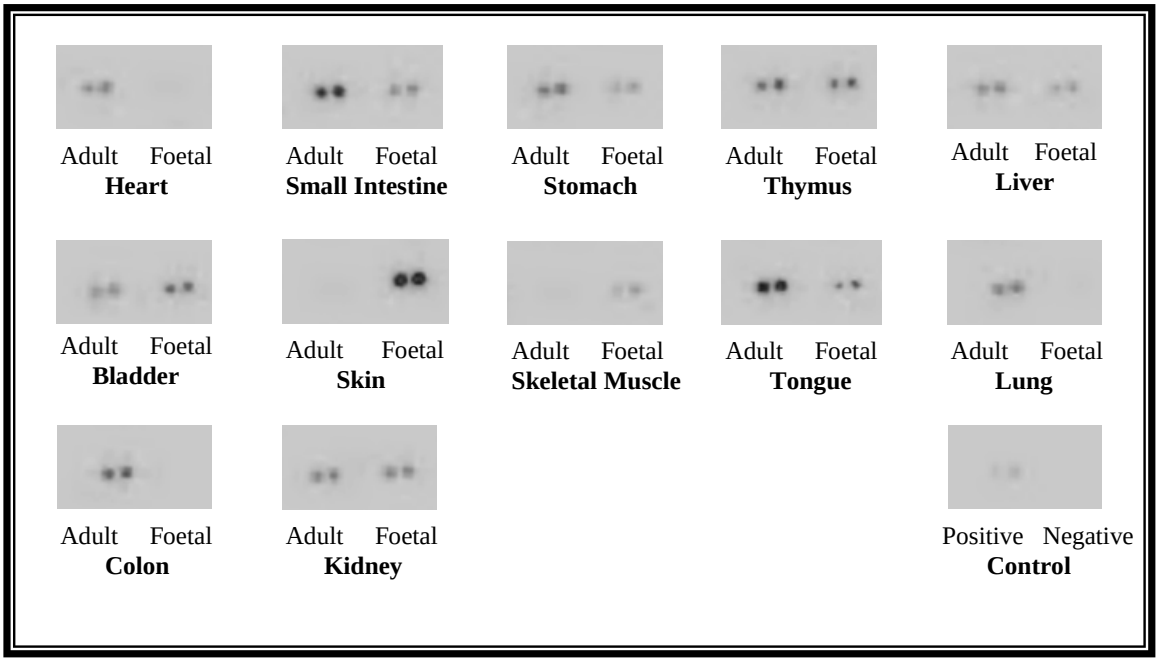


Figure 6.2 ADP-GK detected on an array of total protein from adult and foetal tissue

The total protein array was created from duplicate spots of total protein from adult and foetal tissue spotted onto a nitrocellulose membrane. The human ADP-GK protein was detected using the mouse monoclonal anti-ADP-GK antibody.

The immunoblotting results obtained from the adult and foetal tissue total protein array indicate that there is differential expression of human ADP-GK protein in adult and foetal tissues. ADP-GK expression was observed in all adult tissues, excluding skin and skeletal muscle. ADP-GK expression was observed in all foetal tissues, excluding heart and lung. The ADP-GK appeared to be expressed at higher levels in a greater number of adult tissues than foetal tissues. This was the case for adult heart, small intestine, stomach, liver, tongue, lung, thymus and colon. ADP-GK appeared to be expressed at a greater level in foetal bladder, skeletal muscle and very convincingly in skin. ADP-GK expression seemed to be equivalent in adult and foetal kidney.

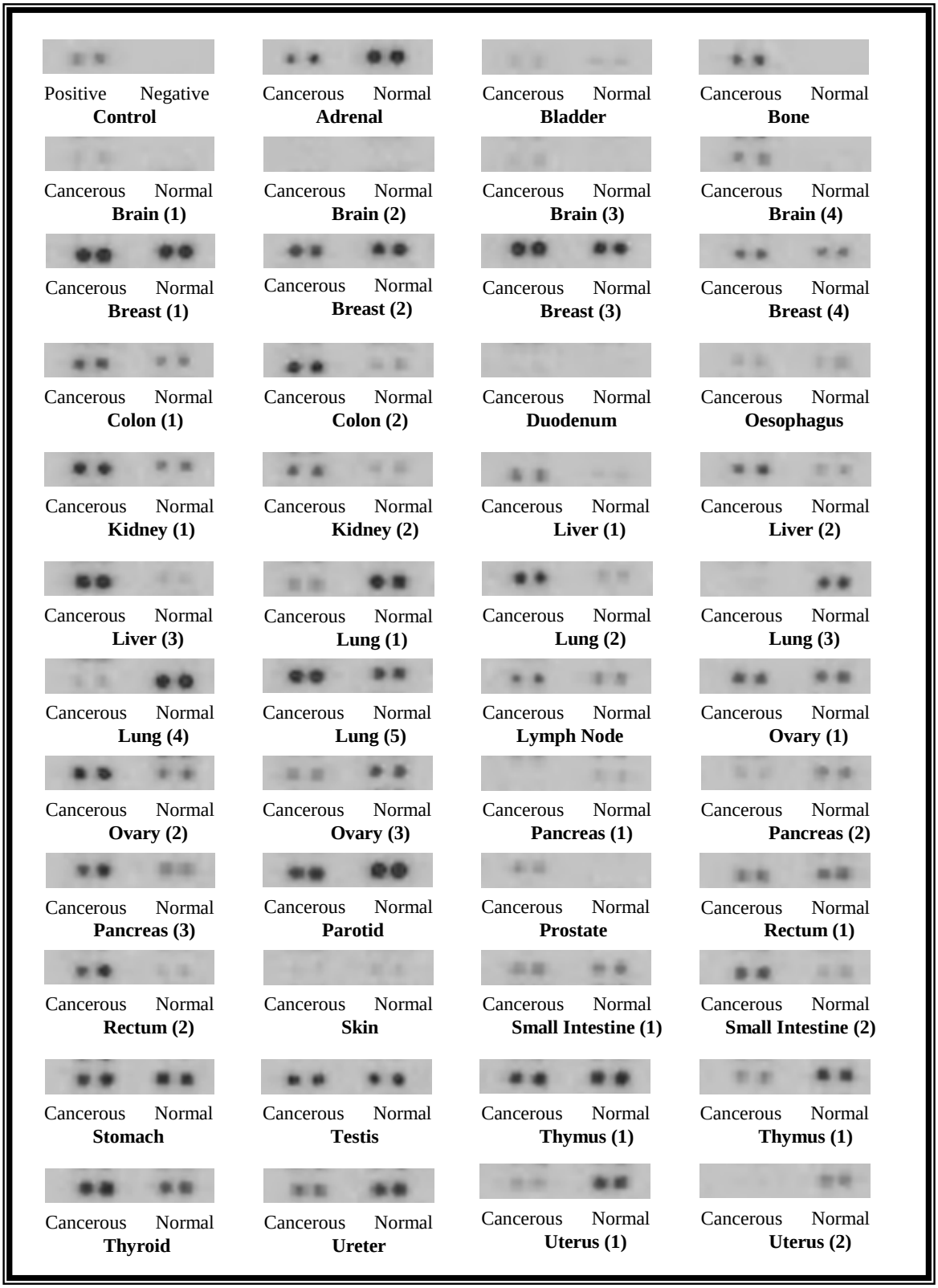


Figure 6.3 ADP-GK detected on an array of total protein from adult normal and cancerous tissue

The total protein array was created from duplicate spots of total protein from adult and foetal tissue spotted onto nitrocellulose membrane. The human ADP-GK protein was detected using the mouse monoclonal anti-ADP-GK antibody.

It was important to address the expression of ADP-GK in cancerous human tissue, as has been hypothesised that ADP-GK may confer an advantage to cells under energy stressed conditions such as hypoxia or ischemia. By using ADP as the energy investment in phase one of the glycolytic cycle instead of ATP, it is predicted that glycolysis could be sustained for longer than if only ATP is invested. It has been proposed that this could be a particularly important function in relation to the survival of cancer cells, as hypoxic regions have been observed within the microenvironments of most cancerous tumours due to insufficient blood supply. This could be of clinical significance as the extent of hypoxia can be correlated with prognosis in a number of tumour types (Talks *et al.*, 2000) and it can also be associated with chemo- and radio-therapeutic resistance (Guppy *et al.*, 2005).

The immunoblot of the normal tissue compared to cancerous tissue protein array was undertaken as a preliminary investigation of whether there is in fact an association between cancerous tissue and increased ADP-GK expression, as hypothesised.

The resulting immunoblot (Figure 6.3) did not produce a clear trend of increased ADP-GK in all cancers, and in some tissues there was no increase at all. However, the result certainly did not negate the hypothesis. It must be noted that the cancerous and normal tissues were not matched; the samples were collected from different individuals.

The array included multiple examples of normal and cancerous brain, breast, lung, pancreas, colon, rectum, ovary, uterus, kidney, liver, small intestine and thymus tissue. In the liver, kidney and colon examples, the expression of ADP-GK was greater in the cancerous tissue than in the normal tissue, in all cases, even though the tumours the cancerous tissue originated from differed between samples of liver (adenocarcinoma,

hepatocellular carcinoma and cholangiocellular carcinoma) and kidney (clear cell carcinoma and granular cell carcinoma) (see Appendix 5).
The situation was not so clear in the lung, breast, pancreas ovary, rectum, and small intestine examples.

The protein array incorporated five different lung samples. In samples (1), (3) and (4), the normal tissue spots appeared to have greater expression of the ADP-GK than the cancerous spots, which in this case were examples of adenocarcinoma and squamous cell carcinoma (Appendix 5). In the last two samples ((2) and (5)) the opposite appeared to be true, with the tumour tissue having greater expression of ADP-GK than the normal tissue.

In the breast examples, the tumour tissue of samples (3) and (4) appeared to have slightly greater expression of ADP-GK than the normal tissue. However, in sample (1), the expression of ADP-GK looked equal in normal and cancerous tissue, and in sample (2) the ADP-GK expression in the normal tissue appeared slightly greater than that of the cancerous tissue.

Of the three ovarian examples, all seemed to have clear expression of the ADP-GK in normal tissue. In the case of sample (1) in which the tumour sample was a clear cell carcinoma, the expression of ADP-GK in the tumour tissue appeared to be about the same as that of the normal tissue. Sample (2) was a tumour sample from a granular cell carcinoma, and this had very strong expression of ADP-GK in comparison to normal tissue. The tumour tissue in sample (3) came from a thecoma and this had slightly lower expression of ADP-GK than in the normal tissue.

Again, in the three samples from pancreatic tissue, the results varied. The expression of ADP-GK in normal tissue in sample (1) was almost undetectable, but in samples (2) and (3), there is clearly detectable expression. All tumour samples were referred to as adenocarcinomas and two were classed as moderately differentiated. In the tumour from sample (1), ADP-GK expression was undetectable. The tumour tissue from sample (2) had

low expression of ADP-GK, and it was less than that of the normal tissue. In sample (3) the tumour expression of ADP-GK was very high in comparison to the normal tissue.
Of the two examples of small intestine tissue, the Leiomyoma tumour of sample (1) had expression of ADP-GK that appeared to be equivalent to that of the normal tissue. In the Non-Hodgekin's lymphoma example, the expression of ADP-GK appeared to be greater than that of the normal tissue.

The rectal tumours were both referred to as moderately differentiated adenocarcinomas, however the tumour tissue in sample (2) had much greater expression of ADP-GK than the tumour tissue of sample (1), which was approximately equivalent to the ADP-GK expression of the normal tissue.

The thymus tissue showed strong expression of the ADP-GK in normal tissue, and the tumour tissue had expression of the ADP-GK equal to that of the normal tissue in sample (1), and slightly less than the normal tissue in sample (2).

The two uterine examples had greater expression of ADP-GK in the normal tissue than in the endometrium adenocarcinoma samples, one of which did not have any detectable expression of ADP-GK.

The brain samples presented an interesting trend: the four normal tissue samples (originating from males aged 22, 26 and 82) presented on this protein array did not appear to have any detectable expression of ADP-GK. However there was a faint signal gained from the expression of ADP-GK in the brain tumours, with the exception of sample (2). The brain tumour samples that had detectable ADP-GK expression originated from astrocytoma, glioblastoma, and neuroepithelioma tumours.

Of the tissues represented by a single sample: the parotid sample displayed very strong expression in both cancerous and normal tissue, but appeared slightly higher in normal. The adrenal sample expressed ADP-GK in both cancerous and normal tissue, but it was greater in normal. This was also the case for the ureter sample.

ADP-GK expression appeared equivalent in the normal and cancerous tissue samples for the bladder, oesophagus, skin, stomach and testis.

ADP-GK expression appeared slightly higher in the cancerous tissue samples for the thyroid, lymph node, prostate and bone. No expression of ADP-GK was detected at all from the duodenal sample.

6.4 Chapter Summary

The conclusion that can be drawn from the immunoblotting results at this point is that the ADP-GK enzyme is expressed widely in pig tissues, and in human tissues at both foetal and adult stages. In addition, there appears to be an increased expression of ADP-GK in cancers that involve the brain, liver, kidney, colon, bone, lymph node, prostate, and thyroid. In some cases (as described) this also seems to be true for cancers involving the lung, breast, ovary, pancreas, rectum, and small intestine.

It was observed that the expression of the ADP-GK in normal tissue collected from different donors but from the same organ had, in some cases, varying degrees of expression. This was not due to the immunoblotting technique or antibody used, as variation between samples from the same tissue was observed when immunoblotting using an antibody against the housekeeping protein GAPDH (see Appendix 5), as was performed by the company the arrays were purchased from.

The differing results obtained from normal tissue from the same organ can be attributed to the variation in protein expression from person to person, and contributing factors such as age, sex and the disease state of the donor. The examples of normal tissue provided on the total protein array presented in this chapter were harvested 4-6 hours post mortem and the examples of cancerous tissue were obtained from biopsy. This may also have influenced the apparent differential expression of protein between normal and cancerous tissues, as a number of changes in protein expression occur after death (Castensson *et al*., 2000). As

such, a protein array such as this will only have a limited ability to illustrate trends in protein expression. Nevertheless, this survey indicates that ADP-GK is widely expressed in humans, suggesting some metabolic role for this enzyme.

Chapter 7: Discussion and future research

7.1 Overview

Glycolysis was the first metabolic pathway to be elucidated and it has been studied rigorously for more than 70 years. For this reason, the discovery in 2004 of an enzyme known as ADP-dependent glucokinase which catalyses the phosphorylation of glucose to glucose-6-phosophate using ADP as the phosphoryl donor, was highly significant. Prior to the discovery of this enzyme, it was believed that the only enzymes responsible for catalysing the phosphorylation of glucose in the phase one energy investment stage of glycolysis were the ATP-dependent hexokinases.

At this point in time, there is no demonstrated metabolic role that can be attributed to the enzymatic activity of a mammalian ADP-GK, other than the presumed role in glucose phosphorylation. Only one study has been published on the enzyme, and this examined the kinetic characteristics of a recombinantly expressed mouse ADP-GK (Ronimus and Morgan, 2004). Using this data, it has been hypothesised that ADP-GK may have a metabolic role in the survival of cells that are experiencing energy stressed conditions, such as those associated with hypoxia. By catalysing the phosphorylation of glucose using ADP instead of ATP when the availability of ATP is low, ADP-GK could aid in sustaining glycolysis.

This project focussed on the human ADP-GK with the intention of expressing and purifying the recombinant protein in order to carry out kinetic characterisation. By comparing the kinetic data of the human ADP-GK to the mouse ADP-GK, more clues as to the metabolic role of the enzyme could have been obtained. However, the project ultimately did not encompass the kinetic characterisation of the recombinant human ADP-GK, as it proved to be more difficult to express and purify in its soluble state than had been expected.

7.2 Summary of results

Cloning and expression of recombinant human ADP-dependent glucokinase

Cloning of the complete coding sequence of the human ADP-GK into the *E. coli* host strains BL21(DE3) and Rosetta, using a pET151/D-TOPO vector was undertaken successfully. The first trials at 25°C and 30°C using IPTG to initiate recombinant protein expression demonstrated that the protein was readily expressed within a few hours of being induced. When the protein harvested from the trials was fractionated however, it became apparent that the greatest proportion of recombinant protein was being expressed insolubly. In an attempt to solve this problem, expression trials were undertaken which manipulated and varied the temperature of expression, the concentration of IPTG used to induce expression, and the strain of *E. coli* used to express the recombinant protein. A trial to induce the expression of molecular chaperones by the addition of benzyl alcohol was undertaken (de Marco *et al.*, 2005), in the expectation that increased chaperone levels may assist in the disaggregation and refolding of insoluble recombinant protein. This did not appear to significantly change the insoluble state of the recombinant human ADP-GK, and neither did the cessation of new protein synthesis by the addition of chloramphenicol, to give the cell's protein folding machinery time to process protein aggregates (de Marco *et al.*, 2007). While it seemed that some improvements were achieved by lowering the expression temperature, and thereby slowing the rate at which the recombinant protein was expressed, no wholesale difference in the proportion of soluble recombinant protein expressed was attained.

Finally, the cell lysis and fractionation conditions were investigated to examine their influence on the apparent distribution of soluble and insoluble recombinant protein. A trial of lysis buffers was carried out which examined the effect of low and high NaCl concentrations, and the presence and/or absence of the non-denaturing detergents, Triton X-100 and CHAPS. The conclusion was reached that the most significant influencing factor on the distribution of recombinant protein between the soluble and insoluble states was the presence of detergent in the lysis buffer, in particular CHAPS (Figures 3.15 and 3.16) .

A number of other approaches to the problem of insoluble expression of recombinant protein could have been taken, but due to time constraints, it was not possible. To begin with, a vector with a C-terminal polyhistidine tag could have been trialled, as the recombinant mouse enzyme was successfully expressed with a C-terminal tag. It is possible that the positioning of the tag could influence protein folding and thus solubility. The co-expression of molecular chaperones may have proved to be successful in the prevention of protein aggregation and also the use of a fusion protein as a solubility enhancing partner, for example the Maltose binding protein (MBP) or N-utilising substance A (Nus-A) may have solved the solubility issue. However, just because a protein is expressed solubly, it does not mean that the protein is folded correctly.

Purification of recombinant human ADP-dependent glucokinase

Although a considerable amount of time was spent using affinity chromatography (Arnau *et al.,* 2006) to purify what little soluble protein was expressed, this approach eventually had to be abandoned. Instead, the insoluble protein was harvested as inclusion bodies and solubilised in 8.0 M urea. A refolding trial was designed based on one used by Vincentelli *et al*., (2004) which utilised a 96 well plate to screen multiple refolding conditions simultaneously. As a result of the trial, the recombinant ADP-GK was refolded in Tris buffer at pH 8.0, which led to the detection of enzyme activity for the first time under standard ADP-GK assay conditions. Enzyme activity confirmed that at least a portion of the recombinant protein in solution had refolded. The refolded protein was then subjected to affinity purification trials using the ÄKTA FPLC system and a HisTrap^{TM} HP 5 mL affinity column. The purification protocol was finally optimised, with elution at 200 mM imidazole yielding the highest amount of recombinant protein, although a number of proteins of smaller molecular mass were present in the elution. These smaller molecular mass proteins appeared to elute with a similar pattern to that of the recombinant ADP-GK. An immunoblot was carried out, and the smaller molecular mass proteins were recognised by the ADP-GK monoclonal antibody. It was therefore concluded that these proteins could actually be truncated expression products caused by the rapid rate of recombinant expression of the human ADP-GK, rather than proteolysis which would have tended to cause a smear instead of distinct size products.

If the removal of the smaller molecular mass species was considered necessary at a later date, size-exclusion chromatography could be used to achieve this. Alternatively, a second step in the purification process such as ion-exchange chromatography could be used. This method was successful in the purification of the recombinant mouse ADP-GK enzyme (Ronimus and Morgan, 2004), so it is likely that it would also be successful with the human version, due to the similarity between the two proteins.

Recombinant protein characterisation

The amount of time spent getting to the point where the recombinant ADP-GK was homogeneous precluded the completion of the enzyme's kinetic characterisation. Instead, it was considered necessary to check for protein secondary structure to confirm that the protein had in fact refolded in the Tris buffer (pH 8.0). The preliminary circular dichroism study suggested that the refolded recombinant human ADP-GK did have significant secondary structure, and the percentages of alpha helix and turns estimated using the K2D2 (Perez-Iratxeta *et al.*, 2008) web server (http://www.ogic.ca/projects/k2d2), demonstrated that the mouse and human recombinant ADP-GK proteins probably have similar overall secondary structure.

Enough protein was available at this stage of the project to undertake a preliminary screen of activity versus pH. This indicated that the human ADP-GK has optimal activity at an acidic pH, and is slowed down as alkalinity increases. At pH 8.5, activity could no longer be detected. This contrasts with the behaviour of the recombinant mouse enzyme, which has been shown to have bi-modal optima between 5.75 and 6.5 and one near 8.75–9.0 (Ronimus and Morgan, 2004).

In order for a complete kinetic characterisation of this protein to be carried out, the first point that will need to be addressed is the stability of the protein. Although the protein appeared to have refolded in the Tris buffer with no additives, it quickly became clear that it was not stable in solution for long, especially at concentrations above 0.2 mg/mL. A screen of buffer conditions and possible additives to assist with protein stability needs to be undertaken before any further work can be done. It is possible that cleaving the

polyhistidine tag may help to stabilise the protein in a folded state. This is a rapid process and should be one of the first approaches trialled.

Tissue screen for ADP-GK protein expression

No investigation into the expression of ADP-GK in mammalian tissues had taken place prior to this project. Although immunoblotting does not divulge information on post-translational modifications or the protein's state of activation, it can give some indication of the protein's importance in the cell.

The first tissue screen was undertaken using tissue collected from freshly culled pigs. The screen encompassed 18 tissues in all, and the presence of ADP-GK was detected in each tissue, and expression at the protein level appeared to be differential. This was followed by a tissue screen using a commercial protein array of whole cell protein harvested from tissue and blotted onto a nitrocellulose membrane. The first array was composed of adult and foetal whole cell protein from specific tissues. It was apparent that the ADP-GK was widely expressed in humans also, with differential expression between the tissues blotted. The ADP-GK appeared to be expressed at higher levels in a greater number of adult tissues than foetal tissues.

It was important to address the expression of ADP-GK in cancerous human tissue also, as it has been hypothesised that ADP-GK may confer an advantage to cells under energy stressed conditions such as hypoxia or ischemia, and tumours are often hypoxic due to poor blood supply. The increased expression of ADP-GK in cancerous tissue when compared to non-cancerous tissue from the same organ was observed in a number of cases, but the trend was not clear. In the protein array screened in this study, there appeared to be an increased expression of ADP-GK in cancers that involved the brain, liver, kidney, colon, bone, lymph node, prostate, and thyroid. In some cases, this also seemed to be true for cancers involving the, lung, breast, ovary, pancreas, rectum, and small intestine.

The main conclusion drawn from this study was that, overall, the ADP-GK enzyme is widely expressed in mammalian tissues. The results in terms of the increased expression in

cancers have to be interpreted carefully, as there is always variation in protein expression from person to person, and contributing factors such as age, sex and the disease state of the donor need to be kept in mind. The examples of normal tissue provided on the total protein array presented in this chapter were harvested 4-6 hours post mortem and the examples of cancerous tissue were obtained from biopsy. This inconsistency may also have influenced the apparent differential expression of protein between normal and cancerous tissues, because a number of changes in protein expression occur after death (Castensson *et al*., 2000).

7.3 Further research

The scope of research that could be undertaken in relation to the ADP-dependent glucokinase is extensive, as so little is currently known about the protein.

The first obvious study that needs to be carried out is the enzymatic characterisation of recombinant human ADP-GK. The specific activity and pH activity curve will need to be confirmed. The specificity the enzyme has for substrates will need to be determined, as well as its kinetic parameters, such as the enzyme's K_m for glucose and ADP. When this has been completed, the data can be compared and contrasted with the data obtained from the study using recombinant mouse ADP-GK (Ronimus and Morgan, 2004).

Solving the structure of the ADP-GK protein will be an important step. The structure should reveal the mechanism that makes the enzyme specific for ADP and not ATP, and to help determine its enzymatic mechanism and substrate specificity.

Due to the expression of ADP-GK in such wide range of mammalian tissues, it will be important to identify any post-translational modifications the ADP-GK protein undergoes. The state of activation of the ADP-GK in tissue is unknown at this point. The identification of post-translational modifications and their importance will need to be determined, and it is likely that these data will give significant insight into the role of this enzyme in the cell, and its relative functional importance. It is very possible that the activity of ADP-GK will

be regulated by post-translational phosphorylation. One way of examining post-translational phosphorylation would be to immunoprecipitate ADP-GK from protein extracted from mammalian tissue samples or cell culture samples, and detect the enzyme's phosphorylation status using specific anti-phosphoserine, anti-phosphothreonine and anti-phosphotyrosine antibodies. With the availability of purified recombinant human ADP-GK, a study could be undertaken into whether recombinant human AMPK, which can be purchased, phosphorylates the ADP-GK *in vitro*, and if so, how this affects the already determined kinetic parameters of the recombinant human ADP-GK. If it is proven that phosphorylation is occurring, a sample of phosphorylated ADP-GK could be analysed by tryptic digestion and MALDI-TOF spectrometry to identify the modification site(s).

The examination of the ADP-GK promoter the sequence for predicted transcription factor binding sites, and analysis by deletion studies to determine the minimal promoter, could give some clues as to the regulation of the expression of this enzyme.

The determination of the kinetic characteristics of this protein, its structure and enzymatic mechanism, and its possible post-translational activation will give much greater insight into the functional role played by this enzyme in the mammalian cell.

Overall, this project has confirmed one method for the production of recombinant human ADP-dependent glucokinase and this opens the door for the next steps in the study of this enzyme.

References:

Aleshin, A. E., Zeng, C., Bartunik, H. D., Fromm, H. J., and Honzatko, R. B. (1998). Regulation of hexokinase I: crystal structure of recombinant human brain hexokinase complexed with glucose and phosphate. Journal of Molecular Biology *282*, 345-357.

Altenberga, B., and Greulichb, K. O. (2004). Genes of glycolysis are ubiquitously overexpressed in 24 cancer classes. Genomics *84*, 1014-1020.

Anfinsen, C, B. (1972). The formation and stabilization of protein structure. Biochem. J. *128*, 737-749.

Arnau, J., Lauritzen, C., Petersen, G. E., and Pederson, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expression and Purification. *48*, 1-13.

Baltrusch, S., Langer, S., Massa, L., Tiedge, M., and Lenzen, S. (2006). Improved Metabolic Stimulus for Glucose-Induced Insulin Secretion through GK and PFK-2/FBPase-2 Coexpression in Insulin-Producing RINm5F Cells. Endocrinology *147*, 5768-5776.

Bedova.F. J., Matschinsky, F. M., Shimizu, T., O'Neil, J. J., and Appel, M. C. (1986). Differential regulation of glucokinase activity in pancreatic islets and liver of the rat. J Biol Chem *261*, 10760-10764.

Biaglow, J., Cerniglia, G., Tuttle, S., Bakanauskas, V., Stevens, C., and McKenna, G. (1997). Effect of Oncogene Transformation of Rat Embryo Cells on Cellular Oxygen Consumption and Glycolysis. Biochemical and Biophysical Research Communications *235*, 739-742.

Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. *7*, 1513-1523.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. *72*, 248-54.

Brahimi-Horn, M. C., and Pouyssegur, J. (2007). Harnessing the hypoxia-inducible factor in cancer and ischemic disease. Biochemical Pharmacology *73*, 450-457.

Buttgereit, F., and Brand, M. (1995). A hierarchy of ATP-consuming processes in mammalian cells. Biochem J *15*, 163-167.

Carew, J., and Huang, P. (2002). Mitochondrial defects in cancer. Molecular Cancer *1*, 1-12.

Carling, D., Zammit, V. A., and Hardie, D. G. (1987). A common bicyclic protein kinase cascade inactivates the regulatory enzymes of fatty acid and cholesterol biosynthesis. FEBS Letters *223*, 217-222.

Castensson, A., Emilsson, L., and Preece, P. (2000). High-resolution quantification of specific mRNA levels in human brain autopsies and biopsies. Genome Res *10*, 1219-1229.

Cline, J., Braman, J. C., and Hogrefe H. H. (1996). PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Res. *18*, 3546-3551

Cooper, P., and Meddings, J. B. (1991). Erythrocyte membrane fluidity in malignant hyperthermia. Biochimica et Biophysica Acta (BBA) - Biomembranes *1069*, 151-156.

Corton, J. M., Gillespie, J. G., and Hardie, D. G. (1994). Role of the AMP-activated protein kinase in the cellular stress response. Current Biology *4*, 315-324.

Danial, N. N., Gamma, C. F., Scorrano, L., Chen-Yu, Z., Krauss, S., Ranger, A. M., Datta, S. R., Greenberg, M. E. Licklider, L. J., Lowell, B. B., *et al.* (2003). BAD and glucokinase reside in a mitochondrial complex that integrates glycolysis and apoptosis. Nature *424*, 952.

Dasari, P., Nicholson, I. C., and Zola, H. (2008). Toll-like receptors. J Biol Regul Homeost Agents. *22*, 17-26.

D'Aurelio, M., Merlo Pich, M., Catani, L., Sgarbi, G. L., Bovina, C., Formiggini, G., Castelli, G.P., Baum, H., Tura S., and Lenaz, G. (2001). Decreased Pasteur effect in platelets of aged individuals. Mechanisms of Ageing and Development *122*, 823-833.

de la Iglesia, N., Mukhtar, M., Seoane, J., Guinovart, J. J., and Agius, L. (2000). The role of the regulatory protein of glucokinase in the glucose sensory mechanism of the hepatocyte. J Biol Chem *275*, 10597-10603.

de Marco, A., Vigh, L., Diamant, S., and Goloubinoff, P. (2005). Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. Cell Stress and Chaperones *10*, 329-339.

de Marco, A., (2007). Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli.* Nature Protocols *2*, 2632-2639.

Garber, K. (2006). Energy Deregulation: Licensing tumours to grow. Science *312*, 1158-1159.

Gloyn, A. L., Odili, S., Zelent, D., Buettger, C., Castleden, H. A. J., Steele, A. M., Stride, A., Shiota, C., Magnuson, M. A., Lorini, R., *et al.* (2005). Insights into the Structure and Regulation of Glucokinase from a Novel Mutation (V62M), Which Causes Maturity-onset Diabetes of the Young. J Biol Chem *280*, 14105-14113.

Gonzalez, C., Ureta, T., Sanchez, R., and Niemeyer, H. (1964). Multiple molecular forms of ATP: Hexose 6-phosphotransferase from rat liver. Biochemical and Biophysical Research Communications *16,* 347-352.

Greenfield, N. J. (2006). Using circular dichroism spectra to estimate protein secondary structure. Nature Protocols *1,* 2876-2890.

Grossbard, L., Weksler, M., and Schimke, R. T. (1966). Electrophoretic properties and tissue distribution of multiple forms of hexokinase in various mammalian species. Biochemical and Biophysical Research Communication *24,* 32-28.

Guppy, M., Brunner, S., and Buchanan, M. (2005). Metabolic depression: a response of cancer cells to hypoxia? Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology *140,* 233-239.

Hardie, G. D., and Hawley, S. A. (2001). AMP-activated protein kinase: the energy charge hypothesis revisited. BioEssays *23,* 1112-1119.

Heikkinen, S., Suppola, S., Malkki, M., and Deeb, S. (2000). Mouse hexokinase II: structure, cDNA promoter analysis, and expression pattern. Mamm Genome *11,* 91-96.

Jetton, T. L., Liang, Y., Pettepher, C. C., Zimmerman, E. C., Cox. F. G., Horvath, K., Matchinsky, F. M., and Magnuson, M. A. (1994). Analysis of upstream glucokinase promoter activity in transgenic mice and identification of glucokinase in rare neuroendocrine cells in the brain and gut. J Biol Chem *269,* 3641-3654.

Kaselonis, G. L., McCabe, E. R. B., and Gray, S. M. (1999). Expression of hexokinase 1 and hexokinase 2 in mammary tissue of nonlactating and lactating rats: Evaluation by RT-PCR. Molecular Genetics and Metabolism *68,* 371-374.

Katzen, H., and Schimke, R. (1965). Multiple forms of hexokinase in the rat: Tissue distribution, age dependency and properties. Proc. Natn. Acad. Sci. U.S. *54,* 1218-1225.

Kelly, S. M., Jess, T. J., and Price, N. C., (2005). How to study proteins by circular dichroism. Biochimica et Biophysica Acta. *1751,* 119–139.

Kengen, S. W., de Bok, F. A., van Loo, N. D., Dijkema, C., Stams, A. J., and de Vos, W. M. (1994). Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by Pyrococcus furiosus. J Biol Chem *269,* 17537-17541.

Ko, Y. H., Pedersen, P. L., and Geschwind, J. F. (2001). Glucose catabolism in the rabbit VX2 tumor model for liver cancer: characterization and targeting hexokinase. Cancer Letters *173,* 83-91.

Labes, A., and Schonheit, P. (2003). ADP-dependent Glucokinase from the hyperthermophilic sulfate-reducing archaeon *Archaeoglobus fulgidus* strain 7324. Arch Microbiol *180*, 69-75.

Laderoute, K., Amin, K., Calaoagan, J., Knapp, M., Le, T., Orduna, J., Foretz, M., and Viollet, B. (2006). 5'-AMP-activated protein kinase (AMPK) is induced by low-oxygen and glucose deprivation conditions found in solid tumour microenvironments. Mol Cell Biol *26*, 5336-5347.

Lu, H., Forbes, R. A., and Verma, A. (2002). Hypoxia-inducible Factor 1 Activation by Aerobic Glycolysis Implicates the Warburg Effect in Carcinogenesis. J Biol Chem *277*, 23111-23115.

Magnuson, M. A., and Shelton, K. D. (1989). An alternate promoter in the glucokinase gene is active in the pancreatic beta cell. J. Biol. Chem. *264*, 15936-15942.

Matchinsky, F. M., Magnuson, M. A., Zelent, D., Jetton, T. L., Doliba, N., Han, Y., Taub, R., and Grimsby, J. (2006). The network of glucokinase-expressing cells in glucose homeostasis and the potential of glucokinase activators for diabetes therapy. Diabetes *55*, 1-12.

Nakashima, R. A., Paggi, M. G., Scott, L. J., and Pedersen, P. L. (1988). Purification and Characterization of a Bindable Form of Mitochondrial Bound Hexokinase from the Highly Glycolytic AS-30D Rat Hepatoma Cell Line. Cancer Res *48*, 913-919.

Pastorino J. G., Shulga, N., and Hoek, J. B. (2002). Mitochondrial binding of hexokinase II inhibits Bax-induced Cytochrome c release and apoptosis. J. Biol. Chem. *277*, 7610-7618.

Perez-Iratxeta, C., and Andrade-Navarro, M. A., (2008). K2D2: estimation of protein secondary structure from circular dichroism sprectra. BMC Structural Biology *8*.

Preller, A., and Wilson, J. E. (1992). Localization of the type III isozyme of hexokinase at the nuclear periphery. Archives of Biochemistry and Biophysics *294*, 482-492.

Rapisarda, A., Uranchimeg, B., Scudiero, D., Selby, M., Sausville, E., Shoemaker, R., and Melillo, G. (2002). Identification of small molecule inhibitors of hypoxia-inducible factor 1 transcriptional activation pathway. Cancer Res *62*, 4316-4324.

Roffe, C., Sills, S., Halim, M., Wilde, K., Allen, M. B., Jones, P. W., and Crome, P. (2003). Unexpected Nocturnal Hypoxia in Patients With Acute Stroke. Stroke *34*, 2641-2645.

Rolfe, D. F., and Brown, G. C. (1997). Cellular energy utilization and molecular origin of standard metabolic rate in mammals. Physiol Rev *77*, 731-758.

Ronimus, R. S., and Morgan, H. W. (2003). Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. Archaea *1*, 199- 221.

Ronimus, R. S., and Morgan, H. W. (2004). Cloning and biochemical characterization of a novel mouse ADP-dependent glucokinase. Biochemical and Biophysical Research Communications *315*, 652-658.

Sebastian, S., Horton, J. D., and Wilson, J. E. (2000). Anabolic function of the type II isozyme of hexokinase in hepatic lipid synthesis. Biochemical and Biophysical Research Communications *270*, 886-891.

Sebastian, S., Edassery, S., and Wilson, J. E. (2001). The human gene for the type III isozyme of hexokinase: Structure, basal promoter, and evolution. Archives of Biochemistry and Biophysics *395*, 113-120.

Shuman, S. (1991) Recombination Mediated by Vaccinia Virus DNA Topoisomerase I in *Escherichia coli* is Sequence Specific. Proc. Natl. Acad. Sci. USA *88*, 10104-10108

Smith, A., and Hall, C. K. (2001). Protein refolding versus aggregation: Computer simulations on an intermediate-resolution protein model. J. Mol. Biol. *312*, 187-202.

Talks, K., Turely, H., Gatter, K., Maxwell, P., Pugh, C., Ratcliffe, P., and Harris, A. (2000). The expression and distribution of the hypoxia-inducible factors HIF-1α and HIF-2α in normal human tissues, cancers and tumour associated macrophages. American Journal of Pathology *157*, 411-421.

Vaupel, P., Mayer, A., and Hockel, M. (2004). Tumour hypoxia and malignant progression. Methods in Enzymology *381*, 335-354.

Vincentelli, R., Canaan, S., Campanacci, V., Valencia, CD., Maurin, D., Frassinetti, F., Scappucini-Calvo, L., Bourne, Y., Cambillau, C., Bignon, C. (2004). High-throughput automated refolding screening of inclusion bodies. Protein Sci. *13*, 2782-2792.

Voet, D., and Voet, J. G. (2004). Biochemistry, 3rd edn (New York: J. Wiley & Sons).

Vogt, C., Ardehali, H., Iozzo, P., Yki-Jarvinen, H., Koval, J., Maezono, K., Pendergrass, M., Printz, R., Granner, D., Defronzo, R., and Mandarino, L. (2000). Regulation of hexokinase II expression in human skeletal muscle in vivo. Metabolism *49*, 814-818.

Warburg, O. (1956). On the origin of cancer cells. Science *123*.

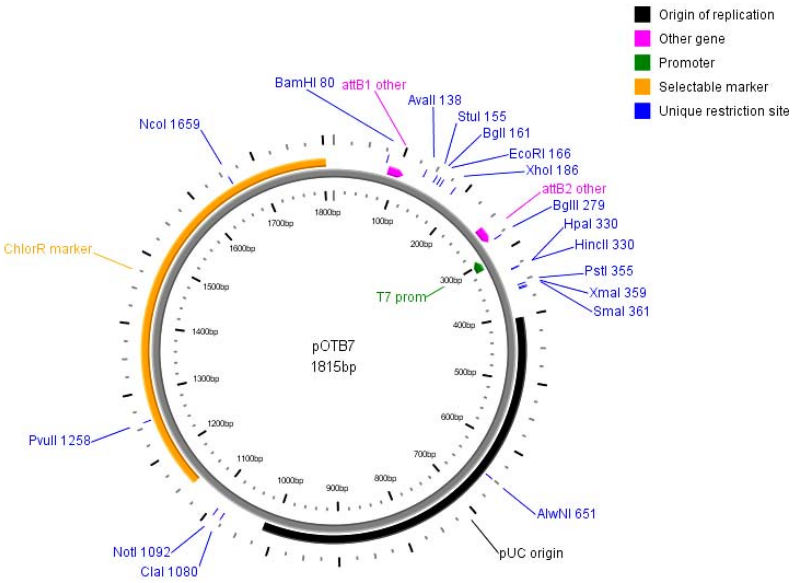
Wilson, J. E. (2003). Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. J Exp Biol *206*, 2049-2057.

Xu, R., Pelicano, H., Zhou, Y., Carew, J., Feng, L., Bhalla, K., Keatin, M., and Huang, P. (2005). Inhibition of glycolysis in cancer cells: a novel strategy to overcome drug resistance associated with mitochondrial respiratory defects. Cancer Res, *65*, 613-621.

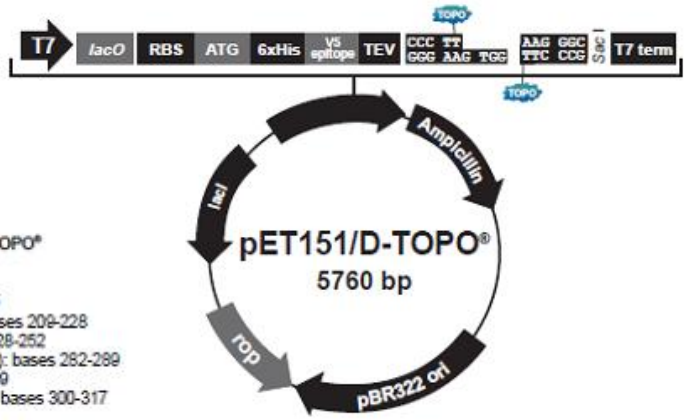
Zelent, D., Golson, M., Koeberlein, B., Quintens, R., Lommel, L., Buettger, C., Weik-Collins, H., Taub, R., Grimbsby, J., Schuit, F., Kaestner, K., and Matschinsky, F. M. (2006). A glucose sensor role for glucokinase in anterior pituitary cells. Diabetes *55,* 1923-1929.

Appendix 1: Vector maps

1.1 pOTB7 Vector



1.2 pET151/D-TOPO vector



Comments for pET151/D-TOPO®
5760 nucleotides

T7 promoter: bases 209-225
T7 promoter priming site: bases 209-228
lac operator (lacO): bases 228-252
Ribosome binding site (RBS): bases 282-289
Initiation ATG: bases 297-299
Polyhistidine (6xHis) region: bases 300-317
V5 epitope: bases 318-359
TEV recognition site: bases 360-380
TOPO® cloning site (directional): bases 387-400
T7 reverse priming site: bases 455-474
T7 transcription termination region: bases 416-544
bla promoter: bases 649-947
Ampicillin (*bla*) resistance gene: bases 948-1808
pBR322 origin: bases 1953-2626
ROP ORF: bases 2997-3188 (complementary strand)
lacI ORF: bases 4500-5612 (complementary strand)

TOPO® Cloning Site of pET151/D-TOPO®

```
121  ATAGGCGCCA GCAACCGCAC CTGTGGCGCC GGTGATGCCG GCCACGATGC GTCCGGCGTA GAGGATCGAG ATCTCGATCC
                T7 promoter/priming site
             T7 promoter                  lac operator
201  CGCGAAATTA ATACGACTCA CTATAGGGGA ATTGTGAGCG GATAACAATT CCCCTCTAGA AATAATTTTG TTTAACTTTA

          RBS                 Polyhistidine region                    V5 epitope
281  AGAAGGAGAT ATACAT ATG CAT CAT CAC CAT CAC CAT GGT AAG CCT ATC CCT AAC CCT CTC CTC GGT CTC
                       Met His His His His His His Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu
                    TEV recognition site
351  GAT TCT ACG GAA AAC CTG TAT TTT CAG GGA ATT GAT CCC TTC ACC ... ... AAGGG CGAGCTCAGA
                                                 GGG AAG TGG ... ...
     Asp Ser Thr Glu Asn Leu Tyr Phe Gln Gly Ile Asp Pro Phe Th
                                    TEV cleavage site
                                                 G TGG
                                                  T7 reverse priming site
411  TCCGGCTGCT AACAAAGCCC GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA ACTAGCATAA
```

Appendix 2: Human ADP-GK coding sequence

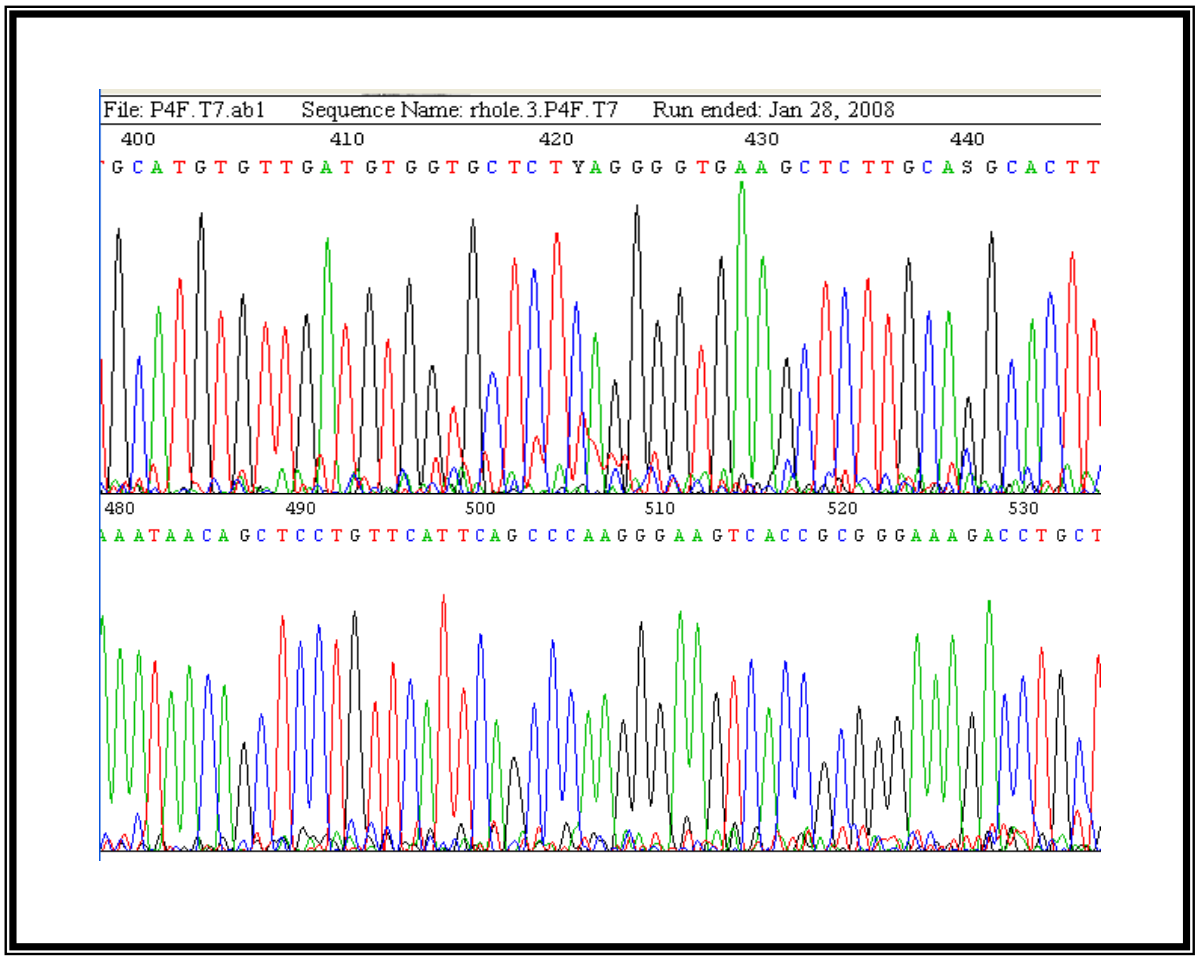
```
FEATURES             Location/Qualifiers
source               1..1635
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:12975 IMAGE:3347312"
                     /tissue_type="Kidney, renal cell adenocarcinoma"
                     /clone_lib="NIH_MGC_14"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1..1635
                     /gene="ADPGK"
                     /note="synonyms: DKFZP434B195, ADP-GK,
                     2610017G09Rik"
                     /db_xref="GeneID:83440"
                     /db_xref="HGNC:25250"
      CDS            37..1530
                     /gene="ADPGK"
                     /codon_start=1
                     /product="ADP-dependent glucokinase"
                     /protein_id="AAH06112.1"
                     /db_xref="GI:13543940"
                     /db_xref="GeneID:83440"
                     /db_xref="HGNC:25250"

/translation="MALWRGSAYAGFLALAVGCVFLLEPELPGSALRSLWSSLCLGPAPAPPGPVSPEGRLAA
AWDALIVRPVRRWRRVAVGVNACVDVVLSGVKLLQALGLSPGNGKDHSILHSRNDLEEAFIHFMGKGAAAERF
FSDKETFHDIAQVASEFPGAQHYVGGNAALIGQKFAANSDLKVLLCGPVGPKLHELLDDNVFVPPESLQEVDE
FHLILEYQAGEEWGQLKAPHANRFIFSHDLSNGAMNMLEVFVSSLEEFQPDLVVLSGLHMMEGQSKELQRKRL
LEVVTSISDIPTGIPVHLELASMTNRELMSSIVHQQVFPAVTSLGLNEQELLFLTQSASGPHSSLSSWNGVPD
VGMVSDILFWILKEHGRSKSRASDLTRIHFHTLVYHILATVDGHWANQLAAVAAGARVAGTQACATETIDTSR
VSLRAPQEFMTSHSEAGSRIVLNPNKPVVEWHREGISFHFTPVLVCKDPIRTVGLGDAISAEGLFYSEVHPHY
"
```

ORIGIN

1 ggagttggtt cgcgcaggtg cggcgcctgg gtcccc**atgg cgctgtggcg cgg**ctccgcg
61 tacgcgggct tcctggcgct ggccgtgggc tgcgtcttcc tgctggagcc agagctgcca
121 ggctcggcgc tgcgctctct ctggagctcg ctgtgtctgg ggcccgcgcc tgcgccccg
181 ggacccgtct cccccgaggg ccggttggcg gcagcctggg acgcgcttat cgtgcggcca
241 gtccggcgct ggcgccgcgt ggcagtggga gtcaatgcat gtgttgatgt ggtgctctca
301 ggggtgaagc tcttgcaggc acttggcctt agtcctggga atgggaaaga tcacagcatt
361 ctgcattcaa ggaatgatct ggaagaagcc ttcattcact tcatggggaa gggagcagct
421 gctgagcgct tcttcagtga taaggaaact tttcacgaca ttgcccaggt tgcgtcagag
481 ttcccaggag cccagcacta tgtaggagga aatgcagctt taattggaca gaaatttgca
541 gccaactcag atttaaaggt tcttctttgc ggtccagttg gtccaaagct acatgagctt
601 cttgatgaca atgtctttgt tccaccagag tcattgcagg aagtggatga gttccacctc
661 attttagagt atcaagcagg ggaggagtgg ggccagttaa aagctcccca tgccaaccga
721 ttcatcttct ctcacgacct ctccaacggg gccatgaata tgctggaggt gtttgtgtct
781 agcctggagg agtttcagcc agacctggtg gtcctctctg gattgcacat gatggaggga
841 caaagcaagg agctccagag gaagagactc ttggaggttg taacctccat ttctgacatc
901 cccactggta ttccagttca cctagagctg gccagtatga ctaacaggga gctcatgagc
961 agcattgtcc atcagcaggt ctttcccgcg gtgacttccc ttgggctgaa tgaacaggag
1021 ctgttatttc tcacccagtc agcctctgga cctcactctt ctctctcttc ctggaacggt
1081 gttcctgatg tgggcatggt cagtgacatc ctcttctgga tcttgaaaga acatgggagg
1141 agtaaaagca gagcctcgga tctcaccagg atccatttcc acacgctggt ctaccacatc
1201 ctggcaactg tggatggaca ctgggccaac cagctggcag ccgtggctgc aggagctcgt
1261 gtggctggga cacaggcctg cgccacagaa accatagaca ccagccgagt gtctctgagg
1321 gcaccccaag agttcatgac ttcccattcg gaggcaggct ccaggattgt attaaaccca
1381 aacaagccag tagtagaatg gcacagagag ggaatatcct tccacttcac accagtattg
1441 gtgtgtaaag accccattcg aactgtaggc cttggagatg ccatttcagc cgaaggactc
1501 ttctattc**gg aagtacaccc tcactattag** gaagattctt aggggtaatt tttctgagga
1561 aggagaacta gccaacttaa gaattacagg aagaaagtgg tttggaagac agccaaaaaa
1621 aaaaaaaaaa aaaaa

Appendix 3: Chromatogram from sequencing of cloned hAPD-GK coding sequence



A portion of the chromatogram obtained during DNA sequencing of the cloned human ADP-GK coding sequence is shown here to confirm the quality of the sequencing data.

Appendix 4: Raw data collected from measuring circular dichroism

Wavelength (nm)	**Mouse ADP-GK** (millidegrees)	**Human ADP-GK** (millidegrees)
260	-0.00639487	-0.00571
259.5	0.011241	0.012737
259	0.0389128	-0.00596
258.5	0.0216235	0.008581
258	0.0184966	0.00139
257.5	0.0061288	-0.03083
257	-0.0852436	-0.07999
256.5	-0.112444	-0.11973
256	-0.121726	-0.12602
255.5	-0.155971	-0.07562
255	-0.178023	-0.0666
254.5	-0.155467	-0.06662
254	-0.161653	-0.07389
253.5	-0.186656	-0.10527
253	-0.22045	-0.11756
252.5	-0.273779	-0.12739
252	-0.328967	-0.16175
251.5	-0.344682	-0.14923
251	-0.313003	-0.13439
250.5	-0.309091	-0.13643
250	-0.315266	-0.1482
249.5	-0.387362	-0.19563
249	-0.429741	-0.2518
248.5	-0.491362	-0.27882
248	-0.463674	-0.27145
247.5	-0.497141	-0.28184
247	-0.536814	-0.28854
246.5	-0.615536	-0.33324
246	-0.638394	-0.39487
245.5	-0.658033	-0.43729
245	-0.662432	-0.48458
244.5	-0.719038	-0.54217
244	-0.746867	-0.60391
243.5	-0.74062	-0.64175
243	-0.795764	-0.69486
242.5	-0.986059	-0.7536
242	-1.16219	-0.88554
241.5	-1.20472	-1.0122
241	-1.24765	-1.13609
240.5	-1.34593	-1.22289
240	-1.52053	-1.33012

239.5	-1.67277	-1.42239
239	-1.81299	-1.54597
238.5	-1.89407	-1.68912
238	-2.07586	-1.84451
237.5	-2.26426	-2.04334
237	-2.4118	-2.24589
236.5	-2.57192	-2.44225
236	-2.78516	-2.62515
235.5	-3.03387	-2.80376
235	-3.30018	-3.01847
234.5	-3.53324	-3.25571
234	-3.73113	-3.48029
233.5	-3.95211	-3.76547
233	-4.23338	-4.04217
232.5	-4.51537	-4.31023
232	-4.8097	-4.59471
231.5	-5.11274	-4.88994
231	-5.36716	-5.17463
230.5	-5.65038	-5.48491
230	-5.91872	-5.78944
229.5	-6.18304	-6.07499
229	-6.43382	-6.31544
228.5	-6.70011	-6.55666
228	-6.99825	-6.82493
227.5	-7.18544	-7.11023
227	-7.44133	-7.42111
226.5	-7.61325	-7.66593
226	-7.85945	-7.83836
225.5	-7.99591	-7.98839
225	-8.23473	-8.17027
224.5	-8.44894	-8.33677
224	-8.66023	-8.51288
223.5	-8.77425	-8.64104
223	-8.89749	-8.73136
222.5	-8.9654	-8.85469
222	-8.96759	-8.93275
221.5	-9.06496	-8.9684
221	-9.20964	-8.99681
220.5	-9.31549	-9.04049
220	-9.3317	-9.08421
219.5	-9.41469	-9.17597
219	-9.45014	-9.2183
218.5	-9.59367	-9.19667
218	-9.63529	-9.18655
217.5	-9.63823	-9.14014
217	-9.63918	-9.10707
216.5	-9.69062	-9.0962
216	-9.6398	-9.13054
215.5	-9.58889	-9.10119
215	-9.62689	-9.12051

214.5	-9.70675	-9.09201
214	-9.78792	-9.06068
213.5	-9.76338	-9.05197
213	-9.73709	-9.08679
212.5	-9.75054	-9.06874
212	-9.80366	-9.11843
211.5	-9.81481	-9.18057
211	-9.92113	-9.24908
210.5	-10.1307	-9.30513
210	-10.1891	-9.36444
209.5	-10.219	-9.29846
209	-10.2297	-9.23284
208.5	-10.2231	-9.16024
208	-9.99094	-9.06359
207.5	-9.74575	-8.86553
207	-9.50237	-8.56578
206.5	-9.21028	-8.10457
206	-8.78663	-7.54665
205.5	-8.27843	-6.86575
205	-7.59338	-6.16133
204.5	-6.79912	-5.37609
204	-5.88353	-4.52003
203.5	-5.13994	-3.58405
203	-4.36755	-2.63932
202.5	-3.48356	-1.58863
202	-2.57759	-0.57299
201.5	-1.66776	0.50937
201	-0.513043	1.59042
200.5	0.553556	2.79226
200	1.56411	3.97121
199.5	2.6564	5.20903
199	4.03729	6.53729
198.5	5.2049	7.85129
198	6.44117	9.14357
197.5	7.64551	10.4186
197	8.71997	11.4758
196.5	9.72283	12.4415
196	10.4046	13.296
195.5	11.0834	13.8767
195	11.7274	14.3347
194.5	12.2444	14.6265
194	12.5474	14.7977
193.5	12.9966	14.9663
193	13.1855	15.2174
192.5	13.7108	15.1662
192	13.637	15.0264
191.5	12.7491	14.7863
191	11.6423	14.4937
190.5	11.393	14.2092
190	10.2025	13.7289

189.5	9.31562	13.2279
189	6.30989	12.9117
188.5	5.61974	11.8572
188	3.94638	13.7863
187.5	4.02786	13.6749
187	-5.70975	16.0934
186.5	-11.6197	17.2538
186	-13.418	14.8423
185.5	-9.20516	18.6801
185	-9.09571	18.0739
184.5	-10.3007	18.0693
184	-9.76468	15.0218
183.5	-9.34634	4.74729
183	-8.75456	-5.05061
182.5	-9.89169	-2.20158
182	-6.90118	-6.31822
181.5	-8.98785	-7.85791
181	-10.3194	-16.1631
180.5	-6.75955	-23.1315
180	-4.28473	-9.62155

Appendix 5: Supplementary data provided for whole cell protein array analysis

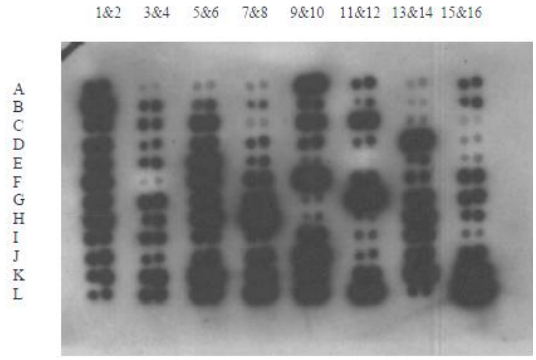
Normal tissue that appears to have greater expression of ADP-GK than cancerous tissue:					
Tissue		**Tumour type**	**Differentiation**	**Age**	**Sex**
Adrenal	**Cancerous**	Malignant fibrous histocytoma	N/A	58	F
	Normal	-	-	65	M
Breast (2)	**Cancerous**	Invasive ductal carcinoma	Poorly	48	F
	Normal	-	-	45	F
Lung (1)	**Cancerous**	Adenocarcinoma	Moderately	67	F
	Normal	-	-	26	M
Lung (3)	**Cancerous**	Squamous cell carcinoma	Moderately	46	F
	Normal	-	-	83	F
Lung (4)	**Cancerous**	Squamous cell carcinoma	Poorly	47	F
	Normal	-	-	26	M
Ovary (3)	**Cancerous**	Thecoma	N/A	22	F
	Normal	-	-	45	F
Ureter	**Cancerous**	Transitional cell carcinoma	N/A	63	M
	Normal	-	-	63	M
Uterus (1)	**Cancerous**	Endometrium adenocarcinoma	N/A	49	F
	Normal	-	-	58	F
Uterus (2)	**Cancerous**	Endometrium adenocarcinoma	N/A	52	F
	Normal	-	-	40	F
Pancreas (1)	**Cancerous**	Adenocarcinoma	N/A	60	M
	Normal	-	-	66	M
Pancreas (2)	**Cancerous**	Adenocarcinoma	Moderately	60	F
	Normal	-	-	56	F
Parotid	**Cancerous**	Adenocarcinoma	N/A	47	M
	Normal	-	-	82	F
Thymus	**Cancerous**	Thymoma (mixed)	N/A	45	M
	Normal	-	-	24	M

Cancerous tissue that appears to have greater expression of ADP-GK than normal tissue:					
Tissue		**Tumour type**	**Differentiation**	**Age**	**Sex**
Bone	**Cancerous**	Osteosarcoma	Well	24	F
	Normal	-	-	26	M
Brain (1)	**Cancerous**	Astrocytoma	N/A	40	M
	Normal	-	-	22	M
Brain (3)	**Cancerous**	Glioblastoma	Poorly	33	F
	Normal	-	-	26	M
Brain (4)	**Cancerous**	Neuroepithelioma	N/A	15	M
	Normal	-	-	82	M
Breast (3)	**Cancerous**	Invasive ductal carcinoma	Poorly	53	F
	Normal	-	-	35	F
Colon (1)	**Cancerous**	Adenocarcinoma	Moderately	58	M
	Normal	-	-	64	M
Colon (2)	**Cancerous**	Adenocarcinoma	Poorly	47	M
	Normal	-	-	26	F
Kidney (1)	**Cancerous**	Clear cell carcinoma	Well	44	F
	Normal	-	-	83	F
Kidney (2)	**Cancerous**	Granular cell carcinoma	Moderately	68	M
	Normal	-	-	62	M
Liver (1)	**Cancerous**	Adenocarcinoma	Poorly	52	M
	Normal	-	-	24	M
Liver (2)	**Cancerous**	Cholangiocellular carcinoma	Poorly	61	M
	Normal	-	-	91	F
Liver (3)	**Cancerous**	Hepatocellular carcinoma	N/A	65	M
	Normal	-	-	26	M
Lung (2)	**Cancerous**	Adenocarcinoma	Poorly	69	M
	Normal	-	-	24	M
Lung (5)	**Cancerous**	Squamous cell carcinoma	Well	72	M
	Normal	-	-	67	F
Lymph node	**Cancerous**	Lymphoma	Poorly	70	F
	Normal	-	-	28	M
Ovary (2)	**Cancerous**	Granular cell carcinoma	N/A	42	F
	Normal	-	-	45	F
Pancreas (3)	**Cancerous**	Adenocarcinoma	Moderately	61	M

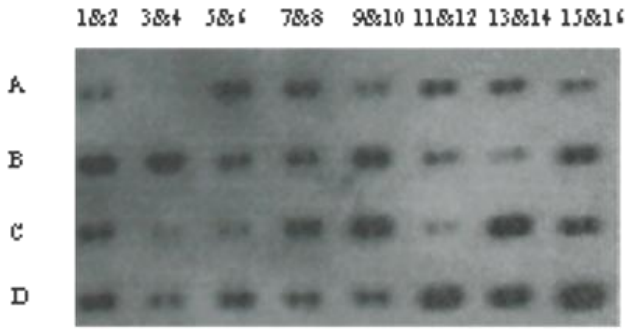
	Normal	-	-	60	M
Prostate	**Cancerous**	Sarcoma	Poorly	51	M
	Normal	-	-	26	M
Rectum (2)	**Cancerous**	Adenocarcinoma	Moderately	66	M
	Normal	-	-	66	M
Small intestine (2)	**Cancerous**	Non-Hodgekin's lymphoma	N/A	35	M
	Normal	-	-	62	M
Thyroid	**Cancerous**	Follicular carcinoma	Well	20	M
	Normal	-	-	25	M

Expression of ADP-GK in cancerous tissue and normal tissue that appears to be the same:					
Tissue		**Tumour type**	**Differentiation**	**Age**	**Sex**
Bladder	**Cancerous**	Transitional cell carcinoma	Moderately	50	F
	Normal	-	-	25	F
Breast (1)	**Cancerous**	Invasive ductal carcinoma	N/A	36	F
	Normal	-	-	36	F
Breast (4)	**Cancerous**	Invasive ductal carcinoma	N/A	26	F
	Normal	-	-	74	F
Oesophagus	**Cancerous**	Squamous cell carcinoma	Moderately	71	M
	Normal	-	-	83	F
Ovary (1)	**Cancerous**	Clear cell carcinoma	N/A	51	F
	Normal	-	-	51	F
Rectum (1)	**Cancerous**	Adenocarcinoma	Moderately	56	M
	Normal	-	-	28	M
Skin	**Cancerous**	Melanoma	N/A	25	M
	Normal	-	-	52	F
Small intestine (1)	**Cancerous**	Leiomyoma	N/A	57	M
	Normal	-	-	57	M
Stomach	**Cancerous**	Adeno-squamous cell carcinoma	Moderately	72	M
	Normal	-	-	60	M
Testis	**Cancerous**	Seminoma	Poorly	45	M
	Normal	-	-	26	M
Thymus (1)	**Cancerous**	Malignant thymoma	Moderately	79	F
	Normal	-	-	28	M

No apparent expression of ADP-GK in cancerous tissue or normal tissue:					
Tissue		**Tumour type**	**Differentiation**	**Age**	**Sex**
Brain (2)	**Cancerous**	Astrocytoma	N/A	27	M
	Normal	-	-	26	M
Duodenum	**Cancerous**	Adenocarcinoma	Moderate-poor	41	F
	Normal	-	-	26	M



Human adult normal and tumour whole cell protein array



Human adult and foetal whole cell protein array

The protein arrays were hybridised with GAPDH antibody. The pattern of variation in intensities among the array spots matches GAPDH expression from western analysis BioChain Institute Inc.).