Molecular Identification and Characterization of Camel Milk Insulin

A thesis submitted to the University of Kent for the degree of PhD in the Faculty of Science, Technology and Medical Studies

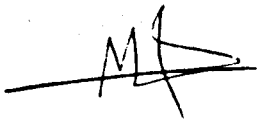
2013

Motasem Ismail

Department of Biosciences

Declaration:

No part of this thesis has been submitted in support of an application for any degree or other qualification of the University of Kent or any other University or Institute of learning.



Motasem Ismail

18/01/2013

Acknowledgements

All thanks are due to God, the creator, who has power over all things and I would like to express my deepest and profound acknowledgement to His Highness Sheikh Mohammed Bin Rashid Al-Maktoum the vice president of United Arab Emirates and The ruler of Dubai for his support to this project. I would like to thank Dr. Ali Ridah, CVRL administrative director and all members of CVRL and MBG laboratory for their support.

I am greatly indebted to my supervisor Professor Mark Smales for his continuous and generous suggestions, encouragement and endless patience throughout this project. My thanks also go to all members of Smales lab, who have helped me during my time at Kent.

I would like to thank Professor Darren Griffin for allowing me to undertake the FISH work in his lab, and also Herwig Riechl and Dr. Mirsolav Hornak for helping me in the Protein isolation and FISH procedures.

Special acknowledgement is owed to Mr. Mutasher Al-Badry the Deputy Manager of the Emirates Industry for Camel Milk and Products Company for providing me with camel milk samples.

In addition I would like to thank my Parents, brother, sisters and my lovely nephew Mohammad for encouraging me to overcome all difficulties. Special thanks goes to my friends; Abd Al-Ali Louhmadi, Dr. Osman Jafer, Dr. Abdullah Al-Mutery, Dr. Fatima Abidi and Maithili Shroff for all their support and encouragement.

Finally, I would like to dedicate this thesis to my brother Towfiq who passed away while I was writing this thesis.

Table of Contents:

Chapter 1 : Introduction

Chapter 2: Materials and Methods

Chapter 3 Camel Insulin Gene Sequencing, SNP, and Chromosome Localization Analysis

Chapter 4: Purification, Characterization and Identification of Camel Milk Insulin

Chapter 5: Production of Recombinant Camel Insulin and Investigations into Camel Milk Insulin Stability

List of Figures:

List of Tables:

List of Abbreviations

Acm	acetamidomethyl
BAC	bacterial artificial chromosome
Boc	*tert*-butoxycarbonyl
Bp	Base pair
CHO	Chinese Hamster Ovary
DEAE	Diethylaminoethanol
DMF	*N, N*-dimethylformamide
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ER	Endoplasmic reticulum
FISH	Fluorescence in situ Hybridization
Fmoc	9-Fluorenylmethoxycarbonyl
FRT	Flp recombination target
HOBt	1-hydroxybenzotriazole
IRS	insulin receptor substrate
kDa	Kilodalton
LB	Luria-Bertani
LC-MS	liquid chromatography-mass spectrometry

PBS	phosphate buffer saline
PCR	Polymerase chain reaction
PVDF	polyvinylidene difluoride
RIA	Radioimmunoassay
SDS	Sodium dodecyl sulphate
SNP	Single Nucleotide Polymorphisms
SPPS	solid phase peptide synthesis
t-Bu	*tert*-butyl
Tris	Tris(hydroxymethyl)aminomethane
UV	Ultra violet

Abstract

The number of people suffering from Diabetes mellitus in both developed and developing nations has increased dramatically over the past two-three decades and this increase is predicted. The condition can result in numerous complications in the patients suffering from the disease if it is not with long-term effects including retinopathy with potential blindness, nephropathy that may lead to renal failure, and/or neuropathy and an increased risk of cardiovascular, peripheral vascular and cerebrovascular disease. Camel milk consumption has been reported to have a positive effect with respect to diabetic severity and complications and as a treatment of diabetes with regular camel milk drinkers requiring lower doses of insulin to control their blood glucose levels. A number of studies have now suggested a direct link between camel milk insulin and the effect of drinking camel milk on diabetic suffers. In this study, insulin from dromedary camel (*Camelus dromedaries*) milk was isolated and characterized in order to investigate the stability and nature of the protein to begin to understand why it may have positive benefits on patients consuming camel milk Initially the camel insulin gene was amplified and sequenced from 32 camels using primers used to amplify the camel proinsulin cDNA. The sequenced fragment was aligned using a BLAST search and the sequence matched a sequence from *lama* with 96% homology. A number of SNPs were identified in key regions of the camel gene sequence, although none in the coding sequence with 3 located in the C-peptide intron, 2 in the 3'-UTR and 1 after the termination signal. When the gene sequence was converted to the corresponding amino acid sequence this revealed that across the 13 species compared the camel insulin has two unique amino acid changes in the signal peptide, one unique amino acid change in the B-chain and two changes in the A-chain compared to human (which are observed in some other species). There is much variation across species in the C-peptide region and in the camel there were various changes compared to other species. Using the sequence, 3D homology modeling of camel proinsulin was undertaken which suggests the formation of different secondary structure compared to that for human insulin which

may impact upon the stability of the molecule. Using FISH, the camel insulin gene was also mapped to the distal (Telomeric) end of the q-arm of camel chromosome 10. In addition to investigating camel insulin at the genomic level, studies were undertaken at the protein level. For protein analysis, fresh camel milk samples were collected from the Emirates dairy farm. The samples were initially defatted and then treated with ethanol in order to precipitate casein which is the major milk protein and other large molecular weight proteins, before the supernatant was passed through an ion-exchange chromatography column of A-25 DEAE Sephadex beads. The bound camel milk insulin was then eluted and further purified by immuno-affinity chromatography using an anti-human insulin antibody. The resulting insulin concentration and activity was measured by radioimmunoassay using a human insulin kit. The camel insulin gene was also cloned into a mammalian expression vector and transfected into a mammalian cell line and attempts made to produce cell lines stably expressing camel insulin. The thermostability of the insulin purified from camel milk was then compared with human and bovine insulin in camel and bovine milk. The results suggest that camel insulin appears to be stabilised due to the unique both amino acid variants in comparison to human insulin but these studies also suggested that camel milk itself provides protection against thermostability which bovine milk does not and this property contributed the most towards the enhanced thermostability of camel milk insulin. The exact nature of this protection and the agents responsible are not currently known, but further elucidation of the mechanism(s) may provide possible routes for the thermostabilisation of other proteins.

Chapter 1
Introduction

1.0 Background and Introduction to Thesis

In the Middle East camel breeding is of great importance and hence information and scientific investigations around the biology of the camel are at the forefront of science in general in the area. The camel is important in this area because of its economic value, as a source of transport, meat and milk. In addition, camel racing is a popular sport in the Arabian Gulf region and racing camels are worth consider sums of money with many valued in the millions of US dollars. The Food and Agriculture Organization (FAO, 1996) reported that there are approximately 18 million camels in the world, which maintain the life of millions of people in arid and semi-arid areas. The high nutritional value of camel milk has made it one of the most important food sources for humans in many parts of the world, particularly in these arid and semi-arid areas where highly nutritional food sources can be limited in their supply or availability (Elagamy *et al.*, 1998; Elagamy, 2000). Of particular interest for this thesis is milk provided by the camel, not for the nourishment perspective but as camel milk and the insulin provided within this, appears to offer benefits to diabetic patients and reduce the dependence of such patients upon the administration of insulin (Agrawal *et al.*, 2007a). Hence, this thesis set out to investigate and characterise camel insulin in more detail.

1.1 The Camel Family and Taxonomy

The Camelids *Camelidae* family are derived from the suborder *Tylopoda* (pad-footed animals) of the order *Artiodactyla* (even-toed ungulates). Camelids are ruminating animals but are not classified under the *Ruminantia* (ruminants) suborder because of

differences in the foot structure, in the gastric system and because of the nonexistence of horns (Schwartz & Dioli, 1992; Fowler, 1998; Wernery, 2003). The *Camelidae* family is further classified into three genera: *Camelus, Lama* and *Vicugna.* In older literature the *Camelidae* family classification was restricted to the first two of these genera. The gunus *Camelus* refers to the old world camel and includes *Camelus dromedaries* and *Camelus bactrianus* as domesticated species. *Camelus ferus*, found in the Gobi desert is known as the wild species of the *camelus* genera (Wilson & Reeder, 2005).

The Dromedary camel (*Camelus dromedarius)* is a one humped camel and is found in the hot deserts of Africa and Asia. The dromedary name is derived from the Greek word "dromeus" that means runner or "droma" running. The Bactrian camel (*Camelus bactrianus*), the two-humped camel, is distributed and typically found in the cold deserts and dry steppes of Asia. The Bacterian name is derived from Bactriana in Central Asia (Rao *et al.*, 1970; Fowler, 1998; Jassim & Naji, 2002).

The general habitat of the camel is therefore in tropical and subtropical dry zones such as North Africa western Asia and North West India. In last century the number of camels in Australia has dramatically increased, as although they are found in the large desert areas in this country there is plentiful food and water resources for the camel to survive and flourish. Camels are also present in the United States of America, Central America, the Caribbean, southern Africa and Europe but in much lower numbers and density (Wilson, 1984, Wilson *et al.*, 1990). The camel is therefore extremely adaptable and able to survive extremes of climate and terrain, making this a very hardy animal able to populate most of those areas globally that would be inhospitable to many other species, but also able to populate areas more favourable to other species.

About five thousand years ago the Dromedary camel (*Camelus dromedaries*) was first domesticated in southern Arabia. This camel has the ability to survive in the harsh desert

climate for up to forty days without drinking and hence has made it a recognised creature of the desert (Bulliet, 1975; Higgins, 1984; Hassan, 1971). For these reasons it became the most important domesticated animal in this region, providing milk, transport across expansive desert terrain, meat and companionship.

1.2 Milk Proteins

Milk is the material generated to feed mammalian infants with all mammals from humans to whales producing milk for this purpose. Around nine thousand years ago animals such as cows, buffalos, sheep, goats and camels were domesticated by man for the purpose of milk and meat production (Eddleman, 1999). The milk is not only used as a direct source of nourishment and protein, but also for the production of other foodstuffs such as cheese and yoghurts.

There has been a huge volume of research undertaken into the milk of many different mammalians and the benefits and roles of the proteins and lipids/fats within these milks and it is beyond the scope of this thesis and introduction to describe these. The milk from many species has significant health benefits if consumed by humans, although in some cases allergies or conditions can result from the consumption that can be life threatening, for example due to lactose intolerance (Qiao *et al.*, 2011). Protein biomarkers in milk have also been suggested as useful markers for disease states and for genomic profiling (Baumgartel and Conley, 2012).

Milk proteins are thus well studied, although much research continues into these economically value proteins. The importance of milk whey proteins is due to many functional aspects, several of which will be discussed here. Firstly, some milk proteins have the ability to be modified, for example to have regions such as the phosphorylation sites in casein that change the properties of the protein (Beg *et al.*, 1984; Beg *et al.*, 1986). Second, some milk proteins are similar to anti-proteases, neurophysins and other

such proteins in their structure and function that indicate the wide evolutionary and functional relationships between these proteins (Dandekar *et al.*, 1972; Hennighausen *et al.*, 1982; Beg *et al.*, 1984; Beg *et al.*, 1986; Urashima *et al.*, 2012). Thirdly, the majority of milk proteins are much like other proteins that are similar in amino acid sequence, appearance and content between different mammalian species and hence are found in the milk of most mammals and are presumed to play the same role across mammals in sustaining infants. However, a minority of milk proteins have been reported as being present or absent in milk from different species and hence there are some species specific proteins present/absent. For example, camel milk is deficient in β-lactoglobulin which is a major protein in other milks (Kessler, 1970; Liberatori, 1979), whilst a number of milk proteins have unidentified function but are suggested to be involved in regulatory processes (Dandekar *et al.*, 1982; Reeds *et al.*, 2000). The deficiency of β-lactoglobulin in camel milk is potentially important as this is one of the major allergens in cow milk and is problematic for a percentage of the population (Adel-Patient *et al.*, 2012) who may therefore be able to drink camel milk in place of cow milk.

For human infants, cows milk is a major source of nutrition and essential nutrients for growth and development, and provides essential proteins, minerals, carbohydrates, fatty acids, immunoglobulins and growth factors where human milk is not available (El-Hatmi, 2007). The existance of specific growth factors in human milk and colostrum, and their activity on different cell types, was first identified in human milk and colostrum in the 1980's, and then subsequently in bovine milk, colostrum and whey (Sylvie *et al.*, 2006). For example, the presence of very active growth related cytokines in bovine colostrum and milk have been identified and characterised by a number of groups in the proceeding years. For example, the G-CSF (Calhoun *et al.*, 1999), IL-1 (Kelly, 2003), IL-1β (Hagiwara *et al.*, 2000), IL-6 and INF-ϒ (Hagiwara *et al.*, 2000, Kelly, 2003), TNF-α (Blum *et al.*, 2000, Hagiwara, 2000; Kelly, 2003; Ontsouka *et al.*, 2003) and IL-18 (Muneta *et al.*, 2005) have all been identified and isolated from bovine milk and/or colostrum and characterised.

The stage of lactation, and the methods used for the quantification of growth factor concentrations in milk explain the fluctuations in the concentrations and presence of these biomolecules. The concentration of growth factors reaches the maximum levels during the first hours after delivery of the new born and then decrease drastically in a time-dependent manner. The total protein concentration in milk and colostrum also decreases during the first four days after delivery (Sylvie *et al.*, 2006). A further important signalling molecule/hormone with regard to this thesis that is particularly found in camel milk is insulin (Agrawal *et al.*, 2005; Wernery *et al.*, 2006). The role of insulin and its presence in milk are discussed in greater detail below.

1.3 Camel Milk

Camel milk is highly nutritious, contains less fat than cow milks, higher levels of potassium, iron and vitamin C and relative to other milks large amounts of insulin (Knoess *et al.*, 1979). The presence of volatile acids such as polyunsaturated fatty acids and linoleic acid increased the value of camel milk compared to other milks and these acids are vital in human nutrition (Rao *et al.*, 1970). Camel milk not only nutritional value, many years ago was proposed to have properties as a 'biodrug' (Rao *et al.*, 1970). For nomads in arid regions camel milk is one of the most valuable food resources and in urban populations the consumption of camel milk is increasing (Farah, 2004).

The visual appearance of camel milk is that of a very white colour with a foamy structure. About 13,500 liters of milk (or even more in some cases) is the volume of milk that can be produced by the typical camel during the lactation period, a period that can last for up to 18 months (Knoess, 1977). The type of the food that a camel eats can affect the taste of the milk and hence the makeup of the milk with the feeding of green fodder producing sweet milk, while certain shrubs and herbs in arid regions can produce salty milk (Rao *et al.*, 1970). The casein content in camel and cow milk is reported to be

very similar, but the whey fraction in the case of the camel is larger than the whey fraction from cow milk. As a result, in the camel the ratio of whey protein to casein is higher than in cow milk but lower than found in human milk. This may explain and clarify the softness of camel milk coagulum compared to that found in cows (El-Agamy, 1983).

As mentioned above, the composition of camel milk is considered very different from that of other ruminants. Haddadin *et al.* (2007) reported a large variation in the concentrations of Arabian dromedary camel milk components over the year, reporting that the concentration of water in camel milk was 877 ± 12.7 g\L, lactose was 39.2 ± 0.5 g\L, fat content was 29.5 ± 2.6 g\L, protein content was 26.9 ± 1.6 g\L and ash content was 8.2 ± 0.3 g\L. The highest concentration of water was observer in August at the peak of the summer while the lowest was observed on January, whilst the concentration of the fat, proteins and ash were lowest in August (Haddadin *et al.*, 2007).

In comparison with the human and bovine milk, the water content in one litre of bovine milk is 878 g/L while in the human milk it is 871 g/L. The total protein contents in bovine milk 32 g/L is and in the human is 13 g/L. The fat in bovine (39 g/L) and human milk (41 g/L) is more the camel milk. The lactose quantity in bovine milk is 48 g/L while in human milk is 72 g/L. Table 1.1 is showing the differences between camel, human and bovine milk samples in the water, fat, lactose and total protein contents (Emmett and Rogers, 1997).

Table 1.1 Comparisons between camel, human and bovine main milk components.

Nutrient	Camel milk	Bovine milk	Human milk
Water	877	878	871
Lactose	39.2	48	72
Fat	29.5	39	41
Proteins	26.9	32	13

As mentioned above, camel milk is deficient in β-lactoglobulin which is the major serum protein available in ruminant milk. Serum albumin, β-lactalbumin, immunoglobulins, lactoferrin and peptidoglycan recognition protein are examples of key proteins in most mammalian milks that have been also been identified in camel milk (Kappeler *et al.*, 2004). Interestingly, the immunoglobulins of camel are different to many other mammals in that there are heavy chain only antibodies with no light chain, so termed single-domain antibodies (Muyldermans *et al.*, 2001), which are now generating much interest as potential biotherapeutic molecules. Finally, the pH of fresh camel milk is high in comparison to other milks, however the pH decreases rapidly when the milk is left to stand at room temperature (Agrawal *et al.*, 2004a).

1.3.1 Camel milk medicinal properties

Initially, milk was only deemed to be a source for particular nutritional factors such as essential amino acids (Hambraeus, 1992). However, more recent studies have shed light on the nutritional and functional characteristics of milk, due to the abundance of bioactive substances which provide various health benefits. It has observed in different parts of the world that both fresh and fermented Dromedary camel milk serves as a potential source of remedy for various illnesses such as autoimmune diseases and metabolic disorders, dropsy, jaundice, tuberculosis, asthma, spleen problems, piles, anemia and diabetes (Rao *et al.*, 1970; Abdelgadir *et al.*, 1998; Shalash, 1984). For example, Mal *et al.* (2001) observed a positive role for camel milk in treating chronic pulmonary tuberculosis, a major disease in developing nations. Camel milk has also been reported to have a therapeutic effect on patients with ulcers and hepatitis. Furthermore, there are no reported lactose intolerance or allergies against camel milk as a result of its consumption (Zagorski *et al.*, 1998; Shabo *et al.*, 2005), a potential issue observed with consumption of cows milk as described above.

1.3.2 Camel milk and allergies

It has previously been shown that camel milk proteins are involved in food allergy prevention and treatment. B-lactoglobulin and different β-casein proteins are responsible for allergies that develop in response to drinking cows milk, however these proteins are absent from camel milk. Camel milk also contains some immunoglobulins that are at least partially very similar to human milk immunoglobulins, which decrease the allergic reactions in children and help in their future tolerance to the food (Shabo *et al.*, 2005). Shabo *et al.* (2005) studied the effect of camel milk on allergy treatment. In this study were included 8 children from parents suffering from acute food allergy who had no response to the conventional allergy treatment, with the children aged between 4 months and 10 years. In the study the parents were advised to feed their children with camel milk and impressive results with this camel milk treatment were observed with a rapid improvement in the children's health and the children being generally less sensitivity/allergic to various foodstuffs (Shabo *et al.*, 2005).

1.3.3 Antimicrobial activity of camel milk

Many studies have described the antimicrobial activity of camel milk. Amongst these it has been reported that camel milk has antimicrobial effect against wide range of Gram negative and Gram positive bacteria, including *Escherichia coli, Listeria monocytogenes, Staphylococcus aureus and Salmonella typhimurium* (El-Agamy *et al.*, 1992; Benkerroum *et al.*, 2004). The antimicrobial activity of camel milk has been found to be due to the presence of lysozyme, hydrogen peroxide, lactoferrin, lactoperoxidase and immunoglobulin's which together all contribute to the many described examples of the antimicrobial activity of camel milk (El Agamy *et al.*, 1992).

1.3.4 Hypocholesterolaemic effect of camel milk

One of the foremost causes of death in the developed world, as well as in many developing countries, is coronary heart disease (Pereira and Gibson, 2002). One of the

key risk factors leading to coronary artery heart disease is the increase in levels of blood and dietary cholesterol (Reddy *et al.*, 1977; Spence *et al.*, 2010). Elayan and his group described in 2008 that administering fermented camel milk with *Bifidobacterium lactis* (BB-12) to rats *in vivo* led to hypocholesterolaemic. The administration of fermented camel milk was effective in lowering the level of total plasma cholesterol, very low density lipoprotein (VLDL) and the low density lipoprotein (LDL) in the tested rats compared with the controls (Elayan *et al.*, 2008). It is postulated that such a benefit might also be present in people who regularly drink camel milk.

1.3.5 Hypoglycaemic effect of camel milk

A significantly reduced occurrence of diabetes has been observed in the Raica community in India and it is thought that this is the result of the consumption of Dromedary camel milk since the prevalence of diabetes is zero among those who commonly drink camel milk (Agrawal *et al.*, 2007b). Camel milk consumption also provides effective management for patients with type 1 diabetes (Agrawal *et al.*, 2003). The elevated concentration of insulin in camel milk is thought to contribute towards this effect (Agrawal *et al.*, 2003; Beg, 1986). The effect of camel milk on diabetics will be discussed further in the following sections, however why there are increased levels of insulin found in camel milk and if this has unique properties compared to other insulin molecules was not known when the work in this thesis was begun.

With regard to the anti-diabetic properties of camel milk, Agrawal *et al.* (2007a) studied the effect of camel milk on type one diabetics in a study that included 50 patients. The patients were separated into two groups based on their age, body mass index, blood glucose, glycated hemoglobin (HbA1c), current insulin dose, plasma insulin, C-peptide level, and concentration of anti-insulin antibodies. For 12 months group 1 was treated with conventional treatments and group 2 received the conventional treatment along with 500 ml of raw camel milk daily. After the treatment period the camel milk

consuming group showed a significant decrease in glucose levels from 115.16 ± 14.50 mg/dl to 100.20 ± 17.40 mg/dl ($p = 0.002$) and a dramatic reduction in the insulin dose required from 30.40 ± 11.97 uU/day to 19.12 ± 13.39 uU/day ($p < 0.001$). There was a slight reduction in the percentage of HbA1c from 8.2 ± 1.44 % to 7.51 ± 0.98 % ($p = 0.06$), and a slight increase in the level of C-peptide observed in the camel milk consuming group. There was no change in the amount of anti-insulin antibodies present in both groups.

Agrawal *et al.* (2007b) also reported a positive effect of camel milk on β-cell function and suggested that there were several possible mechanisms by which this might occur. These were; (1) That by helping to maintain the glucose concentration within normal values may lead to less activity is required by β-cells that ultimately conserves and reserves β-cell function, (2) that the high level of circulating insulin may cause tolerance stimulation in the body, (3) that single domain immunoglobulins in camel milk, which are small in size and weight, may interact with host cell proteins to induce the regulatory cells and eventually discourage the immune system and β-cell salvage, (4) that the existence of half-cysteine whey protein and lactoferrin in camel milk may stimulate β-cell function.

1.4 Insulin and Diabetes

Diabetes as a disease has been recorded for more than 3,500 years, but its cause remained a mystery until the early 1900's. The earliest recorded case of diabetes is on the Papyrus Ebers of Ancient Egyptians in 1500 B.C. describing frequent urination (polyuria) as a symptom. In the first century A.D. Arateus described diabetes as 'the melting down of flesh and limbs into urine' (Lasker *et al.*, 2010). At about the same time a Greek physician, Galen of Pergamum, mistakenly diagnosed diabetes as an ailment of the kidneys. It was not until the 11th century that the term mellitus was added which comes from the Latin word for honey. The common way for diagnosing diabetes was by 'water tasters' who drank the urine of those suspected of having diabetes as it was thought to be sweet tasting (Desbandhu *et al.*, 2011).

In the early 19th century, tests were derived to measure the presence of sugar in the urine. In 1869, Paul Langerhans, a German medical student, reported that the pancreas contained two types of cells; one responsible for the secretion of normal pancreatic juice and the other later identified as the 'islets of Langerhans' (Lasker *et al.*, 2010). In 1909, Jean De Meyer gave the name 'insulin' to the glucose lowering hormone which was still hypothetical at that time and he proposed that the hormone is produced by the islet tissue of pancreas, the name was originated from the Latin word *insula* (*insula* = island) (Lasker *et al.*, 2010).

Insulin was discovered in 1921 by the Canadian scientists Banting, Best, Macleod and Collip, as a substance which reduced the circulating blood sugar level. They subsequently isolated insulin from animal pancreas islands. Until this discovery, type1 diabetes mellitus was often and eventually fatal for those patients suffering from it. Initially they tried administering a pancreas extract on dogs and then they subsequently used such an extract on patients with human diabetes. After several trails they were discovered they were able to use their extract to treat diabetes successfully for the first time (Banting and Best, 1922; Rosenfeld 2002; Joshi *et al.*, 2007; Teuscher, 2007).

In January 1922, a bovine insulin extract was injected into a diabetic human patient. The extract was so impure it caused a 7.5 cm callus at the injection site on the left buttock of the first insulin patient Leonard Thompson. James Collip therefore continued his work to further purify the insulin extract to make it safer for administration and more effective. Every sample of the early isolated insulin had a different effect as a result of the divergence in the purity of the sample and presumably due to differences in insulin concentration. As a result, the early diabetics to be treated with insulin frequently had hypoglycaemic responses, pain and swelling at the injection site (Bliss, 1982; Rosenfeld, 2002; Joshi *et al.*, 2007). Subsequently new purification procedures were developed to generate insulin from swine pancreas and with the advent of recombinant DNA

technology recombinant insulin was developed and expressed in *E.coli* as discussed in more detail below.

Diabetes has a number of effects on those that suffer from the condition, including the long-term damage, dysfunction and failure of various organs if the disease is not controlled (Fowler, 2008). Patients may present with characteristic symptoms including thirst, polyuria, blurring of vision, and weight loss (Ringborg *et al.*, 2009). Often the initial symptoms are not severe and hence pathological and functional damage may proceed for some time before a diagnosis is made. The long-term effects include retinopathy with potential blindness, nephropathy that may lead to renal failure, and/or neuropathy (Ryden *et al.*, 2007). People with diabetes are also at increased risk of cardiovascular, peripheral vascular and cerebrovascular disease (Fowler, 2008; Ringborg *et al.*, 2009).

Diabetes is classified into two broad categories, type 1 and type 2. In type 1 diabetes the cause is a deficiency in insulin secretion and accounts for ~10% of all cases of diabetes. In type 2 diabetes the cause is a combination of resistance to insulin action and an inadequate compensatory insulin secretory response (American Diabetes Association, 2009). In those patients with type 2 diabetes (also referred to as age-onset diabetes), a degree of hyperglycemia sufficient to cause pathologic and functional changes may be present for a long period of time before diabetes is detected. Thus, the majority of diabetes patients are Type 2 diabetics (American Diabetes Association, 2009).

It has been observed that the number of patients with Diabetes mellitus has been increasing dramatically over the past two decades such that the increase is described as an epidemic, leading to many complications in the patients suffering from this disease and a substantial variation in the number suffering worldwide in different countries (King *et al.*, 1998; Diamond, 2003). The figure of the International Diabetes Federation showed a very high number of people diagnosed with diabetes and the number has

increased from 30 million people to more than 246 million people in only the last twenty years (Hossain *et al.*, 2007). In 2000 the global prevalence of diabetes mellitus was estimated for all age groups to be 2.8% and was predicted to rise to 4.4% in 2030 (Wild *et al.*, 2004). The lions share of the drastic increase to date, and of the predicted increase is in developing and third world countries. The developing countries are also predicted to increase further as a consequence of the rapid increase in the aging population these countries will have with improved medication and access to foodstuffs, since the prevalence of diabetes increases with age. It has thus been estimated that by 2025 among those aged 20 years and above the number of diabetics will increase by 35%. India is predicted be the country that will have the largest number of diabetics and the incidence is rising especially in urban areas (Ramchandran *et al.*, 2001). Socio-economic, cultural and behavioural parameters influence the prevalence as well as the proportion of undiagnosed diabetics and thus the number of undiagnosed patients cannot be extrapolated to the whole population (Agrawal *et al.*, 2004a).

The insulin hormone is important for growth and metabolism in most eukaryotes. It is particularly required for normal glucose homeostasis as described above. Insulin deficiency causes diabetes mellitus and due to its clinical significance it is considered an important model protein for structural studies (Steiner *et al.*, 1985). Insulin enhances the uptake of glucose in muscle and adipose tissue by translocating glucose carriers to the cell surface from the intracellular pool (Nelson and Alkon, 2005). Mature human insulin consists of two separate chains, namely the A-chain which consists of 21 amino acids and the B-chain which consists of 30 amino acids. Both chains are joined by specific disulfide bridges and are within a single polypeptide chain from a single gene, which is matured and processed by proteolytic processing to proinsulin and then to mature insulin. The C-peptide is removed from the polypeptide in the process of generating mature insulin, and is the portion of the polypeptide that directly connects the A- and B-chains (Chang *et al.*, 1998). The biosynthesis of insulin therefore includes as the last step the proteolytic conversion of a single chain precursor proinsulin to the double chain

product of insulin (Steiner & Oyer, 1967). A specific number of amino acids, for example 33 and 35 respectively in the cases of porcine and human proinsulin, are excised from the middle of the proinsulin during this process to generate the C-chain or C-peptide (Chance *et al.*, 1968; Ko *et al.*, 1971).

Goeddel *et al.* (1979) synthesized the human insulin gene by recombinant DNA technology. The A- and B- chains were cloned separately, the pBR322 plasmid being used as the cloning vector, to afford efficient transcription and translation and the recombinant genes merged to an *Escherichia coli* β-galactosidase gene. The mature insulin protein was spliced from the β-galactosidase. The insulin chain material was then purified and identified by radioimmunoassay. A complete purification for the individual A-chain and an incomplete purification for the B-chain was then performed, both chains were mixed, reduced and reoxidized to generate human insulin.

1.5 Insulin Biosynthesis

The insulin gene encodes a preproinsulin molecule as a single chain polypeptide; the preproinsulin is proinsulin with the signal peptide that directs the molecule to the endoplasmic reticulum (ER). The ER is where the initial maturation of the insulin biosynthesis process takes place (see Figure 1.1) (Harding and Ron, 2002; Joshi *et al.*, 2007). Once the preproinsulin is synthesised in the cytoplasm, it is translocated into the lumen of the ER through an interaction between the signal recognition particle (SRP) and the signal peptide on the insulin polypeptide, the SRP then binding to the SRP receptor on the ER membrane (Huang and Arvan, 1994).

The signal peptide of preproinsulin gets cleaved off in the ER to produce the proinsulin molecule. In the lumen of the ER, proinsulin completes the process of protein folding and the formation of three disulphide bonds, theses disulphide bridges being essential for insulins stability and bioactivity (Narhi *et al.*, 1993; Dai and Tang 1996; Hober *et al.*,

1997; Weiss *et al.*, 2000; Guo and Feng., 2001; Yan *et al.*, 2003). The properly folded proinsulin is then carried to the Golgi apparatus for packaging in the secretory granules (Huang and Arvan, 1994).

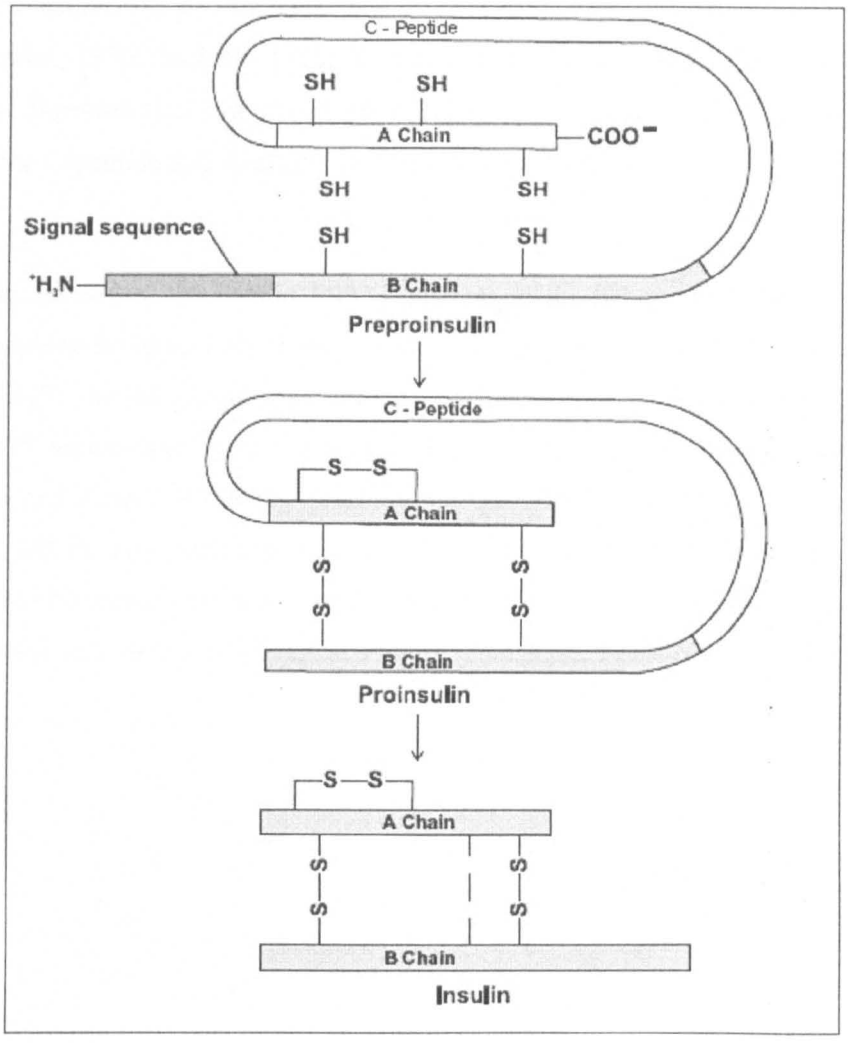


Figure 1.1. Insulin biosynthesis steps from preproinsulin to proinsulin and finally the mature insulin (Joshi *et al.*, 2007).

Proinsulin binds only weakly to the insulin receptor and the bioactivity of the insulin hormone increases massively after proteolytic processing to generate mature insulin (Galloway *et al.*, 1992). The conversion of proinsulin to insulin takes place inside the

secretory granule of the pancreatic β-cell during its maturation after passing through the Golgi apparatus and entering the immature secretory granules (Huang and Arvan, 1994). At this stage the C-peptide is cleaved by a specific prohormone convertase to genertae the mature insulin which is then released by exocytosis along with the C-peptide (Grant and Coombs, 1970; Steiner, 1998; Weiss, 2009). The cleavage takes place at two conserved digestion sites, the first between the B-chain and the C-peptide and the other between the C-peptide and A-chain junctions (Weiss, 2009).

As mentioned above, the insulin folding process in the ER requires the formation of three disulphide bridges. Two of these bridges are between the A and B-chains (A7-B7 and A20-B19), the third bond is an intrachain bridge within the A-chain between the 6^{th} and the 11^{th} amino acids in the chain (see Figure 1.2) (Hua *et al.*, 2006; Narhi *et al.*, 1993; Dai and Tang 1996; Weiss *et al.*, 2000; Guo and Feng., 2001; Yan *et al.*, 2003; Teuscher, 2007). To obtain and maintain the final structure of insulin each disulphide bond should be correct and maintained. Disulphide isomers between the A and B-chain are less stable and have low biological activity (Hua *et al.*, 1995; Hua *et al.*, 2002).

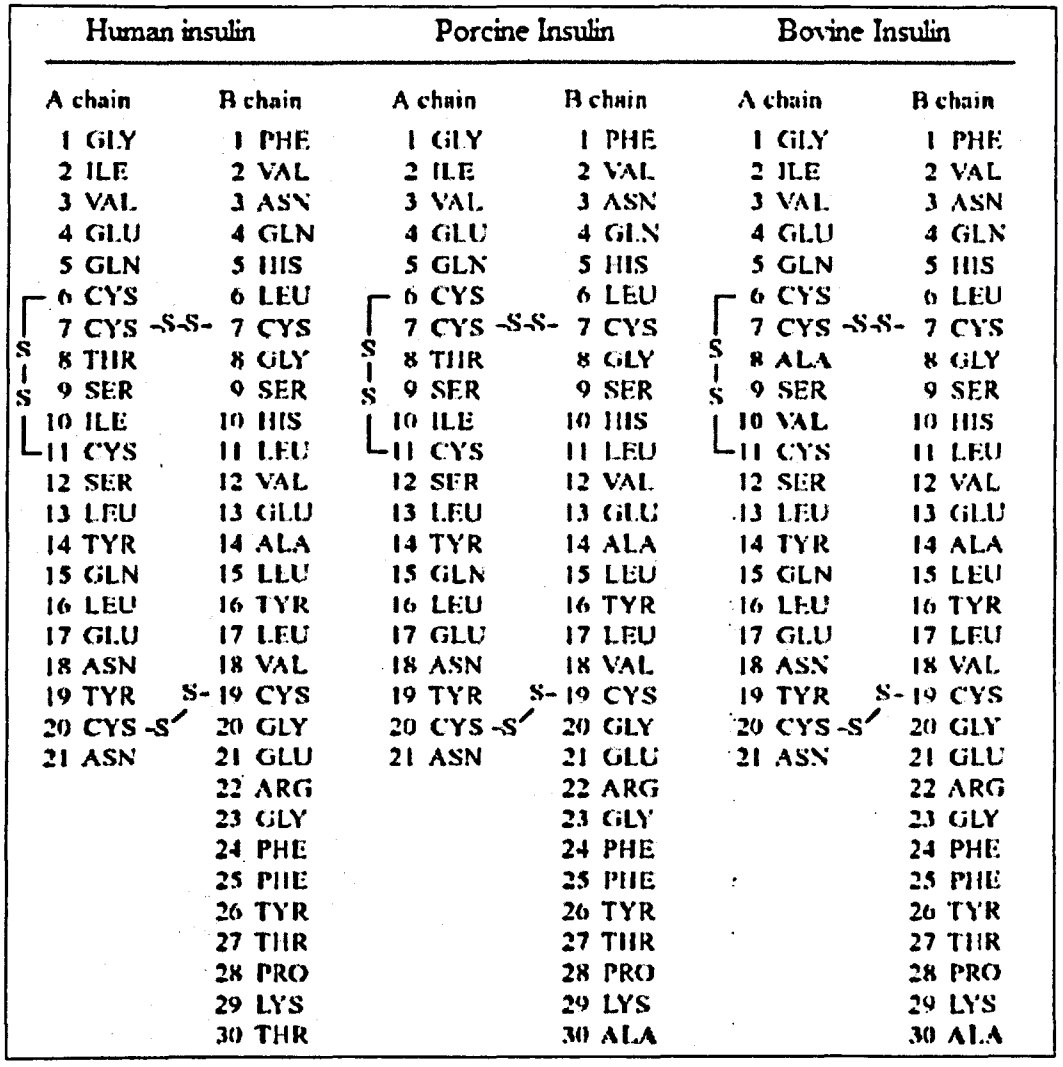


Figure 1.2. The amino acid sequence of human, porcine and bovine insulins and the location of the inter- and intra-chain disulphide bridges in the three insulins (Teuscher, 2007).

1.6 Insulin Hormonal Regulation

The insulin molecule as a hormone has various effects and functions in the body. Within seconds following the sensing of insulin in the blood stream by insulin receptors, 80% of body cells raise glucose uptake and cell membranes become more penetrable to various amino acids, phosphate and potassium ions (Guyton and Hall, 2006). Leftover carbohydrates, which cannot be put into storage as glycogen, are also turned under the

stimulation of insulin into fat and accumulate in the adipose tissue (Guyton and Hall, 2006). Indeed, the rate of glucose uptake in adipose and muscle cells is regulated by insulin. Insulin also stimulates and thereby is involved in the regulation of many pathways. The insulin receptor is embedded in the plasma membrane and is composed of two extracellular alpha subunits and two beta subunits embedded in the plasma membrane (see Figure 1.3 for a schematic representation). The receptor is a tyrosine kinase receptor and once insulin binds to the receptor alpha subunit, the beta subunit catalyses the substrates tyrosine phosphorylation (Chang *et al.*, 2004; Shepherd, 2005; Nadendla, 2008).

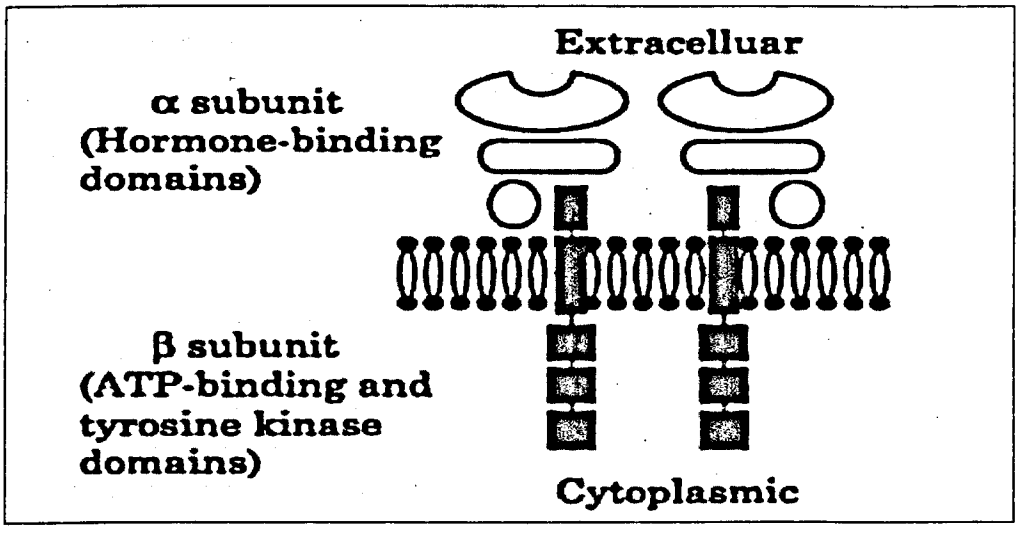


Figure 1.3. The alpha and beta subunits of the insulin receptor and their intra and extracellular locations (Nadendla, 2008).

Tyrosine phosphorylation of the insulin receptor substrate (IRS) results in the stimulation of the phosphatidylinositol 3-kinase (PI-3K) pathway. The PI-3K pathway is responsible for the production of the phosphatidylinositol 3,4,5-tris-phosphate (PIP3) in the plasma membrane. PDK1 protein (3-phosphoinositide-dependent protein kinase) is activated by the presence of PIP3. Then, the PDK-1 phosphorylates and stimulates both

atypical protein kinase C (aPKC) and the protein kinase B (PKB) proteins, which are required for the glucose transporter (Glut 4) translocation to the cell membrane, this process being summarised in figure 1.4 (Chang *et al.*, 2004; Shepherd, 2005; Eriksson, 2007). Protein kinase B (PKB) is a key regulatory protein as it promotes the phosphorylation of many proteins, including pro-apoptotic protein (BAD), glycogen synthase kinase-3 (GSK3β) involved in cellular growth and glycogen synthesis, and also the subclass O of the forkhead family transcription factors Foxo1 which are involved in regulating gene expression (Bhatt *et al.*, 2010). Furthermore, insulin binding to its receptor results in the phosphorylation of the adapter protein APS which activates the G protein TC10 that is foound in the lipid rafts. TC10 is involved in the control of a number of cellular mechanisms including changing the actin in the cytoskeleton, controlling the adapter protein CIP4, and building of the exocyst complex. Together these pathways are important in controlling and recovering the glucose transporter Glut4 (Chang *et al.*, 2004; Kanzaki *et al.*, 2004).

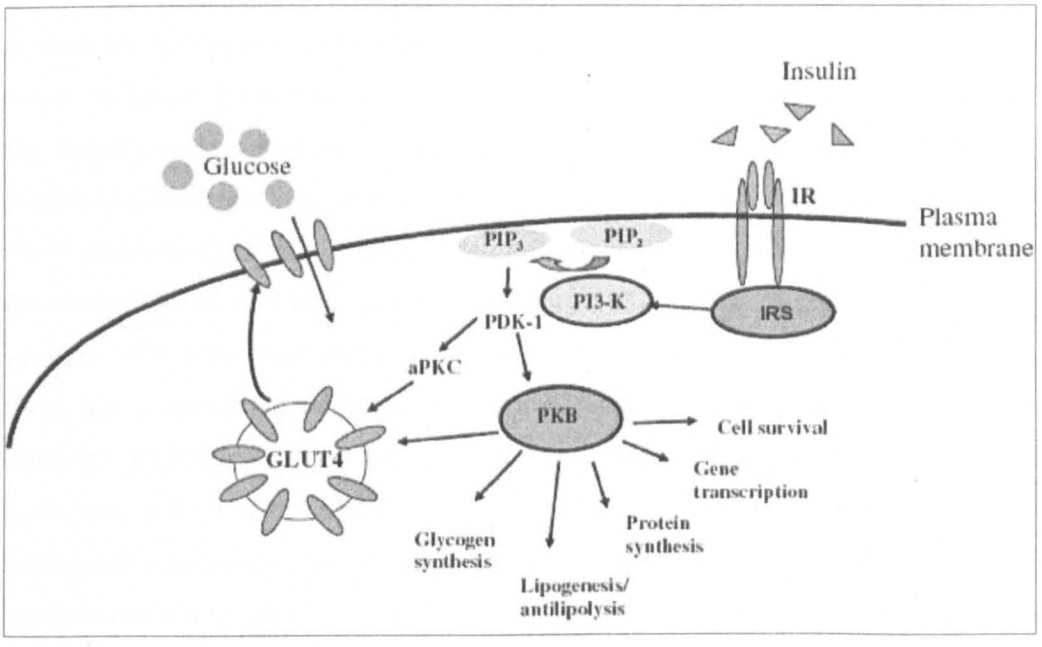


Figure 1.4. Insulin signalling via the PI3K-dependent pathway in adipocytes and skeletal muscle (Eriksson, 2007)

GLUT4 is the insulin responsive glucose transporter and is mainly responsible for insulin motivated/stimulated glucose transport processes in the muscle and adipose tissues (Herman and Kahn 2006; Furtado *et al.*, 2002; Bryant *et al.*, 2002). It is found in intracellular vesicles in the absence of insulin and relocates to the plasma membrane in the presence of insulin stimulation to increases the uptake of glucose (Dugani and Klip 2005). GLUT4 facilitated glucose transport is a rate limiting function in glucose utilization in the muscle and adipose tissues (Zisman *et al.*, 2000; Kim *et al.*, 2001). When insulin stimulation stops, GLUT4 proteins leave the plasma membrane by endocytosis to the intracellular storage compartment and hence are recycled (Dugani and Klip, 2005, Omata *et al.*, 2000; Ramm *et al* 2006; Zhao and Keating, 2007). Along with the translocation regulation of GLUT4, insulin raises the GLUT4 expression in adipocytes in the presence of glucocorticoids (Hajduch *et al.*, 1995; Hernandez *et al.*, 2003; Zhao and Keating, 2007).

1.7 Insulin Variation between Species

A study by Al-Swailem and colleagues (2008) showed that the amino acid sequence of camel proinsulin B showed 100% homology when compared with other species such as pig, dog, horse, sheep and ox. It showed 95.2% homology with man, cemacaque, rabbit, elephant, chihamster and opossum which was not considered a significant variation while mouse and rat showed a highly significant variation (66.6%). On the other hand, the A-chain showed 100% homology with ox, 95.2% with sheep, and a significant variation with man, cemacaque, rabbit, pig, dog and elephant (90.5%). The camel A-chain had a significant variation upon comparison with horse, rat 1, chihamster and opossum (85.7% homology) and a considerably high variation upon comparison with squirmonk and guinea pig (76.2% homology). The above-mentioned differences were considered relevant enough for the authors to suggest that a study into the molecular mechanisms and responses of such different insulins should be undertaken (AL-Swailem *et al.*, 2008).

In terms of amino acids, there are three amino acid changes from porcine insulin to bovine insulin at positions 8, 9 and 10 in the A-chain (Figure 1.2). The immunological reactivity may vary due to these amino acid differences. Human and porcine insulin have a change of a single amino acid at position 30 of the B-chain and it has no known effect on insulins immunological reactivity (Trenkle, 1972).

1.8 Milk Insulin

At the time of early lactation, insulin interacts with the bovine somatotropin (bST) to improve the milk yield and protein synthesis in dairy cows. These effects are most likely mediated through a combination of effects of bST on gluconeogenesis and nutrient mobilization, bST-induced insulin peripheral resistance, and bST/insulin synergism on insulin growth factor-I (IGF-I) secretion and on the metabolic activity of the mammary tissue. Insulin's ability to increase milk protein content, even with a modest increase in the insulin level observed, requires further studies to obtain the goal of manipulating the diet of animals to increase the pancreatic stimulation that leads to increased endogenous release of insulin. Such strategies to get increased insulin secretion in cows with high plasma bST concentrations are of interest to the agricultural industry (Molento *et al.*, 2002).

To obtain lipogenesis similar to those observed during early lactation the co-culture with three hormones insulin, corticosterone and prolactin is required (Strong *et al.*, 1972). The effects of these hormones on cell division and protein synthesis *in vitro* in the mid-pregnant mouse mammary gland are well documented (Topper, 1970; Mills & Topper, 1970; Shushanov, 2011). The epithelial cells are stimulated while culturing with the insulin to synthesize DNA and further division. The daughter cells are then highly sensitive to cortisol which leads to the production of enormous amounts of rough endoplamic reticulum. Only then can the epithelial cells respond to the synergistic actions of insulin and prolactin by synthesizing casein (Forsyth *et al.*, 1972).

1.8.1 Camel Milk Insulin

Wernery *et al.* (2006) reported the mean concentration of insulin in camel serum as 12.77 ± 7.62 µU/ml while the mean insulin concentration in camel milk is 41.9 ± 7.38 µU/ml. Insulin is existing in the milk of other mammalians but in lesser amount, the insulin mean concentration in human milk is 21µIU/ml (Kulski and Hartmann, 1983) while in bovine milk is 16 µIU/ml (Shehadeh *et al.*, 2001) and 17 µIU/ml in porcine milk (Westron *et al.*, 1987). The following chart (figure 1.5) shows the differences in milk insulin concentration among species.

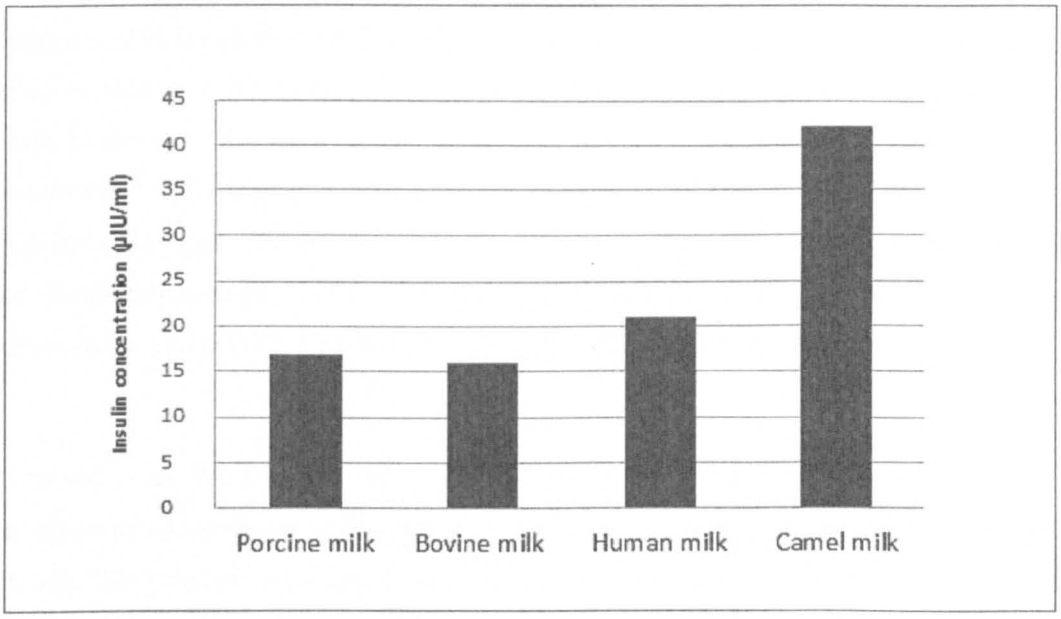


Figure 1.5 Differences in milk insulin concentration (µIU/ml) among camel, human, bovine and porcine.

Agrawal *et al.* (2004b) studied the oral hypoglycemic effect of camel milk on diabetic rats induced by streptozotocin. The study included five different groups. The first 3 groups group were offered raw camel milk, pasteurised camel milk and lactoferrin spiked camel milk, respectively, the 4th group was offered pasteurised raw cattle milk and the 5th group was the control group on a normal diet without streptozotocin treatment. This study showed that in those taking the raw camel milk the mean glucose

levels were found in the normal range of 70 to 90 mg/dl. It was further found that the blood glucose level of all the untreated groups getting raw milk showed a decrease in glucose levels, however the rate of decline was higher in the first group of animals getting camel milk when compared with the groups getting raw cattle milk. From such observations it has been suggested that camel milk can be used as a therapeutic adjunctive option for diabetes mellitus in humans. This study shows that in diabetic rats getting raw camel milk there was a significant decrease in the mean blood glucose level, however there was no apparent added advantages of adding lactoferrin in the raw camel milk. The study also importantly showed that after pasteurization hypoglycemic activity decreases.

Agrawal (2004a) observed a low frequency of impaired fasting glucose (IFG), impaired glucose tolerance (IGT) and diabetes (DM) in the Rica population consuming camel milk routinely. This study in the Raica community in North-Western Rajasthan was informative with regard to the potential beneficial effects of drinking camel milk regularly. The age-adjusted prevalence of diabetes (0.0%), IFG (3.2%), and IGT (8.6%) of camel milk consuming Raica subjects was significantly lower than those consuming other milks where the prevalence is 4.6%, 7.8%, and 20.6% respectively.

Agrawal *et al.* (2007b) also referred to the hypoglycemic activity of camel milk and assigned this property to different factors including, the high concentration of insulin or insulin like substances in camel milk, the positive immune effect on beta cell function and the absence of coagulum formation under the acidic conditions. This also suggested that the nature of the intake of food may have effect on the secreted milk composition. Camels prefer natural vegetation food such as salty plants, desert bushes and herbs like neem. Some herbs may permit secretion of some phyto-chemicals that have anti-diabetic activity.

In a further study the Soroka Teaching hospital collected colostrum and milk from camels, cows, goats, sheep and women two days after birth as part of their normal routine examinations. All samples were kept frozen until subsequent analysis. Additionally, samples of store brought milk (cows) and camel-milk ice cream were obtained and frozen. The insulin concentrations in the milk samples were then determined by specific RIA. The insulin content of the ruminating animals blood was considerably higher than those of their milks. Woman colostrum contained insulin but at levels similar to those of milk of the other animals. The lowest insulin level was found in cow milk. Store milk had the lowest insulin content of all the samples. The study of camel milk on blood glucose levels was informative. In fasted animals administered camels milk, blood glucose levels started to decline after 6 minutes ($P<0.001$) and declined 24% within 8 minutes, then began to rise again to the initial levels after 10 minutes (Zagorski *et al.*, 1998).

Wernery *et al.* (2007b) measured the insulin concentration in camel milk and colostrum over a lactation period for 310 days. The highest insulin concentration was observed in the first 48 hours of lactation with a mean of 286.5 μIU/ml then the concentration decreased to reach a mean value of 40.5 ± 10.7 μIU/ml after 160 days. Before dry off the insulin concentration increased again, suggesting different regulatory mechanism at play and the potential for different levels of insulin in the feed for the calves.

Another more recent study (Agrawal *et al.*, 2009) was carried out to determine the efficacy of camel milk in controlling diabetic nephropathy. Its nature is microvascular and it is an important complication usually found in diabetic patients. In this study the authors managed to study the role of camel milk in controlling microalbuminuria levels in type I diabetic patients. For this study they recruited 24 type I diabetes patients randomly and observed that after adding camel milk to their diet there was a considerable decrease in the rate of microalbuminuria (119.48 ± 1.68 to 22.52 ± 2.68; $p<0.001$). It was suggested that this change may be due to the direct effect of camel milk or due to the better glycemic control as a result of taking camel milk. The actual

mechanisms behind these observations however, are yet to be uncovered. Associated with this, and possibly because of the good nutrition value of camel milk, there was a slight increase in mean Body Mass Index (BMI) (18.52 ± 0.73 to 19.43 ± 0.81) after camel milk consumption, and there was a significant reduction of insulin dose required to maintain glycemic control. The mean dose of insulin/day was reduced to 28.32 ± 2.66 after addition of camel milk into patients treatment regimen, from a mean of 41.61 ± 3.08 before treatment. It was assumed from this study that these changes are the consequence of the insulin or insulin like protein content found in the camel milk. It was also found that there was a countable change in cholesterol, high-density lipoprotein (HDL), low-density lipoprotein (HDL), very-low-density lipoprotein (VLDL) and triglycerides after consuming camel milk. This observation may be due to the low fat content in the camel milk (Agrawal *et al.*, 2009). The fat content in camel milk is 2.49-3.1 g %, while it is 3.79 g % in cows milk (Surwit *et al.*, 1992).

1.9 Stability and Special Features of Camel Proteins

Camel milk proteins themselves have been reported to be unusually stable, possibly as a result of components of the camel milk itself acting as stabilisers. The effect of pasteurization, heat treatment and different storage temperatures on the concentration of insulin in camel milk have been studied by Wernery *et al.* (2006a). A minor decrease of about 7% in the insulin concentration was observed after pasteurization and storage at 4°C for 24 hours with a slightly larger decrease of approximately 13% observed after longer storage at 4°C for 4 days. A much higher reduction of approximately 26% was observed after boiling at 95°C for 5 min. However, overall these studies showed that camel milk insulin us heat stable.

The immune system of camel is different from all other mammals. This has been described by Casterman *et al.* (1993). In the camel, antibodies lack the conventional light chain and the antibody molecule consists only of the two heavy chains (Riechmann *et al.*, 1999) that reduced their average molecular weight to 95 kDa instead of 150-160 kDa for the conventional molecule from other species (Waron *et al., 1987*). The camel

antibody consists of one single antigen binding domain called the VHH, this domain created by merging the variable domains of two heavy chains (VH) (Figure 1.6). The VHH has a strong antigen-binding capacity and offers the potential of interacting to epitopes that are undifferentiated by conventional VH–VL pairs. The absence of the variable light chain is balanced by the existence of the long complementary determining region (CDR3) loop (Muyldermans *et al.,* 2001). The camel HVV site has high neutralizing activity to enzyme antigens (antigens that binds to the enzymes and inhibit their function), whilst conventional antibodies are less likely to have this property. A camel hypervariable domain is a strong and selective inhibitor to hepatitis C enzyme system (Martin *et al.*, 1997).

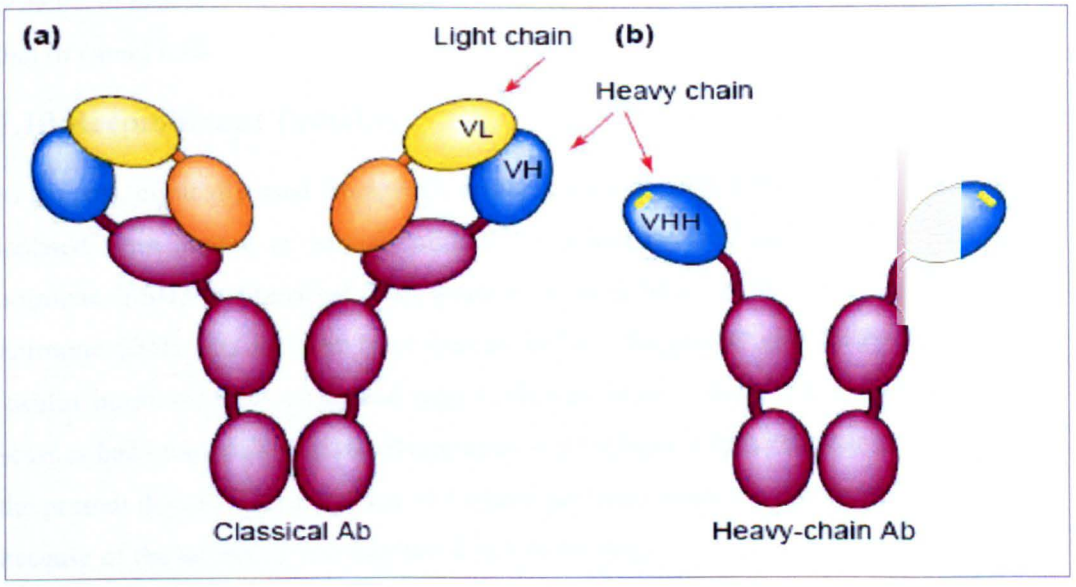


Figure 1.6. Schematic image of (a) classical antibody, (b) heavy-chain antibody of camelids. Abbreviations: Ab, antibody VH, variable domain of heavy chain; VHH, variable domain of heavy chain of heavy-chain antibody; VL, variable domain of light chain (Muyldermans *et al.,* 2001).

Camel and human antibodies have a similar antigen affinity but the camel antibodies are smaller in size and this facilities the antibody access to the target (Jassim & Naji, 2001). Figure 1.5 shows the difference between the classical and camel antibodies.

Elagamy *et al.* (2000) also found that cow and buffalo milk whey proteins are less heat resistant than camel ones. Among camel whey proteins, β-lactoglobulin is more stable than serum albumin and less stable than α-lactalbumin. The concentration of antimicrobial factors like lysozyme (LZ), lactoferrin (LF) and immunoglobulin G (IgG) in camel milk is higher than cow and buffalo milk. It has also been found that the heat resistance of lactoferrin is greater than immunoglobulin G but less than lysozyme. In general the concentration of the antimicrobial factors in camel milk are higher than cow and buffalo milk. The authors concluded that the biological activity of protective proteins in heat-treated cow and buffalo milk proteins at 75°C for 30 minutes is less than that of camel milk.

1.10 Recombinant Proteins

In the past, clinically used therapeutic proteins were purified from natural sources, being isolated from human or animal sources. For instance, the human follicle stimulating hormone (FSH) was purified from womens urine (Dickey *et al.*, 2002), human growth hormone (GH) was isolated from human bodies (Rappaport and Graham, 1987) and insulin hormone from cows and pigs (Feldman *et al.*, 1963). Treatments using such sources had many safety doubts (Rappaport and Graham, 1987; Feldman *et al.*, 1963). In the present day, the animal origin of isolated proteins is rare for biotherapeutic purposes because of the advent of recombinant DNA technology.

In the last thirty years the advancement in the biotechnology filed has resulted in the production of therapeutic proteins from recombinant genetic material using either prokaryotic (bacteria like *Escherichia coli*) or eukaryotic (yeast or mammalian cells) expression systems. The term recombinant is used in order to differentiate them from

naturally produced proteins and to indicate that they are expressed and produced by recombinant genes. The first recombinant protein permitted for human use was insulin produced in *E.coli* in the United States of America in 1984 (Walsh, 2005). Tissue plasminogen activator (tPA) was the first recombinant protein produced by transfected Chinese hamster ovary (CHO) cells as a mammalian cell host in 1986 (Wurm, 2004; Bertschinger, 2006).

Mammalian cells, particularly CHO cells, and *E.coli* cells are the most widely used expression systems for biotherapeutic proteins with almost half the therapeutic protein products available in the market being produced in mammalian cells (Wurm, 2004 and Walsh, 2006).

1.10.1 Production of recombinant protein in Bacteria:

Protein expression in bacterial cells is considered the most commonly utilized expression system for recombinant protein synthesis. The main reasons of this is because that the bacterial cells are quite easy to manipulate, low cost to culture, and the time needed to synthesis a recombinant protein using it is short and it's easy to scale up the protein production (Sahdev *et al.*, 2008). Though, as it is a prokaryotic system, the expressed eukaryotic proteins in this system are not modified properly, hard to facilitate the secretion of enormous amounts of the recombinant protein. Additionally, the largely expressed proteins can may precipitate and form the inclusion bodies (Hunt, 2005; Khow and Suntrarachun, 2012).

Escherichia coli bacterial system is the most frequently used expression system for recombinant industrial and pharmaceutical protein production and large scale production systems are established. However the E. coli may accumulate a pyrogenic lipopolysaccharide for humans and animals. Thus the recombinant therapeutic proteins should be purified twice to produce endotoxin free products (Petsch and Anspach 2000; Terpe, 2006)

1.10.2 Production of recombinant protein in yeast cells: .

Yeast combines the advantage large amount production of recombinant protein in short period of time and the ability to perform the eukaryotic posttranslational modifications, like glycosylation, disulfide crosslinking and phosphorylation. *Saccharomyces cerevisiae* is the first yeast species to be used for the recombinant protein production (Holz and Lang, 2005). The yeast expression systems are well developed and the genetics and physiology of this organism are extremely studied. The yeast expression systems were widely used to understand the function and properties of many of mammalian proteins (Romanos *et al.,* 1992). Different *Saccharomyces cerevisiae* expression systems are available to increase the protein production by eliminating the main proteases or enhance the transcription and expression level (Harashima, 1994). The recombinant proteins generated by this technology either be produced in the cytoplasm or secreted through the secretory pathway (Holz and Lang, 2005).

1.10.3 Production of recombinant protein in mammalian cells

CHO cells are the most widely used mammalian cells for the generation of biotherapeutic proteins for medical applications, however other cell lines such as the human embryo kidney (HEK293), mouse myeloma cells NS0 and the baby hamster kidney (BHK-21) cell lines have also been used for recombinant protein production (Wurm, 2004).

Although the mammalian cell cultures are more expensive and complicated than microbial cells, the mammalian cells are more desired in some cases because of their capability to perform specific post-translation modifications of proteins, mainly glycosylation and disulphide bond formation (Walsh, 2006, Walsh and Jefferis, 2006). Yeast is also able to make disulphide bonds and glycosylate proteins, however the glycosylation in this host differs from that of mammalian cells (Walsh, 2006) therefore efforts have been undertaken to "humanize" the process of glycosylation in non-

mammalian hosts (Wildt and Gerngross, 2005). Table 1.2 summaries the differences between the prokaryotic, yeast and mammalian expression systems.

Table 1.2 Differences between the prokaryotic, yeast and mammalian expression systems

Expression system	Prokaryotic cells	Yeast	Mammalian cells
Commonly used cells	*Escherichia coli*	*Saccharomyces cerevisiae*	Chinese hamster ovary cells
Handling	Straightforward	Straightforward	Requires expertise
Post-translational modifications	No	moderate	Yes
Cost	Cheap	Cheap	Expensive
References	Sahdev *et al.*, 2008	Wildt & Gerngross, 2005	Walsh & Jefferis, 2006

1.10.4 Transfection methods

The generation of recombinant proteins from mammalian cells requires the transfection of DNA, usually in a plasmid, into the target host cell. This may then be expressed transiently or stably if the material is incorporated into the cellular genome. Many different transfection methods have been established for the delivery of DNA into mammalian cells. The selection of the transfection method depends on the cell type and purpose of use. The ideal approach must have low cell toxicity, high transfection efficiency, the least effect on the cell physiology, it should be reproducible and easy to use. The three different methods used are biological, chemical, and physical mediated methods (Kim and Eberwine, 2010).

1.10.4.1 Biological Method

Viral transfection is the most frequently used method in clinical studies and is also termed transduction (Pfeifer and Verma, 2001). Viral transfection is easy to achieve, extremely competent and the expression using this method is sustained due to the viral

nature of integration into the transfected cell genome. For instance, retrovirus murine leukemia virus has been used as a viral vector for stable transgene expression in human (Hacein-Bey-Abina *et al.*, 2002; Roesler *et al.*, 2002). The major disadvantages of this transfection method are immunogenicity and cytotoxicity. The integration of the viral vector could cause an inflammatory reaction and mutation, this being due to the random integration of the viral vector into host genome which may activate oncogenes, disrupt tumour suppressor genes or interrupt vital genes (Woods, 2003). From a recombinant protein perspective, the use of viral vectors carries the potential threat of introducing viral material into the final product.

1.10.4.2 Chemical Method

Chemical transfection methods are the most extensively used methods (Schenborn and Goiffon, 2000). Chemical methods use cationic polymer like calcium phosphate, cationic lipids and cationic amino acids (Schenborn and Goiffon, 2000; Holmen, 1995; Washbourne and McAllister, 2002). The mechanism of all chemical methods is thought to be similar. The positively charged chemicals attract negatively charged nucleic acids to form nucleic acid/chemical complex. This positively charged complex is then attracted to the negatively charged cell membrane. The exact mechanism of cell membrane integration is unknown; it is thought that a endocytosis and phagocytosis processes may play a rule in this process. The transfected DNA must then be delivered to the nucleus to be expressed and once again the exact nature of this process is unknown (Luo and Saltzman, 2000; Kim and Eberwine, 2010). Many factors control the transfection rate when using chemical methods, the efficiency of this method being largely dependent on factors such as nucleic acid/chemical ratio, pH value of the media and the type of cell, hence the method results in less transfection efficiency. However, this technique has advantages of generally low cytotoxicity and no mutagenesis (Kim and Eberwine, 2010).

1.10.4.3 Physical Methods

Direct microinjection and electroporation are commonly used methods in physical transfection. The direct microinjection method is an effective technique in injecting the genetic material into cultured cells or nuclei by fine needle (Cappechi, 1980). This technique was used to produce transgenic organisms after directly transfecting the embryonic stem cells with the required genes (Bockamp *et al.*, 2002). The device used for this method is quite expensive and the technique requires expertise and it is not a method of choice for studies that need a large quantity of the transfected cells (Mellott *et al.*, 2012).

Wong and Neumann in 1982 were the first to use the electroporation technique for gene transfer into mouse cells. The mechanism of this method is based on using an electrical pulse to disturb the cell membrane and form pores to allow the nucleic acid to enter the cell (Shigekawa and Dower, 1988). The technique requires optimisation to balance between pulse duration and strength for each type of cell used. Furthermore, electroporation often requires more cells than the chemical methods because of the enormous cell death during the procedure and the extensive optimization that must be undertaken and is essential to control the transfection efficiency and cell viability. Therefore the physical transfection method is reliable but it needs expensive instruments and causes physical damage to samples (Mellott *et al.*, 2012). However, this is widely used in the generation of cell lines for the industrial production of biotherapeutic proteins.

1.10.5 Transient Transfection

The transient transfection and expression of target proteins is a useful method that allows fast investigation and generation of small amounts of the recombinant protein. The cells are transfected with the gene of interest and then cultured for a limited time to allow for the gene to be expressed (Reeves, 2006). Because the gene is not integrated into the genome, the target protein usually reaches a peak concentration 2-3 days after

transfection, then the cells and culture medium are collected to assay and purify the expressed protein. The transient expression is more often used because it is a fast and simple method (Reeves, 2006). As mentioned previously, the transiently transfected genes are not integrated with the host genome (Figure 1.6B) and are only expressed for a short period of time (Recillas-Targa, 2006). The transiently transfected genetic material can be lost due to environmental factors, digestion of the material, and during the mitosis/cell division (Kim and Eberwine, 2010).

1.10.6 Stable Transfection

The purpose of stable transfection is permanent integration of the gene of interest into the host cell genome (Figure 1.6A). The integration into the genome is a random process, so the expression is influenced by the site of integration and also the number of copies of the targeted gene that stably integrated into the genome (Reeves, 2006). The Flp-In system is a recent technology in this field produced by Invitrogen. This technology allows targeted integration of the gene of interest into the host genome by DNA recombinase. This technology is a quick process and generates reproducible and highly productive cell lines (Sauer, 1994).

To establish a stable cell line, a plasmid cloned with the target gene should be deliver to the host genome and integrated into it. The transfected cells undergo antibiotic resistance selection (the antibiotic resistance gene is carried on the plasmid), hence the untransfected cells and the inefficiently transfected cells will be killed by the antibiotic, while the antibiotic resistant cells will form colonies on the surface of a culturing flask (Kunaparaju *et al.*, 2005). CHO cells are well characterised genetically, and the stable cell lines generated with them have been extensively used for many years to synthesis many of the commercially important recombinant proteins. Bioreactors for large scale cultivation of CHO cell are also well established and hence these are the current gold-standard mammalian expression system used in industry (Reeves, 2006).

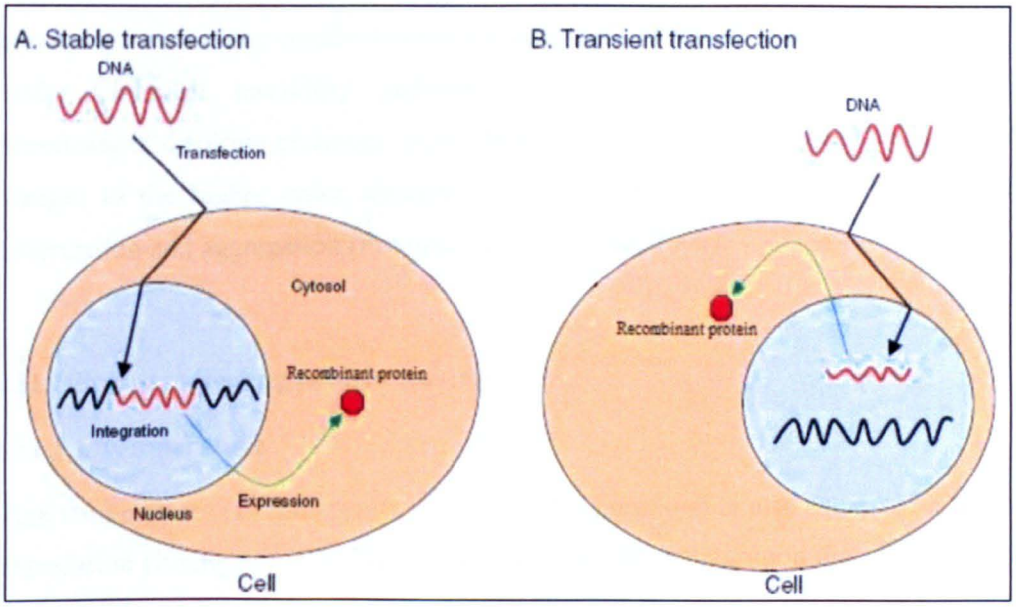


Figure 1.6. The process of two different transfection approaches. (A) Stable transfection. Foreign DNA is carried to nucleus through the cell and nuclear membranes then integrated with the host genome and becomes expressed continuously. (B) Transient transfection. Foreign DNA is delivered into the nucleus to be expressed for a short period of time without genome integration. From (Kim and Eberwine, 2010) with slight modification

1.11 Protein Stability:

Throughout preparation, processing, storage and shipping of the pharmaceutical peptide products, the physical and chemical integrity of the peptide is subjected to many circumstances which may have substantial effects on its. Therefore, many peptides should be transported and warehoused under cold conditions (Zahn *et al.*, 2006).

The unfolding and aggregation of proteins have been involved in some diseases like Alzheimer's and Creutzfeldt Jakob's diseases (Harper and Lansbury, 1997; Dobson, 2001). Thus, understanding the protein stability in solution is significant to assure proper handling for pharmaceutical biotherapeutics (Cheung and Raverkar, 2006).

The main pathways of peptide degradation are generally classified into: chemical and physical instability. Chemical instability is any process altering the peptide by formatting or cleaving covalent bond, creating new chemical objects (Manning *et al.*, 2010). Chemical instability pathways are oxidation, hydrolysis, isomerization, racemization, disulfide exchange, deamidation, and β-elimination (Li *et al.*, 1995). Any changes to the higher order structure can be considered as Physical instability like dimerization and aggregation (Frokjaer and Hovgaard, 2000).

1.11.1 Factors affecting protein stability:

1.11.1.1 Temperature

High temperatures not only result in physical denaturation but also can cause a chemical degradation (Zhang *et al.*, 1995). In some proteins the denaturation that induced by heat is reversible while the most of the protein denaturation cases are irreversible at high temperatures, the reversibility of unfolding can be strongly affected by modifying solution conditions (Mulkerrin and Wetzel, 1989). The used buffer, pH, and protein concentration also can affect the stability against the high temperature denaturation (Mulkerrin and Wetzel, 1989). The cysteine containing proteins fail to form free thiol groups when heated at 100°C (Volkin and Klibanov, 1987).

1.11.1.2 Freeze-Thawing

Some proteins are sensitive to the frequent freeze and thaw which result in extensive aggregation (Chang *et al.*, 1996; Jameel *et al.*, 1997; Papandreou and Fenouillet, 1997). It has been suggested that the frequent freeze and thaw processes is causing interfacial adsorption at ice liquid interfaces which enhance the protein aggregation (Kreilgaard *et al.*, 1998).

1.11.1.3 Agitation and Exposure to Denaturing Interfaces

The proteins in aqueous solution are exposed to agitation or shaking through the handling and shipping of the product (Henson *et al.*, 1970). The protein denaturation and also aggregation that induced by agitation is probably due to protein interactions with surfaces (air or storage vial) (Charman *et al.*, 1993; Henson *et al.*, 1970; Kreilgaard *et al.*, 1998; MacRitchie, 1986; Sluzky *et al.*, 1991).

1.11.1.4 Pressure

The high pressure can induce unfolding intermediates of the protein (Clery *et al.*, 1995; Muller *et al.*, 1981; Wong and Heremans, 1988; Yamasaki *et al.*, 1998). Within certain pressure limits, the denaturation induced by pressure induced is reversible (Zhang *et al.*, 1995). It has been suggested that the pressure denaturation is a result of the destabilization in the hydrophobic core due to the water transfer into it (Hummer *et al.*, 1998; Wolfenden and Radzicka, 1994).

1.11.1.5 The value of pH

The protein stability in aqueous solution is greatly depends on pH. It has been found that at pH value of 1-3 the triptorelin protein generally go through acid catalysed hydrolysis due the deamidation of the C terminal amide and produce free acid deamidated peptide. While in the pH of 5–6, the peptide backbones of triptorelin are hydrolyzed at the N-terminal side of the serine residue [Hoitink *et al.*, 1997 ; Strickley *et al.*, 1990). At pH values above 7, base- catalyzed epimerization is the main pathway of degradation. Serine is most likely involved in base-catalyzed epimerization through a carbanion intermediate (Hoitink *et al.*, 1997)

1.11.1.6 Deamidation of Asn and Gln residues

Deamidation is considered as common chemical degradation pathway for proteins. Peptides consisting glutamine (Gln) and asparagine (Asn) amino acids can go through a

spontaneous deamidation to form glutamic acid (Glu) and aspartic acid (Asp) amino acids under certain physiological environments (Manning *et al.*, 2010).

1.11.1.7 Auto, Photo and Metal ion catalyzed Oxidation pathways

Peptide oxidation is a chemical degradation pathway. The organic molecules oxidation is either a decrease in hydrogen content or an increase in oxygen (Solomons, 1992). The peptides that contain methionine (Met), tyrosine (Tyr), histidine (His), tryptophan (Trp) and cysteine (Cys) amino acids can be impaired due to their extreme interaction with many reactive oxygen species (ROS). Oxidation process can be prompted by light exposure and the contaminating oxidants. Metal ions can enhance the oxidation process during handling and storage. Temperature, pH and buffer composition can influence the oxidation reaction (Payne and Manning, 2009).

In 2007, Kerwin and Remmel studied the light induced degradation in biopharmaceuticals. Photo induced oxidation starts when the protein absorbs a certain light wavelength, this gives energy to raise the molecule to an excited state. The excited molecule can then move that energy to molecular oxygen and convert it to responsive singlet oxygen atoms. By this mechanism the tryptophan, histidine, and tyrosine can modify by oxidation in the presence of light (Li *et al.*, 1995, Grosvenor and Morton, 2010).

1.11.1.8 Disulfide exchange reactions

Disulfide exchange reactions contribute to the formation of dimers and larger aggregates. These reactions may occur at the cysteine–cysteine link. (Windisch *et al.*, 1997).

1.12 Aims of this Study

This project sets out to more fully investigate and define camel milk insulin and test the hypothesis that camel milk insulin is unique in terms of the makeup of the proinsulin molecule, the C-peptide and its sequence, and mature insulin, which is in turn reflected in its stability and structure.

The aims of the project are as follows:

- To identify and sequence the camel insulin gene.
- To undertake SNP analysis and confirm the sequence of the gene encoding camel insulin.
- To map the chromosomal position of camel insulin gene using FISH technique.
- To undertake bioinformatic analysis to compare the camel insulin gene and protein with other species.
- To purify camel insulin from camel milk and investigate the properties of purified camel insulin.
- To express the camel insulin protein in a Chinese hamster ovary (CHO) cell line using recombinant DNA technology.
- To investigate the biophysical properties of camel milk insulin and the protective nature of camel milk on insulin and other protein stabilities upon temperature stress and compare this to other species insulin.

The scope of the project coveres all the aspects mentioned in the proposal, in addition to this; the protein was expressed Mammalian cell line instead of the bacterial expression cells to provide the expressed protein the correct folding and to avoid any insulin isomers.

Chapter 2

Materials and Methods

2.1 Materials

All camel milk and colostrum samples were collected from the Emirates Industry for Camel Milk and Products farm. The blood samples were collected from the camel pen of the Central Veterinary Research Laboratory (CVRL). The pancreatic tissue was collected from the CVRL post-mortem from newly deceased camel carcasses after analysing and determining the reason for death and before sending them for incineration. All collection of samples at the CVRL was undertaken using authorised protocols following the required guidelines.

2.2 Camel Insulin Protein Isolation from Camel Milk

Camel milk samples were defatted and treated with rennet enzyme in order to remove casein which is the major milk protein and problematic in subsequent purification schemes if left in place when trying to purify less abundant proteins. The purification procedure developed for insulin purification involved 4 steps of which the first was defatting, the second casein treatment, the third ion exchange chromatography and finally an immunoaffinity chromatography step. Each of these is described in more detail below.

2.2.1 Defatting of Camel Milk Samples

Camel milk samples were defatted by centrifugation at 4000 g for 10 min at 10°C. A commercial cream separator (Motor Sich-80, Latvia) was used for large scale defatting.

2.2.2 Casein Treatment of Milk and Colostrum

Three different methods were then used to remove the casein from the defatted milk, casein being the major milk protein component. The first method used was acid precipitation. In this method the whey was obtained after acid precipitation of caseins at pH 4.6 by adding 1 ml of 10% acetic acid (v/v) and 1 ml of 1 M sodium acetate to 10 ml of milk or colostrum followed by incubation at 37°C for 10 min and centrifugation at 4000 × g for 10 min at 10°C (El-Hatimi *et al.*, 2007).

The second method used was rennet enzyme treatment. This method is undertaken using rennet enzyme tablets (Hundsbichler, Austria), these tablets being used commercially for cheese making. Rennet tablets (25 mg) were dissolved in 0.1 ml of water then mixed well with one litre of defatted milk before the samples were incubated at 37°C for 1 hour. The whey was then separated by centrifugation at 4000 g for 5 min.

The third method used was the ethanol precipitation method. Ethanol (50% final concentration) was used to precipitate the casein and other higher molecular weight proteins. The whey fraction was isolated by filtration using Whatman filter papers 4 (Cat. No. 1004-240). The insulin concentration in the whey fraction was measured and determined by radioimmunoassay (RIA) as described below.

2.2.3 Radioimmunoassay (RIA) for the Determination of Insulin Concentrations

Radioimmunoassay (RIA) was undertaken throughout this study to determine the concentration of insulin present using a commercially available human insulin kit. During the purification studies all milk fractions were analysed using the human anti-

insulin antibody from the human insulin RIA kit (INS-Irma, KIP1251-KIP 2154, Diasource Europe, Belgium) following the manufacturers instructions. Six insulin controls with a range of activity from 0 to 440 µIU/ml and the samples were briefly vortexed and 50 µl of each sample and control were transferred to the respective antibody coated tubes. Fifty microliters of the tracer (^{125}Iodine labelled anti-insulin) were dispensed into each coated tube. To release any trapped air bubbles, the rack containing the tubes was shaken gently by hand. After 2 hours incubation at room temperature, the total content of each tube was aspirated. The tubes were then washed twice with 2 ml of Working Wash solution (supplied with the kit). The content of each tube was aspirated after each wash. The radioactivity was then counted for 60 sec using a gamma counter LKB wizard 1470 (LKB, Sweden). The standard curve was prepared based on the control results and the concentration of each sample calculated using the generated standard curve. Samples with an insulin concentration above the highest standard were diluted with the zero concentration control and reanalysed.

2.2.4 Western Blotting

The sensitivity of Western blotting to detect human insulin was determined by running serial dilutions of human insulin on a 12% SDS-PAGE gel from 10 IU/mL-10 µIU/mL. The primary antibody used was the commercially available anti-human insulin antibody (Diasource Cat. 51.125.26, Belgium). Ten µL of protein were added to 10 µl of loading buffer (Laemmli buffer, Bio-Rad, USA) and incubated at 70°C for 10 minutes. After incubation, the samples and a protein pre-stained standard marker (SeeBlue Plus2, Invitrogen, USA) were loaded into a 12% NuPAGE gel (Invitrogen) in NuPageMES running buffer (Invitrogen) for protein separation by SDS-PAGE in a XCell SureLock™ mini-cell apparatus (Invitrogen). The gel was run for 75 minutes at 150 V and then blotted onto 0.2 µM polyvinylidene difluoride (PVDF) membrane (Invitrogen) by electro-blotting in an XCell II™ Blot Module (Invitrogen). After 1 hour of blotting at 30 V the membrane was washed in phosphate buffer saline (PBS). The membrane was then incubated in blocking solution (0.1% Tween and 5% non-fat dry milk in PBS) for 1 h

and then incubated overnight at 4°C with the anti-insulin antibody which had been diluted at a ratio of 1:500 in PBS buffer containing 0.1% Tween and 0.1% non-fat dry milk. The following day the membrane was washed 3 times with PBST (0.1% Tween 20 in PBS) buffer for 10 minutes each, and then incubated in secondary antibody (Horse raddish peroxidase conjugated rabbit-anti-mouse IgG, Dako) diluted 1:1000 for 1 hour, then washed 4 times with PBST for 10 minutes each. The membrane was then developed with 2 ml of immobilon Western chemiluminescent HRP substrate (Millipore, USA) and incubated in the dark for 4 minutes. Finally, the immunocomplexes were exposed to lumi-film chemiluminescent detection film (Roche). The films where developed in a compact X4 system (Xograph imaging systems).

2.2.5 Gel Filtration Chromatography

Gel filtration chromatography was undertaken using G-25 Sephadex beads (GE healthcare, USA) using decaseinised colostrum samples where 2.5 ml of sample was passed manually through a 5 ml column. The column was washed with 12 ml of 0.1 M NaCl in 0.02 mM Tris buffer and fractions collected every 1 ml.

2.2.6 Ion Exchange Chromatography

Ion exchange chromatography was performed using A-25 DEAE Sephadex beads (GE healthcare, USA). Several trails were initially undertaken in order to maximise the recovery of insulin. Initially the purification was undertaken at a small scale with a small amount (as detailed below) of milk and colostrum in a 5 ml column, and then the purification expanded to process 100 litres of milk using a 1 litre chromatography column. The trails were carried out in order to optimise the washing and elution conditions. Rennet enzyme treated samples, supernatant 50% ethanol precipitated samples and untreated defatted raw milk and colostrum samples were all investigated to find the best feedstock to load the column for maximum recovery of insulin.

In all, six different ion exchange chromatography trails were performed. In all ion exchange chromatography runs the column was equilibrated with 3 column volumes of 50 mM NaCl in 0.02M Tris buffer and all washing and elution buffers were prepared in 0.02 M Tris buffer. Fractions of 100 µl were collected throughout and stored at -20°C for subsequent insulin quantitation by RIA. In one trail the LP biologic chromatography automated system (BioRad, USA) was also used.

In the first trail, diluted decaseinised colostrum was used where 3 ml of decaseinised colostrum was diluted with 11 ml of equilibration buffer and loaded onto a 5 ml column of DEAE A-25 Sephadex beads. The column was then washed with 6 ml of 10 mM NaCl washing buffer and eluted with three different concentrations of elution buffer (10 ml of 50 mM, 10 ml of 250 mM and 12 ml of 1 M NaCl respectively).

In the second trail 11 ml of decaseinised colostrum sample was used without dilution in the same column as used in the first trail, the column washed with 15 ml of 100 mM NaCl washing buffer and eluted with two different concentrations of elution buffer (20 ml of 500 mM and 20 ml of 1 M NaCl respectively) followed by a 20 ml wash with 1 M acetic acid. In third trail, 45 ml of defatted raw milk was used in the same column as used in the first and second trails, the column washed with 20 ml of 50 mM NaCl washing buffer and eluted with three different concentrations of elution buffer (250 mM, 500 mM and of 1 M NaCl respectively) using 20 ml of each for elution.

The fourth trail was undertaken using 250 ml of defatted decaseinised raw whole milk using a 5 ml column of A-25 DEAE Sephadex. The sample was loaded in 5 batches with 5 washes of 15 ml of 50 mM washing buffer (1 after loading each batch). After loading the entire amount of sample, the column was washed with 30 ml of 250 mM NaCl washing buffer and then eluted with 70 ml of 1 M NaCl.

The fifth trail was undertaken on an ethanol treated defatted raw milk sample in a 5 mL column of A-25 DEAE Sephadex beads. To obtain a 50% ethanol final concentration, 100 ml of absolute ethanol was added gradually to 100 ml of defatted raw milk while stirring. The sample was then incubated at 4°C overnight. The clear supernatant layer of a volume of 100 ml was separated and loaded onto the pre-equilibrated column. The column was then washed with 35 ml of 50 mM NaCl washing buffer and then eluted with 25 ml of 250 mM NaCl elution buffer followed by 65 ml of 1 M NaCl buffer.

The sixth and final trail was also undertaken on an ethanol treated defatted raw milk sample using the biologic LP automated chromatography system (BioRad, USA). For ethanol precipitation the sample was treated as described in the fifth trail. After attaching a 20 ml column of DEAE A-25 Sephadex beads to the chromatography system, the column was equilibrated with 60 ml of equilibration buffer. The sample was then loaded using a flow rate of 2 ml/min. The column was then washed with 80 ml of 80 mM NaCl washing buffer, then eluted with 20 ml of 1 M NaCl elution buffer. Two peaks were observed on the LP data viewer (A280 nm) during the washing step and 3 strong peaks during the elution. The fraction(s) corresponding to each peak were collected and analysed to determine their insulin concentration by RIA.

2.2.7 Immunoaffinity Chromatography Purification of Camel Milk Insulin

The immunoaffinity chromatography column was designed using anti-human insulin antibodies, specifically a monoclonal antibody IgG anti-human recombinant insulin (Diasource Cat. 51.125.26, Belgium) linked to Sepharose 4 beads. The initial immunoaffinity chromatography trail was undertaken on 22 ml of decasienised colostrum which had been treated with rennet enzyme to remove the casein. A three mL column was equilibrated with 12 ml of 50 mM glycine pH 8.0; 0.1% $NaNO_3$ (w/v) and the sample loaded onto the column at a flow rate of approx. 30 ml/hour at 4°C. The column was then washed 10 times with 4 ml of 1.25 M NaCl in 50 mM glycine pH 8.0

before the insulin was eluted with a buffer consisting of 50 mM glycine and 1.5 M NaCl at pH 2.3. For the second and third immunoaffinity purifications, the NaCl concentration was reduced to 0.5 M in the elution buffer. The eluted product was collected in 10 fractions of 2 ml each. The pH was adjusted immediately after every elution with 60 µl of 1 M Tris-buffer for every 2 ml sample.

2.2.8 Purification of Camel Milk Insulin from large quantity of camel milk

The first purification was undertaken with 100 L of camel milk. The milk was defatted by centrifugation in 50 ml tubes at 4000 g for 10 min and the samples then loaded onto a pre-equilibrated 1 L column of DEAE A-25 Sephadex beads. After 11 hours of continuous sample loading the column was washed with 3 L of 0.2 M NaCl and then eluted with 6 L of 1 M NaCl in 3 batches.

In the second purification the supernatant of the 50% ethanol treated sample was used. After defatting of the milk using the cream separator (Motor Sich-80, Latvia), chilled absolute ethanol was added gradually to the sample in a 1:1 ratio with stirring in a 4°C room. After adding 60 L of ethanol to 60 L of defatted milk (two hours), the sample was then incubated overnight at 4°C. The following day the sample was separated into two layers. The transparent upper layer was separated carefully using a pump and then filtered through a 1 µm filter. The sample was then loaded onto an equilibrated 1 L column of DEAE A-25 Sephadex beads over a 6 hour period. The column was then washed with 2 L of 80 mM NaCl buffer and then eluted in 12 fractions where the first 8 fractions were eluted with 700 mM NaCl elution buffer and 1 M NaCl elution buffer used to elute the last 4 fractions.

In the third purification, 45 L of defatted milk was used. The milk was defatted using the cream separator machine (Motor Sich-80, Latvia). The sample was loaded in 6 batches where 7.5 L of the defatted milk was loaded in each batch. After each batch the column

was washed with 2 L of 50 mM NaCl buffer and then eluted with 1 L of 200 mM NaCl followed by 2.5 L of 500 mM NaCl in 5 fractions. The insulin concentration in all fractions was analysed and determined by RIA.

Chromatography fractions with the highest insulin concentrations as determined by RIA were pooled and together the volume was 5.5 L. The sample was then filtered through a 0.2 µm filter unit and then concentrated 38 times using a Pellicon XL 50 ultrafiltration cassette with a 5 kDa. molecular weight cut off (Cat. No. PXB005A50, Millipore, USA). The sample concentration process was undertaken in a 4°C room for 2 days and resulted in 145 ml of concentrated sample.

The concentrated sample (145 ml) obtained from the 2^{nd} and 3^{rd} purifications was divided into two portions (80 and 65 ml) and passed through the immunoaffinity chromatography column separately. The conditions were as described in the trail experiments except for a modification in the salt concentration of the elution buffer which was decreased from 1.5 M to 0.5 M NaCl.

The two purification schemes for the isolation of camel milk insulin are described below in Figure 2.1.

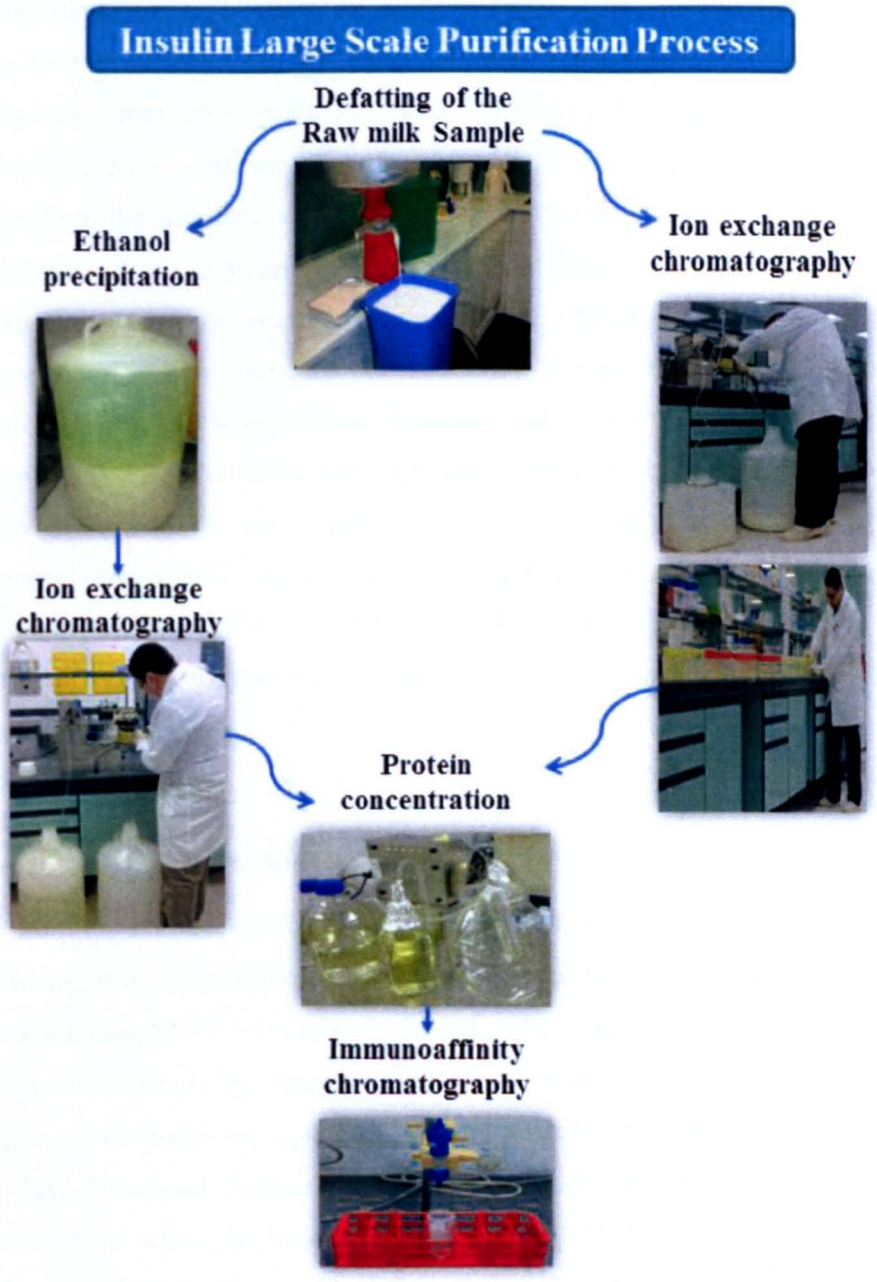


Figure 2.1 The steps of camel milk insulin purification from raw camel milk.

2.3 Analysing the Isolated Camel Milk Protein

Western blotting was undertaken on purified camel milk insulin according to the procedure described in section 2.2.4. The 3rd, 4th and 5th immunoaffinity chromatography elution fractions from the first portion (80 ml concentrated sample) were analysed by western blotting. As a monomer insulin band couldn't be visualized by western blotting, and to confirm the presence and the size of the camel insulin, 50 µL of the 3rd immunoaffinity chromatography elution fractions from the first portion (80 ml concentrated sample) were initially analysed by SDS-PAGE. The resulting gel was cut into 13 horizontal sections and each gel section was cut separately into tiny pieces and then incubated overnight with shaking in passive diffusion buffer consisting of 20 mM TrisHCl (pH 7.5), 150 mM NaCl, 0.1% SDS and 2 mM EDTA. Following day the liquid was separated from the gel pieces and the insulin concentration determined by RIA. To recover more insulin from gel pieces, a second technique was performed using the protocol described by Steven *et al.* (2007) where 100 µl of buffer containing H_2O/2-propanol (2:1 v/v) was used to recover protein from 11 horizontal sections of the SDS-PAGE gel.

2.3.1 In-gel Digestion of the Isolated Camel Insulin

One mL from the immunoaffinity chromatography where the insulin concentration was 90000 µIU/ml, was divided into two tubes and concentrated in a speedvac to a volume of 10 µl. Each sample was then mixed with the same volume of Laemmli sample buffer (Bio-Rad Laboratories). The samples were then run on a 12% Tris base (Invitrogen) precast gel and electrophoresed at 150 V. After protein separation, the gel was stained with Bio-Rad Coomassie Brilliant Blue Staining solution for 1 h with constant shaking. The bottom of gel where the insulin band was expected was horizontally sectioned into four sections. The sections were labelled from A to D starting from bottom as in figure 2.2. Each gel section was then placed in a separate 1.5 ml tube.

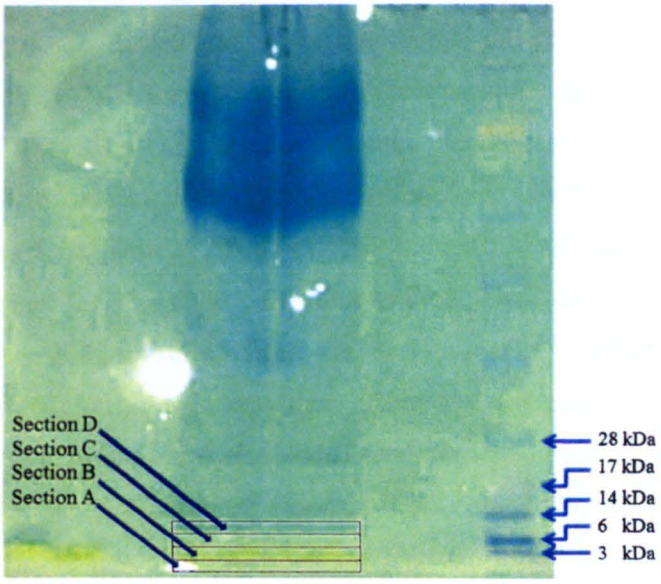


Figure 2.2. Location of the gel sections used for in-gel digestion after SDS-PAGE separation. The prestained standard SeeBlue Plus2 marker (Invitrogen, USA) is on the extreme right side of the image.

The gel pieces were destained with acetonitrile/200 mM ammonium bicarbonate (1:1 ratio) for 15 min at room temperature before the destaining solution was removed and the gel slices washed with fresh destaining solution by pipetting up and down. The solution was then discarded and the gel pieces dried in a SpeedVac for 30 min. The gel sections were then rehydrated with 30 µL of 10 mM dithiothreitol (DTT) dissolved in 5 mM ammonium bicarbonate and the protein reduced for 45 min at 56°C. The DTT was then carefully removed from the sample and discarded. Thirty µL of 50 mM iodoacetamide in 5 mM ammonium bicarbonate was then added and the protein was alkylated for 45 min at room temperature in the dark. The iodoacetamide was then carefully removed from the sample and discarded. The gel pieces were then washed with 5 mM ammonium bicarbonate for 15 min with gentle agitation and the solution removed. One wash with 50% acetonitrile was then performed with gentle agitation and the solution replaced with 30 µL of acetonitrile/25 mM ammonium bicarbonate (1:1

ratio) and the gel sections dried in a SpeedVac. Forty µL of sequencing grade trypsin solution (Promega) was then added to the sample and the gel pieces were allowed to rehydrate at 4°C for 1 h. Excess trypsin solution was carefully removed from the sample and discarded. Thirty µL of 50 mM ammonium bicarbonate was then added to the sample and incubated overnight at 37°C. On the following day the solution from the gel tubes was removed and stored. 30 µL of 50% acetonitrile was then added to the gel pieces and sonicated for 15 min at room temperature. The solution was then aspirated and added to the solution stored from the previous step. Another 30 µL of 50% acetonitrile with 5% formic acid was then added and sonicated for 15 min at room temperature and the solution was also added to the isolated mixture from the previous two steps. Finally, the pool solution from the last 3 steps was completely dried in a SpeedVac. The dried peptides were then reconstituted with 20 µl of 0.1% trifluroacetic acid and analysed by liquid chromatography-mass spectrometry. The mass spectrometry analysis was conducted using a microTOF-Q machine (Bruker Daltonics) using the electrospray ionization mass spectrometry technique.

2.4 PCR Amplification and Sequencing of the Camel Insulin Gene

2.4.1 DNA Extraction

DNA was extracted from camel blood and pancreas tissue using the commercially available MagNA pure LC system (Roche, Germany) using the MagNA pure LC DNA large volume isolation kit (Roche Germany) according to the manufacturer's instructions.

2.4.2 PCR Amplification of the Camel Insulin Gene

The camel insulin gene was amplified using Cam-INS-F2 (5`-TTT GTG AAC CAA CAC CTG TGC GGC TC-3`) and Cam-INS-R2 (5`-CGT CTA GTT GCA GTA GTT CTC CAG CTG-3`) primers, which were optimized to amplify the camel proinsulin

cDNA (AL-Swailem *et al.*, 2008). The PCR reaction volume was 25 µL and consisted of 400 µM dNTPs, 500 nM of appropriate primers, 2 units of Fast start Taq DNA polymerase (Roche, Germany) and 2.5 µL of 10X PCR buffer. A Tetrad thermal cycler from BioRad was used to run all PCR reactions. The cycling conditions were as follows: 40 cycles of denaturing at 94°C for 30 s, annealing at 64°C for 45 s and polymerization at 72°C for 45 s, followed by a final extension at 72°C for 5 min. The PCR products were analysed on a 1.5% agarose gel and the presence of the expected 768 bp band confirmed.

2.4.3 DNA Sequencing of the Camel Insulin Gene

Each 2 µL of PCR product was treated with 0.1 U of exonuclease I and 1 U of shrimp alkaline phosphatase (both from USB, USA) before sequencing the amplified fragment. The amplified fragment was sequenced using the BigDye terminator v3.1 kit (Applied Biosystems, USA) according to the manufacturer's recommendations. Three µL of the treated PCR product from the last step were mixed with 0.3 µL primer and 2 µl BigDye terminator and the reaction volume made up to 5 µl with d.d.H_2O. The samples were cycled in the thermal cycle as follows; 1 cycle at 96°C for 1 min and 25 cycles of 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min.

To clean the sequencing reaction products from the excess BigDye terminator and primers, 1.25 µl of 125 mM EDTA was mixed with sequencing reaction product and 25 µl of cold absolute ethanol was added to each sample and centrifuged at 1650 X g for 30 min, and then the ethanol was removed by quickly spinning the sequencing plate upside down at 185 X g force. The washing step was repeated with 70% ethanol for 15 min followed by centrifugation. Finally, the plate was dried and the sequencing reaction product of each sample was reconstituted in 10 µl of Hi-di Formamide (Applied Biosystems, USA) and loaded into the ABI 3730XL sequencing machine.

The sequencing results were analysed using DNA sequencing analysis software version 5 (Applied Biosystems, USA). Using a BLAST search the sequenced fragment was matched to a BAC clone sequence from *Llama pacos* CH246-515B17 clone with 96% homology. Primers F1 (5'- CAC TGG TCC TTG AGC CTA GC -3') and R2 (5'- CAG GAA AGG TGC AGA TTG GT-3') to amplify the upstream and downstream regions of the insulin gene were then designed using the matched clone, the primers being designed using the primer 3 online software tool (http://frodo.wi.mit.edu/primer3/). The fragment of 1960 bp was amplified using the designed primers in a 25 µL PCR reaction volume, 500 µM dNTPs, 500 nM of appropriate primers, 3 units of Fast start Taq DNA polymerase (Roche, Germany) and 2.5 µL of 10X PCR buffer and with cycling conditions 40 cycles of denaturing at 94°C for 45 s, annealing at 62°C for 45 s and polymerization at 72°C for 2 min's, followed by a final extension at 72°C for 10 min. The PCR products were then analysed on a 1.5% agarose gel. The amplified fragment was then sequenced using the protocol described above. To sequence the entire fragment internal primes R3 (5'-GGC CTT GGG TGT GTA GAA GA-3') and F867 (5'-CTT CGC TAA CCA GCA CCT GT-3') were designed using *Llama pacos* CH246-515B17 clone and the primer3 online software tool (http://frodo.wi.mit.edu/primer3/).

2.5 Identifying the Insulin Gene Regions and Single Nucleotide Polymorphisms (SNP's) of Camel Insulin

The insulin gene regions (signal peptide, A- and B-chains and C-peptide) were identified based on the previous publication of the insulin gene of other species (Bajaj 1986, Brown 1955, Kwok 1983) and by Basic Local Alignment Search BLAST comparison with published sequences. In order to identify the single nucleotide polymorphisms (SNP's) in the camel insulin gene, thirty two different camel samples were sequenced and the resulting sequences then aligned by DNAMAN software version 5.2.2. The SNP's were then identified from the insulin sequence differences among the 32 camels.

2.6 Taxonomic Alignment and 3D Homology Modelling

The complete DNA sequence was analysed with Basic Local Alignment Search Tool BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi). DNAMAN software (Lynnon corporation, Canada) was used to convert the DNA sequence to an amino acid sequence and subsequently align the amino acid sequence across 13 different species. This software was also used to construct a homology tree between the Dromedary camel and the 13 different species compared. The 3-dimentional structure of the camel insulin protein was generated using 3D-JIGSAW protein Comparative Modelling Server (http://www.bmm.icnet.uk/~3djigsaw/). The resulting model was viewed with RasMol software version 2.6 and PyMOL software version1.3.

2.7 Camel Insulin Gene Localization with Chromosomes

The insulin gene localization was undertaken using Fluorescence *in situ* hybridization (FISH) on cultured lymphocyte metaphases. Two methods were used to localize the camel insulin gene on camel chromosomes. The first method used was by hybridizing the amplified PCR product of the camel insulin gene on the camel chromosomes. The second method used was by hybridizing a BAC clone that included the insulin gene. The clone was selected by BLAST search and the presence of the insulin gene was confirmed by PCR.

2.7.1 Cell Culture

Peripheral blood was collected from 3 camels. 8 ml of sterile complete growth medium (RPMI-1640) and 1.5 ml of inactivated fetal calf serum were then put into a 25 cm^2 sterile culture flask and then 0.1 ml of pokeweed mitogen (PWM) added to the flask. 0.5 ml of blood was added to the flask, mixed and incubated in 5% CO_2 and 80% humidity for 96 hours. 70 µl of colcemid (10 µg/ml) was then added to the culture 1 h before harvesting. The cultured cells were then mixed vigorously and transferred to a tube and

centrifuged at 200 g for 10 min. The supernatant was then removed and the cells resuspended in 7 ml of 0.075 M KCl and incubated at 37°C for 30 min before the cells were fixed with 3:1 methanol:glacial acetic acid 3 times for 30 min for each fix.

2.7.2 Probe Labelling

2.7.2.1 PCR product labelling

Two labelling procedures were adapted to label a 1960 base pairs PCR fragment that includes the insulin gene. The fragment was amplified using the primers F1 (5'- CAC TGG TCC TTG AGC CTA GC -3') and R2 (5'- CAG GAA AGG TGC AGA TTG GT-3'). The first method is the nick translation labelling method using Biotin labelled 16-dUTP by Biotin-Nick Translation mix (Cat. No. 11745824910, Roche, Germany), where 1 µg of the PCR product was made up in a 16 µl volume of d.d. H_2O and mixed with 4 µl of Biotin neck translation mix. The mixture was then incubated at 15°C for 90 min. The labelling reaction was terminated by heating at 65°C for 10 min and by adding 1 µl of 0.5 M EDTA (pH 8). The second method was the PCR labelling procedure using the PCR DIG Probe Synthesis Kit (Cat. No. 11636090910, Roche, Germany). In this method the primers F1 and R2 were used to amplify and label the fragment of camel insulin and the surrounding regions with digoxigenin (DIG) labelled dUTP. The labelling PCR reaction was prepared and undertaken according to the manufacturers instructions.

2.7.2.2 BAC clone labelling

BAC clone number CH246-515B17 from *Llama pacos* was selected by BLAST search using the high throughput genomic sequencing (HTGS) database. In order to subculture the BAC clone, a sterile tip was dipped into the original BAC broth culture and transferred to 30 ml of LB broth (10 g Bacto-tryptone, 5 g yeast extract and 10 g NaCl to 1 L H_2O) containing chloramphenicol antibiotic at 20 µg/ml concentration and then

incubated in 37°C shaking incubator overnight at 225-250 rpm. On the following day, the cultured tube was centrifuged at 4000 rpm for 10 min to pellet the cells and the supernatant discarded. BAC DNA was isolated using FlexiPrep kit Cat. No 27928101 (Amersham bioscinces, USA). The pellet was resuspended in 4 ml Buffer P1. To lyse the cells 4 ml of Buffer P2 was added, the tube inverted 6 times and then incubated for 5 min at room temperature. The mixture was then neutralized with 4 ml of P3 buffer, inverted 6 times, and incubated on ice for 10 min. To precipitate the bacterial genomic DNA the tubes were centrifuged at 12,000 g for 15 min at room temperature and the supernatant carefully isolated and transferred to a 15 ml tube. The BAC DNA was then precipitated with 8.4 ml of isopropanol, which was added to the isolated supernatant DNA and then centrifuged at full speed for 30 min. The supernatant was then removed and the pellet reconstituted with 100 µl of H_2O. The presence of the insulin gene in the clone was confirmed by PCR using Cam-INS-F2 and Cam-INS-R2 and the random priming method. The BioPrime Array CGH kit catalogue number 18095-011 (Invitrogen, USA) was used to label the BAC DNA indirectly using biotin labelled dUTP. 300 ng of BAC DNA was used for the labelling procedure to prepare a 15 µl reaction. The random primers and BAC DNA were denatured at 95°C for 5 min and then immediately chilled on ice for 5 min, the dntp's mix containing biotin labelled dUTP and klenow enzyme were then added according to manufacturer recommendation and the volume adjusted to 15 µl by DEPC-treated water. The labelled product was then run on a 2% agarose gel to check the labelling efficiency.

2.7.3 Fluorescence in situ Hybridization (FISH)

2.7.3.1 FISH with the PCR product of the insulin gene

The FISH technique was used to localize the insulin gene on camel chromosomes. Slides were air-dried, then dehydrated with serial alcohol (70%, 90% and 100% ethanol) using 5 min for each. A 3 µl probe cocktail containing indirectly labelled DNA probes and 9

µl of hybridization buffer (Chrombios, Germany) was then overlaid on the slides. A coverslip was placed on each slide and sealed with rubber cement. Slides were then placed on an 85°C hot plate for 6 min and incubated overnight at 37°C in a humidified chamber. Following this, the slides were washed 2 times in 2X SSC solution containing 50% formamide (Q-Biogene, USA) for 5 min at 45°C, soaked in 2X SSC for 5 min at 45°C followed by 2 washes with 4X SSC containing 0.1% Tween 20 for 5 min at room temperature. After blocking by soaking in 4X SSC solution containing 5% non-fat dry milk at room temperature for 5 min, slides were incubated with 100 µl streptavidin-FITC and anti-DIG TRITC cocktail, the dye selected based on the labelling procedure used, then incubated at 37°C in a humidified chamber for 30 min. The slides were then washed 3 times with 4X SSC solution containing 0.1% Tween 20 at room temperature for 5 min each time. After the slides were air-dried, antifade with DAPI (Vectashield, Germany) was applied to the slides. The slides were then visualized using an Olympus BX61 fluorescence microscope (Olympus, Japan). Images were captured using Applied Imaging Cytovision 3.1 software (Applied Imaging, UK).

2.7.3.2 FISH using BAC clone contains the insulin gene

A drop of 12 µl of the fixed cell suspension was spread on a slide exposed to hot steam. The slide was then exposed again to the hot steam for 4 sec and then placed immediately on a heating block at 65°C and the slides dehydrated in serial alcohol (70%, 85% and 96% ethanol) and then aged at 65°C for 1 hour, then overnight at room temperature. The hybridization buffer was prepared and consisted of 5 ml formamide, 1 g dextran sulfate, 1 ml 20xSSC before the volume was adjusted to 7 ml with ddH_2O and pH to 7.0. For probe preparation, 1 µl of the labelled probe was mixed with 7 µl of hybridization buffer, 1 µl salmon sperm DNA and 1 µl ddH_2O. The slides were denatured in denaturation solution which contains (70% formamide / 2xSSC) and then immediately chilled in an ethanol series 70%, 85% and 96% at -20°C. The probe was denatured at 73°C for 10 min and then chilled on ice immediately. 5 µl of the prepared probe was then overlaid on the slides. A cover slip was placed on each slide, and sealed with rubber

cement and allowed to hybridize overnight at 37°C in a moist hybridization chamber. The slides were then soaked in 2XSSC to drop off the cover slips. Following this, the slides were washed with 0.4x SSC/0.3% Igepal, pH 7.0 - 7.2 for 2 min at 73°C, then soaked into storage buffer (4x SSC / 0.05% Igepal) for 15 min at room temperature. After blocking by soaking in storage buffer containing 2.5% bovine serum albumin at room temperature for 25 min, 90 µl of the detection mix (prepared from 50 µl of blocking buffer, 50 µl storage buffer and 0.5 µl streptavidin-Cy3) was added to each slide, covered with a cover slip and incubated at 37°C in a dark humidified chamber for 35 min. The cover slips were then removed and the slides washed 3 times in storage buffer for 3 min each wash with agitation. After rinsing with ddH_2O the slides were air-dried and then antifade DAPI solution (Vectashield, Germany) was applied. The slides were then visualized using an Olympus BX61 fluorescence microscope (Olympus, Japan). Images were captured using Applied Imaging Cytovision 3.1 software (Applied Imaging, UK). To identify the chromosomes, 1 mg of propidium iodide (Sigma-aldrich) was dissolved in 1 ml of phosphate buffered saline (PBS) and 2 µl from the propidium iodide solution were then added to 1 ml DAPI. Finally, 50 µl of the propidium iodide/DAPI solution were dropped onto each slide to produce a banding pattern similar to the GTG banding.

2.8 Synthesis of Recombinant Camel Insulin

2.8.1 Recombinant Insulin Synthesis from Mammalian Cell Lines

To produce camel insulin from a mammalian cell line, recombinant clones using the commercially available pcDNA3.1V5 and pcDNA5 FRT vectors (Invitrogen) were prepared and then transfected into a Chinese hamster ovary (CHO) cell line. Initially separate clones coding for the A-chain and B-chain of insulin were generated using synthetic fragments containing stop codons and phosphorylated sticky ends. The fragments were cloned into the NotI and XhoI sites of each vector. In addition, the pre-

pro-insulin and pro-insulin genes were amplified by PCR using primers with NotI and XhoI restriction sites and subcloned into the same sites of the pcDNA3.1 and pcDNA5 FRT vectors. Due to low recombination efficiency using this pair of sites, the subcloning was repeated using HindIII (which is further upstream in the multiple cloning site (MCS), see Figures 2.3 and 2.4) instead of NotI along with the XhoI restriction site. Following this the pcDNA3.1V5 constructs were used to transiently transfect the appropriate Chinese Hamster Ovary (CHO) cell line as described further below. It was also attempted to produce CHO cell lines stably expressing camel insulin using the commercially available Flp-In CHO cell system (Invitrogen) by transfection of the Flp-In cells with pcDNA5 FRT constructs using the Flp recombinase mediated integration system. All transfection procedures were performed using Lipofectamine 2000 transfection reagent (Invitrogen).

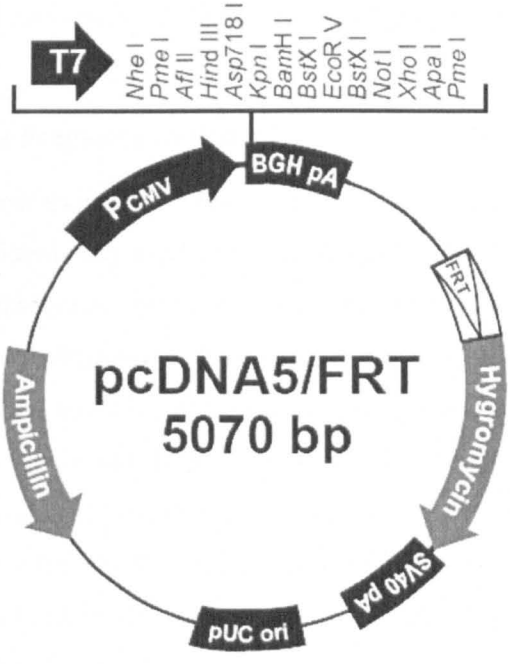


Figure 2.3: The map of pcDNA5 FRT vector adapted from the Invitrogen website, http://tools.invitrogen.com/content/sfs/vectors/pcdna5frt_map.pdf.

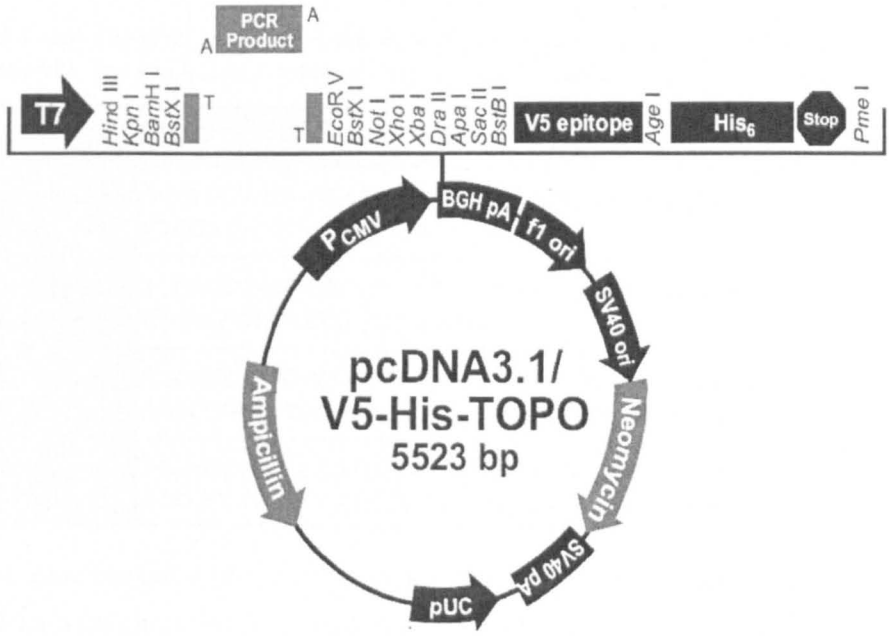


Figure 2.4 : The map of pcDNA3.1 V5His vector adapted from the Invitrogen website, http://tools.invitrogen.com/content/sfs/manuals/pcdna3.1topota_man.pdf.

2.8.2 Primer and Gene Fragment Design

The gene sequences and fragments were initially designed without the insulin signal peptide and phosphorylated after synthesis. The fragments for both the A- and B-chain (see Table 2.1) and subsequent phosphorylation were undertaken commercially by 1st BASE Custom Oligos, Singapore. The sense and antisense DNA strands were synthesized separately and annealed using the following conditions; 1 µl containing 10 nmol of each strand was mixed in the presence of 1 µl of 10X T4 ligase buffer (Promega, USA). This mixture was then heated at 75°C for 10 min before being left to cool to room temperature for 1 h. The fragments were designed to have a NotI sticky site at the 5' end and a XhoI sticky site at the 3' end. Start and stop codon sequences were added to both fragments.

Table 2.1. The gene sequences of the sense and anti-sense strands of the camel insulin A and B chain fragments. The start and stop codons are labelled with red while the sticky ends are labelled with green.

Name	Sequence 5'-->3'
A-chain-Sense Strand	GGCCGCATGGGCATCGTGGAGCAGTGCTGCGCCAGCGTCTGCTCGCTCTACCAG CTGGAGAACTACTGCAACTAGC
A-chain-Anti Sense Strand	TCGAGCTAGTTGCAGTAGTTCTCCAGCTGGTAGAGCGAGCAGACGCTGGCGCAG CACTGCTCCACGATGCCCATGC
B-chain-sense Strand	GGCCGCATGTTCGCTAACCAGCACCTGTGCGGCTCACACCTGGTGGAGGCGCTG TACCTGGTGTGCGGGGAGCGCGGCTTCTTCTACACGCCCAAGGCCTAGC
B-chain-Anti sense Strand	TCGAGCTAGGCCTTGGGCGTGTAGAAGAAGCCGCGCTCCCCGCACACCAGGTAC AGCGCCTCCACCAGGTGTGAGCCGCACAGGTGCTGGTTAGCGAACATGC

The gene sequence of the preproinsulin (proinsulin with signal peptide) and proinsulin cDNA were amplified from pancreatic cDNA as described in section 2.8.3 below.

2.8.3 Synthesis of Camel Prepro- and Pro-insulin from cDNA

2.8.3.1 Extraction of Total RNA from Pancreas

Total pancreatic RNA was isolated as follows. 100 mg of fresh pancreatic tissue was frozen in liquid nitrogen and pounded into powder using a mortar and pestle. After transferring the tissue powder to a 2 ml tube, 1 ml of TRIzol reagent (Invitrogen, USA) was added, mixed thoroughly with the sample and incubated at room temperature for 5 min. A volume of 0.2 ml of chloroform was then added to the suspension and the tube was agitated for 15 sec followed by 3 min incubation at room temperature before being centrifuged at 12,000 g for 15 min at 4°C. The upper aqueous phase was then transferred into a new 1.5 ml tube. To precipitate the RNA, 0.5 ml of isopropanol was then added and the solution mixed by inversion before being incubated for 10 min at room temperature and then centrifuged at 12,000 g for 15 min at 4°C. The supernatant was then discarded and the pellet washed with 1 ml of 70% ethanol before being centrifuged at 12000 g for 5 min and the supernatant discarded again. The pellet was finally air dried

before being dissolved in 50 µl of diethylpyrocarbonate (DEPC)-water. The extracted RNA was then stored at -20°C until required for experiments.

2.8.3.2 Reverse Transcription cDNA Synthesis

Insulin cDNA was prepared using the commercially available ImProm-II Reverse transcription system (Cat.No. A3800, Promega, USA) by mixing 1 µg of isolated pancreatic RNA with 1 µl of random primer, 20 nmol of each deoxynucleoside triphosphate (dNTPs), 4 µl of ImProm-II 5X Reaction buffer, 3 mM magnesium chloride, 0.5 µl of recombinant RNasin ribonuclease inhibitor and 1 µl of ImProm-II reverse transcriptase in a total volume of 20 µl. The reaction mix was then incubated for 5 min at 25°C followed by a further 1 h at 42°C and then finally at 70°C for 15 min. In an initial trial experiment 1 µl of Oligo (dT) primer was used instead of the random primer.

2.8.3.3 Polymerase Chain Reaction to Synthesise Prepro- and Pro-insulin

The gene sequences of the Preproinsulin (proinsulin with signal peptide) and Proinsulin cDNA were amplified from pancreatic cDNA generated as described in section 2.8.3.1 using the primers described in Table 2.2 below.

Table 2.2. The sequences of the primers used to amplify the camel prepro-insulin and pro-insulin. All primes have a restriction site compatible with the vector into which the subsequent PCR product was to be cloned into and extra nucleotides at 5' end (overhang) for the restriction enzyme.

Oligo Name	Sequence (5`-->3`)
SP-INS-F	AAT GCG GCC GCA TGG CCC TGT GGA CAC GCC
Proinsulin-F	AAT GCG GCC GCA TGT TCG CTA ACC AGC ACC
Proinsulin-R	AAT CTC GAG GTT GCA GTA GTT CTC CAG CTG
SPHind-F	ACG AAT AAG CTT ATG GCC CTG TGG ACA CGC CTG
PROHind-F	ACG AAT AAG CTT ATG TTC GCT AAC CAG CAC CTG
INSSCXho-R	ACG AAT CTC GAG CTA GTT GCA GTA GTT CTC CAG

To enhance the restriction enzyme digestion efficiency, the primers used to amplify the prepro- and pro-insulin were designed to have 3-6 extra bases prior to the restriction site. Primers SP-INS-F and Proinsulin-R (Table 2.2) were used to amplify the camel preproinsulin gene while Proinsulin-F and Proinsulin-R (Table 2.2) were used to amplify camel proinsulin. Both forward primer sets had a NotI restriction site whilst the reverse primers had a XhoI restriction site. The primers SPHind-F and PROHind-F (Table 2.2) were used separately with the INSSCXho-R primer to amplify the camel preproinsulin and camel proinsulin respectively but with a Hind III restriction site at the 5' end and stop codon and XhoI site at the 3' end. The primers SPHind-F and PROHind-F were also used with Proinsulin-R to amplify preproinsulin and proinsulin fragments without a stop codon in order to express the protein with a His tag when cloned into the pcDNA3.1 vector. The total PCR reaction volume was 100 µL and consisted of 500 µM dNTPs, 500 nM of each of the appropriate primers, 5 Units of Fast start Taq DNA polymerase (Roche, Germany) and 10 µL of 10X Roche PCR buffer. A Tetrad thermal cycler from BioRad was used to run all PCR reactions. The cycling conditions were as follows: initial denaturation at 95°C for 5 min, 40 cycles of denaturing at 94°C for 30 sec, annealing at 58°C for 30 sec and polymerization at 72°C for 35 sec, followed by a final

extension at 72°C for 5 min. 5 µL of the resulting PCR products were analysed on a 2% agarose gel and the presence of a band of the expected size confirmed in this way. Prior to restriction digestion, all samples were purified using the commercially available wizard® SV gel and PCR clean-up system (Cat. No. A9281, Promega, USA) following the manufacturers instructions.

2.9 Restriction Enzyme Digestion of Plasmid DNA and PCR Products

Double digestion with NotI and XhoI restriction enzyme (New England Biolabs (NEB), UK) was initially undertaken. However, the restriction site was later changed and NotI replaced with HindIII to increase the digestion efficiency. One microgram (1 µg) of purified PCR product was digested with 10 units of each restriction enzyme in the recommended buffer (buffer 3 was used for NotI and XhoI double digestion while buffer 2 was used for HindIII and XhoI) in the presence of 1X bovine serum albumin (BSA) (NEB, UK) in a 50 µl volume and incubated for 3 h at 37°C. The enzymes were then heat inactivated at 65°C for 15 min. 2 µg of vector DNA (pcDNA3.1 and pcDNA5 FRT vectors from Invitrogen, USA) were digested with 20 units of restriction enzyme in the recommended buffer in a 100 µl volume overnight at 37°C followed by an enzyme inactivation step at 65°C for 15 min. The correct digestion of plasmid DNA was confirmed by conventional agarose gel electrophoresis. The digested PCR products and plasmids were then gel purified on a 1.5% agarose gel. DNA from the gel pieces was recovered using the commercially available wizard® SV gel and PCR clean-up system (Cat. No. A9281, Promega, USA). The gel pieces were dissolved in a 1:1 ratio *v:wt* membrane binding solution by heating at 55°C for 15 min until they melted completely before applying the resulting solution to the column provided in the kit. The column with the mixture loaded was then centrifuged at 12,000 g for 1 minute. The flow through was discarded and the DNA was washed with 750 µl of membrane washing solution provided with the kit for 1 min. The washing step was repeated with a 500 µl volume for 5 min and the column was again centrifuged at 12,000 g. The column was then placed

into a sterile 1.5 mL microfuge tube and the DNA was eluted from the column using warm (50°C) dd. H_2O.

2.9.1 Ligation of Plasmid DNA and PCR Products

All ligation reactions were undertaken in a total volume of 10 µl using 3 Units of T4 DNA ligase (Promega, USA) in 2X rapid ligation buffer (Promega, USA). The molar ratio of insert to vector was 3 to 1. Ligation reactions were then left to proceed at 4°C overnight.

2.9.2 Transformation of Ligation Mixtures into Competent *E.coli* Cells

Four microliters (4 µL) of ligation mix was used to transform JM109 *E.coli* competent cells (Promega, USA). Bacterial transformation was performed using the heat shock method at 42°C for 30 sec then immediately chilling the mixture on ice for 5 min. The cells were then left to grow in 400 µl of SOC media (Invitrogen, USA) for 1 h before spreading the transformed cells onto Luria-Bertani (LB) plates which contained ampicillin antibiotic (GIBCO, USA) at a 100 µg/ml final concentration.

2.9.3 Colony Selection and Screening Following Transformation

Colonies were picked from the culture plates and sub-cultured individually in 1 ml of Luria-Bertani LB culture media containing ampicillin at a concentration of 100 µg/ml. These cultures were incubated at 37°C for 6 h with shaking at 200 rpm in microfuge tubes. The plasmid DNA from all colonies was isolated using a crude DNA method whereby 10 µl of 0.2 M sodium hydroxide was added to 10 µl of the cultured cells, incubated for 5 min at 95°C, and the solution then neutralized with 35 µl of Tris-HCl (pH 7.4) before the solution was centrifuged for 30 sec at 12000 g. One microliter from the supernatant was then used for the subsequent PCR to confirm the presence of the expected gene insert. All colonies were screened by PCR using a T7 forward primer 5'-

TAATACGACTCACTATAGGG-3' and a BGH reverse primer 5'-TAGAAGGCACAGTCGAGG-3'. The total PCR reaction volume was 20 µL and consisted of 100 µM of dNTPs, 100 nM of each of the appropriate primers, 1 Unit of Fast start Taq DNA polymerase (Roche, Germany) and 2 µL of 10X Roche PCR buffer. A Tetrad thermal cycler from BioRad was used to run all PCR reactions. The cycling conditions were as follows: initial denaturation at 95°C for 5 min, 35 cycles of denaturing at 94°C for 30 sec, annealing at 53°C for 30 sec and polymerization at 72°C for 35 sec, followed by a final extension at 72°C for 5 min. Five microliters of the PCR products were then analysed on a 2% agarose gel. The positive colonies were identified by the PCR product size difference in comparison with the control wild type vector band. The cloning site and insert of the selected positive clones were sequenced using a T7 forward primer and following the sequencing procedure described in section 2.4.3. The sequencing results were also confirmed by a second sequencing reaction using the BGH reverse primer. The positive colonies were then sub-cultured and plasmid DNA generated and purified from these for subsequent experiments as described below in section 2.10.

2.10 Generation of Recombinant Plasmid DNA for Transfection Experiments

Following the generation of the cloned camel insulin plasmid material, it was necessary to generate sufficient material for transfection into mammalian cells.

2.10.1 Small Scale (Mini-Prep) and Large Scale Preparation (Maxi-Prep) of Plasmid DNA for Transfection Experiments

For all plasmid DNA preparations, anion ion exchange columns were used, in both small and large purification scales from commercially available kits as described below. The first constructs were purified at a small scale using mini-prep columns, later maxi-prep columns were used to produce large amount of the construct.

For plasmid small scale purification (mini-prep), 100 µl of culture was initially used to inoculate 5 ml of a Luria-Bertani LB culture containing ampicillin and the subsequent solution incubated at 37 °C for 16 h. On the following day the plasmid DNA was isolated using the commercially available Wizard® Plus SV Miniprep DNA Purification System (Promega, USA) following the manufacturers instructions. Briefly, the bacterial cells were pelleted by centrifugation at 4000 g for 10 min and the pellet then resuspended in 250 µl of cell resuspension solution. The cells were then lysed by adding 250 µl of Cell Lysis Solution and the tubes inverted 4 times for mixing. After 5 min at room temperature, proteins were digested using 10 µl of alkaline protease solution. The tubes were incubated for 5 min at room temperature after 5 inverts. Following this, 350 µl of neutralization solution was added and the tubes inverted immediately 6 times for mixing. The cell debris and genomic DNA were then pelleted by centrifugation at 12000 g for 10 minutes. The supernatant was transferred onto a Spin Column provided with the kit and centrifuged at 12,000 g for 1 min to bind the plasmid DNA to the column. The bound DNA was washed twice using the Washing Solution (using 750 µl initially for 1 min and the second wash was with 250 µl for 5min) and then the DNA eluted with 100 µl of warm (50°C) Nuclease Free Water by centrifugation at 12000 g for 1 min after 5 min incubation at room temperature.

A similar procedure was undertaken for the large-scale purification but using the commercially available Qiagen kit. For large scale purification (maxiprep), 500 µl of a positive culture was added to 200 ml of Luria-Bertani LB culture medium containing ampicillin at a concentration of 100 µg/ml and incubated at 37°C for 16 h with shaking at 200 rpm. On the following day the plasmid DNA was isolated using QIAGEN Plasmid Maxi Kit (cat. No. 12162, Qiagen, USA) following the manufacturers instructions. Briefly, the bacterial culture was harvested by centrifugation for 10 minutes at 4000 g at 4°C. The supernatant was then discarded and the pellet was resuspended in 10 ml of Buffer P1 (Resuspension Buffer containing RNase A). A volume of 10 ml of Buffer P2 (Lysis Buffer) was then added to the resuspended pellet, mixed gently and

incubated for 5 min at room temperature. Ten millilitres of chilled Buffer P3 (Neutralisation Buffer) was then added and mixed immediately. The lysate was then poured into a QIAfilter Cartridge, which had been closed with a cap on its outlet noozle and left for 20 min at room temperature. Using 10 ml of QBT Buffer the QIAGEN-tip 500 had been equilibrated by applying the buffer and letting the column empty by gravity flow. The plunger supplied with the kit was then gently inserted into the QIAfilter Cartridge and the lysate filtered through the previously equilibrated QIAGEN-tip 500. The QC buffer was then used to wash the QIAGENtip 500 twice with 30 ml used for each wash. Subsequently the DNA was eluted with 15 ml of Buffer QF (Elution Buffer) and then precipitated using 10.5 ml of isopropanol (Merck). After mixing the precipitated DNA was centrifuged immediately at 4000 g for 30 min at 4°C and the supernatant removed carefully. To remove the precipitated salt, 5 ml of room temperature 70% ethanol (Merck) was added to the pellet and the tube was then centrifuged at 4000 g for 20 minutes at 4°C. The supernatant was then completely removed and the pellet air dried and re-dissolved in 500 µl of nuclease free H_2O. To check the purification efficiency, approximately 5 µl of the sample was run on an agarose gel. The remaining sample was stored at -20°C until required for further experiments.

2.11 Transfection of CHO Cells with Plasmid DNA

A Chinese hamster ovary cell line (CHO-K1, sourced from ECACC) was transiently transfected with camel insulin A-chain and B-chain, proinsulin and preproinsulin recombinant pcDNA vectors in order to determine if insulin could be produced from these constructs The commercially available Flp-In CHO cell line (Invitorgen) was also transfected to try and generate stably expressing cell lines.

2.11.1 Transient Transfection of CHO-K1 Cells

Insulin A-chain and B-chain constructs with stop codon in the pcDNA3.1V5 vector were used in this transfection procedure. The procedure was repeated with prepro-insulin and pro-insulin pcDNA3.1V5 constructs with and without stop codon. Prior to transfection 2.0×10^5 CHO-K1 cells were seeded in each well of a 6-well tissue culture plate in 2.5 ml of CHO-K1/DMEM medium (which contained Dulbecco's Modified Eagle's Medium/Nutrient F-12 Ham (DMEM/F12) (Invitrogen), 10% Fetal calf serum, 6 mM L-Glutamine, 1x non-essential amino acids (Gibco), 7 mg/L each of adenosine, guanosine, cytidine, uridine, 2.4 mg/L of thymidine and 60 mg/L each of L-glutamic acid and asparagine) and cultured overnight at conditions of 5% CO_2, 37°C. The following day, the cells were transfected using the commercially available Lipofectamin 2000 (Invitrogen, USA) cationic lipid reagent whereby two solutions were initially prepared – (a) 250 µl OptiMEM and 12 µl Lipofectamine 2000 were mixed in a microfuge tube and (b) a further 250 µl of OptiMEM was added to 4 µg of vector construct DNA and mixed. Both mixes were then incubated at room temperature for 5 min. Following this 5 min incubation, the two solutions were mixed and incubated for an additional 20 min at room temperature. Subsequently, the solution was added to the cells gently and drop wise. The cultures were then left for 5.5 h at 37°C, 5% CO_2. Afterward this period the medium was replaced with 2.5 mL of fresh CHO-K1/ DMEM media and the cells incubated at 37°C, 5% CO_2 until harvested to analyse for insulin.

2.11.1.1 Cell lysis and Collection following Transfection

Forty eight hours after transfection, the medium from transient transfection experiments was collected and transferred to 15 ml tubes. The tubes were centrifuged at 200 g for 5 min to separate any cells and the supernatant then stored at -80°C. Each well of the adherent transfected cells was washed with 3 ml of phosphate buffer saline (PBS). After removing the PBS, 200 µl of extraction buffer (20 mM HEPES, 100 mM NaCl, 1% Triton_X 100, 10 mM β-glycerol phosphate, 50 mM sodium fluoride, 50 mM sodium vanadate and one tablet of Complete Protease Inhibitor Cocktail (Roche) per 10 ml

buffer) were added to each well. The cells were then scraped using a cell scraper and the scraped cells transferred to mircrofuge tubes, vigorously vortexed and incubated on ice for 30 min. The tubes were then centrifuged at 4000 g for 10 min at 4°C, the supernatant/lysate separated and stored at -80°C until required for further experiments.

2.11.2 Stable Transfection of CHO Flp-In Cells

Constructs of the pcDNA5 FRT vector with prepro-insulin and pro-insulin were used in the transfection procedure of the commercially available FRT Flp-In cells from Invitrogen. Each construct was designed to have and to lack a stop codon (that lacking reading into a distal his-tag), therefore a total of 4 vectors were used in this procedure. An additional two controls were also used, one was transfection of the Flp-In cells with POG44 vector alone and the other was the cell line without transfection. Flp-In CHO-K1 cells were maintained in Flp-In medium which contain Ham's F12 medium (Cat. No. 21765-029, Invitrogen, USA) containing 10% Fetal Bovine Serum (FBS) (Cat. No. 16000-044, Invitrogen, USA) and 0.1% zeocin. Prior to transfection, 5 x 10^5 cells were seeded into a T25 culture flask with Flp-In medium without Zeocin cultured in 5% CO_2 at 37°C. Twenty four hours later, the cells were transfected using Lipofectamin 2000 (Invitrogen, USA) more-or-less as outlined in 2.11.1 above. Briefly, 625 µl of OptiMEM and 30 µl of Lipofectamine 2000 were mixed in a microfuge tube whilst in another tube 625 µl of OptiMEM was used to dilute 9 µg of the pOG44 vector and 1 µg of the vector construct DNA. Both mixes were then incubated at room temperature for 5min before these were themselves mixed and incubated for an additional 20 min at room temperature. Subsequently, the solution was added to the cells in the T25 flask that were prepared the previous day. The cultures were then allowed to transfect for 5.5 h at 37°C, 5% CO_2. Following this incubation period, the cells were washed with PBS and the transfection medium was then replaced with 6 ml of Flp-In medium without zeocin. After 48 h the Flp-In medium was replaced with fresh medium supplemented with Hygromycin of 500µg/ml final concentration. The medium was aspirated and replaced with Hygromycin containing medium every 48 h hours. Eight days after transfection

when distinct colonies clearly appeared the cells were washed with 5 ml PBS, followed with 3 min trypsin-EDTA (Invitrogen) treatment with 500 µl volume per T25 flask at 37°C. Twenty millilitre of hygromycin containing Flp-In medium was added to inactivate the trypsin. Cells were then split into 2 T25 flasks and cultured at 5% CO_2, 37°C. The hygromycin containing Flp-In medium was changed 3 days later and the cells were transferred to a T75 flask and after another 3 days the cells were transferred to a T175 flask with fresh hygromycin containing medium. During each transfer half of the cells were cultured and one quarter taken for analysis and the remaining quarter was stored in liquid nitrogen as below.

2.12 Cryopreservation of Cells in Liquid Nitrogen

To preserve CHO cells in liquid nitrogen, the cell cultures were centrifuged at 250×g for 5 minutes. The supernatant was then removed and the cells were resuspended in the Flp-In medium consisting of 10% FCS and 90% Ham's F-12 at a cell density of 5×10^6 cells/ml. The cryoprotective agent, dimethyl sulfoxide (DMSO) was then added to the cell suspension at a ratio of 1:9. The cell suspension was then dispensed into sterile, screw capped, plastic cryotubes in 1 ml aliquots. The ampoules were placed in a -80°C freezer for 3 h before being transferred to liquid nitrogen storage.

2.13 Cell Lysis using RIPA Buffer

One quarter of the stable transfected cells were analysed after each trypsin treatment. The trypsin treated cells were transferred to 15 ml tubes and then to remove the culture medium centrifuged at 200 X g for 5 min. Cells were then washed with 5 ml of sucrose containing solution. After removing the sucrose, 400 µl of RIPA buffer (Brenner, 1997) (25 mM HEPES, 0.2% SDS, 0.5% deoxycholic acid, 125 mM NaCl, 1% Triton_X 100, 10 mM sodium fluoride, 10mM sodium pyrophosphate, 10 mM sodium orthovanadate and one tablet of protease inhibitor per 10 ml buffer) was added to lyse the cells. The cells were passed through a needle and syringe several times to breakdown the cells and

DNA, followed by incubation at 4°C for 30 min. The tubes were then vortexed and incubated again for another 30 min at 4°C and finally centrifuged at 4000 g for 10 min at 4°C, the supernatant was separated and stored at -80°C until required for analysis

2.14 Analysis of Material from Transfected CHO Cells

In order to investigate the transfected cell lines western blotting and radioimmunoassay were undertaken.

2.14.1 Western Blotting

All supernatants and cell lysate were analysed by western blotting for the presence of insulin. Ten microliters of protein were added to 10 µl of loading buffer (Laemmli buffer, Bio-Rad, USA) and incubated at 70°C for 10 minutes. The western blotting was undertaken following the western blotting procedure described in section 2.2.4.

2.14.2 Radioimmunoassays

The cell lysate and the culture supernatants were analysed using a human anti-insulin antibody using the human insulin RIA kit (INS-Irma, KIP1251-KIP 2154, Diasource Europe, Belgium). The radioimmunoassay was undertaken according to the procedure described in section 2.2.3.

2.15 Chemical Synthesis of the Insulin Peptides

Human and camel insulin A and B-chains were chemically synthesized using a peptide synthesizer and the resulted polypeptides analysed by liquid chromatography mass Spectrometry (LC/MS) and purified using reverse phase high performance liquid chromatography (HPLC).

2.15.1 Solid Phase Peptide Synthesis

Four different peptides were synthesised by solid phase peptide synthesis which was conducted on PSSM-8 peptide synthesizer from SHIMADZU Corporation, Japan. The peptides synthesised are described in Table 2.3 below.

Table 2.3. Amino acid sequences of the 4 synthetic peptides generated during this study

Peptides	N - terminus → C - terminus	Species
Insulin A-Chain	G I V E Q C C A S V C S L Y Q L E N Y C N	Camel
Insulin A-Chain	G I V E Q C C T S I C S L Y Q L E N Y C N	Human
Insulin B-Chain	F A N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K A	Camel
Insulin B-Chain	F V N Q H L C G S H L V E A L Y L V C G E R G F F Y T P K T	Human

Peptides A and B of camel and human insulin were synthesized at 20 µmol scale using a 0.2 mmol/g of NovaSyn®TGT resin (Novabiochem) and 0.16 mmol of 9-fluorenylmethoxycarbonyl (Fmoc) amino acids (8 equivalents). In the synthesis process 3 different types of cysteine were used, each cysteine was protected with different side chain protecting groups. The first cysteine in B-chain and the second cysteine of A-chain were protected with the acetamidomethyl (Acm) protecting group, while the last in both chains cysteine was protected with a triphenylmethyl (Trt) protecting group. A-chains have two more cysteine residues involved in intrachain disulphide bridge formation, therefore A-chain peptides were synthesised to have the first and the third cysteine residues with tert-butyl (tbu) protecting groups. Prior to synthesis the resins were washed with DMF for 3 minutes. To start the synthesis, the first resin-bound amino acid, Fmoc was deprotected using piperidine where two resin volumes of freshly prepared 20% piperidine in DMF were recirculated through the resin two times for 7.5 min each. The reaction mixture was drained and the resin washed five times for 1 min each with excess DMF. For the next amino acid to be coupled, 160 µmol of the relevant Fmoc amino acid was added along with 500 µmol of each of the coupling reagents (both

coupling and activating reagents were prepared in DMF) N-Hydroxybenzotriazole(HOBt)/ 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium (HBTU) and 1 mmol of the activating reagent *N, N*-diisopropylethylamine (DIPEA). The reaction mixture was incubated for 45 min to couple the amino acid. Between each coupling and deprotection step the excess amino acids were removed by 5 washes with 600 µl DMF for each wash, then washed with 5 ml of DMF (See Appendix 2 and 3 for more details about the synthesis calculations).

To release the final peptide product from the resin, the synthetic peptides were treated with Trifluroacetic Acid (TFA) and two scavengers, ethanedithiol and triisopropylsilane, to react preferentially with the free radical and other reactive species released during the cleavage reaction. At the end of the synthesis, the resins were washed twice with methanol and allowed to dry in a speed vacuum concentrator for 1 h. To cleave the peptide from the resin, 750 µl of freshly prepared ice-cold cleavage mixture (1% triisoproplysilane, 2.5% ethanedithiol, 2.5% H2O, 94% TFA) was added, and the mixture was incubated for 1 h, then another 750 µl of the ice-cold cleavage mixture was added and incubated again for 1 h. The reaction mixture was collected from the resins by draining it into 30 ml of cold diethyl ether in a 50 ml centrifuge tube and the resins were washed again for 5 minutes with 1 ml of undiluted TFA. The washing solution was collected from the resin by draining it again into the cold diethyl ether which was used for the previous collection. The peptide was precipitated immediately by centrifugation at 4000 g for 10 min. The supernatant was decanted and precipitates were washed twice with 10 ml of diethyl ether for each wash. The peptides were then dried in the air to remove diethyl ether and redissolved in 1 ml of 80% acetonitrile containing 0.1% TFA and mixed well with 1 ml more of 0.1% TFA added. For insulin A-chain peptides, 1.2 g of urea was added to enhance the solubility of the peptide. Five microliters of each of the 4 synthetic peptides were diluted in 45 µl of 0.1% TFA and analysed by liquid chromatography-mass spectrometry. The mass spectrometry analysis was conducted

using a micrOTOF-Q machine (Bruker Daltonics) using the electrospray ionization mass spectrometry technique.

2.15.2 HPLC Purification of Synthetic Insulin Peptides

High pressure reverse phase liquid chromatography was used to purify the cleaved crude peptides from the scavengers and other synthesis by-products. HPLC was undertaken using a Waters 600 HPLC system and a C18 reversed phase column. Each synthetic peptide was diluted in a 1:1 ratio with 0.1% TFA to reduce the concentration of the acetonitrile to 20% and then injected in 8 separate injections with a batch volume of 0.5 ml. The peptides were injected onto the C18 reversed phase preparative column and then eluted at a flow rate of 4 ml/min using a gradient of aqueous acetonitrile containing 0.1% formic acid. The elution fraction of the B-chain peptide was collected 27 min after injection while the A-chain peptides were collected after 31 min. The fractions of each peptide were pooled and stored at -20°C after lyophilisation.

2.16 Investigating Insulin Stability in Milk

Bovine and human insulin were dissolved in 50 mM Tris-glycine buffer (pH 2.5) and incubated for 2 min at room temperature whereupon the solution was immediately neutralized with 1 M Tris buffer. The two insulin solutions, along with the isolated camel milk insulin (the fraction where the insulin concentration was 90000 µIU/ml) were then diluted in bovine and camel milk samples to give a final insulin concentration between 250-300 µIU/ml. Prior to the thermal stability experiment, the insulin concentration in the diluted samples was measured by RIA to assure the required concentration. All samples, along with plain milk (unspiked), were divided into two aliquots. One was incubated at 60°C and the other at 72°C, both for one hour. Following the thermal (heat) stress, the insulin concentration was measured by RIA after the thermal treatment and compared with original concentration.

Chapter 3

Camel Insulin Gene Sequencing, SNP, and Chromosome Localization Analysis

3.1 Introduction

The insulin gene is present in a single copy in most species, however in the mouse and the rat there are two copies of the insulin gene present. The human insulin gene is localised on the short arm of chromosome 11 at position p15 (Harper *et al.*, 1981). On-the-other-hand, both copies of the rat insulin gene are located on chromosome 1 at the q-arm and the genes are approximately 100,000 kilobases distant from each other (Soares *et al.*, 1985), while in mouse one copy of the insulin gene (insulin 1 gene) is localised on chromosome 19 (Davies *et al.*, 1994) and the second copy of the insulin gene (insulin 2 gene) is positioned on chromosome 7 (Duvillie *et al.*, 1998). The human insulin gene and the rodent insulin 2 gene contain three exons and two introns, while the rodent insulin 1 gene lacks the second intron (Melloul *et al.*, 2002) (see Figure 3.1).

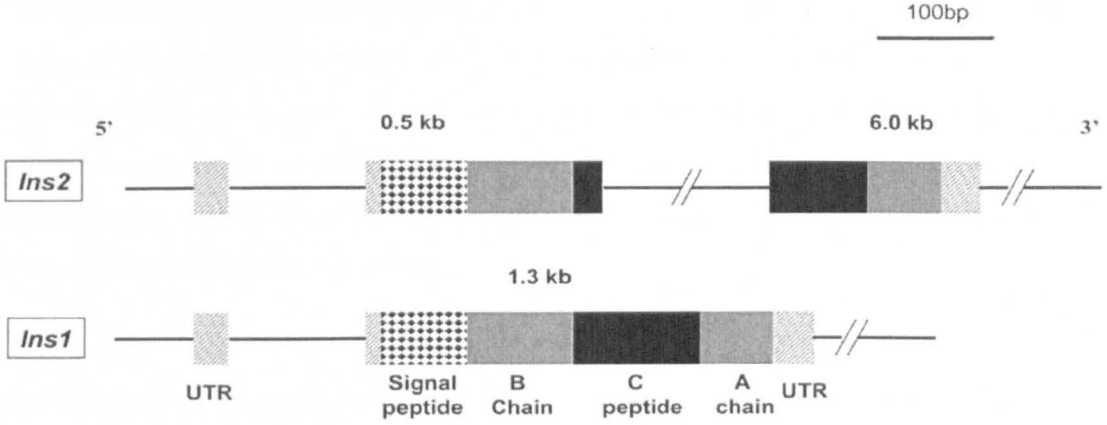


Figure 3.1 The gene structure of Ins2 and Ins1 genes in the mouse. Boxes show the exon regions while the solid lines indicate the intronic and the flanking regions of the gene (Shiao *et al.*, 2008).

The insulin protein is present in invertebrates all the way up to humans making the study of it's evolution an important and interesting subject (Falkmer and Ostberg, 1976). Studies on the structure of crystalline Hagfish *(Myxine glutinosa)*, thought to have diverged from the main line of vertebrate evolution about five hundred thousand years ago, have shown that even with 37% difference at the amino acid level (Peterson *et al.*, 1975) the main peptide chain sequence matches porcine insulin (Cutfield *et al.*, 1979). This shows that the basic structure of insulin evolved was in the early vertebrates and has subsequently been largely unchanged with only small changes during subsequent evolution leading to man (Shu *et al.*, 1981). The existence of insulin like molecules in insects, molluscs and echinoderms (Falkmer and Ostberg, 1976) as well as the presence of other insulin like proteins in vertebrates, all suggest that insulin is an ancient protein (Shu *et al.*, 1981).

It has been suggested (de Haen *et al.*, 1976) that insulin and other structurally related proteins like relaxin, NGF (nerve growth factor) and IGF (insulin-like growth factor) may have evolved from the serine protease proteins such as trypsinogen and chrymotrypsinogen. This route of evolution is supported by a number of observations; (1) the presences of an intron within the C-peptide region of the insulin gene that corresponds to the removal of approximately 140 residues from the central region of the serine protease structure to leave a shortened connecting peptide similar to that found between the B- and A-chains of insulin (Shu *et al.*, 1981); (2) it is well known that some proteases can act in an insulin like manner and can stimulate many of the cellular functions that insulin stimulates such as like cell replication (Holley, 1975; Glenn and Cunningham, 1979; Shu *et al.*, 1981).

As described earlier in this thesis, the insulin protein is produced as a single precursor molecule termed preproinsulin. The preproinsulin is post translationally processed and modified to produce proinsulin and finally cleaved to produce the mature insulin hormone (Wentworth et. al., 1986). Glucose is the main regulator of insulin gene

expression. Insulin biosynthesis is mainly regulated by glucose levels (Tillmar *et al.*, 2002) with insulin biosynthesis being controlled at the level of protein synthesis and elongation rates (Wesh *et al.*, 1986). In addition, glucose can also enhance the production of insulin mRNA (Brunstedt and Chan, 1982) by the selective stimulation of the transcription of the insulin gene (Nielsen *et al.*, 1985) and by increasing the stability of insulin mRNA (Wesh *et al.*, 1985; Tillmar *et al.*, 2002). The half-life of insulin mRNA is approximately 29 h at low glucose concentration while at high glucose concentration the mRNA is stabilized such that the half-life is extended to approximately 77 h (Wesh *et al.*, 1985). Such studies have therefore proposed that the insulin mRNA levels in insulin generating cells is rapidly increased upon an increase in glucose concentrations by the stimulation of insulin gene transcription and changes in mRNA stability (Leibiger *et al.*, 1998; Leibiger *et al.*, 2000; Tillmar *et al.*, 2002).

The control of insulin expression and its regulation is therefore a complex process that includes the stimulation and control of transcriptional and post-transcriptional mechanisms (Permutt and Kipnis, 1972). The insulin gene in adults is expressed only in the β-cells of pancreatic Langerhans islands (Marshak *et al.*, 2001). Insulin production is controlled at the transcription level of the gene encoding preproinsulin (the insulin gene), rather than at the level of preproinsulin mRNA translation (Iype *et al.*, 2005). Hence it is transcriptional activity and its control that regulates the levels of insulin made and there is little translational control. The amount of insulin produced is governed by the levels of mRNA present, which in turn are regulated by transcriptional control elements as discussed above.

With regard to the genetics of the camel insulin gene, very few studies have investigated this. The karyotypes of different camelid species have been studied previously by many groups (Taylor *et al.*, 1968; Koulischer *et al.*, 1971; Bunch *et al.*, 1985; Bianchi *et al.*, 1986; Graphodatsky, 2006; Di Berardino *et al.*, 2006). All studies indicated that the

diploid number of chromosome in camelid species is 74, but no conformity on chromosome nomenclature has been reached (Balmus *et al.*, 2007).

Balmus and co-workers (2007) prepared a set of chromosome painting probes from the dromedary camel (*Camelus dromedarius*) using flow sorting and degenerate oligonucleotide primed PCR techniques. The painting probes were used to identify the karyotypes of the dromedary camel (*C. dromedarius*), the bactrian camel (*C. bactrianus*), the guanaco (*L. guanicoe*), the alpaca (*L. pacos*) and the hybrid of dromedary and guanaco karyotypes. In these studies Blumus and colleagues established a GTG banding karyotype and idiogram for *C. dromedaries* as shown in Figure 3.2. Furthermore, the cross species chromosome painting between camel, cattle, pig and human with painting probes from the camel and human allowed the establishment of genome wide comparative maps (Balmus *et al.*, 2007). The study showed that the dromedary karyotype consists of one metacentric, three sub-metacentric and 32 acrocentric autosomal chromosomes. Both X and Y chromosomes are metacentric but X is a large chromosome while Y is a small chromosome. With the small chromosomes (from 22 to 36) the correct pairing and identification can be established only by FISH, this being because the GTG-banding patterns are too similar to clearly differentiate these chromosomes (Balmus *et al.*, 2007).

This chapter investigates and discusses the composition of the camel insulin gene. In this chapter, for first time, the full camel insulin gene was sequenced along with part of the upstream and downstream regions. Five different single nucleotide polymorphisms (SNP's) were identified by sequencing the insulin gene of 32 camels. The camel insulin sequence was compared with the insulin sequence from 13 different species and the potential structure of camel proinsulin investigated and predicted using 3D-JIGSAW software. Ultimately the gene was mapped onto camel chromosomes by the use of fluorescence *in situ* hybridisation (FISH) technique.

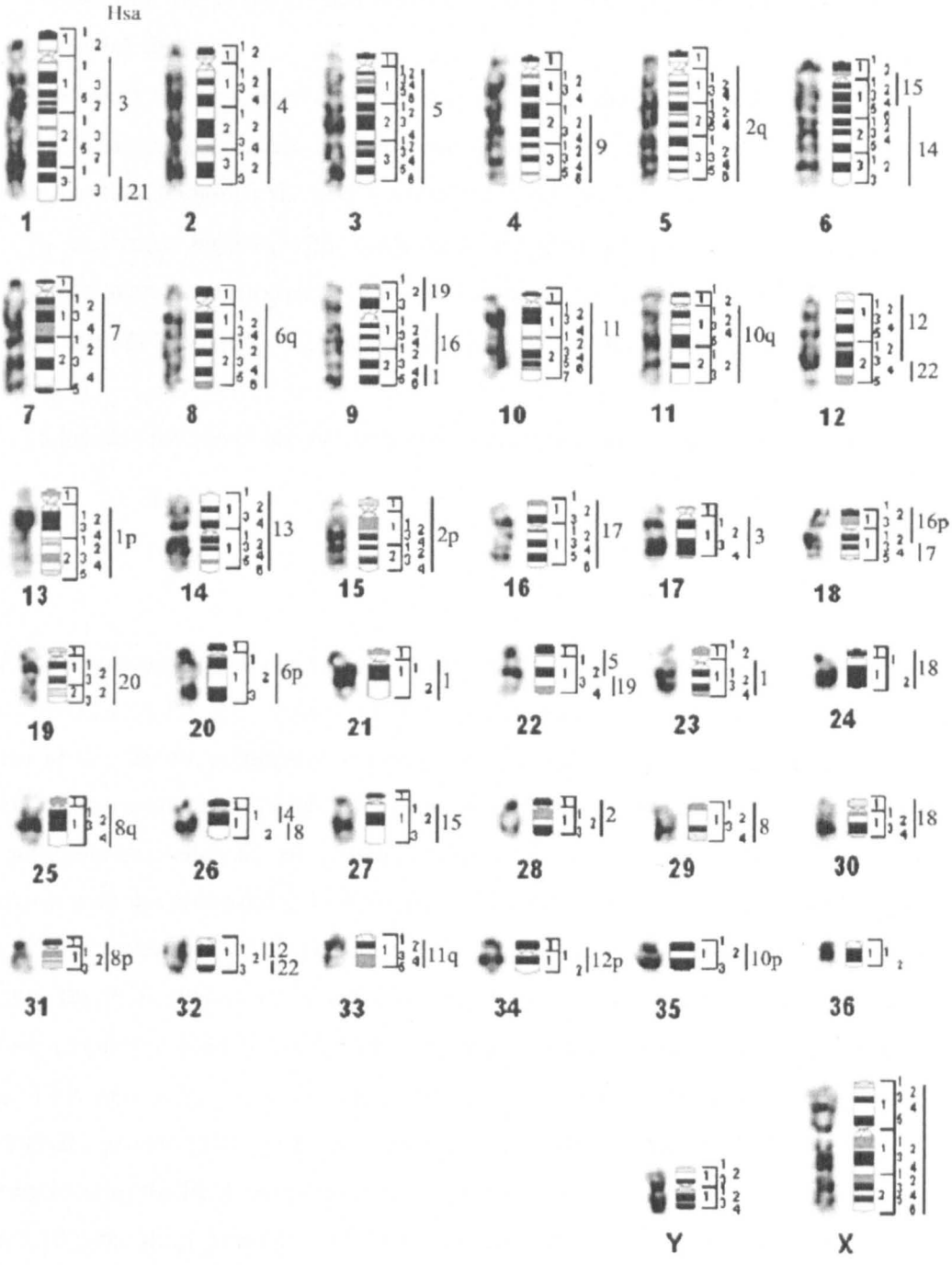


Figure 3.2. GTG banding karyotype and the idiogram of Dromedary camel indicating the chromosomal homology with human (HSA) (Balmus *et al.*, 2007).

3.2 Aims:

- To amplify the camel insulin gene and part of the upstream and downstream regions of the gene.
- To identify the regions of the functional regions of the camel insulin gene.
- To sequence the camel insulin gene across multiple camels to confirm the sequence and identify the SNP's within the camel insulin gene.
- To undertake bioinformatic analysis to compare the camel insulin gene and protein with other species and identify the sequence differences.
- To the 3D structure of camel proinsulin protein and compare it with human model.
- To localise the camel insulin gene on camel chromosomes using FISH technique.

3.3 Results

3.3.1 PCR Amplification of the Insulin Gene from Camel Genomic DNA

Primers CAM-INS-F2 and CAM-INS-R2, which had been previously used (AL-Swailem *et al.*, 2008) to amplify camel proinsulin cDNA, successfully amplified a fragment of approximately 772 bp from camel genomic DNA as determined by agarose gel electrophoresis analysis of the resulting PCR products (see figure 3.3). In comparison with the published cDNA fragment, the intronic region accounts for 507 bp of the whole fragment and a single and specific PCR product was achieved after increasing the PCR stringency. The PCR stringency was increased by increasing the annealing tempreture to 64°C and reducing the $MgCl_2$ concentration in the PCR reaction. Figures 3.3A and 3.3B show the electrophoresis analysis of the CAM-INS-F2 and CAM-INS-R2 primer PCR reactions before and after optimizing the PCR conditions. Before optimising the PCR reaction a series of up to 5 bands or products were observed (figure 3.3A), the most prominent of which corresponded to the desired product. This

was the only band observed after the PCR conditions were optimised as described above (see figure 3.3B)

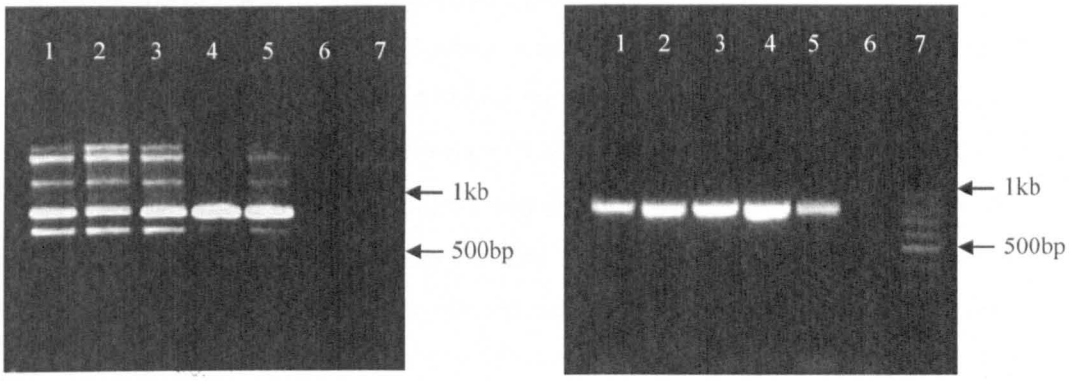


Figure 3.3 A) Lanes 1-5 different camel DNA amplified by CAM-INS-F2 and CAM-INS-R2 primers lane 6 is the negative control and lane 7 is the molecular 100 bp size marker. A: shows the result by using the cDNA amplification conditions for genomic DNA. B) the results after optimizing the PCR conditions.

3.3.2 DNA Sequencing of the PCR Amplified Camel Insulin Gene

The 772 bp PCR amplified fragment described in section 3.2.1 above was sequenced and aligned using BLAST search. The sequence is provided in Appendix 1 of this thesis. The sequenced fragment matched a BAC clone sequence of *Llama pacos* CH246-515B17 clone with 96% homology. Figure 3.4 shows the BLAST result and alignment. Primers F1 and R2 (see section 2.4.3 for sequence of these primers) were designed from the matched clone amplified fragment of 1960 base pair and 1849 base pair sequence of the pancreatic insulin and surrounding regions and have been sequenced from the camel genomic DNA. Figure 3.5 shows the agarose gel electrophoresis analysis when using primers F1 and R2 to amplify the genomic DNA, whereby a band of 1960 bp in size was clearly present.

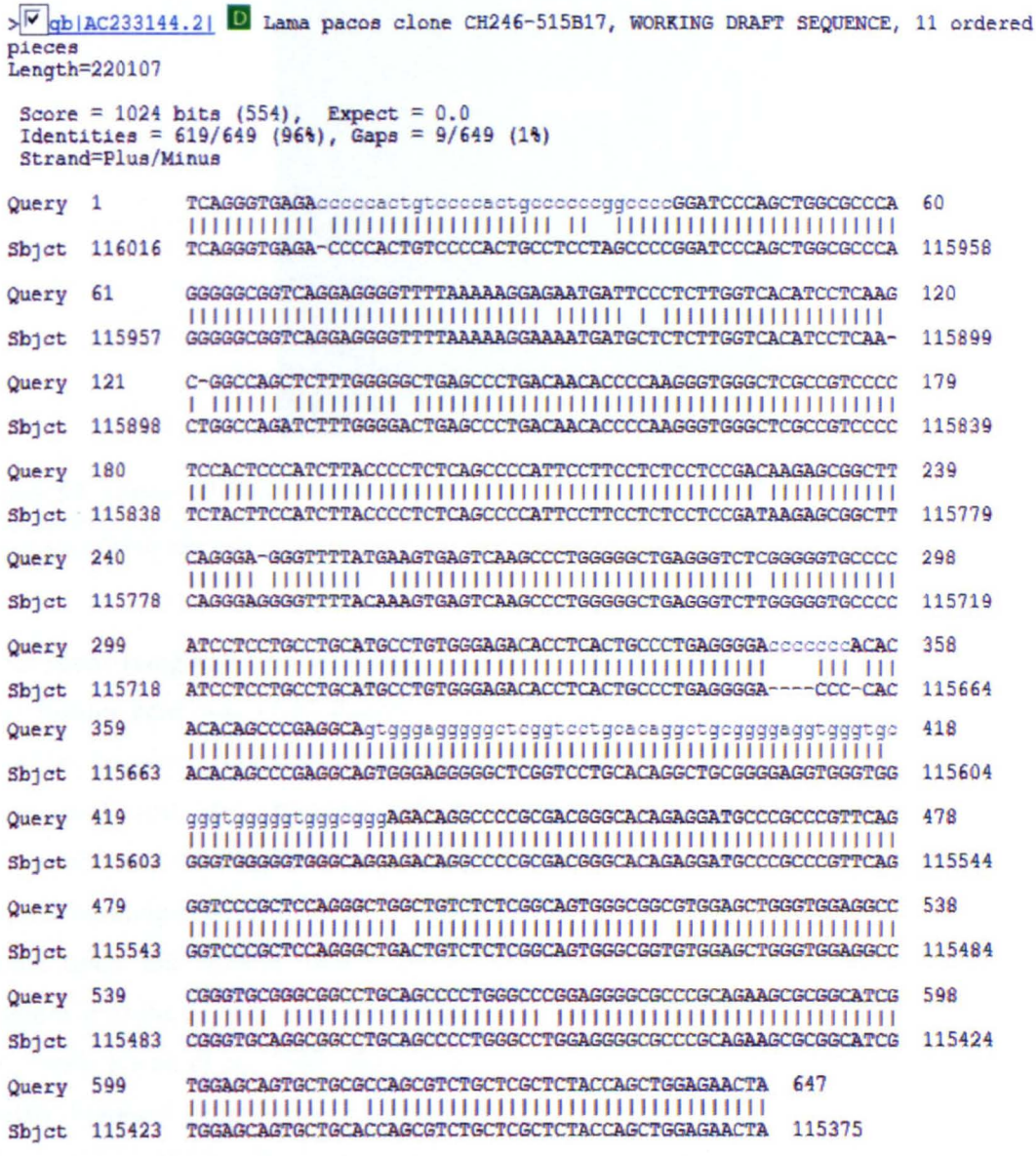


Figure 3.4 First hit in BLAST results of the sequence of CAM-INS-F2 and CAM-INS-R2 primers product. The results shows 96% homology between the sequenced dromedary camel insulin gene and the clone number CH246-515B17 from Lama pacos.

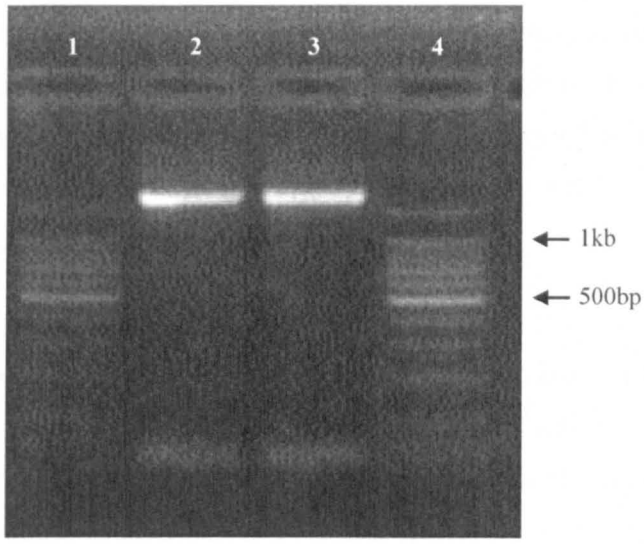


Figure 3.5 Agarose gel electrophoresis analysis of the PCR fragmenta generated from genomic camel DNA as described in section 3.2.1 above. Lane 1 and 4 are 100 bp size marker, lane 2 and 3 the insulin gene band of 1960 base pair fragment.

3.3.3 Identifying and Characterising the Camel Insulin Gene Regions

The insulin gene was PCR amplified from the genomic DNA of 32 camels and the resulting sequence compared across the camel samples to determine the camel insulin gene start and stop regions and to identify any potential single nucleotide polymorphisms (SNPs). Figure 3.6 shows the sequencing results of the insulin gene and the surrounding regions from the 32 camel samples. The gene regions were identified based upon the BLAST results by comparing the obtained camel sequence after alignment of the 32 sequences with other published results (Bajaj *et al.,* 1986, Brown *et al.,* 1955; Kwok *et al.,* 1983; Buck *et al.*, 2001; Hay and Docherty, 2006; Stoy *et al.,* 2010). Figure 3.6 highlights the signal peptide, A and B chains, C-peptide, Goldberg-Hogness box (TATA Box), cleavage site, stop (TAG) codon, and termination signal. The results show that the camel insulin gene contains a 72 bp length signal sequence, a 90 bp insulin B-chain, a 93 bp C-peptide, a 63 bp insulin A-chain and a 507 bp intron within the C-peptide. A 217 bp sequence was present between the TATA box and the ATG start codon while a 48 bp length of sequence was present between the stop codon and the termination signal. The sequence between the stop codon and termination signal

is rich in guanine and cytosine nucleotides (Figure 3.6). The data confirms that sections of the insulin gene are conserved, in camel insulin with other species as shown by comparison with conserved regions in the A and B-chains as reported in previously published data (Bajaj *et al.*, 1986; Brown *et al.*, 1955; Kwok *et al.*, 1983; Buck *et al.*, 2001 ; Hay and Docherty, 2006; Stoy *et al.*, 2010).

In addition to defining the regions of the camel insulin gene, the data generated from the genomic DNA sequence across the 32 camels allowed the identification of potential SNPs. Five single nucleotide polymorphisms were observed within the camel insulin gene across the sequence data from the 32 camels (see Figure 3.6 which highlights the SNPs observed). Three SNP's were located within the intronic region of the C-peptide sequence whilst a further 2 SNP's were located in the 3' untranslated region (UTR) after the stop codon and before the termination signal (Figure 3.6). One additional SNP was located after the termination signal (Figure 3.6). Thus there were no SNPs located within the coding exon regions across all the camels investigated and the camel insulin coding sequence was highly conserved. The SNPs found within the intron will be absent in the mature mRNA and therefore whether these have any biological effect is unknown. However, the SNPs in the 3' UTR could potentially influence mRNA stability, turnover and translational efficiency (Proudfoot, 2011) although this was not studied in this investigation. One more SNP was found after the termination signal.

For the three identified SNP's within the intronic region of C-peptide; out of the 32 sequenced camels the first SNP was found heterozygous (GA) in 12 samples and homozygous (AA) in 5 samples, the second SNP was heterozygous (CT) in 11 samples and homozygous (TT) in 2 samples and for the third SNP 16 samples were heterozygous (AC) and only one homozygous (CC). Less SNP frequency was observed for the 3'UTR SNP's in comparison with C-peptide intron SNPs; in the first SNP only 5 samples were heterozygous (AC) and one homozygous (CC) while in the last SNP only 4 samples were heterozygous (CT) and no homozygous samples were observed.

GGGGCCGTGAGCCGAGGAGACCGTGCCTCACTGCCCCCCGCCCCGGCCTTGGACCGCCCCAGGGCCTCTCTCCCACC
CCACTCCCTGGATCAAGTGCGGAAGAGGCACGAGCGAGGCCGGAGGGGCTGGTCCCTGCGGACGCTGGCCCCCAGC
TCCGAGCTGGGGGGCTGGGCCAGCAGGCCCCTCATTAAGGCTCTAATGACGCAGCCTCAGGAGGTGCTGACGGCCA
CGGAGATGTTCCCGCACACCCAGCATCAAGGAAATGGCCCAGAAACTGCAGCCTCAGCCCCCACCCACCCCCGCCATC
TGCCACCCACTTGGAGGCCCTAAGGAGCCAGGAGGGGGCCGGGGGCTATAAAGCCCGCAGCGCGGCCCAGCAGCC
CCCAGCCCTCTGGATCAGCTGCGTTCCGAGGCTGCCAGCAAGCAGGTGTGTCCCGGAGGGCTCCGAGCTCTGCTGAC
CGGCAGCAGGGTGTGGGCCCCTCCTCCTGGAGCCCTCGGGGGTGGGCGCTGGCGGGGGTCCCGGGGGTCCCTGCAC
GGCCTTAACCCTGCCCGTCTGCCAGGCCTCGCCCCGCCGCCATGGCCCTGTGGACACGCCTGCTGGCCCTGCTGGCCC
TGCTGGCCCTCGGGGCGCCCACCCCCGCCCGGGCCTTCGCTAACCAGCACCTGTGCGGCTCACACCTGGTGGAGGCG
CTGTACCTGGTGTGCGGGGAGCGCGGCTTCTTCTACACGCCCAAGGCCCGCCGGGAGGTGGAGGACACTCAGGGTG
AGACCCCCACTGTCCCCACTGCCCCCCGGCCCCGGATCCCAGCTGGCGCCCAGGGGGCGGTCAGGAGGGGTTTTAAA
AAGGAGAATGATTCCCTCTTGGTCACATCCTCAAGCGGCCAGCTCTTTGGGGGCTGAGCCCTGACAACACCCCAAGG
GTGGGCTCGCCGTCCCCTCCACTCCCATCTTACCCCTCTCAGCCCCATTCCTTCCTCTCCTCCGACAAGAGCGGCTTCAG
GGA(G/A)GGGTTTTATGAAGTGAGTCAAGCCCTGGGGGCTGAGGGTCT(C/T)GGGGGTGCCCCATCCTCCTGCCTGC
ATGCCTGTGGGAGACACCTCACTGCCCTGAGGGGACCCCCCC(A/C)CACACACAGCCCGAGGCAGTGGGAGGGGGC
TCGGTCCTGCACAGGCTGCGGGGAGGTGGGTGCGGGTGGGGGTGGGCGGGAGACAGGCCCCGCGACGGGCACAG
AGGATGCCCGCCCGTTCAGGGTCCCGCTCCAGGGCTGGCTGTCTCTCGGCAGTGGGCGGCGTGGAGCTGGGTGGAG
GCCCGGGTGCGGGCGGCCTGCAGCCCCTGGGCCCGGAGGGGCGCCCGCAGAAGCGCGGCATCGTGGAGCAGTGCT
GCGCCAGCGTCTGCTCGCTCTACCAGCTGGAGAACTACTGCAACTAGGCGGCCGCCCCCGACCCGGC(A/C)CCCCGG
CCCCC(C/T)GCGCCCCCGCTCTGCAATAAACCCCCTGAATGAGCCCTGCTGGTGTCGTCTGTGTGGCCTGGGGGCTGC
GGGGCTGGGGGAAGGGGGCGAGGAAGGCTGGGCCTACCTGCCCAAGACCCCCTTCTGCTCTCTCTGCGCTGGCCCC
TCCAAACTTTCTCTACCTGCACGGGCGCCCACGGGCACGTGGGCGGACGGGGACGTGGGCCCGCGGCAGTGCCCCC
CAAGGAGGCGAGCTCCTGTCCTCAAGGGTCCTGACGCCCCCCTCCCCAGCCCCCATGCCGCACACGGCGGCTCCGTG
GGCCGCCCGCCCCGCTGCTGCTGGCAGACCCGAGGGCCCCAGGG(T/C)TGCCTGGGCTGGCCGTGGCCTGGGAGGT
CAGAGGGTGACTTGGCTTACCAAGGCCTG

Single Nucleotide polymorphisms (SNP's)
Goldberg-Hogness Box (TATA Box)
Signal peptide
B-Chain
Cleavage site
C-peptide
A-Chain
Stop Codon
Termination signal

Figure 3.6. The full sequence of camel insulin gene as determined by alignment of the sequence data from 32 camels. The sequence is annotated using colour to define the signal peptide (blue), which contains the start ATG codon, the A/B/C chains, the TAG stop codon, the termination signal and the TATA box. The SNPs observed are also highlighted with 5 being located within the insulin gene and one beyond the termination signal.

3.3.4 Taxonomic Alignment and 3D Homology Modelling of Camel Insulin

Based on the genomic sequence of the camel insulin gene obtained and reported in section 3.2.3 above, the cDNA sequence of camel preproinsulin (signal peptide, A-chain, B-chain and C-peptide) was compared with human (Sures *et al.*, 1980) and bovine (D'Agostino *et al.*, 1987) preproinsulins using DNAMAN software (Figure 3.7). The analysis showed a number of changes at the DNA level between the three species. Interestingly, the changes in the A- and B-chains are mainly non-sense and occurred mostly in the third nucleotide of the codon and did not change the corresponding amino acid sequence. Thus, in the B-chain there is only a single amino acid difference between camel and bovine while 4 nucleotides differences exist and there were 12 nucleotide differences between camel and human but only 2 amino acid changes. At the protein level of the A-chain there is no amino acid difference between the camel and bovine chains while the alignment shows 2 nucleotides differences (Figure 3.7). Two amino acids differences are present between the human and camel A-chain and seven nucleotides changes are present.

In the case of signal and C-peptides there were changes distributed among and across the sequence. In the signal peptide there are 15 nucleotide differences between human and camel with 6 amino acid changes and 9 nucleotide differences between camel and bovine with 4 amino acid changes. In the C-peptide there are 20 nucleotide differences between human and camel with 8 amino acid changes. It is noted that the bovine C-peptide is 5 amino acids shorter than the human and camel equivalent due to a 15 nucleotide deletion (D'Agostino *et al.*, 1987) within the C-peptide sequence. Because of this the bovine sequence was not included in the C-peptide alignment with human and bovine shown in figure 3.7.

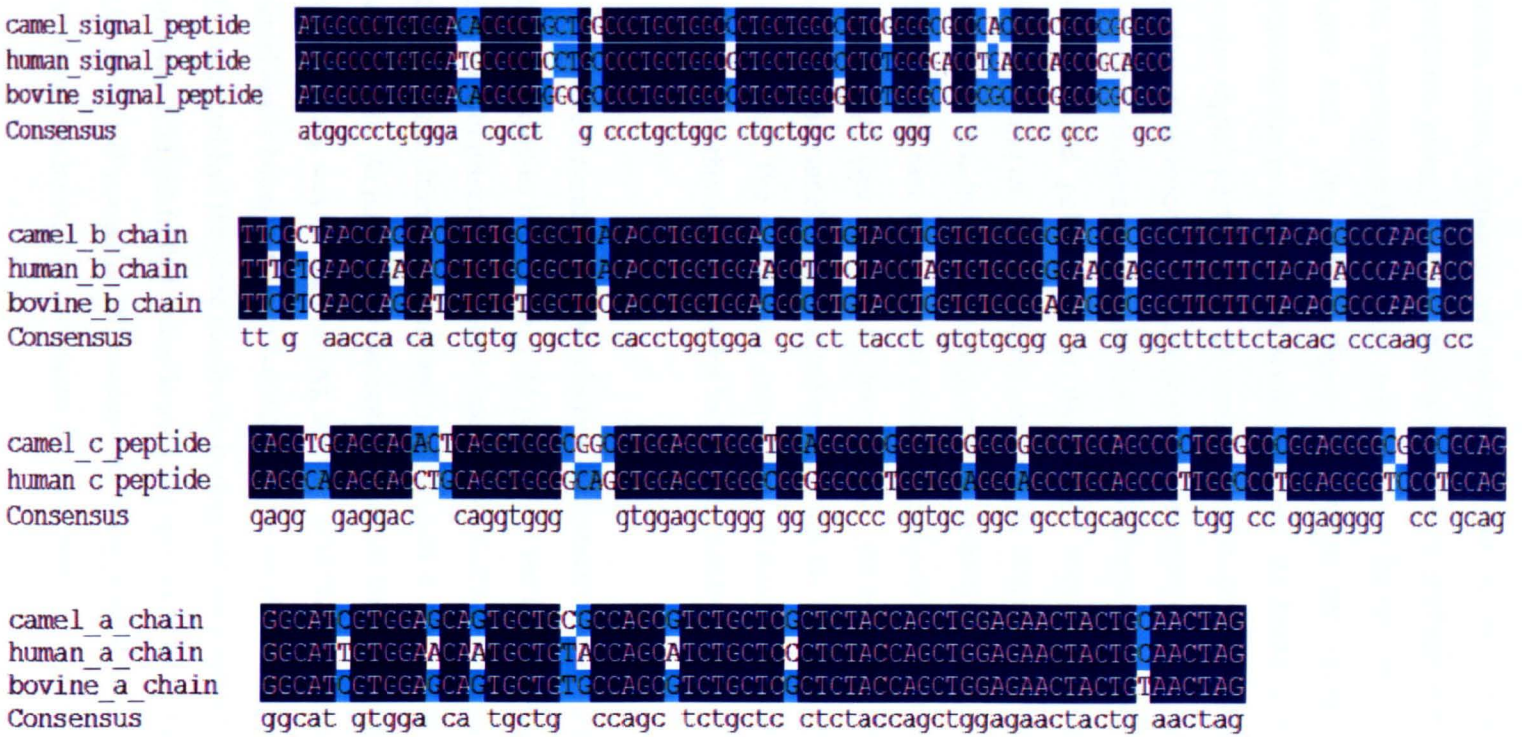


Figure 3.7. DNA sequence alignment of preproinsulin chains (Signal peptide, B-chain, C-peptide and A-chain) between camel, bovine and human, the alignment of c-peptide shows the camel and human sequences only because in bovine the c-peptide has a 5 amino acid deletion.

Knowing the camel insulin genomic sequence it was possible to predict the camel insulin DNA sequence by undertaking and *in silico* transcription and translation prediction using the online available software package http://insilico.ehu.es/translate/. The resulting predicted amino acid sequence for the camel insulin gene is reported in figure 3.8. The data in figure 3.8 reports the amino acid sequence of camel preproinsulin in comparison to other species. As found in the other species, the camel insulin signal peptide was 24 amino acids in length. The signal peptide is the first section of the nascent polypeptide chain that emerges from the ribosome and is bound by the signal recognition particle (SRP) which arrests translation elongation and directs the ribosome and polypeptide to the ER by recognising SRP receptors on the ER (Hatsuzawa *et al.*, 1997). Signal peptides are characterised by their highly hydrophobic nature and usually contain a high number of leucine residues (Zanen *et al.*, 2005) as is found in the case of the camel insulin signal peptide and the other sequences this was compared to (see Figure 3.8). Interestingly, there were a large number of differences between the camel insulin signal peptide and the other insulin signal peptides this was compared to. The camel has a unique alanine residue at position 9 in the signal peptide whereas all other signal peptides had a proline in this position (Figure 3.8).

There were a number of other obvious differences between the camel signal peptide and the other species investigated. At position 17 the camel sequence has a glycine residue that was not present in any other species, most of which had a tryptophan (W) residue in this position although rabbit and squirrel had a cysteine (C) and leucine (L) residue respectively. At position 20 the camel sequence has a threonine (T) residue which in the other species was an alanine (A), aspartic acid (D) or asparagine (N) residue, thus the camel again being unique in having a threonine (T) residue at this position. At position 22 the camel had the same residue as in dog, cat and bovine (an arginine (R) residue) but this was different to that found in all other species where either a glutamate (Q), alanine (A), proline (P) or serine (S) residue was found. Finally there were also other differences between species across the signal peptide, notably at positions 5 and 18. There was

therefore considerable diversity in the insulin signal sequence amino acid sequence with the camel signal sequence having several unique residues compared to the other species investigated.

Compared to the variability observed in the signal peptide sequence of insulin, the sequence of the B-chain was highly conserved across the species compared (Figure 3.8). The B-chain is 30 amino acids in length with the most obvious change being a single unique change between camel and all other species investigated at position 2 of the B-chain whereby the camel sequence contained an alanine (A) residue and all other species a valine (V) residue (Figure 3.8). The other notable feature of the B-chain was that camel, as in bovine, dog, cat and pig insulin contained the presence of an alanine residue in the last amino acid of the B-chain which was either a threonine (T) or serine (S) residue in all other species. The A-chain, like the B-chain was highly conserved across the species investigated. The A-chain is 21 amino acids in length and there are two sites that show prevalence for change in this section of the protein. These two differences in the A-chain are found at positions 8 and 10 (threonine to alanine in camel) and (isoleucine to valine in camel), these 2 differences also existing in cat and bovine insulin (Figure 3.8).

Finally, there was much diversity in the amino acid sequence of the C-peptide across the species investigated (Figure 3.8). The C-peptide is 31 amino acids in length and among the many differences in C-peptide there were 5 unique to the camel C-peptide sequence. , These were an alanine to glycine and leucine to proline substitutions at the 25^{th} and 26^{th} positions of the C-peptide, respectively. In the camel C-peptide there are also threonine (T), glycine (G) and R residues found at positions 5, 9, and 29 respectively whilst various other residues were present in these positions in the other species investigated (Figure 3.8). At residue 30 in the camel there was a proline (P) residue, also found in bovine and pig whilst this was a different residue in all other species. There were various other sequence differences across the C-peptide between the species investigated showing that the C-peptide, much like the signal peptide, is not well conserved across the species present and shows considerable species specificity.

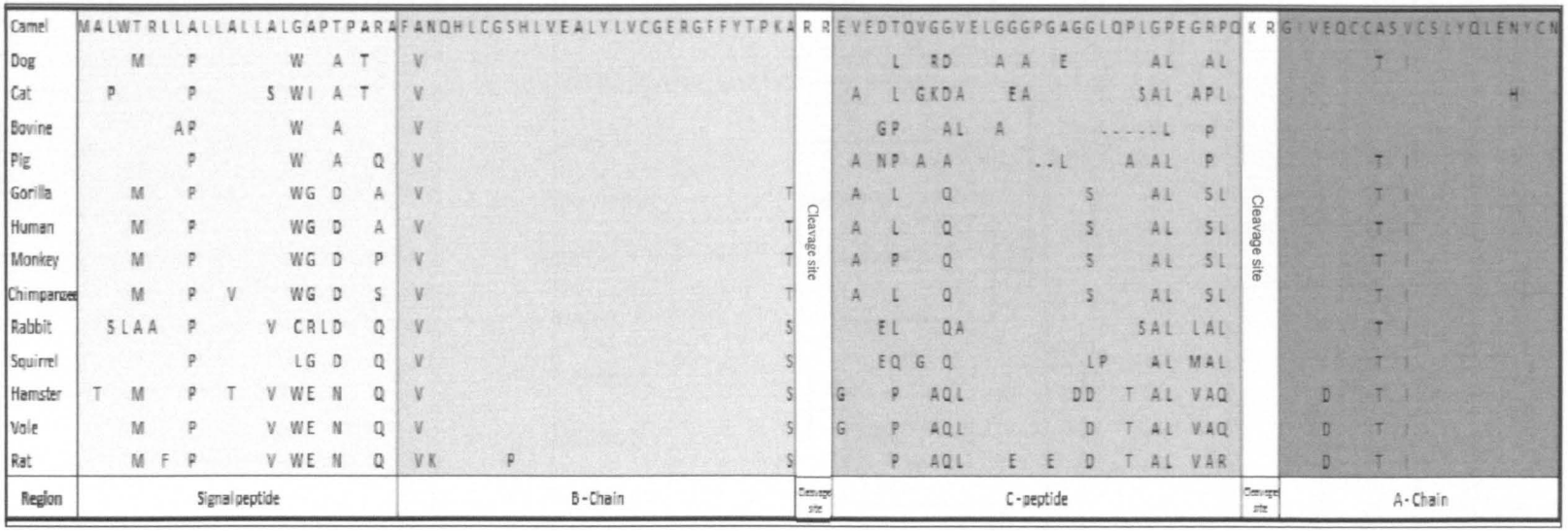


Figure 3.8. The amino acid sequence, domain/chain regions of preproinsulin and the sequence alignment of the deduced camel sequence with 13 different species. The preproinsulin sequences were obtained from (Kwok *et al.,* 1983; Bajaj *et al.,* 1986; Buck *et al.*, 2001 ; Stoy *et al.,* 2010).

Using the sequence data presented in Figure 3.8, a homology tree was created using DNAMAN software version 3.2.2) and the resulting diagram is shown in figure 3.9 below. Based on the results of the drawn homology tree (Figure 3.9) of camel preproinsulin, the dog preproinsulin is the closest to camel while rabbit preproinsulin has the most number of differences and is the less homologous to the camel preproinsulin.

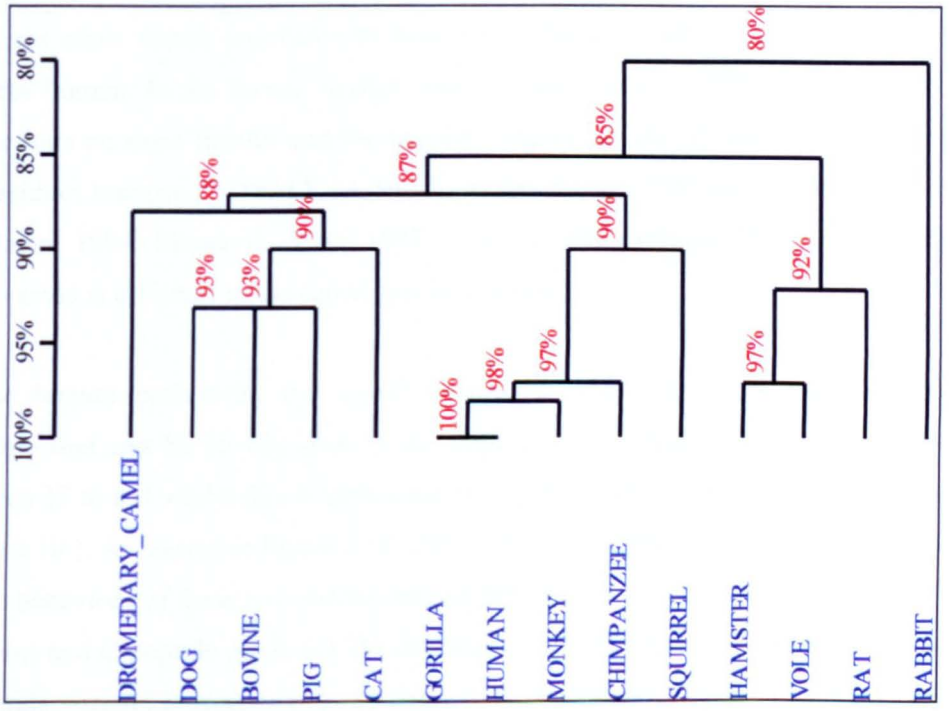


Figure 3.9 The homology tree of the species described in figure 3.8, prepared using DNAMAN software. The percentage numbers refer to the percentage homology at the amino acid level between species.

Using the amino acid sequence deduced from the genomic DNA sequence of camel insulin the 3-dimensional structure of the molecule was predicted using commercially available software (figures 3.10; 3.11; 3.12). The 3D-JAIGSAW software works by comparing with protein models available in the Protein Data Bank (PDB), PFAM and structural classification of Protein (SCOP) databases. The predicted 3-dimentional structure of both camel and human proinsulin, based upon the amino acid sequences, highlights predicted structural differences between the two species. The predicted structures show camel proinsulin to have more alpha helical structure in comparison with the human. In the human insulin structure the residues which are involved in the interactions between insulin and the insulin receptor are the 4^{th} and 5^{th} of the A-chain and residues number 12, 16, 17, 24, and 26 of the B-chain (DeMeyts *et al.*, 1990; White and Kahn, 1994; Mynarcik *et al.*, 1997 ; Yip and Ottensmeyer, 2003). None of these amino acids is different in the camel insulin molecule.

Unlike human proinsulin, the camel proinsulin molecule has a predicted α-helix structure that can be clearly seen in the upper loop of camel proinsulin (amino acid residues 27 to 61) whilst this is unstructured or β-sheet like in in human equivalent (see figure 3.10). As shown in figure 3.10, this could potentially have a very dramatic effect on the behaviour of these two molecules and their stability. Figure 3.11 shows the A and B-chains and C-peptide predicted 3D structure of human and camel proinsulins were the C-peptide appears to have mainly an alpha helical structure. Figure 3.12 shows that the existence of alanine in the junction between the B-chain and C-peptide in camel instead of threonine in human generates a predicted proinsulin molecule with more coiled structure. Furthermore, the alanine in the second position of the camel proinsulin results in a predicted structure that is twisted at the beginning of the camel proinsulin more than the human.

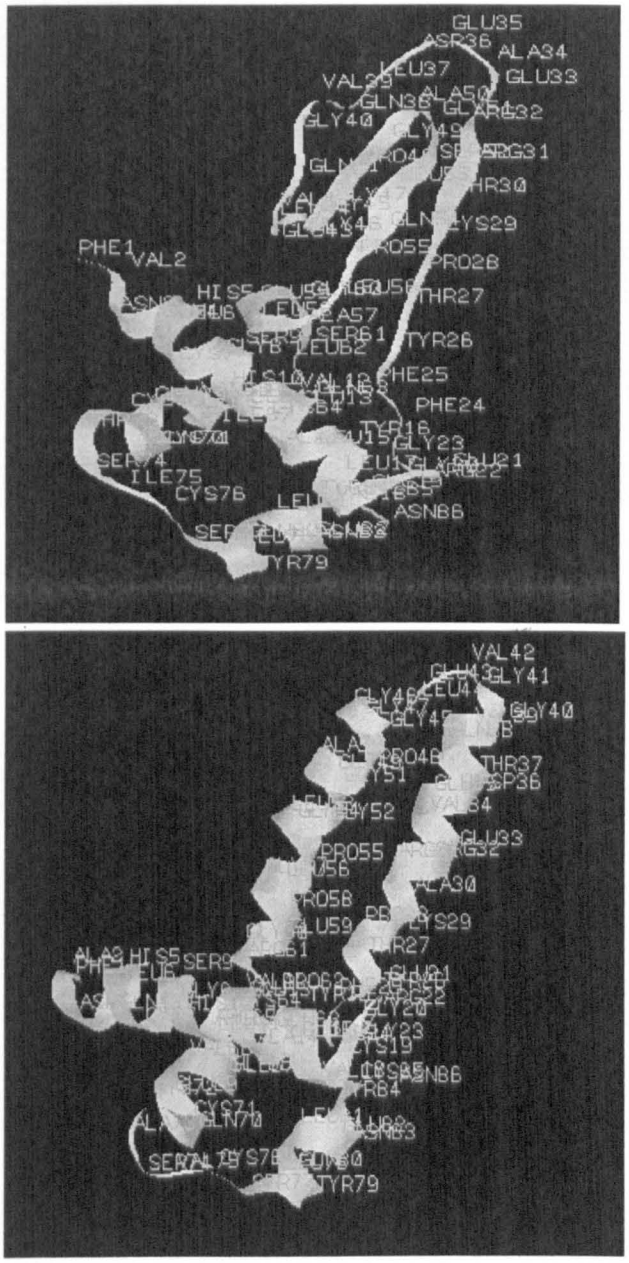


Figure 3.10. The predicted three dimensional (3D) structure of human proinsulin (upper figure) and camel proinsulin (lower figure) using 3D-JIGSAW comparative modeling server. The models were generated and viewed using the freeware RASMOL software.

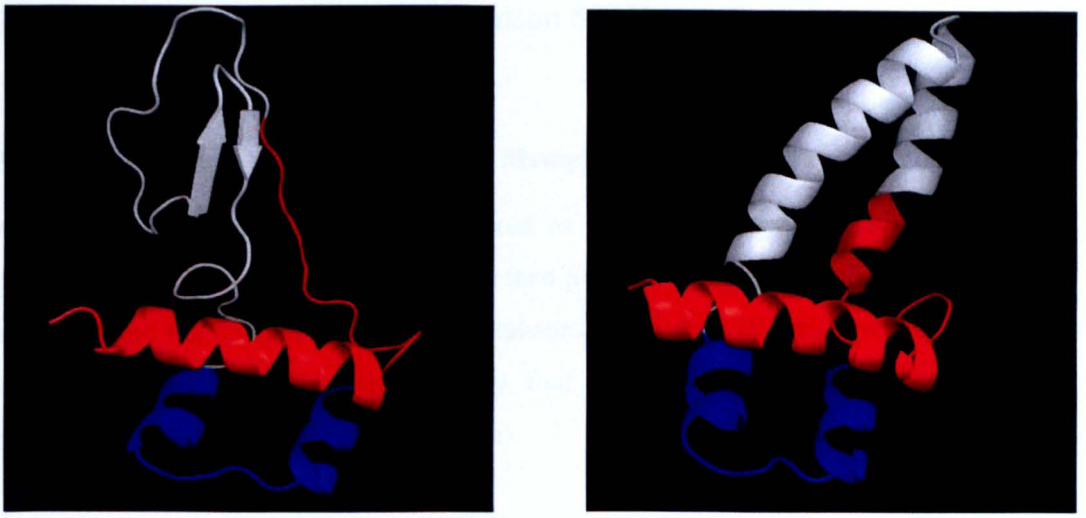


Figure 3.11. The predicted three dimensional (3D) structure of human proinsulin (left) and camel proinsulin (right) using 3D-JIGSAW comparative modeling server, the models being viewed by pymol software. In both proinsulins, the A-chain is shown in blue, C- peptide in gray and B-chain in red.

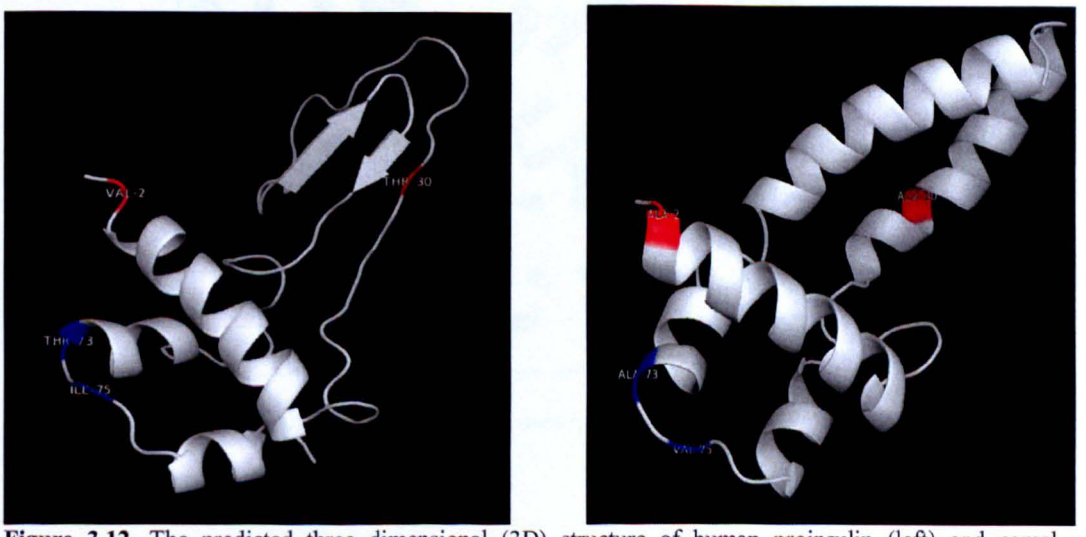


Figure 3.12. The predicted three dimensional (3D) structure of human proinsulin (left) and camel proinsulin (right) using 3D-JIGSAW comparative modeling server, the models being viewed by pymol software. The predicted differences in the A- and B-chains between human and camel are shown in blue for the A-chain and red for the B-chain.

3.3.5 Camel Insulin Gene Localization Studies

3.3.5.1 Cell Culture of Camel Cells and Metaphase Preparation

A suspension of camel cells was prepared as described in the methods section and metaphase cells harvested. The cells were then placed on slides and stained with Giemsa stain to score the number of cells in metaphases. More than 10 well spread metaphases were found on each slide and seventy four chromosomes were counted in each metaphase (see figure 3.13 for an example).

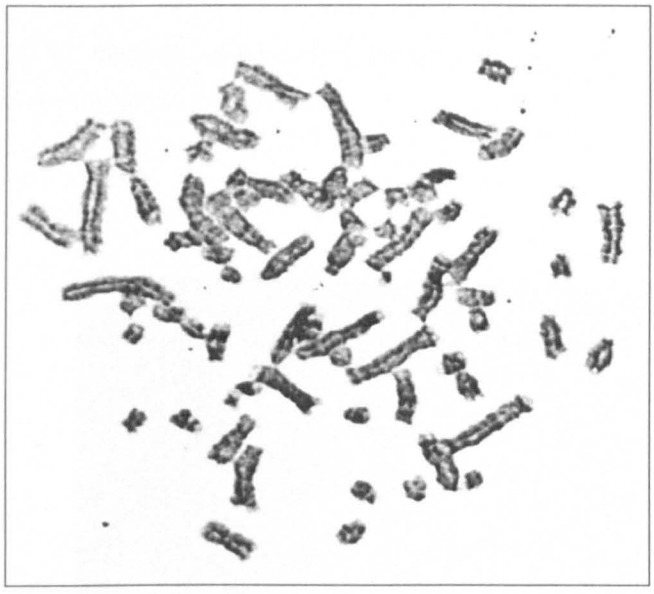


Figure 3.13 Metaphase chromosomes of dromedry camel.

3.3.5.2 FISH Analysis of the Camel Insulin Gene with a PCR Amplified Product Probe

The PCR product of the camel insulin gene was labelled by two methods for FISH analysis. The first method used a biotin nick translation technique. When this labelling method was used, for subsequent hybridisation with the labelled probes there was more than 2 signals on the telomeric regions of a number of chromosomes (Figure 3.14), however, after increasing the wash stringency, no signals were observed whatsoever. As such this approach was not considered appropriate to visualise and analyse the positioning of the camel insulin gene. A second labelling approach was also attempted with a DIG-PCR labelled probe and in this case no signals were once again observed. This approach to the mapping of the camel insulin gene was therefore not successful in identifying the location of the insulin gene.

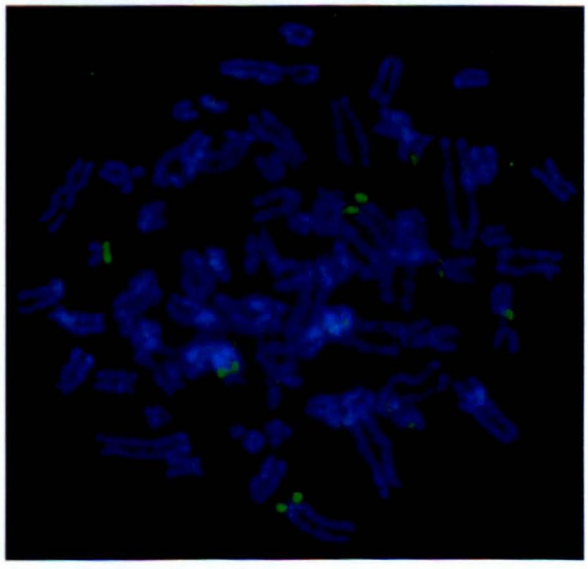


Figure 3.14 FISH results using FITC avidin labeled by nick translation technique.

3.3.5.3 FISH Analysis of the Camel Insulin Gene using a BAC Clone

As the PCR probe approach to the FISH analysis of the insulin gene was not successful, FISH was repeated using *Llama pacos* CH246-515B17 BAC clone. Thirty micrograms of BAC DNA was isolated and used for FISH analysis. A band matching the camel insulin gene band was observed after amplifying the BAC DNA using camel specific insulin primers, confirming the presence of the camel insulin gene in the BAC clone (figure 3.15). When the DNA was used for labelling, 0.3 µg was used for the labelling procedure. After labelling and agarose gel electrophoresis analysis the resulting bands were observed with smearing of the labelled product between 150-250 bp as shown in figure 3.16.

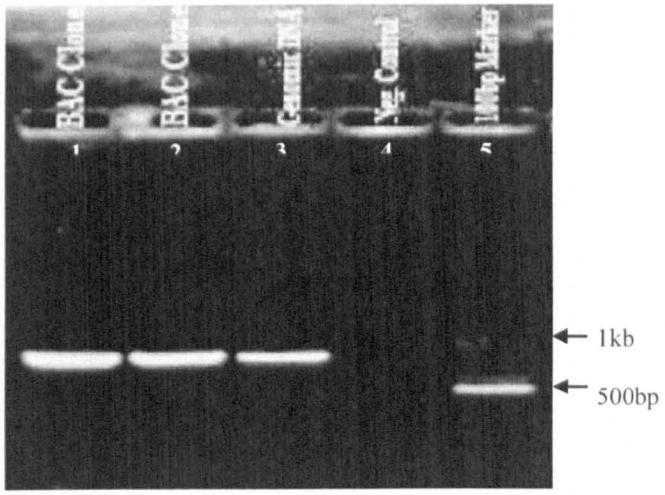


Figure 3.15 Agarose gel electrophoresis anlysis of the PCR amplification of the camel insulin gene from the BAC (lane 1 and 2) clone and control genomic (lane 3) DNA. Negative control loaded in lane 4 and 100 bp size weight marker in lane 5

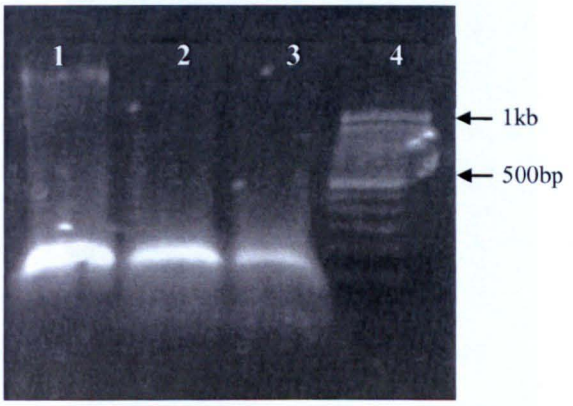


Figure 3.16 Agarose electrophoresis of random priming labeled product. Lanes 1-3 the product labelled, lane 4 is the 100 bp size marker..

A good number of well spread metaphase camel cells were then observed on the slides which were used for the FISH experiment. After hybridizing the slides with the labelled BAC probe, two strong and sharp specific cy3 signals were observed on the interphase lymphocytes (figure 3.17 and 3.18).

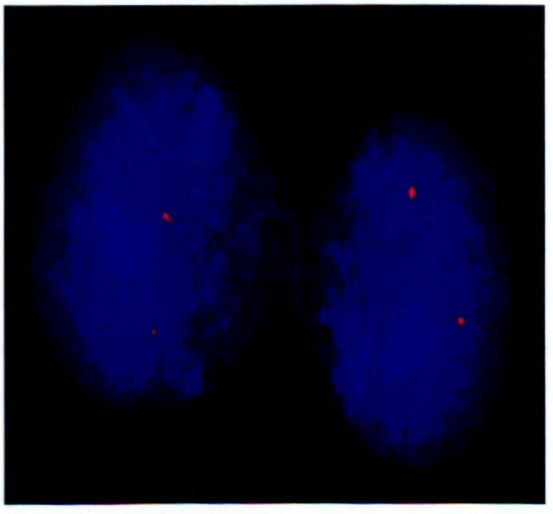


Figure 3.17 FISH analysis for camel insulin gene using a BAC probe showing two interphase cells of dromedery camel lymphocytes labled with cy3 avidin.

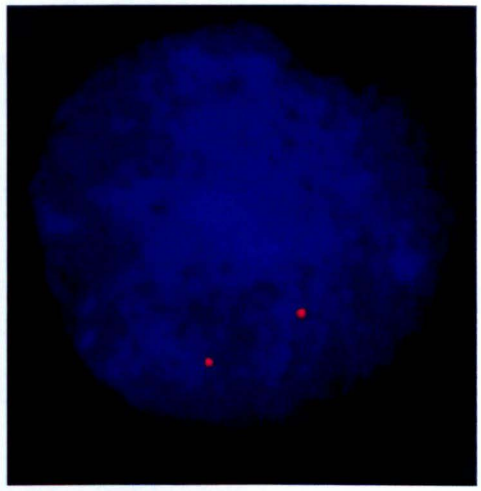


Figure 3.18 FISH analysis for camel insulin gene using a BAC labelled probe on interphase cells of dromedery camel lymphocytes labled with cy3 avidin.

Strong and sharp Cy3 signals were also observed on one homologous pair of chromosomes as shown in figures 3.19A, 3.20, 3.21 and 3.22A. Once appropriate labelling and FISH images had been obtained it was necessary to identify upon which chromosome(s) the probe was hybridising to and hence identify the chromosome location of the camel insulin gene. To identify the chromosomes slides were stained after the FISH analysis with propidium iodide and DAPI. The two dyes gave a uniform fluorescence of chromosomal arms and heterochromatin regions which produced a banding pattern similar to the GTG banding as in figures 3.19b, 3.20, and 3.22b.

To identify the chromosome where the camel insulin gene is located, the karyotype of the first 16 camel chromosomes along with the X-chromosome (figure 3.19c) was prepared using Adobe Photoshop software according to the published camel karyotype. By comparison with the published camel karyotype, and according to the chromosome size, banding pattern, and centromere position, it was deduced that the insulin gene of dromedary camel is localized on the distal (Telomeric) end of the q-arm of chromosome 10.

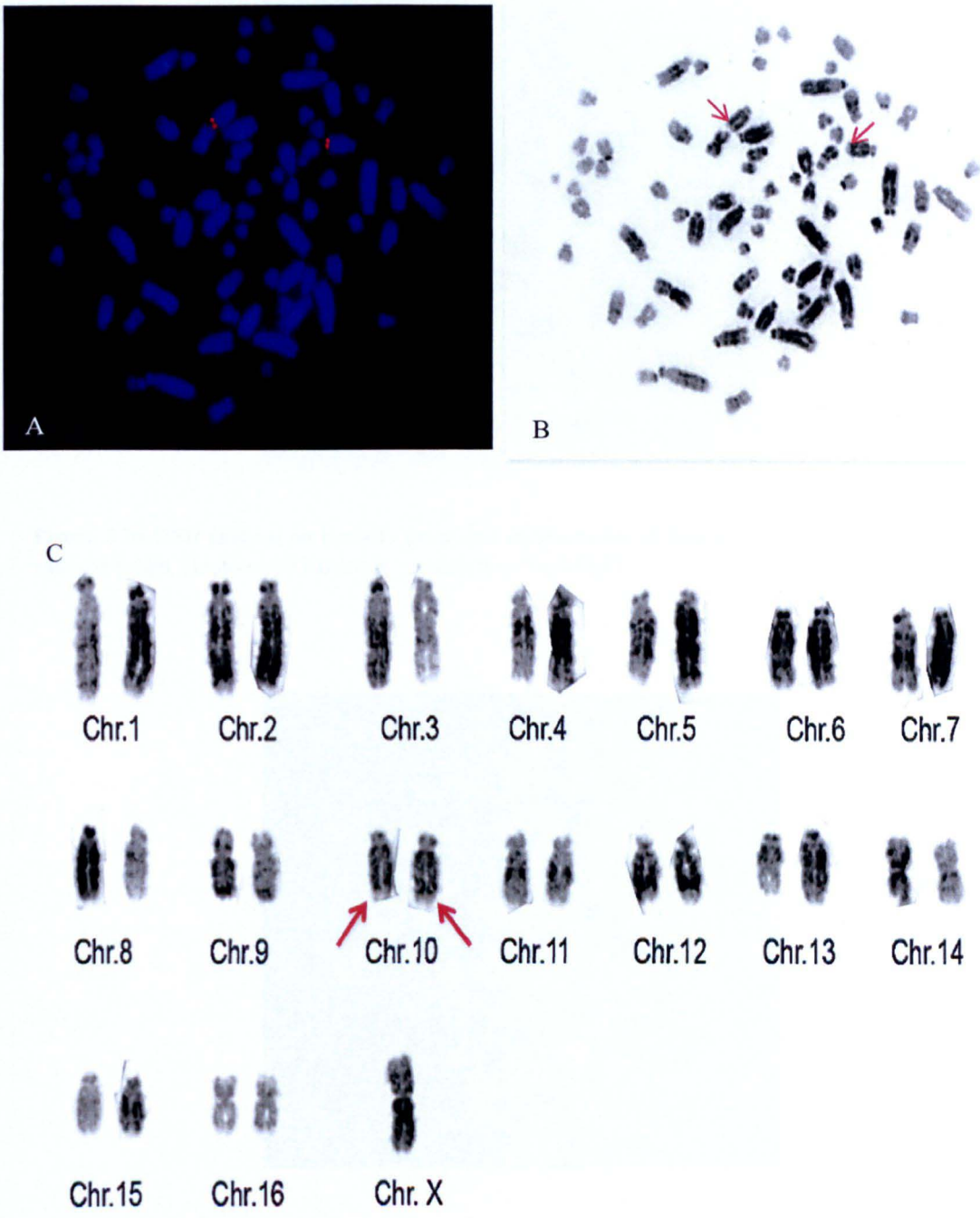


Figure 3.19 (a) FISH analysis on metaphase of dromedery camel lymphocytes labelled with cy3 avidin. (b) banded metaphse chromosomes stained and banded by propidium iodide/DAPI. (c) the karyotype of first 16 camel chromosomes of (b) metaphase, chromosome X is also identified. The arrow shows the location of the insulin gene.

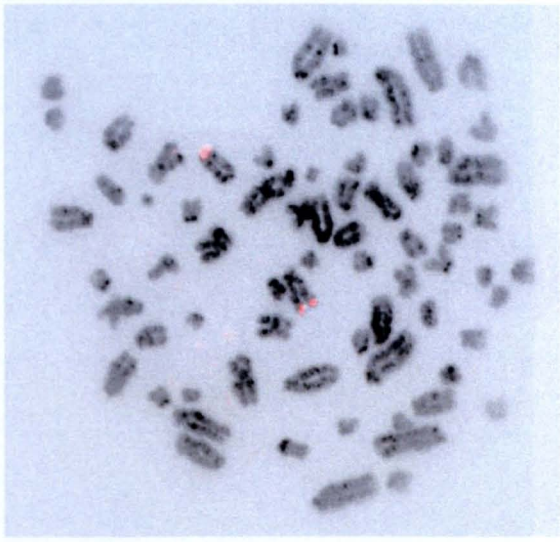


Figure 3.20 FISH analysis on banded metaphase chromosomes of dromedery camel lymphocytes labelled with cy3 avidin, chromosome banded by propidium iodide/DAPI.

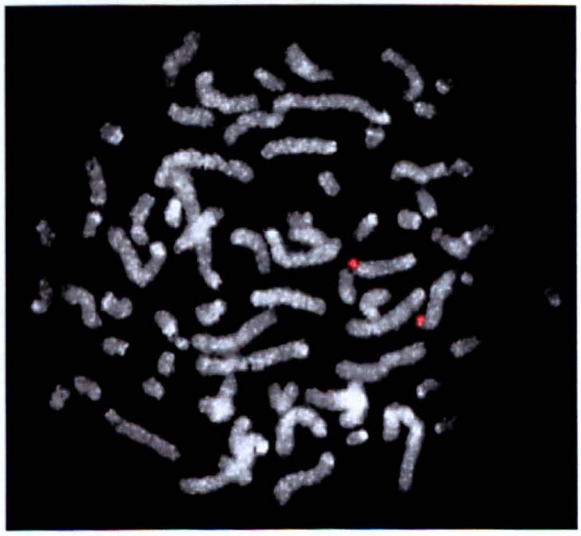


Figure 3.21 FISH analysis on metaphase chromosomes of dromedery camel lymphocytes labled with cy3 avidin.

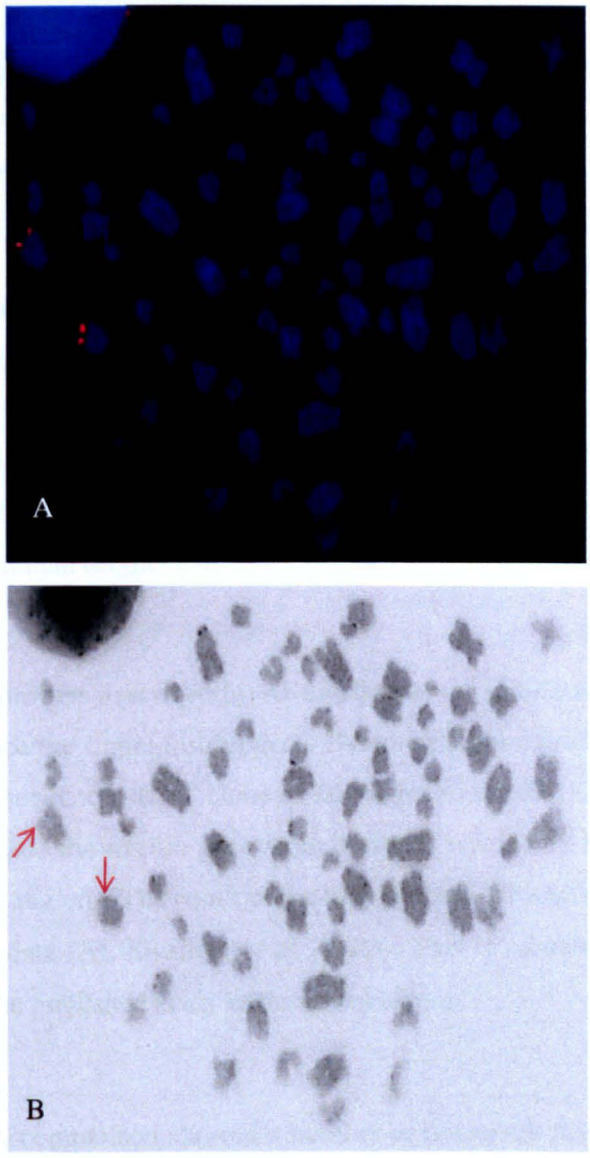


Figure 3.22 (a) FISH analysis on metaphase of dromedery camel lymphocytes labled with cy3 avidin. (b) the banded metaphse chromosome as shown by propidium iodide/DAPI. Arrows show the insulin gene location.

3.4 Discussion

In this chapter the genomic DNA sequence of the camel insulin gene was deduced and compared to other species. Following this FISH approaches were used to characterize and locate the chromosomal positioning of the camel insulin gene.

In 2008 Al-Swailem and colleagues amplified the camel proinsulin cDNA using primers designed initially to amplify the human insulin gene (Al-Swailem *et al.*, 2008). Al-Swailem used a forward primer which hybridized to the B-chain and a reverse primer which hybridized to the A-chain. Due to the sequence difference in the B-chain between human and camel, the start region of the camel published sequence (see figure 3.7 and section 2.4.2) is of human origin.

In this project, the primers described by Al-Swailem *et al.* (2008) were initially used to amplify and sequence the camel insulin gene. Based on the sequencing results obtained, which matched the sequenced BAC clone of *Llama pacos* number CH246-515B17, new primers, which flanked the insulin gene were designed to identify the full insulin gene and the surrounding regions. The result of the new primers showed sequence differences with the published data (AL-Swailem *et al.*, 2008). This is probably due to the use of human primers in the published study as described above.

The results of DNA comparison showed a number of non-sense nucleotide changes that did not affect the amino acid composition of the molecule when compared to other species. Insulin is an important cellular protein and changes in the protein structure that affect the function of the insulin molecule would have potential function consequences. This probably explains why the A and B-chains are conserved among species while the signal and C-peptides (which are not directly affecting the binding and the function of the protein) are polymorphic among the species investigated.

To confirm the sequencing results and to identify any single nucleotide polymorphisms (SNP's), the camel insulin gene along with the flanking regions of 32 camels was sequenced and aligned for comparison. After aligning the insulin sequences of the 32 camels, five SNP's were discovered, three of them in the intronic region of the C-peptide and two SNP's in the 3' untranslated region of the gene. After the stop codon and before the termination sequence, these two SNP's may play a role in gene expression, particularly at the level of translational control.

The predicted translated amino acid sequence of camel insulin gene showed a number of differences when compared with other species as shown in Figure 3.8. The existence of alanine instead of valine in 2nd amino acid of the B-chain only in camel may play a role in the stability of the mature insulin molecule. For example, Dao-Pin *et al.* (1990) showed higher thermostability of protein T4 phage lyoszyme after valine had been substituted to alanine. The authors suggested that the gain in thermostability could be due to reduction in strain which is caused by the rigidness of an alpha helix containing valine and found that the alaine T4 phage lysozyme mutant is more thermostable than the wild type. In a further study Lyu and co-workers (1990) studied the stability of alpha helical structures formed by different amino acids. The study was performed by substituting a single residue in a polypeptide sequence with different amino acids using a chemical peptide synthesis technique. In all cases the alpha helical stability was higher in alanine containing helicies than in valine containing helices. This unique amino acid change in camel insulin might therefore result in a more stable insulin molecule than in other species and could, at least partially, explain the higher stability of camel insulin observed in other studies compared to that from other species (Wernery *et al.*, 2006a).

The stability of the zwitterionic form of different amino acids has been extensively studied, but of particular interest is an investigation by Wu and McMahon (2007). The authors studied isomers of each amino acid and of relevance to this study concluded that

the thermochemical stability of alanine isomers is higher than valine isomers (Wu and McMahon., 2007).

In comparison with human insulin (B-chain and A-chain), camel insulin has 4 amino acid differences. Of these differences three out of the four are changes to alanine residues. The first substation is in the 2nd residue of the B-chain and is a change from valine to alanine as has been discussed above. The second and third differences are in the last amino acid of the B-chain and in 8th residue of A-chain and both are threonine to alanine substitutions. According to Lyu *et al.* (1990), alanine residues are more alpha helically stable than valine and threonine. Only the 4th camel substitution results in a change to a less alpha helically stable amino acid (isoleucine in human to valine in the camel). The inclusion of 3 additional alanine residues in camel insulin compared to the human insulin sequence suggests that camel insulin contains more alpha helical stable residues and hence the structure of camel insulin may be more alpha helical and that these helices are of a greater stability. This would be required to be tested further and could be done so by sequential point mutation of the human amino acids to alanine and to complement this mutate the camel residues sequentially to valine residues and determine the effect on helix formation and insulin thermostability. Further, the alpha helix structure is known to be more stable than beta sheets (Chakrabartty & Baldwin, 1995), (Serrano & Fersht, 1989) (Nicholson *et al.*, 1991) and hence these observations may, at least partially, explain the greater stability of camel insulin compared to its human counterpart.

Many publications have reported on studies around the interaction between human insulin and it's receptor (DeMeyts *et al.*, 1990; White and Kahn, 1994; Mynarcik *et al.*, 1997 ; Yip and Ottensmeyer, 2003). The results obtained here for camel insulin indicate that none of the unique amino acids that exist only in camel insulin are involved in the interaction between the insulin and insulin receptor and hence this function is most likely conserved in camel as in other species. The second and last amino acids in the B-

chain may have an important role in stabilising the insulin molecule by providing the insulin molecule more helical structure.

In agreement with the substitution with alanine residues discussed above, the predicted 3-D structure of camel proinsulin showed more alpha helical structure than beta sheet in comparison with the human proinsulin structure (Figure 3.12). Proinsulin is the precursor of insulin and plays an important role in controlling the structure of the mature insulin protein (Orci, 1986). The .C-peptide plays a vital role in insulin biosynthesis as it links the A- and B-chains in a manner that allows correct folding and assists the formation of disulphide bridges (Steiner and Tager, 1995, Wahren *et al.*, 2000). The C-peptide region falls between two conserved regions of insulin A- and B-chains. The DNA sequence of the C-peptide was found to be polymorphic among the species investigated. The alpha helical predominant structure of the C-peptide in camel may play a role in protecting the proinsulin protein. Thus, if the camel has a more stable precursor molecule this may result in more mature insulin being produced with the correct folding although this hypothesis remains to be tested. Finally, insulin production is a complex process that includes the control of transcriptional and post-transcriptional mechanisms (Permutt and Kipnis, 1972; Bhatt *et al.*, 2010).

With regard to the signal peptide, the 9th amino acid in the signal peptide of camel preproinsulin is alanine. The signal peptide binds to the signal recognition particle (SRP) which arrests translation elongation and directs the ribosome and polypeptide to the ER by recognising SRP receptors on the ER (Hatsuzawa *et al.*, 1997). Signal peptides are characterised by their highly hydrophobic nature (Zanen *et al.*, 2005). Alanine is a hydrophobic amino acid while proline is ambivalent so this could enhance the signal peptide binding to the SRP receptor (Biswas *et al.*, 2003). Furthermore, alanine residues can help form a stable alpha helical structure (Lyu *et al.*, 1990), while the proline restricts the conformation and disfavour the alpha helical structure (Williamson, 1994) and it is well-known that the signal peptides can insert impulsively into lipid membranes

and adopt a predominately alpha helical structure (Zheng and Nicchitta, 1999). Thus, the alanine change in the camel signal peptide may have an influence on how the polypeptide is transported to and transferred into the ER.

Here, FISH analysis was also undertaken to determine the localization of the camel insulin gene on camel chromosomes. Initially, an amplified PCR product was used as a probe to hybridise with camel chromosomes. The results showed many nonspecific and weak hybridisations onto the camel chromosomes. This could be due to the length of the PCR product being too small to give a strong signal. Using this PCR product it was also observed that there was strong signals on four chromosomes. These extra signals could be the result of the hybridisation of insulin psedogenes or insulin like growth factor. Therefore, FISH was performed using a *Llama pacos* BAC clone CH246-515B17 (longer DNA sequence), which contained the insulin gene. The hybridization with this BAC clone gave a specific and clear result. Sharp and unambiguous signals were observed on the telomiric region of the q arm of camel chromosome 10. The human insulin gene was previously localised on the distal end of the short arm of human chromosome 11 (Harper *et al.*, 1981). Balmus *et al.* (2007) studied the camel chromosomes and compared between camel, cattle, pig and human chromosomes by using a cross-species chromosome painting technique. Interestingly, the results showed that chromosome 11 in human is similar to chromosome 10 in camel. This finding supports the results presented here for the localization of the camel gene and provides further evidence that the gene location has been correctly identified in camel i.e. the insulin in human is at the distal end of p-arm while in camel it is at the distal of the q-arm.

3.5 Conclusion:

From the results obtained in this chapter it can be concluded that, the full length of camel insulin gene is 772 bp including intronic region accounts for 507 bp and the sequence matched (using the BLAST search) a sequence from *lama* with 96% homology. Three different SNPs were identified in the none-coding sequence of C-peptide intron, 2 in the 3'-UTR and 1 after the termination signal. Across 13 species compared with corresponding amino acid sequence, the camel insulin has two unique amino acid changes in the signal peptide, one unique amino acid change in the B-chain and two changes in the A-chain compared to human (which are observed in some other species). Many variations were observed across species in the C-peptide region of camel insulin compared to other species. The 3D homology modeling of camel proinsulin showed more alpha helical structure compared to the human insulin which may influence the molecule stability. Ultimately, the camel insulin gene was mapped to the distal (Telomeric) end of the q-arm of camel chromosome 10.

Chapter 4

Purification, Characterization and Identification of Camel Milk Insulin

4.1 Introduction

In 1922 Insulin was isolated from pancreas for the first time by Banting and Best (Banting and Best, 1922). In the same year the isolated insulin was successfully used in medical applications. An insulin extract was injected into a 14 year old diabetic patient. The extract was effective in reducing the blood glucose level and improved the patient's condition (Banting *et al.*, 1922). Ever since there has been considerable research and medical interest on the control of insulin expression, the isolation of insulin for medical use, and subsequently the expression and purification of recombinant insulin for administration to patients. The current state of insulin research is such that so called 'designer insulin' is now available with tailored medical purposes (Danne and Bolinder, 2011).

The discovery of insulin and the subsequent understanding of its role in glucose metabolism led to its medical applications and was one of the leading medical innovations of the last century. As such, there have been many studies undertaken to understand the function, structure, and synthesis of insulin. In the last few decades this has become increasingly important with the global rise in the number of people suffering from diabetes and who require insulin treatment to control this condition (Mudaliar and Edelman, 2001).

As described in the introduction chapter of this thesis, the mature insulin molecule is a 51 amino acid long peptide hormone. The mature insulin hormone consists of two polypeptide chains that are covalently attached whereby two disulphide bonds connect the A- and B-chains. Like other peptide hormones, insulin matures through a proteolytic cleavage modification from proinsulin. The proinsulin molecule consists of three parts; the A-, B-chain and C-peptide (see introduction chapter for a more indepth discussion on the molecule). As a result of the proteolytic modification, the C-peptide which is the central portion of the proinsulin molecule is cleaved and then released in equimolar amounts with the mature insulin (Orci, 1986). As described in the introduction, the mature insulin molecule plays a vital role in many metabolic processes by enhancing the storage of glycogen in liver and increasing the glucose uptake in muscle and adipose tissue which subsequently reduce the blood glucose levels. Insulin is also involved in the membrane transport of amino acids and various ions and stimulates the synthesis of various RNA and protein molecules (Steiner, 1977; Bhatt *et al.*, 2010).

Camel milk contains approximately 52 microunits/ml of insulin and it is thought that this is probably why there is a lower requirement or dependence for insulin in diabetic patients receiving camel milk (Agrawal *et al.*, 2005). These relatively high levels of insulin in camel milk and the resulting insulin-like activity are proposed to elicit regulatory and immune modulatory effects on beta cells in patients and partially explain camel milk anti-diabetic activity (Agrawal *et al.*, 2005). Surprisingly, Agrawal *et al.* (2007a) found that camel milk can help in reducing further destruction of β-cells in the pancreas and may therefore be one of the mechanisms by which camel milk reduces the need for insulin in diabetic patients.

As there are a limited number of studies on camel milk derived insulin, this chapter describes the methods used to isolate the insulin protein from camel milk. The camel insulin was isolated using a number of different purification techniques, starting with milk defatting, then the large molecular weight proteins were separated by ethanol

precipitation leaving insulin and other small molecular weight proteins in the supernatant. The insulin then was partially purified by ion exchange chromatography and finally isolated using immunoaffinity chromatography. This material was then available for further study to characterize the stability and properties of insulin isolated directly from camel milk.

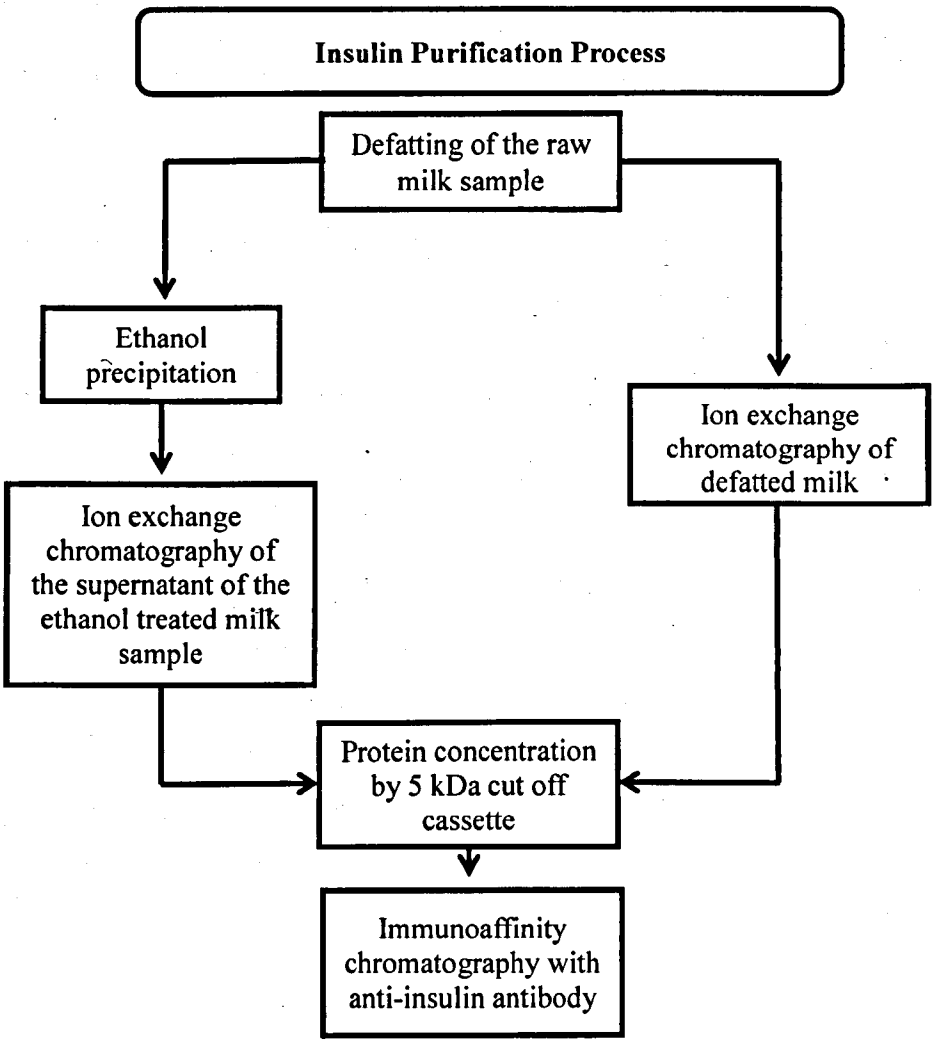
4.2 Aims:

- To purify the camel insulin protein from camel milk using the following methods.
  - Defatting of the raw camel milk
  - Casein treatment of the defatted milk sample.
  - Ion exchange chromatography for the raw and decaseinsed sample
  - Immuno-affinity chromatography of the product of ion exchange chromatography.

- To characterise the isolated protein by:
  - liquid chromatography coupled mass spectrometry (LC/MS) analysis of the purified sample
  - Western blotting
  - SDS-PAGE followed by Radioimmuno assay for the separated proteins.

4.3 Results

Camel milk insulin was purified/enriched from camel milk following the purification scheme detailed in the schematic flow diagram below.



4.3.1 Isolation and Purification of Camel Milk Insulin from Camel Milk

4.2.1.1 Defatting of Raw Camel Milk

The camel milk was obtained fresh as described in the materials and methods section and hence required defatting. This is required to prevent fouling and blocking of subsequent purification steps. In the defatting step a thin fat layer was observed in colostrum samples, while in normal milk this layer was thicker and accounted for approximately 2-3% of the total milk.

4.3.1.2 Casein Treatment of Camel Milk

Treatment of milk and colostrum with rennet enzyme recovered the highest concentration (71%) of insulin as determined by RIA in the whey fraction (see Table 4.1). When using the acid precipitation technique 89% of total insulin in the samples was lost during this step. Figure 4.1 shows the difference between the acid treated and untreated samples as shown by SDS-PAGE analysis. By comparison, there was only a 30% loss in the total insulin activity observed as determined by RIA when using the rennet enzyme treatment method (see Table 4.1). The insulin concentration when using the ethanol precipitation technique was 50% of the original concentration as determined and measured by radioimmunoassay. The concentration of ethanol used in the ethanol precipitation method was varied and optimized and the highest recovery obtained from an initial 50% ethanol precipitation to precipitate high molecular weight proteins (including casein) followed by a second 90% ethanol precipitation step to precipitate the insulin. Table 4.1 reports the concentration of insulin recovered using the different initial treatment methods.

Table 4.1. The concentration of insulin at different treatment stages during purification from camel milk

Method / Fractions	Acid precipitation method	Rennin enzyme treatment	Ethanol precipitation (50% followed by 90% Ethanol concentration)
Raw milk	64 µIU/mL	98 µIU/mL	100 µIU/mL
Defatted milk	59 µIU/mL	91 µIU/mL	92 µIU/mL
After treatment	7 µIU/mL	70 µIU/mL	50 µIU/mL

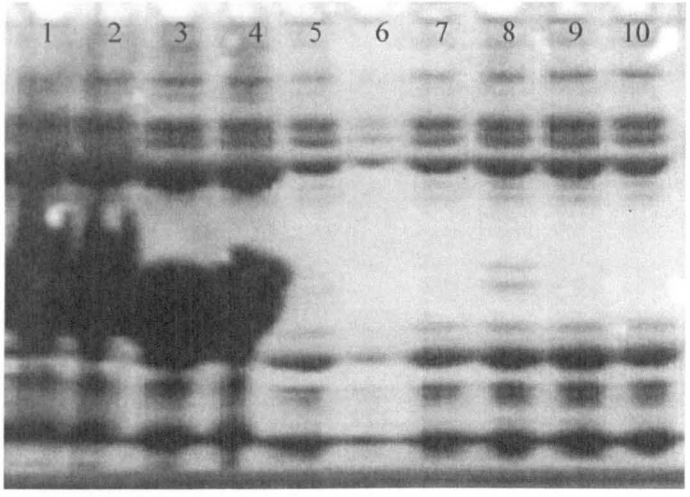


Figure 4.1. SDS-PAGE analysis of acid precipitation method for purification of camel insulin from camel milk. Lanes 1- 4 lwithout casein treatment, Lanes from 5-10 after casein treatment. Protein was loaded and run and then protein bands present visualized using Coomassie blue staining reagent.

4.3.1.3 Determining the Sensitivity of Western Blotting for the Characterisation and Detection of Insulin

In order to follow the purification of camel milk insulin and allow the relative amounts in different samples to be analysed rapidly via gel electrophoresis without undertaking RIA, western blotting of samples was investigated. As an antibody specifically to camel insulin was not available, an antibody to human insulin was used instead. As the sequence of the A-chain and B-chain of camel and human insulin are similar (see

Chapter 3) it was anticipated that the human insulin antibody would be appropriate to detect camel insulin. To define the sensitivity of the western blot approach for insulin, human insulin at different concentrations were prepared, run on an SDS-PAGE gel, transferred to a PDVF membrane and then western blotted for insulin using the anti-human insulin antibody as described in the materials and methods section (Figure 4.2). This analysis showed that the minimum concentration of human insulin detected by Western blotting was 10 mIU/mL, although the band at this concentration was very weak. Hence the western blotting approach cannot be used to analyse and detect insulin at concentrations lower than 10 mIU/mL.

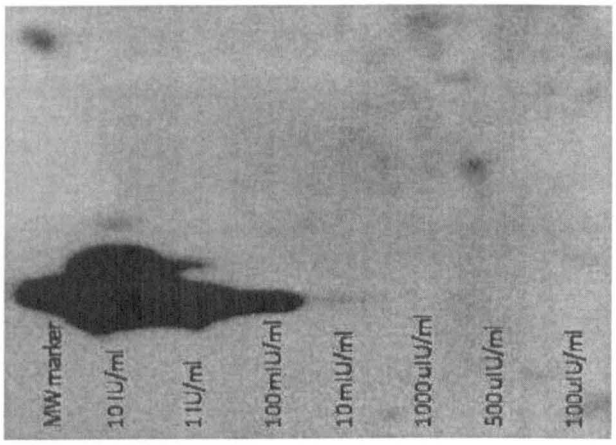


Figure 4.2. Western blotting of serial diluted insulin concentrations, the minimum concentration detected is 10 mIU/mL. The concentrations loaded into each lane are indicated.

4.3.1.4 Use of Gel Filtration to further Purify Insulin after Defatting and Ethanol Precipitation

As insulin is a small molecule, gel filtration was used as a first chromatography step to further purify the camel milk insulin whereby larger proteins would be expected to be efficiently removed. For this purpose a G-25 Sephadex bead column was used and the insulin in the resulting 1 mL fractions that were collected deteremined by RIA. The

results of the RIA analysis of the gel filtration chromatography fractions collected showed that the maximum concentration of insulin was recovered in the 3rd ml of washing where 30.4% of total insulin loaded was observed, however 31.6% of the total protein loaded was also found in this same fraction (Table 4.2). The 4th ml of washing showed the best result in terms of purification where 16.5% of the insulin was recovered with only 0.35% of total protein present in this fraction. However, this represents only a low recovery yield of the total insulin loaded with most being found in farctions 2 and 3 (Table 4.2). The gel filtration approach was therefore not very efficient at further purifying the insulin from the remainig total protein.

Table 4.2. Radio immunoassay and total protein concentration results of gel filtration test with G-25 beads to purify insulin from camel milk. Total protein was determined using a western blot, RIA was used to measure insulin activity in the various fractions.

Fractions	Total Protein (mg/ml)	Insulin RIA (µIU/mL)	Volume (ml)
Decaseinised colostrum sample	89.18	230	2.5
1st ml washing	30.43	42	1
2nd ml washing	85.56	150	1
3rd ml washing	70.63	175	1
4th ml washing	0.79	95	1
5th ml washing	1.44	25	1
6th ml washing	0.43	11.5	1
7th ml washing	0.15	8.2	1
8th ml washing	0.3	6.5	1
9th ml washing	0	5.5	1
10th ml washing	0.05	4.6	1
11th ml washing	0.04	4	1
12th ml washing	0.07	3.7	1

4.2.1.5 Ion Exchange Chromatography as an Alternative for Purification and Recovery of Insulin from Camel Milk

Ion exchange chromatography was also investigated as a means to rapidly purify insulin from camel milk. The results of ion exchange chromatography using A-25 DEAE Sephadex beads as determined by analysing the insulin activity recovered by RIA is presented in Tables 4.3-4.9 and confirmed good binding (as little insulin activity was found in the flow through samples) and subsequent good recovery of the bound insulin. Tables 4.3 and 4.4 report the results of ion exchange chromatography of decasienised colostrum samples diluted and from the stock respectively. This purification approach recovered approximately 62% of the camel insulin when the sample was eluted with buffer that contained 500 mM or more NaCl as determined by RIA. As an alternative to using decasienated milk, defatted raw milk was also subjected to ion exchange chromatography. However, when this approach was used 46% of insulin activity was found in the 500 mM elution fraction and 14% of the activity was recovered in the 1 M elution fraction. When both elutions were merged a total of 60% recovery was obtained using the defatted raw milk (Table 4.5).

Table 4.3. Ion exchange chromatography of diluted decaseinised colostrum (3/11) colostrum / (equilibration Buffer + H2O) with A-25 DEAE Sephadex beads in 5 mL column.

Fractions	Total Protein (mg/mL)	Insulin RIA (µIU/mL)	Volume (mL)
Decaseinised colostrum sample	89.18	180	3
1st Flow through	6.72	5	4
2nd Flow through	16.82	5.2	4
3rd Flow through	18.43	4.5	4
4th Flow through	18.65	5.5	4
Washing with 10 mM NaCl	3.89	4.8	6
Elution with 50 mM NaCl	0.21	4.7	10
Elution with 250 mM NaCl	1.59	6.2	10
Elution with 1 M NaCl	0.8	28	12

Table 4.4. Ion exchange chromatography decaseinised colostrum using A-25 DEAE Sephadex beads in 5 mL column.

Fractions	Total Protein (mg/ml)	Insulin RIA (µIU/mL)	Volume (ml)
Decaseinised colostrum sample	89	230	11
Flow through	80.75	13.6	11
Washing with 100 mM NaCl	4.9	55	15
Elution with 500 mM NaCl	1.42	80	20
Elution with 1 M NaCl	0.17	4	20
Elution with Acetic Acid	0.1	0.8	20

Table 4.5. Ion exchange chromatography of defatted raw whole milk trial with A-25 DEAE Sephadex beads in 5 mL column.

Fractions	Total Protein (mg/ml)	Insulin RIA (µIU/mL)	Volume (ml)
Raw milk Sample	35.3	20	45
Flow through	26.9	1	13
Flow through	25.1	1.4	13
Flow through	30.4	1.9	15
Last drop of flow through	32	1.6	N/A
1st wash 50 mM NaCl	15.8	2.8	5
2nd wash 50 mM NaCl	0.53	5	10
3rd wash 50 mM NaCl	0.93	6	5
Elution with 250 mM NaCl	1.89	5.5	20
Elution with 500 mM NaCl	0.63	21	20
Elution with 1M NaCl	0.18	6.3	20

As these small scale preliminary results were promising in terms of recovering insulin activity, a purification trail was attempted using 100 L of milk. The milk was initially defatted using centrifugation and then the milk was loaded onto a 1 L DEAE A-25

Sephadex bead column. After applying more than 10 column volumes of sample the column became saturated such that after 1 hour of loading the insulin concentration in the flow through returned to the original concentration as shown in Table 4.6 Thus, no additional insulin was binding at this time.

Table 4.6. Ion exchange chromatography, the trail of scale started with 100 L of partially defatted raw milk with A-25 DEAE Sephadex beads in 1L column.

Fractions	Total Protein (mg/ml)	Insulin RIA (μIU/ml)	Volume (L)
Raw milk	-	72	100
Flow through during 1st hour	-	18	-
Flow through during 2nd hour	-	69	-
Flow through during 3rd hour	-	68	-
Flow through during 4th hour	-	69	-
Flow through during 5th hour	-	70	-
Flow through during 6th hour	-	71	-
Flow through during 7th hour	-	68	-
Flow through during 8th hour	-	69	-
Flow through during 9th hour	-	67	-
Flow through during 10thhour	-	69	-
Flow through during 11thhour	-	68	-
Wash with 0.2 M NaCl	0.18	6	3
1st Elution 1 M NaCl	1.21	4	2
2nd Elution 1 M NaCl	1.55	5	2
3rd Elution 1 M NaCl	0.53	4	2
After elution wash	0.9	2	3

Further, purification approaches were undertaken with rennet enzyme casein treated sample, but in this case intermediate column washes were included in the protocol. However, the column again became saturated in this case with 5 column volumes of

milk when the casein treated milk with rennet enzyme was used to load the column as shown in table 4.7. Other pre-treatments of the camel milk were also investigated in terms of the activity of insulin recovered following ion-exchange chromatography. Table 4.8 and 4.9 report the insulin activity and profile following ion exchange of defatted raw milk samples when treated with 50% ethanol and shows the recovery improved from 36% in table 4.8 to 68% in table 4.9 whereby a 20 mL column was used in the later study and a 5 mL column in the first. An example chromatogram showing the absorbance profile at A280 nm across the purification procedure is shown in figure 4.3. The majority of the bound insulin was eluted in the 2nd and 3rd elution fractions (figure 4.3) as determined by the RIA activity assay (see table 4.9) where 61% of the total activity loaded was found.

Table 4.7. Ion exchange chromatography of defatted decaseinised raw whole milk trial using A-25 DEAE sephadex beads with intermediate washing in 5 mL column.

Fractions	**Total Protein (mg/ml)**	**Insulin RIA** (μIU/mL)	**Volume** (ml)
Decaseinised raw milk	7.54	69	250
Flow through 30 mL	6.23	35	-
Flow through 75 mL	6.94	43	-
Flow through 125 mL	6.73	46	-
Flow through 200 mL	7.07	56	-
Flow through 250 mL	7.18	68	-
Wash 50 mM	1.23	15.2	30
Elution 250 mM NaCl	0.23	13	20
Elution 1 M NaCl	1.02	34	20
Elution 1 M NaCl	0.76	33	50

Table 4.8. Ion exchange chromatography of 50% ethanol treated defatted raw milk trial using A-25 DEAE Sephadex beads in 5 mL column.

Fractions	Total Protein (mg/ml)	Insulin RIA (μIU/mL)	Volume (ml)
Defatted raw milk sample	30	107	100
1st wash 50mM NaCl	2.37	10	10
2nd wash 50mM NaCl	1.43	65	25
Elution 250mM NaCl	1	106	25
Elution 1M NaCl	0.5	28	25
Elution 1M NaCl	0.11	8	40

Table 4.9. Ion exchange chromatography of 50% ethanol treated defatted raw milk trial using A-25 DEAE Sephadex beads in 20 mL column using LP biologic.

Fractions	Total Protein (mg/ml)	Insulin RIA (μIU/mL)	Volume (ml)
Defatted raw milk sample	33	90	100
Ethanol treated sample	0.74	-	200
Flow through after 100 ml	0.32	-	-
Flow through after 200 ml	0.44	-	-
Wash 80 mM NaCl 1st peak	0.65	6.7	4.5
Wash 80 mM NaCl 2nd peak	0.78	7	5
Elution 1 M NaCl , 1st peak	2.97	96	6.5
Elution 1 M NaCl, 2nd peak	3.83	750	4.5
Elution 1 M NaCl, 3rd peak	4.48	430	5

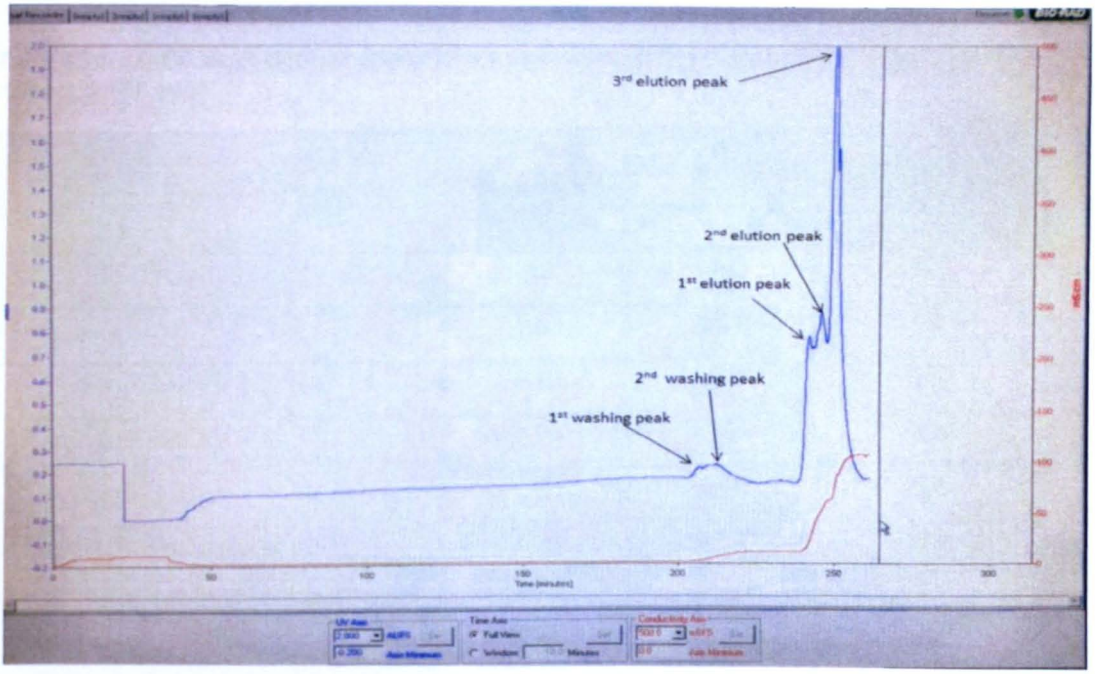


Figure 4.3. Ion exchange A280 nm chromatogram of the data reported in table 4.9 generated from data viewer of LP biologic automated chromatography machine (BioRad, USA). Samples were collected according to the peaks marked on the chart. The chromatogram shows the UV curve at A280nm in blue, while the red curve is the conductivity (mS/cm).

At this stage a second purification was attempted at scale starting with 65L whereby the defatted milk was subjected to 50% ethanol treatment as described above. In this case approximately 2 kg of cream was initially separated from 65 L of milk and the subsequent ethanol treatment produced 57 L of clear supernatant containing 59% of the insulin activity of the original sample as determined by RIA. The results from the subsequent ion exchange purification, reported in table 4.10, show that the binding was very efficient during the first 1.5 hours, but after this there was an increase in the insulin activity found in the flow through suggesting breakthrough and material not binding to the column due to saturation (Table 4.10). Washing the column with ethanol free washing buffer resulted in the elution of the majority of the insulin during the washing step. The majority of the remaining insulin bound to the column was eluted in the first three elution fractions with 0.7 M NaCl and increasing the washing stringency to 1 M NaCl did not elute more insulin.

Table 4.10. Ion-Exchange chromatography for the supernatant of 50% ethanol treated samples of scale stated with 65L milk:

Fractions	Total Protein (mg/ml)	Insulin RIA (μIU/mL)	Volume (L)
Raw milk (Defatted)	32	69	65
Sample	1.62	40.8	57
Flow through after 1.5 hour	1.13	1.2	NA
Flow through after 3 hours	1.23	20	NA
Flow through after 5 hours	1.38	24.8	NA
Flow through after 6 hours	1.43	47.2	NA
Total Flow through	1.09	16	55
Wash1 (80 mM)	5.79	208.32	1
Wash 2 (80 mM)	2.19	155	1
1st elution (700 mM)	1.76	124	0.5
2nd elution (700 mM)	3.87	150	0.5
3rd elution (700 mM)	2.55	77	0.5
4th elution (700 mM)	0.83	39	0.5
5th elution (700 mM)	0.49	20	0.5
6th elution (700 mM)	0.67	17	0.5
7th elution (700 mM)	0.2	6	0.5
8th elution (700 mM)	0.22	5	0.5
9th elution (1 M)	0.64	6	0.5
10th elution (1 M)	0.22	4	0.5
11th elution (1 M)	0.09	4	0.5
12th elution (1 M)	0.08	3	0.5

A third purification run was also undertaken using ion-exchange chromatography of untreated but defatted camel milk. From 45 L of milk approximately 1.5 kg of cream was separated and then the 45 L of untreated defatted milk processed in 6 independent batches. In each batch 7.5 L of milk was loaded onto the ion-exchange column and the insulin eluted and after each batch the column was regenerated. The resulting RIA analysis of insulin activity, detailed in table 4.11, show that much of the insulin was not binding to the column and was passing straight through the column.

4.3.2 Concentration of Insulin Purified by Chromatography

Following purification of the insulin and determination of those fractions that contained insulin by RIA, it was necessary to concentrate the purified material. The high insulin activity fractions from the second and third purification runs were therefore combined together and resulted in 5 L of sample. After filtering the entire 5 L sample using an 0.2 µm filter, the sample was concentrated using a 5 kDa cut off cassette to reduce the volume to 145 mL. The 145 mL sample concentration as determined by RIA was 12000 µIU/ml. This material was then subjected to further purification using imunoaffinity chromatography.

Table 4.11. Insulin purification by Ion-Exchange chromatography for defatted milk sample run in 6 batches. Insulin activity was determined across all fractions by RIA.

Sample	Batch	Insulin (µIU/mL)	Volume (L)
Defatted milk	Total	94	45
last 10 ml of flow through before washing (2nd batch)	2nd batch	132	0.01
last 10 ml of flow through before washing (3rd batch)	3rd batch	180	0.01
last 10 ml of flow through before washing (4th batch)	4th batch	143	0.01
last 10 ml of flow through before washing (6th batch)	6th batch	156	0.01
Wash with 2 liters for each batch (50mM Nacl)	Total	88	13
Elution with 200 mM Nacl	1st batch	162	1
1st Elution with 500 mM NaCl	1st batch	135	0.5
2nd Elution with 500 mM NaCl	1st batch	92	0.5
3rd Elution with 500 mM NaCl	1st batch	62	0.5
4th Elution with 500 mM NaCl	1st batch	33	0.5
5th Elution with 500 mM NaCl	1st batch	18	0.5
Elution with 200 mM NaCl	2nd batch	100	1
1st Elution with 500 mM NaCl	2nd batch	103	0.5
2nd Elution with 500 mM NaCl	2nd batch	107	0.5
3rd Elution with 500 mM NaCl	2nd batch	78	0.5
4th Elution with 500 mM NaCl	2nd batch	51	0.5
5th Elution with 500 mM NaCl	2nd batch	31	0.5
Elution with 200 mM NaCl	3rd batch	45	1
1st Elution with 500 mM NaCl	3rd batch	62	0.5
2nd Elution with 500 mM NaCl	3rd batch	63	0.5
3rd Elution with 500 mM NaCl	3rd batch	48	0.5
4th Elution with 500 mM NaCl	3rd batch	28	0.5
5th Elution with 500 mM NaCl	3rd batch	20	0.5
Elution with 200 mM NaCl	4th batch	16	1
1st Elution with 500 mM NaCl	4th batch	14	0.5
2nd Elution with 500 mM NaCl	4th batch	17	0.5
3rd Elution with 500 mM Nacl	4th batch	16	0.5
4th Elution with 500 mM NaCl	4th batch	10	0.5
5th Elution with 500 mM NaCl	4th batch	8	0.5
Elution with 200 mM NaCl	5th batch	95	1
1st Elution with 500 mM NaCl	5th batch	85	0.5
2nd Elution with 500 mM NaCl	5th batch	76	0.5
3rd Elution with 500 mM NaCl	5th batch	52	0.5
4th Elution with 500 mM NaCl	5th batch	31	0.5
5th Elution with 500 mM NaCl	5th batch	19	0.5
Elution with 200 mM NaCl	6th batch	63	1
1st Elution with 500 mM NaCl	6th batch	70	0.5
2nd Elution with 500 mM NaCl	6th batch	67	0.5
3rd Elution with 500 mM NaCl	6th batch	46	0.5
4th Elution with 500 mM NaCl	6th batch	28	0.5
5th Elution with 500 mM NaCl	6th batch	17	0.5

4.3.3 Immunoaffinity Chromatography Purification of Insulin following Ion-Exchange Chromatography and Concentration of Insulin Containing Fractions.

To further purify the camel milk insulin following ion-exchange chromatography, the concentrated insulin was subjected to immunoaffinity chromatography purification. The immunoaffinity chromatography was undertaken using a 3 ml column that containing an anti-human insulin antibody linked to Sepharose 4 beads. An initial trail was undertaken using this column and decasienised colostrum to load onto the column (so as not to use the ion-exchange purified material on optimisation studies) and this work showed that after elution from the immunoaffinity chromatography column approximately 87% of the total insulin loaded was recovered as detected by RIA (Table 4.12). The RIA activity data from the fractions collected from this experiment showed that the majority of the insulin activity was found in the 4th and 5th elution fractions when using 1.5 M NaCl in the elution buffer.

Table 4.12. Immuno affinity chromatography of decasienised colostrum in 3 mL column.

Fractions	Total Protein (mg/ml)	Insulin RIA (μIU/mL)	Volume (ml)
Decasenised colostrum sample	92	540	22
1st elution fraction 1.5 M NaCl	0.01	6.3	2
2nd elution fraction 1.5 M NaCl	0.01	6.5	2
3rd elution fraction 1.5 M NaCl	0.85	15	2
4th elution fraction 1.5 M NaCl	4.65	2300	2
5th elution fraction 1.5 M NaCl	3.2	2100	2
6th elution fraction 1.5 M NaCl	0.62	640	2
7th elution fraction 1.5 M NaCl	0.15	124	2

Following demonstration that the affinity chromatography column approach was working, the concentrated camel milk insulin from the ion-exchange purification was

subjected to immunoaffinity chromatograph with the 145 mL of sample being processed in two batches, one of 80 mL and one of 65 mL. Table 4.13 reports on the RIA insulin activity results from the immunoaffinity chromatography of 80 ml of the concentrated product from the ion-exchange runs described and reported in in table 4.10 and 4.11. The highest insulin activity/concentration was observed in (a) the 3rd fraction where 0.18 IU of insulin was eluted, this amount equals 8.19 µg of protein, and (b) the 4th fraction which contained 0.48 IU (2.2 µg) of insulin. The insulin binding to the column was efficient as less than 0.4% of the insulin activity was detected in the flow through (Table 4.13). The remaining 65 ml of sample was then run in a second batch (table 4.14) where most of the insulin activity was in the 2nd and 3rd fraction where 0.14 IU (6.3 µg) was detected. Less than 0.5% of the insulin activity was in the flow through (Table 4.14).

Table 4.13. Immunoaffinity chromatography of the ion-exchange purification concentrated sample in 3 mL column (80 ml).

Fractions	**Total Protein (mg/ml)**	**Insulin RIA (µIU/mL)**	**Volume** (ml)
Sample	23.25	12000	80
1st Flow Through	19.65	48	40
2nd Flow Through	24.82	27	40
1st elution fraction 0.5 M NaCl	0.18	640	2
2nd elution fraction 0.5 M NaCl	0.1	670	2
3rd elution fraction 0.5 M NaCl	0.93	90000	2
4th elution fraction 0.5 M NaCl	0.34	24000	2
5th elution fraction 0.5 M NaCl	0.09	2850	2
6th elution fraction 0.5 M NaCl	0.05	1300	2
7th elution fraction 0.5 M NaCl	0.04	900	2
8th elution fraction 0.5 M NaCl	0.04	800	2
9th elution fraction 0.5 M NaCl	0.03	700	2
10th elution fraction 0.5 M NaCl	0.01	60	2

Table 4.14. Immunoaffinity chromatography of the ion-exchange purification concentrated sample in 3 mL column (65 ml).

Sample ID	Total Protein (mg/ml)	Insulin RIA(µIU/ml)	Volume(ml)
Sample	23.25	12000	65
1st Flow Through	21.36	56	50
2nd Flow Through	24.07	61	15
1st Elution 0.5 M NaCl	0.04	4	2
2nd Elution 0.5 M NaCl	0.4	10600	2
3rd Elution 0.5 M NaCl	0.44	68800	2
4th Elution 0.5 M NaCl	0.11	3700	2
5th Elution 0.5 M NaCl	0.03	600	2
6th Elution 0.5 M NaCl	0.03	400	2
7th Elution 0.5 M NaCl	0.01	300	2
8th Elution 0.5 M NaCl	0.02	150	2
9th Elution 0.5 M NaCl	0.04	100	2
10th Elution 0.5 M NaCl	0.02	80	2

4.3.4 Characterisation of the Purified and Isolated Camel Milk Insulin

The protein recovered after the ion-exchange and immunoaffinity purification procedure was then analysed. To do this 10 µl of the elution product from fraction no 2, 3, 4 and 5 in table 4.14 above were loaded onto a 12% SDS-PAGE system and after running blotted onto PVDF membrane. Human insulin was also run as a control. The resulting membrane was then western blotted using the anti-human insulin antibody and the resulting blot is shown in Figure 4.4. As can be seen in Figure 4.4, in the control lane containing human insulin a clear single band was detected, however no such corresponding band was observed in the fractions eluted from the immunoaffinity column. Instead a number of bands were observed at a much higher molecular weight in all the elution fractions whilst there was no band at the monomer insulin size expected (see figure 4.4). It was possible that these higher molecular weight bands were multimers of insulin that were covalently linked via disulphide bridges scrambled during purification. They could not be non-covalently associated, as these would have been destroyed in the presence of the SDS.

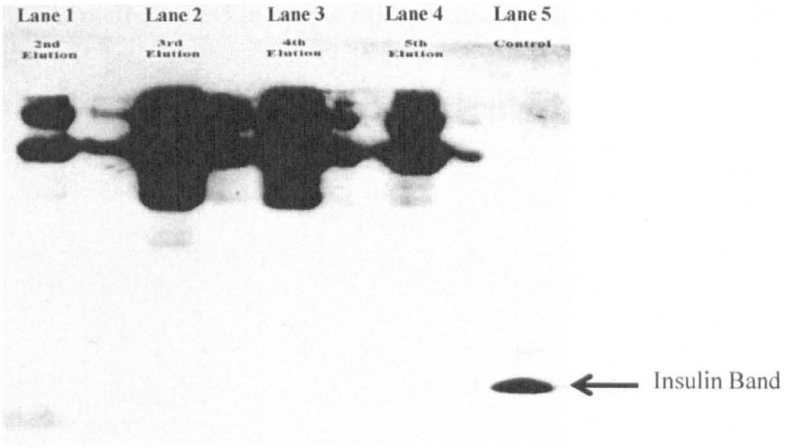


Figure 4.4 Western blotting analyses of elution fractions from the immunoaffinity chromatography fractions reported in table 4.14. Lane 1 = 2nd elution fraction , Lane 2 = 3rd elution fraction, Lane 3= 4th elution fraction, Lane 5= 5th elution fraction, Lane 6= Insulin control. The arrow indicates the position of insulin on the blot.

As no insulin band was clearly visualize and identified by western blotting, and in order to confirm the size of the purified insulin when analysed on the SDS-PAGE gel, the gel was cut into horizontal sections as shown in figure 4.5. The protein was then eluted from each horizontal gel section and subjected to RIA analysis to determine the regions of insulin activity. The resulting RIA data indicated the presence of active insulin in the gel section between 3-6 kDa (see Figure 4.5). Thus, active camel milk insulin was present in the region where the control human insulin was detected by western blotting and there was little to no activity where the higher molecular weight species were observed. The higher molecular weight species are therefore either not insulin and are simply cross-reactivity with the antibody, or contain inactivated insulin which is not detected by the RIA technique. The antibody used for RIA detection was raised against the insulin B-chain and the samples were mixed with loading buffer containing a reducing agent, so the results indicate the size of insulin B-chain only. The same amount of the 3rd elution fraction from table 4.13 was loaded in the two gels shown in Figure 4.5 below. The RIA results show that more insulin was recovered from the gel (46-fold greater) in figure 4.5B where 3:2 H_2O/isopropanol method was used to recover the insulin in comparison with the passive diffusion method used to recover the insulin from the gel used in figure 4.5A.

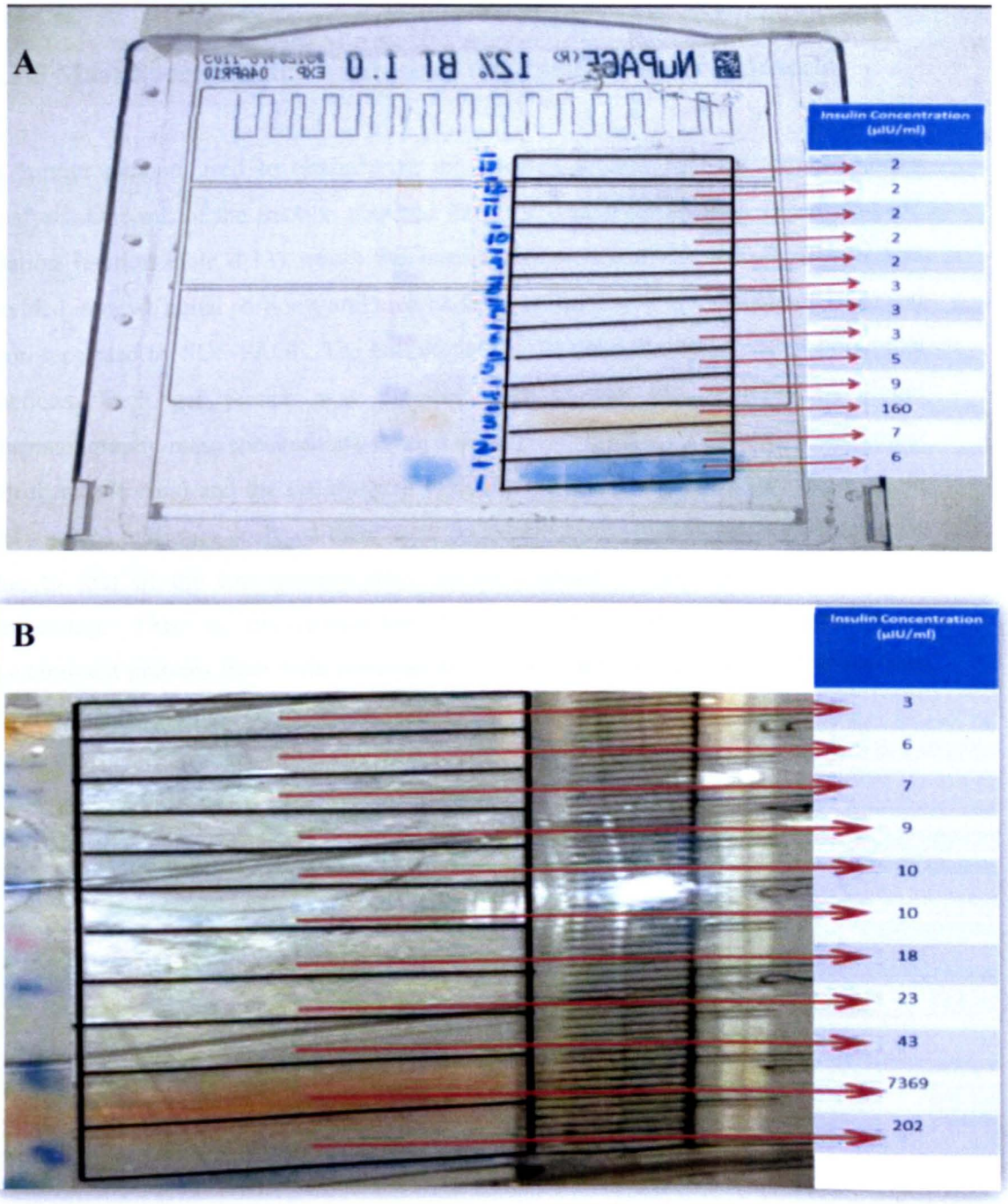


Figure 4.5. SDS-PAGE analysis of the 3rd immunoaffinity elution fraction described in table 4.13. The resulting gel was then sectioned into horizontal sections and the protein in each section recovered and subjected to insulin activity analysed by RIA. In (A) protein was recovered from the gel by using a passive diffusion method while in (B) the protein present in the excised sections were recovered with a 1:2 H_2O/isopropanol solution method.

4.3.5 Mass Spectrometry Analysis of the Purified Insulin Material

A further method used to characterize the isolated protein was mass spectrometry analysis. One mL of the fraction obtained from the insulin purification experiment (3rd elution fraction table 4.13) where the insulin concentration was 90000 µIU/ml was divided into two equal portions, and each half was concentrated to a volume of 10 µl and then separated by SDS-PAGE. The bottom part of the gel was divided into 4 horizontal sections. Each gel section was digested with trypsin and analysed by liquid chromatography-mass spectrometry using a micrOTOF-Q mass spectrometry instrument (Bruker Daltonics) and the electrospray ionization mass spectrometry technique. In all gel sections that were analysed there were no insulin peptides detected, which could be due to loss of the low concentration insulin peptides during the in-gel digestion processing. This is discussed in the following section. However, a number of contaminant proteins from milk were identified confirming that the insulin fraction was not 100% pure as might be expected.

4.4 Discussion

The isolation of proteins from natural sources such as milk is not an easy process. This is due to the presence of other proteins present in large quantities and often the low abundance of the target protein as in the case of this study. For the isolation of insulin from camel milk different purification steps were used including defatting, casein treatment, ion exchange chromatography and immunoaffinity chromatography. Such steps are commonly used in the chromatographic purification of target proteins from natural sources such as milk.

The casein treatment of the camel milk with rennet enzyme was shown to be very inefficient, although in the initial trails the rennet treatment was very effective and produced a clear whey fraction. Even when using newly manufactured enzyme the casein treatment did not improve the subsequent purification steps. This may be due to seasonal changes in camel milk components (Haddadin *et al.*, 2007), which can affect the curdling process. The initial trial was performed in winter while the trails where the treatment was not efficient were undertaken in summer.

The casein treatment using acid precipitation methods resulted in an approximate 89% loss of insulin in the whey fraction. In the early insulin isolation trails in 1923, Doisy developed a method for insulin isolation (Doisy *et al.*, 1923). The method was based on a proteins property to precipitate at a pH value near the isoelectric focusing point (pI) and hence where insulin precipitated under slightly acidic conditions. Reducing the pH in the studies here by using acid precipitation of casein reduced the insulin solubility and precipitated most of the insulin with the casein.

An alternative precipitation method was also investigated, the use of ethanol. When radioimmunoassay was used to monitor the fate of insulin upon ethanol precipitation, these data showed that ethanol precipitation did not precipitate the majority of the

insulin at 50% and the insulin was precipitated at 90%. This was useful as at 50% many other proteins precipitated and hence were purified away from the camel milk insulin and also suggested, as expected, that the camel milk insulin was a small molecular weight protein as for human insulin. These results were in agreement with other previous studies where in earlier insulin isolation attempts it was found that the precipitate formed from 90% concentration alcohol contained the majority of the insulin protein that was being attempted to be recovered (Rosenfeld, 2002).

Gel filtration using G-25 beads was used to try and remove large proteins from the small insulin molecule. Using this methodology the insulin was eluted late, indicating that the insulin entered the beads pores and was of a small size as expected. The gel filtration method can be used for insulin polishing but is probably not appropriate as a main purification procedure. It is reported that insulin can go through irreversible aggregation in the aqueous environment. The aggregation reactions are influenced by pH, the presence of ions including Zn^{2+} and Cu^{2+}, the buffer used, temperature, ionic strength and the protein concentration while at equilibrium it is reported that insulin can exist as dimers, tetramers, and hexamers in the aqueous environment (Jeffrey PD and Coates 1966; Pekar and Frank 1972; Darrington and Anderson, 1994; Darrington and Anderson, 1995; Birnbaum *et al.*, 1997). This association is likely to be caused by hydrophobic interactions between insulin molecules (Yu *et al.*, 2006; Yu *et al.*, 2008). Furthermore, in gel filtration method the sample needed to be loaded in small fractions and eluted with large amounts of elution buffer, making this method not very practical for a large scale purification.

As stated above, insulin is a small molecular weight protein and is negatively charged at neutral pH. For this reason ion-exchange was investigated as part of the purification methodology. The A-25 DEAE Sephadex beads (GE healthcare, USA) combine the benefits of ion exchange chromatography with gel filtration chromatography as the beads are positively charged and have small pores. The use of A-25 DEAE Sephadex as

both an anion exchanger and gel filtration bead system has been previously described for the purification and isolation of insulin from other sources (Chen *et al.*, 1995) and hence was investigated in this study.

It has been reported that eluting the insulin from an ion exchange chromatography column gives high insulin recovery when using 10 mM ammonium acetate (pH 8) as an elution buffer (Dave *et al.*, 2008). Using such a buffer in the large-scale isolation of insulin from milk might not be efficient for several reasons. Firstly, the acidic elution may result in denaturing traces of casein in/on the column which may block the column. Secondly, in the large-scale purification, keeping the isolated insulin at low pH conditions for long time may alter the structure of the isolated protein and lead to aggregation or the generation of misfolded material.

Immunoaffinity chromatography has been reported in the literature as a method to isolate and then study the insulin from samples that contain insulin at low concentrations such as urine and serum (Kuuranne *et al.*, 2008; Hes *et al.*, 2012). Kuuranne and colleagues (2008) successfully isolated horse insulin and recombinant insulin from racing horse urinary samples and then differentiated between the two insulins by mass spectrometry. Hess and co-workers (2012) also managed to isolate and analyse traces of insulin and insulin analogues from human serum using the immunoaffinity chromatography purification technique. The anti-insulin antibody in the immunoaffinity column used in this project is different than the one used by Hess *et al.* (2012). The antibody used in this study has previously been shown to have good binding ability to camel milk insulin in the insulin RIA kit (Diasource, Belgium).

The A-25 DEAE Sephadex bead ion-exchange methodology provided purification in initial smaller scale trails with decaseinised milk, but this was not reflected in the first purification run which started with 100 L milk. The first purification failed for two main

reasons. Firstly, the fat content in the milk could not all be stripped out due to the large sample volume and the lack of a defatting machine at that time, and hence a white layer of fat accumulated over the column and blocked the beads. Secondly, the large quantity of casein protein in the milk can bind to A-25 DEAE Sephadex beads (Yaguchi & Rose 1971) and saturate them rapidly. Further, rennet enzyme treated milk samples which had small amounts of the leftover digested casein can result in the casein binding to beads more than in the untreated samples; these smaller casein molecules can enter the beads pores and bind more specifically and saturate the column rapidly rendering it unable to bind the much less concentrated insulin in the sample. Casein treatment and ethanol precipitation with 50% ethanol can help elevate this problem by not only removing the casein but also by precipitating the large molecular weight proteins.

The second ion-exchange purification (scale started with 65L of milk) combined the ethanol precipitation with ion exchange chromatography. In this approach the column binding capacity improved and there was almost full insulin recovery during the first 90 min of sample loading through the column. In this second purification run, most of the insulin was eluted during the washing step; the bed volume of DEAE A-25 Sephadex beads in the ion exchange chromatography column was increased during the washing step. The loaded sample had 50% ethanol, however the washing was undertaken with ethanol free buffer.

Following the recovery of insulin by ion-exchange chromatography, the insulin containing fractions were combined and concentrated. These were then applied to an affinity chromatography column which resulted in high binding efficiency and good recovery of the insulin. The purification of the insulin by a combination of ion exchange chromatography using A-25 DEAE Sephadex and immunoaffinity chromatography with an anti-human insulin antibody (Diasource diagnostics, Belgium) resulted in a good recovery of the purified insulin from the initial sample.

The purified material was then analysed by western blotting. The western blotting method was optimised using human insulin and the anti-human insulin antibody. The western blotting sensitivity of human insulin using the anti- human insulin antibody was tested and the lowest activity/concentration of insulin that could be detected was at 10 mIU/ml insulin concentration. Although more than this amount of camel insulin was obtained, the camel insulin band was not observed, even at the highest concentration of 90 mIU/ml. However, higher molecular weight bands were observed. It was not determined if these were insulin aggregates that were inactive but this appeared to represent the majority of material recovered. It is possible that the human anti-insulin antibody may not recognize the camel insulin due to the small changes in amino acid sequence as described in Chapter 3 although this seems unlikely as the insulin obviously bound to the immunoaffinity column. It seems more likely that much of the insulin aggregated and was inactivated and little monomer was left in solution.

To determine the location of the isolated insulin protein the SDS-PAGE gel was excised in strips and each strip was analysed by RIA according to size. The RIA showed insulin activity in a position matching the human insulin control, proving that the isolated protein was insulin. However, there was no significant activity observed in the area corresponding to the higher molecular weight bands.

To further characterize the purified insulin, mass spectrometry analysis was undertaken on the purified sample. The purified material was subjected to tryptic digestion and then analysed by LC-MS/MS. The LC/MS results of the in-gel tryptic digested isolated product didn't show any of the expected insulin fragments, however a number of milk proteins were identified as being present in the purified material. The lack of any insulin protein being observed is probably due to the tryptic digestion reducing the insulin molecule to very small pieces and the low abundance of insulin compared to some of the

contaminating proteins that remained in the solution. This observation has been described and explained by Muller and colleagues (2010) who describe the reasons for failing to detect low molecular weight (LMW) proteins by mass spectrometry (Muller *et al.*, 2010). They described that this can be due to the low numbers of proteolytic peptides formed by tryptic digestion as well as the propensity of the protein to be lost in protein separation, concentration and desalting procedures. Protein loss may also happen during destaining, by adsorption on surfaces of pipette tips and digestion tubes, by drying of samples in the speedvac (vacuum concentrator) and incomplete recovery of peptides from the gel after digestion (Speicher, 2000 and Stewart, 2001).

4.5 Conclusion:

It can be concluded that combining ethanol precipitation, ion exchange and immuno-affinity chromatography was efficient in isolating insulin from camel milk. Approximately 17 µg of camel insulin was purified using the above mentioned methods. The purified camel insulin was similar in size as the control human insulin after separating the sample by SDS-PAGE and analysing the separated protein by Radioimmuno assay. The conditions of in-gel digestion need to be optimised for camel insulin analysis.

Chapter 5

Production of Recombinant Camel Insulin and Investigations into Camel Milk Insulin Stability

5.1 Introduction

For many years insulin purified from porcine and bovine pancreas has been used to treat diabetes (Kumar and Bhat, 2003). Purification was achieved in Canada by Banting, Best, Macleod and Collip in 1921 (Banting *et al.,* 1922). Limitations in insulin production, supply and a concern about the transfer of viruses and incomplete removal of contaminating additional protein, encouraged scientists to find alternatives. In the 1960s human insulin was chemically synthesized by peptide synthesis of the A- and B-chains independently. The synthesis process was followed by random combination and folding of the two synthetic chains to try and produce 'native' human insulin. However, there were low yields in these intial efforts to obtain correctly folded insulin protein from the direct oxidative assembly of the synthetic A- and B-chains (Meienhofer *et al.*, 1963, Kung *et al.*, 1966; Katsoyannis *et al.*, 1966). Since these initial investigations, much research has been undertaken in the intervening years to develop thiol protection, cleavage and oxidation schemes for the stepwise formation of the correct disulphide bonds to form native insulin (Sieber *et al.*, 1974, Akaji *et al.*, 2002 Moroder *et al.*, 2003). The first landmark report into insulin chemical synthesis was performed by the successful production of correctly folded human insulin using regioselective cysteine-pairing procedures (Akaji *et al.*, 1993).

Human insulin for medical purposes is now produced by recombinant DNA technology. This technology was initially utilised by cloning the DNA fragments of the human insulin A- and B-chains into two separate vectors then transforming the vectors into

Escherichia coli (*E.coli*) cells. The insulin chains are then harvested and assembled to generate the appropriate disulphide bonds to yield native human insulin (Goeddel *et al.*, 1979).

An alternative to the production of insulin in *E.coli* is to use yeast based technology. Yeast based technology offers structural and protein folding advantages over the use of *E.coli*. These include the fact that yeast is a eukaryotic cell that contains an ER and the correct chaperone machinery to fold and assemble insulin that *E.coli* lacks. The technology is based upon producing a folded proinsulin protein which is then converted into mature insulin by enzymatic cleavage to generate insulin (Markussen *et al.*, 1986).

An alternative to yeast technology is the use of mammalian cells for the production of insulin. Approximately 70% of the recombinant therapeutic proteins are made using mammalian cell biotechnology and cell culture whereby the product of interest is secreted into the cell culture media (Reichert *et al.*, 2005). For mammalian cell technology the Chinese hamster ovary (CHO) cell line is the most extensively used mammalian cell type which can produce glycosylated and correctly folded recombinant proteins (Walsh, 2006, Peter and David , 2008).

This chapter describes attempts at the production of recombinant camel insulin using mammalian cell biotechnology and the chemical synthesis of the insulin peptides using solid phase peptide synthesis.

5.2 Aims:

- To generate camel cDNA from pancreatic RNA.
- To amplify the preproinsulin and proinsulin from camel cDNA
- To clone both amplified preproinsulin and proinsulin (with and without stop codon) into both pcDNA3.1v5His and pcDNA5 FRT vectors.
- To transiently transfect the cloned pcDNA3.1v5His vectors (preproinsulin and proinsulin with and without stop codon) in Chinese's hamster ovary (CHO) cell line.
- To stably transfect the cloned pcDNA5 FRT vectors (preproinsulin and proinsulin with and without stop codon) in CHO cell line using the commercially available flp-in system.
- To chemically synthesis the camel insulin A and B chains by solid phase peptide synthesis technique using the peptide synthesiser machine.

5.3 Results

5.3.1 Cloning of Camel Insulin Gene

5.3.1.1 Cloning of the camel insulin gene using synthetic DNA fragments

In order to generate recombinant camel insulin, the insulin A- and B-chains were initially cloned into both the pcDNA3.1v5His and pcDNA5 FRT commercially available mammalian expression vectors. The sense and anti-sense strands of A- and B-chains were commercially synthesised by 1st BASE Custom Oligos and were successfully annealed as shown in figure 5.1. Figure 5.1 shows the synthetic DNA strands as analysed on a 2% agarose gel before and after annealing. After annealing it was observed that the size of the band observed was slightly shifted as expected (Figure 5.1).

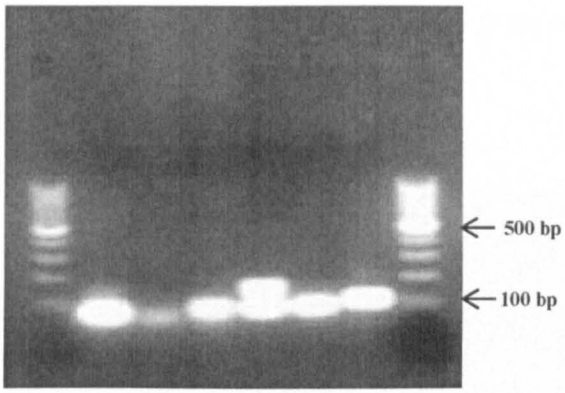


Figure 5.1. Agarose gel electrophoresis analysis of synthetic A- and B-chain DNA fragments. Lane 2 is sense strand of insulin A-chain, lane 3, the antisense A-chain. Lane 4 the annealed A-chain fragment. Lane 5 sense strand of B-chain, Lane 6 antisense B-chain, Lane 7 annealed B-chain fragment. Molecular size marker of 100 bp size was loaded in 1st and 8th lanes.

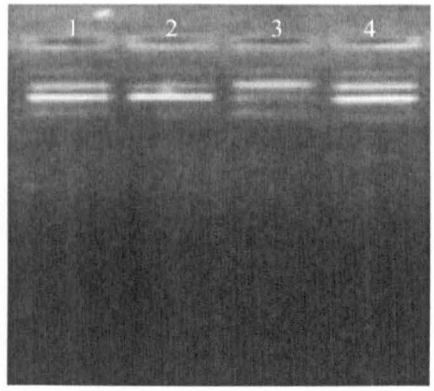


Figure 5.2. Agarose gel electrophoresis analysis of ligation reaction products. The pcDNA5 FRT/A-chain, pcDNA5 FRT/B-chain, pcDNA3.1V5 His/A-chain and pcDNA3.1v5 His/B-chain were loaded respectively into the 4 lanes.

The pcDNA3.1v5His and pcDNA 5FRT vectors were digested with XhoI and NotI restriction enzymes and then gel purified before being used in a subsequent ligation reaction. The ligation reaction was analysed on a 1.5% agarose gel and the result is shown in figure 5.2 and shows 4 different bands for each reaction which indicates the presence of super coiled, linear, wild type and recombinant vectors (Figure 5.2).

After transformation there were numerous colonies observed on all LB/ampicillin culture plates. Thirty colonies from each transformation were sub-cultured in 5 ml LB/ampicillin. To identify the positive subcultures, DNA from all sub-cultures was crudely extracted and then PCR amplified using T7 and BGH primers. PCR screening of the sub-cultures showed an increase in the size of the PCR product in some of the clones which was indicative of positive ligation. However, the recombination rate of this approach to include the target DNA sequence was low. From the thirty sub-cultures screened, only one positive vector construct was obtained of the pcDNA5 FRT/A-chain (Figure 5.3) and pcDNA3.1/B-chain (Figure 5.4), while two positive constructs were

found with pcDNA3.1 v5His/A-chain (Figure 5.5) and pcDNA5 FRT/B-chain (Figure 5.6).

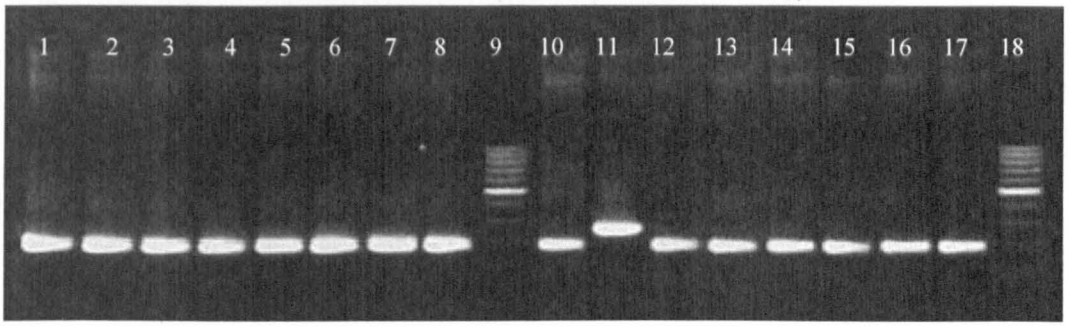


Figure 5.3. PCR amplification screening of clones after the transformation of pcDNA 5 FRT/A-chain ligation. Lane 11 shows one positive clone, 100 bp molecular size marker was loaded in lane 9 and 18.

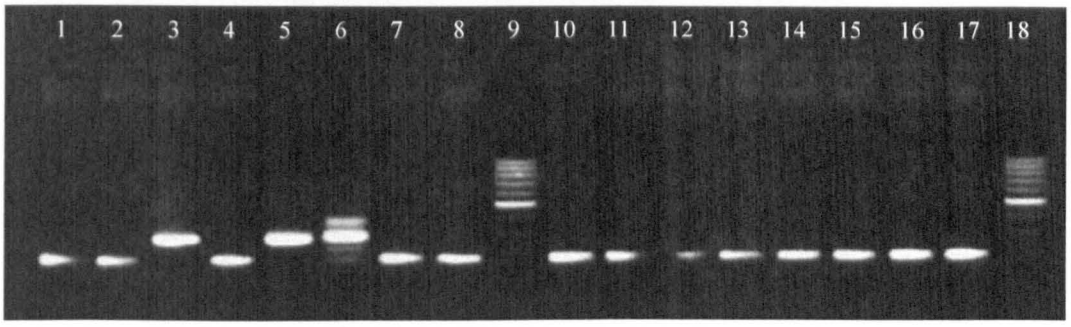


Figure 5.4. PCR amplification screening of clones after the transformation of pcDNA 5 FRT/B-chain ligation. Lane 3 and 5 show two positive clones, 100 bp molecular size marker was loaded in lane 9 and 18.

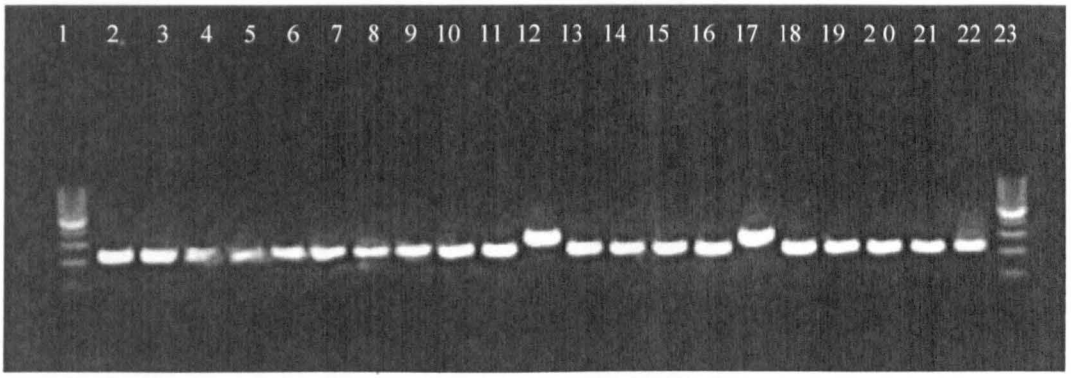


Figure 5.5 PCR amplification screening of clones after the transformation of pcDNA 3.1V5His/A-chain ligation. Lane 12 and 17 show two positive clones, 100 bp molecular weight marker was loaded in lane 1 and 23.

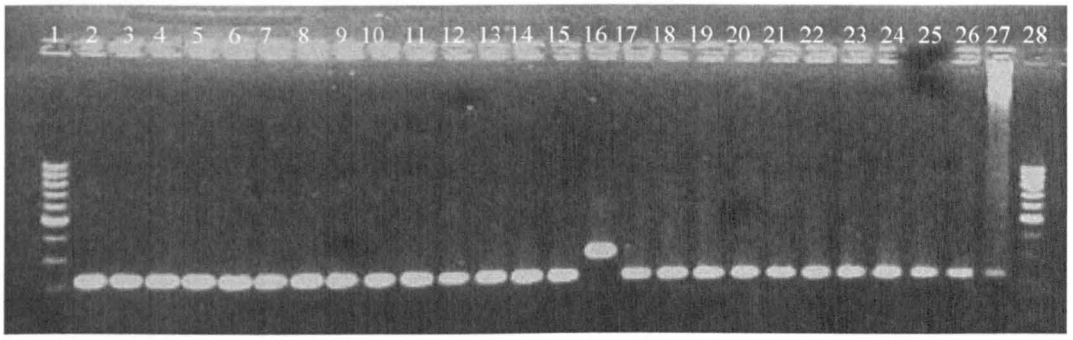


Figure 5.6. PCR amplification screening of clones after the transformation of pcDNA 3.1V5His/B-chain ligation. Lane 16 shows one positive clone, 100 bp molecular size marker was loaded in lane 1 and 28.

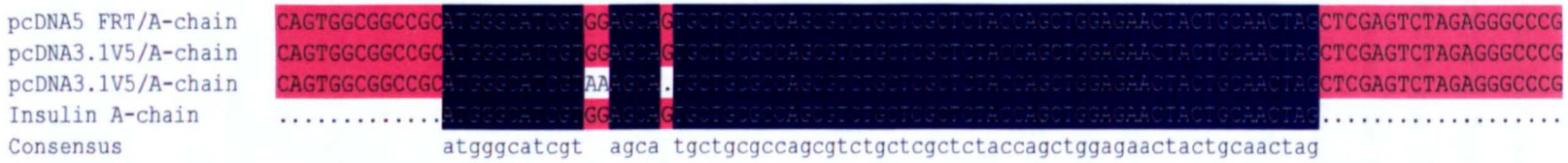


Figure 5.7 The alignment of the sequencing results of the A-chain constructs with the camel insulin sequence obtained and reported in Chapter 3 of this thesis using DNAMAN software. The second construct of pcDNA3.1v5/A-chain shows a mutation in the 12th and 13th nucleotides. The 18th nucleotide was also missing and hence this construct was discarded.

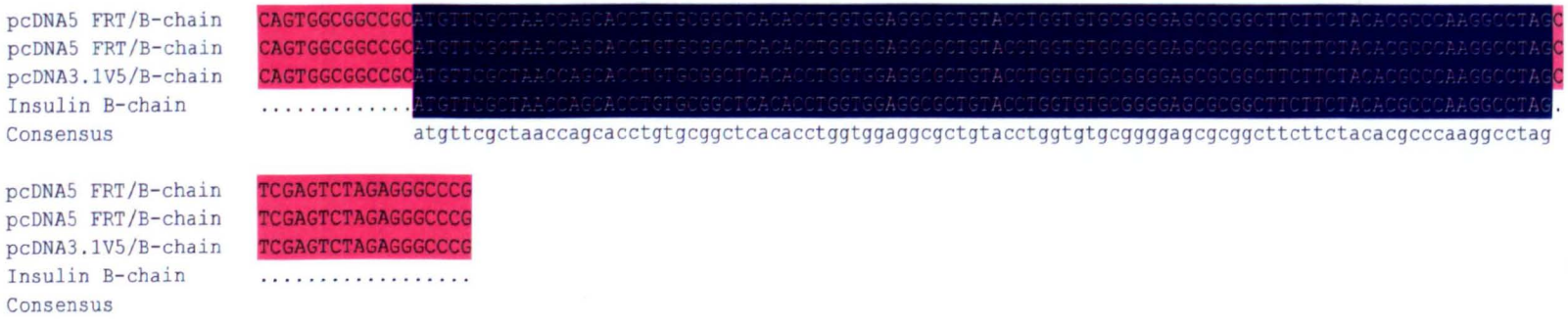


Figure 5.8 The alignment of the sequencing results of the B-chain constructs with the camel insulin gene sequence obtained and reported in Chapter 3 of this thesis using DNAMAN software.

The positive clones as shown in Figures 5.3-5.6 were sequenced using a forward T7 primer and the sequencing results were also confirmed by another sequencing reaction with a reverse BGH primer. The forward and reverse sequences matched each other in all cases. The sequencing results of the constructs were aligned using DNAMAN software (version 5.2.2) the result was resembled to the camel insulin sequence reported in chapter 3 and the results are shown in figure 5.7 and 5.8.

From the sequencing results it was found that one of the pcDNA3.1v5His/A-chain constructs contained a mutation in the 12^{th} and 13^{th} nucleotides, two guanine nucleotides were replaced with two adenine nucleotides, the 18^{th} nucleotide was also missing thus this construct was excluded. The successful and correct vector constructs were then purified from the positive sub-cultures having the correct sequence using the commercially available Wizard® Plus SV Minipreps DNA Purification System (Promega, USA) and approximately 10 μg of vector DNA was purified from each construct.

5.3.1.2 Cloning of camel insulin using a PCR product

As an alternative to using synthetic oligos to clone and produce recombinant camel insulin, a PCR product was also generated and used. Nine micrograms of total RNA isolated from camel pancreatic tissue with TRIzol reagent. cDNA was then prepared from this RNA using the commercially available ImProm-II Reverse transcription system. Preproinsulin and proinsulin DNA fragments were amplified from the pancreatic cDNA described above. Based on the designed primers the amplified products were such that they either contained a stop codon of were without a stop codon and hence would read into the V5 tag of the vector and produce a tagged version of the protein.

The cloning of the insulin gene into the pcDNA3.1v5His and pcDNA FRT vectors was undertaken into two restriction multiple cloning sites. The first cloning attempt was to

insert the PCR product into the NotI and XhoI restriction sites, however the cloning efficiency was very low. A second attempt was undertaken to clone the PCR product into the HindIII and XhoI restriction site using a new set of primers that gave a HindIII overhang.

5.3.1.2.1 Cloning of the camel insulin gene using the NotI and XhoI restriction sites

The PCR primers used were designed to have a NotI restriction site in the forward primer and a XhoI restriction site in the reverse primer. The primers included 3 extra nucleotides in each primer prior to the restriction site in order for the restriction enzyme to cut efficiently. A single band of approximately 350 bp (the actual band size is 353 bp which includes 333 bp from the preproinsulin and 20 bp for the restriction sites and the extra 3 bases added to each end) was observed from the PCR for the preproinsulin when analysed by agarose gel electrophoresis (Figure 5.9). The proinsulin primers produced a fragment of size around 280 bp following PCR (the actual band size is 284 bp which includes 261 bp from the proinsulin, 3 bp for the start codon and 20 bp for the restriction sites and the extra 3 bases added to each end). Figure 5.9 shows the analysis of the amplified PCR products of preproinsulin and proinsulin with and without stop codon by agarose gel electrophoresis.

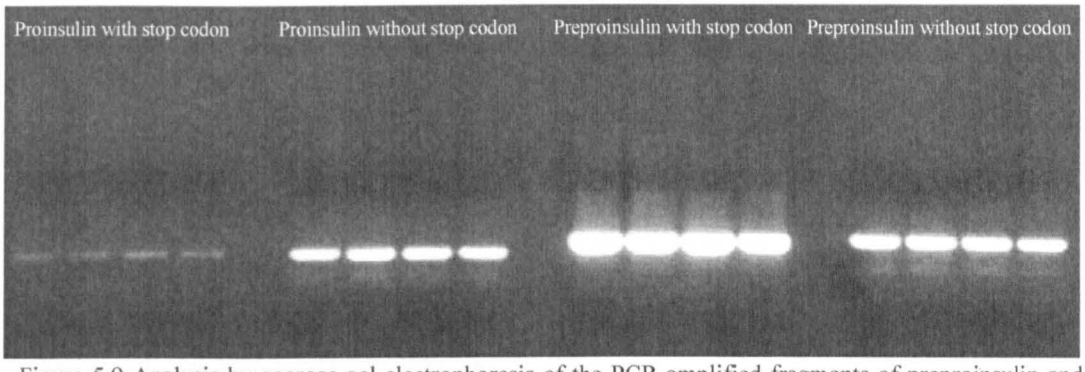


Figure 5.9 Analysis by agarose gel electrophoresis of the PCR amplified fragments of preproinsulin and proinsulin with and without stop codon from camel pancreatic cDNA.

As a first step to verify the specificity and correct amplification of the desired target of the PCR, the PCR product was digested with the SmaI restriction enzyme. The SmaI enzyme will cut the preproinsulin in two places, cutting the preproinsulin DNA fragment at position 67 and 215 to produce 3 digest fragments of size 77, 127 and 148 bp. Although the SmaI restriction enzyme digests at position 67, the observed product in this case is 77 bp in size due to the presence of the additional nucleotides. Figure 5.10 shows the restriction digestion of the preproinsulin PCR product with SmaI when analysed by agarose gel electrophoresis. The 127 and 148 bp bands are clearly visible but diffused, however the 77 bp band and a faint band of undigested product are also observed (Figure 5.10).

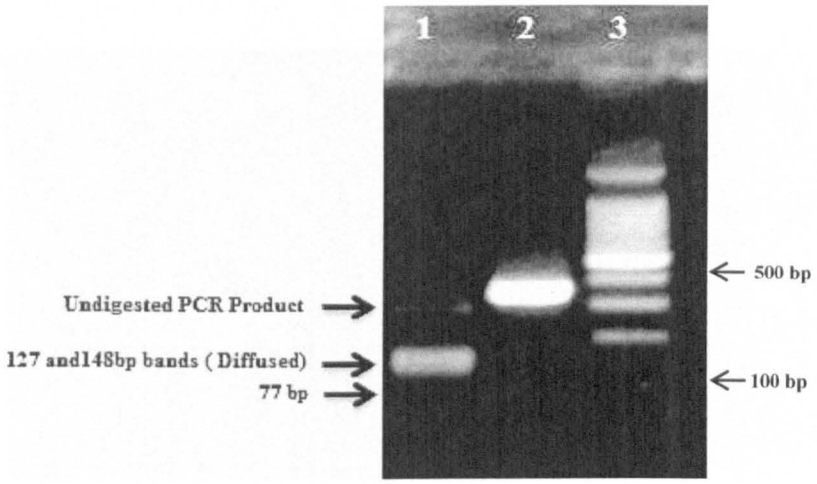


Figure 5.10. Restriction digestion of preproinsulin fragment with Sma I enzyme for sequence confirmation. In 1st column the digested product, 2nd column preproinsulin fragment and 3rd column is 100 bp molecular size marker.

The vector and PCR products were then double digested with the NotI and XhoI enzymes and gel purified. The preproinsulin and proinsulin fragments with and without stop codon were then ligated into both the pcDNA3.1v5His and pcDNA5 FRT vectors. As a result 8 different constructs were prepared which were; pcDNA3.1v5His Proinsulin with stop codon, pcDNA3.1v5His Proinsulin without stop codon, pcDNA3.1v5His Preproinsulin with stop codon, pcDNA3.1v5His Preproinsulin without stop codon, pcDNA5 FRT Proinsulin with stop codon, pcDNA5 FRT proinsulin without stop codon,

pcDNA5 FRT Preproinsulin with stop codon and pcDNA5 FRT/Preproinsulin without stop codon.

After transformation numerous colonies were observed on all culture plates and a number of colonies were sub-cultured in LB-broth. The DNA from the sub-cultured colonies was crudely extracted and screened using PCR with T7 and BGH primer set. Initially, 48 colonies were sub-cultured from each of the 8 vector combinations. From the 48 sub-cultures of each of the vector combinations, only one positive vector construct was found in pcDNA5 FRT Preproinsulin without stop codon (Figure 5.11).

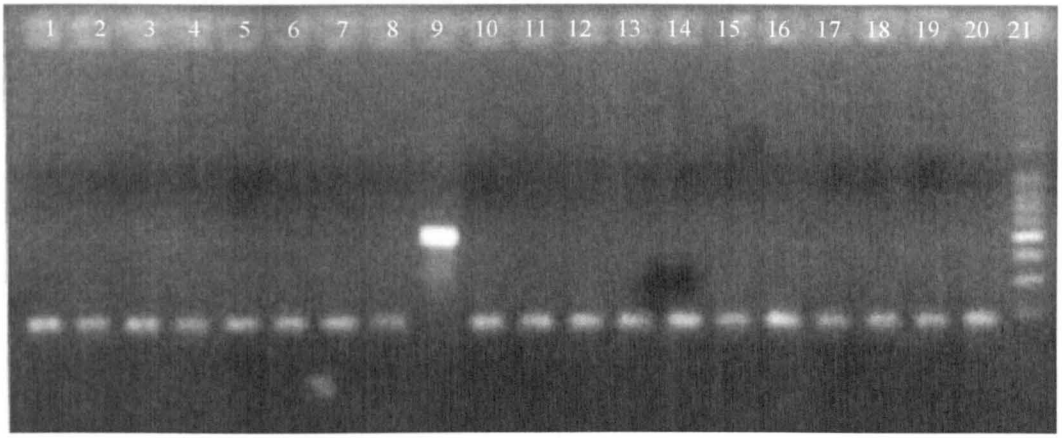


Figure 5.11. Screening by PCR for positive clones containing the preproinsulin insert with pcDNA 3.1 FRT vector without a stop codon (9^{th} lane positive clone). In the last lane 100 bp molecular size marker.

Forty eight more colonies from each of the remaining 7 construct combinations were also sub-cultured and screened for the presence of the correct insert by PCR. From this large screen only one further positive vector construct for the pcDNA3.1V5 His Preproinsulin without stop codon was found (Figure 5.12). Below the expected band a further band was also present (Figure 5.12). Both bands (the expected size band and the band below it) were gel purified and sequenced. The sequencing results of the weak band confirmed that it was a nonspecific PCR product.

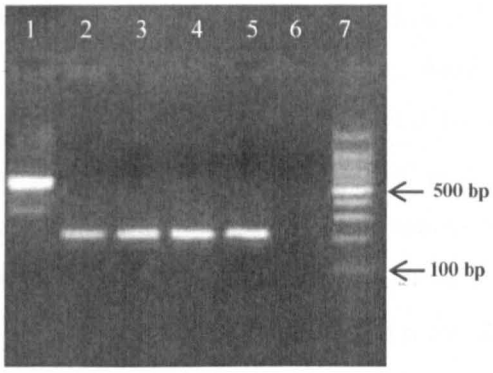


Figure 5.12. PCR screening for presence of correct insulin gene sequence into vector contructs. In lane 1 is a positive vector construct of pcDNA 3.1V5 Preproinsulin without stop codon. The weak band in lane 1, below the band of interest (~350 bp band) is a non-specific PCR product as confirmed by sequencing. One hundred base pair molecular size marker is loaded in the last lane.

One final screen of an additional 48 colonies from each of the remaining 6 potential constructs were then sub-cultured and analysed by PCR screening for the presence of the correct insert. Among all the sub cultures, only one further positive vector construct was found - the pcDNA5 FRT Proinsulin without stop codon. Figure 5.13 shows the screen that identified this positive clone, the size of the positive construct PCR band being more than 260 bp higher than the wild type vector.

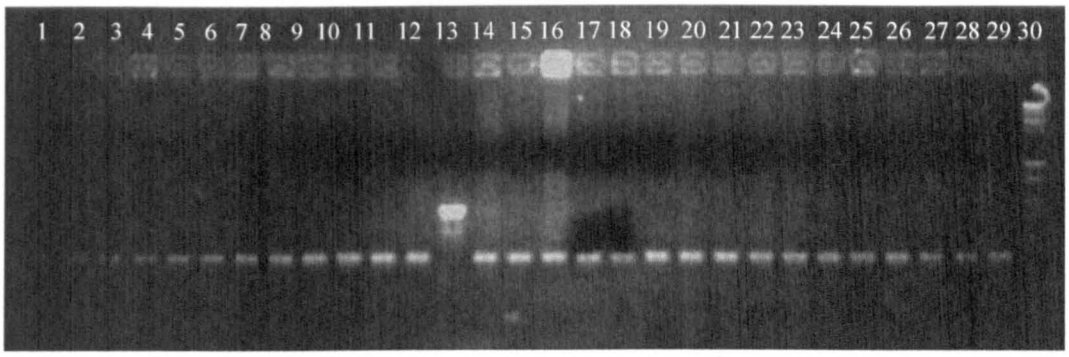


Figure 5.13. PCR screening of colonies for the presence of correct insulin gene inserts. In the 13th lane is a positive vector construct of pcDNA5 FRT Proinsulin without stop codon. One hundred base pair molecular size marker is loaded in the last lane.

To further confirm the correct seqeunce was present in the positive vector constucts, the cloning site and insert of the 3 positive vector constructs were sequenced using a forward T7 primer. The sequencing results were confirmed by another sequencing reaction with a BGH reverse primer. The sequence of camel preproinsulin which was obtained and described in chapter 3 was aligned with the sequencing results from the preproinsulin positive vector constructs of pcDNA5 FRT and pcDNA3.1V5 His using DNAMAN software version 5.2.2. The seqencing and alignment showed no sequence differences between the cloned gene and the genomic DNA sequence obtained in chapter 3 confirming the presence of the correct seqeunce (Figure 5.14).

The 3 positive vector constructs were sub-cultured in 10 ml LB-broth. Approximately 10 µg of DNA of each of the 3 sub cultures was then subsequently purified using the commercially available Wizard® Plus SV Miniprep DNA Purification System (Promega, USA). Based on the low number of positive colonies obtained from screening many colonies, the recombination rate using the NotI and XhoI restriction sites is most likely very low under the conditions used during this study.

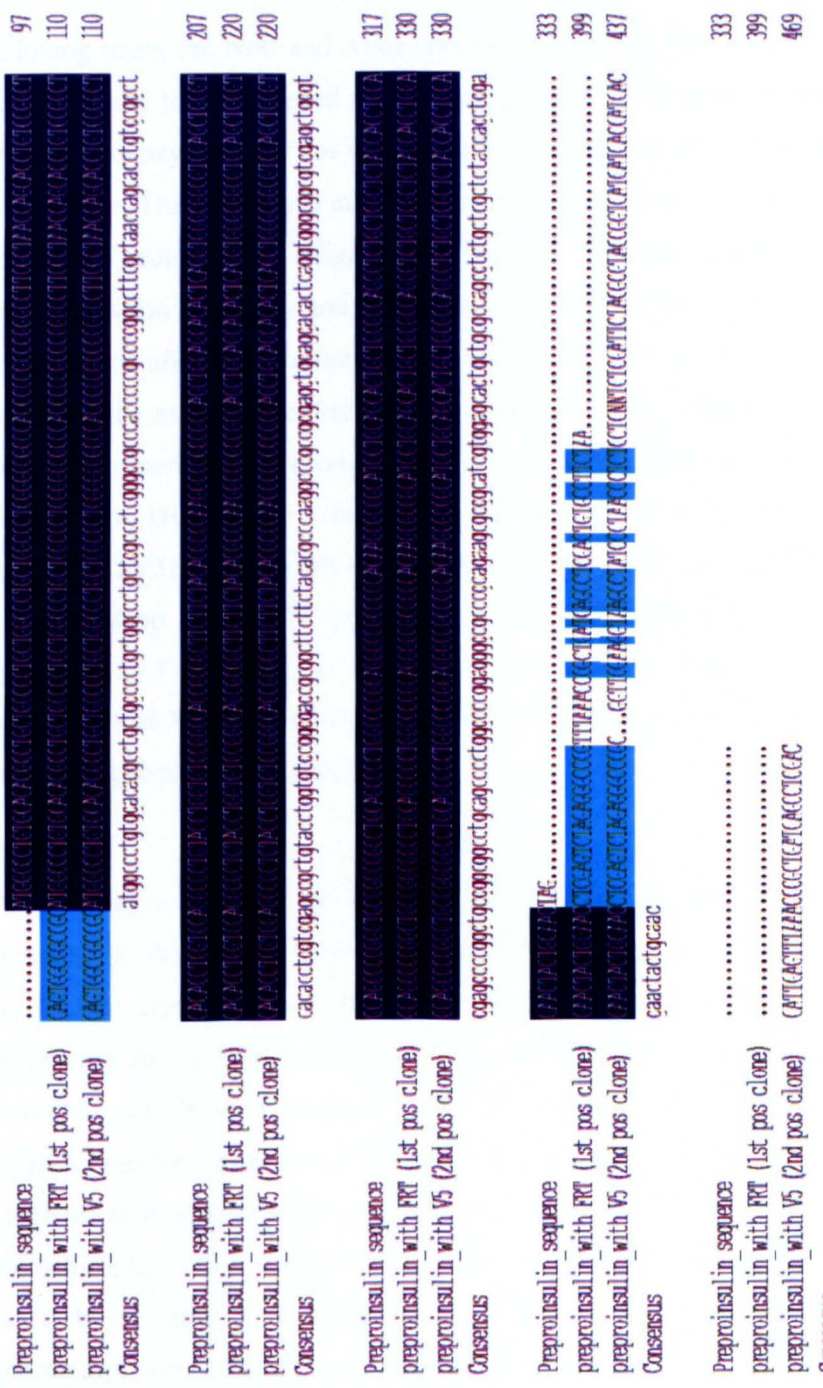


Figure 5.14 DNA sequence alignment of the camel preproinsulin sequence and the preproinsulin positive vector construct of pcDNA5 FRT and pcDNA3.1V5 His using DNAMAN software.

5.3.1.2.1 Cloning of the camel insulin gene using HindIII and XhoI restriction sites

Cloning using the NotI and XhoI restriction sites was inefficient and large numbers of colonies had to be screened to find any positive, PCR insert containing colonies. The restriction enzymes may not work efficiently if the two restriction sites are very close to each other. This may leave many acceptor vectors cut at only a single site and the single digested vectors are religated during the ligation reaction. This reduces the recombination efficiency and results in many of the observed colonies being of the wild type vector after transformation. The XhoI and NotI restriction sites are next to each other on the multiple cloning site of pcDNA3.1V5His and pcDNA5 FRT vectors. In an attempt to increase the recombination efficiency, the NotI cloning site was not used and instead the HindIII site was used. According to the vector sequence of the circular pcDNA3.1V5His and pcDNA5 FRT, the distance between HindIII and XhoI restriction sites is 68 bp in the pcDNA5 FRT. The distance between the two sites in the circular pcDNA3.1V5 His vector is approximately 50 bp. This should allow the two enzymes (Hind III and XhoI) to cut efficiently and remove any potential limitations resulting from using neighbouring restriction sites.

In addition to changing the cloning sites used, the new primers for Hind III and XhoI cloning of the insulin gene were designed to have 6 extra nucleotides prior to the restriction site instead of the 3 nucleotides previous used to increase the digestion efficiency of the PCR products. The PCR was undertaken using the pancreatic cDNA template which been prepared previously. Figure 5.15 shows the camel pancreatic cDNA amplified fragments of preproinsulin with stop codon and the proinsulin with and without stop codon. A band of size approximately 350 bp (the actual size is 337 bp including the 333 bp from the preproinsulin and 24 bp for the restriction sites and the extra bases added to each end) was observed for the preproinsulin PCR product. The proinsulin primers produced a fragment approximately 280 bp (the actual size is 288 bp, 261 bp from the proinsulin, 3 bp from the start codon and 24 bp for the restriction sites and the extra bases added to each end).

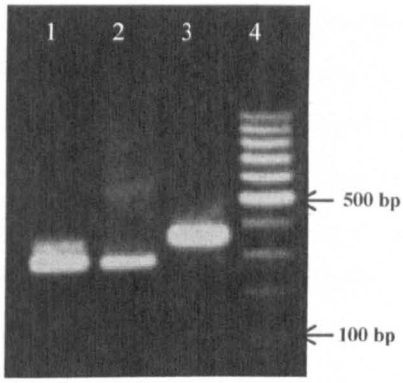


Figure 5.15. The PCR amplified fragments of proinsulin with stop codon in lane 1, proinsulin without stop codon in lane 2, preproinsulin with stop codon in lane 3. PCR products were amplified from camel pancreatic cDNA. In the 4th lane is the 100 bp molecular size marker.

The acceptor vector and PCR products were double digested with HindIII and XhoI enzymes and subsequently gel purified. The preproinsulin with stop codon and the proinsulin fragments with and without stop codon were then cloned into both pcDNA3.1v5 His and pcDNA5 FRT vectors. Thus, the constructs prepared using this cloning strategy/set of restriction sites were; pcDNA3.1v5 His Proinsulin with stop codon, pcDNA3.1v5 His Proinsulin without stop codon, pcDNA3.1v5 His Preproinsulin with stop codon, pcDNA5 FRT Proinsulin with stop codon and pcDNA5 FRT Preproinsulin with stop codon.

After transformation only a few colonies were observed on all culture plates and these were sub-cultured in LB-broth. The DNA from the sub-cultured colonies was crudely extracted and then screened using PCR with the T7 and BGH primer set. The recombination efficiency was much higher than observed with the NotI/XhoI restriction combination. Figure 5.16 shows the results of the cloning using HindIII and XhoI restriction sites as determined by PCR screening of colonies for the presence of the inserts.

To verify the correct inserts were present in the clones obtained, PCR products from 2 of each of the positive vector constructs were treated with exonuclease and shrimp alkaline

phosphatase to remove the extra primers and dntp's and then directly sequenced from both directions using T7 forward primer and BGH reverse primer. The resulting sequences were compared with the sequence of the full insulin gene obtained in chapter 3 and all sequences were found to be correct. One colony from each of the verified (sequenced) colonies was selected and sub-cultured in 250 ml LB-Broth. The vector constructs were then isolated using the commercially available Maxi-prep Kit (Qiagen). Around 50 µg of vector DNA were isolated from each of the five vector constructs.

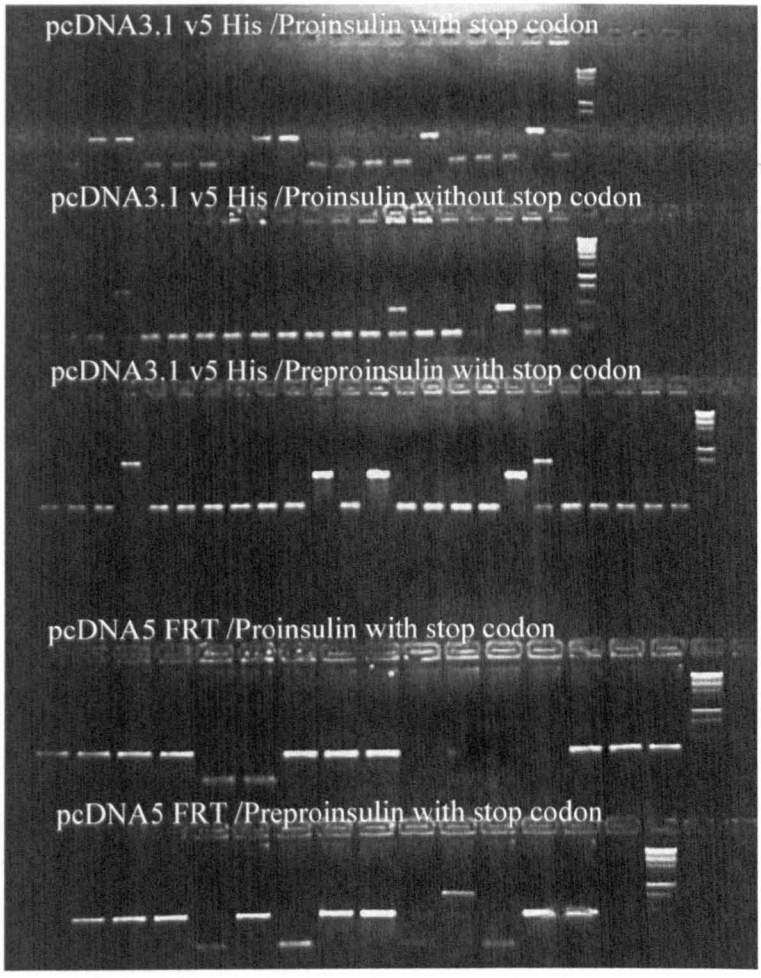


Figure 5.16. PCR screening of colonies from the cloning of pcDNA3.1V5 and pcDNA5 FRT vectors with proinsulin and preproinsulin using HindIII and XhoI restriction sites.

5.3.2 The expression of camel insulin in the Chinese hamster ovary (CHO) cell line

In this project the expression of recombinant camel insulin from CHO cells was attempted. Three different transfection experiments were performed, two transient transfections - one with insulin A- and B-chains and the other with proinsulin and preproinsilin; and one stable transfection with proinsulin and preproinsulin.

5.3.2.1 Transient transfection and expression of camel milk insulin

5.3.2.1.1 Transient transfection of insulin A- and B-chains

Four expression vectors consisting of pcDNA3.1V5 and pcDNA5 FRT vector constructs with insulin A- and B-chains were transiently transfected in Chinese hamster ovary (CHO) cells using Lipofectamine 2000 (Invitrogen) transfection reagent in 6 well culture plates. The transfections of A- and B-chains were setup separately in either the pcDNA3.1 or pcDNA5 format. As these chains do not contain signal sequences they would be expected to be synthesised in the cytosol and not secreted out of the cell.

Western blotting analysis was performed to determine whether the insulin chains were being produced in the transiently transfected cells. The transfected cells were trypsinised and then lysed in RIPA buffer 48 hours after transfection. The cell lysate and its respective cell culture media were analysed by western blotting using an anti-insulin antibody (Diasource, Belgium). Figure 5.17 shows the results of western blotting where no signal was observed matching the control insulin using this approach. This agreed with the results from the purification of insulin from camel milk whereby the low levels of insulin present were not detected but could be detected by radioimmunoassay (RIA). RIA was therefore also used to determine if the insulin chains were being produced and to measure the insulin concentration in the cell lysate and the respective culture media. The sensitivity of the radioimmuno assay is higher than western blotting and can detect as little as 1 μIU/ml of insulin.

The RIA results are presented in Table 1 and confirm the lack of insulin chain production in the cell lysate and cell culture medium in the all transfected cell lines.

Table 5.1: The concentration of insulin as determined by radioimmunoassay for the transiently transfected cell lines with pcDNA3.1 V5 and pcDNA5 FRT vector constructs with camel insulin A- and B-chains.

Transient transfection	Vector	Source	Insulin (µIU/ml)
Insulin A-chain	pcDNA 3.1 V5	Cell lysate	2
Insulin B-chain	pcDNA 3.1 V5	Cell lysate	1
Insulin A-chain	pcDNA 5 FRT	Cell culture media	2
Insulin B-chain	pcDNA 5 FRT	Cell culture media	2
Untransfected cells control	NA	Cell lysate	1
Untransfected cells control	NA	Cell culture media	2

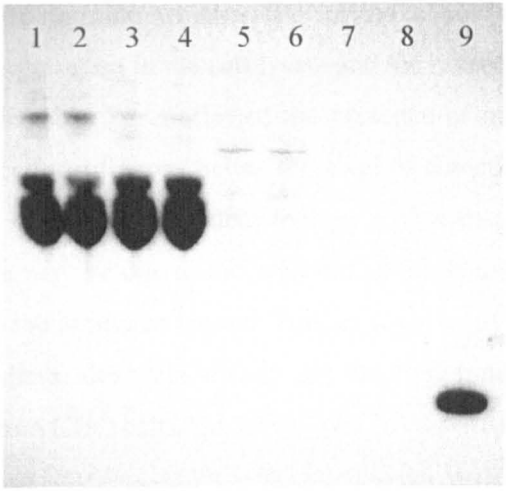


Figure 5.17: Western blotting analysis using anti-insulin antibody of the 4 transiently transfected experiments with pcDNA3.1V5 and pcDNA5 FRT vector constructs with insulin A- and B-chains. Lane's 1-4 cell culture media, Lane's 5-8 cell lysate, lane 9 is insulin positive control.

5.3.2.1.2 Transient transfection and expression of camel proinsulin and preproinsulin in CHO cells

Four expression vectors of pcDNA3.1V5 vector constructs consisting of preproinsulin and proinsulin with and without stop codon were transiently transfected into an Chinese hamster ovary (CHO) cell line using Lipofectamine 2000 (Invitrogen) transfection reagent in 6 well culture plates. Western blotting was used to deteremine whether insulin production was present in the transiently transfected cells. Forty eight hours after transfection culture media was aspirated and the transfected cells were scraped and lysed in extraction buffer. The cell lysate and its respective cell culture media were analysed by western blotting using a commercially available anti-insulin antibody (Diasource, Belgium). As with the A- and B-chain experiment, no signal was observed matching the size of the control insulin although a higher molecular weight band was observed (Figure 5,18).

RIA was also used to determine whether the insulin chains were produced and to measure the insulin concentration in the cell lysate and the respective culture media. The RIA results presented in table 5.2 confirmed the presence of insulin production in the cell culture. The produced insulin was below the level of detection by western blotting, hence explaining why no insulin was detected using this method. The low level of insulin in the cell lysate may be due to the presence of an insulin degrading enzyme in the cell which degrades the produced insulin. Further work would need to be undertaken to confirm this. Regardless, this data shows, for the first time, that it is possible to express camel insulin from CHO cells.

Table 5.2. The concentration of insulin as determined by radioimmuno assay for the transiently transfected cell lines with pcDNA 3.1 V5 vector constructs with Proinsulin and Preproinsulin.

Transient transfection	Vector	Source	Insulin (µIU/ml)
Proinsulin with stop codon	pcDNA 3.1 V5	Cell lysate	2
Proinsulin without stop codon	pcDNA 3.1 V5	Cell lysate	2
Proinsulin with stop codon	pcDNA 3.1 V5	Cell culture media	47
Proinsulin without stop codon	pcDNA 3.1 V5	Cell culture media	42
Preproinsulin with stop codon	pcDNA 3.1 V5	Cell lysate	1
Preproinsulin without stop codon	pcDNA 3.1 V5	Cell lysate	3
Preproinsulin with stop codon	pcDNA 3.1 V5	Cell culture media	18
Preproinsulin without stop codon	pcDNA 3.1 V5	Cell culture media	70
Untransfected cells control	NA	Cell lysate	2
Untransfected cells control	NA	Cell culture media	1

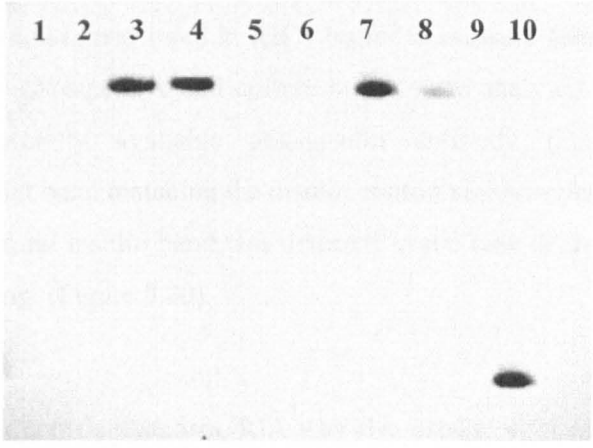


Figure 5.18. Western blotting analysis using anti insulin antibody of the 4 transiently transfected cell lines with pcDNA3.1 V5 vector constructs of Proinsulin and Preproinsulin with and without stop codon. Lane's 1, 2, 5 and 6 the cell lysate, Lane's 3, 4, 7 and 8 cell culture media, lane 9 is the negative control cell lysate and in lane 10 is insulin positive control.

5.3.2.2 Stable transfection of CHO cells with proinsulin and preproinsulin

Whilst the transient studies described above confirmed that camel insulin could be expressed in CHO cells, the amounts observed were low and therefore as an alternative to transient expression stable cell lines were generated. To generate stable cell lines which expressed camel insulin constitutively, four expression vectors of pcDNA5 FTR vector with preproinsulin and proinsulin with and without (to read through to the tag) stop codon were stably transfected into the Chinese hamster ovary (CHO) Flp-In cell line along with the pOG44 vector using Lipofectamine 2000 (Invitrogen) transfection reagent in T25 culture flasks. Hygromycine selection was then applied to the transfected cells to select for those transfected cells that had the plasmid DNA inserted into the recombination site. Distinct cell colonies appeared 6 days after the hygromycim selection and 3 days later the transfected cells were trypsinised, pooled and cultured in T75 flask in the presence of hygromycin. Three days later the cells were transferred to T175 tissue culture flasks.

To confirm that the stably transfected and selected CHO cells were producing insulin, the culture media was collected during each transfer to a larger flask. The transfected cells were also trypsinised and lysed in RIPA buffer to measure for intracellular insulin. The cell lysate and its respective cell culture media were analysed by western blotting using the commercially available anti-insulin antibody (Diasource, Belgium). Encouragingly, a faint band matching the insulin control size was observed (Figure 5.19) in the cell lysate but no insulin band was detected in the case of the cell culture media when western blotting (Figure 5.20).

As in the case of the transient studies, RIA was also used to confirm insulin production and to measure the insulin concentration in the cell lysate and the respective culture media. Table 5.3 reports the results of the RIA measuremenrs. Low levels of insulin were detected in both the cell culture and cell lysate, lower than observed in the transient

studies. The low insulin concentration is again probably due to presence of an insulin degrading enzyme. The time between sample processing and RIA testing was longer than in the transient studeis which increases the opportunity for sample degradation even in cell culture.

Table 5.3. The concentration of camel insulin as determined by radioimmuno assay for the stably transfected cell lines with pcDNA 3.1 FRT vector constructs with proinsulin and preproinsulin.

Stable transfection	Vector	Source	Insulin (μIU/ml)
Proinsulin with stop codon	pcDNA 5 FRT	Cell lysate	3
Proinsulin without stop codon	pcDNA 5 FRT	Cell lysate	4
Proinsulin with stop codon	pcDNA 5 FRT	Cell culture media	4
Proinsulin without stop codon	pcDNA 5 FRT	Cell culture media	5
Preproinsulin with stop codon	pcDNA 5 FRT	Cell lysate	4
Preproinsulin without stop codon	pcDNA 5 FRT	Cell lysate	4
Preproinsulin with stop codon	pcDNA 5 FRT	Cell culture media	5
Preproinsulin without stop codon	pcDNA 5 FRT	Cell culture media	4
Untransfected cells control	NA	Cell lysate	1
Untransfected cells control	NA	Cell culture media	1

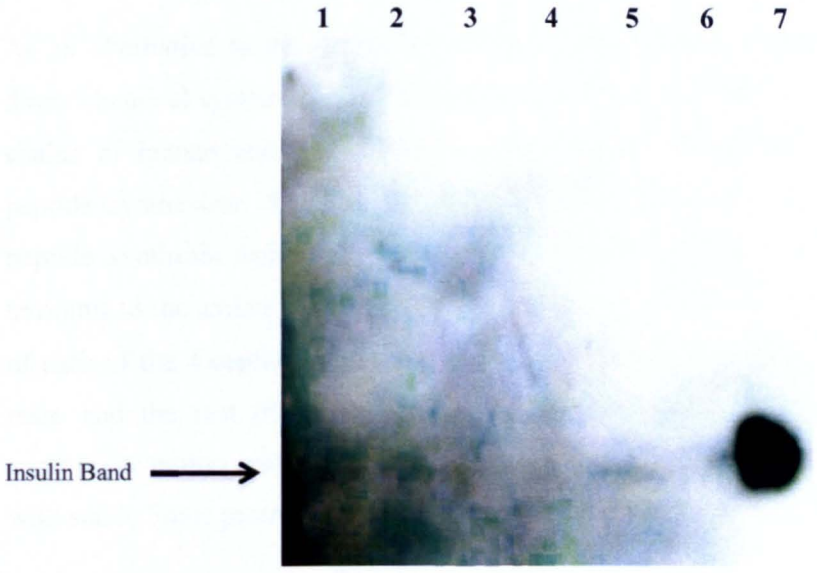


Figure 5.19. Western blotting using anti-insulin antibody of cell lysate of the 4 stably transfected cell lines with pcDNA 5 FRT vector constructs of proinsulin and preproinsulin with and without stop codon. A faint insulin band (as indicated by the arrow) was present in lanes 1, 2, 3 and 5 where the cell lysate of proinsulin with stop codon, proinsulin without stop codon, preproinsulin with stop codon and preproinsulin without stop codon were loaded respectively. Insulin positive control is in lane 7.

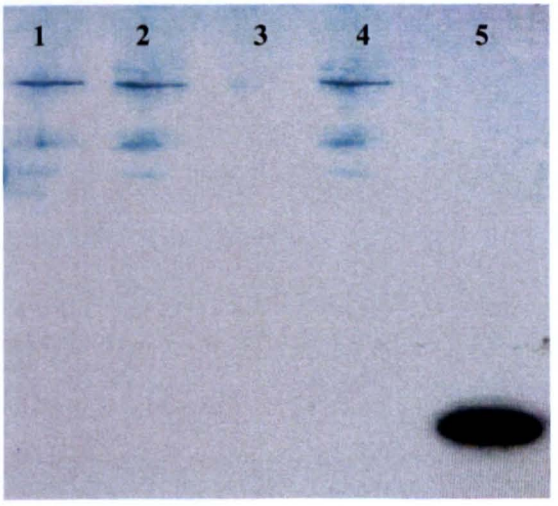


Figure 5.20. Western blotting using an anti-insulin antibody of cell culture media (lane 1-4) of the 4 stably transfected cell lines with pcDNA 5 FRT vector constructs of proinsulin and preproinsulin with and without stop codon. Insulin positive control is in lane 5.

5.3.3 Chemical synthesis of camel insulin peptides

As an alternative to the production of camel insulin from cultured mammalian cells, direct chemical synthesis of the individual chains was also investigated. The A- and B-chains of human and camel insulin were therefore chemically synthesized using a peptide synthesizer. The four different peptides were produced by the solid phase peptide synthesis technique. The peptides synthesis process was from the carboxyl terminus to the amino terminus. The carboxyl group of the carboxy terminal amino acid of each of the 4 peptides was covalently bonded to the solid support of NovaSyn®TGT resin and the rest of the chain was then synthesised by stepwise addition of the appropriate amino acids. The reactive functional group of the amino acids was protected with stable Fmoc protecting group to prevent the formation of secondary chains.

5.3.3.1 Production of thiol-protected insulin B-chain

The protected human and camel insulin B-chains were constructed on NovaSyn®TGT resin by the Fmoc-based solid-phase synthesis method. The thiol group of the cysteine at position 7 in the both human and camel insulin B-chain was protected by the acid-stable acetamidomethyl (Acm) group. Triphenylmethyl (Trt) as an acid-labile protecting group was used to protect the cysteine thio group at position 19 of both peptides.

5.3.3.2 Production of thiol-protected insulin A-chain

For the synthesis of the insulin A-chain, the cysteine thiol group in both camel and human chains was protected by the acid labile Trt group at amino acid position number 20. The trifluroacetic (TFA) stable protecting group Acm was used to protect the cysteine at position 7 and another TFA stable protecting group tert-butyl (t-Bu) was used at positions 6 and 11.

5.3.3.3 Peptide cleavage and purification

After the synthesis process, the 4 peptides were washed with methanol and then cleaved from the resin using TFA and two scavengers; ethanedithiol and triisopropylsilane. The TFA cleaved the peptide from the resin while the scavengers preferentially reacted with free radical and other reactive species released during the cleavage reaction and washed them off. De-protection of the 4 synthetic peptides with TFA in the presence of scavengers produced insulin A and B-chains with a free thiol group present in the last cysteine in all chains.

Following this, after adding diethyl ether, a white precipitate appeared indicating the presence of the peptide. The precipitate was dried after two washes with the diethyl ether. The human and camel B-chains were easily reconstituted in 1 ml of 80% acetonitrile containing 0.1% TFA. The insulin A-chain of both human and camel didn't dissolve in 1 ml of 80% acetonitrile containing 0.1% TFA so a further 1 mL of 0.1% TFA was added. Solid urea was then also added to give an 8 M final urea concentration, the urea still only partially dissolving the insulin A-chain.

The peptides were then analysed and the presence of the correct peptide verified by liquid chromatography coupled mass spectrometry (LC/MS). A large and distinct peak was observed on the reverse phase HPLC chromatogram for each peptide. In the case of the insulin B-chain of both camel and human, the large peak was observed at a retention time of 21 min. The subsequent mass specta collected using electrospray ionization mass spectrometry confirmed the large peak as the full insulin B-chain peptide. The calculated mass of the positively double charged ion of the human insulin B-chain with the protecting groups is 1750.37, while the calculated mass for camel B-chain is 1721.35. The extracted ion chromatogram (Figure 5.21 (Human) and Figure 5.22(Camel)) confirmed the presence of the double charged ion of human and camel insulin B-chain at the retention time matching the large peak reverse phase HPLC chromatogram.

Many peaks were observed on the reverse phase HPLC chromatogram in the case of the insulin A-chain of both camel and human (Figure 5.23 and 5.24, respectively). The mass spectra obtained using electrospray ionization mass spectrometry confirmed the presence of the full insulin A-chain peptide corresponding to one of the many peaks observed in the UV trace from the HPLC (Figures 5.23 and 5.24). The calculated mass of the positively double charged ion of the human A-chain with the protecting groups is 1283.59, while the calculated mass for camel A-chain is 1261.58. The extracted ion chromatogram shown in Figure 5.23 and 5.24 confirmed the presence of the double charged ion of the insulin A-chain of human and camel, respectively, at the retention time of 23 min, which matched a small but broad peak on the reverse phase HPLC chromatogram.

Following confirmation of synthesis of the correct peptide by mass spectrometry, the insulin B-chain of camel and human were purified by reveres phase high performance liquid chromatography (HPLC). The synthetic peptides were injected onto a C_{18} reversed phase preparative column for this purpose and then eluted with acetonitrile. The retention time of the large peak corresponding to the insulin B-chain in the case of the preparative column was 27 min. The fraction of the large peak was collected. Fractions from ten HPLC runs were pooled and lyophilised by freeze drying. From the purification and lyophilisation approximately 12 mg of camel and human insulin B-chain peptide was produced. The lyophilised product was then stored at -20°C until required for further use.

The A-chain insulin peptides were very insoluble and could not be purified by HPLC. Due to a lack of time the A-chain could not be correctly purified and solubilised and hence the process of *in vitro* insulin folding was not performed and the stability study on insulin was carried out with the insulin from camel milk isolated. The details of insulin isolation from camel milk are described in detail in chapter 4.

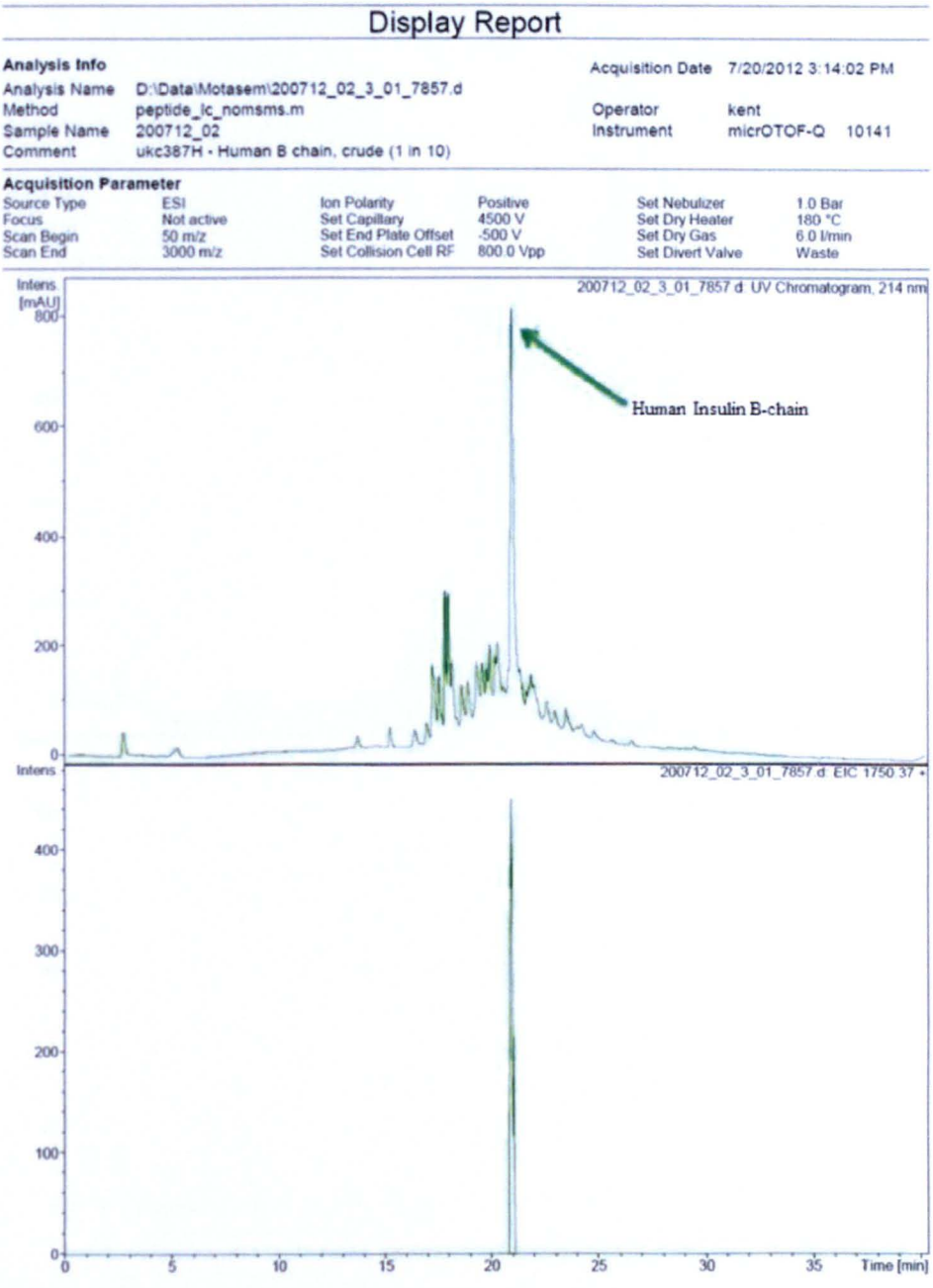


Figure 5.21. Reverse phase HPLC chromatogram and the extracted ion chromatogram at mass 1750.37 of human insulin B-chain. The extracted ion chromatogram confirms the presence of human insulin B-chain at the HPLC retention time of 21 min.

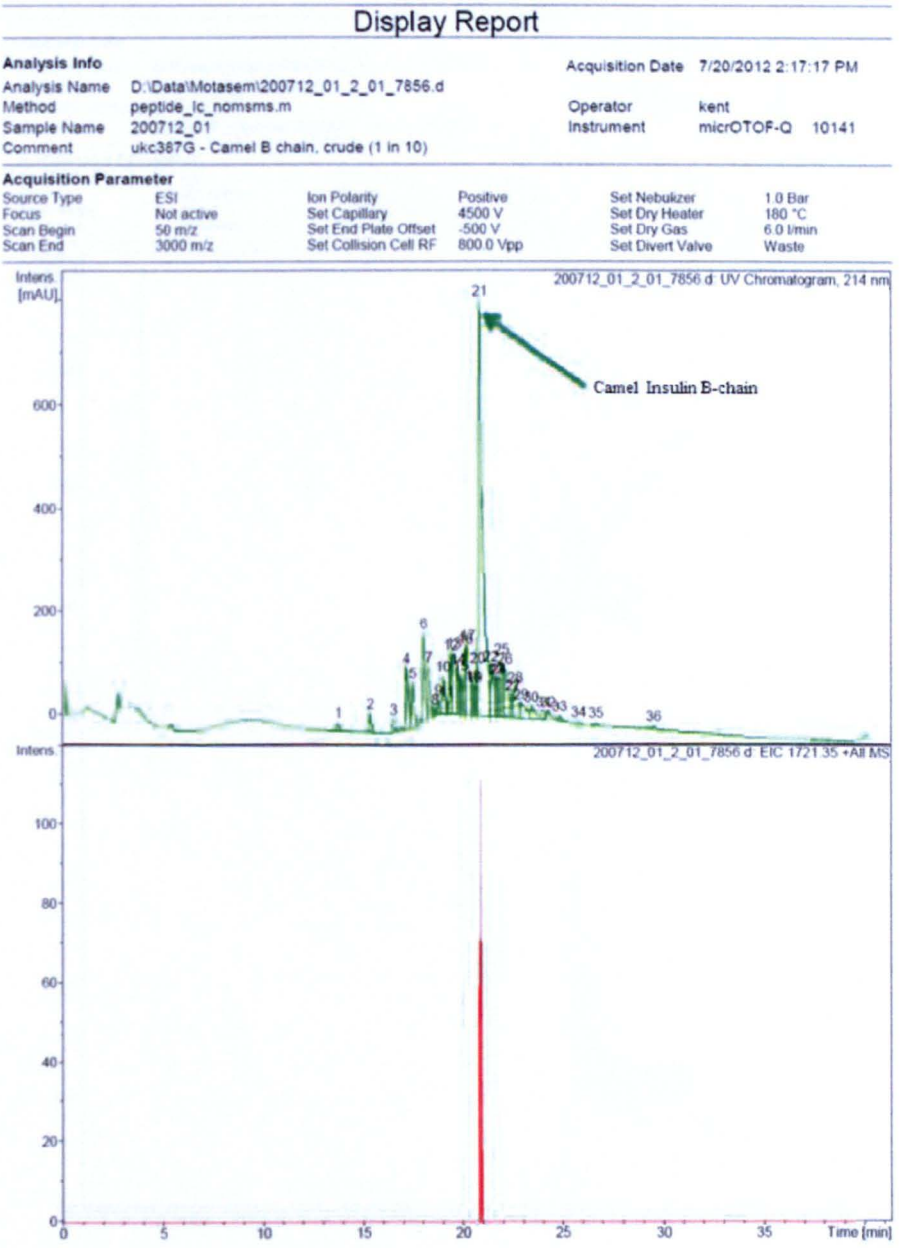


Figure 5.22. Reverse phase HPLC chromatogram and the extracted ion chromatogram at mass 1721.35 of camel insulin B-chain. The extracted ion chromatogram confirms the presence of the camel insulin B-chain at HPLC retention time of 21 min.

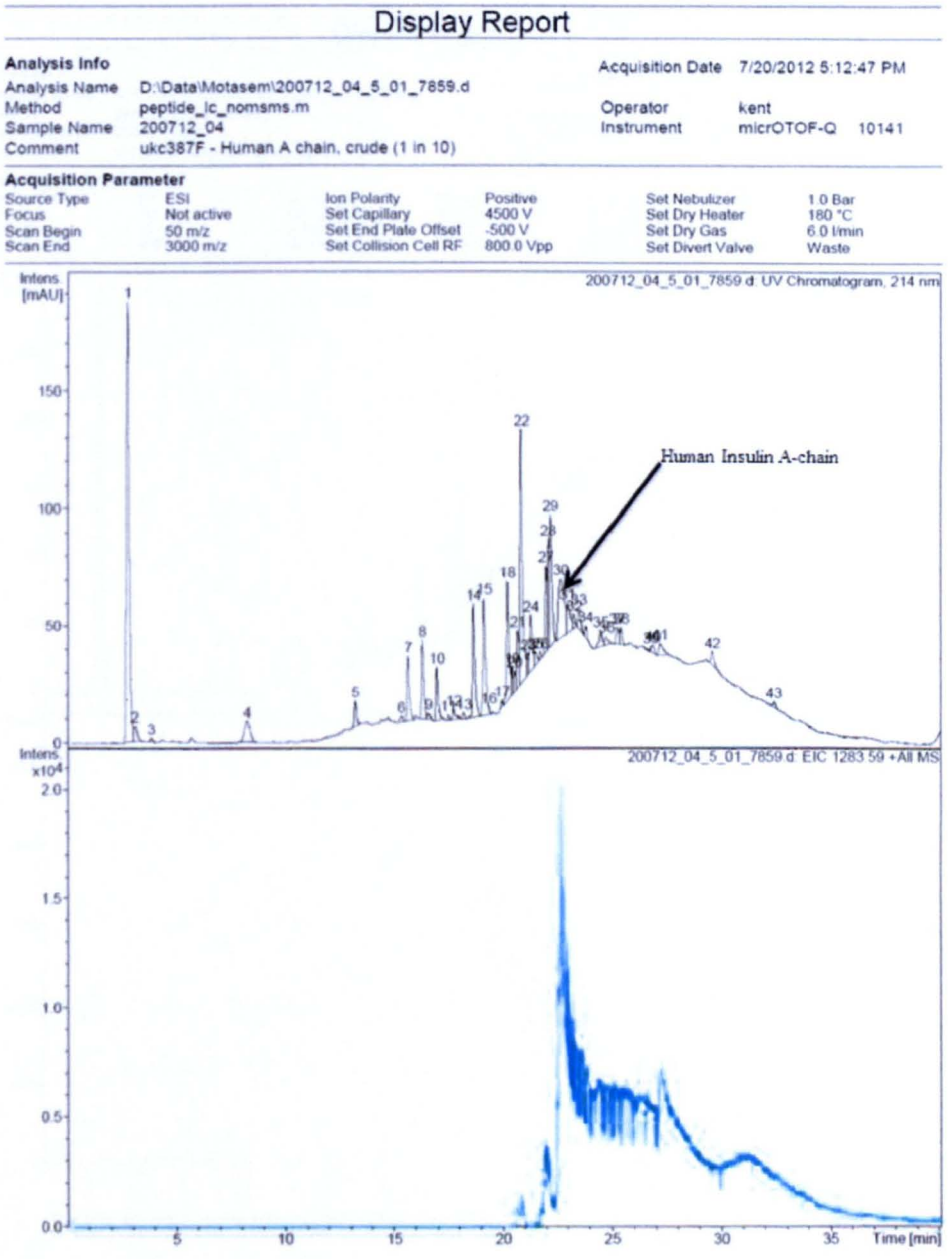


Figure 5.23. Reverse phase HPLC chromatogram and the extracted ion chromatogram at mass 1283.59 of human insulin A-chain. The extracted ion chromatogram confirms the presence of the human insulin A-chain at HPLC retention time of 23 min.

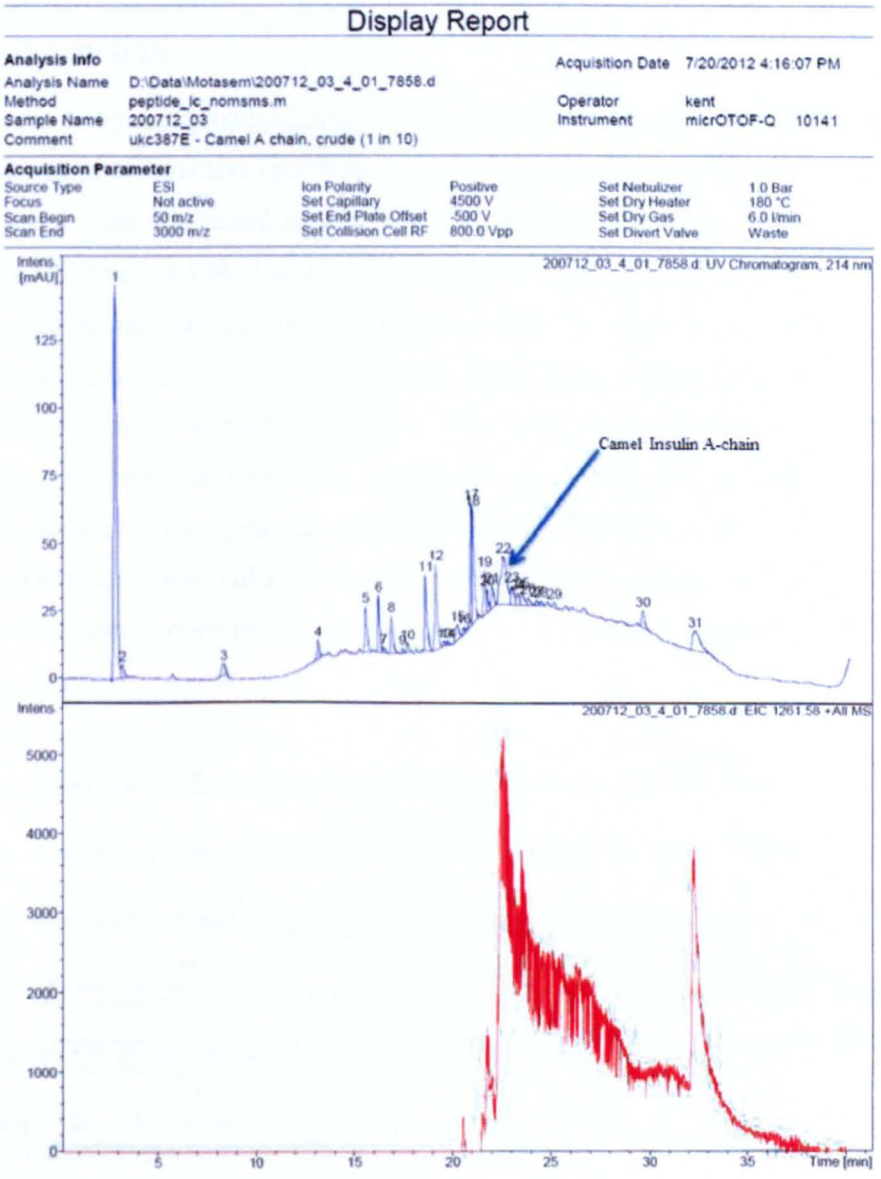


Figure 5.24. Reverse phase HPLC chromatogram and the extracted ion chromatogram at mass 1261.58 of camel insulin A-chain. The extracted ion chromatogram confirms the presence of camel insulin A-chain at HPLC retention time of 23 min.

5.4 Discussion

In order to try and produce camel insulin from an *in vitro* cultured mammalian cell line, mammalian cell vectors (pcDNA3.1 v5His and pcDNA 5 FRT) were used as acceptor vectors to clone the camel insulin gene and insulin peptide DNA fragments. The cloning using the XhoI and NotI restriction sites in the multiple cloning site was difficult and very low recombination efficiency was observed. The XhoI and NotI restriction sites are close to each other in the vector (figure 5.26). This reduces the efficiency of double digestion with the restriction enzymes. The single digested vectors can religate easily without the insert and result in many wild type vector colonies after transformation. After replacing the upstream cloning site with HindIII instead of NotI, the cloning efficiency increased and a few colonies were observed on plates. Furthermore, adding additional bases to the primer allowed the restriction enzymes to cut more completely.

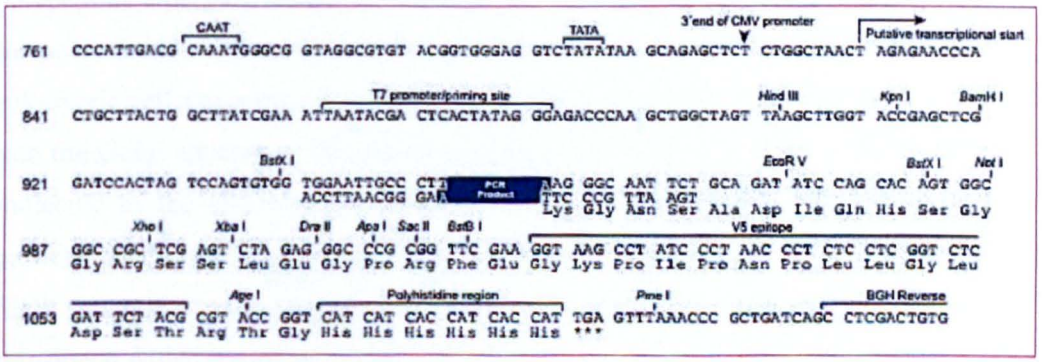


Figure 5.26. The pcDNA3.1v5His vector multiple cloning site. From the Invitrogen website http://tools.invitrogen.com/content/sfs/manuals/pcdna3.1topota_man.pdf

The CHO cell line was initially transiently transfected with insulin A and B-chains and subsequently transiently and stably transfected with proinsulin and preproinsulin. The production of insulin from mammalian cells turned out to be a difficult process and many attempts failed to produce large amounts of recombinant insulin from insulin transfected mammalian cells. Such difficulty in expressing insulin from mammalian cell

lines has been previously reported by others (Groskreutz *et al.*, 1994; Quin *et al.*, 1991 and Vollenweider *et al.*, 1992).

The main reason given for the low yield of insulin from mammalian cell lines is due to the presence of the insulin degrading enzyme in mammalian cells. Insulin degrading enzyme is a neutral thiol metalloprotinases which is involved in switching off the insulin signal by rapidly degrading the excess insulin (Duckworth, 1988). To confirm this was indeed the issue, siRNA knockdown experiments could be undertaken to knockdown the levels of this enzyme and determine whether this results in increased levels of insulin being observed. If this was the case, potentially a knockout cell line could be made to produce insulin in the absence of the degrading enzyme.

The first transfection attempt was undertaken with vectors cloned with insulin A- and B-synthetic chains without the signal sequence. The signal peptide directs the nascent polypeptide emerging from the ribososome to the endoplasmic reticulum (ER) apparatus for post-translational modification and secretion. For the protein to be secreted from a eukaryotic cell, the protein should first enter the endoplasmic reticulum then be transport into the Golgi apparatus. The signal peptide at the beginning of the proteins directs the molecule to the endoplasmic reticulum lumen via binding to the signal recognition particle, SRP. The signal recognition particle then binds to the SRP then binds to the SRP receptor on the surface of the ER and passes the polypeptide to the Sec61 translocon which transports the peptide into the ER. The signal peptide is cleaved off by signal peptidase after serving its purpose of directing the protein to endoplasmic reticulum and is not present in the mature protein (Stern *et al.*, 2007).

The subsequent transfections were performed with preproinsulin and proinsulin cloned vectors. Proinsulin is the precursor of insulin and consists of the insulin B-chain, C-peptide and A-chain while preproinsulin is proinsulin with the signal peptide. For each

vector construct one was made with a stop codon and the other was without. That without should read through to a stop further downstream with a his tag between the end of the insulin and new stop. The tag could have been used for purification purposes should sufficient insulin have been made in the mammalian system.

The RIA results of transient transfection with preproinsulin and proinsulin indicated the presence of insulin in culture media more than the cell lysate. The reduced amount of insulin in the lysate may be due to the presence of insulin degrading enzyme in the cell lysate being at a higher concentration. The western blotting results didn't show an insulin specific band, most likely due to the insulin concentration being below the detection level of insulin. Despite this, the RIA result show that insulin can be produced in CHO cells but that degradation likely rapidly occurs afterwards.

The CHO cell line was used to successfully produce stably transfected cells with preproinsulin and proinsulin using the commercially available flp-in system. Using this system, healthy cells were observed one week after transfection. The expression level of insulin from the transfected cells is likely to be higher than the amount detected by RIA and western blotting due to a number of reasons including:

- Low post-translational modification from proinsulin to insulin; the antibody used in the RIA kit and western blotting is specific to the mature insulin.
- Low secretion levels of insulin, the intracellular insulin is probably degraded rapidly.
- The insulin degrading enzyme my also be secreted out of the cell and degrade the secreted insulin. Even if it is not secreted, many proteins are released into the media during cell death and as such it is likely to be released into the media and result in insulin degradation.

- Potential for technical problems, for instance; the samples were stored at -20°C before the RIA test instead of -80°C and they been shipped to the hospital for RIA analysis at 4°C.

Several methods could be used to enhance the level of the expressed insulin:

- Transfect the insulin gene in a cell line generated from β-cells of pancreas.
- Collect the culture media of the insulin transfected cells after 24 hours of culturing. The insulin peak concentration was found 24 hours after culturing the cells in other studies (Pak *et al.*, 2002)
- Include bacitracin antibiotic with the culture media. Bacitracin is an antibiotic which also inhibits the insulin degrading enzyme. This antibiotic was used by Pak and colleagues in 2002, after the bacitracin treatment the level of the recombinant insulin produced by CHO transfected cells was enhanced.
- Knockdown the degrading enzyme by RNAi to reduce/eliminate the potential problem. Ultimately generate a knockout of the enzyme of the knockdown approach was shown to be viable. This is likely to be the best solution for the long term elimination of the potential problematic enzyme.

As an alternative to production from mammalian cells, solid phase peptide synthesis was successfully used to synthesise the insulin peptides (A and B-chains). LC/MS verified the correct peptides had been synthesised. The broad A-chain peak in the HPLC and the extracted ion chromatograms were due to the low solubility of the insulin A-chain in the acidic conditions used which proved problematic in the purification and handling of the peptide. Although the A-chain peptide was solubilised by pH adjustment to basic condition before LC/MS analysis, the HPLC purification requires acidic buffers including acetonitrile containing triflouroacitic acid.

5.5 Conclusion:

From the results of this chapter it can be concluded that the camel insulin can be produced by Chinese's hamster ovaries cell lines using the transient and stable transfection techniques but in limited amount due to the presence of the insulin degrading enzymes in the mammalian cells. The camel insulin production using solid phase peptide synthesis is hard due to the difficulties in purifying the camel insulin A-chain after synthesis.

Chapter 6

Investigating the thermal stability of camel insulin and insulin from other species in camel milk

6.1 Introduction:

Camel milk proteins have been reported to be remarkably stable Elagamy (2000). Previous studies suggested that camel milk insulin is more stable than insulin from other species (Wernery *et al.*, 2006a). From the work described in chapter 3, there are a few but potentially important changes in the amino acid sequence of camel insulin compared to human which may account for some of this enhanced stability as described in chapter 3. However, the stability of camel milk insulin could be due to the amino acid makeup of insulin alone or due to other factors in the milk that protect the insulin in the camel milk.

In this chapter, the stability of camel, bovine and human insulin in camel and bovine milk was studied under different thermal conditions to determine the relative thermal stability of the insulin from these species and to determine if camel milk itself offered any thermal protect to the insulin proteins.

6.2 Aims:

- To investigate the biophysical properties of camel milk insulin and the protective nature of camel milk on insulin upon temperature stress.

- To study the thermal stability of camel, bovine and human insulins in camel and bovine milk under different temperatures.

6.3 Results:

Insulin from bovine and human, along with camel insulin isolated from camel milk were spiked into camel and bovine milk and the thermal stability of the insulin molecules investigated. To do this the samples were divided into three groups; the first was used as a control, the second was heated at 60°C for one hour and the third group was incubated at 72°C for one hour. The plain milk (unspiked) samples were also treated under the same conditions to determine the stability of the milk insulin itself. The insulin concentration remaining was then measured by radioimmunoassay.

The results from this study (Tables 6.1 and 6.2) show that the reduction of camel and bovine insulin concentration (10% and 12% respectively) was much less than human insulin (31%) in camel milk after incubating the samples at 60°C for one hour. This suggests that both camel and bovine insulin are more stable than human insulin. There was no further reduction in the insulin concentration after incubating the human insulin (34%) with camel milk at 72°C for one hour in comparison with 60°C incubation. However the camel and bovine insulin (21% and 22% reduction, respectively) showed higher insulin concentrations being maintained than the human insulin.

Interestingly, when the camel insulin was in bovine milk it maintained higher insulin concentrations (10% reduction) after incubating the sample at 60°C for one hour and the percentage of reduction was similar to that with camel milk and less than the other two insulin's investigated. The bovine and human insulin showed a similar reduction percentage in bovine milk (32% and 31% respectively) after being at 60°C for one hour.

The camel, bovine and human insulin concentrations dropped drastically when they were incubated at 72°C for one hour in bovine milk. These data suggest that the camel insulin is the most stable of the insulins investigated and is the most thermal resistant as shown in the bovine milk suggesting that the amino acid differences between camel and the other insulins provide some of the additional stability. However, it is also apparent that the camel milk itself provides some protection or 'stability' to the different insulin molecules and hence the results here suggest that the enhanced thermal stability of camel milk insulin is due to both the amino acid makeup of the camel insulin molecule and the presence of unknown factors that provide stability in the camel milk.

The changes in insulin concentrations in the plain milk samples after the thermal treatment were close to the changes that occurred with the insulin spiked milk samples. Tables 6.1 and 6.2 summarise the results obtained of insulin concentrations before and after incubating the samples at 60°C and 72°C. Figure 6.1 illustrates the results obtained in a chart format.

Table 6.1 Insulin concentrations and the reduction of insulin before and after thermal treatment at 60°C for one hour in either camel or bovine milk

Milk	Insulin	Original Insulin Concentration (µIU/ml)	Insulin concentration after incubation at 60°C / 1h (µIU/ml)	Reduction percentage after incubation at 60°C / 1h (%)
Camel Milk	**Camel**	307	276	10
	Human	262	181	31
	Bovine	292	256	12
	Milk	60	56	7
Bovine Milk	**Camel**	260	235	10
	Human	227	157	31
	Bovine	256	173	32
	Milk	36	26	28

Table 6.2 Insulin concentrations and the reduction of insulin before and after thermal treatment at 72°C for one hour in either camel or bovine milk.

Milk	Insulin	Original Insulin Concentration (µIU/ml)	Insulin concentration after incubation at 72°C / 1h (µIU/ml)	Reduction percentage after incubation at 72°C / 1h (%)
Camel Milk	**Camel**	307	243	21
	Human	262	173	34
	Bovine	292	228	22
	Milk	60	48	20
Bovine Milk	**Camel**	260	10	96
	Human	227	6	97
	Bovine	256	8	97
	Milk	36	3	92

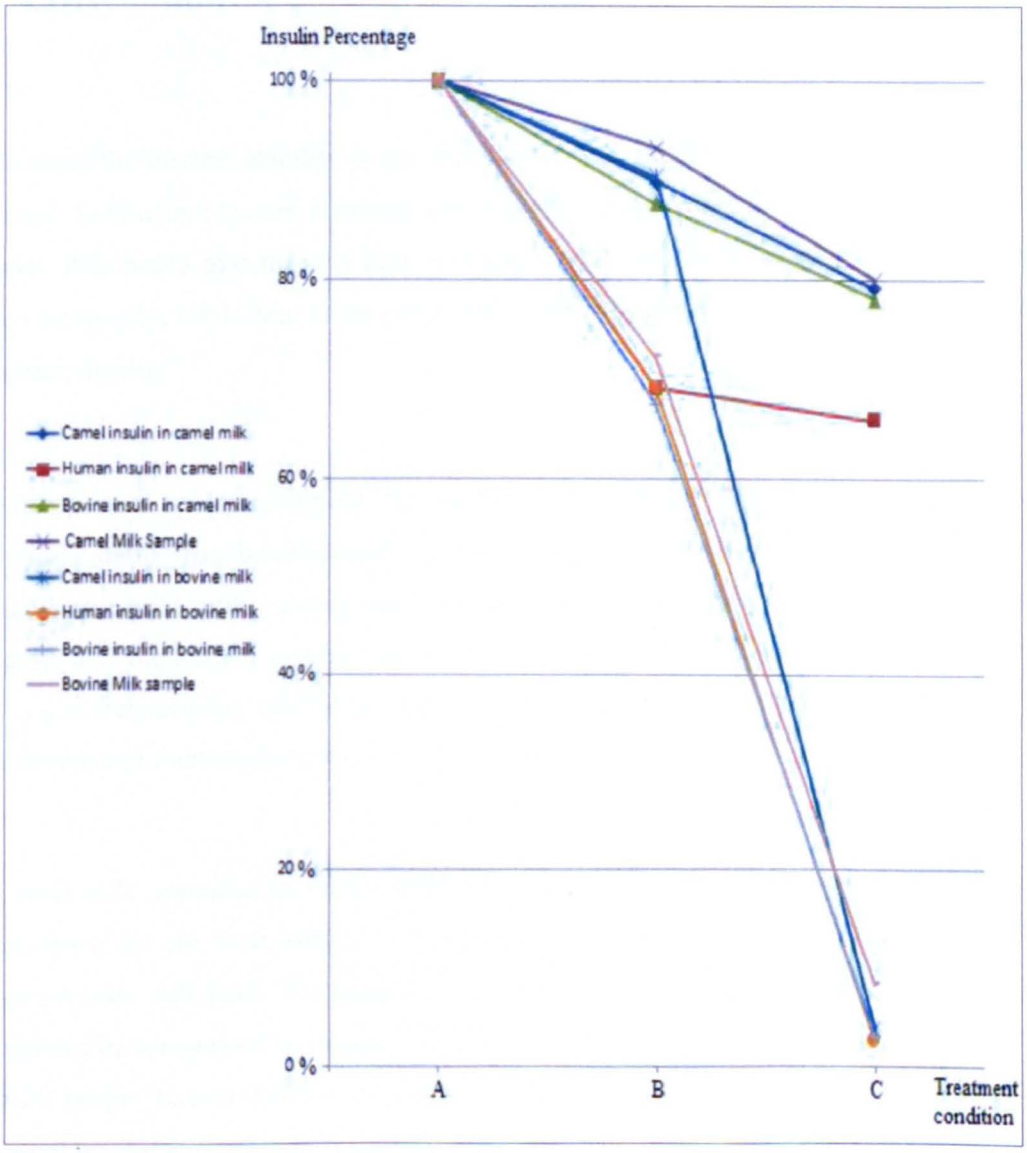


Figure 6.1 The change in insulin concentration after the thermal treatment of different insulins in either camel or bovine milk. A is original sample without treatment (denoted as 100%), B is the percentage of insulin remaining after incubating at 60°C for one hour, C is the insulin concentration remaining after incubating at 72°C for one hour. The large reduction in the insulin concentration after the 72°C incubation time point in bovine milk is very clear.

6.4 Discussion:

To study the thermal stability of camel insulin, camel and bovine milks were spiked with camel, bovine and human insulin's and then thermally treated at 60°C and 72°C for one hour. It is noted that there is only a single amino acid difference between camel and bovine insulin, while four amino acid differences are present between the camel and the human insulin.

After thermal treatment at 60°C for one hour, camel insulin appeared to be slightly more thermal stable than bovine insulin in bovine milk and all insulin's were more thermal stable in camel milk. Among the three insulin's, camel insulin was the most thermal stable and the human insulin was the less thermal stable. This reflects the amino acid changes between the species insulin molecules which seems to account for some of the difference in thermal stability observed, but not all.

Camel milk appeared to offer stability to the insulin molecules, even at 72°C thermal treatment for one hour, whilst the bovine milk lost any ability to stabilise insulin at this temperature and time. The camel insulin appears to be stabilised by both amino acid variants in comparison to human (mild effect) and by unknown properties of camel milk itself (major factor). Bovine insulin, where only a single amino acid difference with the camel insulin is present, was the next most thermal stable of the insulin molecules whilst human insulin was the least stable and is the most different at the amino acid level. The insulin stability results therefore suggest that the camel milk insulin thermal stability is largely a property of the camel milk but that the amino acid composition also adds to the observed thermal stability, although this has a minor effect compared to the effect of the camel milk itself.

Similar to the findings presented here, Elagamy (2000) studied the effect of heat treatment on camel, bovine and buffalo milk. In this research he used a radial immune diffusion assay to quantify the concentration of lysozyme, lactoferrin and immunoglobulin G (IgG) following thermal insult and showed that heating milk at 65°C for 30 min had no significant effect on lysozyme and lactoferrin concentration. However, the activity of IgGs was significantly reduced in the three milk types. The total activity of IgG in cow and buffalo milk was lost at 75°C for 30 min while the loss was only 68.7% of the total activity of camel IgG. The complete activity of lactoferrin was lost at 85°C after 30 min incubation in all types of milk, but at this temperature, the reduction in the activity of the lysozyme was 56% for camel and 74% and 81.7% for cow and buffalo milk, respectively (Elagamy, 2000). On the other hand Omidfar and colleagues (2007) investigated the thermal denaturation stability using ELISA and Circular Dichrosim (CD) spectra analysis to study the thermal stability of camel antibodies at different temperature to determine the reversibility and binding activity after heat denaturation. The VHH of the camel nanobodies and both IgG2–IgG3 displayed good binding ability even at 90°C. At 90°C the protein maintained up to 50% of its activity due to its moderately stable conformation and preservation of more than 50% of its secondary structure. Conesa and colleagues (2008) also found that the lactoferrin in camel was the most active lactoferrin against *E. coli* 0157:H7, while alpaca and human lactoferrins had minimum activity. The antimicrobial activity was calculated by measuring the minimum bactericidal concentrations (MBCs), this result was obtained after testing the lactoferrin of sheep, goat and elephant as well.

The results presented in this chapter, and other published results, therefore strongly suggest that camel milk proteins in general are more heat resistant than the proteins of human and other species like cow. The exact reason for this is not clear at this stage. In the case of insulin the amino acid makeup of the protein appears to play some part in this increased thermal stability, but the major role is played by unknown factors in the camel milk. In the future further studies to investigate the constituents of camel milk that might

be responsible for this stability or 'protection' would be interesting. Identification of these constituents could help in the engineering of more thermal stable proteins or allow for the use of these constituents to protect other proteins from thermal stress, including the formulation of biotherapeutic proteins and live vaccines for use in 3rd world countries and in hot climates.

6.5 Conclusion:

From the results obtained from this chapter it can be concluded that camel insulin seems to be stabilised due to the amino acid variants in comparison to human insulin furthermore this study also suggested that the components of camel milk itself gives greater thermostability in comparison with bovine milk. Thus, the greater thermal stability of camel milk insulin is due to both, the amino acid composition of the camel insulin protein and also the existence of unknown factors that provide higher stability in camel milk.

Chapter 7

Final Discussion

Diabetes is a condition associated with a patient's inability to control their glucose levels as a result of insulin deficiency. This is usually treated by regular application of recombinant insulin produced in *E.coli* cells. As described earlier in this thesis, a number of studies have now reported on the observation that diabetic patients who are regular drinkers of camel milk have a reduced dependency on insulin and that in some communities the incidence of diabetes is zero amongst regular camel milk drinkers (Agrawal *et al.*, 2007b). Compared to the milk from other species, camel milk has a high concentration of insulin and it has been suggested that this can even help activate B cells in camel milk drinkers in addition to providing insulin (Agrawal *et al.*, 2007a). There is currently an overwhelming and exponential increase in type II diabetics in the Middle East area and beyond which shows no sign of abating at present. As camel milk insulin has high thermostability and is a widely taken dietary supplement across the Middle Eastern population, the consumption of camel milk may help prevent, reduce and/or even treat diabetics.

The prevalence of diabetics in the Middle East is amongst the highest in the world and the rate is increasing rapidly, contributing to the concern of limiting the epidemic and the financial burden of health care of the countries in this area (FAO, 2010). The prevalence rate of the disease in most of these countries, especially in the gulf council countries (GCC), is such that they rank between the 2nd and 12th in the world in terms of the incidence within the population with the United Arab Emirates being the 2nd in the world in terms of diabetes prevalence within the population (Alhyas *et al.*, 2012). The treatments for this diseases costs these countries many millions of dollars and requires many medical staff. The disease also contributes to other serious and costly

complications, such as early stage eye disease, heart attack, coronary heart disease and non-traumatic lower limb amputation (Ringborg *et al.*, 2009). Therefore, it is considered to be extremely important that these countries, and the world overall, find new preventive methods and treatments/cures for this disease which do not just rely upon recombinant insulin treatment.

For many years, insulin has remained the best and most widely used treatment for diabetes. Many diabetic patients require multiple injections of insulin daily. Patients often resist taking insulin therapy in the early stages of their disease for a variety of reasons including a belief that the need for insulin reflects personal failure and concerns that insulin therapy will be complicated, hypoglycaemic consequences correlated with the injection, and the usage inconvenience. Many disadvantages of injectable insulin therapy could overcome, such as these social issues, by using alternative methods of delivering recombinant insulin, oral delivery being the most preferred method (Polonsky and Jackson 2004; Brunton *et al.*, 2006; Meece, 2006; Brod *et al.*, 2009). However, to date a reliable oral delivery system for insulin has not been devised.

Diabetes is a genetic and immunological disorder and it is unfeasible to change the genetics of a population; however it is feasible to change the lifestyle of the population as in many cases lifestyle issues, particularly diet, underpin the development of the disease. Thus, alongside a change of diet, encouraging the consumption of camel milk as a regular dietary supplement across the Middle East could be an important step in combating diabetes in the area. This thesis set out to investigate the properties of camel insulin in more detail and to try and understand why camel milk has higher insulin concentrations than other species and if this was potentially a property of the camel insulin molecule or a property of the milk itself. Camel and camel milk are available widely in the Middle East and so a further understanding of the insulin within this that contributes to a positive diabetic benefit will aid in understanding the mechanism behind this observation. It has been shown in this thesis that the insulin in camel milk is more

thermostable under harsher elevated temperature conditions compared to the insulin found in cow or human. To further investigate why this is the case, the genomic DNA sequence of the camel insulin gene was deduced and compared to other species. The sequencing results were confirmed after sequencing 32 camels.

Although Al-Swailem and colleagues sequenced camel proinsulin in 2008 the start region of the sequence was not of camel origin (Al-Swailem *et al.*, 2008). This was due to their choice of primers, which were driven from the human insulin sequence. Initially, using the primers from this study and subsequently by designing new primers from the BAC clone of *lama pacos* number CH246-515B17 flanking region of the insulin gene, the full sequence of camel insulin was obtained. The result of the new primers showed sequence differences with the published data (AL-Swailem *et al.*, 2008). The comparison of the sequences of the 32 sequenced camels showed five single nucleotide polymorphisms within the insulin gene, three of which are in the intronic region of the C-peptide and two which were found in the 3' untranslated region of the gene after the stop codon and before the termination sequence. These two SNP's found in the 3'-UTR may play a role in gene expression although this has not been investigated in this study.

The sequencing showed, as expected, that the camel insulin gene sequence is highly conserved between individuals. However, when comparing the insulin gene of camel with human and bovine counterparts, this revealed a number of non-sense differences at the DNA level of the A- and B-chains while in the signal and C-peptides there were changes distributed across the sequence. This suggests that the A- and B-chains, which constitute the final active insulin molecule are still highly conserved at the amino acid level and reflect insulin function and its importance, whilst the remainder of the gene and prepro-molecule is less well conserved, these section of the protein not being part of the final functional insulin protein molecule.

The predicted amino acid sequence from the camel insulin gene showed a number of unique differences at the amino acid level when compared with other species as shown in figure 3.8 One of the major differences between camel insulin and other species is the amino acid change of valine in all other species investigated to alanine in the 2nd amino acid of the B-chain of the camel. This change could well contribute to the increased thermostability for the camel insulin molecule shown in this thesis compared with other species. This hypothesis is supported by work from Dao-Pin *et al.* (1990) that showed the substitution of valine to alanine has implications for high thermostability in the T4 phage lysozyme protein (Dao-Pin *et al.*, 1990). In addition to this Lyu and co-workers (1990) reported the variation in stability of an alpha helical structure after performing a substitution of a single residue in a polypeptide sequence with different amino acids using a chemical peptide synthesis technique. The authors showed in this study a higher stability of alpha helical structure in alanine containing helicies than in valine containing helices. Thus, this unique amino acid change in camel insulin might result in a more stable insulin molecule than in other species and explain the higher thermostability of camel insulin when compared with the human and bovine insulins under different thermal conditions. This could be simply tested by mutating the alanine reside back to valine recombinantly, and then assessing the effect on thermostability of the resulting molecule.

In comparison with human insulin (B-chain and A-chain), camel insulin has 4 amino acid differences; furthermore three out of the four are changes to alanine residues. The first substation is in the 2nd residue of the B-chain as discussed above, and is a change from valine to alanine. The second and third differences are in the last amino acid of the B-chain and in 8th residue of the A-chain, both being threonine to alanine substitutions. According to Lyu *et al.* (1990), alanine residues are more alpha helically stable than valine and threonine. Camel insulin contains three additional alanine residues compared to the human insulin sequence. Thus, a molecule containing more alpha helical stable residues could increase the stability of camel insulin. The predicted 3-D structure of

camel proinsulin also predicts more alpha helical structure than beta sheet in comparison with the human proinsulin structure (Figure 5.10) which may further explain the greater stability of camel insulin compared to its human counterpart. Proinsulin is the precursor of insulin and plays an important role in controlling the structure of the mature insulin protein (Wahren *et al.*, 2000) and if the camel insulin has a more stable precursor molecule this may result is in more mature insulin with correct folding.

The protein isolation process used in this project was efficient in purifying and concentrating the insulin protein. Combining three different isolation methods facilitated the purification of insulin from a natural source rich with other proteins using large quantity of milk. Ethanol purification was initially used to remove large molecular weight proteins and casien, whilst ion exchange chromatography was used to bind and purify negatively charged proteins, insulin being one of these proteins. Finally, immunoaffinity chromatography using human anti insulin antibodies was effective in purifying the insulin protein form most of the remaining proteins carried through the previous steps. SDS-PAGE and radioimmuno assay confirmed that the isolated protein contained insulin whereby the RIA results showed insulin activity in a position matching the human insulin control and there was no significant activity observed in the area corresponding to the higher molecular weight bands. Despite this, other contaminant proteins were still clearly present as shown by cross-reaction on western blots and from mass spectrometry analysis of the purified material.

To further study camel insulin and to produce it in a larger quantity, the recombinant expression of camel insulin was investigated. To overcome the incorrect folding problems and lack of, or less human like post-translational modifications associated with bacterial and yeast production systems (O'Callaghan and James, 2008), the camel insulin gene was cloned into a mammalian expression vector and transiently and stably transfected into a Chinese hamster ovary (CHO) cell line. The transfected cell line did successfully produce camel insulin but at very low concentrations. The reason behind

the limited production of insulin by mammalian cells is most likely due to the production of insulin degrading enzyme in such cells as explained by Pak *et al.* (2002). Potentially, transfecting the camel insulin in a cell line derived from pancreatic cells or into a knockout cell line whereby the insulin degrading enzyme has been removed could result in enhanced production of insulin in the absence of the degrading enzyme.

This project confirmed the previously published data by Wernery *et al.* (2006a) that camel milk insulin is a highly thermostable molecule, Further, in this study it has been shown that camel milk components appear to have the greatest effect on the stability of camel milk insulin, contributing more to the relatively (relative to the thermostability of insulin from other species) high thermostability of the camel insulin molecule than the changes at the amino acid level. The thermostability study investigated camel, bovine and human insulin and their thermostability under different thermal conditions in bovine and camel milks. The results, after heating the spiked bovine and camel milks with different insulin samples at 60°C for one hour, confirmeds that among the three insulin's, camel insulin was the most thermal stable whilst the human insulin was the least thermal stable. This strongly suggests that amino acid changes between the insulin species accounts for some of the difference in thermal stability observed. Camel milk appeared to confer thermostability to all three of the insulin molecules, even after heating at 72°C for one hour, whereas the bovine milk lost any ability to stabilise insulin at this temperature and time. Generally, camel insulin appears to be stabilised by unknown properties of camel milk components and to a lesser degree by the amino acid variants in comparison to human and bovine insulin.

The thermostability studies also showed that bovine insulin, where only a single amino acid difference with the camel insulin is present, was the next most thermal stable of the insulin molecules whilst human insulin was the least stable and is the most different at the amino acid level. This again suggests that whilst camel milk insulin thermostability is a property of the camel milk itself, the amino acid composition also contributes to the

observed thermal stability, even though this has a lesser effect in comparison with the effect of the camel milk itself.

In the future, the camel milk components could be fractionated by different chromatographic, filtration and ultra-filtration techniques and then each fraction of milk mixed with insulin then heat stressed to find the milk fraction(s) conferring this stabilising effect. Each fraction could then be further investigated to identify the makeup of this fraction in order to try and identify the components responsible. These could be protein in nature or lipid based. The stability could be due to the interaction with other milk proteins or due to the minerals (e.g. trace metal ions) available in the milk.

Finding and identifying the component(s) which stabilise insulin in camel milk could help the insulin and generally recombinant protein industry in a variety of ways. First, such a component could be added to recombinant insulin vials to help in stabilising the insulin during shipping and transport. Second, as previously reported by Agrawal *et al.* (2007a, 2007b) where drinking camel milk helped in reducing the blood glucose level, insulin could be produced in a tablet form with the stabilising agent and given to diabetics orally. This would help millions of diabetic patients who suffer from insulin injections. Thirdly, understanding the mechanism(s) by which these components stabilise insulin may help in the development of stabilisation formulations and approaches for other recombinant protein molecules and/or vaccines for use globally. As such, there are still many exciting opportunities for the improvement of insulin therapy, diabetic treatment and development of stabilisation approaches for other therapeutics that might arise from the study of camel milk and camel insulin.

References:

Abdelgadir W. S., Ahmed T. K., & Dirar H. A. (1998) The traditional fermented milk products of the Sudan. *International Journal of Food Microbiology* (44): 1-13.

Adel-Patient K, Nutten S, Bernard H, Fritsché R, Ah-Leung S, Meziti N, Prioult G, Mercenier A, Wal JM (2012) Immunomodulatory potential of partially hydrolyzed β-lactoglobulin and large synthetic peptides *J Agric Food Chem.* 60(43):10858-10866.

Agrawal R.P., Beniwal R., Sharma S., Kochar D.K., Tuteja F.C., Ghorui S.K. and M.S. Sahani (2005) EFFECT OF RAW CAMEL MILK IN TYPE 1 DIABETIC PATIENTS: 1 YEAR RANDOMISED STUDY *Journal of Camel Practice and Research* 12(1) 27-35.

Agrawal R.P., G.Singh, K.C.Nayak, D.K.Kochar, R.C.Sharma, R.Beniwal, P.Rastogi and R.Gupta. (2004a) Prevalence of Diabetes in Camel Milk Consuming 'RAICA' Rural Community of North West Rajasthan. *Int. J. Diab. Dev. Countries* (24): 109-114.

Agrawal R.P., .Kochar D.K., Sahani M.S., Tuteja F.C., Ghouri S.K. (2004b) Hypoglycemic Activity of Camel Milk in Streptozotocin Induced Diabetic Rats. *Int. J. Diab. Dev. Countries* (24) 47-49.

Agrawal R.P., Rutba Dogra, Niranjana Mohta, Raksha Tiwari, Sushma Singhal, Surender Sultania (2009) Beneficial effect of camel milk in diabetic nephropathy. *Acta Biomed* (80):131-134.

Agrawal R.P., S.Budania, P.Sharma, Rajeev Gupta, D.K.Kochar, R.B.Panwar, M.S.Sahani. (2007b) Zero prevalence of diabetes in camel milk consuming Raica community of north-west Rajasthan, India *Diabetes Research and Clinical Practice* (76) 290-296.

Agrawal R.P., Sanjay Saran, Poornima Sharma, Gupta R.P., Kochar D.K. Sahani M.S. (2007 a) Effect of camel milk on residual b-cell function in recent onset type 1 diabetes. *Diabetes Research and Clinical Practice.* (77): 494-495.

Alhyas L, McKay A, Majeed A (2012) Prevalence of Type 2 Diabetes in the States of The Co-Operation Council for the Arab States of the Gulf: A Systematic Review *PLoS ONE* 7(8): e40948. doi:10.1371/journal.pone.0040948.

Al-Swailem A. M. , Al-Fageeh M. B. , Alyamani E. J., Shehata M. M. and Al-Shammari T. A. (2008) Characterization of recombinant Arabian camel (Camelus dromedarius) insulin . *African Journal of Biotechnology* 7 (18): 3389–3394.

American Diabetes Association (2009) Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care.* 32(1): S62-S67.

Baile C. A., Z. Glick and J. Mayer. 1969. Effects of secretin and cholecysto kinin pancreozymin on pancreatic juice and insulin secretion of goats. *J. Dairy Sci.*, 52: 513-517.

BALMUS G., TRIFONOV V.A., BILTUEVA L.S., O'BRIEN P.C.R., ALKALAEVA E.S., FU B., SKIDMORE J.A., ALLEN T., GRAPHODATDSKY A.S. and YANG F., and Ferguson-Smith M.A. (2007) Cross-species chromosome painting among camel, cattle pig and human: further insights into the putative Cetartiodactyla ancestral karyotype. *Chromosome Research* (15): 499-514.

Banting FG, Best CH, Collip JB, Campbell WR, Fletcher AA. (1922) Pancreatic Extracts in the Treatment of Diabetes Mellitus. *Can Med Assoc J.* 12(3):141-6.

Banting, F.G. and Best, C.H., The internal secretion of the pancreas (1922) *J. Lab. Clin. Med.* 7, 251-266

Baumgartel KL and Conley YP. (2012) The Utility of Breastmilk for Genetic or Genomic Studies: A Systematic Review *Breastfeed Med.* (Epub ahead of print).

Beg O.U., Von Bahr-Lindstrom H, Zaidi ZH, Jornvall H. (1986) A camel milk whey protein rich in half cystine. Primary structure, assessment of variations, internal repeat patterns, and relationships with neurophysin and other active polypeptides. *Eur J Biochem* (159):195-201.

Beg OU, von Bahr-Lindstrom H, Zaidi ZH, Jornval H (1984) A small camel-milk protein rich in cysteine/ half-cysteine. *Biosci Rep* (4): 1065-1070.

Belgi, Alessia; Akhter Hossain, Mohammed; W. Tregear, Geoffrey; D. Wade, John 2011 The Chemical Synthesis of Insulin: From the Past to the Present *Immunology, Endocrine & Metabolic Agents - Medicinal Chemistry* 11 (1) 40-47

Benkerroum, N., Mekkaoui, M., Bennani, N., & Kamal, H. (2004). Antimicrobial activity of camel's milk against pathogenic strains of Escherichia coli and Listeria monocytogene. *International Journal of Dairy Technology* (57): 39-43.

Bhatt AP, Bhende PM, Sin SH, Roy D, Dittmer DP (2010) Dual inhibition of PI3K and mTOR inhibits autocrine and paracrine proliferative loops in PI3K/Akt/mTOR-addicted lymphomas. *Blood* (115): 4455–4463.

Bianchi NO, Larramendy ML, Bianchi MS, Cortes L (1986) Karyological conservatism in South American camelids *Experientia* 42: 622-624.

Birnbaum DT, Kilcomons MA, DeFelippis MR, Beals JM. (1997) Assembly and dissociation of human insulin and LysB28ProB29-insulin hexamers: a comparison study. *Pharm Res.* 14(1):25–36.

Biswas K M; DeVido D R; Dorsey J G (2003) Evaluation of methods for measuring amino acid hydrophobicities and interactions *Journal of Chromatography A* (1000): 637–655.

Bliss M: The Discovery of Insulin. Chicago, University of Chicago Press, (1982), p 155.

Blum J.W. and Hammon H. (2000) Colostrum effects on the gastrointestinal tract, and on nutritional, endocrine and metabolic parameters in neonatal calves. *Livest. Prod. Sci.* (66): 151–159.

Bockamp, E., Maringer, M., Spangenberg, C, Fees, S., Fraser, S., Eshkind, L., Oesch, F., and Zabel, B. (2002) Of mice and models: Improved animal models for biomedical research. *Physiol. Genomics* (11) 115–132.

Brod M, Kongso JH, Lessard S, Christensen TL. (2009) Psychological insulin resistance: patient beliefs and implication for diabetes management *Qual Life Res.* (18):23-32.

Brunstedt, J., and Chan, S. J. (1982) Direct effect of glucose on the preproinsulin mRNA level in isolated pancreatic islets *Biochem. Biophys. Res. Commun.* 106, 1383–1389.

Brunton SA, Davis SN, Renda SM. (2006) Overcoming psychological barriers to insulin use in type 2 diabetes. *Clin Cornerstone.* 8 (suppl 2):519-526.

Bryant, N.J., Govers, R., James, D.E. (2002) Regulated transport of the glucose transporter GLUT4. *Nat. Rev. Mol. Cell. Biol.* (3): 267-277.

Bulliet RW (1975) The Camel and the Wheel. Harvard University Press, Cambridge, USA.

Bunch TD, Foote WC, Maciulis A (1985) Chromosome banding pattern homologies and NORs for the Bactrian camel, guanaco and llama. *J Hered* (76): 115-118.

Calhoun D.A., Lunoe M., Du Y., Staba S.L. and Christensen R.D (1999) Concentrations of granulocyte colony-stimulating factor in human milk after in vitro simulations of digestion. *Pediatr. Res.* (46): 767–771.

Cappechi, M.R. (1980) High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. *Cell* (22): 479–488.

Casterman, C., T. Atarhouch, S. Muyldermans, G. Robinson, C. Hamers, E. Bajyana, Songa, N. Bendahman, and R. Hamers (1993) Naturally occuring antibodies devoid of light chains. *Nature* (363): 446-448.

Chance, R.E., Ellis R.M, and W.W. Bromer (1968). Porcine proinsulin: characterization and amino acid sequence. Science (Wash. D. C.). (161): 165-167.

Chang L, Chiang SH, Saltiel AR (2004). Insulin signaling and the regulation of glucose transport. *Molecular Medicine*.10(7–12):65–71.

Chang SG, Kim DY, Choi KD, Shin JM, Shin HC (1998) Human insulin production from a novel mini-proinsulin which has high receptorbinding activity. *Biochem. J.* (329): 631-635.

Chang, B. S., Kendrick, B. S. and Carpenter, J. F. (1996) Surface-induced denaturation of proteins during freezing and its inhibition by surfactants. *J Pharm Sci* (85):1325-1330.

Charman, S. A., Mason, K. L. and Charman, W. N. (1993) Techniques for assessing the effects of pharmaceutical excipients on the aggregation of porcine growth hormone. *Pharm Res* (10):954-962.

Chen JQ, Zhang HT , Hu MH and Tang JG (1995) Production of human insulin in an E.coli system with Met-Lys Human Proinsulin as the Expressed Precursor. Applied Biochemistry and Biotechnology (55): 5-15.

Cheung J.K., and Raverkar P.S. (2006) Analytical model for studying how environmental factors influence protein conformational stability in solution. *THE JOURNAL OF CHEMICAL PHYSICS* (125) 224903.

Clery, C., Renault, F. and Masson, P. (1995) Pressure-induced molten globule state of cholinesterase. *FEBS Lett.* (370):212-214.

Conesa C, Sánchez L, Rota C, Pérez MD, Calvo M, Farnaud S, Evans RW. (2008) Isolation of lactoferrin from milk of different species: calorimetric and antimicrobial studies. *Comp Biochem Physiol B Biochem Mol Biol.* 150(1):131-9.

Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Emdin, S. F., and Reynolds, C. D. (1979) Structure and biological activity of hagfish insulin. *J. Mol. Biol.* (132): 85-100.

D'Agostino J, Younes MA, White JW, Besch PK, Field JB, Frazier ML (1987) Cloning and nucleotide sequence analysis of complementary deoxyribonucleic acid for bovine preproinsulin *Mol Endocrinol.* 1(4):327-331

Dai, Y., and Tang, J. G. (1996) Characteristic, activity and conformational studies of [A6-Ser, A11-Ser]-insulin *Biochim. Biophys. Acta* (1296): 63–68.

Dandekar, A. M., Robinson, E. A,, Appella, E. & Qasba, P. K. (1982) Complete sequence analysis of cDNA clones encoding rat whey phosphoprotein: Homology to a protease inhibitor. *Proc. Natl Acad. Sci. USA* (79): 3987 - 3991.

Danne, T. and Bolinder, J. (2011) New insulins and insulin therapy. *Int J Clin Pract Suppl.* (170): 26-30

Dao-Pin S, Baase WA, Matthews BW. (1990) A mutant T4 lysozyme (Val 131→Ala) designed to increase thermostability by the reduction of strain within an alpha-helix. *Proteins* (7):198–204.

Darrington R. T. and Anderson B. D. (1995) Effects of insulin concentration and self-association of its A-21 cyclic anhydride intermediate to desamido insulin and covalent dimer. *Pharm. Res.* (12):1077–1084.

Darrington, R.T. and Anderson, B.D. (1994) The Role of Intramolecular Nucleophilic Catalysis and the Effects of Self-Association on the Deamidation of Human Insulin at Low pH *Pharm. Res.* (11) 784-793.

Dave N, Hazra P, Khedkar A, Manjunath HS, Iyer H, Suryanarayanan S. (2008) Process and purification for manufacture of a modified insulin intended for oral delivery J Chromatogr A. 1177(2):282-286.

Davies, P. O., Poirier, C., Deltour, L., and Montagutelli, X. (1994). Genetic reassignment of the insulin-1 (Ins1) gene to distal mouse chromosome 19. *Genomics* (21) 665-667.

de Haen, C , Swanson, E., and Teller, D. C (1976) The evolutionary origin of proinsulin. Amino acid sequence homology with the trypsin-related serine proteases detected and evaluated by new statistical methods. *J. Mol. Biol.* (106): 639-661

DeMeyts, P., Gu, J. L., Shymko, R. M., Kaplan, B. E., Bell, G. I., and Whittaker, J. (1990) Identification of a ligand-binding region of the human insulin receptor encoded by the second exon of the gene *Mol. Endocrinol.* 4, 409–416.

Desbandhu J., Parthiv C., Kundan P., Mithilesh P., Sanjai T. (2011) A Review on Diabetes *International Research Journal of Pharmacy* 2 (8): 22-25.

Di Berardino D, Nicodemo D, Coppola G King AW, Ramunno L, Cosenza GF, Iannuzzi L, Di Meo GP, Balmus G, Rubes J (2006) Cytogenetic characterization of alpaca (Lama pacos, fam. Camelidae) prometaphase chromosomes *Cytogenet Genome Res* (115): 138-144.

Diamond J,. (2003) The double puzzle of diabetes, Nature (423): 599–602.

Dickey, R. P., Thornton, M., Nichols, J., Marshall, D. C., Fein, S. H., Nardi, R. V., and Bravelle IVF, S. G. (2002) Comparison of the efficacy and safety of a highly purified human follicle-stimulating hormone (Bravelle (TM)) and recombinant follitropin-beta for in vitro fertilization: a prospective, randomized study, *Fertil. Steril.* 77(6): 1202-1208.

Dobson C.M. (2001) The structural basis of protein folding and its links with human diseases. *Philos. Trans. R. Soc. London, Ser. B* (356):133-145.

Doisy EA, Somogyi M, Shaffer PA (1923) Some properties of an active constituent of pancreas (insulin) [Proceedings]. J Biol Chem (55)

Duckworth, W.C. (1988). Insulin degradation: mechanisms, products, and significance. *Endocrine Rev.* (9), 319–345.

Dugani, C.B., and Klip, A. (2005) Glucose transporter 4: cycling, compartments and controversies. *EMBO Rep.* (6): 1137-1142.

Duvillie, B., Bucchini, D., Tang, T., Jami, J., and Paldi, A. (1998) Imprinting at the mouse Ins2 locus: evidence for cis- and trans-allelic interactions. *Genomics* (47): 52-57.

Eddleman H. (1999) Study of the lactoperoxidase system and its functions. President, Indian Biolab,14045 Huff St., Palmyra, pp:112-115.

El Hatmi H. , Giradet J.M. , Gaillard J.L. , Yahyaoui MH, Attia H (2007) Characterization of whey proteins of camel (Camelus dromedarius) milk and colostrums. *Small Rum. Res.* (70): 267-271.

Elagamy, E. I. (2000) Effect of heat treatment on camel milk proteins with respect to antimicrobial factors: a comparison with cow's and buffalo milk proteins. *Food Chemistry* (68): 227–232.

Elagamy, E. I., Abou-Shloue, Zeinab, I., & Abdel-Kader, Y. I. (1998) Gel electrophoresis of proteins, physicochemical characterization and vitamin C content of milk of different species. *Alex. J. Agric.Res* 43(2): 57-70.

El-Agamy, E. I., Ruppanner, R., Ismail, A., Champagne, C. P., Assaf, R. (1992) Antibacterial and antiviral activity of camel milk protective proteins. *Journal of Dairy Research* (59) 169-175.

El-Agamy, E.I. (1983). Studies on camel's milk. M.Sc. Thesis. Alexandria University, Egypt, pp. 40–43.

Elayan, A. A., Sulieman, A. E., & Saleh, F. A. (2008). The hypocholesterolemic effect of Gariss and Gariss containing Bifidobacteria in rats fed on a cholesterol-enriched diet. Asian Journal of Biochemistry (3): 43-47.

El-Hatmi H., J.M. Girardet, J.L. Gaillard, M. H. Yahyaoui and Attia H. (2007) Characterisation of whey proteins of camel (Camelus dromedarius) milk and colostrums. *Small Rumi. Res* (2): 267-271.

Emmett P. M. and Rogers I. S. (1997) Properties of human milk and their relationship with maternal nutrition. *Early human development* (49): S7-S28.

Eriksson, J.W. (2007) Metabolic stress in insulin's target cells leads to ROS accumulation a hypothetical common pathway causing insulin resistance. *FEBS Lett.* (581): 3734-3743.

Falkmer, S., and Ostberg, Y. (1976) Comparative morphology of pancreatic islets in animals. *In* The Diabetic Pancreas. Volk, B. W., and Wellman, K. F., Eds. New York, Plenum Publ. Co.

FAO (1996). Statistics year-book, FAO, Rome.

FAO (2010) Burden: mortality, morbidity and risk factors. Global status report on noncommunicable diseases 2010 (1):9-31.

Farah, Z. (2004): An introduction to the camel. In: FARAH, Z. & A. FISCHER (eds): Milk and meat from the camel - Handbook on products and processing. vdf Hochschulverlag, Zürich, Switzerland, 15 – 28 pp.

Feldman R., Kohout F. W., Grodsky, G. M., and Mcwilliams, N. B. (1963) Immunologic studies in a diabetic subject resistant to bovine insulin but sensitive to porcine insulin, *Am. J. Med.* (35): 411-417.

Forsyth, I. A., C. R. Strong, and R. Dils (1972) Interactions of insulin, corticosterone and prolactin in promoting milk-fat synthesis by mammary explants from pregnant rabbits. *Biochem. J.* (129): 929-935.

Fowler M.J. (2008) Microvascular and Macrovascular Complications of Diabetes. *Clinical Diabetes.* 26 (2): 77-82.

FOWLER, M. E. (1998) Medicine and surgery of South American camelids. Iowa State University Press, Ames, Iowa, USA.

FOWLER, M. E. (1998) Medicine and surgery of South American camelids. Iowa State University Press, Ames, Iowa, USA.

Frokjaer S and Hovgaard L. (2000) Pharmaceutical Formulation Development of Peptides and Proteins, CRC Press, United States America.

Furtado, L.M., Somwar, R., Sweeney, G., Niu, W., Klip, A. (2002) Activation of the glucose transporter GLUT4 by insulin. *Biochem. Cell. Biol.* (80): 569-578.

Galloway, J. A., Hooper, S. A., Spradlin, C. T., Howey, D. C., Frank, B. H.,Bowsher, R. R., and Anderson, J. H. (1992) Biosynthetic human proinsulin: review of chemistry, in

vitro and in vivo receptor binding, animal and human pharmacology studies, and clinical trial experience. *Diabetes Care* (15): 666–692.

Glenn, K. C, and Cunningham, D. D. (1979) Thrombin-stimulated cell division involves proteolysis of its cell surface receptor. *Nature* (278): 711-714.

Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K.,. and Riggs, A. D. (1979) Expression in Escherichia coil of chemically synthesized genes for human insulin. *Proc. Nail. Aco.d. Sd. USA* (76):106-110.

Grant, P. T., and T. L. Coombs. (1970) Proinsulin, a biosynthetic precursor of insulin. In Essays in Biochemistry. P. N. Campbell and G. D. Greville, editors. Academic Press, Inc., Ltd. London. (6): 69-92.

Granvogl, B., Gruber, P., Eichacker, L.A. (2007) Standardization of rapid in-gel digestion by mass spectrometry, Proteomics, 7(5): 642-654

Graphodatsky A.S. (2006) Camelus bactrianus. In O'Brien SJ, Menninger JC, Nash WG, eds. Atlas of Mammalian Chromosomes. New York: Wiley-Liss p. 547.

Groskreuiz DJ, Sliwkowski MX and Gorman CM (1994) Genetically engineered proinsulin constitutively processed and secreted as mature, active insulin. *J Biol Chem* (269):6241–6245

Grosvenor A.J., Morton J.D.D. (2010) Profiling of residue-level photooxidative damage in peptides. *Amino Acids* (39): 285-296.

Guo, Z. Y., and Feng, Y. M. (2001) Effects of cysteine to serine substitutions in the two inter-chain disulfide bonds of insulin. *Biol. Chem.* (382): 443–448

Guyton, A. C. and J. E. Hall. (2006) Textbook of Medical Physiology. 11th ed. Elsevier Saunders, Philadelphia, PA. p.p 870-871

Hacein-Bey-Abina S, Le Deist F, Carlier F, Bouneaud C, Hue C, De Villartay JP, Thrasher AJ, Wulffraat N, Sorensen R, Dupuis-Girod S, Fischer A, Davies EG, Kuis W, Leiva L, Cavazzana-Calvo M. (2002) Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N Engl J Med* 346(16):1185–1193

Haddadin M.S., Gammoh S.I. and Robinson R.K. (2007) Seasonal variations in the chemical composition of camel milk. *Jordan. J Dairy Res* (31):1-5.

Hagiwara K., Kataoka S., Yamanaka H., Kirisawa R. and Iwai H. (2000) Detection of cytokines in bovine colostrum, Vet. Immunol. Immunopathol. (76): 183–190.

Hajduch, E., Hainault, I., Meunier, C., Jardel, C.; Hainque, B., Guerre-Millo, M., Lavau, M. (1995) Regulation of glucose transporters in cultured rat adipocytes: synergistic effect of insulin and dexamethasone on GLUT4 gene expression through promoter activation. *Endocrinology* (136): 4782-4789.

Hambraeus L. (1992) Nutritional Aspects of Milk Proteins. In: Fox PF, editor. Advanced Dairy Chemistry–1: Proteins. London: Elsevier Applied Science; pp. 457–490.

Harashima, S. (1994) Heterologous protein production by yeast host-vector systems. *Bioprocess Technol.* (19):137-158.

Harding H. P. and Ron D., (2002) Endoplasmic reticulum stress and the development of diabetes: a review *Diabetes.* 51 (3): 5455–5461.

Harper J. D. and Lansbury P. T. (1997) Models of amyloid seeding in Alzheimer's disease and scrapie. *Annu. Rev. Biochem.* (66):385-407.

Harper, M. E., Ullrich, A., and Saunders, G. F. (1981). Localization of the human insulin gene to the distal end of the short arm of chromosome 11. *Proc Natl Acad Sci U S A* (78): 4458-4460.

Hassan Y. M. (1971) Note of The Effect of Dehydration on a Camel. *Sudan Journal of Veterinary Science and Animal Husbandry* (12): 111–112.

Hatsuzawa K, Tagaya M, Mizushima S. (1997) The hydrophobic region of signal peptides is a determinant for SRP recognition and protein translocation across the ER membrane *J Biochem.* 121(2):270-277.

Hay and Docherty (2006) Comparative Analysis of Insulin Gene Promoters Implications for Diabetes Research *Diabetes* (55): 3201-3213

Hcrtelendy, F., K. Takahashi, L. J. Machlin, and D. M. Kipnis. (1970) Growth hormone and insulin secretory response to arginine in the sheep, pig and cow. *Gen. Comp. Endocrinol.* (14):72-77.

Hennighausen, L. G., Sippel, A. E., Hobb, A. A. & Rosen, J. M. (1982) Comparative sequence analysis of the mRNAs coding for mouse and rat whey protein. *Nucleic Acids Res.* (10): 3733-3744.

Henson, A. F., Mitchell, J. R. and Musselwhite, P. R. (1970) The surface coagulation of proteins during shaking. *J Colloid Interface Sci* (32):162-165.

Herman, M.A., Kahn, B.B. (2006) Glucose transport and sensing in the maintenance of glucose homeostasis and metabolic harmony. *J. Clin. Invest.* (116): 1767-1775.

Hernandez, R., Teruel, T., Lorenzo, M. (2003) Insulin and dexamethasone induce GLUT4 gene expression in foetal brown adipocytes: synergistic effect through CCAAT/enhancer-binding protein alpha. *Biochem. J.* (372): 617-24.

Hess C, Thomas A, Thevis M, Stratmann B, Quester W, Tschoepe D, Madea B, Musshoff F. Simultaneous determination and validated quantification of human insulin and its synthetic analogues in human blood serum by immunoaffinity purification and liquid chromatography-mass spectrometry *Anal Bioanal Chem.* 2012 (6-7):1813-1822

Higgins N. (1984) The Camel: Health and Diseases. Baillier Tinall London.

Hober S., Uhlen M, Nilsson B. (1997). Disulfide Exchange Folding of Disulfide Mutants of Insulin-like Growth Factor I *in vitro. Biochemistry* (36): 4616-4622.

Hoitink M.A, Beijnen J.H., Boschma M.U.S., Bult A., Hop E., Nijholt J., Versluis C., Wiese G, Underberg,W.J.M., (1997) Identification of the degradation products of gonadorelin and three analogues in aqueous solution. *Anal. Chem.* (69): 4972-4978.

Holley, R. W. (1975) Factors controlling the growth of 3T3 cells and transformed 3T3 cells. *In* Proteases and Biological Control. Reich, E., Rifkin, D. B., and Shaw, E., Eds. Coldspring Harbor Laboratory, pp. 777-784.

Holmen SL, Vanbrocklin MW, Eversole RR, Stapleton SR, Ginsberg LC (1995) Efficient lipid-mediated transfection of DNA into primary rat hepatocytes. *In Vitro Cell Dev Biol Anim* 31(5):347–351.

Holz C. and Lang C. (2005) Expression of cDNA in Yeast. *Cell biology* (2005), 3rd edition (1): 57-62.

Horino, M., L. J. Maehlin, F. Hertelendy, and D. M. Kipnis. (1968) Effect of short chain fatty acids on plasma insulin in ruminant and nonruminant species. *Endocrinology* (83):118-128.

Hossain, P., B. Kawar, and M. El Nahas, (2007) Obesity and diabetes in the developing world--a growing challenge. *N Engl J Med* 356(3): 213-215.

Hua, Q. X., Gozani, S. N., Chance, R. E., Hoffmann, J. A., Frank, B. H., and Weiss, M. A. (1995) Structure of a protein in a kinetic trap *Nat. Struct. Biol.* (2): 129–138

Hua, Q. X., Mayer, J. P., Jia, W., Zhang, J., and Weiss, M. A. (2006) The folding nucleus of the insulin superfamily: a flexible peptide model foreshadows the native state *J. Biol.Chem.* (281): 28131–28142.

Hua, Q.X., W. Jia, B.H. Frank, N.F. Phillips, and Weiss M.A. (2002) A protein caught in a kinetic trap: structures and stabilities of insulin disulfide isomers. *Biochemistry.* (41):14700–14715.

Huang X. F. and Arvan P. (1994) Formation of the insulin containing secretory granule core occurs within immature β- granules *Journal of Biological Chemistry* 269 (33): 20838–20844.

Hummer, G., Garde, S., Garcia, A. E., Paulaitis, M. E. and Pratt, L. R. 1998. The pressure dependence of hydrophobic interactions is consistent with the observed pressure denaturation of proteins. *Proc Natl Acad Sci U S A* (95):1552-1555.

Hunt I. (2005) From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Protein Expression and Purification* (40): 1–22.

Irvin, B., and A. Trenkle. (1971): Influences of age, breed and sex on plasma hormones in cattle. *J. Animal Sci.* (32): 292-295.

Iype, T., Francis, J., Garmey, J. C., Schisler, J. C., Nesher, R., Weir, G. C., Becker, T. C., Newgard, C. B., Griffen, S. C., and Mirmira, R. G. (2005). Mechanism of insulin gene regulation by the pancreatic transcription factor Pdx-1: application of pre-mRNA analysis and chromatin immunoprecipitation to assess formation of functional transcriptional complexes. *J Biol Chem* (280) 16798-16807.

Jameel, F., Bogner, R., Mauri, F. and Kalonia, D. (1997) Investigation of physicochemical changes to Lasparaginase during freeze-thaw cycling. *J Pharm Pharmacol* (49):472-477.

Jassim S. A. A. and Naji M.A. (2001) Camel immune system and activity of milk. *Biologist* (48):268–272.

JASSIM, S. A. A. & M. A. NAJI (2002) The desert ship: heritage and science - unique camel antibodies. *Middle East Lab.* (5): 6 – 11.

Jeffrey PD, Coates JH. (1966) An equilibrium ultracentrifuge study of the self-association of bovine insulin. *Biochemistry.* 5(2):489–498.

Joshi SR, Parikh RM, Das AK. (2007) Insulin – history, biochemistry, physiology and pharmacology. J Assoc Physicians India (55)(Suppl):19–25

K. Akaji, K. Fujino, T. Tatsumi, Y. Kiso; (1993)Total Synthesis of Human Insulin by Regioselective Disulfide Formation Using the Silyl Chloride-Sulfoxide Method J. Am. Chem. Soc. (115): 11384-11392

K. Akaji, Y. Kiso, in: Houben-Weyl, (2002)Methods of Organic Chemistry, *Synthesis of Peptides and Peptidomimetics Vol. E22b*, M. Goodman, A. Felix, L. Moroder, C. Toniolo, eds., Thieme, Stuttgart, , pp 101-141

Kanzaki M, Mora S, Hwang JB, Saltiel AR, Pessin JE. (2004) Atypical protein kinase C (PKCzeta/lambda) is a convergent downstream target of the insulin-stimulated phosphatidylinositol 3-kinase and TC10 signalling pathways. *J. Cell Biol.* (164): 279-90.

Katsoyannis PG, Tometsko A, Zalut C (1966) Insulin peptides XII. Human insulin generation by combination of synthetic A and B chains. J Am Chem Soc (88):186.

Kelly G.S. (2003) Bovine colostrums: a review of clinical uses. *Altern. Med. Rev.* (8): 378–394.

Kerwin B.A. and Remmele R.L. (2007) Protect from light: photodegradation and protein biologics. *J. Pharm. Sci.* (96): 1468- 1479.

Kessler, E. & Brew, K. (1970) The whey proteins of pig's milk: isolation and characterization of a. beta lactoglobulins in human colostrum and milk.. *Biochim. Biophys. Acta* (200): 449-458.

Khow O. and Suntrarachun S., (2012) Strategies for production of active eukaryotic proteins in bacterial expression system. *Asian Pacific Journal of Tropical Biomedicine* (2012): 159-162.

Kim, J.K., Zisman, A., Fillmore, J.J., Peroni, O.D., Kotani, K., Perret, P., Zong, H., Dong, J., Kahn, C.R., Kahn, B.B., Shulman, G.I. (2001) Glucose toxicity and the development of diabetes in mice with muscle-specific inactivation of GLUT4. *J. Clin. Invest.* (108): 153-160.

Kim, T. K.; Eberwine, J.H. (2010) Mammalian cell transfection: the present and the future *Anal Bioanal Chem* (397): 3173–3178

King H., Aubert, R.E.. Herman W.H, (1998) Global burden of diabetes, 1995–2025. *Diabetes Care* (21) 1414–1431.

Knoess K. H. (1977) The Camel as a meat and milk animal. *World Animal Review* (22): 39–44.

Knoess K.H (1979) Milk production of the dromedary. *In*: Camels. IFS Symposium, Sudan, 201-214.

Ko, A. S. C., Smyth D. G., Markussen J., and F. Sundby (1971) The amino acid sequence of the C-peptide of human proinsulin. *Eur. J. Biochem.* (20): 190-199.

Koulischer L, Tijskens J, Mortelmans J (1971) Mammalian cytogenetics. IV. The chromosomes of two male Camelidae: Camelus bactrianus and Lama vicugna. *Acta Zool Pathol Antverp* (52): 89-92.

Kreilgaard, L., Jones, L. S., Randolph, T. W., Frokjaer, S., Flink, J. M., Manning, M. C. and Carpenter, J. F. (1998) Effect of Tween 20 on freeze-thawing- and agitation-induced aggregation of recombinant human factor XIII. *J Pharm Sci* (87):1597-603.

Kreilgaard, L., Jones, L. S., Randolph, T. W., Frokjaer, S., Flink, J. M., Manning, M. C. and Carpenter, J. F. (1998) Effect of Tween 20 on freeze-thawing- and agitation-induced aggregation of recombinant human factor XIII. *J Pharm Sci* (87):1597-1603.

Kulski J. K. and Hartmann P. E. (1983) Milk insulin, GH and TSH: relationship to changes in milk lactose, glucose and protein during lactogenesis in women. *Endocrinol Exp.* 17(3-4):317-326.

Kumar, K.M.P., Bhat, G.K. (2003) Animal insulins - current status. *Int J. Diab. Dev. Countries* (23): 6-9

KUNAPARAJU R.; LIAO M ; SUNSTROM N.A. (2005) Epi-CHO, an episomal expression system for recombinant protein production in CHO cells *BIOTECHNOLOGY AND BIOENGINEERING* 91, 6, 670 – 677.

Kung YT, Da YC, Huang WT, Chen CC, Ke LT (1966): Total synthesis of crystalline insulin. *Sci Sin* (15):544.

Kuuranne T, Thomas A, Leinonen A, Delahaut P, Bosseloir A, Schänzer W, Thevis M (2008) Insulins in equine urine: qualitative analysis by immunoaffinity purification and liquid chromatography/tandem mass spectrometry for doping control purposes in horse-racing. *Rapid communications in mass spectrometry* 22 (3): 355-362

L. Moroder, H.-J. Musiol, N. Schaschke, L. Chen, B. Hargittai, G., Barany, in: Houben-Weyl, (2003) Methods of Organic Chemistry, *Synthesis of Peptides and Peptidomimetics, Vol. E22a*, M. Goodman, A. Felix, L. Moroder, C. Toniolo, ed., Thieme, Stuttgart, pp. 384-423

Laleye, L. C., Jobe, B. & Wasesa, A. A. H. (2008). Comparative Study on Heat Stability and Functionality of Camel and Cow's Milk Whey Proteins. Journal of Dairy Science, 91, 4527-4534.

Lasker S.P.; McLachlan C.S.; Wang, L.; Ali S.M.K.; Jelinek, H.F. (2010) Discovery, treatment and management of diabetes *Journal of Diabetology* 1(1): 1-8.

Leibiger, B., Moede, T., Schwarz, T., Brown, G. R., Kohler, M., Leibiger, I. B., and Berggren, P. O. (1998) Short-term regulation of insulin gene transcription by glucose *Proc. Natl. Acad. Sci. U. S. A.* 4, 9307–9312.

Leibiger, B., Wahlander, K., Berggren, P.O., and Leibiger, I. B. (2000) Glucose-stimulated insulin biosynthesis depends on insulin-stimulated insulin gene transcription *J. Biol. Chem.* 275, 30153–30156

Li S, Schoneich C, Borchardt R.T. (1995) Chemical instability of protein pharmaceuticals: Mechanisms of oxidation and strategies for stabilization. *Biotechnol. Bioeng.* (48): 490-500.

Liberatori, J., Morisio Guidetti, L., Conti, *A.* & Napolitano, L. (1979) Immunological evidence of Betalactoglobulin in human colostrum and milk. *Boll. SOC. Ital. Biol. Sper.* (55): 1369-1373.

Luo D. and Saltzman W. Synthetic DNA delivery systems. (2000) Nature Biotechnology 18(1) 33–37.

Lyu.P.C, Liff.M.I., Marky.L.A and Kallenbach.N.R.(1990) Side Chain Contributions to the Stability of Alpha-Helical Structure in Peptides *Science.* 250, 669-673.

MacRitchie, F. (1986). Spread monolayers of proteins. *Adv Colloid Interface Sci* (25):341-385.

Mal G., Sena D.S., Jain V.K. and Sahani M.S. (2001) Therapeutic utility of camel milk as nutritional supplement in chronic pulmonary tuberculosis. *Livestock International.* 4-8.

Manning M C, Chou D K, Murphy B M, Payne R W, Katayama D S. (2010) Stability of protein pharmaceuticals: an update. *Pharm. Res.* (27) 544-575.

Markussen J, Damgaard U, Diers I, Fiil N, Hansen MT, Lassen P, Norris F, Norris K, Schou O, Snel L, Thim L, Voigt HO (1986). Biosynthesis of human insulin in yeast via single chain precursors. *Diabetologia* (29): 568-569.

Marshak, S., Benshushan, E., Shoshkes, M., Leibovitz, G., Kaiser, N., Gross, D., Bertuzzi, F., Cerasi, E., and Melloul, D. (2001). beta-cell-specific expression of insulin and PDX-1 genes. *Diabetes* 50 Suppl 1, 5131-5132.

Martin F, Volpari C., Steinkuhler C.and Dimas N. (1997) Affinity selection of a camelized V (H) domain antibody inhibitor of hepatitis C virus NS3 protease. *Protein Eng* (10):607–14.

MeAtee, J. W., and A. Trenkle. (1971) Metabolic regulation of plasma insulin levels in cattle. *J. Animal Sci.* (33):438-442.

Meece J. (2006) Dispelling myths and removing barriers about insulin in type 2 diabetes. *Diabetes Educ.*(32):95-185.

Meienhofer J, Schnabel E, Bremer H, Brinkhoff O, Zabel R, Sroka W, Klostermeyer H, Brandenburg D, Okuda T, Zahn H (1963) Synthese der Insulinketten und ihre Kombination zu insulactiven Praparaten. *Z Naturforsch* (18):1120-1121.

MELLOTT, A. J., FORREST, M. L., DETAMORE, M. S. (2012) Physical Non-Viral Gene Delivery Methods for Tissue Engineering. *Annals of Biomedical Engineering* 10.1007/s10439-012-0678-1

Melloul, D., Marshak, S., and Cerasi, E. (2002) Regulation of insulin gene transcription. *Diabetologia* (45) 309-326.

Mills E. S. and Topper J. Y. (1970) Some Ultra structural effect of insulin, hydrocortisone and prolactin on mammary gland explants . *J. Cell Biol.* (44): 310-328.

Molento, C.F.M., E. Block, R.I. Cue, and D. Petitclerc (2002) Effects of insulin, recombinant bovine somatotropin, and their interaction on insulin-like growth factor-I secretion and milk protein production in dairy cows. *J. Dairy Science* (85):738-747.

Mudaliar S, Edelman SV (2001) Insulin therapy in type 2 diabetes. *Endocrinol Metab Clin.* 30 (4): 935-982.

Mulkerrin, M. G. and Wetzel, R. (1989) pH dependence of the reversible and irreversible thermal denaturation of gamma interferons. *Biochemistry* (28):6556-6561.

Muller SA, Kohajda T, Findeiss S, Stadler PF, Washietl S, Kellis M, von Bergen M, Kalkhof S. Optimization of parameters for coverage of low molecular weight proteins. Anal Bioanal Chem. 2010 Dec;398(7-8):2867-81. Epub 2010 Aug 28. PubMed PMID: 20803007; PubMed Central PMCID: PMC2990009.

Muller, K., Ludemann, H. D. and Jaenicke, R. (1981) Pressure-induced structural changes of pig heart lactic dehydrogenase. *Biophys Chem* (14):101-110.

Muneta Y., Yoshihara K., Minagawa Y., Nagata R., Yasuyuki M., Yamaguchi T., Takehara K. (2005) Bovine IL-18 ELISA: detection of IL-18 in sera of pregnant cow and newborn calf, and in colostrums. *J. Immunoassay Immunochem.* (26) 203–213.

Muyldermans, S., Cambillau, C., and Wyns, L. (2001) Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. *Trends Biochem. Sci.* (26): 230-235.

Mynarcik, D. C., Williams, P. F., Schaffer, L., Yu, G. Q., and Whittaker, J. (1997) Identification of Common Ligand Binding Determinants of the Insulin and Insulin-like Growth Factor 1 Receptors. INSIGHTS INTO MECHANISMS OF LIGAND BINDING *J. Biol. Chem.* 272, 18650–1865

Nadendla R.R. (2008) Principles Of Organic Medicinal Chemistry. New Age International Pvt Ltd Publishers, p 274

Narhi, L. O., Hua, Q. X., Arakawa, T., Fox, G. M., Tsai, L., Rosenfeld, R., Holst, P., Miller, J. A., and Weiss, M. A. Role of native disulfide bonds in the structure and

activity of insulin-like growth factor 1: genetic models of protein-folding intermediates (1993) *Biochemistry* (32): 5214–5221

Nielsen, D. A., Welsh, M., Casadaban, M. J., and Steiner, D. F. (1985) Effects of glucose and cyclic AMP on the transcription of insulin mRNA *J. Biol. Chem.* 260, 13585–13589.

Nolan, C., E. Margoliash, J. D. Peterson, and D. F. Steiner. (1971) The structure of bovine proinsulin. *J. Biol. Chem.* (246): 2780-2795.

O'Callaghan P M and James D C (2008) Systems biotechnology of mammalian cell factories *Briefings in functional Genomics and Proteomics* 7(2) 95-110

Omata, W., Shibata, H., Li, L., Takata, K., Kojima, I. (2000). Actin filaments play a critical role in insulin-induced exocytotic recruitment but not in endocytosis of GLUT4 in isolated rat adipocytes. *Biochem. J.* (346): 321-328.

Omidfar, K., Rasaee, M.J., Kashanian, S., Paknejad, M. & Bathaie, Z. (2007). Studies of thermostability in Camelus bactrianus (Bactrian camel) single-domain antibody specific for the mutant epidermal-growth-factor receptor expressed by Pichia. *Biotechnology and Applied Biochemistry.* (46): 41-49

Omidfar, K., Rasaee, M.J., Kashanian, S., Paknejad, M. & Bathaie, Z. (2007). Studies of thermostability in Camelus bactrianus (Bactrian camel) single-domain antibody specific for the mutant epidermal-growth-factor receptor expressed by Pichia. *Biotechnology and Applied Biochemistry* (46): 41-49

Ontsouka C.E., Bruckmaier R.M. and Blum J.W. (2003) Fractionized milk composition during removal of colostrum and mature milk. *J. Dairy Sci.* (86): 2005–2011.

Orci L. 1986 The insulin cell: its cellular environment and how it processes (pro)insulin. *Diabetes Metab Rev.* 2(1-2):71-106.

Pak S.C.O., Hunt, S.M.N. Sleigh, M.J. Gray P.P. (2002) Expression of recombinant human insulin in Chinese hamster ovary cells is complicated by intracellular insulin degrading enzymes New Developments and New Applications in Animal Cell Technology, pp 59-67

Papandreou, M. J. and Fenouillet, E. (1997) Effect of various glycosidase treatments on the resistance of the HIV-1 envelope to degradation. *FEBS Lett* (406):191-195.

Payne R.W. and Manning M.C. (2009) Peptide formulation: challenges and strategies. *Innov. Pharm. Technol.* (28):64-68.

Pekar A.H., Frank B.H., Conformation of proinsulin. A comparison of insulin and proinsulin self-association at neutral pH *Biochemistry* 11 (1972) 4013-4016.

Pereira, D. I. A., and Gibson, G. R. (2002) Effects of consumption of probiotics and prebiotics on serum lipid levels in humans. *Critical Reviews in Biochemistry and Molecular Biology* (37): 259-281.

Permutt, M. A., and Kipnis, D. M. (1972). Insulin biosynthesis. II. Effect of glucose on ribonucleic acid synthesis in isolated rat islets. *J Biol Chem* (247) 1200-1207.

Peter M.O'Callaghan and David C. James; Systems biotechnology of mammalian cell factories, *BRIEFINGS IN FUNCTIONAL GENOMICS AND PROTEOMICS.* VOL 7; NO 2; 95-110

Peterson, J. D., Steiner, D. F., Emdin, S. O., and Falkmer, S. (1975) The amino acid sequence of the insulin from a primitive vertebrate, the Atlantic hagfish *(Myxine glutinosa). J. Biol. Chem. 250:* 5183-5191.

Petsch D, Anspach F.B. (2000) Endotoxin removal from protein solutions. *J Biotechnol* (76):97–119.

Pfeifer, A. and Verma IM (2001) Gene therapy: promises and problems. Annu Rev Genomics Hum Genet (2):177–211.

Polonsky WH, Jackson RA. (2004)What's so tough about taking insulin? Addressing the problem of psychological insulin resistance in type 2 diabetes. *Clin Diabetes.* 2004 (22):147-150.

Proudfoot N, (2011) Ending the message: poly (A) signals then and now *Genes & Dev* (25): 1770-1782.

Qiao R, Huang C, Du H, Zeng G, Li L, Ye S. (2011) Milk consumption and lactose intolerance in adults *Biomed Environ Sci.* 24(5):512-517

Quin D, Orci L, Ravazzola M and Moore H-PH (1991) Intracellular transport and sorting of mutant human proinsulin that fails to form hexamers. *J Cell Biol* (113):987–996

Ramchandran A., C. Snehalatha, A. Kapur, V. Vijay, V. Mohan, A.K. Das, Rao PV, Yajnik CS, Prasanna Kumar KM, Nair JD (2001) High prevalence of diabetes and impaired glucose tolerance in India: National Urban Diabetes Survey. *Diabetologia* (44) 1094–1101.

Ramm, G., Larance, M., Guilhaus, M., James, D.E. (2006) A role for 14-3-3 in insulin-stimulated GLUT4 translocation through its interaction with the RabGAP AS160. *J. Biol. Chem.* (281): 29174-29180.

Rao M.B., Gupta R.C. and Dastur N.N. (1970) Camel's milk and milk products. *Indian Journal of Dairy Science.* (23): 71-8.

Rappaport , E. B. and Graham, D. J. (1987) Pituitary growth hormone from human cadavers: Neurologic disease in ten recipients. *Neurology* 37 (7): 1211-1213.

Recillas-Targa, F. (2006) Multiple strategies for gene transfer, expression, knockdown, and chromatin influence in mammalian cell lines and transgenic animals. *Mol Biotechnol* 34(3):337–354.

Reddy, B. S., Mastromarini, A., & Wynder, E. (1977) Diet and metabolism: large bowel cancer. *Cancer* (39): 1815-1819.

Reeds PJ, Burrin DG, Davis TA, Fiorotto ML, Stoll B, van Goudoever JB.(2000) Protein nutrition of the neonate. *Proc Nutr Soc.* 59(1):87-97

Reeves P., (2006) Structural Biology of Membrane Proteins *The Royal society of Chemistry* 17-19.

Reichert JM, Rosensweig CJ, Faden LB, Dewitz MC (2005) Monoclonal antibody successes in the clinic. *Nat Biotechnol* (23): 1073–1078.

Riechmann L and Muyldermans S. (1999) Single-domain antibodies: comparison of camel V H and camelised human V H domains [Review]. *J Immun Methods* (231): 25–38.

Ringborg A, Yin DD, Martinell M, Stalhammar J, Lindgren P. (2009) The impact of acute myocardial infarction and stroke on health care costs in patients with type 2 diabetes in Sweden. *Eur J Cardiovasc Prev Rehabil.* 16(5):576-582

Roesler,J., Brenner, S., Bukovsky, A. A., Whiting-Theobald, N., Dull, T., Kelly, M., Civin, C. I., and Malech, H. L. (2002) Third-generation, self-inactivating gp91 (phox) lentivector corrects the oxidase defect in NOD/SCID mouse-repopulating peripheral blood-mobilized CD34+ cells from patients with X-linked chronic granulomatous disease. *Blood* 100(13): 4381–4390

Romanos, M. A., Scorer, C. A., and Clare, J. J. (1992). Foreign gene expression in yeast: A review. *Yeast.* (8):423-488.

Rosenfeld L., (2002). Insulin: Discovery and Controversy, *Clinical Chemistry* (48):12 2270–2288

Rydena L, Standla E, Bartnika M, Van den Berghea G, *et al.,* (2007) Guidelines on diabetes, pre-diabetes, and cardiovascular diseases. *Eur Heart J.* 28(1):88-136.

Sahdev S., Khattar S. K., and Saini K. S. (2008) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies *Mol Cell Biochem* (307):249–264.

Sauer B. (1994) site-specific recombination developments and applications *Curr Opin Biotechnol* 5(5):521-7.

Schenborn ET, Goiffon V (2000) DEAE-dextran transfection of mammalian cultured cells. *Meth Mol Biol* (130):147–153.

Schwartz, H. J. & M. DIOLI (1992) The one-humped camel (*Camelus dromedarius*) in eastern Africa: a pictorial guide to diseases, health care and management. Verlag Josef Markgraf, Weikersheim, Germany.

Shabo Y, Barzel R, Margoulis M and Yagil R. (2005) Camel Milk for Food Allergies in Children. *Immunology and Allergies* 7(12): 796–798.

Shabo, Y., R. Barzel, M. Margoulis and R. Yagil, (2005) Camel milk for food allergies in children. *Isr. Med. Assoc.* J., (7): 796-798.

Shalash, M. R. (1984). The production and utilization of camel milk. In W. R. Cockrill (Ed.), The Camelid: An all-purpose animal. Uppsala, Sweden: Scandinavian Institute of African Studies. pp. 196-208

Shehadeh, N.; Gelertner, L.; Blazer, S.; Perlman, R.; Solovachik, L.; Etzioni, A., (2001) Importance of insulin content in infant diet: Suggestion for a new infant formula. *Acta Paediatrica.* 90(1): 93-95.

Shepherd PR. (2005) Mechanisms regulating phosphoinositide 3-kinase signalling in insulin-sensitive tissues. *Acta Physiol. Scand.* (183):3-12.

Shiao MS, Liao BY, Long M, Yu HT (2008) Adaptive evolution of the insulin two-gene system in mouse. *Genetics* (178): 1683–1691

Shigekawa, K. and Dower, W.J. (1988) Electroporation of eukaryotes and prokaryotes: A general approach to the introduction of macromolecules into cells. *BioTechniques* (6): 742–751.

SHU J. CHAN, SIMON C. M. KWOK, AND DONALD F. STEINER (1981) The Biosynthesis of Insulin: Some Genetic and Evolutionary Aspects. *DIABETES CARE.* 4 (1) 4-10.

Shushanov SS. (2011) Insulin-like growth factors 1 and 2 regulate expression of β-casein in vitro in mouse mammary epithelial cells *Bull Exp Biol Med.* 152(2):202-205

Sieber P., Kamber B., Hartmann A., Jöhl A., Riniker B., Rittel W.; (1974) Total synthesis of human insulin under directed formation of the disulfide bonds*Helv. Chim. Acta* (57) 2617-2621

Simon SA (1998) Modulation of the binding of signal peptides to lipid bilayers by dipoles near the hydrocarbon-water interface. *Biochemistry* (37): 12241-12252

Sluzky, V., Tamada, J. A., Klibanov, A. M. and Langer, R. (1991) Kinetics of insulin aggregation in aqueous solutions upon agitation in the presence of hydrophobic surfaces. *Proc Natl Acad Sci USA* (88):9377-9381.

Soares, M. B., Schon, E., Henderson, A., Karathanasis, S. K., Cate, R., Zeitlin, S., Chirgwin, J., and Efstratiadis, A. (1985). RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon. Mol Cell Biol 5, 2090-2103.

Solomons T.W.G. (1992) Organic Chemistry, 5th ed., John Wiley and Sons, New York.

Speicher, K. D., Kolbas, O., Harper, S., Speicher, D. W. (2000) Systematic Analysis of Peptide Recoveries from In-Gel Digestions for Protein Identifications in Proteome Studies *J. Biomol. Tech.* (11), 74–86

Spence JD, Jenkins DJ, Davignon J. (2010) Dietary cholesterol and egg yolks: not for patients at risk of vascular disease *Can J Cardiol* 26(9): 336-339.

Steiner D, Bell G, and Tager H (1995) Chemistry and biosynthesis of pancreatic protein hormones. In: *Endocrinology,* edited by L DeGroot. Philadelphia, PA: Saunders p. 1296–1328.

STEINER D. F., CHAN S. J., WELSH J. M., KWOK S. C. (1985) Structure and evolution of the insulin gene. *Ann. Rev. Genet.* (19): 463-484.

Steiner, D. F. (1998) The proprotein convertases *Curr. Opin. Chem. Biol.* (2): 31–39.

Steiner, D. F., 0. Hallund, A. Rubenstein, S. Cho, and C. Bayliss. (1968) Isolation and properties of proinsulin, intermediate forms, and other minor components from crystalline bovine insulin. *Diabetes* (17): 725-736.

Steiner, D. F., and. Oyer P. E. (1967) The biosynthesis of insulin and a probable precursor of insulin by a human islet cell adenoma. *Proc. Natl. Acad. Sci. U. S. A* (57): 473-480.

Steiner, D. F., S. Cho, P. E. Oyer, S. Terris, J. D. Peterson, and A. H. Rubenstein. (1971) Isolation and characterization of proinsulin C-peptide from bovine pancreas. *J. Biol. Chem.* (246): 1365-1374.

Stern J. S., Baile C. A., and J. Mayer. (1970) Are propionate and butyrate physiological regulators of plasma insulin in ruminants? *Amer. J. Physiol.,* (219):84.

Stern, B., Olsen, L.C., Tröbe, C., Ravneberg H. & Pryme I.F. (2007) Improving mammalian cell factories : The selection of signal peptide has a major impact on

recombinant protein synthesis and secretion in mammalian cells. *Trends in Cell & Molecular Biology* (2): 1-17

Stern, B., Olsen, L.C., Tröbe, C., Ravneberg H. & Pryme I.F. (2007) Improving mammalian cell factories : The selection of signal peptide has a major impact on recombinant protein synthesis and secretion in mammalian cells. *Trends in Cell & Molecular Biology* (2): 1-17

Stewart, I. I., Thomson, T., Figeys, D., (2001) 18O labeling: a tool for proteomics *Rapid Commun. Mass Spectrom.*, (15): 2456–2465.

Strickley R.G., Brandl M., Chan K.W., Straub K., Gu L. (1990) High-performance liquid chromatographic (HPLC) and HPLC mass spectrometric (MS) analysis of the degradation of the luteinizing hormonereleasing hormone (LH-RH) antagonist RS-26306 in aqueous solution. *Pharm. Res.* (7):530-536.

Strong, C. R., Forsyth, I. A. and Dils, R. (1972) The effects of hormones on milk-fat synthesis in mammary explants from pseudopregnant rabbits. *Biochemn. J.* (128): 509-519.

Sures I., Goeddel D.V., Gray A., Ullrich A. (1980) Nucleotide Sequence of Human Preproinsulin Complementary DNA *Science* 208:57-59

Surwit R.S., Schneider M.S., Feinglos M.N., (1992) Stress and diabetes mellitus. Diabetes Care; (15): 1413-1422.

Sylvie F. Gauthier, Yves Pouliot, and Jean-Louis Maubois (2006) Growth factors from bovine milk and colostrum: composition, extraction and biological activities. *Le Lait* 86 (2): 99-125.

Taylor KM, Hungerford DA, Snyder RC, Ulmer FA (1968) Uniformity of karyotypes in the Camelidae *Cytogenetics* (7): 8-15.

Terpe K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems 2006. *Appl Microbiol Biotechnol* 72:211–222

Teuscher A. (2007) Insulin: A Voice for Choice. S Karger Pub. pp 4-13

Topper, Y. J. (1970) Multiple hormone interactions in the development of mammary gland in vitro. *Recent Prog. Horm. Res.* (26):287-303.

Trenkle A. (1972) Radioimmunoassay of plasma hormones: review of plasma insulin in ruminants. *J Dairy Sci.* 55(8):1200–1211.

Trenkle, A. (1970) Effects of short-chaln fatty acids, feeding, fasting and type of diet on plasma insulin levels in sheep. *J. Nutrition* (100):1323-1330.

Trenkle, A. (1971) Postprandial changes in insulin secretion rates in sheep. *J. Nutrition* 101:1099-1104.

Urashima T, Fukuda K, Messer M. (2012) Evolution of milk oligosaccharides and lactose: a hypothesis *Animal* 6(3):369-374.

Volkin D B, Klibanov A M. (1987) Thermal Destruction Processes in Proteins Involving Cystine Residues. *J. Biol. Chem.* (262): 2945-2950.

Vollenweider F, Imiinger JC, Gross DJ, Villa KL and Halban PA (1992) Processing of proinsulin by transfected hepatoma (FAO) cells. *J Biol Chem* (267):14629–14636.

Wahren J, Ekberg K, Johansson J, Henriksson M, Pramanik A, Johansson BL, Rigler R, Jörnvall H.(2000) Role of C-peptide in human physiology *Am J Physiol Endocrinol Metab* (278): 759-768

Walsh G. Biopharmaceutical benchmarks 2006. *NatBiotechnol* (24):769–76.

Walsh, G. (2005) Therapeutic insulins and their large-scale manufacture. *Appl. Microbiol. Biotechnol.* (67): 151-159.

Walsh, G. (2006) Biopharmaceutical benchmarks 2006, *Nat. Biotechnol.* (24): 769-775.

Walsh, G. and Jefferis, R. (2006) Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* (24): 1241-1252.

Waron,U. H., E. Elias, A. Gluckman, and Trainin Z. (1987) Dromedary IgG: Purification, characterization and quantification in sera of dams and newborns. *Israel J. Vet. Med.* (43):198–203.

Washbourne, P., McAllister, A.K. (2002) Techniques for gene transfer into neurons. *Curr Opin Neurobiol* 12(5):566–573.

Weiss M. A., (2009) Proinsulin and the genetics of diabetes mellitus. *Journal of Biological Chemistry*, 284 (29): 19159–19163.

Weiss, M. A., Hua, Q. X., Jia, W., Chu, Y. C., Wang, R. Y., and Katsoyannis, P. G. (2000) Hierarchical protein "un-design": insulin's intrachain disulfide bridge tethers a recognition alpha-helix. *Biochemistry* (39): 15429–15440

Welsh, M., Nielsen, D. A., MacKrell, A. J., and Steiner, D. F. (1985) Control of insulin gene expression in pancreatic beta-cells and in an insulin-producing cell line, RIN-5F cells. II. Regulation of insulin mRNA stability *J. Biol. Chem.* 260, 13590–13594.

Welsh, M., Scherberg, N., Gilmore, R., and Steiner, D. F. Translational control of insulin biosynthesis Evidence for regulation of elongation, initiation and signal-recognition-particle-mediated translational arrest by glucose (1986) *Biochem. J.* 235, 459–446.

Wentworth, B. M., Schaefer, I. M., Villa-Komaroff, L., and Chirgwin, J. M. (1986).Characterization of the two nonallelic genes encoding mouse preproinsulin. *J Mol Evol* (23) 305-312.

WERNERY, U. (2003) New observations on camels and their milk. Dar Al Fajr Printing and Publishing, Abu Dhabi, United Arab Emirates.

Wernery, U., B. Johnson and W.T. Ishmail (2006b) Insulin content in raw dromedary milk and serum measured over one lactation period, *J. Camel Pract. and Res.* 13(2):89-90.

Wernery, U., P. Nagy, I. Bhai, W. Schiele and B. Johnson (2006a) The effect of heat treatment, pasteurization and different storage temperatures on insulin concentration in camel milk, *Milchwissenschaft* 61(1):25-28.

Westrom BR, Ekman R, Svendsen L, Svendsen J, Karlsson BW (1987) Levels of immunoreactive insulin, neurotensin, and bombesin in porcine colostrum and milk. *J Pediatr Gastroenterol Nutr.* 6(3):460-465.

White, M. F., and Kahn, R. (1994) The insulin signaling system. *J. Biol. Chem.* 269, 1–4

Wild S., G. Roglic, A. Green, R. Sicree, H. King (2004) Global prevalence of diabetes estimates for the year 2000 and projections for 2030. *Diabetes Care* 27 1047–1053.

Wildt, S. and Gerngross, T. U. (2005) The humanization of N-glycosylation pathways in yeast, *Nature Reviews Microbiology* (3): 119-128.

WILLIAMSON, M .P. (1994) The structure and function of proline-rich regions in proteins *Biochem. J.* (297): 249-260

WILSON, D. E. & D. M. REEDER (eds) (2005) Mammal species of the world (3rd ed.). Smithsonian Institution Press, Washington, USA.

Wilson, R.T. (1984) The camel, London, Longman Group Ltd. 233 pp.

Wilson, R.T., Araya, A. & Melaku, A. 1990. The one-humped camel Technical Papers Series No. 3. New York, UNSO. p 300.

Windisch V., DeLuccia F., Duhau L., Herman F., Mencel J.J., Tang S.Y., Vuilhorgne M. (1997) Degradation pathways of salmon calcitonin in aqueous solution. J. Pharm. Sci. (86): 359-364.

Wolfenden, R. and Radzicka, A. (1994) The probability of finding a water molecule in a nonpolar cavity. *Science* (265):936-937.

Wong, P. T. and Heremans, K. (1988) Pressure effects on protein secondary structure and hydrogen deuterium exchange in chymotrypsinogen: a Fourier transform infrared spectroscopic study. *Biochim Biophys Acta* (956):1-9.

Woods, N.B. , Muessig, A. , Schmidt, M., Flygare, J. , Olsson, K. , Salmon, P., Trono, D. , Von Kalle, C., Karlsson, S. (2003) Lentiviral vector transduction of NOD/ SCID repopulating cells results in multiple vector integrations per transduced cell: risk of insertional mutagenesis. *Blood* 101 (4):1284–1289.

Wu, R. H.; McMahon, T. B. (2008) Stabilization of Zwitterionic Structures of (Gly, Ala, Val, Leu, Ile, Ser and Pro) Amino Acids by Ammonium Ions in the Gas Phase, *J. Am. Chem. Soc.* 130(10): 3065-3078

Wurm, F. M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells, *Nat. Biotechnol.* (22): 1393-1398.

Yamasaki, K., Taniguchi, Y., Takeda, N., Nakano, K., Yamasaki, T., Kanaya, S. and Oobatake, M. (1998) Pressure-denatured state of Escherichia coli ribonuclease HI as monitored by Fourier transform infrared and NMR spectroscopy. *Biochemistry* (37):18001-18009.

Yan, H., Guo, Z. Y., Gong, X. W., Xi, D., Feng, Y. M. A (2003) peptide model of insulin folding intermediate with one disulfide *Protein Sci.* (12): 768–775

Ye, G.N. Daniell H, Sanford JC. (1990) Optimization of delivery of foreign DNA into higher-plant chloroplasts. *Plant Mol. Biol.* (15): 809–19.

Yip C.C. and Ottensmeyer P. (2003) Three-dimensional Structural Interactions of Insulin and Its Receptor 278 (30): 27329–27332

Yu CM, Chin CY, Franses EI, Wang NH. (2006) In situ probing of insulin aggregation in chromatography effluents with spectroturbidimetry. *J Colloid Interface Sci.* 299(2):733-739.

Yu CM, Mun S, Wang NH (2008) Phenomena of insulin peak fronting in size exclusion chromatography and strategies to reduce fronting *J Chromatogr A.* 1192(1):121-129

Yu, S. S., and A. E. Kitabchi, (1973) Biologic activity of proinsulin and related polypeptides in the fat tissue. *J. Biol. Chem.* (248): 3753-3761.

Zagorski, O., Maman, A., Yaffe, A., Meisels, A., van Creveld, C. and Yagil, R. (1998) Insulin in milk - a comparative study. *International Journal of Animal Science* (13): 241-244.

Zagorski, O., Maman, A., Yaffe, A., Meisels, A., van Creveld, C. and Yagil, R. (1998) Insulin in milk - a comparative study. International Journal of Animal Science (13): 241-244.

Zahn M, Kallberg P.W., Slappendel G.M., Smeenge H M. (2006) A risk-based approach to establish stability testing conditions for tropical countries. *J. Pharm.* Sci. (95): 946-965.

Zanen G, Houben EN, Meima R, Tjalsma H, Jongbloed JD, Westers H, Oudega B, Luirink J, van Dijl JM, Quax WJ.(2005) Signal peptide hydrophobicity is critical for early stages in protein export by Bacillus subtilis *FEBS J.* 272(18):4617-4630.

Zhang, J., Peng, X., Jonas, A. and Jonas, J. (1995) NMR study of the cold, heat, and pressure unfolding of ribonuclease A. *Biochemistry* (34):8631-8641.

Zhao F.Q. and Keating A.F. (2007) Functional Properties and Genomics of Glucose Transporters. *Current Genomics.* (8): 113-128.

Zisman, A., Peroni, O.D., Abel, E.D., Michael, M.D., Mauvais- Jarvis, F., Lowell, B.B., Wojtaszewski, J.F., Hirshman, M.F., Virkamaki, A., Goodyear, L.J., Kahn, C.R., Kahn, B.B. (2000) Targeted disruption of the glucose transporter 4 selectively in muscle causes insulin resistance and glucose intolerance. *Nat. Med.* (6): 924-928.

Appendices:

Appendix 1: The Sequence of camel insulin Gene using CAM-INS-F2 and CAM-INS-R2 Primers:

Camel number 1 forward sequence:

>GGGGAGCGCGGCTTCTTCTACACGCCCAAGGCCCGCCGGGAGGTKGAGGACACTCAGG
GTGAGACCCCCACTGTCCCCACTGCCCCCCGGCCCCGGATCCCAGCTGGCGCCCAGGGG
GCGGTCAGGAGGGGTTTTAAAAAGGAGAATGATTCCCTCTTGGTCACATCCTCAAGCGG
CCAGCTCTTTGGGGGCTGAGCCCTGACAACACCCCAAGGGTGGGCTCGCCGTCCCCTCC
ACTCCCATCTTACCCCTCTCAGCCCCATTCCTTCCTCTCCTCCGACAAGAGCGGCTTCA
GGGAAGGGTTTTATGAAGTGAGTCAAGCCCTGGGGGCTGAGGGTCTGGGGGTGCCCCAT
CCTCCTGCCTGCATGCCTGTGGGAGACACCTCACTGCCCTGAGGGGACCCCCCCCACAC
ACAGCCCGAGGCAGTGGGAGGGGGCTCGGTCCTGCACAGGCTGCGGGGAGGTGGGTGCG
GGTGGGGGTGGGCGGGAGACAGGCCCCGCGACGGGCACAGAGGATGCCCGCCCGTTCAG
GGTCCCGCTCCAGGGCTGGCTGTCTCTCGGCAGTGGGCGGCGTGGAGCTGGGTGGAGGC
CCGGGTGCGGGCGGCCTGCAGCCCCTGGGCCCGGAGGGGCGCCCGCAAAGCGCGGCATC
GTGGAGCAGTGCTGCGCCAGCGTCTGCTCGCTCTACCAGCTGGAGA

Camel number 2 forward sequence:

>GGGGAGCGCGGCTTCTTCTACACGCCCAAGGCCCGCCGGGAGGTKGAGGACACTCAGG
GTGAGACCCCCACTGTCCCCACTGCCCCCCGGCCCCGGATCCCAGCTGGCGCCCAGGGG
GCGGTCAGGAGGGGTTTTAAAAAGGAGAATGATTCCCTCTTGGTCACATCCTCAGCGGC
CAGCTCTTTGGGGGCTGAGCCCTGACAACACCCCAAGGGTGGGCTCGCCGTCCCCTCCA
CTCCCATCTTACCCCTCTCAGCCCCATTCCTTCCTCTCCTCCGACAAGAGCGGCTTCAG
GGAGGGTTTTATGAAGTGAGTCAAGCCCTGGGGGCTGAGGGTCTCGGGGGTGCCCCATC
CTCCTGCCTGCATGCCTGTGGGAGACACCTCACTGCCCTGAGGGGACCCCCCCCACACAC
ACAGCCCGAGGCAGTGGGAGGGGGCTCGGTCCTGCACAGGCTGCGGGGAGGTGGGTGCG
GGTGGGGGTGGGCGGGAGACAGGCCCCGCGACGGGCACAGAGGATGCCCGCCCGTTCAG
GGTCCCGCTCCAGGGCTGGCTGTCTCTCGGCAGTGGGCGGCGTGGAGCTGGGTGGAGGC
CCGGGTGCGGGCGGCCTGCAGCCCCTGGGCCCGGAGGGGCGCCCGCAGAAGCGCGGCAT
CGTGGAGCAGTGCTGCGCCAGCGTCTGCTCGCTCTACCAGCTGGAGAACTA

Appendix 2: The protocol used for solid phase peptide synthesis on PSSM-8 peptide synthesizer from SHIMADZU Corporation

Protocol :

step	times	volume(micro-L)	bubbling time(sec)
INITIAL RESIN WASH	1	DMF 600	180
DeBLOCK	2	PIP 500	450
NEEDLE WASH(PIP)	ON(1)		
REMOVE PIP	5	DMF 600	60
ACTIVATE AA			180
COUPLING			45(min)
NEEDLE WASH(AA)	ON(1)		
REMOVE EXCESS AA	5	DMF 600	60
DUMMY RESIN WASH		DMF 500	
EXCESS VOLUME		5.0(mL)	

Appendix 3: Solid Phase peptide synthesis calculation sheet for Camel insulin B-chain.

```
Strategy : Fmoc/HBTU

Peptide  : Camel B Chain

Sequence : Phe-Ala-Asn-Gln-His-Leu-Cys-Gly-Ser-His-Leu-Val-Glu-Ala-
           Leu-Tyr-Leu-Val-Cys-Gly-Glu-Arg-Gly-Phe-Phe-Tyr-Thr-Pro-
           Lys-Ala-OH

         : C155 H228 N40 O41 S2 ( Mwt : 3371.87 )
         : ( Note : calculation as Cys = SH free )

Resin    : TGT ( C-Term AA on Resin )

   Resin substitution :    0.200[mmol/g]
   Resin quantity     :      100[mg]
   Free peptide       :    67.44[mg]

Amino acid station ( Excess  8.0 fold ) :    160.00[micro-mol]
                                               Mwt       [mg]
      1 : Fmoc-Lys(Boc)-OH ---------------- 468.55 ---- 74.97
      2 : Fmoc-Pro-OH --------------------- 337.38 ---- 53.99
      3 : Fmoc-Thr(tBu)-OH ---------------- 397.47 ---- 63.60
      4 : Fmoc-Tyr(tBu)-OH ---------------- 459.54 ---- 73.53
      5 : Fmoc-Phe-OH --------------------- 387.44 ---- 62.00
      6 : Fmoc-Phe-OH --------------------- 387.44 ---- 62.00
      7 : Fmoc-Gly-OH --------------------- 297.31 ---- 47.57
      8 : Fmoc-Arg(Pmc)-OH.IPE ------------ 764.99 ---- 122.40
      9 : Fmoc-Glu(OtBu)-OH --------------- 425.48 ---- 68.08
     10 : Fmoc-Gly-OH --------------------- 297.31 ---- 47.57
     11 : Fmoc-Cys(Trt)-OH ---------------- 585.72 ---- 93.72
     12 : Fmoc-Val-OH --------------------- 339.39 ---- 54.31
     13 : Fmoc-Leu-OH --------------------- 353.42 ---- 56.55
     14 : Fmoc-Tyr(tBu)-OH ---------------- 459.54 ---- 73.53
     15 : Fmoc-Leu-OH --------------------- 353.42 ---- 56.55
     16 : Fmoc-Ala-OH --------------------- 311.34 ---- 49.82
     17 : Fmoc-Glu(OtBu)-OH --------------- 425.48 ---- 68.08
     18 : Fmoc-Val-OH --------------------- 339.39 ---- 54.31
     19 : Fmoc-Leu-OH --------------------- 353.42 ---- 56.55
     20 : Fmoc-His(Trt)-OH ---------------- 619.72 ---- 99.16
( 1) 21 : Fmoc-Ser(Trt)-OH ---------------- 568.66 ---- 90.99
( 2) 22 : Fmoc-Gly-OH --------------------- 297.31 ---- 47.57
( 3) 23 : Fmoc-Cys(Acm)-OH ---------------- 414.48 ---- 66.32
( 4) 24 : Fmoc-Leu-OH --------------------- 353.42 ---- 56.55
( 5) 25 : Fmoc-His(Trt)-OH ---------------- 619.72 ---- 99.16
( 6) 26 : Fmoc-Gln(Trt)-OH ---------------- 610.71 ---- 97.72
( 7) 27 : Fmoc-Asn(Trt)-OH ---------------- 596.68 ---- 95.47
( 8) 28 : Fmoc-Ala-OH --------------------- 311.34 ---- 49.82
( 9) 29 : Fmoc-Phe-OH --------------------- 387.44 ---- 62.00
```