Isolation and Characterisation of the 5' Region Sequence for the Bovine ATP-Citrate Lyase Gene

A thesis presented in fulfilment of the requirements for the degree of
Master of Science in Biochemistry
at Massey University

Xingzhang Tong

2001

Abstract

ATP-citrate lyase (ACL) is one of the major lipogenic enzymes. It catalyzes the synthesis of acetyl-CoA from citrate in the cytosol. This is the first committed step towards the conversion of carbohydrate precursors into fatty acids. Acetyl-CoA serves as the major precursor for lipogenesis and cholestogenesis. Examination of this pathway shows that the rate of fatty acid synthesis from glucose is dependent on the activity of ACL. In rats the activity of this enzyme can be increased by feeding high carbohydrate diet and reduced to low levels by fasting. These changes are regulated at the transcriptional level.

The ruminant provides a good model to study the regulation of expression of ACL. The levels of this enzyme are high in young ruminants, but fall to very low levels once a functional rumen is developed. In adult ruminants, acetyl-CoA for fatty acid synthesis is produced directly from acetate formed by microbial fermentation in the rumen and carried to the peripheral tissues. The down-regulation of this enzyme can be reversed by the administration of glucogenic precursors by a route that bypasses their fermentation to volatile fatty acids in the rumen. An understanding of the regulation of expression of ACL in the adult ruminant and a comparison with monogastric animals will provide significant new information about the regulation of the conversion of carbohydrate into fat.

A probe containing exon 2 to exon 3 of the rat ACL gene was prepared. Its specificity to bovine genomic DNA was verified and the probe was then used to screen a bovine λ genomic library. A 17 kb clone was isolated. The restriction map of this clone was determined with several enzymes. A part of this clone (9490 base pairs) was sequenced and shown to consist of a 3 kb promoter region and doenstream seqence as far as intron 3 of bovine ACL. The transcription start sites were determined by 5'RACE. Several important features of this gene were discovered by computer analysis of the sequence. Two key transcription factor binding sites were found in the promoter region. This work provided a solid basis for further investigation towards elucidating the mechanism of the transcriptional regulation of bovine ACL and the process of lipogenesis.

Acknowledgement

I sincerely thank my supervisor, Associate Professor John Tweedie, for his remarkable guidance and supervision. His encouragement, patience and enthusiasm throughout the course of this project were deeply appreciated. I would also like to gratefully thank Carole Flyger for her help in preparing the probe and general guidance, Carolyn Young for her help in the library screening, and all members of the Twilight Zone for their help and advice in the lab.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abbreviations

ACC	acetyl CoA carboxylase
ACP	acyl carrier protein
ACS	acetyl CoA synthetase
ACL	ATP citrate lyase
ADD-1	adipocyte determination and differentiation factor-1
ATP	adenosine triphosphate
b/HLH/LZ	basic/helix-loop-helix/leucine zipper
ChoRE	carbohydrate response element
cDNA	complementary DNA
CoA	coenzyme A
cpm	counts per minute
ddNTP	dideoxynucleotide triphosphate
DEPC	diethylpyrocarbonate
dH_2O	deionised water
Dnase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
EDTA	ethylenediamine tetraacetic acid
EEO	electroendosmosis
FAS	fatty acid synthase
GlRE	glucose response element
GLUT	glucose transporter
GSP	gene-specific oligonucleotide
HEPES	N-2-hydroxyethyl piperazine-N'-2-ethane sulfonic acid
HMC-CoA	3-hydroxy-3-methylglutaryl-CoA
IPTG	isopropyl β-D-thiogalactoside
LDL	low density lipoprotein
L-PK	L-type pyruvate kinase
λ	bacteriophage lambda

mRNA	messenger RNA
NADPH	nicotinamide adenine dinucleotide phosphaste, reduced form
NLS	n-lauryl sarcosine
nt	nucleotide
PCR	polymerase chain reaction
pfu	plaque forming units
Pol II	RNA polymerase II
PUFA	polyunsaturated fatty acids
RACE	Rapid Amplification of cDNA Ends
RNase	ribonuclease
rpm	revolutions per minute
RT	reverse transcriptase
SCD	stearoyl-CoA desaturase
SDS	sodium dodecyl sulphate
SRE	sterol regulatory element
SREBP	sterol regulatory element binding protein
SSC	sodium chloride and sodium citrate solution
TAE	tris-acetate buffer containing EDTA
Taq	*Thermus aquaticus*
TBP	TATA box binding protein
TdT	terminal deoxynucleotidyl transferase
TE	tris-HCl buffer containing EDTA
TSS	transcription start site
USF	upstream stimulating factor
UTR	untranslated region
UV	ultraviolet light
VFA	volatile fatty acids

CHAPTER ONE: INTRODUCTION

1.1 Overview of Lipogenesis

Lipids are a class of biological molecules playing a variety of cellular roles. They are the principal form of stored energy in most organisms, as well as the major constituents of cell membranes. Specialized lipids serve as pigments, cofactors, detergents, transporters, hormones, extracellular and intracellular messengers and anchors for membrane proteins. The ability to synthesize a variety of lipids is therefore essential to all organisms.

Fatty acids are important for the function of all cells because they are the major components of both triacylglycerols and phospholipids, are substrates for prostanoid production, and are essential for the modification of some proteins and complex carbohydrates. As a consequence, most tissues synthesize fatty acids at a low rate, but that rate is not regulated by diet or hormones. In contrast, the synthesis of fatty acids in liver and adipose tissue can occur at rates 10-1,000 times faster than in other tissues and is regulated by diet and hormones. Lipogenesis is defined as the conversion of glucose to triacylglycerol. Lipogenesis in liver and adipose tissue converts dietary carbohydrate to triacylglycerol if the carbohydrate is in excess of that required for immediate energy needs. Glucose derived from the metabolism of dietary carbohydrate is the primary substrate for fatty acid synthesis in most tissues in the body. The reactions of glycolysis, the citric acid cycle, and fatty acid synthesis *per se* give rise to the carbon backbone of the fatty acids and glycerol; reactions that generate NADPH provide hydrogen atoms for the fatty acids. When glucose is the main substrate used for fatty acid synthesis, the enzymes of the glycolytic pathway can be considered as an extended part of the lipogenic pathway (Stryer, 1995).

Glycolysis is the set of reactions that converts glucose into pyruvate in the cytosol. Under aerobic conditions, the pyruvate dehydrogenase complex in mitochondria catalyses the conversion of pyruvate to acetyl-CoA, which is then completely oxidized to CO_2 by the citric acid cycle, or transferred from mitochondria to the cytosol for fatty acid synthesis. Mitochondria, however, are not readily permeable to acetyl-CoA. The barrier to acetyl-

CoA is bypassed by production of citrate, which carries acetyl groups across the inner mitochondrial membrane. Citrate is formed in the mitochondrial matrix by the condensation of acetyl-CoA with oxaloacetate. When present at high levels, citrate is transported to the cytosol, where it is cleaved by ATP-citrate lyase (ACL) to release oxaloacetate and acetyl-CoA which then serves as the precursor for fatty acid and cholesterol synthesis. Fatty acid synthesis starts with the carboxylation of acetyl-CoA to malonyl-CoA, the committed step. This ATP-driven reaction is catalysed by acetyl- CoA carboxylase. The intermediates in fatty acid synthesis are linked to an acyl carrier protein (ACP), specifically to the sulphydryl terminus of its phosphopantetheine prosthetic group. Acetyl-ACP is formed from acetyl-CoA, and malonyl-ACP is formed from malonyl-CoA. Acetyl-ACP and malonyl-ACP condense to form acetoacetyl-ACP, a reaction driven by the release of CO_2 from the activated malonyl unit. This is followed by a reduction, a dehydration, and a second reduction. The butyryl-ACP formed in this way is ready for a second round of elongation, starting with the addition of a two-carbon unit from malonyl-ACP. Seven rounds of elongation yield palmitoyl-ACP, which is hydrolyzed to palmitate. In higher organisms, the enzymes carrying out fatty acid synthesis are covalently linked in a multi-functional enzyme complex called fatty acid synthase. The flexible phosphopantetheinyl unit of ACP carries the substrate from one active site to another in this complex. Fatty acids and glycerol-3-phosphate, through a series of reactions, ultimately form triacylglycerols as the main form of energy storage (Stryer, 1995).

The relative importance of the liver and adipose tissue as sites of lipogenesis is different in different organisms. Fatty acid synthesis occurs primarily in the liver of humans and birds. Rodents can synthesize fatty acids at high rates in both liver and adipose tissue, whereas in pigs and ruminants, adipose tissue is the primary site (Hillgartner *et al.*, 1995).

1.2 Regulation of Lipogenesis

Mammals and other higher species have evolved complex mechanisms to maintain a constant supply of metabolic energy in the face of highly variable food supplies.

Although carbohydrate, fat, or protein in the diet may exceed the daily caloric requirements, the main form of energy storage is triglycerides in adipose tissue. Several mechanisms are known to be involved in this regulation.

Nutritional regulation. The amount and composition of macronutrients in the diet, particularly fat and carbohydrate, regulate the rate of fatty acid synthesis. In general, starvation decreases, and refeeding a high-carbohydrate low-fat diet increases, the activities of the lipogenic enzymes. The rate of lipogenesis changes as much as 100-fold during the transition. The key pathways involved in regulating flux of glucose and other precursors to acetyl-CoA act similarly. Animals fed diets that are high in carbohydrate, and contain little or no fat, have higher rates of lipogenesis than those fed diets rich in fat and low in carbohydrates. Diets with fructose as the exclusive source of carbohydrate cause higher rates of fatty acid synthesis and higher activities of the lipogenic enzymes than diets containing equivalent amounts of glucose. In contrast, the addition of fat, in particular polyunsaturated fatty acids, reduces the rate of fatty acid synthesis and the activities of the lipogenic enzymes (Hillgartner *et al.*, 1995).

Hormonal factors. Qualitative and quantitative changes in diet are communicated to the body's organs by circulating hormones and fuels. Altered concentrations of these mediators signal the liver and other organs to increase or decrease the activities of metabolic enzymes and proteins that regulate the transport, storage, and catabolism of endogenous or exogenous nutrients. With respect to lipogenesis, insulin and thyroid hormone are positive effectors; elevated levels of these hormones are characteristic of the fed state. Glucagon is a negative effector; elevated levels are characteristic of the starved state. Growth hormone, glucocorticoids, and growth factors can also regulate the activities of lipogenic enzymes under some conditions; the roles of these agents in the dietary regulation of lipogenic enzyme activities are unknown (Hillgartner *et al.*, 1995).

Nutrients and metabolic products. Recent experiments suggest that circulating fuels such as glucose and fructose (positive effectors) and long-chain fatty acids (negative

effectors) also communicate nutritional status to tissues, but the mechanisms involved are only beginning to be unraveled (Towle *et al.*, 1997).

1.3 The Molecular Mechanisms of the Regulation of Lipogenesis

The activity of the lipogenic pathway is dependent on nutritional conditions, both in liver and adipose tissue. Consumption of a diet rich in carbohydrates stimulates the lipogenic pathway, whereas starvation or consumption of a diet rich in lipids and poor in carbohydrates decreases its function. The flow of molecules in most metabolic pathways is determined primarily by the amounts and activities of certain enzymes rather than by the amount of substrate available. Regulation of substrate flow through a metabolic pathway is imposed at two levels. Short-term changes in flux are initiated by altering the catalytic efficiency of pace-setting enzymes. Both covalent modifications, such as phosphorylation-dephosphorylation, and allosteric mechanisms are involved. These rapidly reversible modifications allow pathway flux to respond quickly to changes in the concentrations of humoral factors. Long-term adjustments of enzyme activity occur when stimuli for altered flux are sustained for prolonged periods and generally involve changes in the concentrations of regulatory enzymes (Hillgartner *et al.*, 1995).

Enzymes central to the process of lipogenesis are those that catalyze fatty acid biosynthesis: acetyl-coenzyme A carboxylase (ACC); fatty acid synthase (FAS); and ATP-citrate lyase (ACL), which plays a role in the transfer of acetyl-CoA from the mitochondrion to the cytosol, where fatty acid synthesis occurs. All three of these enzymes are induced by diets favoring lipogenesis, and the extent of the responses of these enzymes and their mRNAs is quite dramatic. Increases of 25- to over 100-fold have been reported between the fasted and fasted/refed states (Elshourbagy *et al.*, 1990; Paulauskis and Sul, 1988, 1989). In the majority of cases examined, transcriptional regulation is the major site for regulation of mRNA levels caused by a high-carbohydrate diet. Examples include FAS (Paulauskis and Sul, 1988, 1989) and ACL (Kim *et al.*, 1992). Apart from transcriptional regulation, the activity of ACC is also subject to covalent modifications (Kim 1997).

1.3.1 General Model of Gene Transcription

The control of transcription appears to be the most important component in gene regulation. For most genes, the primary control point is the regulation of transcriptional initiation. In eukaryotes, DNA is assembled into chromatin, which maintains genes in an inactive state by restricting access to RNA polymerase II (Pol II) and its accessory factors. During the process of development, genes are turned on and off in a pre-programmed fashion. In a typical gene, a DNA sequence called the core promoter is located immediately adjacent to and upstream of the gene. A typical core promoter (Figure 1) encompasses DNA sequences between approximately –40 and +50 relative to a transcription start site (Smale 1994). In most cases, the core promoter elements do not play a direct role in regulated transcription. A unifying principle is that transcription factors have primary responsibility for recognizing the characteristic sequence elements of any particular promoter, and they serve in turn to bind the Pol II and to position it correctly at the startpoint at each type of promoter. The initiation complex is assembled by a series of reactions in which individual factors join or leave the complex.

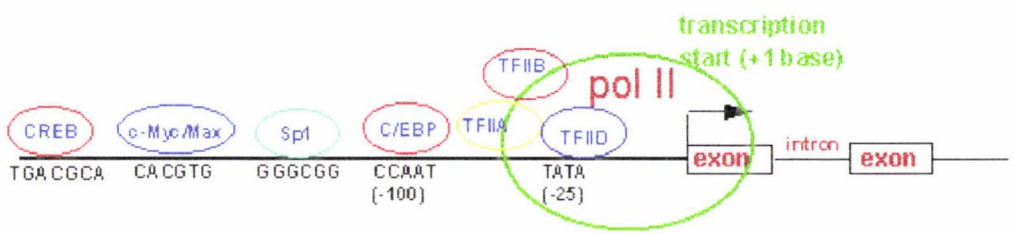


Figure 1: A general model of gene transcription.

The TATA box near the start point, and the initiator region immediately at the startpoint, are responsible for selection of the exact startpoint at promoters for Pol II. TBP (TATA box binding protein) binds directly to the TATA box when there is one; in TATA-less promoters TBP is located near the startpoint by other means. The general transcription factors for Pol II assemble the basal transcription apparatus at the promoter, and are

mostly released when Pol II begins elongation. The initiation complex that binds the core promoter comprises two classes of factors: (1) the general transcription factors including Pol II, TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH; and (2) coactivators and corepressors that mediate response to regulatory signals (Smale 1994).

At least one class of promoter appears to lack a TATA box but instead contains several transcription initiation sites, a high G/C content, and multiple binding sites for the ubiquitous mammalian transcription factor Sp1 (Smale 1997). On these promoters, which often are associated with "housekeeping genes", Sp1 directs the formation of initiation complexes to a region 40 to 100 bp downstream of its binding sites.

Some regulatory promoter elements are present in many genes and are recognized by ubiquitous factors; others are present in a few genes and are recognized by tissue-specific factors. Elements that uniquely identify particular groups of genes that are regulated in response to certain transcription factors are called response elements. The elements occur in different combinations in individual promoters. Binding of factors to specific sequences is followed by protein-protein interactions with other components of the general transcription apparatus. Transcription factors often have a modular construction, in which there are independent domains responsible for binding to DNA. Transcription factors can influence transcription levels by either enhancing (activators) or antagonizing (repressors) the assembly or activity of the basal transcription machinery (Carey and Smale, 1999).

1.3.2 Glucose/Insulin Regulation in Lipogenic Gene Expression

Despite the fact that changes in gene expression are an important component in the response of adipose and other tissues to altered nutritional states, the precise mechanisms responsible for this are not known. Each component of the diet contributes to the regulation of lipogenic gene expression in a different manner. Glucose, the most abundant monosaccharide in nature, provides a very good example of how organisms have developed regulatory mechanisms to cope with a fluctuating level of nutrient supply. In mammals the response to dietary glucose combines effects directly related to glucose

metabolism itself and effects secondary to glucose-dependent hormonal modifications, mainly pancreatic stimulation of insulin secretion and inhibition of glucagon secretion. In the pancreatic cells, glucose is the primary physiological stimulus for the regulation of insulin synthesis and secretion. In the liver, glucose, in the presence of insulin, induces expression of genes encoding glucose transporters and glycolytic and lipogenic enzymes (Towle *et al.*, 1997).

In most glucose-sensitive tissues, glucose entry is mediated through specific glucose transporters; GLUT2 in the liver and β-cells, and GLUT4 in adipocytes and muscle (Olson and Pessin, 1996). After its entry into liver, adipocytes, and β-cells, glucose has to be metabolized to generate an intracellular signal that allows for transcriptional regulation of metabolic genes. Several glucose metabolic pathways and metabolic intermediates have been investigated. Although it is now clear that the first step of glycolysis, namely phosphorylation of glucose to glucose 6-phosphate, is instrumental in glucose-dependent regulation (either positive or negative), the subsequent steps are still disputed and the exact nature of the active intermediate remains obscure (Towle *et al.*, 1997).

Study of the promoters of the lipogenic enzymes would greatly aid efforts to unravel their regulation in response to changes in glucose metabolism. One potential means of achieving this is to identify the sequences in promoters of lipogenic enzyme genes that are required for regulation. These response elements would be the binding sites for transcription factors that are either directly modulated in response to glucose metabolism or a part of a complex including such a protein. DNA sequences and DNA binding complexes involved in glucose-regulated gene expression have been characterized recently in several lipogenic genes, *e.g.* L-type pyruvate kinase (L-PK), S14, ACC, and FAS, and repressive genes of the gluconeogenic pathway, such as the phosphoenolpyruvate carboxykinase (Girard *et al.*, 1997).

Examination of the ChoRE (carbohydrate response element)/GlRE (glucose response element) sequences of the L-PK and S14 genes reveals some striking similarities (Figure 2). Both genes contain two copies of a motif (5'-CACGTG-3') related to the consensus

binding site known as the E-box which binds the upstream stimulating factor (USF). USF belongs to the basic/helix-loop-helix/leucine zipper (b/HLH/LZ) family, and mediates glucose regulation. In both the L-PK and S14 ChoREs, the two CACGTG motifs are separated by 5 bp. Both the spacing and the orientation of CACGTG motifs are critical to control. USF is present in nuclear extracts from liver and binds to the ChoREs of either the L-PK or the S14 genes (Towle *et al.*, 1997). USF also binds to a site in the FAS promoter that is critical for a response to insulin and/or glucose in mouse 3T3-L1 adipocytes. Nevertheless, it is likely that the transcription complex that accounts for glucose responsiveness is not organized in the same way in each of these three genes (Girard *et al.*, 1997). In addition, the tissue distribution of USF is ubiquitous and does not seem to be altered by the nutritional or hormonal status. No link has been established between glucose activation of gene expression and a change in the transcriptional activity of USF (Kaytor *et al.*, 1997). These conflicting results could indicate that a factor other than USF is involved in the glucose responsiveness of glycolytic and lipogenic genes.

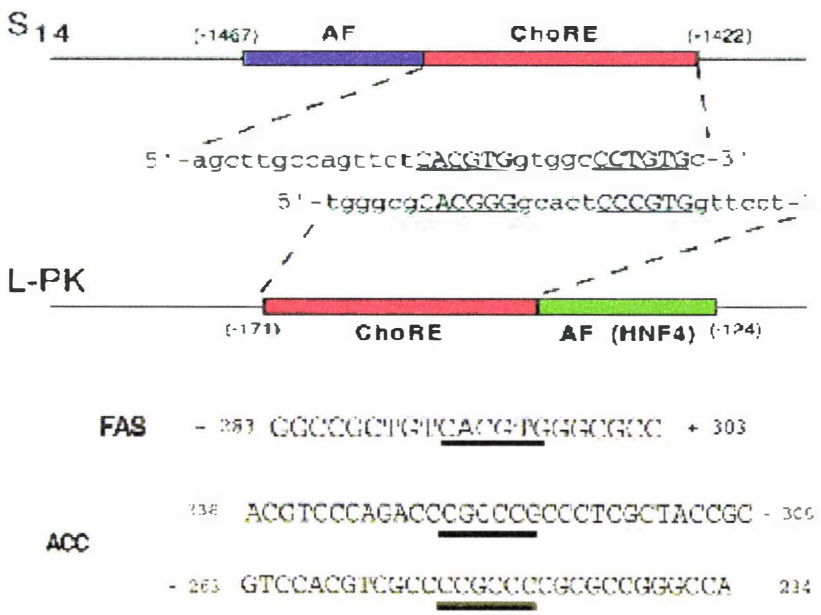


Figure 2: Minimal sequences from the L-PK, S14, FAS, and ACC genes that are able to confer glucose responsiveness on their functionality. Pictures are cited from Girard (1997) and Towle (1997).

The possibility that multiple transcription factors are involved in the carbohydrate regulation of gene expression has been suggested by studies of the ACC gene. Expression of the ACC gene from its PII promoter is induced by glucose in the mouse 30A5 adipocyte cell line. The ACC PII promoter has the structural features of a housekeeping gene (no CCAAT or TATA box) and is constitutively expressed in all tissues. Unlike the L-PK and S14 promoters, the PII promoter of ACC gene does not bind one of the USF family of transcription factors, even though it has the appropriate sequences. By using deletional analysis, the region from - 340 to - 182 of the ACC PII promoter was shown to be necessary and sufficient for supporting a glucose response. This region contains two GC boxes, the consensus binding site for the Sp1 transcription factor. Specific mutation of either GC box disrupted the response of the ACC PII promoter to glucose. The increased binding activity of Sp1 was thought to be due to dephosphorylation of Sp1 (Numa and Tanabe, 1984; Luo *et al.*, 1989; Kim 1997).

Insulin regulates metabolism by altering the synthesis, stability, or translation of specific mRNA or by producing activity-altering modifications of preexisting enzyme molecules. Many classic actions of insulin (e.g., glucose transport and glycogen synthesis) occur very rapidly and require the mediation of rapid changes in the state of protein phosphorylation or dephosphorylation. It is established that, alongside these rapid non-genomic effects, changes in gene expression play critical roles in insulin action in adipose cells, liver, and other tissues. In the fed state, the insulin level in the blood is higher than that in the starved state. Insulin is a likely signal for the onset of feeding because its response to dietary carbohydrate is large and rapid. Refeeding a high-carbohydrate or glucose diet after a period of starvation elicits a greater insulin response than feeding a diet of mixed composition. In each case, circulating insulin levels correlate positively with rates of lipogenesis and the activities of the lipogenic enzymes in liver and adipose tissue (O'Brien and Granner, 1996). It is known that insulin acts in a permissive fashion with glucose, to increase the transcription of the liver isoform of pyruvate kinase (L-PK), S14, FAS and ACC genes. In contrast, the glucokinase gene is activated by insulin through a direct intracellular signaling mechanism that is independent of extracellular glucose (Moustaid *et al.*, 1994; Girard *et al.*, 1997).

The available data suggest that insulin mediates its action on gene transcription by two basic mechanisms: (1) through plasma membrane-initiated signaling cascades that result in changes in nuclear protein phosphorylation/dephosphorylation, and (2) more directly via transcription factors. Several *cis*-acting elements that mediate the effect of insulin on gene transcription have recently been defined (O'Brien and Granner, 1996). In contrast to other hormone response elements, however, no consensus insulin response element has been identified to date. In addition, the identity of transcription factors responsible for mediating key insulin actions on gene expression has been obscure until recently. Several recent studies provide important new information on the identity of an insulin-responsive transcription factor that appears to be linked to regulation of hepatic glucose and lipid metabolism. This transcription factor, known alternatively as ADD-1 (adipocyte determination and differentiation factor-1) or SREBP-1c (sterol regulatory element binding protein-1c), was previously shown to regulate the expression of key genes of fatty acid and triglyceride metabolism in fibroblasts, adipocytes, and the livers of transgenic mice (Foretz *et al.*, 1999). Taken together with recent data that ADD-1/SREBP-1c might mediate insulin effects on gene expression in adipose cells (Azzout-Marniche,2000), ADD-1/SREBP-1c now appears to be a strong candidate to be a general mediator of the action of insulin to regulate metabolism via effects on gene expression. The details of this pathway promise to be important for future studies of insulin action and diabetes pathogenesis.

1.3.3 The Role of SREBPs in Lipogenesis

1.3.3.1 Members of the SREBP Family

Sterol regulatory element-binding proteins (SREBPs) are members of b/HLH/LZ family of transcription factors. They are synthesized as precursors that are threaded into membranes of the endoplasmic reticulum and nuclear envelope in a hairpin orientation such that the amino and carboxyl tails both face the cytoplasm. The amino-terminal half of the precursor is clipped out of the membrane in two steps responding to regulatory cues that signal the need for increased cellular cholesterol. The released amino-terminal fragment, which contains the transcriptional activation and DNA binding domains, is

targeted to the nucleus where it activates expression of SREBP target genes (Wang *et al.*, 1994; Brown and Goldstein, 1997).

Currently, there are three forms of SREBP that have been characterized. SREBP-1a and -1c are derived from a single gene through the use of alternate promoters that give rise to different first exons, The SREBP-1c isoform is a much weaker activator of gene expression than 1a because it lacks 29 acidic amino acids present in 1a. Transgenic mouse studies have shown that SREBP-1 plays a more active role in regulating the transcription of genes involved in fatty acid synthesis than those involved in cholesterol synthesis. SREBP-2 is derived from a different gene and is known to be actively involved in the transcription of cholesterogenic enzymes. It has been shown that all cultured cells analysed to date exclusively express SREBP-2 and the -1a isoform of SREBP-1, whereas most organs, including the liver, express predominantly SREBP-2 and the -1c isoform of SREBP-1 (Shimano *et al.*, 1997; Shimomura *et al.*, 1997).

1.3.3.2 Target Genes of SREBPs

Several distinct genes of both cholesterol and fatty acid metabolism are directly activated by SREBPs in studies performed in cultured cells. Genes involved in cholesterol synthesis and regulated by SREBPs include those encoding the low density lipoprotein (LDL) receptor, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase, 3-hydroxy-3-methylglutaryl-CoA reductase, squalene synthase, and SREBP-2 (Hua *et al.*, 1993). Fatty acid synthetic genes such as ACC, FAS, stearoyl-CoA desaturase-1 and -2 (SCD), ACL, S14, and glycerol-3-phosphate acyl-transferase are also directly activated by SREBPs (Figure 3) (Foretz *et al.*, 1999; Magaña *et al.*, 2000; Shawn *et al.*, 1996).

The synthesis of fatty acids requires a second substrate, NADPH. The livers of SREBP-1a transgenic mice contain significantly elevated levels of the mRNAs for all three NADPH producing enzymes, indicating both pathways for NADPH synthesis are activated. Therefore, SREBP-1a not only activates the genes directly responsible for fatty acid synthesis, ACC and FAS, but also activates the genes responsible for the production

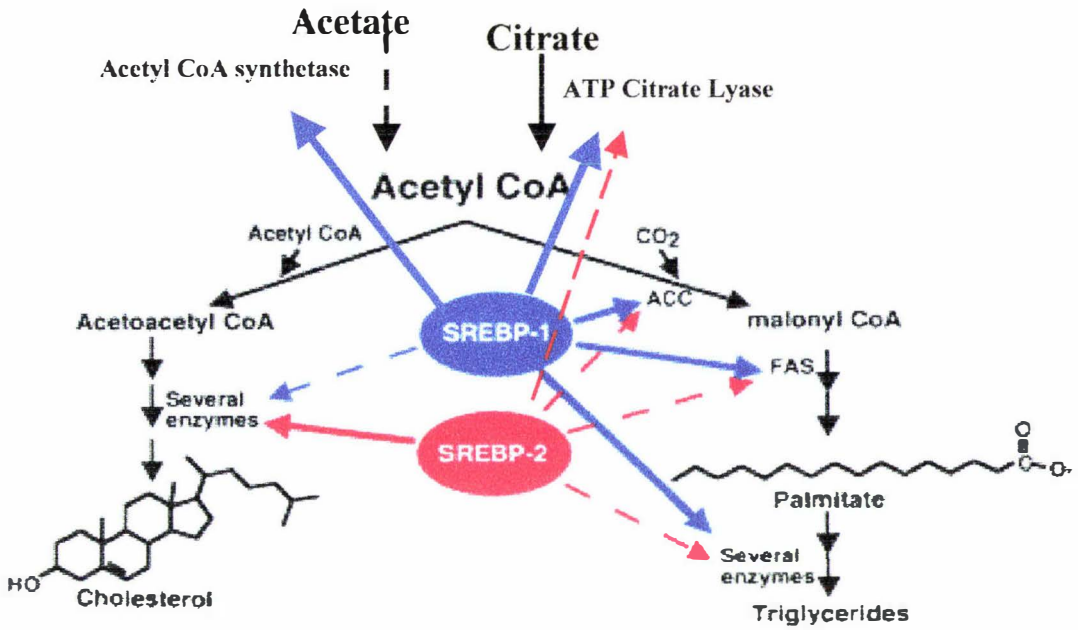


Figure 3: An outline of the target genes of SREBPs. Modified from Osborne (2000).

of the acetyl-CoA and NADPH required for fatty acid synthesis (Kim and Spiegelman, 1996; Shimomura, 1998).

It is not surprising that ATP-citrate lyase and acetyl-CoA synthetase are also regulated by SREBPs (Luong, 2000). These two enzymes represent two different pathways for providing the precursor acetyl-CoA for both fatty acids synthesis and cholesterol synthesis.

1.3.3.3 Binding Specificity of SREBPs

As members of the bHLH family of DNA binding proteins, SREBPs form dimers that recognise the E-box 5'-CANNTG-3', which was found in the ChoRE/GlRE sequence of several genes. However, SREBPs also bind to the direct repeat sterol regulatory element (SRE) 5'-TCACNCCAC-3' or to related sites (Millinder-Vallett *et al*, 1996). When the direct repeat SRE in the LDL gene was converted into an E-box identical to the existing E-box in this sterol-regulated promoter, SREBPs were still able to bind and activate this promoter, however, sterol regulation was lost (Kim et al., 1995). The results are

consistent with the mutant promoter being a target for promiscuous activation by constitutively expressed E-box binding bHLH proteins that are not regulated by cholesterol. Kim and coworkers (1995) demonstrated that the dual DNA binding specificity of SREBPs is caused by a specific tyrosine in the conserved basic region of the DNA binding domain that corresponds to an arginine in all other bHLH proteins that recognize only E-boxes. These data suggest an evolutionary mechanism where a DNA binding protein along with its recognition site have coevolved to ensure maximal specificity and sensitivity in a crucial nutritional regulatory response.

1.3.3.4 SREBP as the Major Mediator of Insulin

It is known that more than 100 genes can be regulated by insulin, and both activation and repression effects can occur within a single cell (O'Brien and Granner 1996). It is inferred that the effect of insulin on the regulation of lipogenic gene expression is mediated by other factors. As mentioned in Section 1.3.2, there has been a long effort to explore the mediator of the transcriptional effect of insulin. Kim *et al.* (1998) demonstrated that mRNA encoding SREBP-1c was suppressed by fasting and restored by feeding in adipose tissue of the mouse, and that this closely paralleled the expression of two adipocyte genes that are likewise regulated by nutrition, FAS and leptin. To complete the mechanistic loop, those authors made the critical observation that insulin could mimic these effects when added to adipocytes in culture, and that SREBP-1c could transactivate both FAS and leptin promoters. Insulin's stimulatory effect on the FAS promoter was mapped to an E-box mitif contained within a sequence previously identified as the major insulin response element of this gene. Thus, it was established that SREBP-1 was a transcription factor involved both in adipogenesis and in the regulation of insulin-responsive genes in adipose cells.

The two recent reports extend these findings considerably. Foretz *et al.* (1999) have used adenoviral vectors expressing either a dominant negative SREBP-1c or a dominant positive version of this protein. In the isolated hepatocyte model, they show that dominant negative SREBP-1c blocks the ability of insulin to induce the glucokinase gene, and a dominant positive version stimulates glucokinase gene expression in the absence of

insulin, which is otherwise necessary for expression of this gene. Shimomura *et al.* (1999) bring this finding to the *in vivo* level, by demonstrating that SREBP-1c (but not SREBP-1a) mRNA is specifically diminished in the liver of mice made diabetic with the β-cell toxin streptozotocin and is rapidly induced after insulin therapy. This specific effect was reproduced *in vitro* with isolated hepatocytes. Taken together, SREBP-1c appears to be a transcription factor poised to mediate many of the actions of insulin on expression of genes involved in both lipid and carbohydrate metabolism in two key tissues, fat and liver.

The investigation of the mechanism by which insulin regulates SREBP is progressing. Azzout-Marniche *et al.* (2000) showed that insulin, through the PI3-kinase pathway, activates the synthesis of the precursor form of SREBP-1c whereas a high glucose concentration has no significant effect. This precursor form is then rapidly cleaved, possibly by a constitutive process (Brown and Goldstein, 1997). Insulin also rapidly activated the transcription of the lipogenic genes, in which an increased amount of the mature form of SREBP-1c in the nucleus was not a prerequisite, suggesting that additional actions of the hormone are involved, such as the activation of the nuclear form of SREBP-1c or of an unidentified SREBP-1c partner (Azzout-Marniche *et al.*, 2000).

1.3.3.5 Regulation of SREBP by Glucose and Fatty Acid

Previous studies have highlighted the importance of insulin in the induction of SREBP-1 leading to regulation of downstream lipogenic enzymes. A recent study demonstrated that in well differentiated H2-35 cells, glucose also plays an important role in SREBP-1 expression (Hasty *et al.*, 2000). The SREBP-1 induction by glucose is dose- and time-dependent and is at the level of transcription. The change in SREBP-1 concentrations occurs only after 12 h of incubation in the corresponding media, suggesting that the induction might be mediated through some metabolite of glucose and that it takes some time to have an effect. Interestingly, induction of some lipogenic genes were independent of insulin in this study. It is necessary to further study the glucose metabolism pathways responsible for the carbohydrate response of SREBP-1 and lipogenic genes, and the mechanisms controlling SREBPs cleavage by different signal transduction pathways.

Because SREBPs are regulated directly by cholesterol and they are involved in both cholesterol and fatty acid metabolism, it was important to determine whether they are directly regulated by fatty acids as well. To date, many hepatic genes have been shown to be regulated by the fat component of the diet. The levels of expression of genes encoding rodent malic enzyme, ACC, ACL, L-PK, FAS, glucose transporter 4 (GLUT4), S14 protein and stearoyl-CoA desaturase (SCD1) are all known to be decreased (60-90%) by dietary polyunsaturated fatty acids (PUFA) (Sessler and Ntambi, 1998). Studies on SCD1, GLUT4 in adipocytes and the Δ-9 desaturase 2 (SCD2) gene in lymphocytes have shown that the effect of PUFA on gene expression can be at the level of both gene transcription and mRNA stability (Sessler *et al.* 1996). The mechanisms involved are just beginning to be unraveled.

Worgall *et al.* (1998) analysed the effects of PUFA on genes having sterol regulatory elements (SRE) in their promoter-regulatory regions. In transfected cells they found that oleate and PUFAs reduced expression of SRE-containing genes by decreasing the level of SREBP. Saturated fatty acids were ineffective. It was later reported that PUFAs decrease SREBP-1 gene expression and the nuclear content of SREBP-1 in liver and in primary hepatocytes but not in 3T3-L1 adipocytes (Xu *et al.*, 1999; Marter *et al.*, 1999). This observation is corroborated by results of experiments with transgenic mice either overexpressing SREBP-1 (Yahagi *et al.*, 1999) or knocked out for this gene (Shimano *et al.*, 1999).

1.3.3.6 Coregulatory Factors

In all promoters for SREBP target genes studied thus far, SREBP-dependent regulation requires an additional generic coregulatory DNA binding factor(s) for efficient expression (Dooley *et al.*, 1999). The identity of the coregulatory factor and the position of its binding site relative to the binding site(s) for the SREBPs differs from promoter to promoter.

The gene that encodes the key protein of cholesterol uptake, the low density lipoprotein (LDL) receptor, has a promoter that contains a single SREBP site flanked on either side

by a coregulatory site for the generic Sp1 protein. SREBP and Sp1 activate the LDL receptor promoter synergistically (Sanchez *et al.*, 1995, Naar *et al.*, 1998). As sterol levels fall, the increased processing and nuclear accumulation of SREBPs results in an increased association of Sp1 with the LDL receptor promoter DNA *in vivo*. SREBP recruitment of its generic coregulatory factor is not specific for Sp1 in the LDL receptor promoter because similar increased recruitment of NF-Y and CREB coregulatory factors required for the HMG-CoA reductase promoter, was also observed on depletion of intracellular cholesterol (Bennett and Osborne, 1999). In other reports, SREBPs were shown to interact directly with Sp1, NF-Y, and CREB in solution in the absence of DNA (Naar *et al.*, 1998, Xiong et al., 1999).

A very recent study showed that NF-Y is an essential coactivator of SREBPs in the cholesterol response of human FAS promoter I. Human FAS promoter I also contains two Sp1-binding sites. These sites function redundantly in cholesterol response, and Sp1 alone cannot activate the sterol response in human FAS promoter I (Xiong *et al.*, 2000).

Thus, the increased recruitment of generic transcription factors to sterol-regulated promoters by a direct protein-protein interaction is likely to be a common aspect of regulated gene expression mediated by the SREBP proteins. The common use of SREBP provides a mechanism for coordinate regulation. However, the unique coregulatory factors and subtle differences in promoter architecture provide the opportunity for more subtle promoter-specific regulatory effects that integrate other cellular signaling pathways with simple nutrient sensing to provide optimal control of cellular lipid levels.

1.3.3.7 Analysis of the SREBP-1c Promoter

Because of the importance of SREBPs in the regulation of lipid metabolism, it is interesting how SREBPs themselves are regulated. A recent report on the SREBP-1c promoter provided new information on this issue (Amemiya-Kudo *et al.* 2000). A cluster of putative binding sites for several transcription factors composed of an NF-Y site, an E-box, an SRE 3, and an Sp1 site were located at –90 bp in the SREBP-1c promoter. Luciferase reporter gene assays indicated that this SRE complex was essential to the

basal promoter activity, and conferred responsiveness to activation by nuclear SREBPs. Deletion and mutation analyses suggested that the NF-Y site and SRE3 in the SRE complex were responsible for SREBP activation, although the other sites were also involved in the basal activity. Gel mobility shift assays demonstrated that SREBP-1 binds to the SRE3. These findings implicated a positive feedback control of SREBP-1c through the SRE complex. Furthermore, reporter assays using larger upstream fragments indicated another region that was inducible by addition of sterols. Unexpectedly, this study failed to establish a connection between the glucose/insulin regulation and SREBP-1c expression, and the authors suggested that this was due to the inappropriate experiment system. Therefore, further study is absolutely necessary.

1.4 ATP-Citrate Lyase

1.4.1 Role of ATP-Citrate Lyase

ATP-citrate lyase (ACL) was first discovered in chicken liver (Srere & Lipmann, 1953), and it has subsequently been isolated from a variety of mammalian and plant sources. It functions in the cytosol as a tetramer of four identical subunits (Elshourbagy *et al.,* 1990; Elshourbagy *et al.,* 1992), with a total molecular weight of 440 kDa (Singh *et al.,* 1976). As stated above (Section 1.1), it catalyzes the formation of acetyl-CoA, which serves as an important biosynthetic precursor in lipogenesis, as well as cholesterogenesis. Further details in this pathway are given in Figure 4.

When glucose serves as the source of lipid synthesis, after glycolysis, the product pyruvate is transported across the mitochondrial membrane, where it is oxidised to acetyl CoA and CO_2 (Spencer and Lowenstein, 1962). Acetyl-CoA must be transferred from the mitochondrion into the cytosol, where the enzymes involved in fatty acid synthesis are situated. The inner mitochondrial membrane is impermeable to acetyl-CoA, so the acetyl-CoA is converted to citrate, and then transported out of the mitochondrion by the citrate translocator in the mitochondrial membrane (Halperin *et al.,* 1975). The cytosolic citrate is converted to acetyl-CoA and oxaloacetate by the action of ACL (Srere, 1959), in

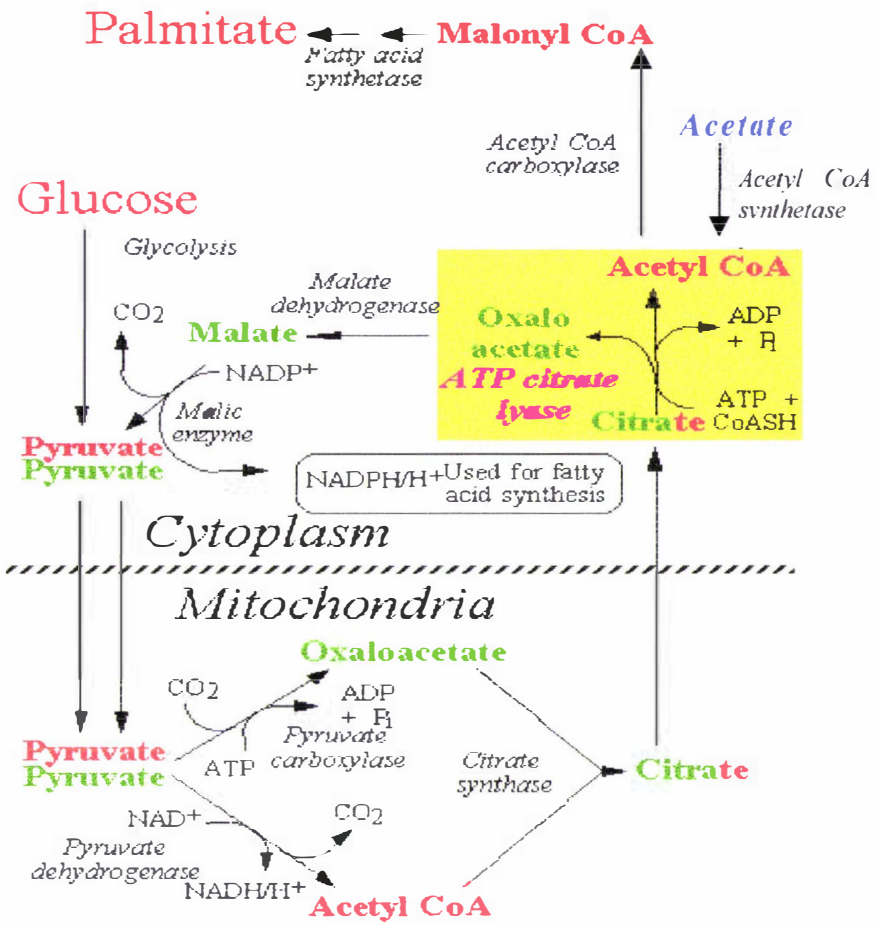


Figure 4: The role of ATP citrate lyase and an outline of lipogenesis

an ATP-dependent reaction. Acetyl-CoA is then available in the cytosol as a precursor for the biosynthesis of fatty acids and cholesterol.

Oxaloacetate that is formed in the cleavage reaction catalysed by ACL must be returned to the mitochondrion. The inner mitochondrial membrane is also impermeable to oxaloacetate, and so oxaloacetate undergoes a series of reactions to form pyruvate, which readily diffuses into the mitochondrion. These reactions result in the production of cytosolic NADPH, which is then available for the synthesis of fatty acids. Therefore, the action of this enzyme is involved in the recycling of oxaloacetate back into the mitochondrion, and the concomitant production of cytosolic NADPH which is required for the production of fatty acids (Pentyala and Benjamin, 1995).

1.4.2 Regulation of ATP-Citrate Lyase

Like other lipogenic enzymes, the activity of ACL varies in accordance with the nutritional and hormonal state of the animal. It is suppressed on starvation and restored on re-feeding. The increase in enzyme activity depends on the diet. It is largest on diets high in carbohydrate and low in fat, and smallest on diets high in fat. Intermediate increases are obtained with balanced diets. These changes are due to alterations in the rate of ACL biosynthesis (Kim *et al.*, 1994). The changes in activity of ACL correlate with the parallel changes in other lipogenic enzymes and the rate of fatty acid synthesis, providing evidence for the involvement of this enzyme in fatty acid synthesis.

In a fast/refeeding experiment, northern blot analyses demonstrated strong expression of ACL gene in liver and adipose tissue of rat and weak expression in brain, heart, small intestine and muscle. After refeeding a carbohydrate/protein diet to fasted rats, the transcriptional rate had already increased within 2 h, the mRNA concentration reached a maximal level of an approximately 30-fold increase in 16 h, and the enzyme induction increased 6-fold in 48 h (Elshourbagy *et al.*, 1990). In diabetic rats fed on a glucose diet, the transcriptional rate, mRNA concentration and enzyme level were very low in comparison with normal rats. The administration of insulin restored the ACL expression to response to carbohydrate diet. Park *et al* (1994) suggested that the increase in the amount of ACL by insulin was primarily due to the increase in the transcriptional activity of the gene in nuclei, which results in the subsequent increase in the amount of mRNA for the biosynthesis of ACL in the cytosol.

1.4.3 Phosphorylation of ATP-Citrate Lyase

In addition to ACL being regulatory in long-term control of lipid synthesis, it could also be important in its short-term regulation, along with acetyl-CoA carboxylase (ACC) activity that is regulated by both allosteric effectors and phosphorylation. ACC is phosphorylated by both cAMP-dependent and cAMP-independent protein kinases with its activity increasing with decreasing phosphorylation. ACL is also phosphorylated by both cAMP-dependent and cAMP-independent protein kinases. No functional role has

been assigned to ACL phosphorylations even though in fat and liver cells nearly all ACL undergoes phosphorylation changes that are complex (Benjamin *et al.*, 1994).

ACL has three phosphorylation sites, each four amino acids apart (threonine 446 and serines 450 and 454). Serine 454 is phosphorylated by cAMP-dependent protein kinase and an insulin-stimulated kinase. Threonine 446 and serine 450 are phosphorylated only by GSK-31. The sequence about these phosphorylation sites strongly resembles those in glycogen synthase. Their phosphorylation is hormonally regulated (Ramakrishna *et al.*, 1990).

β-Adrenergic agents increase phosphorylation of all three ACL phosphorylation sites whereas insulin increases phosphorylation at serine 454 but decreases phosphorylation at threonine 446 and serine 450 (Ramakrishna *et al.*, 1990). Insulin increases phosphorylation of serine 454 presumably by both increasing the activity of an insulin stimulated kinase and decreasing phosphorylation at sites 446 and 450. This result in turn decreases GSK-3 activity, making serine 454 a better substrate for further phosphorylation (Ramakrishna *et al.*, 1989).

Unlike ACC, the available data so far suggests that the short-term regulation of ACL is not important in the regulation of fatty acid synthesis.

1.4.4 Gene Structure

The gene encoding ACL was mapped to human chromosome 17q12-q21 and to rat chromosome 10 (Kim *et al.*, 1994; Park *et al.*, 1997). The rat ACL gene was cloned and the complete exon-intron organization of the gene identified (Elshourbagy *et al.*, 1990; Kim *et al.*, 1994). The rat ATP-citrate lyase gene, spanning about 55 kb, is divided into 29 exons that range in size from 30 to 986 base pairs. The sequences bordering the splice site junctions universally follow the GT/AG rule. The reverse transcription-polymerase chain reaction showed two forms of ACL mRNA, one containing complete exons and another lacking exon 14, were found in the brain, kidney, mammary gland, lung and liver. Also, restriction fragment length polymorphisms were observed at intron 10 and

intron 11. The sequence of the 5'-region has also been determined (Kim *et al.,* 1994). Primer-extension analysis showed that mRNA transcription is initiated 4407 nucleotides upstream from the translation initiation site. The promoter used for transcription is identical in mammary gland, lung, liver, brain and kidney. Southern-blot analysis showed that the rat ACL gene exists as a single copy. The 5' flanking region contains several consensus sequences defined as promoter elements. These include a CAAT box and Sp1-binding sites. However, this promoter lacks a TATA box.

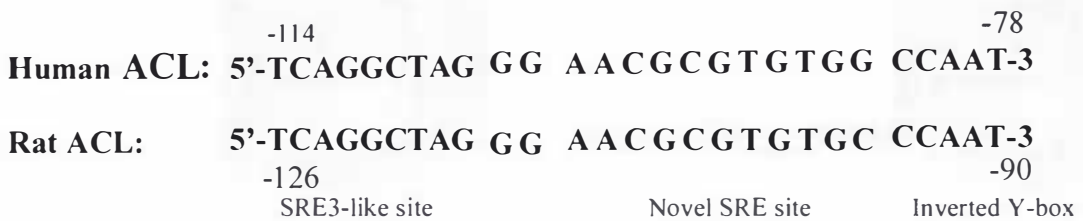
A full-length cDNA clone of 4.3 kb encoding the human ACL enzyme has been isolated (Elshourbagy *et al.*, 1992). Nucleic-acid sequence data indicate that the cDNA contains the complete coding region for the enzyme, which is 1105 amino acids in length with a calculated molecular mass of 121,419 Da. Comparison of the human and rat ATP-citrate lyase cDNA sequences reveals 96.3% amino acid identity throughout the entire sequence. The 5'-flanking region of human ACL gene was identified by screening a human placental genomic DNA library (Park *et al.*, 1997). The 5' flanking region of ACL had a CAAT box at -92 bp from the transcription initiation site (+1), however, a TATA box was not found. The sequences of 5'-flanking region (1.5 kb) of human ACL showed 60% identity with that of rat; however, the identity was less in the exon 1 and intron 1 regions. Several consensus sequences, including four Sp1 binding sites, were found in the 5' flanking region of this gene. More sequence information of human ACL will be available from the Human Genome Project.

1.4.5 Molecular Mechanism of the Regulation of ATP-Citrate Lyase

Based on the information about the ACL gene structure, several studies have been carried out to investigate the molecular mechanism involved in the regulation of the ACL gene expression. The early studies on the ACL promoter demonstrated that the region from -64 to -41 of the rat ACL gene and from -213 to –128 of the human ACL gene was responsible for stimulation due to insulin/glucose. The stimulation was suppressed by polyunsaturated fatty acids, and Sp1 may be involved in the regulation (Fukuda *et al.*, 1996, 1997; Park *et al.*, 1997).

While this project was in progress, some reports described the role of SREBPs in the regulation of the ACL gene. These studies showed that, in mouse liver, the entire coordinated program of fatty acid biosynthesis, including the ACL, can be activated by SREBPs (Shimomura *et al.*, 1998; Shimano *et al.*, 1999). ACL gene expression is enhanced by SREBP-1 in CHO cells, the liver of transgenic mice, and human Hep G2 cells. An SRE-3 site is located at –114 to –106 (TCAGGCTAG) in the human ACL gene and binds to SREBP-1a in CHO-487 cells. The binding of NF-Y to a CCAAT site 13 bp downstream of the SRE site is important for the activation by SREBP (Sato *et al.*, 2000).

Figure 5: A comparison of SRE sites and inverted Y-box in human and rat ACL.



In the rat ACL promoter, the region between nucleotides -114 and -60 was the major determinant for the activation by SREBPs (Moon *et al.*, 2000). Recombinant SREBP-1a bound strongly to three SRE sites. The first SRE at –71 to -62 (GTGAGCTGA), its inversed sequences are similar to SRE1, which has been found in many target genes of SREBPs (Magana, 2000). Mutations at this region significantly reduced the responsiveness to SREBPs without decreasing basal activity of the ACL promoter. The second SRE in -104 to -95, the region upstream of the inverted Y-box, showed a novel sequence (5'-AACGCGTGTG-3') differing from known consensus SREs. Mutations at this region prevented the binding of SREBP-1a without changing NF-Y binding and also decreased the stimulation of the ACL promoter by SREBP-1a. The third SRE at –126 to –117 (TCAGGCTAG) has a sequence highly homologous to SRE3, which was reported in the glycerol 3-phosphate acyltransferase promoter (Ericsson *et al.*, 1997). However, mutations in this region did not alter the responsiveness to SREBPs, even though SREBP-1a showed strong affinity to this region and its binding to the mutant probe was

significantly reduced (Moon *et al.*, 2000). This is interesting because such a phenomenon has not been observed in the study of human ACL (Sato *et al.*, 2000), even though the sequence of these two ACL promoter have high homology in this region (Figure 5). Furthermore, a SRE3 site and NF-Y site with similar spacing was thought to be necessary for the stimulation of SREBP target genes by SREBP binding. It might be worthwhile to study further whether this is specific for the ACL genes.

The importance of Sp1 binding sites for the action of SREBPs has been reported in many promoters, such as the LDL receptor, ACC, and FAS promoters. But in the case of the rat ACL promoter, although deletion of five upstream Sp1 binding sites decreased the basal activity of this promoter, it did not effect SREBPs activation. On the other hand, the responsiveness to SREBPs was markedly reduced when the binding of NF-Y was disturbed by the mutation at the inverted Y-box or by the overexpression of the dominant negative form of NF-YA (Moon *et al.*, 2000). These results indicate that the regulation of the ACL promoter by SREBPs requires the neighboring binding of NF-Y, as was observed for the farnesyl-pyrophosphate synthase, squalene synthase, HMG-CoA synthase, glycerol-3-phosphate acyltransferase, and SREBP-2 genes (Moon *et al.*, 2000).

Acetyl-CoA is the precursor for both fatty acid and cholesterol synthesis. It is reasonable to assume that ACL would be regulated by both SREBP-1 and SREBP-2. The mechanism by which SREBPs activate their prefered target genes has not been studied thus far. In this study, the authors observed that SREBP-2 showed much less affinity to SREs than SREBP-1 in the rat ACL promoter. Therefore, further study to determine how SREBPs differentially and/or synergistically regulate ACL will provide useful information in understanding the regulation of lipogenesis and cholesterogenesis.

1.5 ATP-Citrate Lyase in Ruminants

Ruminants are distinguished from simple-stomached or monogastric animals by the development of a series of pouches anterior to their true gastric stomach. Of these pouches, the rumen is the largest and metabolically the most important. Carbohydrates make up the major portion of the diet of ruminant animals. In the rumen of an animal fed

at a low level, essentially 100% of the readily available carbohydrates will be fermented by the rumen microorganisms. The principle end products of fermentation are the volatile fatty acids (VFA; primarily acetic, propionic, and butyric), bacterial proteins, nucleotides, carbon dioxide, methane, and heat. The animal, in turn, uses the volatile fatty acids as a source of energy for its life processes. As a result of this conversion of plant cellular constituents into VFA and microbial cells, the metabolism of the ruminant animal is different from that of the simple-stomached animal (Bergman *et al.*, 1965).

1.5.1 Glucose Metabolism in Ruminants

The glucose requirements of ruminant animals are largely the same as in other mammals. Glucose is the major circulating form of carbohydrate and is present in the blood and extracellular fluids in a far higher concentration than any other sugar. It plays a vital role in cell metabolism and adequate blood concentrations must be maintained. Glucose is required in large amounts by several specific tissues including the brain, adipose tissue, muscle, fetuses and mammary gland (Bergman, 1973).

Glucose metabolism varies in different animal species. In monogastric animals, dietary carbohydrates are digested to glucose and other simple sugars, absorbed into the portal blood, and then act as sources of tissue glycogen and glucose for the body's metabolism. In between meals, glucose is released by the liver and new glucose must be produced (gluconeogenesis) as the glycogen stores become depleted. In ruminants, however, carbohydrates are fermented in the rumen to VFAs that usually furnish 70-80% of the animal's total caloric requirements (Bergman *et al.*, 1965). Thus it is evident that insufficient glucose is absorbed by ruminants and that gluconeogenesis must be continuous and of prime importance for metabolism.

In ruminants, the liver accounts for approximately 85% of the glucose production and the kidneys virtually all of the remainder. The turnover of glucose in the whole body is variable and its utilization usually is directly proportional to its concentration in blood (Bergman *et al.*, 1973). Propionate is the major substrate for gluconeogenesis, accounting

for 54% of the glucose synthesized. Other precursors include lactate, pyruvate, and glycerol (Young , 1977).

Infant ruminants are essentially non-ruminants, their food is not fermented and they require dietary vitamins and amino acids. Their blood glucose level is high, like most non-ruminants. As the rumen develops, their metabolism adjusts towards the adult condition with a decline in blood glucose (Vernon, 1981). This seems to be an evolutionary adaptation of the adult ruminant to the perpetual necessity of glucose conservation and gluconeogenesis.

Adult ruminants maximize glucose conservation, and they have a number of metabolic adaptations not seen in non-ruminants that help in the process. Among the metabolic changes associated with glucose conservation is the virtual elimination of ATP-citrate lyase and NAD-dependent malate dehydrogenase activities. This prevents the direct conversion of glucose to cytosolic acetyl-CoA and fatty acids in ruminant tissues, thus conserving glucose for its most essential functions, particularly supplying energy for brain tissue. Although the brain of ruminants depends mainly on glucose for its metabolism, its smaller size, relative to many other species, ensures a lower total glucose need. In addition, ruminant tissues are able to take up glucose in the face of hypoglycemia that would produce coma in non-ruminants (Van der Walt and Linington, 1989).

Another difference between mature ruminants and non-ruminants is the response of blood glucose to fasting. In non-ruminants gluconeogenic mechanisms maintain blood glucose concentrations throughout a fast; in ruminants, however, concentrations decline within a day or two. The relative tolerance of ruminant tissues to low blood glucose may be possible because tissues can utilize other energy substrates. There is no great rise in the level of blood glucose after a meal in ruminants, in contrast to the case in non-ruminants. After a meal gluconeogenesis is generally supressed in non-ruminants, while it is increased in ruminants because of increased substrate availability (Harmon 1992).

1.5.2 Fatty Acid Synthesis in Ruminants

In non-ruminant animals, the major source of carbon for fatty acid synthesis is carbohydrate. The *de novo* synthesis of fatty acids from acetyl-CoA occurs in the cytosol of the cell. In non-ruminants most of the acetate is produced in the mitochondria, then being transported from the mitochondria to the cytosol as citrate which is cleaved to oxaloacetate and acetyl-CoA in the cytosol by the ACL. This sequence of synthesis is also found in fetal and neonatal sucking ruminants where appreciable quantities of exogenous carbohydrate are available in the tissues. However, once the rumen begins to function, most dietary carbohydrates will be fermented in the rumen, and the importance of glucose absorbed from the gut decreases with time until at weaning virtually no absorbed glucose is available. Glucose is now at a premium and has to be formed from gluconeogenesis (Ballard *et al.*, 1969).

Volatile fatty acids are absorbed from the rumen as the free acids. Acetic, propionic, and butyric acids are produced in the approximate ratio of 70:20:10. Butyrate and propionate are largely metabolized by the rumen wall and liver, respectively, and acetate is largely utilized in the peripheral tissues. Acetate may enter fatty acid synthesis as acetyl-CoA or through carboxylation to malonyl-CoA in adipose tissue; alternatively, it may enter the citric acid cycle through condensation with oxaloacetate. Butyrate is partly metabolized by the rumen epithelium to ketone bodies, principally acetoacetate and β-hydroxybutyrate, which are interconverted in the liver. Propionate carbon enters the citric acid cycle in the form of succinate and in a few steps is converted to the important intermediate oxalacetate. Several routes of metabolism are open for oxalacetate. It may be directly converted to glucose via a reversal of the glycolytic pathway; it may be condensed with acetyl-CoA to form citrate; or the respective keto acids may be reductively aminated to form the glycogenic nonessential amino acids (van der Walt and Linington, 1989).

The three major inputs of precursors for triacylglycerol synthesis (lipogenesis) in the bovine adipocyte are fatty acids released from chylomicrons and very-low-density lipoproteins by the action of extracellular lipoprotein lipase within the adipose tissue,

acetate, and glucose from the blood (Baldwin and Smith, 1971). Acetate is the major precursor for fatty acid synthesis, and glucose is the major precursor for glycerol of the triacylglycerols. Lactate also is converted readily to fatty acids, but, because its availability is less than that of acetate, lactate is a secondary rather than a primary precursor (Robertson *et al.*, 1981).

The reason for relatively low rates of glucose conversion to long-chain fatty acids remains unknown. Hanson and Ballard (1967) demonstrated that ruminant adipose tissue differs markedly from that of non-ruminants in the very low activities of two key enzymes involved in the translocation of acetyl-CoA from mitochondria to cytosol, namely ACL and NADP malate dehydrogenase, which converts malate to pyruvate. In consequence ruminant adipose tissue has a very limited capacity to utilize glucose carbon for fatty acid synthesis, but the large amounts of acetate formed by rumen fermentation can be utilized, after being converted to acetyl-CoA by the cytosolic acetyl-CoA synthetase.

Even though glucose is an insignificant precursor for fatty acid synthesis in ruminants, glucose stimulates the incorporation of acetate and lactate into long-chain fatty acids in bovine adipose tissue (Van der Walt and Linington, 1989). The stimulatory effect of glucose probably is the result of NADPH generation especially by the pentose phosphate cycle. Activities of glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase in the pentose phosphate cycle are sufficiently great to meet the NADPH requirement for fatty acid synthesis. In lipogenic tissues from non-ruminant animals, such as in rat adipose tissue, nearly one half of the NADPH requirement seems to be provided by NADP-malate dehydrogenase (Anderson and Bridges, 1984). On a theoretical basis, one NADPH is generated by this enzyme for each acetyl-CoA that exits the mitochondrion and incorporation of each acetyl-CoA into fatty acids requires two NADPHs. Because of very low NADP-malate dehydrogenase activity in bovine adipose tissue and because the supply of cytosolic NADH limits the conversion of oxaloacetate to malate, NADPH generation for lipogenesis by the NADP-malate dehydrogenase in bovine adipose tissue seems insignificant. Another source of reducing equivalents for

fatty acid synthesis is the isocitrate cycle. The activity of cytosolic NADP-isocitrate dehydrogenase is extremely high in ruminant adipose tissue relative to its activity in non-ruminants. Baldwin and Smith (1971) estimated that 36 to 60% of the NADPH required for fatty acid synthesis was derived from the activity of this cycle in bovine adipose tissue. The advantage of the isocitrate cycle to ruminants is that acetate can be utilized to generate NADPH and spare glucose.

The main site of lipogenesis in ruminants is the adipose tissue, where about 90% of fat synthesis occurs. The liver, which is the major site in some non-ruminant species, accounts for only 5% of fat synthesis. Concentrating lipid synthesizing ability in the adipose tissue will be of distinct advantage to the ruminant as it will allow the liver to be concerned mainly with gluconeogenesis and there will be no competition for energy, substrate or reducing equivalent (Bauman, 1976). Of the VFAs reaching the liver from the rumen, propionate is extracted almost quantitatively and used for gluconeogenesis whereas acetate is not removed but passes to the peripheral circulation and is used for energy production and fat synthesis.

1.5.3 Regulation of Lipogenesis in Ruminants

The regulation of lipogenesis in ruminants is not as well understood as it is in non-ruminants. Fatty acid synthesis from acetate, although a more direct pathway than that from glucose, is still subject to a variety of controls. In the long-term, the activities and probably the amounts of most of the associated enzymes adapt to changing physiological conditions in the ruminant as in the non-ruminant. The mechanisms responsible for these adaptations have not been investigated in the ruminant, but it is certain that insulin is involved (Harmon 1992).

With regards to short-term regulation of fatty acid synthesis from acetate, plasma acetate concentration would seem to be a critical determinant in ruminants. In addition, acetate inhibited fatty acid synthesis from glucose in adipose tissue of ruminant animals. In contrast, acetate stimulated fatty acid synthesis from glucose in rat adipose tissue (Saggerson, 1972). The available data indicate that increasing acetate concentration lead

to increased acetate oxidation to CO_2 whether insulin is present or not. Increasing acetate concentrations led to increased incorporation of acetate into fatty acids only when insulin was present (Baldwin and Smith, 1971). It is surprising that at very high concentrations acetate elicits only negligible changes in plasma insulin levels in ruminants (de Jong, 1982).

Glucose availability for NADPH and glycerol 3-phosphate synthesis may also become rate-limiting for fatty acid synthesis in ruminants. Addition of glucose increased the rate of fatty acid synthesis from acetate in adipose tissue (Ballard *et al.*, 1972). Like the non-ruminants, dietary fat can lead to decreased rates of fatty acid synthesis. Stearic acid was the most potent inhibitor, while palmitic, oleic, linoleic and linolenic acids had similar effects. Growth hormone appears to be an insulin-antagonist. The role of glucagon was uncertain (Harmon, 1992).

The mechanisms by which insulin stimulates fatty acid synthesis in the ruminant are thought to be essentially the same as those in the non-ruminant, namely a decreased rate of lipolysis and an increased glucose uptake leading to increased NADPH and glycerol 3-phosphate production. But this effect of insulin stimulation is generally smaller than that observed in non-ruminants. The reason for this limited response to insulin is not clear. It has been rationalized on the grounds that the rumen ensures a slow influx of nutrients into the animal and buffers it against sudden changes in nutrient availability, but the biochemical mechanism is still obscure (Bauman 1976). It is still not clear how acetate and glucose regulate insulin activity and how these factors regulate lipogenesis in the ruminant.

Of the various lipogenic enzymes, the low activity and properties of ACC have led to suggestions that it is the normal rate-limiting enzyme, although with suboptimal concentrations of acetate, ACC could well become rate-limiting (van der Walt and Linington, 1989). A low activity of ACL in ruminant adipose tissue, as in the fasted rat, may be a result rather than the cause of the restricted utilization of glucose for fatty acid synthesis.

1.5.4 ATP Citrate Lyase in Ruminants

In the absence of glucose, the utilization of acetate for fatty acid synthesis in ruminant adipose tissue is 10 times high than that in rat adipose tissue. This incorporation is slightly increased by addition of glucose and insulin to the incubation medium. ACL and NADP malate dehydrogenase activities were low in ruminant adipose tissue, but acetyl-CoA synthetase activity was higher in the ruminant than in the rat. Acetyl-CoA synthetase allows the ruminant to form acetyl-CoA directly in the cytoplasm of the fat cell, thus bypassing the need for citrate to move acetyl-CoA across the mitochondrial membrane (Hanson and Ballard, 1967).

In ruminants with a functional rumen, glucose incorporation into fatty acids and activities of relevant enzymes can be increased substantially by infusing glucose postruminally or intravenously. The infusion of glucose into lambs dramatically increased glucose utilization for lipogenesis relative to acetate with 9- to 44-fold increases in the activities of ACL and NADP-malate dehydrogenase in both adipose tissue and liver (Ballard *et al.*, 1972; Baldwin *et al.*, 1966).

In young lambs the pattern of carbohydrate metabolism has been shown to resemble that of non-ruminant in many aspects. Glucose is the major source for energy and lipogenesis. Fatty acids synthesis occurs mainly in the liver. The liver ACL activity was about half of that of rat liver under the comparable assay conditions. In addition, significant activity of ACL has been observed in brain, and retinal tissue of the ruminants. Lack of the enzyme activity in the liver may be regarded not to be due to the genetic defects but probably to some regulatory mechanisms related to the changes in the nutritional environments (Muramatsu *et al.*, 1969).

According to Hardwick (1966), extracts of goat liver or mammary gland did not inhibit rat liver ACL activity. Therefore, the low activity of this enzyme in mature ruminants does not seem to be due to the presence of inhibitor but to a repression of enzyme synthesis.

As in the non-ruminants, ACL in ruminants is also subject to nutritional and hormonal control, Robertson (1981) observed a two-fold increase in ACL activity after 48 hours incubation of tissue slices from sheep liver in a medium containing insulin, but its molecular mechanism was not elucidated. Furthermore, the molecular mechanisms for regulation of lipogenic enzymes in the ruminant have not yet been investigated.

1.6 Aim of This Study

In summary, lipogenesis is an important physiological process in energy metabolism. Animals have evolved different regulation patterns to adapt to changes in the environment and food availability. This pattern in ruminants is demonstrated by the fact that acetate rather than glucose is the major lipogenic precursor and the activity of the lipogenic enzymes reflects this feature. This specificity in the ruminant provides an opportunity to explore the regulatory mechanism of lipogenesis. As the enzyme catalysing the first committed step of the conversion of carbohydrate precursors into fatty acids, ACL has the potential to play an important role in lipogenesis. Previous work demonstrated that ACL is regulated at the transcriptional level. The mechanism of this regulation might vary between different species, as indicated by the comparison of the promoter sequences between human and rat ACL genes. It is reasonable to assume that there is a novel regulatory mechanism involved in the expression of ACL in the ruminant, corresponding to its development patterns. Therefore, the study of this enzyme and the regulation of its expression will shed light on the complex regulation of enzymes involved in fatty acid synthesis.

The questions asked in this project and in the future will be: (1) what is the molecular mechanism involved in the regulation of ACL expression during ruminant development? Do young ruminants share similar mechanisms of ACL regulation with non-ruminants? What is the mechanism involved in the repression of ACL expression in adult ruminants? (2) Although glucose/insulin are less important in lipogenesis in ruminants, do they also have a role in the regulation of ACL expression? Do SREBPs also mediate the effects of insulin in regulation of genes encoding lipogenic enzymes ? (3) As the major precursor

for fatty acid synthesis, does acetate have a role in regulation of lipogenesis and ACL expression? (4) In contrast to the low level of ACL expression in adult ruminants, the expression of the enzymes of fatty acid synthesis from cytosolic acetate are present at high levels. The mechanisms involved in the co-ordinated regulation of these two pathways are still to be resolved.

The isolation and characterization of the ACL promoter will be a starting point from which to answer these questions. The alignment of the coding sequences between rat and human ACL showed a high level of identity. It is likely that this identity would extend to ruminants. This information was used to design a probe, which was used to isolate the 5' region of the ACL gene from bovine genomic library, this is the primary aim of this project. Once the gene had been isolated, its promoter was characterized by restriction mapping, sequencing and bioinformatic analysis. The further study of *cis*-elements and *trans*-factors will give useful information to understand the regulation of ACL expression in ruminant.

CHAPTER TWO: MATERIALS AND METHODS

2.1 Materials

Restriction endonucleases and DNA modifying enzymes were obtained from the following companies: Life Technologies Inc., MD, USA; Promega Corporation, WI, USA; Boehringer Mannheim, West Germany; Stratagene, La Jolla, CA, USA; New England Biolabs Inc., MA, USA.

The following reagents and chemicals were purchased from Life Technologies Inc., MD, USA: Proteinase K, 1 kb plus DNA molecular size ladder, *Taq* DNA polymerase, Reaction Buffer and magnesium chloride, guanidinium thiocyanate, cesium chloride, yeast extract, 5' RACE System Kit.

X-gal, IPTG, low EEO type I-A agarose, ficoll, polyvinylpyrrolidone, PEG 8000, SDS, lysozyme, DNase, RNase, ampicillin, ethidium bromide, maltose, bovine serum albumin, and mineral oil were purchased from Sigma Chemical Company, St. Louis, MO, USA.

The λ DASH II bovine genomic libraries were obtained from Stratagene, La Jolla, CA, USA.

The cloning vectors pGEM 3Zf(-), pGEM-T and pGEM-T Easy and the Wizard Genomic DNA Purification System, were purchased from Promega Corporation, WI, USA.

Radioisotopes were purchased from ICN Biomedicals Inc, New England Nuclear Research Products, Boston, MA, USA or Amersham Pharmacia Biotech (UK). Ready-To-Go DNA Labelling Beads and Probe-Quant™ G-50 Micro Columns were purchased from Amersham Pharmacia Biotech (UK).

Colony/Plaque Screen™ Hybridization Transfer Membranes was obtained from NEN™ Life Science Products, Inc., MA, USA. Other hybridisation filters (82 and 134 mm) nitrocellulose and nylon membranes were purchased from Amersham Pharmacia Biotech, UK, Schleicher and Schuell, West Germany.

Oligonucleotides were manufactured by Sigma Chemical Company, St. Louis, MO, USA and Life Technologies Inc., MD, USA.

X-ray film was obtained from Eastman Kodak, NY, USA or Fuji Photo Film Company Ltd., Japan. Photographic developer and fixer was purchased from Eastman Kodak, NY, USA.

QIAEX II Agrose Gel Extraction Kit was purchased from QIAEX Pty Ltd, Australia.

All other chemicals and reagents were of analytical grade or similar quality.

Bacterial strains of *Escherichia coli.* in this study are shown in Table 1.

Table1: Genotype of *Escherichia coli* strains used in this study.

Strain	**Genotype**
XL-1 Blue	*SupE44, hsdR17, recA1,endA1,gyrA96, thi-1,relA1,* *lac*[F' *proAB*$^{+}$,*lacI*^{q}, *lacZΔM15* Tn10 (tet^{r})]
SRB	*RecJ, sbcC201, uvrC, umuC::Tn5 (kan^{r}), mcrA, mcrB, mrr, lac,* *Δ(hsdRMS), endA1, pro, gyrA96, thi-1, relA1, supE44* *{F', proAB, laclq ZΔM15}*

2.2 Methods

All general techniques for DNA manipulation such as ethanol precipitation, cloning of DNA fragments and phenol/chloroform extraction of DNA were carried out according to standard protocols (Ausubel *et al.,* 1989; Sambrook *et al.,* 1989).

2.2.1 Maintenance and Storage of Bacterial Strains and Phage

All bacterial strains and phage stocks were maintained, cultured and stored using standard protocols (Miller, 1987).

2.2.2 Preparation of Plasmid DNA

Small amounts of plasmid DNA for general use were prepared by extraction of DNA from a bacterial culture using the rapid-boil technique according to Sambrook *et al.*, (1989). Plasmid DNA for sequencing was prepared using the MiniPrep Kit from Life Technologies or Bio-Rad.

2.2.3 Preparation of Phage DNA

Phage DNA was prepared using either the plate lysate or liquid lysate methods (Ausubel *et al.,* 1989; Sambrook *et al.,* 1989).

2.2.4 Preparation of Genomic DNA

Genomic DNA from bovine white blood cells, lamb liver tissue and rat liver tissue was extracted with the Wizard Genomic DNA Purification System (Promega) by following the manufacturer's instructions.

2.2.5 DNA Amplification

Specific DNA sequences was amplified by PCR using a standard protocol. Reactions were carried out in 0.5 ml tubes overlaid with mineral oil, and typically contained lx reaction buffer (20 mM Tris-HCI (pH 8.4), 50 mM KCl), 2 mM MgCl, 3 mM deoxynucleotide triphosphates, 250 ng (20 pmol) of each oligonucleotide primer, 1-10 ng of template DNA, and 2.5 U *Thermus aquaticus (Taq)* DNA polymerase. Plasmid DNA prepared by the rapid boil method was diluted 100 fold to reduce the concentration of any

potential inhibitors before the addition of 1-2 µl to the PCR reaction. Amplification was carried out using a thermal cycler program for an initial 4 minute denaturation at 95'C followed by 30 cycles of 95 ^{0}C for 30 seconds, 65-68^{0}C for 45 seconds, and 72^{0}C for 1 minute. Reactions were stored at 4°C prior to analysis.

For the DNAs which were hard to amplify due to possible secondary structure or primer dimer formation, touch-down PCR was performed by using a gradient of annealing temperatures, in which the annealing temperature in the first cycle was about 5 ^{0}C higher than the calculated annealing temperature.

2.2.6 Quantitation of DNA

An estimation of the quantity and purity of the DNA was made by spectrophotometric measurements of diluted DNA samples. DNA was quantified by A_{260}, given that 1.0 A_{260} unit represents 50 µg/ml of double-stranded DNA (Sambrook *et al.* 1989). The purity was estimated from the ratio of the absorbencies at 260 and 280 nm. Alternatively, an aliquot of DNA was analysed by agarose gel electrophoresis alongside quantitation standards (10-100 ng), prepared by digesting a known amount of vector with a restriction endonuclease. Fragments were visualised under UV illumination, and the concentration was determined by comparison to the quantitation standards.

2.2.7 DNA Digestion and Agarose Gel Electrophoresis

All restriction endonuclease digests were carried out with appropriate buffers using conditions recommended by the supplier. Electrophoresis of DNA fragments was performed in low electroendosmosis grade agarose containing ethidium bromide (0.5 mg/ml) and 1 x TAE buffer (0.04 M Tris, 0.02 M acetate, 1 mM EDTA, pH 8.0) according to Sambrook *et al.* (1989). DNA fragment size was confirmed using the 1 kb plus molecular size ladder and/or λ Hind III standard (GIBCO BRL Life Technologies Ltd).

2.2.8 Digestion of Genomic DNA

Digestion of genomic DNA was carried out in sterile 0.5 ml microcentrifuge tubes according to the following protocol.

Genomic DNA	10 µg
Appropriate 10x buffer	10 µl
Sterile water up to	100 µl
Restriction enzyme	20 unit

Reactions were incubated at 37°C overnight. An aliquot (5 µl) was analysed by agarose gel electrophoresis to determine the efficiency of the digestion. Reactions which required further digestion were incubated with additional restriction enzyme (2 unit/µg DNA) for another 2-4 hours. The correct concentration of buffer was maintained in these incubations.

2.2.9 Electrophoresis of Genomic DNA

Prior to electrophoresis, digested genomic DNA was concentrated by ethanol precipitation according to Sambrook *et al.* (1989). Samples were resuspended in a small volume of sterile TE buffer (pH 8.0) and heated to 65^0C for 10 minutes with loading dye before electrophoresis in 0.7% agarose gel in TAE buffer. Electrophoresis was carried out for 16 hours at 22 V or until the dye front was ~4 cm from the end of the gel tray. The DNA fragments were stained with ethidium bromide (0.5 µg/ml), visualised, and photographed under UV illumination.

2.2.10 Purification of Fragments from Agarose Gels

DNA fragments separated by gel electrophoresis were excised under illumination by long wavelength UV light (366 nm), and then extracted from the agarose using the QIAEX II Agrose Gel Extraction Kit.

2.2.11 Preparation of Vectors for Subcloning

Digested vector to be used for subcloning (1-5 µg) was treated with 1µl calf alkaline phosphatase for 10-20 minutes at 37 ^{0}C. The reaction mix was then treated with 2 µl proteinase K (10 mg/ml) and 2 µl 20% SDS for 60 minutes at 37^0C. DNA was extracted with phenol/ chloroform, ethanol precipitated and resuspended in TE buffer (20-100 µl).

2.2.12 Ligation of Vector with Insert DNA

Ligations were carried out with various ratios of vector to insert DNA, by incubation with T4 DNA ligase and ligase reaction buffer at room temperature for 3 hours. When a high efficiency and low background was needed, samples were incubated at 4 ^{0}C for 16 to 22 hours

2.2.13 Transformation of Competent Cells

Competent *E. coli* XL-1 Blue cells suitable for transformation were prepared using calcium chloride (Sambrook *et al.*, 1989), All competent cells were transformed by heat shock according to Sambrook *et al.* (1989). Agar plates used for plating cells after transformation were supplemented with appropriate antibiotics

2.2.14 Labelling DNA Probes with ^{32}P

DNA probes were labelled with ^{32}P-dCTP using Ready-To-Go DNA Labelling Beads (Amersham). The DNA to be labelled is first denatured and then mixed with oligodeoxyribonucleotides of random sequence. These "random oligomers" anneal to random sites on the DNA and then serve as primers for DNA synthesis by a DNA polymerase. With labelled nucleotide(s) present during this synthesis, highly labelled DNA is generated. Radioactively labelled probes were purified using Probe-Quant^{TM} G-50 Micro Columns (Amersham) according to the manufacturers instructions. The specific activity and degree of incorporation of radioactively labelled DNA was determined by scintillation counting in a Beckman LS 8000 scintillation counter.

2.2.15 Southern Transfer

DNA fragments separated by gel electrophorsis were transferred to nitrocellulose filter by capillary transfer according to Sambrook *et al.*(1989). Filters were then dried between two pieces of Whatman paper by baking at a 80^{0}C oven for two hours.

2.2.16 Southern Hybridisation

Prehybtidisation

All colony lifts and Southern blots were treated in the same manner. The membranes were baked in a 80^0C oven for two hours and then prehybridised for 2-3 hours in a rotary oven or a shaking water bath at 68^0C in 10-20 ml of prehybridisation solution (6x SSC containing 5x Denhardts [1% ficoll (Type 400), 1% polyvinylpyrrolidine, 1% bovine serum albumin] and 100ug/ml herring sperm DNA).

Hybridisation

Hybridisation was performed in a rotary oven at 68^0C overnight. Radiolabelled probe was added at a concentration of 9 x 10^6 cpm/ug DNA per ml of hybridisation solution (1M NaCl, 50 mM $NaH_2PO_4 \cdot 2H_20$, 2 mM EDTA, 0.5% SDS, 5x Denhardts [1% ficoll (Type 400), 1% polyvinylpyrrolidine, 1% bovine serum albumin], and 100 ug/ml herring sperm DNA). The labelled DNA probe was boiled for 5 minutes and placed on ice before addition to the prehybridised lifts or blots.

Washing

Hybridisation solution was discarded and the filters washed as follows, unless otherwise stated. Filters were washed twice [6x SSC, 1% SDS] for 1 hour at the hybridisation temperature, followed by two washes for 30 minutes in 6x SSC, 0.5% SDS at the same temperature, and finally a high stringency wash for exactly 30 minutes in 1x SSC which had been prewarmed to the hybridisation temperature.

2.2.17 Autoradiography

All membranes and blots were wrapped in plastic film to prevent drying out during exposure to X-ray film in autoradiography cassettes. Autoradiography was carried out in the presence of intensifying screens at -70'C or at room temperature. Films were developed using Kodak developer.

2.2.18 Screening the Bacteriophage Library

The bovine genomic library was screened as described by the supplier. The library was first titered to determine its plaque forming units (pfu). The host cell was the SRB strain of *E.coli* supplied by Stratagene. The first round of screening was performed with a bacteriophage density of approximately 50,000 pfu per plate (134 mm). This was decreased to approximately 300 pfu for second round screening and further decreased to approximately 100 pfu per plate (82 mm) for third round screening. Dilutions of phage stocks for titering and plating were prepared in SM buffer (0.1 M NaCl, 8 mM $MgSO_4$, 20 mM Tris-HCI pH 7.5, 0.1% gelatin). Duplicate lifts of the plates were transferred to nylon membranes or nitrocellulose filters which were subsequently hybridised with ^{32}P-labelled DNA probe.

2.2.19 DNA Sequencing

All DNA sequencing was done in the DNA Sequencing Facility of IMBS at Massey University. Sequencing reactions were carried out using ABI PRISM BigDye™ Terminator Cycle Sequencing Ready Reaction Kits following the supplied protocol. Results were analysed in an ABI PRISM™ 377 DNA Sequencer (PE bioSystems).

2.2.20 Isolation of Total Cellular RNA

The total cellular RNA from bovine liver was extracted with guanidinium thiocyanate followed by centrifugation in a cesium chloride solution. This method was described by Chirgwin *et al.* (1979). Guanidinium thiocyanate is used to disrupt the cells, and the resulting homogenate is then layered on a cushion of a dense solution of CsCl. During centrifugation, the RNA forms a pellet on the bottom of the tube, while the DNA and protein float above the CsCl solution. The RNA pellet was purified by phenol/chloroform extraction and ethanol precipitation.

2.2.21 5' RACE System for Rapid Amplification of cDNA Ends

5'RACE is a procedure for amplification of nucleic acid sequences from a messenger RNA template between a defined internal site and unknown sequences at the 5' -end of the mRNA. First strand cDNA synthesis was primed using a gene-specific antisense oligonucleotide (GSP1). This permitted cDNA conversion of specific mRNA, and

maximizes the potential for complete extension to the 5' -end of the message. Following cDNA synthesis, TdT (Terminal deoxynucleotidyl transferase) was used to add homopolymeric tails to the 3' ends of the cDNA. Tailed cDNA was then amplified by PCR using a mixture of three primers: a nested gene-specific primer (GSP2), which anneals 3' to GSP1; and a combination of a complementary homopolymer-containing anchor primer and corresponding adapter primer which permits amplification from the homopolymeric tail. This allows amplification of unknown sequences between the GSP2 and the 5'-end of the mRNA. 5' RACE products were then cloned into pGEM-T Easy for subsequent sequencing.

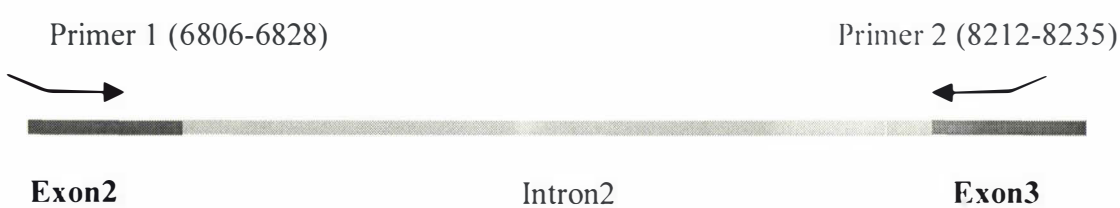
CHAPTER THREE: RESULTS AND DISCUSSION

3.1 Probe Preparation

Comparison of the translated human and rat ACL protein sequences (GI:603073 [GenBank], GI:949989 [GenBank]) revealed 96.3% amino acid identity throughout the entire sequence. It was reasonable to expect that this high level similarity would extend to the bovine protein. The 5'-region sequences of rat and human ACL (GI:4850175 [GenBank], GI:436002 [GenBank]) showed 61% identity. Virtually no similarity was found in exon 1 (untranslated) and in the untranslated regions of exon 2 between human and rat ACL, and this sequence divergence possibly extends to the gene for bovine ACL. To isolate the 5'-region of this gene, the probe should have close sequence identity and be as close as possible to the 5'-region. As the coding sequence of bovine ACL may also span tens of kb in the genome as rat ACL does, the ideal region for the probe would be exon 2. Previous work done in our laboratory had shown that a probe based solely on exon 2 of rat ACL did not work well. We decided to use a longer probe which covered the whole of intron 2 and parts of exon 2 and exon 3 sequences from rat genomic DNA. The length of this probe length was 1.4 kb.

Primers were designed in exon 2 and exon 3, respectively, as shown in Figure 6. With rat genomic DNA (previously prepared in our laboratory) as the template, a 1.4 kb product was obtained by using touch-down PCR. This product was then cloned into the pGEM 3Z f (-) vector. Restriction mapping and partial sequence analysis confirmed that this product contained the anticipated rat ACL genomic sequences (data not shown).

Figure 6: Probe Position on Rat ACL Gene



3.2 Verification of the Probe Specificity

Verification of the probe specificity was done by hybridizing this probe to genomic DNA from a variety of sources. DNA was prepared from bovine white blood cells, lamb and rat liver tissue using the Wizard Genomic DNA Purification System (Promega) or the method from Sambrook *et al* (1989). Both methods of genomic DNA preparation gave DNA with good yield and quality except from rat liver tissue. Only partially degraded DNA could be obtained from the rat DNA preparations with either method. It is very likely that the tissue stored had decayed despite its outward appearance. DNA's from bovine and lamb formed a single, sharp, high molecular weight band on agarose gel, but rat DNA formed a smear of lower molecular weight DNA fragments.

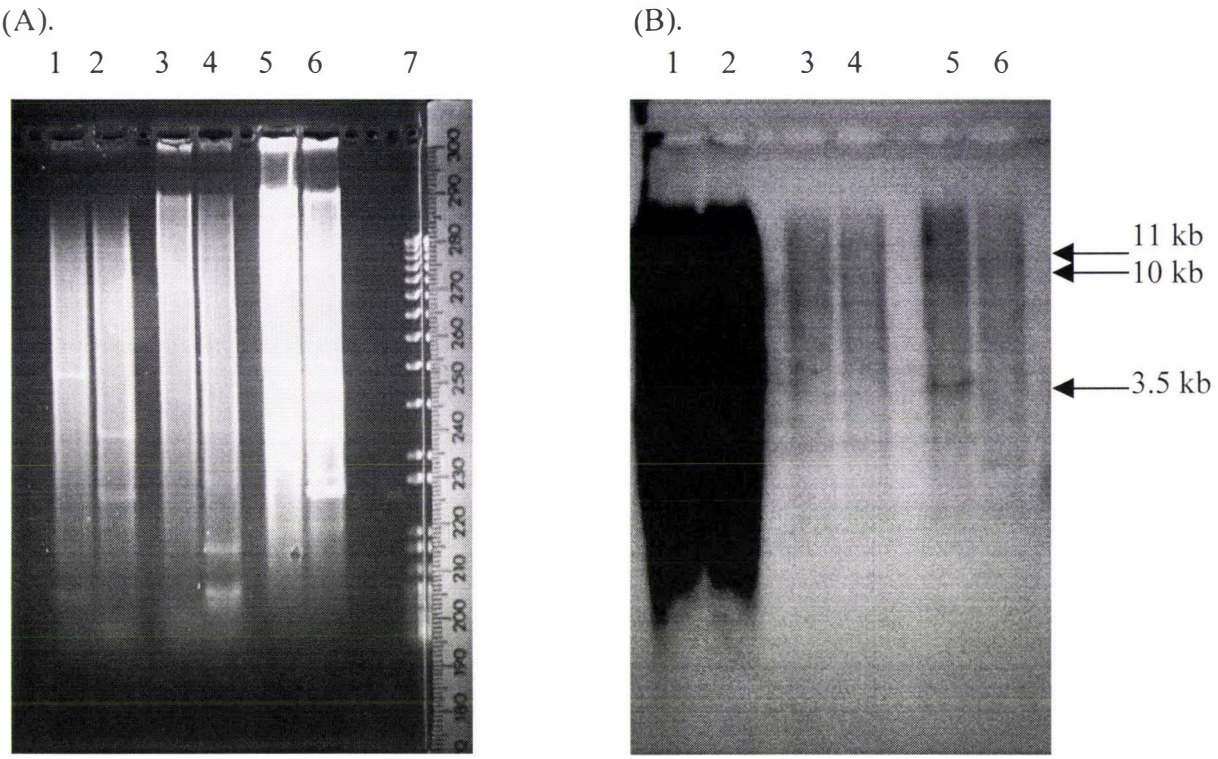
DNA was digested with *EcoR* I and *Hind* III, respectively. The digested DNA was run on a 0.7% agarose gel at 22 volts for 16 hr. Southern blotting and hybridization was performed following standard protocols. The degraded rat DNA showed a smear of hybridizing fragments. Although the background was high, several bands can be seen with bovine and lamb DNA (Figure 7). This result suggested that this rat-based probe hybridises with bovine DNA and could be used for screening the bovine library.

3.3 Screening the Library

A λ DASH II unamplified bovine genomic library (Stratagene) was screened using the verified rat probe. The library was first titered to determine the plaque forming units (pfu). About 50,000 pfu was plated on a 134 mm NZY agarose petri dish using the recommended SRB host strain of *E.coli* (Stratagene). A total of approximately 1.1×10^6 plaques in 22 plates were screened in the first round. Duplicate lifts were taken (Sambrook *et al.*, 1989) and hybridized to the radiolabelled rat probe.

The first round of screening produced a high background combined with signals of varying intensity. Each plate contained from zero to ~10 putative positive plaques. Twelve agarose plugs corresponding to the positive plaques from this screening were

Figure 7: Verification of the Probe Specificity



Genomic DNA's (~10 μg) from rat, lamb and bovine were digested with *EcoR* I and *Hind* III. Six samples were run on a 0.7% agarose gel containing 1x TAE at 22 volts for 16 hr. The gel was stained with ethidium bromide and visualised with an UV transilluminator. DNA was Southern transferred to a nitrocellulose filter and hybridised with the radiolabeled rat probe as described in section 2.2.

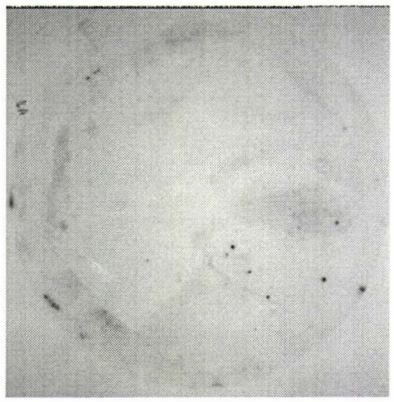
(A) : Agarose gel photo

1. rat DNA digested with *Hind* III
2. rat DNA digested with *EcoR* I
3. lamb DNA digested with *Hind* III
4. lamb DNA digested with *EcoR* I
5. bovine DNA digested with *Hind* III
6. bovine DNA digested with *EcoR* I
7. 1 kb plus DNA ladder

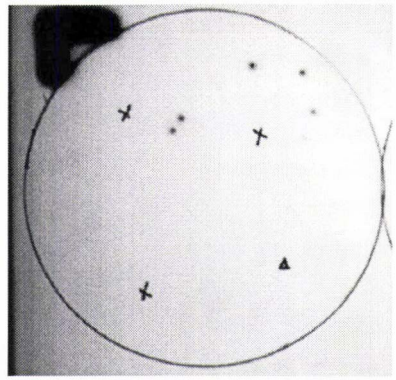
(B): Hybridization photo

1. rat DNA digested with *Hind* III
2. rat DNA digested with *EcoR* I
3. lamb DNA digested with *Hind* III
4. lamb DNA digested with *EcoR* I
5. bovine DNA digested with *Hind* III
6. bovine DNA digested with *EcoR* I

1. First round of screening.



2. Second round of screening



3. Third round of screening

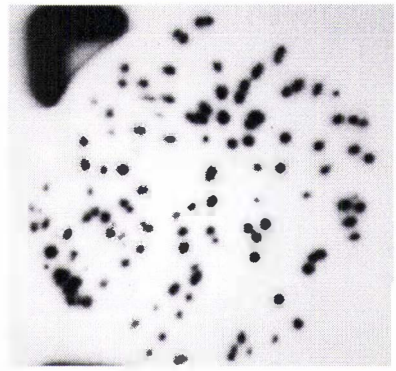


Figure 8: Autoradiographs of hybridisation filters used in the screening of the λ DASH II library

The library was plated and lifts prepared following the supplier's instructions. The filters were hybridised to a radiolabelled rat probe, washed under conditions of high stringency, and autoradiographed. The figure shows autoradiographs of representative filters from first, second and third rounds of screening of λ TW5 clone.

picked and separately plated in a 82 mm petri dish at a density of 300 pfu for the second round of screening. Two plates among them gave ~10 clear positive signals. One agarose plug from each plate was picked from the second round screen and separately plated on a 82 mm petri dish for the third round screening. Every visible plaque (~100 plaques per plate) gave positive hybridysation signal with a similar intensity on third round screening. One of the plaques from each plate was picked and DNA amplified as described in Sambrook *et al* (1989). These two clones were designated as λ TW5 and λ TW6. The hybridised filters from which λ TW5 was selected are shown in Figure 8.

3.4 Characterisation of two λ Clones

3.4.1 Restriction Mapping of two λ Clones

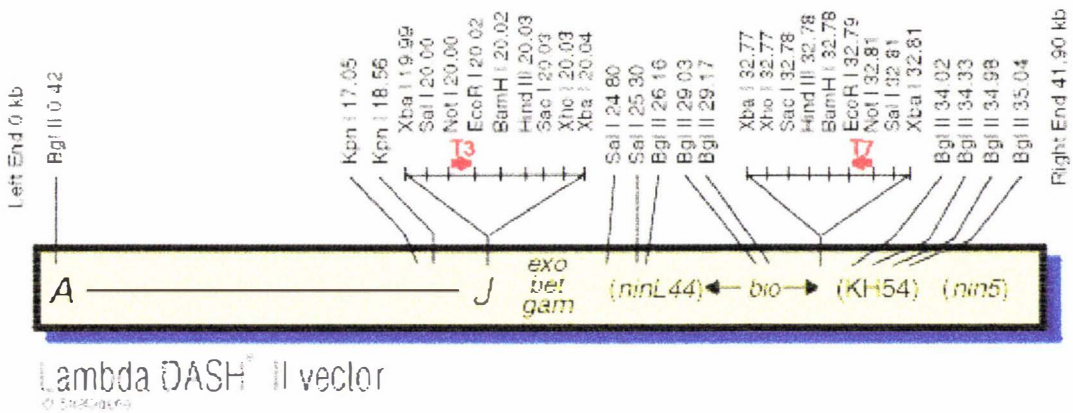
λ TW5 and λTW 6 were characterised by restriction endonuclease mapping. Restriction enzymes, either alone or in various combinations, were used to digest DNA. Following separation by gel electrophoresis and Southern transfer, the fragments were hybridised with the radiolabelled rat probe in an attempt to interpret the complex fragmentation patterns produced by each digestion. The region with which the rat probe hybridised was also assumed to be the corresponding region in the bovine ACL gene (i.e. containing exon2 and/or intron2 and/or exon3 sequences).

A map of the λ DASH II cloning vector is shown in Figure 9. The insert was cloned between the *BamH* I sites. The map shows that *Bgl* II, *EcoR* I, *Kpn* I, *Not* I, *Sal* I and *Xba* I flank the cloning junctions of the insert. These enzymes were used to digest λ TW5 and λ TW6 DNA. The sizes of digestion products were obtained by aligning with the λ Hind III ladder and the BRL 1kb plus DNA ladder (Life Technology). This information is summarised in Tables 2 and 3. The partial restriction maps for both λ TW5 and λ TW6 were then deduced by comparison of each digest and hybridisation patterns.

The λ DASH II vector is 41.90 kb in length. With the insert between the *BamH* I sites, the enzymes *EcoR* I, *Not* I, *Sal* I and *Xba* I have only one site flanking each arm. Digestion with these enzymes will produce a 20 kb left arm, a 9.1 kb right arm and

different sizes of other bands. As showed in Figure 10, λ TW5 digestion gave clearly resolved bands that allowed the construction of a partial restriction map when combined with hybridisation analysis. *Not* I digests gave two bands of 20 kb and 9.1 kb, corresponding to the two arms of the vector, and a third band of ~17 kb which hybridised to the probe and represented the cloned genomic insert. Most of the large fragments in the *EcoR* I, *Kpn* I, *Sal* I and *Xba* I digests can be visualised but some small fragments might be lost because of the amount of DNA loaded. These small fragments were resolved when subclones were mapped. *Bgl* II digestion yielded too many small bands to analyse.

Figure 9: λ DASH II Vector Map



The λ vector map shows two *Kpn* I sites in the left arm, at 17.05 and 18.56 kb. There is no site in the right arm. *Kpn* I digestion gave bands of 21.30, 17, 6.4 and 1.5 kb. The probe hybridised to the 21.30 kb band which therefore contains 9.1 kb of right arm and ~11 kb of insert. The 6.4 kb fragment probably represents the rest of the insert up to the *Kpn* I site at 18.56 kb in the left arm. *EcoR* I digestion gave bands of 20.10, 9.1 kb (vector arms) and two bands of 5.8 kb and 10 kb, of which the 10 kb band hybridised to the probe. Combined with the results of the *Kpn* I digestion, this suggests that the 10 kb fragment is next to the right arm of the vector. A double digest of *EcoR* I/*Kpn* I confirmed this deduction as the 21.30 kb *Kpn* I band disappeared and a 10 kb band appeared which hybridised to the probe. *Sal* I digestion yielded two potential insert bands with similar sizes, 8.6 kb and 8.4 kb respectively. The 8.6 kb *Sal* I fragment hybridised to

the probe, suggesting that this fragment was next to the λ right arm by reference to the *EcoR* I digest. *EcoR* I/*Sal* I double digests confirmed this deduction. *Xba* I digestion yielded three insert bands, 3.4, 3.5 and 10.2 kb. The latter two fragments hybridised to the probe. The 3.5 kb fragment was assumed to be next to the λ right arm. If the 10.2 kb fragment was next to the λ right arm, the 3.5 kb fragment should not hybridise to the probe by reference to the hybridisation patterns of *EcoR* I and *Sal* I. This deduction was later confirmed by subclone mapping. The results of the restriction and hybridisation analysis of λ TW5 are shown in Figure 10 and the analysis is summarised in Table 2 and Figure 11.

The λ TW6 clone was analysed in the same way. This clone also appeared to have an ~17 kb insert (Figure 12). However an unambiguous restriction map was not obtained for this clone because *Kpn* I and *Xba* I failed to digest to completion. This was probably be due to poor λ DNA quality. *Not* I and *Sal* I gave only one band in addition to the two vector arms. However one important deduction could be made from this digestion and hybridisation. In the λ TW5 *Sal* I digest the hybridising band is next to the right arm of the vector. Since λ TW6 appears to have no *Sal* I site apart from those in the vector adjacent to the cloning site, the insert in λ TW6 probably extends further to the right direction of λ TW5. This result means that, irrespective of whether the 5'-region sequence are to the right or left of the region complementary to the probe, the two clones together, containing ~10 kb flanking each side of the hybridising region (exon 2, intron 2, exon 3), should be long enough to contain at least 5 kb of DNA upstream of exon 1. This is based on the assumption that the gene structure is conserved between human, rat and bovine. The length of intron 1 in rat is only 4.3 kb.

Further characterisation was limited to λ TW5 because of restriction digestion problems with λ TW 6. Fragments from λ TW5 were subcloned into pGEM 3Z f (-) (Promega) to facilitate further analysis. The fragment initially chosen for subcloning was the 10 kb fragment from *EcoR* I digestion. This was the largest hybridising fragment, apart from those resulting from *Not* I (17 kb) and *Kpn* I digestion (21.30 kb) which were too large for subcloning into pGEM 3Z f (-).

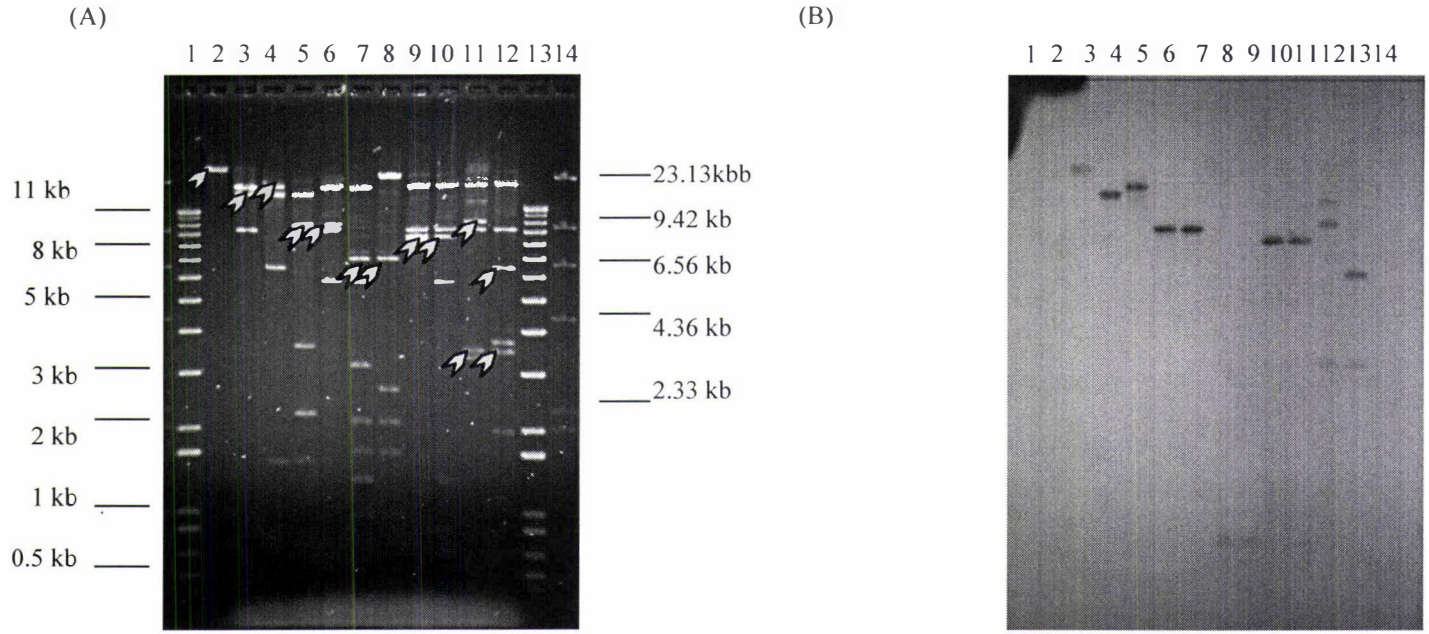


Figure 10: Agarose gel electrophoresis of digested λ TW5 DNA and autoradiograph of the agarose gel after Southern blotting and hybridisation to the rat ACL probe.

(A). λ TW5 DNA was digested with the restriction endonucleases listed blow for one hour at 37 ^{0}C. The samples were separated by gel electrophoresis througy a 0.7% agarose gel in 1xTAE and photographed under UV illumination. Arrowheads indicate the fragments which hybridised to the probe (B).

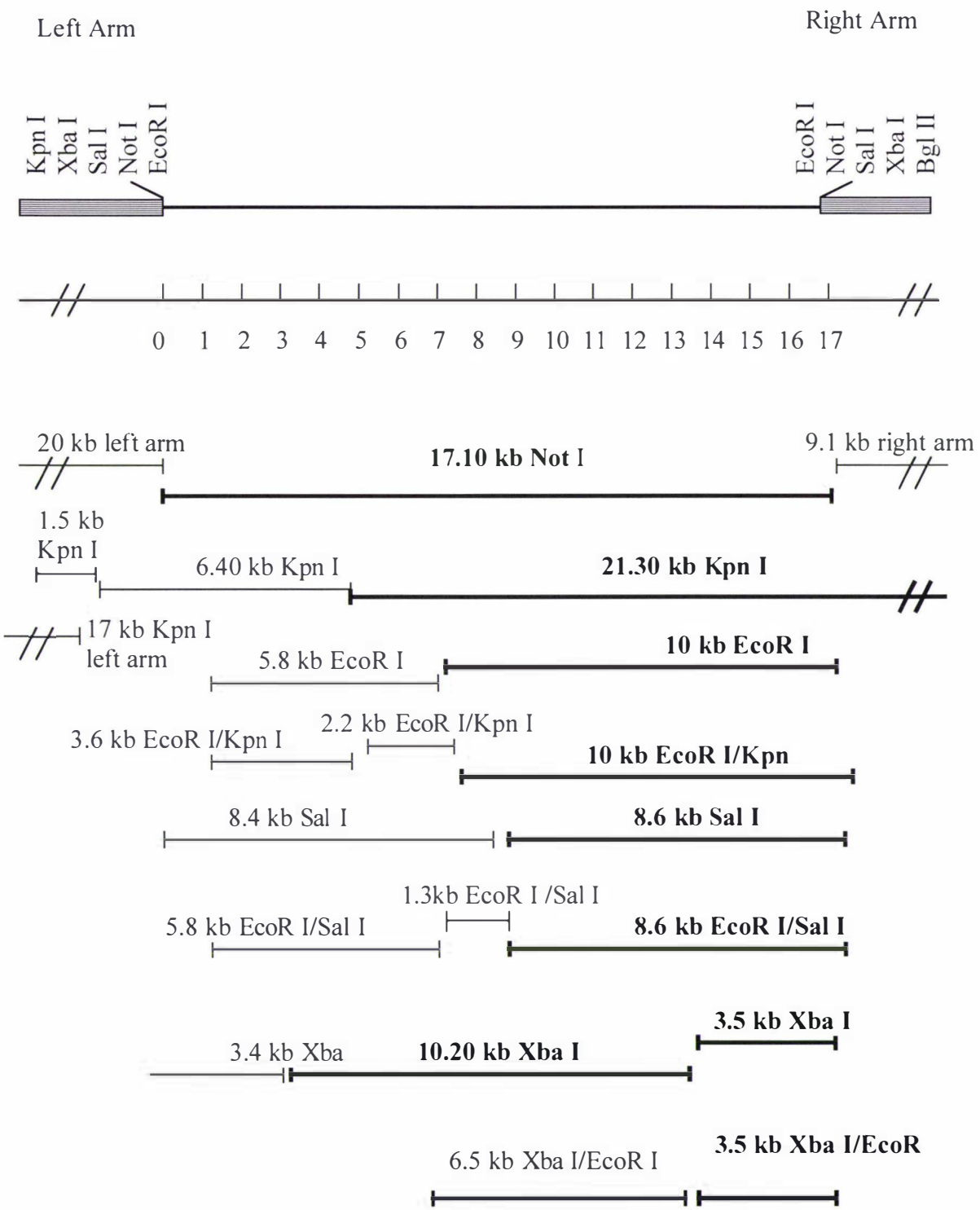
(B) DNA from the gel shown in (A) was transferred to nitrocellulose filter by Southern transfer. The dried filter was hybridised to the rat ACL probe as described in section 2.2.

(1). BRL 1 kb plus DNA ladder; (2). Uncut λ TW5 DNA; λ TW5 DNA was digested with: (3). *Not* I; (4). *Kpn* I; (5). *Kpn* I/*EcoR* I; (6). *EcoR* I. (7); *EcoR* I/*Bgl* II; (8).*Bgl* II; (9). *Sal* I; (10). *Sal* I/*EcoR* I; (11). *Xba* I; (12). *Xba* I/*EcoR* I; (13). BRL 1 kb plus DNA ladder; (14). BRL λ Hind III markers.

Not I (kb)	*Kpn* I (kb)	*EcoR* I/ *Kpn* I (kb)	*EcoR* I (kb)	*EcoR* I/ *Bgl* II (kb)	*Bgl* II (kb)	*Sal* I (kb)	*EcoR* I/ *Sal* I (kb)	*Xba* I (kb)	*EcoR* I/ *Xba* I (kb)
20.10	**21.30**	17.00	20.10	20.60	>25	20.10	20.10	20.10	20.60
17.10	17.00	**10.00**	**10.00**	**7.00**	**7.00**	9.10	9.10	**10.20**	9.20
9.10	6.40	9.10	9.10	5.80	2.70	**8.60**	**8.60**	9.10	**6.50**
	1.50	3.60	5.80	3.10	2.10	8.40	5.80	**3.50**	3.70
		2.20		2.10	1.60		1.30	3.40	**3.50**
		1.50		1.60	1.00				2.00
		0.70		1.00	0.70				0.70
				0.70					
46.30	46.20	44.10	45.60	41.60	39.10	46.20	44.90	46.30	46.20

Table 2: Sizes of fragment resulting from restriction endonuclease digestion shown in Figure 10. Hybridising fragments are shown in bold and underlined.

Figure 11: Restriction map of the λ TW5 clone deduced from the results shown in Figure 10 and Table 2. The fragments which hybridise to the rat ACL probe are shown in bold.



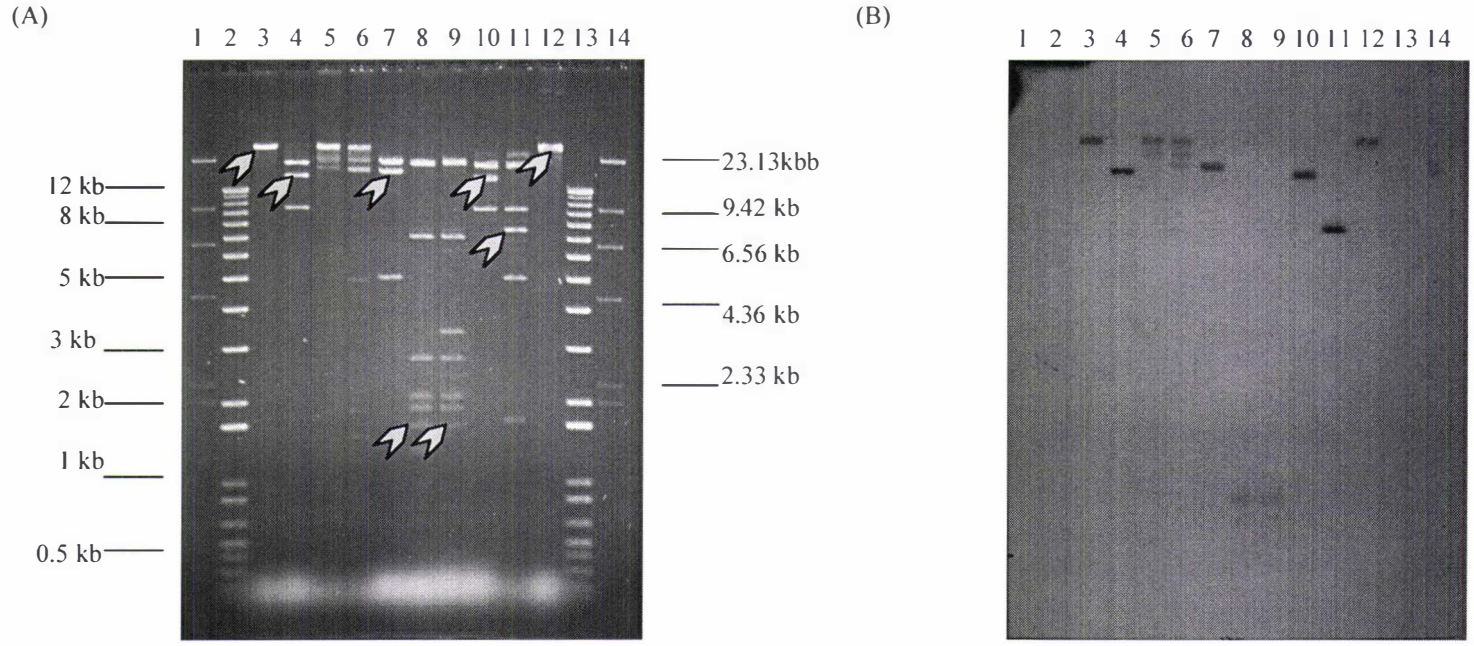


Figure 12: Agarose gel electrophoresis of digested λ TW6 DNA and autoradiograph of the agarose gel after Southern blotting and hybridisation to the rat ACL probe.

(A). λ TW 6 DNA was digested with the restriction endonucleases listed blow for one hour at 37 ^{0}C. The samples were separated by gel electrophoresis througy a 0.7% agarose gel in 1xTAE and photographed under UV illumination. Arrowheads indicate the fragments which hybridised to the probe (B).

(B) DNA from the gel shown in (A) was transferred to nitrocellulose filter by Southern transfer. The dried filter was hybridised to the rat ACL probe as described in section 2.2.

(1).and (14): BRL λ Hind III markers; (2). and (13): BRL 1 kb plus DNA ladder; (3). Uncut λ TW5 DNA; λ TW5 DNA was digested with: (4). *Not* I; (5). *Kpn* I; (6). *Kpn* I/*EcoR* I; (7). *EcoR* I. (8); *EcoR* I/*Bgl* II; (9).*Bgl* II; (10.) *Sal* I; (11.) *Sal* I/*EcoR* I; (12). *Xba* I.

3.4.2 Characterisation of the 10 kb EcoR I Subclone

Subcloning was performed as described in section 2.2. The DNA from the λ TW5 clone and from pGEM 3Z f(-) was digested with *EcoR* I, and the pGEM 3Z f(-) DNA was treated with calf alkaline phosphatase. Ligation was carried out at 4 ^{0}C overnight using T4 DNA ligase. Aliquots of the ligation reaction were used to transform the XL-1 blue strain of *E.coli*. Twelve colonies were picked from appropriate selective plates and overnight cultures prepared. Plasmid was extracted by the rapid-boiling method. Plasmid DNA was digested with *EcoR* I and the resulting bands visualised after agarose gel electrophoresis. Of the 12 colonies, 4 contained a 10 kb insert, 6 contained a 5.8 kb insert and 2 appeared to have no insert. One of the clones with the 10 kb insert was selected for further analysis. Maps of pGEM 3Z f(-) and the 10 kb *EcoR* I construct are shown in Figure 13.

To facilitate further subcloning, only restriction enzymes which have a site in the Multiple Cloning Site in pGEM3Z f(-) were selected for digestion of the 10 kb *EcoR* I Subclone. Restriction enzymes, either alone or in various combinations, were used to digest the DNA. Following separation by agarose gel electrophoresis (two gels) and Southern transfer, the fragments were hybridised with the radiolabelled rat probe. The results are shown in Figure 14 and Figure 16. To distinguish some ambiguous restriction patterns, further restriction digests were done as shown in Figure 18.

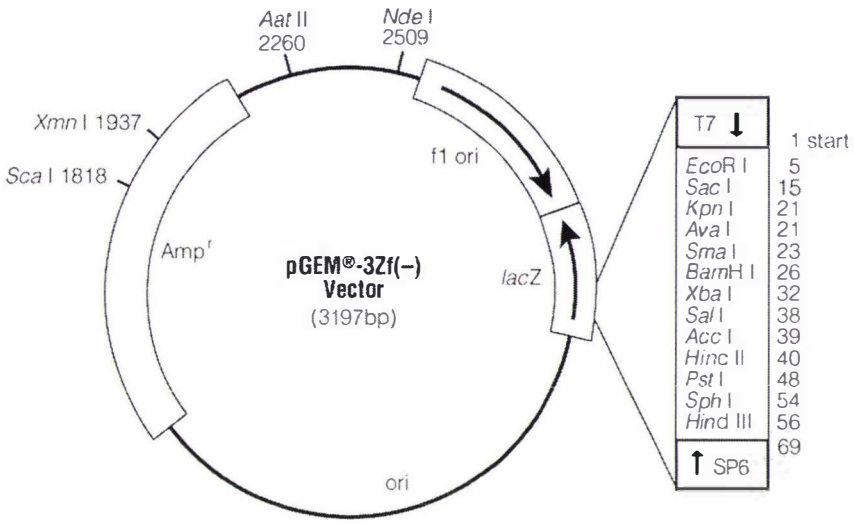
Sal I digestion produced a 4.5 kb and a 8.7 kb fragment. When *Sal* I digestion was combined with *EcoR* I, the 4.5 kb fragment was replaced by a 1.3 kb and a 3.2 kb fragment which was the size of the vector. This suggested that the 1.3 kb fragment was located at the T7 side and the 8.7 kb fragment was located at the Sp6 side. In the same way, it can be deduced from *Xba* I and *Xba* I/*EcoR* I digestion that the 9.6 kb *Xba* I fragment (containing the vector) was at the T7 side and the 3.6 kb *Xba* I fragment was adjacent to the Sp6 side. As it is already known from the λ TW5 digestion map that the 3.6 kb *Xba* I fragment and the 8.7 kb *Sal* I fragment were in the right side of λ TW5, it

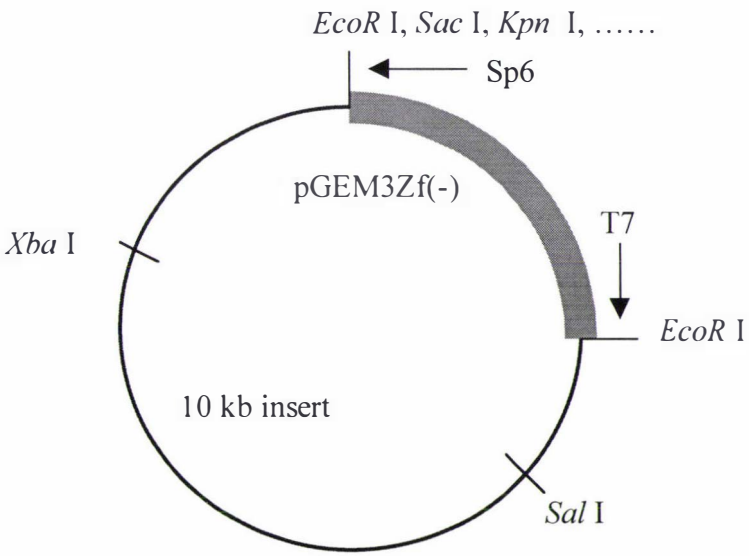
can be concluded that the orientation of the 10 kb insert from T7 to Sp6 corresponded with the left-to-right direction of λ TW5.

BamH I digestion yielded four fragments of 4.9 kb, 4.2 kb, 2.0 kb and 1.9 kb. No smaller band was observed. When combined with *EcoR* I digestion, the 4.2 kb fragment was replaced by 1.0 kb and 3.2 kb fragments. This suggests that the 4.2 kb fragment contained the vector and 1.0 kb of insert. In *BamH* I/*Sal* I digestion, the 4.9 kb fragment was replaced by 0.3 kb and 4.6 kb fragments. This suggested that the 4.9 kb fragment was on the right of the 1.0 kb fragment, as the *Sal* I site was 1.3 kb away from the left end of the insert. In *BamH* I/*Xba* I digestion, the 1.9 kb fragment was replaced by 1.4 kb and 0.5 kb fragments. Since the *Xba* I site was 3.6 kb away from the right end of the insert, this 1.9 kb fragment was then on the right side of the 4.9 kb fragment. Therefore, the 2.0 kb *BamH* I fragment was on the right end of the insert. Therefore, all of the four *BamH* I fragments have been mapped. Because of their relatively average size, each of the four *BamH* I fragments was subcloned into pGEM 3Z f(-) for further analysis.

The restriction analysis of the *BamH* I fragments (not shown here) helped the further determination of the restriction maps using other enzymes. Since the size of fragments from the *Sac* I digestion of the 10 kb *EcoR* I clone was convenient for sequencing, all of these fragments were subcloned into pGEM 3Z f(-). Due to the presence of many small fragments, the maps of *Acc* I, *Ava* I, *Pst* I, and *Sma* I were only partly defined. *Hinc* II and *Sph* I digestions were not investigated further.

Figure 13: Graphic maps of the pGEM 3Zf(-) vector and the 10 kb EcoR I subclone construct





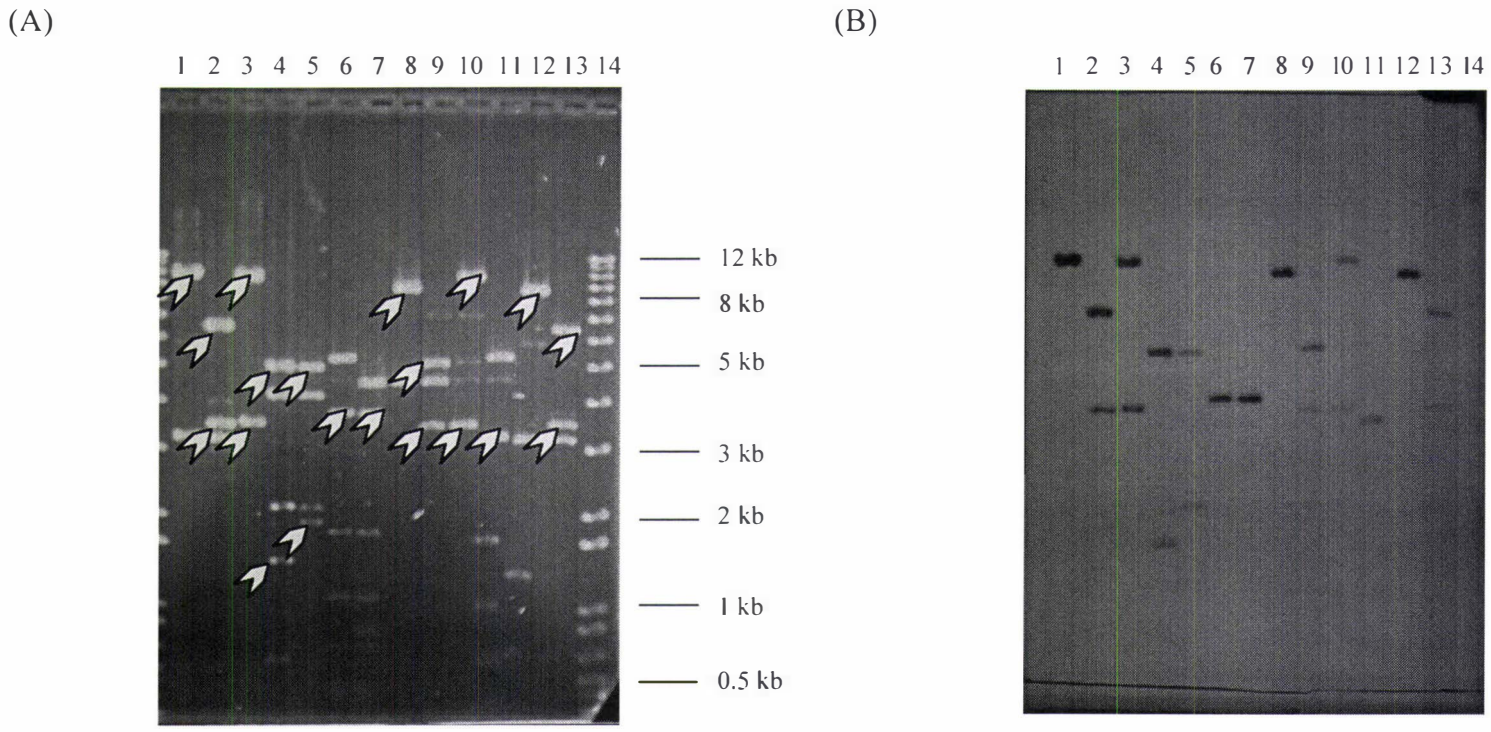


Figure 14: First group of restriction digests and Southern blot of the 10 kb *EcoR* I subclone.

(A). Plasmid DNA was digested with the restriction endonucleases listed below for one hour at 37 ^{0}C. The samples were separated by gel electrophoresis through a 0.7% agarose gel in 1xTAE and photographed under UV illumination. Arrowheads indicate the fragments which hybridised to the probe.

(B) Digested samples were subjected to electrophoresis(A) and Southern transferred to nitrocellulose filter, which was hybridised overnight at 68 ^{0}C and washed under conditions of high stringency. The resultant blot was wrapped in Saranwrap and exposed to film.

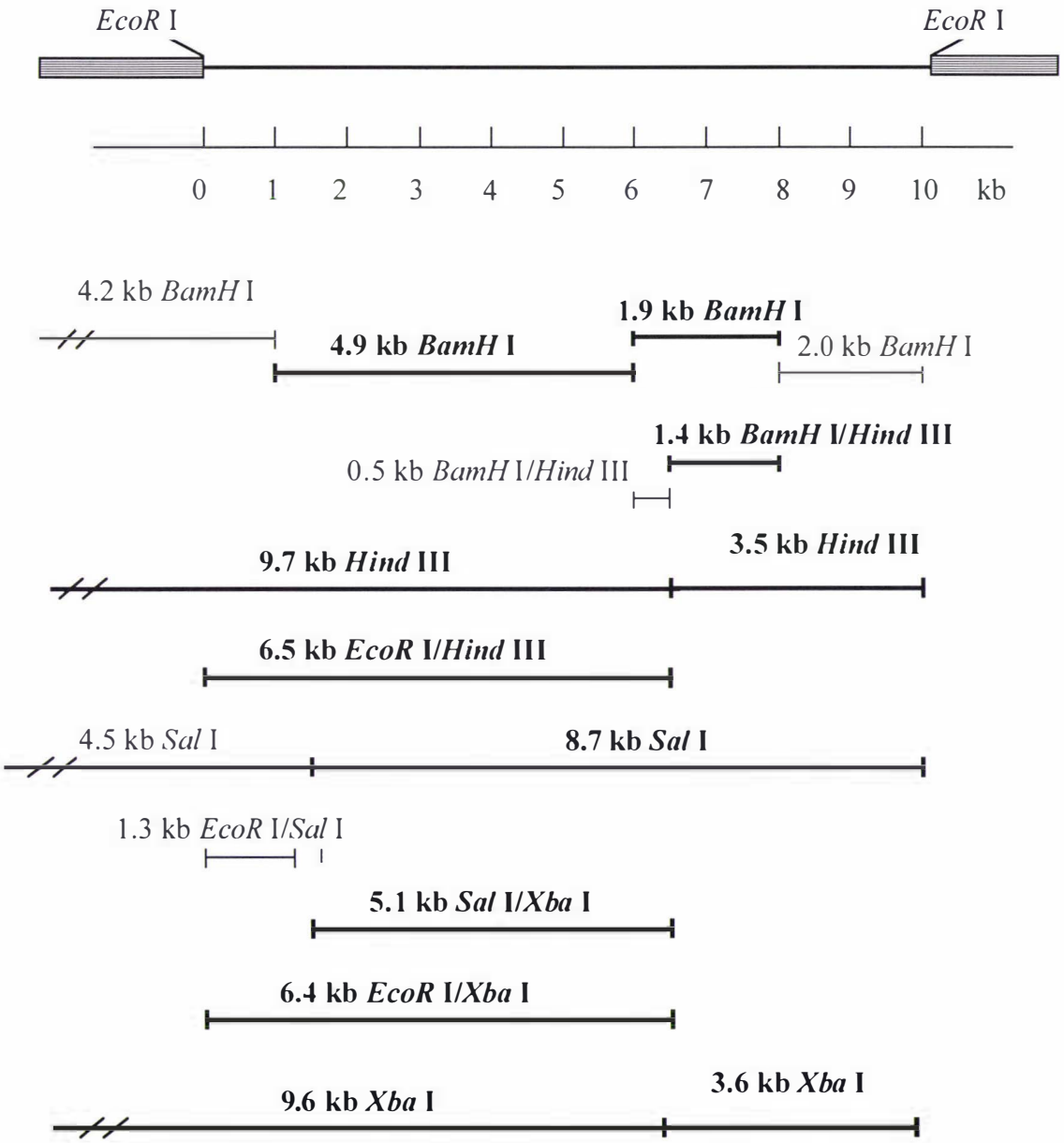
Plasmid DNA was digested with: (1). *EcoR* I; (2). *EcoR* I/*Hind* III; (3). *Hind* III; (4). *Hind* III/*BamH* I; (5). *BamH* I; (6). *Pst* I; (7). *Pst* I/*Sal* I; (8). *Sal* I; (9). *Sal* I/*Xba* I; (10). *Xba* I; (11). *Xba* I/*Pst* I; (12). *EcoR* I/*Sal* I; (13). *EcoR* I/*Xba* I; (14). BRL 1 kb plus ladder.

EcoR I (kb)	*EcoR* I/ *Hind* III (kb)	*Hind* III (kb)	*Hind* III/ *BamH* I (kb)	*BamH* I (kb)	*Pst* I (kb)	*Pst* I/ *Sal* I (kb)	*Sal* I (kb)	*Sal* I/ *Xba* I (kb)	*Xba* I (kb)	*Xba* I/ *Pst* I (kb)	*EcoR* I/ *Sal* I (kb)	*EcoR* I/ *Xba* I (kb)
10.00	**6.50**	**9.70**	**4.90**	**4.90**	5.20	4.50	**8.70**	**5.10**	**9.60**	5.20	**8.70**	**6.40**
3.20	**3.50**	**3.50**	4.20	4.20	**3.75**	**3.75**	4.50	4.50	**3.60**	**3.50**	3.20	**3.60**
	3.20		2.00	2.00	1.80	1.80		3.60		1.80	1.30	3.20
			1.40	**1.90**	1.10	1.10				1.10		
			0.50		0.50	0.70				0.70		
					0.30	0.50						
						0.30						
13.20	13.20	13.20	13.00	13.00	12.65	12.65	13.20	13.20	13.20	12.20	13.20	13.20

Table 3: Sizes of fragment resulting from restriction endonuclease digestion shown in Figure 14. Hybridising fragments are shown in bold and underlined.

Figure 15: Restriction map of the 10 kb *EcoR* I subclone deduced from the results shown in Figure 14 and Table 3. The fragments which hybridise to the rat ACL probe are shown in bold.

The orientation of the insert is arranged to correspond with that in λ clone. The multiple cloning sites are at right side behind *EcoR* I site. Hybridising bands are shown in bold.



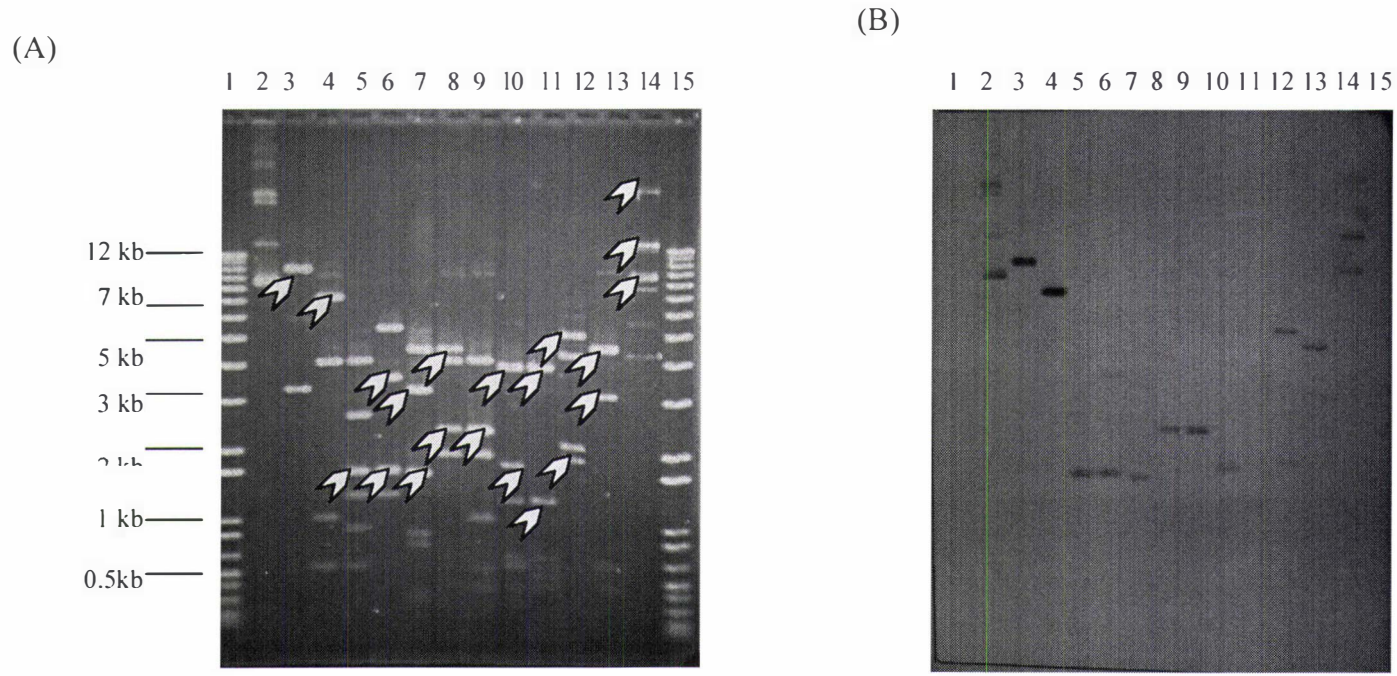


Figure 16: Second group of restriction digests and Southern blot of the 10 kb *EcoR* I subclone

(A). Plasmid DNA was digested with the restriction endonucleases listed below for one hour at 37 ^{0}C. The samples were separated by gel electrophoresis through a 0.7% agarose gel in 1xTAE and photographed under UV illumination. Arrowheads indicate the fragments which hybridised to the probe.

(B) DNA from the gel shown in (A) was transferred to nitrocellulose filter by Southern transfer. The dried filter was hybridised to the rat ACL probe as described in section 2.2.

(1) and (15), .BRL 1 kb plus DNA ladder; (2). Uncut DNA; plasmid DNA was digested with: (3). *EcoR* I; (4). *Acc* I; (5). *Acc* I/*Sac* I; (6). *Sac* I; (7). *Sac* I/*Sma* I; (8). *Sma* I; (9). *Sma* I/*Acc* I; (10). *Ava* I; (11). *Ava* I/*BamH* I; (12). *BamH* I; (13). *Hinc* II; (14). *Sph* I.

EcoR I (kb)	*Acc* I (kb)	*Acc* I/ *Sac* I (kb)	*Sac* I (kb)	*Sac* I/ *Sma* I (kb)	*Sma* I (kb)	*Sma* I/ *Acc* I (kb)	*Ava* I (kb)	*Ava* I/ *BamH* I (kb)	*BamH* I (kb)	*Hinc* II (kb)	*Sph* I (kb)
10.00	**7.00**	4.00	5.20	4.40	**4.40**	3.90	3.80	3.80	**4.90**	4.70	incomplete digestion
3.20	4.00	2.40	**3.40**	**3.20**	3.90	**2.30**	**3.70**	**3.70**	4.20	**4.60**	
	1.00	1.70	1.70	1.75	**2.40**	2.20	**1.90**	**1.30**	2.00	**3.20**	
	0.60	**1.65**	**1.65**	**1.65**	2.00	1.90	1.30	0.70	**1.90**	0.65	
		1.25	1.25	0.90		1.10	0.70	0.65			
		0.80		0.80		0.70	0.65				
		0.60				0.50					
13.20	12.60	12.40	13.20	12.60	12.70	12.60	12.05	10.15	13.00	13.15	

Table 4: Sizes of fragment resulting from restriction endonuclease digestion shown in Figure 16. Hybridising fragments are shown in bold and underlined.

Figure 17: Restriction map of the 10 kb *EcoR* I subclone deduced from the results shown in Figure 12 and Table 4. The fragments which hybridise to the rat ACL probe are shown in bold.

The orientation of the insert is arranged to correspond with that in λ clone. The multiple cloning sites are at right side behind *EcoR* I site. Hybridising bands are shown in bold.

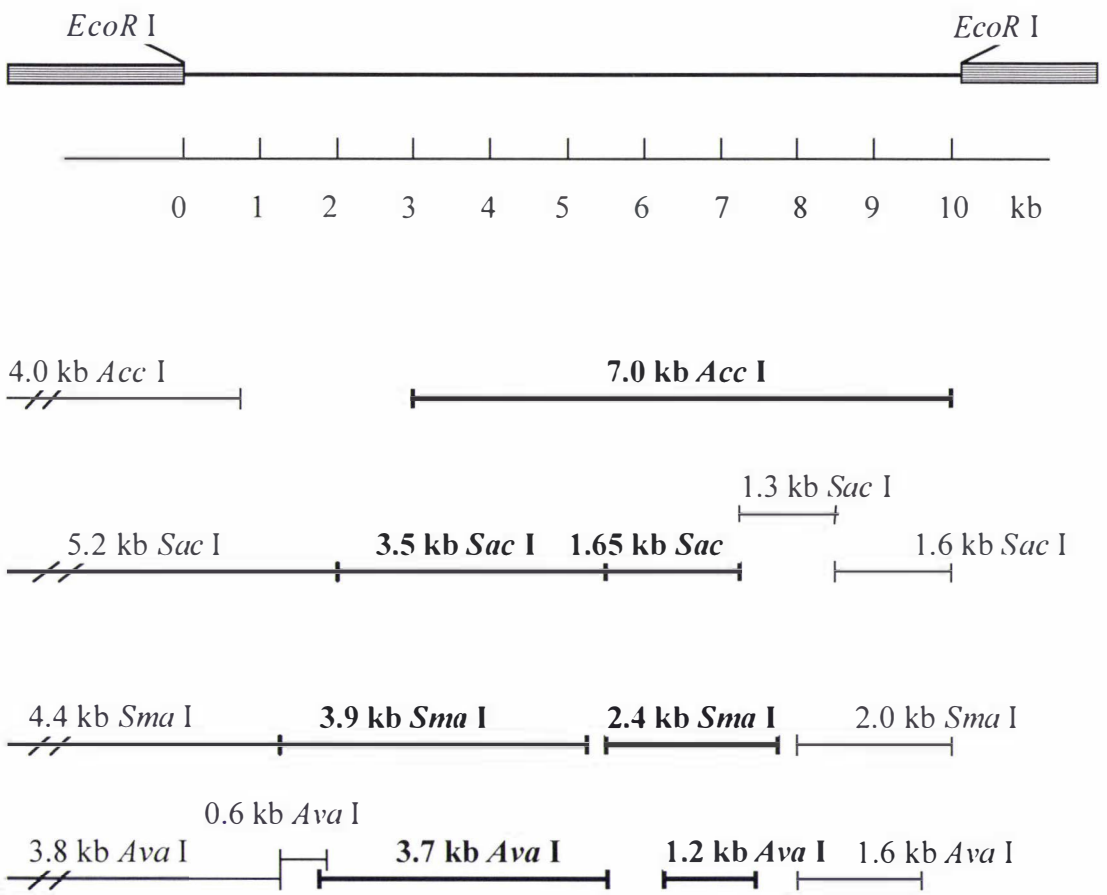
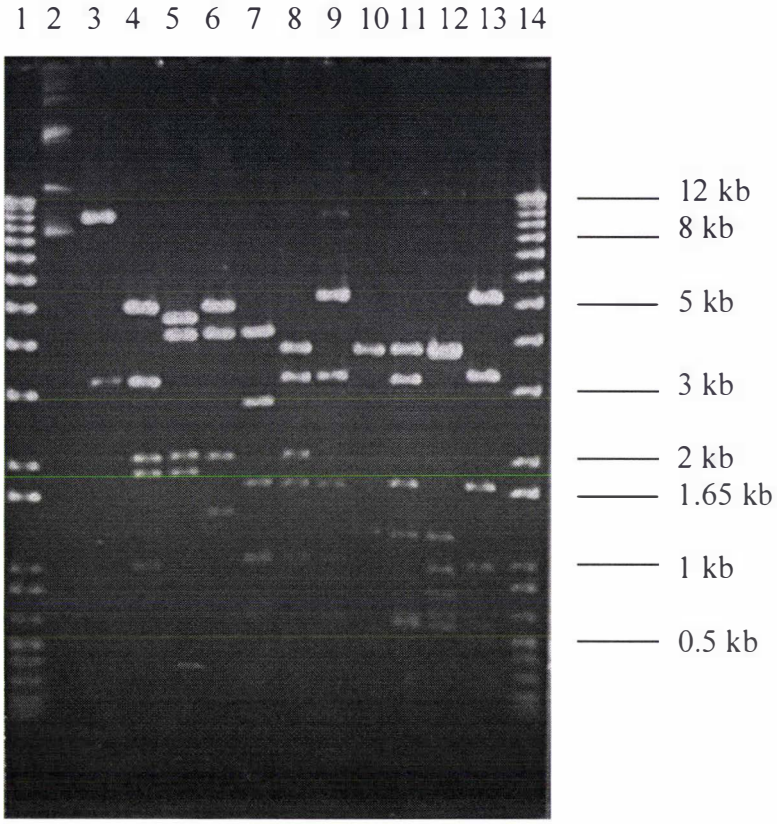


Figure 18: Third group of restriction digests of the 10 kb *EcoR* I subclone



Plasmid DNA was digested with the restriction endonucleases listed below for one hour at 37 ^{0}C. The samples were separated by gel electrophoresis through a 0.7% agarose gel in 1xTAE and photographed under UV illumination.

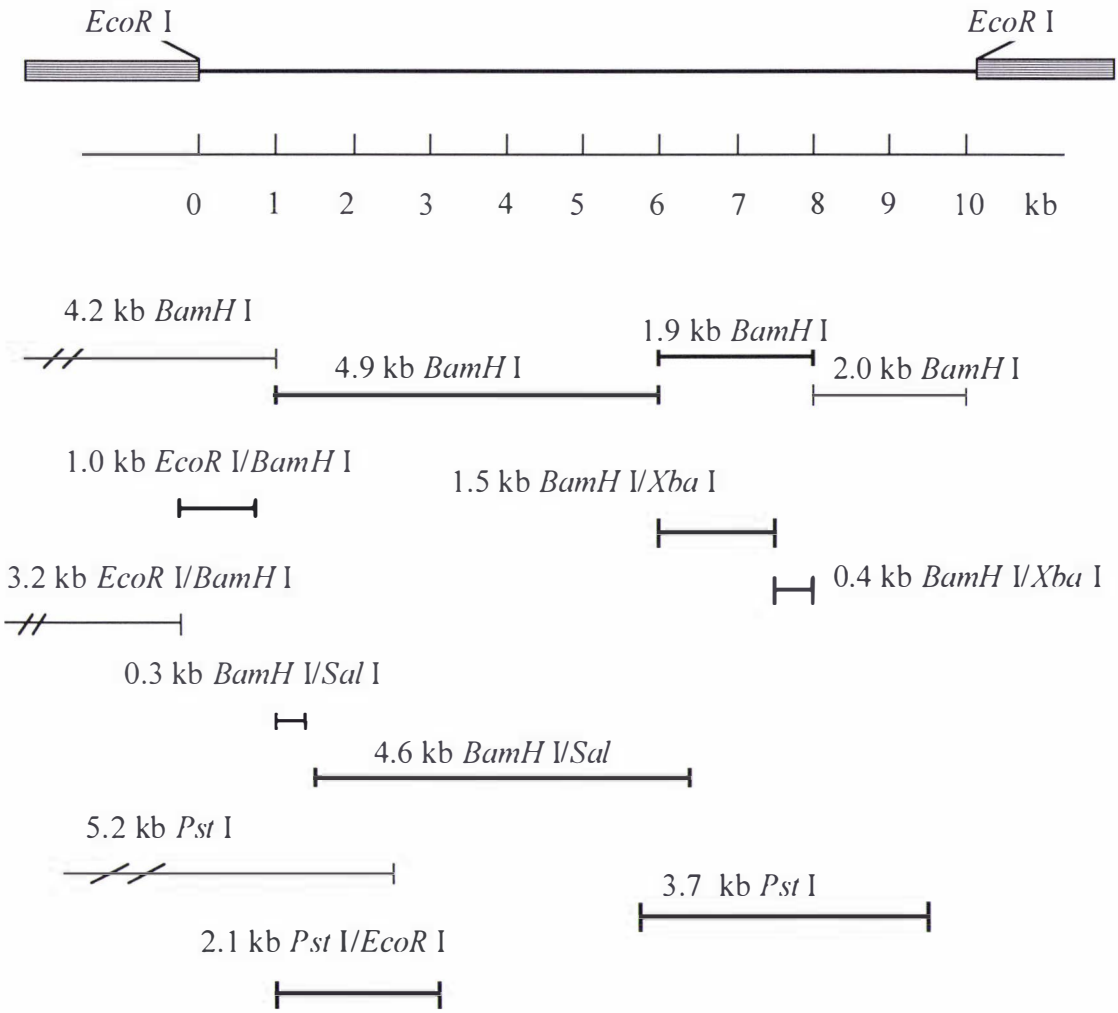
1. 1 kb plus DNA ladder
2. uncut DNA
3. *EcoR* I digest
4. *EcoR* I/*BamH* I digest
5. *BamH* I/*Sal* I digest
6. *BamH* I/*Xba* I digest
7. *BamH* I/*Pst* I digest
8. *Pst* I/*EcoR* I digest
9. *Pst* I/*Xba* I digest
10. *Ava* I/*BamH* I digest
11. *Ava* I/*EcoR* I digest
12. *Ava* I/*Hind* III digest
13. *Hind* III/*Pst* I digest
14. 1 kb plus DNA ladder

EcoR I (kb)	*EcoR* I/ *BamH* I (kb)	*BamH* I/ *Sal* I (kb)	*BamH* I/ *Xba* I (kb)	*BamH* I/ Pst I (kb)	*Pst* I/ *EcoR* I (kb)	*Pst* I/ *Xba* I (kb)	*Ava* I/ *BamH* I (kb)	*Ava* I/ *EcoR* I (kb)	*Ava* I/ *Hind* III (kb)	*Hind* III/ *Pst* I (kb)
10.00	4.90	4.60	4.90	4.30	3.90	5.50	3.80	3.80	3.80	5.20
3.20	3.20	4.20	4.20	2.90	3.40	3.20	3.70	3.20	3.70	3.20
	2.10	2.10	2.10	1.80	2.10	1.80	1.20	1.70	1.30	1.70
	1.90	1.90	1.50	1.20	1.80	1.20	0.60	1.30	1.00	1.00
	1.00	0.30	0.40	1.10	1.20	0.60	0.50	0.60	0.80	0.60
				0.70	0.60	0.50	0.40	0.50	0.65	0.50
				0.60	0.20	0.20	0.30	0.30	0.60	0.20
				0.30				0.20	0.30	
				0.20					0.20	
13.20	13.10	13.10	13.10	13.10	13.20	13.00	10.50	11.60	12.35	12.40

Table 5: Sizes of fragment resulting from restriction endonuclease digestion shown in Figure 18.

Figure 19: Restriction map of the 10 kb EcoR I subclone deduced from the results shown in Figure 18 and Table 5.

The orientation of the insert is arranged to correspond with that in the λ clone. The multiple cloning sites are at the right side behind the *EcoR* I site.



3.4.3 Characterisation of the 8.4 kb *Sal* I Subclone

Sequencing of the 10 kb *EcoR* I subclone revealed that this fragment contained the sequence from part of intron 3 through exon 1 and ~800 bp of promoter region. To obtain additional promoter sequence, it was necessary to select another subclone.

The 8.4 kb *Sal* I fragment of λ TW5 was subcloned in pGEM 3Z f (-) and mapped without Southern hybridisation (Figure 21, 22). This fragment was next to the left arm of λ TW5 and has a 1.3 kb region that overlaps with the 10 kb *EcoR* I fragment as shown in Figure 10 and 11. *EcoR* I digestion yielded four fragments of 0.7 kb, 1.3 kb, 3.8 kb and 5.8 kb. The 1.3 kb fragment was certainly the overlapping region with the 10 kb *EcoR* I fragment and on the T7 side of pGEM 3Z f (-). This suggests that the orientation of the 8.4 *Sal* I fragment from Sp6 to T7 correspond to the left-to-right direction of λ TW5. *Sal* I/*EcoR* I digestion cuts the 3.8 kb *EcoR* I fragment to give a 0.6 kb fragment and a 3.2 kb fragment which is pGEM 3Z f (-). This suggests that the 0.6 kb fragment was on the Sp6 side of this *Sal* I clone and therefore on the left end of the λ TW5 insert. *Xba* I digestion gave 5.0 kb and 6.60 kb fragments. *Xba* I/*Sal* I digestion produced a 3.4 kb and a 3.2 kb fragment (vector) which replaced the 6.60 kb *Xba* I fragment, suggesting that the 6.60 kb fragment was on the left end the λ TW5 insert. *EcoR* I/*Xba* I digestion gave 3.8 kb, 3.7 kb, 2.1 kb, 1.3 kb and 0.7 kb fragments. Combining this with *EcoR* I and *Xba* I digestion, the results suggested that *Xba* I cut the 5.8 kb *EcoR* I fragment to give 3.7 kb and 2.1 kb fragments, *EcoR* I cut the 5.0 kb *Xba* I to a 3.7 kb fragment and a 1.3 kb fragment which was on the T7 end of this *Sal* I subclone. Therefore the 5.8 kb *EcoR* I fragment was on the left of 1.3 kb fragment, 0.7 kb *EcoR* I fragment was then left the 5.8 kb fragment, then 0.6 kb left end (Sp6 side) of the λ TW5 insert. The two small *EcoR* I fragments were not observed in λ TW5 mapping (Figure 10), this is possibly due to the amount of DNA loaded.

BamH I digestion produced three fragments of 0.3 kb, 3.6 kb and 7.7 kb. According to the map of the 10 kb *EcoR* I clone, there was a *BamH* I site just 0.3 kb away from the right end of this *Sal* I insert. This places the 0.3 kb *BamH* I on the T7 side. *BamH* I/*Xba* I digestion gave a 6.6 kb and a 1.1 kb fragment which replaced the 7.7 kb *BamH* I

fragment. This suggested that the 7.7 kb fragment was on the Sp6 side (left end of the λ clone), and the 3.6 kb fragment was between the 0.3 kb and the 7.7 kb fragments.

Digestion with Sac I and *SacI/EcoR* I did not give enough information which confirmed the *Sac* I map of the 8.4 kb *Sal* subclone. This was also true of the *Pst* I map with the digestions of *Pst* I and *Pst* I/*EcoR* I. From the *Pst* I digestion, it is clear that the 5.8 kb fragment was the only one bigger than the 3.2 kb vector and thus contained the vector. This fragment must be on the T7 side and the *Pst* I site was then 2.6 kb left away from the T7 end producing a 1.3 kb *Pst* I/*EcoR* I fragment which was subcloned for sequencing.

The 5.8 kb *EcoR* I fragment from λ TW5, which was within the 8.4 kb *Sal* I subclone and next to the 10 kb *EcoR* I fragment, was also cloned and mapped (data not show). Pst I restriction sites were finalised. A 1.1 kb *Pst* I fragment adjacent to the 1.3 kb *Pst* I/*EcoR* I was also cloned and sequenced.

Figure 20: The 8.4 kb *Sal* I subclone.

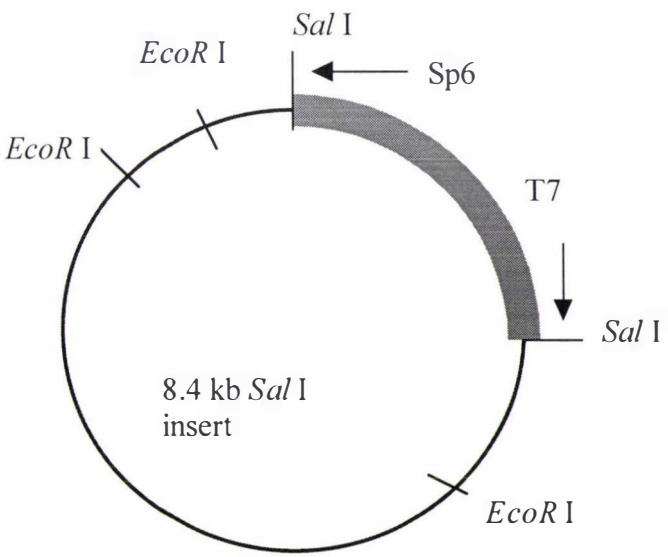
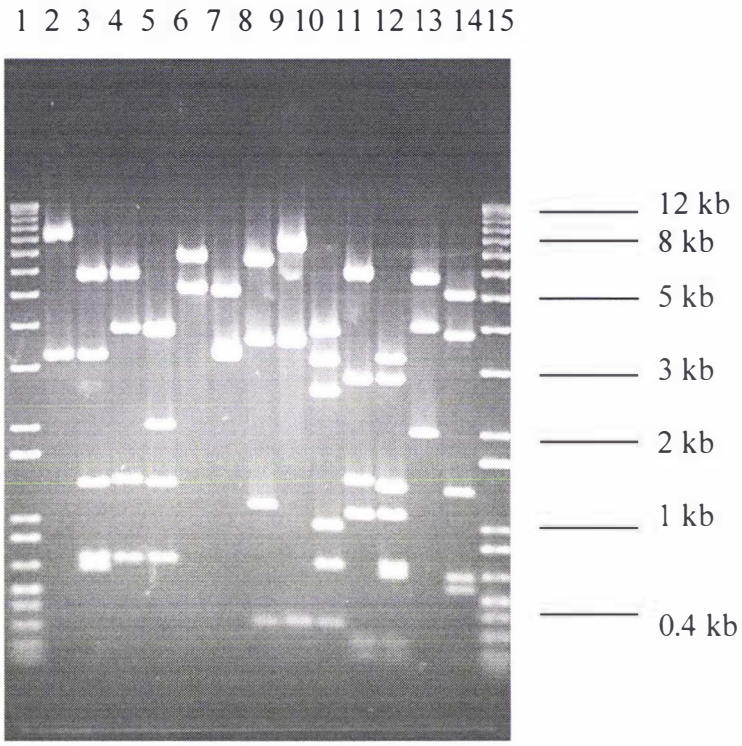


Figure 21: Restriction digests of the 8.4 kb *Sal* I subclone



Plasmid DNA was digested with the restriction endonucleases listed below for one hour at 37 ^{0}C. The samples were separated by gel electrophoresis through a 0.7% agarose gel in 1xTAE and photographed under UV illumination.

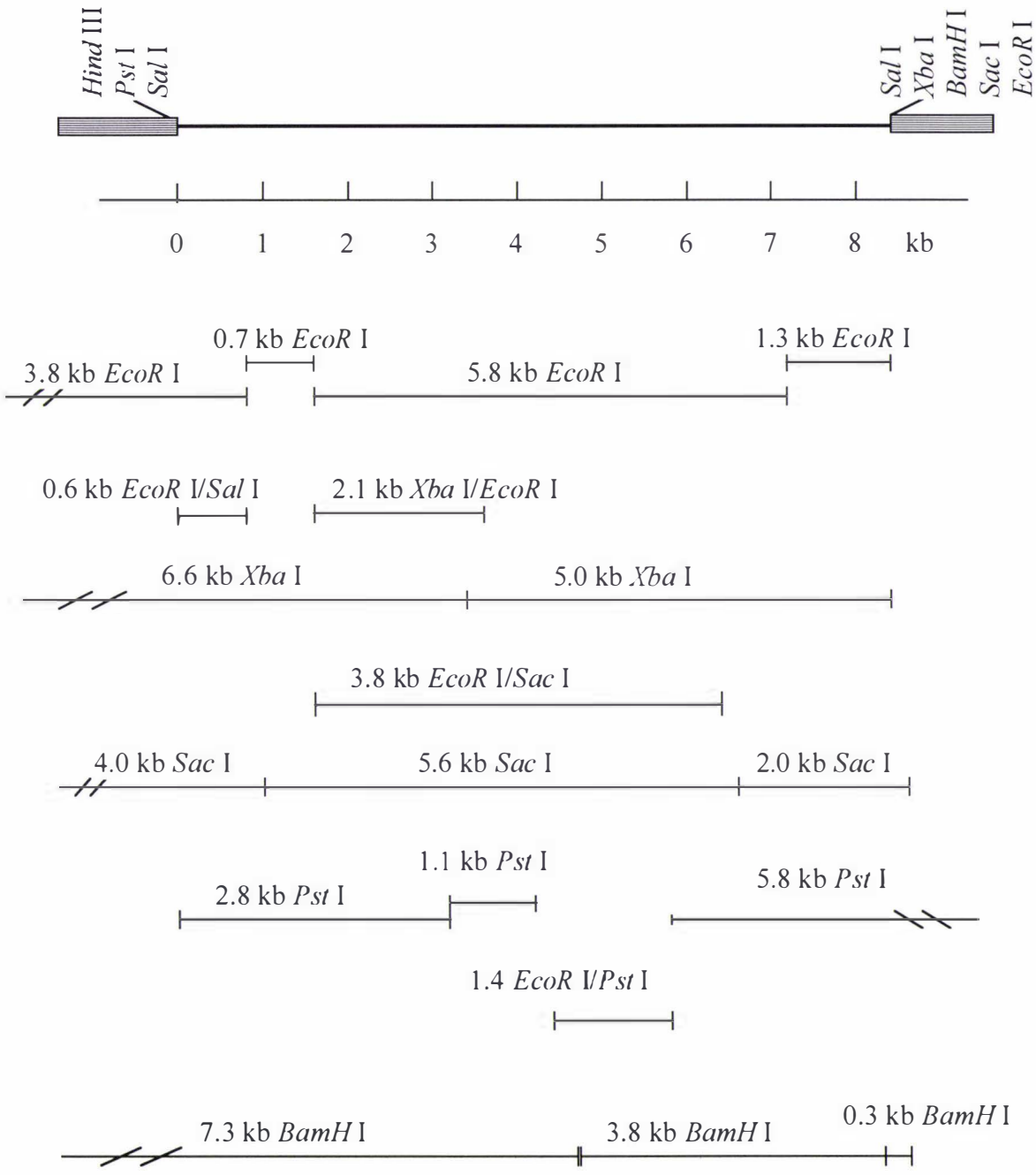
1. 1 kb plus DNA ladder
2. *Sal* I digest
3. *Sal* I/*EcoR* I digest
4. *EcoR* I digest
5. *EcoR* I/*Xba* I digest
6. *Xba* I digest
7. *Xba* I/*Sal* I digest
8. *Xba* I/*BamH* I digest
9. *BamH* I digest
10. *BamH* I/*EcoR* I digest
11. *Pst* I digest
12. *Pst* I/*EcoR* I digest
13. *Sac* I digest
14. *Sac* I/*EcoR* I digest
15. 1 kb plus DNA ladder

Sal I (kb)	*Sal* I/ *EcoR* I (kb)	*EcoR* I (kb)	*EcoR* I/ *Xba* I (kb)	*Xba* I (kb)	*Xba* I/ *Sal* I (kb)	*Xba* I/ *BamH* I (kb)	*BamH* I (kb)	*BamH* I/ *EcoR* I (kb)	*Pst* I (kb)	*Pst* I/ *EcoR* I (kb)	*Sac* I (kb)	*Sac* I/ *EcoR* I (kb)
8.40	5.80	5.80	3.80	6.60	5.00	6.60	7.70	3.80	5.80	3.20	5.60	5.15
3.20	3.20	3.80	3.70	5.00	3.40	3.60	3.60	3.20	2.80	2.80	4.00	3.80
	1.30	1.30	2.10		3.20	1.10	0.30	2.60	1.50	1.40	2.00	1.30
	0.70	0.70	1.30			0.30		1.00	1.10	1.30		0.65
	0.60		0.70					0.70	0.30	1.10		0.60
								0.30	0.20	0.70		
										0.60		
										0.30		
										0.20		
11.60	11.60	11.60	11.60	11.60	11.60	11.60	11.60	11.60	11.70	11.60	11.60	11.50

Table 6: Sizes of fragment resulting from restriction endonuclease digestion shown in Figure 21.

Figure 22: Restriction map of the 8.4 kb *Sal* I subclone deduced from the results shown in Figure 21 and Table 6.

The orientation of the insert is arranged to correspond with that in λ clone.



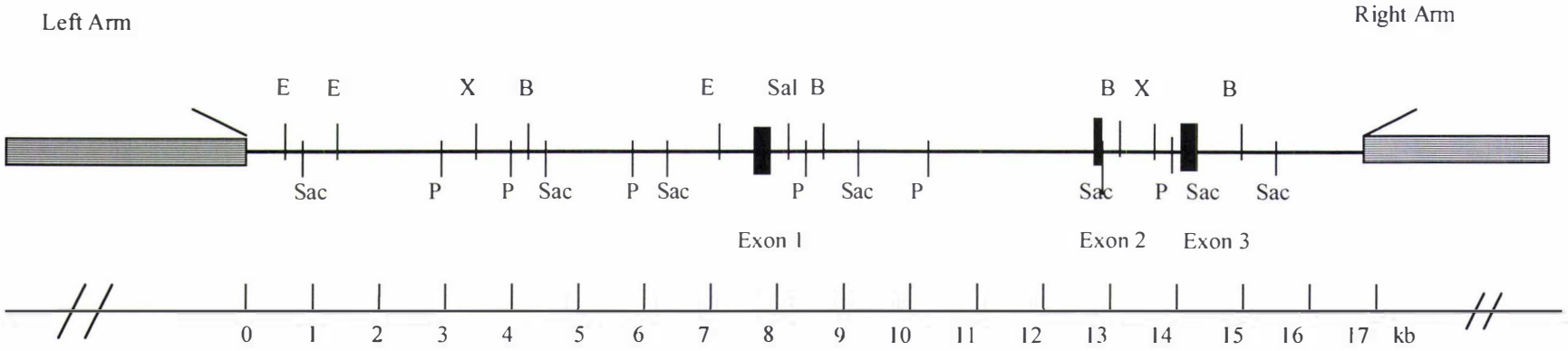
3.5 Sequencing

All sequencing was done in the DNA Sequencing Facility of IMBS at Massey University. Sequencing reactions were carried out using ABI PRISM BigDye^{TM} Terminator Cycle Sequencing Ready Reaction Kits. Results were analysed using an ABI PRISM^{TM} 377 DNA Sequencer (PE BioSystems). A series of subclones was constructed in pGEM 3Zf(-). Each subclone was sequenced from two directions with T7 polymerase primer and Sp6 polymerase primer. The oligonucleotide primers subsequently used were designed from the sequences previously obtained. Sequencing of a subclone was carried out till sequences overlapping with the next subclone were obtained. The sequences of all subclones were assembled manually. The sequencing strategy and the total length sequenced is shown in Figure 24.

Initially, the 1.65 kb *Sac* I fragment from the 10 kb *EcoR* I subclone was sequenced. A BLAST search of the NCBI database showed that this fragment contained the sequences homologous with those of exon 2 and exon 3 from the rat and human ACL genes. The sequence information also indicated that the orientation of λ TW5 from left to right was in the 5' to 3' direction of the sense strand of this gene. The other *Sac* I fragments upstream from the 1.65 kb fragment, i.e the 3.5 kb *Sac* I fragment and the 5.2 kb *Sac* I fragment, were then subcloned into pGEM 3Zf(-) and sequenced. Further BLAST searches showed that the 5.2 kb *Sac* I fragment, which was on the left end of the 10 kb *EcoR* I subclone, contained the putative exon 1 and about 0.8 kb of the promoter region.

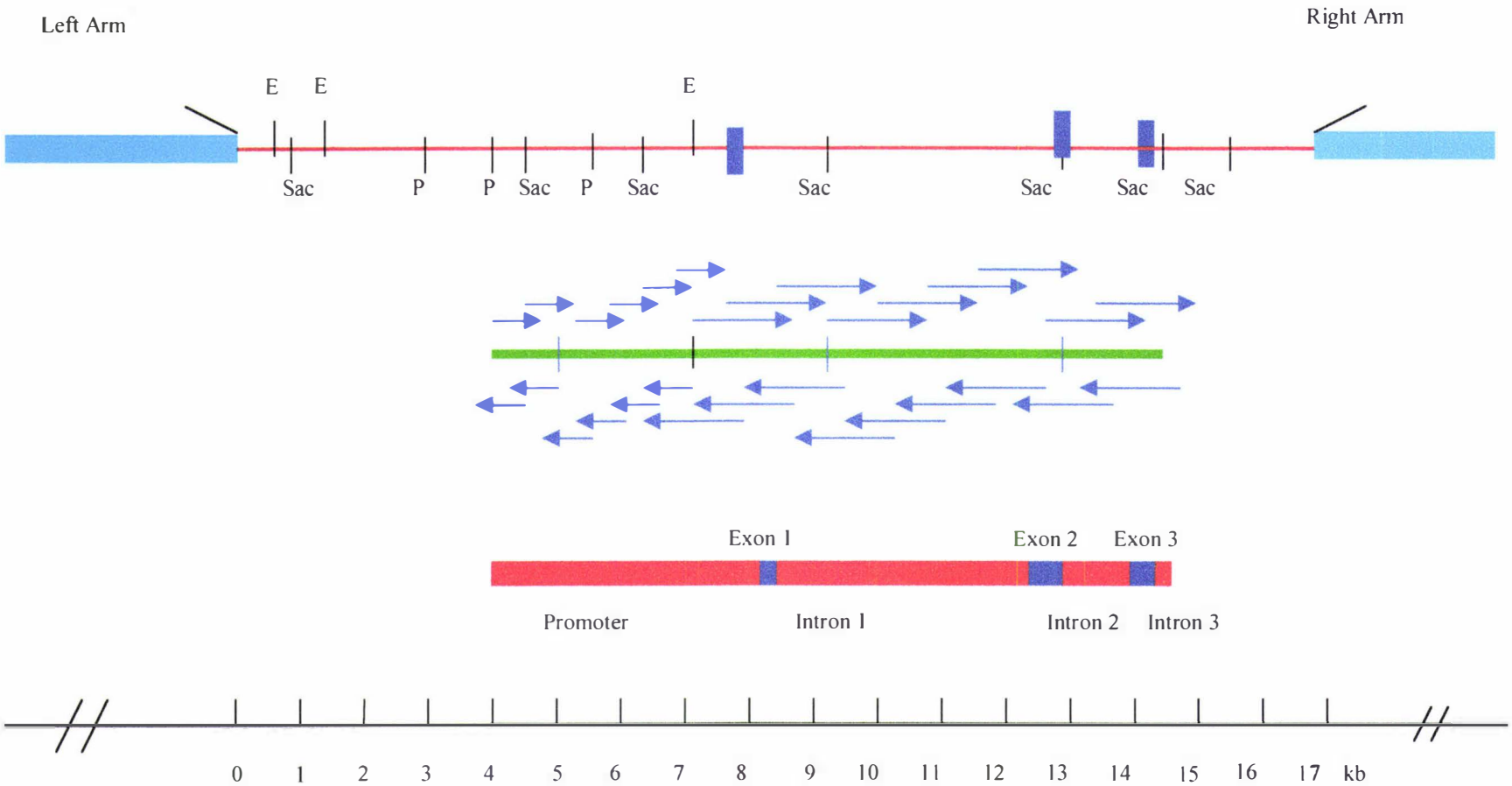
To understand fully the details of how expression of the ACL gene is regulated, a longer promoter sequence was needed. Within the 8.4 kb *Sal* I subclone, an 1.1 kb *Pst* I fragment and an 1.3 kb *EcoR* I/*Pst* I fragment, which was in conjunction with the left end of the 10 kb *EcoR* I subclone, as shown in Figure 20 and 21, were also cloned and sequenced. Thus, about 3.2 kb of DNA sequence upstream from exon 1 was obtained. In all a total of 9.5 kb was sequenced, containing 3.2 kb of the promoter region, and extending through intron 1 to the beginning of intron 3 as shown in Figure 24.

Figure 23: Summary restriction map of λ TW5



E=*EcoR* I X=*Xba* I B=*BamH* I Sal=*Sal* I Sac=*Sac* I P=*Pst* I

Figure 24: Map of sequenced fragments and sequencing strategy



3.6 Determination of the Transcription Start Point

The transcription start site for a gene is determined by identifying the 5' end of the encoded mRNA. It is generally assumed that the sequence at the 5' end of an mRNA corresponds to the DNA sequence at which transcription initiates.

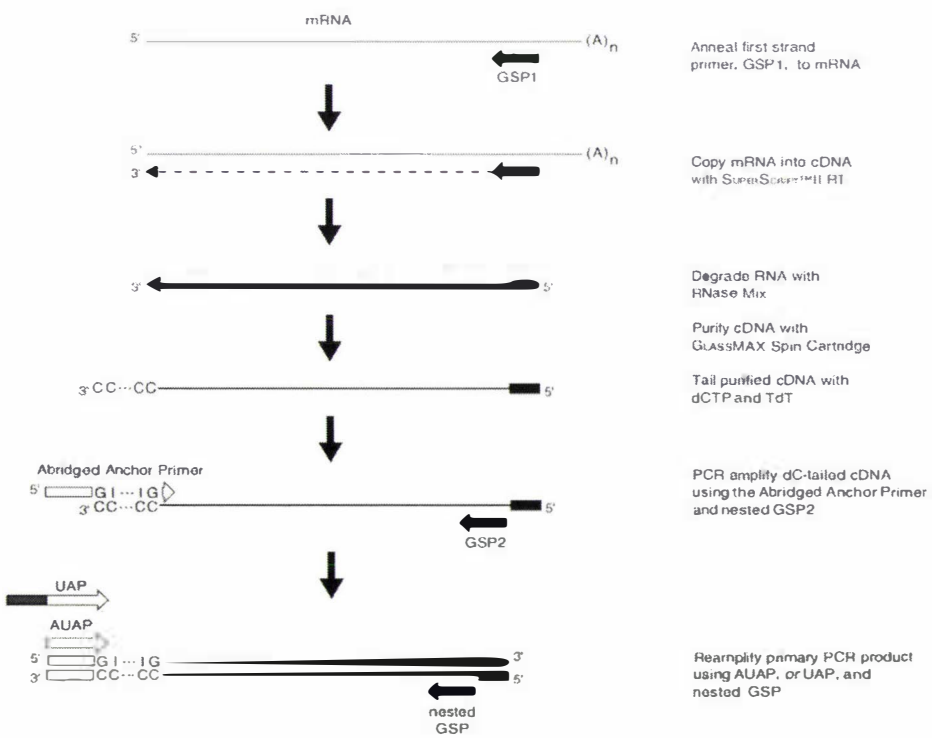
5' RACE was used to determine the transcription start site for the bovine ACL gene. 5'RACE is a procedure for amplification of nucleic acid sequences from a mRNA template between a defined internal site and unknown sequences at the 5'-end of the mRNA. Sequencing the amplified products would produce the sequence of the most 5'-end and the presumptive transcription initiation site. An overview of the procedures used is summarised in Figure 25. First strand cDNA was synthesized using a gene-specific primer (GSP) which anneals to an internal site of the specific mRNA. The 5' end cDNA products of first-strand synthesis were tailed and amplified with universal primers and nested gene-specific primers. From the sequences already determined, BLAST searches of the NCBI database indicated that the sequenced genomic fragment contained the promoter to intron 3 including regions which were homologous with rat, and human ACL exon 2 and exon 3. Three gene-specific primers were then designed in the putative exon 3 region.

Total cellular RNA was obtained from bovine liver tissue by the guanidinium thiocyanate-CsCl method as described in section 2.2.20. This method gave a good yield of high quality RNA. Clear 28S and 18S rRNA bands seen on formaldehyde gel electrophoresis indicated that the total RNA was not degraded (Figure 26A).

First strand cDNA was synthesized from total RNA using a gene-specific primer (GSP1) and SUPERSCRIPT™ II, an RNase H-derivative of Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT), by incubating at 42^0C for 50 min. After first strand cDNA synthesis, the original RNA template was removed by treatment at 37 ^{0}C for 20 min with the RNase Mix (mixture of RNase H, which is specific for RNA:DNA heteroduplex molecules, and RNase T1). Unincorporated dNTPs, GSP1, and proteins

were separated from cDNA using a PCR Clean-up Spin Cartridge (Promega). A homopolymeric tail was then added to the 3'-end of the cDNA using TdT and dCTP by incubating at 37 ^{0}C for 10 min. PCR amplification was accomplished using *Taq* DNA polymerase, a nested, gene-specific primer (GSP2) that anneals to a site located within the cDNA molecule, and a novel deoxyinosine-containing anchor primer (Abridged Anchor Primer, AAP). To increase the specificity and quantity of the PCR product, the PCR reaction was performed once more with a primer provided with the kit (Abridged Universal Amplification Primer, AUAP) and a farther upstream, nested, gene-specific primer (GSP3). PCR reactions were performed under the following conditions: 94^{0}C for 2min, 30 cycles of 94^{0}C for 30s, 55^{0}C for 45s, 72^{0}C for 60s, and one final extension step at 72^{0}C for 5min. Two different sizes of PCR product were seen on agarose gel electrophoresis (Figure 26B). They were 0.32 kb and 0.27 kb, respectively. These two clear bands suggested that the products were highly specific. They were recovered from the gel and were cloned into the pGEM-T Easy vector for subsequent sequencing.

Figure 25: Overview of the 5'RACE procedure



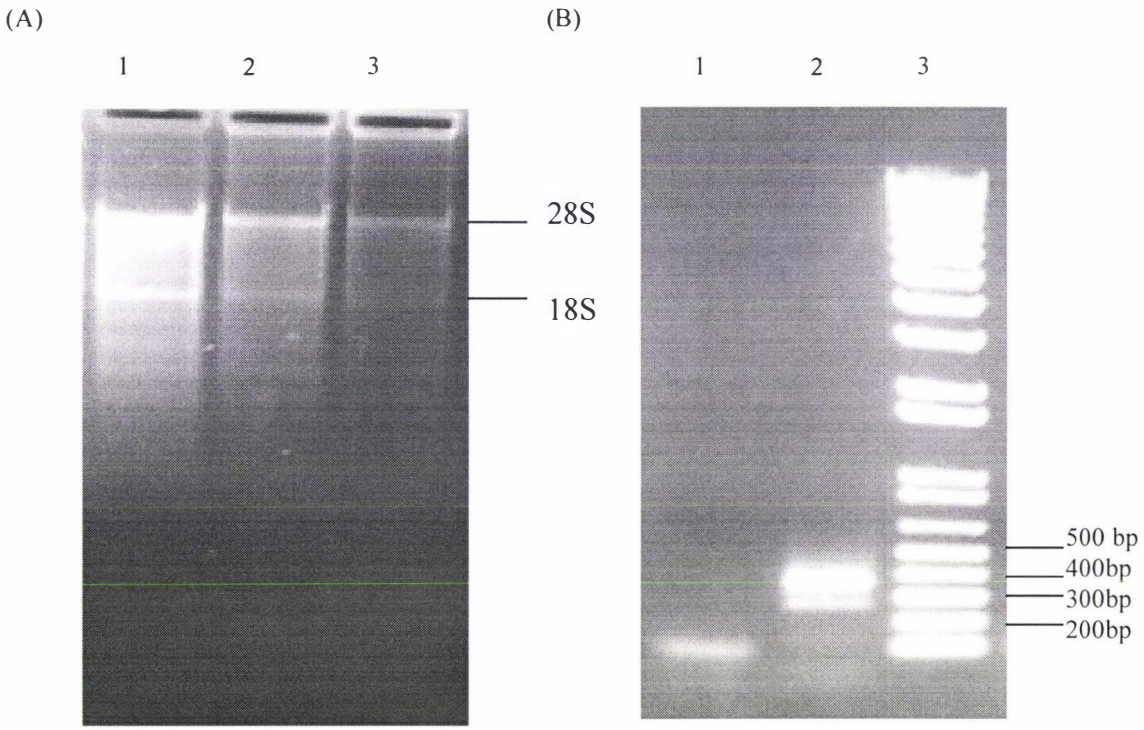


Figure 26: Gel electrophoresis of total cellular RNA isolated from bovine liver tissue (A) and 5'RACE products (B).

(A): Gel details for total RNA: 1.5% agarose gel containing 2.2 M formaldehyde, stained with ethidium bromide and visualized by UV. Electrophoresis was carried out at 44V for 3 hours.

Loading details: (1). 8.0 µg total RNA, (2). 4.0 ug total RNA, (3). 2.0 µg total RNA

(B): 5'RACE product: after two rounds of PCR amplification of first strand cDNA, the products were electrophoresed on a 0.7% agarose gel as described in section 2.2.

Loading details: (1). negative control, (2). 5'RACE products, (3). BRL 1 kb plus DNA ladder.

A potential disadvantage of 5'RACE system is that PCR can preferentially amplify products that do not represent the authentic 5' ends of the transcripts. During the initial reverse transcriptase step, the reverse transcriptase may have terminated with low frequency at specific nucleotides before reaching the 5' ends of the transcripts due to secondary structure of the template RNA, or suboptimal reverse transcription conditions. These shorter copies may be preferentially amplified as the dominant product, making them appear to represent the true and major transcription start site (Carey 1999). This seems not to be the case with these 5'RACE products. As shown in Figure 26, the shorter product band was much weaker than the longer one, and no further longer band can be observed. This suggested that the result from 5'RACE system was reliable and there are indeed two transcription start sites and the major one was the longer one.

Two clones (designated as 5RACE11 and 5RACE8) with different sized inserts were sequenced. A BLAST analysis of these two sequences against the genomic sequence showed that both contained different lengths of exon1, and identical complete exon 2 and partial exon 3 sequences. 5RACE11 contained 106 bp of exon1 corresponding to positions 3251-3356 in the 9490 bp genomic sequence. The sequence of 5RACE8 corresponding to exon 1 was 50 bp shorter. It corresponded to positions 3301-3356 in the genomic sequence. This result suggests that there are two transcription start sites. One is at 3251, and another is at 3301. The sequence downstream of these sites is shown below and the two transcription start sites are underlined.

3251 **A**CAAAAGCCA GGTCTCCGGC GGCTGCGGGA GGCTGGAGCG CTCCTAGCAG

3301 **T**GGGCGGTGA CTTGGGTTCT GTAGACTGAA CGCCGCGCAT AGACTTTCTC

3351

3.7 Analysis of the Sequence

There are many useful computer tools to do DNA sequence analysis, none of these are 100% accurate and so it is useful to be able to compare the results of many programs which use different methods. The bovine genomic sequence was analysed by using a number of computer programs. Most useful were the NIX DNA analysis program at the UK Human Genome Mapping Project Resouce Centre, the GCG (Genetics Computing Group) package at eBioinformatics. Inc. and the BLAST search program at the NCBI (National Center for Biotechnology Information, US).

3.7.1 An Overview of the Whole Sequence

NIX is a WWW tool for viewing the results of running many DNA analysis programs. It provided an excellent overview for the analysed sequence. As shown in Figure 27, the green fat line in the centre is the bovine ACL sequence. The analysis was carried out on both forward sequence (above the sequence line) and reverse sequence (below the sequence line). Regions that share some common property were shown as segmented blocks and linked by a line. Each block can be further studied using the linked programs. Several significant features of the 9490 bp bovine ACL gene sequence were discovered by this program.

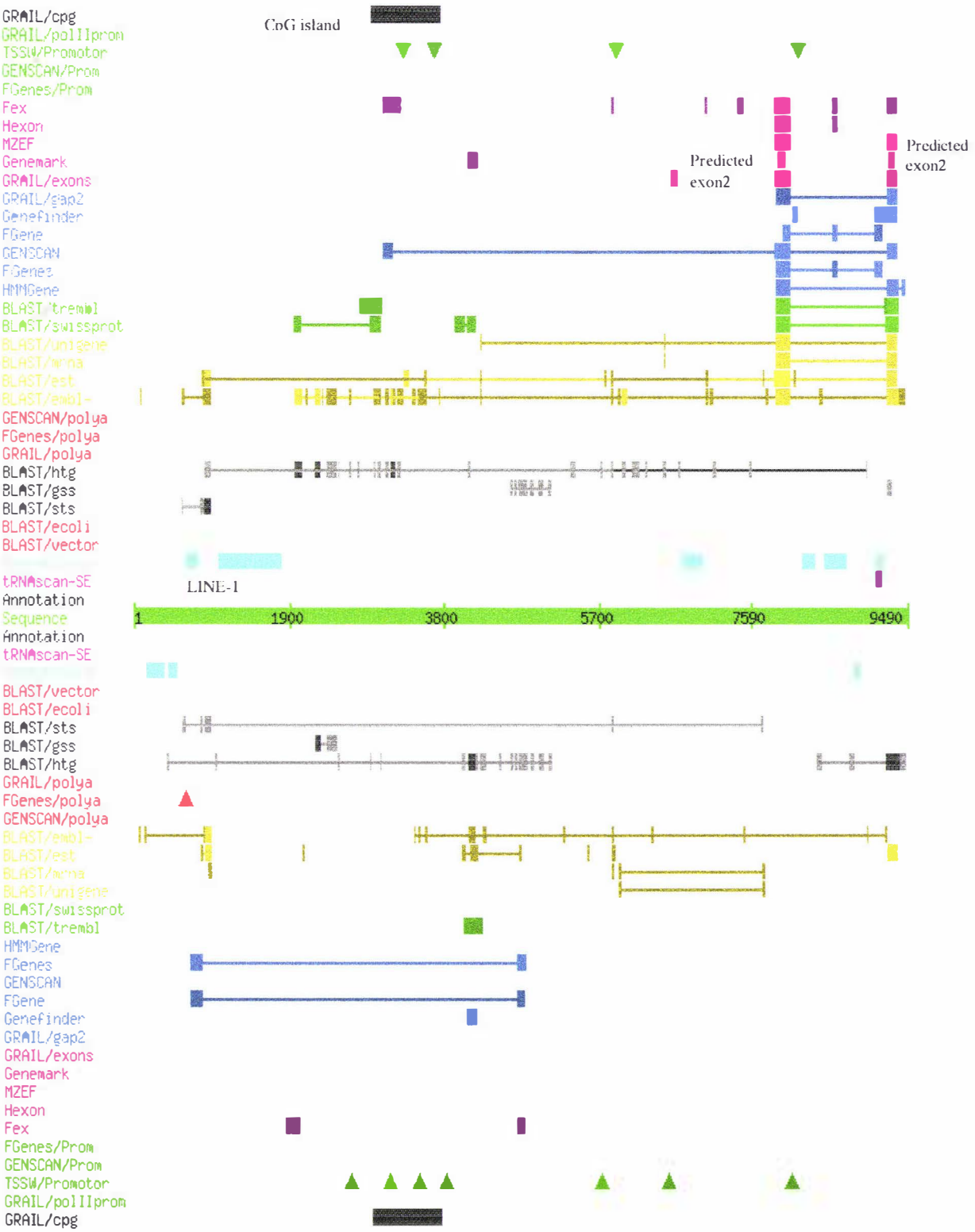
- GRAIL found a CpG island.
- RepeatMasker found two LINE-1 elements.
- 14 programs precisely predicted the exon 2 and exon 3.
- Several BLAST programs described the homology of bovine ACL with other sequence.

These results, combined with the results and knowledge from other sources, are discussed below.

Homology with other sequences:

Comparison of the 9490 bp bovine ACL gene sequence with the corresponding region (1.6 kb upstream of exon 1 to the beginning of intron 3) of the rat ACL genomic

Figure 27: Analysis of the 9490 bp bovine ACL gene sequence using the program NIX



sequence showed a 51.67% identity. In the Highthroughput (htg) Database at NCBI, there is a working draft sequence (GI:6730720 [GenBank]) of human chromosome 17, in which the human ACL gene is located. Although some similarity was found between this sequence and our sequence data, with many gaps and unordered pieces, the accurate identity can not be obtained. The other sequences at the NCBI database that are highly homologous with 9490 bp bovine ACL are: the human ACL mRNA complete sequence (GI:603073 [GenBank]), the human ACL promoter sequence (GI:4850175 [GenBank]), the rat ACL mRNA complete sequence (GI:949989 [GenBank]) and the rat ACL genomic DNA sequence of exon 1 to exon 7 (GI:436002 [GenBank]). The sequences with marginal similarity are not discussed here.

Organisation of exons and introns:

Several BLAST programs gave similar results for the organisation of exons and introns. Exon 2 and 3 were perfectly predicted, the results matched exactly with that obtained from the 5'RACE products. Their sizes and locations were very similar to that of exon 2 and 3 in the human and rat ACL sequence. Exon 1 was not found by these programs. A comparison of the sizes of these exons and introns is shown in Table 7.

Table 7: The length of exons (1, 2 and 3) and introns (1 and 2) from the bovine, human and rat ACL genes.

	Bovine ACL	Human ACL	Rat ACL
Exon 1	106 bp/66 bp	64 bp	76 bp
Exon 2	182 bp	~177 bp	182 bp
Exon 3	123 bp	~128 bp	122 bp
Intron 1	4490 bp	Unknown	4307 bp
Intron 2	1191 bp	Unknown	1193 bp

LINE element:

RepeatMasker is a program that screens DNA sequences for low complexity DNA sequences and interspersed repeats known to exist in mammalian genomes. It identified

that the sequences 626-766 and 1012-1799 as belonging to the L1 family. The L1 element (**LINE**-1, long interspersed repeated DNA) is the mammalian version of the non-long terminal repeat class of transposable elements that replicate via an RNA intermediate (retrotransposons). Every modern mammalian species studied to date contains a distinctive L1 family consisting of tens of thousands of members, which are interspersed throughout the genome (Furano and Usdin, 1995). The significance of the existence of L1 element in this gene is unknown. The sequence further upstream did not show any other significant features. Its role in promoter function remains to be determined. The short L1-like element was also found in the corresponding region of the rat ACL gene but not in the human one. The long one does not appear in the corresponding region of the human or the rat genome.

CpG island:

GRAIL is a suite of analysis tools for sequence exploration and gene discovery. If presented with genomic DNA, GRAIL collect important auxillary information such as putative polyA sites, Pol II promoters, CpG Islands. The algorithm to determine CpG Islands is based on the definition of CpG Islands by Gardiner-Garden and Frommer (1987). The sequence at position 2894-3747 was found to be a CpG Island. This predicted CpG island is 854 bp in length, containing a part of the promoter region, exon 1 and a part of intron 1. Its GC content is 67.35%. A CpG island of similar size and location is also present in human and rat ACL genes.

The presence of CpG islands in the 5' regions of some genes is connected with the effect of methylation on gene expression. These islands are detected by the presence of an increased density of the dinucleotide sequence, CpG. They have an average G·C content of ~60%, compared with the 40% average in bulk DNA. They take the form of stretches of DNA typically 1-2 kb long. The nucleosomes at the islands have a reduced content of histone H1, the other histones are extensively acetylated, and there are hypersensitive sites. In several cases, CpG islands begin just upstresm of a promoter and extend downstream into the transcribed region before petering out. All of the housekeeping genes that are constitutively expressed have CpG islands, this accounts for about half of

the islands altogether (Kundu and Rao, 1999). The other half of the islands occur at the promoters of tissue-regulated genes, but only a minority of these genes have islands. In these cases, the islands are unmethylated irrespective of the state of expression of the gene. The presence of unmethylated CpG islands may be necessary, but therefore is not sufficient, for transcription. Thus the presence of unmethylated CpG islands may be taken as an indication that a gene is potentially active, rather than inevitably transcribed. This may be the case with the bovine ACL gene, but no research had been carried on this issue so far.

3.7.2 Analysis of the mRNA Sequence

Comparison of the bovine ACL sequences from the 5'RACE products with human and rat ACL mRNA sequences showed a high degree of identity and confirmed that the 5'RACE products contained exon 1, 2, and a part of exon 3 sequence. By comparison with the 9490 bp bovine ACL sequence, the positions of these three exons in th genomic sequence was established. The full length of exon 3 was determined based on the structure of exon-intron boundaries (GT-AG rule) and the homology with exon 3 of the human and rat ACL. These results were consistent with the predictions from several computer programs. A potential translation start codon was not found in exon 1. There were two ATG sequences in exon 2. By comparison with human and rat sequences, it appeares that the first ATG in exon 2 was the translation start codon. This mRNA sequence is shown below (with different colors for each exon):

```
1    ACAAAAGCCA GGTCTCCGGC GGCTGCGGGA GGCTGGAGCG CTCCTAGCAG TGGGCGGTGA
61   CTTGGGTTCT GTAGACTGAA CGCCGCGCAT AGACTTTCTC GGAGAGGTAG TGCAGGGCCT
121  TCTGTAACCATGTCGGCCAA GGCGATTTCT GAGCAGACAG GCAAAGAGCT CCTCTACAAG
181  TACATCTGCA CCACCTCGGC CATCCAGAAC CGCTTCAAGT ATGCCCGGGT CACGCGTGAC
241  ACAGACTGGG CCCGCCTGTT GCAGGACCAC CCGTGGCTGC TCAGCCAGAG CTTGGTGGTC
301  AAGCCAGACC AGCTGATCAA ACGGCGTGGA AAGCTGGGCC TAATTGGGGT CAACCTCACT
361  CTGGATGGAG TCAAGTCCTG GCTGAAGCCA CGCCTGGGAC AGGAAGCTAC A
```

The sequences of exon 1 to 3 of the three species were aligned and shown in Figure 28. Among bovine, human and rat genes, the three exons together showed 70.21% identity. Exon 1 alone only had 50.31% identity. But for exon 2 and 3 together, there was 93.31% identity and this reached 96.49% at the amino acid level (Figure 29). These results suggested that the function of the ACL gene product in these three species is highly conserved, although the mechanisms of regulation in expression might be divergent.

Figure 28: Comparison of the mRNA sequences of exon 1, 2, 3 from the bovine, human and rat ACL gene.

Figure 29: Alignment of the protein sequences of exon 2 and 3 from bovine, human and rat

```
                MSAKAISEQTGKELLYK5ICTTSAIQNRFKYARVTPDTDWArLLQDHPWL
Rat         :   MSAKAISEQTGKELLYKYICTTSAIQNRFKYARVTPDTDWAHLLQDHPWL  :   50
Human       :   MSAKAISEQTGKELLYKFICTTSAIQNRFKYARVTPDTDWARLLQDHPWL  :   50
Bovine      :   MSAKAISEQTGKELLYKYICTTSAIQNRFKYARVTPDTDWARLLQDHPWL  :   50

                LSQsLVVKPDQLIKRRGKLGL6GVNL3LDGVKSWLKPRLGqEAt
Rat         :   LSQSLVVKPDQLIKRRGKLGLVGVNLSLDGVKSWLKPRLGHEA-T      :   94
Human       :   LSQNLVVKPDQLIKRRGKLGLVGVNLTLDGVKSWLKPRLGQEAT-      :   94
Bovine      :   LSQSLVVKPDQLIKRRGKLGLIGVNLTLDGVKSWLKPRLGQEAT-      :   94
```

3.7.3 Analysis of the Promoter Region

The promoter sequence alignment for the bovine, human and rat ACL is shown in Appendix 3. A common starting point for analyzing gene regulation at the level of transcription initiation is to identify its promoter and characterize the *cis*-acting sequence elements and *trans*-acting proteins responsible for promoter activity. The promoter includes the DNA sequence elements in the vicinity of the start site that direct activation or repression of transcription. The transcription start sites were determined by the 5'RACE experiment and sequence analysis described earlier. The *cis*-acting sequence elements were mapped using computer programs. These programs predicted potential transcription start sites (TSS) using a linear discriminant function that combines functional motifs and oligonucleotide composition across the query sequence. Selected transcription factor binding sites were detected using the **TSSW, TSSG** or **TESS** programs (from Baylor College of Medicine, Houston, US) and the currently supported functional site database created by E. Wingender (1994), in addition to the parameters of Prestridge's method (1995).

Most promoters have a sequence called the TATA box, usually located ~25 bp upstream of the transcription start point. It is a crucial positioning component of the core promoter. Although several programs respectively gave predictions for the TATA box position at 3130, 3173 and 3249 of this 9490 bp bovine ACL sequence, none of them seems to be the true TATA box because these site are upstream far away from the transcription start

sites determined by the 5'RACE. It is not surprising that bovine ACL has a TATA-less promoter, since this is the case with the human and rat ACL genes. It could also explain the observation that this gene has two transcription start sites.

GC boxes are a relatively common promoter component. Often multiple copies are present in the promoter, and they occur in either orientation. Six GC-boxes were found at 3221, 3119, 3002, 2978, 2884 and 2486 in this sequence, located respectively 30, 142, 249, 273, 367 and 765 bp upstream from the major transcription start site (at 3251). These findings correlate with a previous report that clustered GC boxes are required for efficient transcription from TATA-less promoters (Smale, 1997).

The inverted Y-box (CCAAT) is another common element in many promoters. It is often located at about –80, but it can function at distances that vary considerably from the start point. It does not appear to play a direct role in promoter specificity, but its inclusion increases promoter strength. A recent report suggested the adjacent inverted Y-box played a critical role in SREBP transactivation (Moon, 2000). Two inverted Y-boxes were found, one is at 2365, another at 3187 which is just 74 bp upstream from the major transcription start site.

Sterol regulatory element binding proteins (SREBPs) are the transcription factors which regulate the transcription of many genes involved in cholesterol and fatty acid synthesis. Recent reports suggested that SREBP-1 might be a potential transactivator mediating lipogenic enzyme gene regulation during the fasting/refeeding cycle. The action of SREBPs on the promoters of genes for several lipogenic enzymes is discussed in Section 1.3.3. It was suggested that SREBPs stimulate the ACL promoter through their binding to an upstream promoter region and that the activation requires NF-Y binding on the inverted Y-box (see Section 1.4.5). Interestingly, the region containing SRE sites in the bovine ACL gene is very similar to the corresponding region in the human and rat ACL as shown in Figure 30. This region is at 3162 to 3192 in the 9490 bp bovine ACL sequence.

Compared to human and rat, expression of the adult bovine ACL gene is repressed under normal physiological conditions. However, it can be induced by infusing glucose postruminally or intravenously (Muramatsu *et al.*, 1969). Considering the role of SREBPs in other mammalian systems and that the expression pattern of lipogenic enzymes in young ruminants is similar to that in non-ruminants, the conserved SRE sites in the bovine ACL gene may be responsible for this glucose induction. This raises questions as to whether the SRE sites are responsible for repression of the bovine ACL gene and what mechanism is involved?

Figure 30: An alignment of the SRE sites in the bovine, human and rat ACL genes.

```
                TCAGGCTAGGGAAcGCGTGTGgCCAAT
bovine      :   TCAGGCTAGGGAATGCGTGTGGCCAAT   :   27
human       :   TCAGGCTAGGGAACGCGTGTGGCCAAT   :   27
rat         :   TCAGGCTAGGGAACGCGTGTGCCCAAT   :   27
```

Repression mechanisms are less well understood than activation mechanisms. In general, transcriptional repression can be divided into three broad categories. First, repression can occur by inactivation of an activator, which can be accomplished by several distinct mechanisms. Second, repression can be mediated by proteins that associate tightly with general transcription factors and thereby inhibit the formation of a initiation complex. The third category of repression is mediated by a specific DNA element and DNA-binding protein, which act dominantly to repress both activated and basal transcription of a given gene. To date, there is no study on repression of any gene encoding a lipogenic enzyme. The bovine ACL gene may be a good candidate for such a study.

It would be worthwhile to ask if inactivation of SREBP is a possible mechanism responsible for the repression of the bovine ACL gene and then what factors mediate this inactivation. Whether there is another intrinsic mechanism, for example, silencer and repressor involved? If bovine acetyl-CoA synthetase (ACS) was activated by SREBPs under physiological condition (It may also be true for ACC and FAS), how are the

activation and repression of SREBPs for ACL and ACS mediated in a single cell? Analysis of our bovine ACL promoter sequence would be able to start to answer these questions.

In addition, a number of sequence elements similar to the hepatocyte nuclear factors (HNF), fat-specific element (FSE), cAMP response element (CRE), thyroid response element (TRE) and the glucocorticoid response element (GRE), were found in this promoter region. The predictions from the computer programs showed that there is also a lot of other possible transcription factor binding sites (Appendix 2).

The predictions of transcription factor binding sites in the ACL promoter will facilitate further investigation, but experimental confirmation would be required before any conclusions can be drawn.

CHAPTER FOUR: FUTURE DIRECTIONS

The successful isolation of the bovine ATP-citrate lyase promoter provides a basis for further investigations which will provide an insight into the mechanisms controlling the expression of the bovine ACL gene. This information will not only increase our understanding of the mechanisms involved in the regulation of acetyl-CoA synthesis but will also provide insights into the molecular mechanisms involved in the regulation of the expression of lipogenic enzymes in general. In addition, understanding of the molecular mechanisms involved in repression of the bovine ACL gene will complement the current knowledge of eukaryotic gene expression and is likely to benefit the understanding and controlling of clinical problems such as obesity and diabetes which are closely related to lipogenesis and its regulation.

4.1 Confirmation of Transcription Start Sites

Accurate mapping of the transcription start site is extremely important for the success of subsequent promoter analysis. Although the results obtained by 5'RACE were very clear and convincing, more than one technique should be used to increase the reliability of the results. Any one of the methods described below can be chosen to confirm the results from 5'RACE.

4.1.1. Nuclease Protection

The various methods of mapping RNA by nuclease protection share a common principle. A population of RNA is hybridized to a radioactively labelled probe that is complementary, along part or all of its length, to the specific RNA being analysed. Probe molecules that fail to hybridize, and those regions of the probe that are not annealed to the target RNA, are then removed by nuclease digestion. Finally, the digestion products are analysed by gel electrophoresis. The two commonly used methods, nuclease S1 mapping and RNase protection, differ only in the nature of the probe and the nuclease that they employ. The RNA-RNA hybrids formed during RNase protection are

thermodynamically stable, much more so than the RNA-DNA hybrids formed during nuclease S1 mapping. The results obtained from genomic sequencing and 5'RACE provide information for construction of appropriate probes for nuclease protection analysis.

4.1.2 Primer Extension

Primer extension analysis is the direct converse of nuclease mapping. An end-labeled oligonucleotide derived entirely from within the gene is hybridized to RNA and is utilized as a primer by reverse transcriptase in the presence of deoxynucleotides. The best region for designing the probe is exon 1 or 2 which sequences are already known. The RNA is thus reverse transcribed into cDNA and is analyzed on a denaturing polyacrylamide gel. The length of the cDNA reflects the number of bases between the labeled nucleotide of the primer and the 5′-end of the RNA; the quantity of cDNA product is proportional to the amount of targeted RNA. This method uses similar technology to 5'RACE and may be subject to the same limitation.

4.2 Determination of the Minimal Promoter of the Bovine ACL Gene

Before extensive characterisation of the bovine ACL promoter can be undertaken, it will be necessary to define the minimal promoter region required for transcriptional activity. This can be achieved by progressively reducing the length of a functional ACL promoter until the ability to support transcription in a transient reporter gene expression system ceases.

Functional assays involve the introduction of reporter gene constructs, containing the promoter region of interest, into tissue culture cells and monitoring the transfected cells for the expression of the reporter gene. Initially the fragments tested could be based on the presence of suitable restriction endonuclease sites in this region of the gene. Subsequent constructs will be prepared by PCR once the initial transcriptional activity of the region has been established. Suitable fragments of the 5' region of the bovine ACL

gene could be subcloned into the pGL3™ series of luciferase reporter gene plasmids (Promega). The transcriptional activity of the constructs could be tested by transfection into HepG2 cells (a human hepatocyte cell line) and differentiated 3T3-Li cells. The latter cell line differentiates to form adipocytes under appropriate conditions.

4.3 Binding Sites for Transcription Factors

Promoters are normally composed of clusters of *cis*-acting DNA sequences which are situated distal and proximal to the transcriptional start point. The binding of *trans*-acting protein factors to these DNA-motifs and protein-protein interactions between the bound transcription factors form the transcription initiation complex which facilitates gene expression. Computer sequence homology searches of the 5'-flanking regions of the bovine ACL promoter sequences (on 2500-3500 of the 9490 bp sequence) have identified some putative DNA-binding elements (Appendix 1 and 2). This has revealed the presence of putative consensus binding elements for several known transcription factors as described in section 3.7.3. Combined with the results from reporter gene assays, it should be possible to indicate the transcription factors that may be important in the regulation of this gene. These results will be supported by DNase footprinting assay and electrophoretic mobility shift assays.

4.3.1 DNase I footprinting

DNase I footprinting allows short sequences representing protein-binding sites to be identified within a relatively large DNA fragment. A double stranded fragment of the promoter region which is radiolabelled at one end, is mixed with either purified protein factors or crude nuclear extract. Protein bound to the DNA will protect specific regions from enzymatic attack by Dnase I. Maxam and Gilbert sequencing reactions of the same DNA fragment are separated by denaturing polyacrylamide gel electrophoresis alongside the DNase I cleavage products to allow the specific nucleotides involved in the protein interaction to be determined.

4.3.2 Electrophoretic mobility shift assays (EMSA)

EMSA provides a method for investigating the ability of protein factors, either purified or in a crude nuclear extract, to specifically interact with a radiolabelled, double stranded oligonucleotide probe. The probe generally consists of a short DNA sequence (20-30 bp) which represents a putative binding site for a transcription factor. Protein factors bound to the radiolabelled oligonucleotide cause it to migrate more slowly than the free probe when subjected to electrophoresis through a non-denaturing polyacrylamide gel. These result in the appearance of a retarded band upon autoradiography allowing the presence of bound protein factors to be detected. The specificity of protein-DNA interactions can be tested by the use of competitor oligonucleotides representing specific and non-specific DNA sequences within the promoter region being investigated.

Following these experiments one would be in a position to either identify the protein factors involved in regulation if they are those already known or to design a protocol for isolation of a new factor if one is shown to be involved.

4.4 Expression of Bovine ACL in Tissues During Development

Enzyme activity and mRNA levels of various lipogenic enzymes has been followed in various tissues of rats maintained under different dietary regimes (fasting, and refeeding high carbohydrate and high fat diets). However, most of this work has been focused on ACC and FAS (Towle, 1997). It would be informative to examine the expression of ACL mRNA by quantitative northern blotting or ribonuclease protection assays and to assay ACL activity in homogenised cell supernatants from the same tissues of young and developing ruminants. As ACL is regulated at the level of transcription, the levels of ACL mRNA will reflect the rate of transcription. The two cDNA products with different sizes obtained from 5'RACE indicated that there are two sizes of mRNA, thus also two transcription start sites. The concentrations of these two mRNAs will reflect the difference in the function of two promoters, it is very likely that these differences will be in a tissue specific patterns.

4.5 Long-term Aims

Once the different expression patterns of bovine ACL during development have beeb determined, the roles of the *cis*-elements and *trans*-factors involved in this changes will be studied. The regulatory roles of glucose, PUFA, insulin, and acetate (which has specific importance for ruminants) will also be investigated.

The role of SREBPs in lipogenesis has aroused increasing interest. The highly conserved SRE sequences in the bovine ACL promoter indicate that SREBPs may have important roles in lipogenesis in bovine, as well as human and rat. Although some common functions for SREBPs may be shared by these animals, the activation of bovine ACL by SREBPs, if this is the case, may involve a different mechanism since bovine ACL expression is usually low and can be activated under certain conditions. It may be that bovine ACL expression is usually low because SREBPs functions are normally repressed in adults ruminants. This possibility would be unique to ruminants and will be further investigated.

The studies described above could in future be extended to other lipogenic enzymes such as ACC and FAS, but acetyl-CoA synthetase (ACS) will be the primary focus. Compared to the low level of ACL expression in adult ruminants, ACS has a high expression level to utilize the acetate produced in the rumen of adult ruminants. This is just the opposite state to the situation in young ruminants. The regulation of these two enzymes must be co-ordinated. Analysis of the switch in the expression patterns of these two enzymes will complement knowledge of lipogenesis regulation.

REFFERENCES

Anderson, J. W. and Bridges, S. R. (1984). Short chain fatty acid fermentation products of plant fiber affect glucose metabolism of isolated rat hepatocytes. *Proceedings of the Society for Experimental biology and Medicine.* **177:**372-376.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1989). *Current Protocols in Molecular Biology.* Greene Publishing Associates and Wiley-Interscience.

Azzout-Marniche, D., Dominique, B., Colette, D., Marc, F., Pascal, F. and Fabienne, F. (2000). Insulin effects on sterol regulatory-element-binding protein-1c (SREBP-1c) transcriptional activity in rat hepatocytes. *Biochemistry.* **350**:389–393.

Baldwin, R. L. and Smith, N. E. (1971). Intermediary aspects and tissue interactions of ruminant fat metabolism. *Journal of Dairy Science.* **54**:583-595.

Baldwin, R. L., Ronning, M., Radanovics, C. and Plange, G. (1966). Effect of carbohydrate and fat intakes upon the activities of several liver enzymes in rats, guinea pigs, piglets and calves. *Journal of Nutrition* **90**:47-55.

Ballard, F. J., Filsell, O. H. and Jarrett, I. G. (1972). Effects of carbohydrate availability on lipogenesis in sheep. *Biochemical Journal* **226:**193-200.

Bauman, D. E. (1976). Intermediary metabolism of adipose tissue. *Federation procedings* **35**: 2308-2313.

Benjamin, W. B., Pentyala, S. N., Woodgett, J. R., Hod, Y. and Marshak, D. (1994). ATP citrate lyase and glycogen synthase kinase-3 P in 3T3-L I cells during differentiation into adipocytes. *Biochemical Journal.* **300:**477-482.

Bennett, M. K., Lopez, J. M., Sanchez, H. B. and Osborne, T. F. (1995). Sterol regulation of fatty acid synthase promoter. *Journal of Biological Chemistry* **270**:20078-20083.

Bennett, M. K., Ngo, T. T., Athanikar, J. N., Rosenfeld, J. R. and Osborne, T. F. (1999). Co-stimulation of promoter for low density lipoprotein leceptor lene by sterol regulatory element-binding protein and Sp1 is specifically disrupted by the Yin Yang I protein. *Journal of Biological Chemistry.* **274**:13025-13032.

Bergman, E. N., Reid, R. S., Murray, M. G., and Whitelaw, F. G., (1965). Interconversions and production of volatile fatty acids in the sheep rumen. *Biochemical Journal,* **97**:53-58.

Bergman, E.N. (1973). Glucose metabolism in ruminants as related to hypoglycemia and ketosis. *Cornell Veterinary.* **63**:341-382.

Brown, M. S. and Goldstein, J. L. (1997). The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell.* **89**: 331-8.

Carey, M. and Smale, S. T. (1999). *Transcriptional regulation in eukaryotes: concepts, strategies, and techniques.* Cold Spring Harbor Laboratory Press.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry.* **18**:5294-5299.

De Jong, A. (1982). Patterns of plasma concentrations of insulin and glucagon afer intravascular and intraruminal administration of volatile fatty acids in the goat. *Journal of Endocrinolog.* **92:**357-370.

Dooley, K. A., Bennett, M. K. & Osborne, T. F. (1999). A critical role for cAMP response element-binding protein (CREB) as a co-activator in sterol-regulated transcription of 3-hydroxy-3-methylglutaryl coenzyme A synthase promoter. *Journal of Biological Chemistry.* **274**:5285-5291.

Elshourbagy, N. A., Near, J. C., Kmetz, P. J., Sathe, G. M., Southan, C., Strickler, J. E., Gross, M., Young, J. F., Wells, T. N. C. and Groot, P. H. E. (1990). Rat ATP citrate-lyase: molecular cloning and sequence analysis of a full-length cDNA and mRNA abundance as a function of diet, organ, and age. *Journal of Biological Chemistry.* **265**:1430-1435.

Elshourbagy, N. A., Near, J. C., Kmetz, P. J., Wells, T. N. C., Groot, P. H. E., Saxty, B. A., Hughes, S. A., Franklin, M. and Gloger, I. S. (1992). Cloning and expression of a human ATP-citrate lyase cDNA. *European Journal of Biochemistry* **204:**491-499

Ericsson, J., Jackson, S. M., Kim, J.B., Spiegelman, B. M. and Edwards, P. A. (1997). Identification of glycerol-3-phosphate acyltransferase as an adipocyte determination and differentiation factor 1- and sterol regulatory element binding protein-responsive gene..*Journal of Biological Chemistry* **272**: 7298-305.

Flier, J. S and Hollenberg, A. N. (1999). ADD-1 provides major new insight into the mechanism of insulin action. *Proceedings of the National Academy of Sciences (USA)*. **96**:14191-2.

Foretz, M., Guichard, C., Ferré, P. & Foufelle, F. (1999). Sterol regulatory element binding protein-1c is a major mediator of insulin action on the hepatic expression of glucokinase and lipogenesis-related genes. *Proceedings of the National Academy of Sciences (USA)*. **96**:12737-12742.

Foretz, M., Pacot, C., Dugail, I., Lemarchand, P., Guichard, C., Le Liepvre, X., Berthelier-Lubrano, C., Spiegelman, B., Kim, J. B., Ferre, P. and Foufelle, F. (1999). ADD1/SREBP-1c is required in the activation of hepatic lipogenic gene expression by glucose. *Molecular and Cellular Biology*. **19:**3760-8.

Fukuda, H., Iritani, N., Katsurada, A. and Noguchi, T. (1996). Insulin and polyunsaturated fatty acid responsive region(s) of rat ATP citrate-lyase gene promoter. *FEBS Letters* **380:**204-207.

Fukuda, H., Iritani, N. and Noguchi, T. (1997) Transcriptional regulatory region for expression of the rat ATP citrate lyase gene. *European Joumal of Biochemistry*. **247:** 497-502.

Furano, A. V. and Usdin, K. (1995). DNA "fossils" and phylogenetic analysis. Using L1 (LINE-1, long interspersed repeated) DNA to determine the evolutionary history of mammals. *Journal of Biological Chemistry*. **270:**25301-4.

Girard, J., Ferre, P. and Foufelle, F. (1997). Mechanisms by which carbohydrates regulate expression of genes for glycolytic and lipogenic enzymes. *Annual Review of Nutrition*. **17**:325-52

Halperin, M. L., Cheema-Dhaldi, S., Taylor, W. M. and Fritz, L. B. (1975) Role of the citrate transporter in the control of fatty acid synthesis. *Advances in Enzyme Regulation.* **13:** 435-445.

Hanson, R. W. and Ballard, F. J. (I 967) The relative significance of acetate and glucose as precursors for lipid synthesis in liver and adipose tissue from ruminants. *Biochemical Journal.* **105:** 529-536.

Hardwick, D. C. (1966). The fate of acetyl groups derived from glucose in the isolated perfused goat udder. *Biochemical Journal.* **99:**228-31.

Harmon, D. L. (1992). Impact of nutrition on pancreatic exocrine and endocrine secretion in ruminants: a review. *Journal of Animal Science.* **70:**1290-1301.

Hasty, A. H., Shimano, H., Yahagi, N,, Amemiya-Kudo, N., Perrey, T., Yoshikawa, T., Osuga, J., Okazaki, H., Tamura, Y., Iizuka, Y., Shionoiri, F., Ohashi, K., Harada, K., Gotoda, K., Nagai, R., Ishibashi, S. and Yamada, N. (2000). Sterol regulatory element-binding protein-1 is regulated by glucose at the transcriptional level. *Journal of Biological Chemistry.* **275**:31069-31077.

Hillgartner, F. B., Salati, L. M. and Goodridge, A. G. (1995). Physiological and molecular mechanisms involved in nutritional regulation of fatty acid synthesis. *Physiological Review.***75(1):**47-76.

Hua, X., Yokoyama, C., Wu, J., Briggs, M. R., Brown, M.S., Goldstein, J.L. and Wang. X. (1993). SREBP-2, a second basic-helix-loop-helix-leucine zipper protein that stimulates transcription by binding to a sterol regulatory element. *Proceedings of the National Academy of Sciences (USA).* **90**:11603-11607.

Kaytor, E. N., Shih, H. and Towle, H. C. (1997). Carbohydrate regulation of hepatic gene expression. *Journal of Biological Chemistry.* **272**: 7525-7531.

Kim, J. B. and Spiegelman, B. M. (1996). ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. *Genes & Development..* **10**:1096-1107.

Kim, J. B., Spotts, G. D., Halvorsen, Y. D., Shih, H. M., Ellenberger, T., Towle, H. C. and Spiegelman, B. M. (1995). Dual DNA binding specificity of ADD1/SREBP1 controlled by a single amino acid in the basic helix-loop-helix domain. *Molecular and Cellular Biology* . **15**:2582-2588.

Kim, J. B., Pasha. S., Wright, M., Yao, K. M., Mueller, E., Solanes, G., Lowell, B. B. and Spiegelman, B. M. (1998). Nutritional and insulin regulation of fatty acid synthetase and leptin gene expression through ADD1/SREBP1. *The Journal of Clinical Investigation.***101**:1-9.

Kim, J. B., Spotts, G. D., Halvorsen, Y. D., Shih, H. M., Ellenberger, T., Towle, H. C. and Spiegelman, B. M. (1995). Dual DNA binding specificity of ADD1/SREBP1 controlled by a single amino acid in the basic helix-loop-helix domain. *Molecular and Cellular Biology* . **15**, 2582-2588.

Kim, K. H. (1997). Regulation of mammalian acetyl-CoA carboxylase. *Annual Review of Nutrition* .**17:**77-99.

Kim, K. S., Park, S. W. and Kim, Y. S. (1992). Regulation of ATP-citrate lyase at transcriptional and post-transcriptional levels in rat liver. *Biochemical and Biophysical Research Communications* . **189**:264–71.

Kim, K. S., Park, S. W., Moon, Y. A. and Kim, Y. S. (1994). Organisation of the 5' region of the rat ATP citrate lyase gene. *Biochemical Journal.* **302:**759-764.

Kundu, T. K. and Rao, M. R. (1999). CpG islands in chromatin organization and gene expression. Journal of Biochemistry (Tokyo). **125(2):**217-22.

Lopez-Casillas, F., Ponce-Castaneda, M. V. and Kim K. H. (1991). *In vivo* regulation of the activity of the two promoters of the rat acetyl coenzyme-A carboxylase gene. *Endocrinology.* **129**:1049–58.

Luo, X., Park, K., Lopez-Casillas, F. and Kim, K. H. (1989). Structural features of the acetyl-Co A carboxylase gene: mechanisms for the generation of mRNAs with 50 end heterogeneity. *Proceedings of the National Academy of Sciences (USA).* **86**:4042--46.

Luong, A., Hannah, V. C., Brown, M. S., Goldstein, J. L. (2000). Molecular characterization of human acetyl-CoA synthetase, an enzyme regulated by sterol regulatory element-binding proteins. *Journal of Biological Chemistry.* **275:**26458-26466.

Magana, M. M. and Osborne, T. F. (1996). Two tandem binding sites for sterol regulatory element binding proteins are required for sterol regulation of fatty-acid synthase promoter. *Journal of Biological Chemistry.* **271**:32689–94.

Magaña, M., Koo, S. H., Towle, H.C and Osborne, T. F. (2000). Different sterol regulatory element-binding protein-1 isoforms utilize distinct co-regulatory factors to activate the promoter for fatty acid synthase. *Journal of Biological Chemistry.* **275:** 4726-4733.

Marter, M. K., Thelen, A. P., Pan, D. A., and Jump, D. B. (1999). Sterol response element-binding protein 1c (SREBP1c) is involved in the polyunsaturated fatty acid suppression of hepatic S14 gene transcription. *Journal of Biological Chemistry.* **274**: 32725-32732.

Millinder-Vallett, S., Sanchez, H. B., Rosenfeld, J. M., and Osborne, T. F. (1996). A direct role for sterol regulatory element binding protein in activation of 3-hydroxy-3-methylglutaryl coenzyme A reductase gene. *Journal of Biological Chemistry.* **271**: 12247-12253.

Moustaid, N., Beyer, S R. and Sul, H.S. (1994). Identification of an insulin response element in the fatty acid synthase promoter. *Journal of Biological Chemistry.* **269**:5629-5634.

Nåår, A. M., Beaurang, P. A., Robinson, K. M., Oliner, J. D., Avizonis, D., Scheek, S., Zwicker, J., Kadonaga, J. T. & Tjian, R. (1998). Chromatin, TAFs, and a novel multiprotein coactivator are required for synergistic activation by Sp1 and SREBP-1a in vitro. *Genes and Development.* **12**, 3020-3031.

Numa, S. and Tanabe, T.(1984). Acetyl-CoA carboxylase and its regulation. In *Fatty Acid Metabolism and Its Regulation,* pp. 1–27. New York: Elsevier Science.

O'Brien, R. M. and Granner, D. K. (1996). Regulation of gene expression by insulin. *Physiological Reviews.* **76**:1109-1161.

Olson, A. L. and Pessin, J. E. (1996). Structure, function, and regulation of the mammalian facilitative glucose transporter gene family. *Annual Review of Nutrition.* **16**:235-256.

Osborne, T. F. (2000). Sterol regulatory element-binding proteins (SREBPs): Key regulators of nutritional hhomeostasis and insulin action. *Journal of Biological Chemistry.* **275:** 32379-32382.

Park, S. W., Kim, K. S., Whang, S. K., Kim, J. S. and Kim, Y. S. (1994). Induction of hepatic ATP-citrate lyase by insulin in diabetic rat--effects of insulin on the contents of enzyme and its mRNA in cytosol, and the transcriptional activity in nuclei. *Yonsei Medical Journal.* **35:**25-33.

Park, S., Moon, Y., Kim, K., Ahn, Y. and Kim, Y. (1997). Cloning and characterization of the 5' flanking region of human ATP-citrate lyase gene. *Biochimica et Biophysica Acta.* **1353**:236-40.

Paulauskis, J. D. and Sul, H. S. (1988). Cloning and expression of mouse fatty acid synthase and other specific mRNAs. Developmental and hormonal regulation in 3T3-L1 cells. *Journal of Biological Chemistry.* **263**:7049–54.

Paulauskis, J. D. and Sul, H. S. (1989). Hormonal regulation of mouse fatty acid synthase gene transcription in liver. *Journal of Biological Chemistry.* **264**:574–77.

Pentyala, S. N. and Benjamin, W. B. (1995). Effect of oxaloacetate and phosphorylation on ATP-citrate lyase activity. *Biochemistry* **34:**10961-9.

Prestridge, D. S. (1995). Predicting Pol II promoter sequences using transcription factor binding sites. *Journal of Molecular Biology.* **249:**923-32.

Ramakrishna, S., Murthy, K. S. and Benjamin, W. B. (1989). Effect of insulin on ATP-citrate lyase phosphorylation: regulation of peptide A and peptide B phosphorylations. *Biochemistry.* **28:**856-60.

Ramakrishna, S., D'Angelo, G. and Benjamin, W. B. (1990). Sequence of sites on ATP-citrate lyase and phosphatase inhibitor 2 phosphorylated by multifunctional protein kinase (a glycogen synthase kinase 3 like kinase). *Biochemistry.* **29:**7617-24.

Robertson, J. P., Faulkner, A. and Vernon, R. G. (1981). Lactate as a source of carbon for fatty acid synthesis in adult and foetal sheep. *Biochimica et Biophysica Acta*, **65:**511-518.

Roder, K., Klein, H., Kranz, H., Beck, K. F. and Schweizer, M. (1994). The tripartite DNA element responsible for diet-induced rat fatty acid synthase (FAS) regulation. *Gene.* **144**:189–95.

Saggerson, D. E. (1972). The regulation of glyceride synthesis in isolated white-fat cells. The effects of acetate, pyruvate, lactate, palmitate, electron-acceptors, uncoupling agents and oligomycin. *Biochemical Journal.* **128:**1069-78.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Second Edition, Cold Spring Harbour Laboratory Press.

Sanchez, H. B., Yieh, L. and Osborne, T. F. (1995). Cooperation by Sterol Regulatory Element-binding Protein and Sp1 in Sterol Regulation of Low Density Lipoprotein Receptor Gene. *Journal of Biological Chemistry.* **270**: 1161-1169.

Sessler, A. M. and Ntambi, J. M. (1998). Polyunsaturated fatty acid regulation of gene expression. *Journal of Nutrition.* **128:**923-6.

Sessler, A. M., Kaur, N., Palta, J. P. and Ntambi J. M (1996). Regulation of stearoyl-CoA desaturase 1 mRNA stability by polyunsaturated fatty acids in 3T3-L1 adipocytes. *Journal of Biological Chemistry.* **271:**29854-29858.

Shawn, M. V., Sanchez, H. B., Rosenfeld, J. M. and Osborne, T. F. (1996). A direct role for sterol regulatory element binding protein in activation of 3-hydroxy-3-methylglutaryl coenzyme A reductase gene. *Journal of Biological Chemistry.* **271:**12247-12253.

Shimano, H., Horton, J. D., Shimomura, I., Hammer, R. E., Brown, M. S., and Goldstein, J. L. (1997) Isoform 1c of sterol regulatory element binding protein is less active than isoform 1a in livers of transgenic mice and in cultured cells. *The Journal of Clinical Investigation.*. **99:**846-854.

Shimano, H., Yahagi, N., Amemiya-Kudo, M., Hasty, A. H, Osuga, J., Tamura, Y., Shionoiri, F., izuka, Y., Ohashi, K., Har, K., Gotoda, T., Ishibashi, S. and Yamada, N. (1999). Sterol regulatory element-binding protein-1 as a key transcription factor for nutritional induction of lipogenic enzyme genes. *Journal of Biological Chemistry.* **274**:35832-9.

Shimomura, I., Shimano, H., Horton, J. D., Goldstein, J. L. and Brown, M. S. (1997). Differential expression of exons 1a and 1c in mRNAs for sterol regulatory element binding protein-1 in human and mouse organs and cultured cells. *The Journal of Clinical Investigation.*. **99:** 838-845.

Shimomura, I., Shimano, H., Korn, B. S., Bashmakov, Y. and Horton, J. D. (1998). Nuclear sterol regulatory element-binding proteins activate genes responsible for the entire program of unsaturated fatty acid biosynthesis in transgenic mouse liver. *Journal of Biological Chemistry* **273**:35299-306.

Shimomura, I., Bashmakov, Y., Ikemoto, S., Horton, J. D., Brown, M. S. & Goldstein, J. L. (1999). *Proceedings of the National Academy of Sciences (USA)* **96**: 13656-13661.

Singh, M., Richards, E. G., Mukherjee, A. and Srere, P. A. (1976). Structure of ATP citrate lyase from rat liver: Physiochemical studies and proteolytic modification. *Journal of Biological Chemistry* **251**:5242-5250 1976.

Smale, S. T. (1994). Core promoter architecture for eularyotic protein-coding genes. In *Transcription: Mechanisms and regulation, pp.63-81.* Raven Press, New York..

Smale, S. T. (1997). Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochimica et Biophysica Acta.* **1351(1-2):**73-88.

Spencer, A. F. and Lowenstein, J. M. (1962). The supply of precursors for the synthesis of fatty acids. *Journal of Biological Chemistry* **237**:3640-3648.

Srere, P. A. and Lipmann, F. (1953). An enzymatic reaction between citrate, ATP and coenzyme A. *Journal of the American Chemical Society.* **75:** 4874-4879.

Srere, P. A. (1959). The citrate cleavage enzyme: distribution and purification. *Journal of Biological Chemistry.* **234:**2544-2547.

Stryer, L. (1995). *<<Biochemistry>>. 4th edition. pp.603-628.* W.H.Freeman and Company, New York.

Towle, H. C., Kaytor, E. N., and Shih, H. M. (1997) Regulation of the expression of lipogenic enzyme genes by carbohydrate. *Annual Review of Nutrition.* **17**: 405-433.

Van der Walt, J. G. and Linington, M. J. (1989). A review of energy metabolism in producing ruminants. *Journal of the South African Veterinary Association* **60(4):**223-227.

Vernon, R. G., Robertson, J. P., Clegg, R. A. and Flint, D. J. (1981). Aspects of adipose-tissue metabolism in foetal lambs. *Biochemical Journal.* **196**:819-824.

Wagner, P. D. and Vu, N. D. (1995). Phosphorylation of ATP-citrate lyase by nucleoside diphosphate kinase. *Journal of Biological Chemistry.* **270**:21758-64.

Wang, X., Sato, R., Brown, M. S., Hua, X. and Goldstein, J. L. (1994). SREBP1, a membrane-bound transcription factor released by sterol-regulated proteolysis. *Cell.* **77**:53-62.

Wingender, E. (1994). Recognition of regulatory regions in genomic sequences. *Journal of Biotechnology.* **35(2-3):**273-80

Worgall, T. S., Sturley, S. L., Seo, T., Osborne, T. F., and Deckelbaum, R. J. (1998). Polyunsaturated fatty acids decrease expression of promoters with sterol regulatory elements by decreasing levels of mature sterol regulatory element-binding protein. *Journal of Biological Chemistry.* **273**:25537-25540.

Xiong, S., Chirala, S. S., Wakil, S. J. (2000). Sterol regulation of human fatty acid synthase promoter I requires nuclear factor-Y- and Sp-1-binding sites. *Proceedings of the National Academy of Sciences (USA).* **94:** 12975-12980 .

Xu, J., Nakamura, M. T., Cho, H. P., and Clarke, S. D. (1999). Sterol regulatory element binding protein-1 expression is suppressed by dietary polyunsaturated fatty acids. *Journal of Biological Chemistry.***274**,23577-23583.

Yahagi, N., Shimano, H., Hasty, A. H., Amemiya-Kudo, M., Okazaki, H., Tamura, Y., Iizuka, Y., Shionoiri, F., Ohashi, K., Osuga, J., Harada, K., Gotoda, T., Nagai, R., Ishibashi, S., and Yamada, N. (1999). A crucial role of sterol regulatory element-binding protein-1 in the regulation of lipogenic gene expression by polyunsaturated fatty acids. *Journal of Biological Chemistry.* **274**:35840-35844.

Young, J. W. (1977). Gluconeogenesis in cattle: significance and methodology. *Journal of Dairy Sci*ence. **60(1):**1-15.

Appendix 1

Sequence of the 5'-Region of the Bovine ATP-Citrate Lyase Gene

```
   1  CTCTGGTCAG CACCTCTCCT CTTTCACTCC TTGGATAGTA ACTTTCTTGT
  51  TGTTCTGCCC AGGATTAGTC AAACCTGGGT CCAGCAACAT CGAGATGGGG
 101  AACTTAATCA CTCAGCAAAC ACTTACTGAC TAGGTGCATG CATGCGAAGT
 151  TGCTTCAGTC ATGTCTGGCT CTTTGTGACA CTGTGGACTG TAGCCTACCA
 201  GGCTTCTCTG TCCATGGGAT TTTGCAGGCA AGAATTTTGG AGTGGGTTGC
 251  CATGCCTCCT CCAGGGGATC TTCCTGACCC AGGGATTGAA CCCACATCTC
 301  TTACATCTTC TGCCTTGGCA AGAGGGTTCT TTACCGTTAG CACCACCTGG
 351  CAAGCCCATT GACTAGGTAC CAGGCCCTAT TGAAAGTGCT GGAGAGAGCA
 401  GCCAGAAAAT ACATACATCT GCAATATATT GAAGGATAAG TTGTTAGAAA
 451  TGGAATTGCT GGGGCTGTGG ACACGTGCAT TTAAANATTT TCTCAATATT
 501  GTCAAATAGA TGTACAAAGT TTATGACAGC CTAGATGCCT GTTTCTGGAT
 551  GTGTAATTTT TTTTTAAATT TTATTGAAGT ATGGTTGATT TGAATTTTAT
 601  TGAAGAATTT TATTGAAGTA TGGTTGATGG CTGGATGGCA TCACTGACTC
 651  GATGGACATG AGTCTCAGTG AACTCCGGGA GTTGGTGATG GACAGGGAGG
 701  CCTGGCGTGC TGCGATTCGT GGGGTTGCAA AGAGTCGGAC ACGACTGAGC
 751  AACTGATCTG ATCTGATCTG ATGGTTGATT TGCAAAGTGT TAATTTCTGC
```

LINE-1

```
 801  TTACAACAAA GTGATTGTTA TACATATATA TAATTTTTTT ACAGTCTTTT
 851  CCATTATGGC TTATCAGAGG ATATTGAATA TAGTTCCCTA TGCTATACAG
 901  TAGGACCTTG TTGTTTGGCT GGTTGTATCT CTCTAACCAG GGGGATAATT
 951  GCTACACAGT AGATTCTCCC TGACTGCAGT CTTCCCACTA GCCCACAAAG
1001  ATCCTGAGTT GCTCAGATAT GCAGATGACA CCACCCTTAT GGCAGAAAGT
1051  GAAGAGGAAC TAAAAAGCCT CTTGATGAAA GTGAAAGAGG AGAGTGAAAA
1101  AGTTGGCTTA AAGCTCAACA TTCAGAAAAC AAAGATCATG GCATCTGGTC
1151  CCATCACTTC ATGGGAAATA GATGGGGAAA CAGTGGATTT ATTTTTTGGG
1201  GCTCCAAAAT CACTGCAGAT GGGACTGTAG CCATGAAATT AAAAGACGCT
1251  TACTCCTTGG AAGGAAAGTT ATGACCAACC TAGATAGCAT ATTGTAAAGC
1301  AGAGACGTTA CTTTGCCAAC AAAGGTCTGT CTAGTCAAGG CTATAGTTTT
```

1351 TCCAGTGGTC ATGTATGGAT GTGAGAGTTG GACTGTGAAG AAAGCTGAAT

1401 GCCGAAGAAT TGATGCTTTT GAACTGTGGT GTTGGAGAAG ACTCTTGAGA

1451 GTCCCTTGGA CTGCAAGGAG ATCCAACCAG TCCATTCTGA AGGAGATCAG

1501 TCCTGGGATT TCTTTGGAAG GAATGATGCT AAAGCTGAAA CTCCAGTACT

1551 TTGGCCACCT CATGCGAAGA TTTGACTCAT TGGAAAACAC TCAAATGCTG

1601 GGAAGGATTG GGGGCAGGAG ATTGGGGGCA GGAGGAGAAG GGGACGACAG

1651 TGGATGAGAT GGCTGGATGG CATCACTGAC TCGATGGACG TGAGTCTGGG

1701 TGAACTCCGG GAGTTGGTGA CGGACAGGGA GGCCTGGCGT GCTGTGATTC

1751 ATGGGGTCGC AAAGAATCGG ACATGACTGA GTGACTGAAC TGAACTGAAC

LINE-1

1801 TGAAGGCCAG AAAAGGAGTA CCTTCACAGG AGGTGTTTTC TTTAGCTGTG

1851 GAGAGGTAGT CCTTCTCCAA GATGGGTGCT ATTTGGGGAG ATTCTGAGTA

1901 GGGGTGGAGC TCATTCCATG TAGGAGTGGA AGGCTCTCTA TGACCAACAG

1951 GAGAGAGGAG CTGAGGAAAT GCTGTTTCAA AAAGATGCCT GTGAGGGAAG

2001 GGACCAGAAA CTGACATCAT CAGGGATCAC TTTGTTCAGG CATGGCCAAC

2051 AAGTGTATTG GGGGGAGAAG GATTTGGCCA GGTTTTTGTG TACAGCAATG

2101 TGGTAGCACT GTTACAGTGT GACCCCAGCG CCTGGCAGGA ACACAGACAT

2151 GTGGTCTGTT TGCACACTGC GCTGACCCTT TCACCTGAGC TGCTCAGCGC

2201 ATGATCCTCC TTCCTGGAGC TAGACAGCTG TTAGGGTCAG GGAGCAGCTC

2251 GAGAGACCTA GAGGACTGGG ATGGACATAG GAAAGCTCCA GGATGTCTCA

2301 GCAGTAAATC TATGTCTTGG TTTCTAGGAC TGTCTGATTT TAAAAATTCT

2351 GTCCCACAAG TTCACCAATA CCTGTCAGAC CAGGTGTCCC AGTTTTTGAT

2401 TTGGAAAATG CACTCTGTTT CCCCGGTAAT GGAGAAGGGG AAGGCAGCAG

2451 TTCATTCATT CCTCTAGTCT AAGCGTGAGA TAGTGGGGGC AGGGCGGAGA

GC-box

2501 GCCCGAGCAG CAGAGATGGC CCCAGCCTTG CGCTAGAATC TCTCATTGAA

2551 TTCTCTGCTT CTGTGCGCCC AGGTGAGTGG ATGAGTGGTG GTGACTGACG

2601 CGAGGTGGTG ATTGACCCCT CCTAGCCCCA CCTGCTCACG CGGTCCTCCG

2651 CCCTATCCCA TTTACAGGGG GAGGTCTCAT CTATCTCTGC TCTCGGCCTC

2701 GCTCCACCAG CTTAAGGAAG AGAGTCCACG TGCAAATGGG CAAGCAGTCT
2751 TGCAAAAAGT AGGTCTAAGC AACTGGGTTC GATGTGGCCG ATTTCAAGCT

2801 GCCAGTTTCT CCCAAAGCGA GATGGGAAGG AGACTCTGGG TCCCCAAGGC

2851 TGCGATACAG CCCATTCGGC CAACACCCCA GCCCCCACTC CCCAACACCC

GC-box

2901 CATCGCAGCG GCTCTAGGCC CACCTTCTAA GCGATCAGGC CATGGCCCCC

2951 AGCCTCGTAA GCTCCCCTCC CAGCGTAGCC CGGCCCATCT CCCCACCGGA

GC-box

3001 GGCCCCGCCC CCGACCCCTC CCTGGCTTGG CAGGCTCCGG AGAGGCGGTG

GC-box

3051 CCCGCCAGCC GTCCGAGGCA GACGCACCAG CGGGACTACA AGTTCCAGCA

3101 AACCCTGGGG CCTGGCCTCG GGGGCGGGGC TATATCAGGC AGCGAATTGG

GC-box

3151 GAGGAGGTCT GGCGCTCAGG CTAGGGAATG CGTGTGGCCA ATGGCCCGGC

SRE site Inverted Y-box

3201 GGCTCGCGCT GTGTGCCGAT GGGGGCGGGG AGAAGCCCGC CAGCGCCCGG

GC-box

3251 ACAAAAGCCA GGTCTCCGGC GGCTGCGGGA GGCTGGAGCG CTCCTAGCAG

Transcription start site

3301 TGGGCGGTGA CTTGGGTTCT GTAGACTGAA CGCCGCGCAT AGACTTTCTC

Transcription start site 2 Exon 1

3351 GGAGAGGTGA GTGGCCGACG GCTGTTTTCG CGAGATGGAC CCCACCTGGT

3401 CCAGGAGCTT CCTGCTTCCC TCCCCCAACC CGGAACCGCC TCCTTCGAAG

3451 CCAGCTTGGG GCTCACGCGC GCTCCCTGCA CCTGCTTGGA GTTTTGGGGG

3501 TCTTAGTTTT AGGTTAGCCT TCGCTGGGAT CCCGGCCGGT GGCCTCGTGT

3551 CGCCCTCCCC TCCCCCGCCC CAACTCCTCC ACATTCTAGT TCCCGGAGTT

3601 TGCACAAGGC TGGCCGCCGT TGTTGCCTGC CTTCCCTCCG CCGTATTACA

3651 TCGCCCCTAA CTGGCTCCCC TCATCCCCAC CCCGGGACTA GTCGCCGGGC

3701 CTCATCGGGC TTTCTGTGGC CACCTCTTGC CCCGATCTGG GTCCCTCCAC

3751 CCCCGCCATC GGATGCCCGG CAAAGGAGGT AGAGCGGTGG CTAGTCTCCC

3801 CACTCCTACT CCTCACCAGT CGGTTCAGGC CCTCCCTTTT AGTTCTGTGG

3851 AGAGTCGACT GAGCCCCGCG CCAAAGTGGA AGGAGCGAGC AGGGCGCGCA

3901 GCCGCCGAAG ACCACACTTG GCAGGGCCAA CCCAGGCCTC GCCACGAGGA

3951 ACTATCCTTG GATGGGCGGT GCAGGGAGGA CTCTCCTAGC TCTGGGGGTG

4001 CCGTCAGGCC CAGCCGCTAA GCGACAGATG GCCGGGGTGC CCGTCTCTCT

4051 CCCCACCCTT GGCGTTGATG CCGAGCTTGG CCGAGACAGG GTCCTCCAGG

4101 GGCTGGAAGA TGCTGCCGCC ACCCTGCTGT CGGGATTGGC AGAGGGCCAG

4151 TCCCAGCGCC TCCTGCTCTC CAACCTTGGC GGATATCAAG GAGTTCTACC

4201 ATTAGGTGGT CCTCCGGAGT CTGCCTGGGC AGAGGATTTT GCGGGGGGAA

4251 GCAGTGCATA GCACAGTTTG GTCAACCAGC CCACTCCTCT CTTCACCCAT

4301 TAACTTTGCA CTGTTTCAGC CAGCTCAGAT CATTGGATTT CTGACATTCT

4351 GAAATGTTGG GTGAAGGACA CCGGGTAGAC CCAGCTGCCC AGAGGGGTGT

4401 TTTCACTTCT TAAAGAGCTA CTGCTGCTTT TTTTTTTTTT TTTTAATATT

4451 TAAAGAGCTT CTTAAACAGC TTGGAAAATG ATAATGGGAA CAGCTCTTGA

4501 GCTCAGCTGG GTGGTAGGTG TCTCCACTCC CATTCTGCAT TTTGTTTGCT

4551 GCAGTAGGGG ACCCTCAGAA TTTTGTCCCT CTCTTTACTC CAGGGATTGC

4601 CTGGCCTTTC CAAAAAAAAA AAAAACACCA CTCCTGTGGA TACTAACCTT

4651 AAAATAAGTG CTCTTTGTCC TCCTCACAGA GACCCTACAT AGCATCCCCA

4701 GTGGTTCCTC TCCACGGGCT GTGTCCCAGG GGTACAGTCC TGGGGAACAA

4751 ACTCCTTTCA TTCAGCCTTC CCTGTGGGCT ACTTCTGCCC TCCCGTGGCT

4801 GCTGTGGATG AACTCAATTC TGGGTCACTT ATCCTGTCTG CAGAAAAGAC

4851 CCCATTCAGC TCAGAAGGGC CCTCCTTTCG ATTGAGAGAG TGTGTGTGTG

4901 TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTTCAGGTC CCCTGGCATG

4951 ATTTTCCCCA AATTTACCTA GGTAACAGTT CCTGGGACCC CAAAGTGGAC

5001 CCCTTGGATC TTAGCTCTGC CCTGAAATCA TCTCTGACCA GTCTGACAGC

5051 ATATTCAGTT ATTCAGTTAC TGGGGAGGCA GCCCAGGACT CTGCAGTGGT

5101 GTGTGTGTGT GTGTTGTGTT CCCAGCCCAG ACTGGTTTAC ACTTTAAGTC

5151 CCTTTGTCAC TTGAGCCAGT GTTCTTAGCT ATGGGTCTCT GGGCTGTCAG

5201 ATCATGGTCA CAGGTCACCA TCTTGAAGGG ATAAGCGAAA CAGGGTATGT

5251 GTACATTTTT CTGGGCTTTT CTTCTGGTGC TCGAAAGAGT CTGACCTCAA

5301 ATAGGTTAAG ATCCACCACC TTCAAGAGAC TCCCCCGAAT GATTAAACTT

5351 GTGATTTATC TCTACCCAGT CTACTCCCTG GCTTTGTGAG CCTTCTTACC

5401 CTGTCTTGAT AAATATGCAG AGGCCAGATT CTTTCCCTGT CTTTTTACTC

5451 TGTCCTGTAA ATTCTCTGTC TTGGCCTTGA GAGGGCCTGC CCCATCTATG

5501 CCAAAGTGCT CTGATGGCCA GTCTGTGTTT TGGTGAGCTG GCAGAGCTTT

5551 TCTATCTTGT GCCATTCCCC TGTTGTCAGC TCTCCCATCC CTCTCCTGAG

5601 CACCAGAGAT GAGTTCTTGG ATTTGAGATC TGAACTAAGG CTCAAGGAAA

5651 TTGTGATTTT TTTTCAGCTG CAGAAGGTCC CTTTTCTCAT CCAGGTTCTT

5701 TCCAAAGTTG GGTCTTTGTA TGGGGGTAGG GGGCGGTCGT TGTTCTGGGA

5751 GTGTCTCTCT GCTGGAGAGG CTTAACCGGG GTTGGGGGGG GCGGGGGCGG

5801 GGCGTCCTTT CTCTTGGAAT TTGGGATTAG CTTGTGAGCG ACTGAACTGA

5851 ACTGAACTGA TGTGTCCTAA CCACTAAGAG TGGTCGGCGG CTGCTCTCTC

5901 CTTTATCTGC TTGACTGCCC TGGGTTTCAT TCTGTTTGGG AGAGGGTAGG

5951 GGCACATTTA TGTACCCATC TGATTTCAGG ACCAGTTTTT CTGGTCTGAG

6001 TCAAGAAGGT TGGGAAATGG AGCTGGAGTG GGTGGGGGAT GGAGGGGTAT

6051 GGGGGAGTGG GGCTGATTGA GGCGTCTCCT TTGGAAACTC ATATTGAGAA

6101 GGAGCCGGAA GTCATCCCAG GGGAAGGCGA GAAACGGGAA CGTTTCCTCT

6151 GAGCCCCTTC CCAGCTACCC GCCTTTGGGC CAGACACCAC ACTGAATGCA

6201 CCCTTTGTGC CTTGCTGGAG GATAGCTGGG AGGCTCAGCC AGCCAGTCAA

6251 GCCAGGATTA GAACCTGAAG CCCAGCCTGG AGAATTGGAA AGGAGGTGCA

6301 GAGTGGCCCC AGAAACTCTG GCCTTCCCTG AGCTAGAAAT GGAAGACACC

6351 AACCTTCCTG AGTATGTAAA GCAACCAGCC CTTCAGCTCC ACCCCTCCTT

6401 TTTTACACTT GGCTTTTCTC TCCCAGTAAA TTCAACAACC TCAATTCTGA

6451 ATCTTTGTGA GATAGGAGTG ACTTTTCCTG ATCTGGCCAA GCTCTTGTCA

6501 GGAATGACTT AGTTTCCTTT CCCTTCCTGT CTCCCCGCCC TTTATCTCCC

6551 CAAATTGCTC CAGGCACTAC CCCAAGGGAG AAGGGGTGTG TCAGGACAAA

6601 GTGACAGGGA GGGAAGGTGA AGTTGTCATC ACAGATCCTT CTGTGAGGTG

6651 AATGGGGGAA ATGGCCTGGC TCCTGGACAA GTGGTCTTTT CTGTCAGGCT

6701 GACTGTGGGA TTTTCTAGGG ACCAGGAGTG AGAGGGTTTG AGTTCAAATC

6751 TCAGGACCCC TGCATATTAA TTAGGCTTCT TGGGCAAGTT GCCTTAACTC

6801 CTCTGAGCTT TTATTCCTCA TCTAAGAAAT AGGGCTGTAG GAACTTCCCT

6851 GGCGGTCCAG TGGTTAAGAA TCTGTGCTTC CACTGCGGAG GGCATGGGTT

6901 TGATCTCTGG TTGGGGAACT AAGGTCGTGC ATGCCACGTG GTGTGGCCAA

6951 AAATAAAAAT AAGAAATAGG ACTCTCAATC TCCATGCACC TCCTAGGGCT

7001 CACCTACATG TGACCCAAGA TGGCTTCTGC CTCTGTTCCC CTAGTTCTGG

7051 CTCGGTTAGA AATGACCGGT ATGGGTGGTT CCCAGAGAGG ACTCTGTTTT

7101 GCTCTAAATG AGTAGCTGGA TGACTCAGTG GGGTGGGTGG GATGCACTGA

7151 AGGTCACAGT TCCCCGCCCT GTGTCAGCTC TGAACTTGGG TGGGAGTCTG

7201 GTGTCCTGGG AGAAATGGGC TTTGACATAG CTCTTCCTCT TCACTCCTGG

7251 GGGGCTTGGT CCTGCGTGAG GCTAAATAGA CTTGTTTACC CAACCTTTGG

7301 GCTGGTGATG TGTGTGAGTG TGTGAAATGA GTGAGTGAGT GTGTTGGAGA

7351 CCAGACATGG AGGAGGGGTA GGGCCAGTCT TACACTATTT AACTGTGTTC

7401 CCCGCAGGTC ACAGAACTAG AGGGTGCATT TCAGGGTCTT TGTGCCCATC

7451 TCGGGTGGGA CAGGCACTCA GGTTTGGGTC ACATGCCTGC TATTATGGTC

7501 TCCCCATCCA GCTTTCTTTG CTCCAGGGTG AATGTTCGTG AATGTACTCT

7551 TGGGTGAGCT AGTTTGCCTA GAGAGTAGAG GCCACGAGTC CCCGGGGAGG

7601 TAACTGCCAT CTTTGTTTTT CTCCTCTATC CTCACCATCA TTTCCTAATC

7651 CGCACTGTCT GTGGTCAGTG CTGTTTGAGA TCTCCCTTCC ATCCTTCCTT

7701 CTGGGGCAGC CCTCTTCTCC CTCTGGAGAA GCACTGAGAT CAGGTTTTAT

7751 GGTCTGGGGG CAGCTGGGTG GGGGTCTGTC CTGCTGGAGA AGATAGCACC

7801 TCCATCTCCC CGCACCCTTT GACCAGCCCC TCTCTCCCCA ACAG**GTAGTG**

7851 **CAGGGCCTTC TGTAACCATG TCGGCCAAGG CGATTTCTGA GCAGACAGGC**

7901 **AAAGAGCTCC TCTACAAGTA CATCTGCACC ACCTCGGCCA TCCAGAACCG**

7951 **CTTCAAGTAT GCCCGGGTCA CCCCTGACAC AGACTGGGCC CGCCTGTTGC**

Translation start codon

8001 **AGGACCACCC GTGGCTGCTC AGCCAG**GTGA GCCCTCCCCG CCCCCCCACC

Exon 2

8051 CCACCCCTTC TTCAGGCCCA GATGTCCCCT CAGAACAAGT GATGGCAGAT

8101 GGCCTGCCTG GGTGGGCAGG GTGCAGGGGG CTGAGCGCGT TTTGTGTTCA

8151 GCCCTGTGTG CCGCATGCTG AGGTGCTGAG CCTGGACTGG AACTGGACAA

8201 CTGAAGTCCG GGTAGCGGCT CTCTGGTTAT TAACCATGTG GCCTGGGGCA

8251 AGCCATTTCA CCTCTCTTGA GTTTAAGGAC TAGTAATAAT GCCTGGAAAG

8301 TCTTTAGTCT GCTGCCTGGC ATTTGGTAAA TGCTCAGTAA AAAGCACCGG

8351 CTCCATCAGG AATATGTCAG GGCCATGAGG AAATGTGATG CTGTACTTTG

8401 TGTAGATGGC TGATAGAATC CACCCAGAGT AGGGGATCCT CTCTGTAGAT

8451 CCATTTAGAC GTCTCTGGCA GGTCCCAGAG AACTAGGACC AGTGAGGGGG

8501 GTGAGTGGGA GGGTGTTCAG GGATGCAATT TTCCTTCCTC TATGAGGCAG

8551 CCCAAACCCC AAACCACAAC GGGAGTTGGT GAACTTCCTG GTACTGGAGG

8601 TATGTGAGCA GTGGCTAGGG GCTGCTTGTA GATCTTGTAA ATGGAGTTCA

8651 AGTACAGGGA TAGGGTTGAC CAGATGTCTT CTATGGTGCT TTGTCACCCT

8701 AGGGCTCTGT CTTTCTAGTC TTTCTTTTTG GAATGTGGGT GCCAATCCAT

8751 ATTCTGAGAT GTTGAATGAA GCTAAGATAC CACATAGGGA GGGTAGGGAA

8801 TCCTTTACTC ATTCAGCAGC TATTCAATGA GCACCTACTC TGTGCCAGAC

8851 CCTATCCTGA GCCACGACAG CTCTAGATTC CCGGTTGTCA AAGATGTTTG

8901 GCTGGGATAG GTGAAGGAAG CTTCTGGGAG CTGACCTTGA AGATGGGTGG

8951 AATTTACCTT GGGGGTGGGG TGAATTGGAA GCCACCCAGG GTATCCTGTG

9001 CTTTGGGGCC CTAGACTATT AACTCTTAAC TTTCCCCTCT CCTTCTTTGC

9051 TCTAAATACA GACCCCTCTT GGAGATTCCC TGGTGGTCCG GGGGTTAGGA

9101 CTCCAGGCTC TCAATGCCAA GAGCCTGGGT TCAATTCCTG GTCGGAGAAC

9151 TAGATCCCAC AAGCCACATG GCTCAGCCAA AACACAAAAC AAGCAAACCA

9201 ACTCTCTTGT TCCAG**AGCTT GGTGGTCAAG CCAGACCAGC TGATCAAACG**

9251 **GCGTGGAAAG CTGGGCCTAA TTGGGGTCAA CCTCACTCTG GATGGAGTCA**

9301 **AGTCCTGGCT GAAGCCACGC CTGGGACAGG AAGCTACA**GT GAGTCTGGGT

Exon 3

9351 GTAGGGTTGG GGTGGATGTG CTGTGGGAAC AGAAGCAAAC ATGGTTGCTT

9401 TCAGATCTGC AGCCCTATGG ATCACCCCCA TGACTGGTCC GATTGCAGCC

9451 TGGGGAGGAC AGGAGTCCTG CAGAGCTGGT TTATGTAGTA

Appendix 2

Potential transcription factor binding sites in the bovine ACL promoter

— predicted by the TESS program from Baylor College of Medicine, USA

----- -------(13.3264) Sp1 Q00161
----- -----(12.6797) Sp1 S02139
===== =====(15.8496) Sp1 S02149
----- -----(12.6797) MAZ,Sp1,TBP,TFIID,Sp1 S02472, S02481
===== =====(15.8496) Sp1 S02470
---- ------(12.6797) GCF S02443
---------(12.6797) Sp1 S02473
----···(12.6973) GAGA_factor Q00066

02501 GCCCGAGCAG CAGAGATGGC CCCAGCCTTG CGCTAGAATC TCTCATTGAA

···---------(12.6973) GAGA_factor Q00066
--- ---------- --(13.3056) YY1,delta_factor S02836
= ========(12.6797) TFIID S02391

02551 TTCTCTGCTT CTGTGCGCCC AGGTGAGTGG ATGAGTGGTG GTGACTGACG

-------- --(12.6797) Sp1 S01948
----------···(12.6797) NFe S02736

02601 CGAGGTGGTG ATTGACCCCT CCTAGCCCCA CCTGCTCACG CGGTCCTCCG

···-(12.6797) NFe S02736
--------- -(12.6797) AP-1,c-Fos S01962
- ---------(12.6797) Sp1 S02206

02651 CCCTATCCCA TTTACAGGGG GAGGTCTCAT CTATCTCTGC TCTCGGCCTC

- ----------···(12.7444) GAGA_factor Q00066

02701 GCTCCACCAG CTTAAGGAAG AGAGTCCACG TGCAAATGGG CAAGCAGTCT

···--(12.7444) GAGA_factor Q00066
==========(12.6797) kappaY_factor S01733
- ---------(12.6797) NF-GMa S01909
--------- -------(13.0541) Arnt M00236
---···(12.6797) ANF-1 S03696

02751 TGCAAAAAGT AGGTCTAAGC AACTGGGTTC GATGTGGCCG ATTTCAAGCT

···-------(12.6797) ANF-1 S03696

02801 GCCAGTTTCT CCCAAAGCGA GATGGGAAGG AGACTCTGGG TCCCCAAGGC

=== ======(14.2647) LyF-1 S03757
=== ======(13.5660) LyF-1 M00141

02851 TGCGATACAG CCCATTCGGC CAACACCCCA GCCCCCACTC CCCAACACCC

====== ===(12.6797) TFIID S02391
----- -----(12.6797) Sp1 S01230
------- ---(12.6797) NF-kappaB,Sp1 S02466

02901 CATCGCAGCG GCTCTAGGCC CACCTTCTAA GCGATCAGGC CATGGCCCCC

=====···(12.6797) TFIID S02391

02951 AGCCTCGTAA GCTCCCCTCC CAGCGTAGCC CGGCCCATCT CCCCACCGGA

···=====(12.6797) TFIID S02391
--- -------(12.6797) Sp1 S02472, S03412
---- ------(12.6797) SREBP-1 S00722
=···(17.6371) Sp1 Q00161
=···(16.3337) V$GC_01 M00255

03001 GGCCCCGCCC CCGACCCCTC CCTGGCTTGG CAGGCTCCGG AGAGGCGGTG

···==========-(17.6371) Sp1 Q00161
···==========-===(16.3337) V$GC_01 M00255
---------(12.6797) GCF S02443
==========-===(17.3728) V$SP1_Q6 M00196
--------- -(12.6797) Sp1 S01230, S01861, S02130, S02141, S02149, S02238
=========-(12.0403) Sp1 M00008
=========-(15.8496) Sp1 S02139
=========-===(16.5216) Sp1 Q00158
-------- -(12.6797) Sp1,octamer-binding_factor,Sp1 S01941, S02470, S02472, S02474, S02476, S02483, S03412

=(15.8496) MAZ,Sp1,TBP,TFIID S02481
--(12.6797) HiNF-C,Sp1 S01946, S02489
=(14.2647) Sp1 S02480
=(12.6797) Sp1 S02478
=(12.6797) Sp1 S01251
----(12.6797) EGR2 S01289
==(12.6797) AP-1,Sp1 S01287
==(12.6797) Sp1 S02492
-------(12.6797) Sp1 S01230
----(12.6797) CACCC-binding factor S01214
···(12.6797) Sp1,octamer-binding factor S01941
···(12.6797) Sp1 S02239

03051 CCCGCCAGCC GTCCGAGGCA GACGCACCAG CGGGACTACA AGTTCCAGCA

···-(12.6797) Sp1,octamer-binding factor S01941
···-(12.6797) Sp1 S02239

03101 AACCCTGGGG CCTGGCCTCG GGGGCGGGGC TATATCAGGC AGCGAATTGG

------(12.6797) Sp1 S02472
-------(12.6797) EGR2 S01289
==========(16.5216) Sp1 Q00158
==========--(15.4335) V$GC_01 M00255
==========-(16.8101) V$SP1_Q6 M00196
=======(12.6797) Sp1 S02492
=======(12.6797) AP-1,Sp1 S01287
---------(12.6797) HiNF-C,Sp1 S01946, S02489
==========--(15.4847) Sp1 Q00161
---------(12.6797) Sp1,octamer-binding factor,Sp1 S01941, S02470, S02472,

S02474, S02476, S02483, S03412

========(12.6797) Sp1 S02478
========(14.2647) Sp1 S02480
==========(15.8496) Sp1 S02139
==========(15.8496) MAZ,Sp1,TBP,TFIID S02481
----------(12.6797) Sp1 S01230, S01861, S02130, S02141, S02149, S02238
==========(12.0403) Sp1 M00008
========(12.6797) Sp1 S01251
--------- -(12.6797) GCF S02443
---------(12.6797) En S01128

03151 GAGGAGGTCT GGCGCTCAGG CTAGGGAATG CGTGTGGCCA ATGGCCCGGC

========-----(12.6959) V$CAAT_01 M00254
------ -----(12.4782) V$NFY_Q6 M00185
···(12.6797) GCF S02197

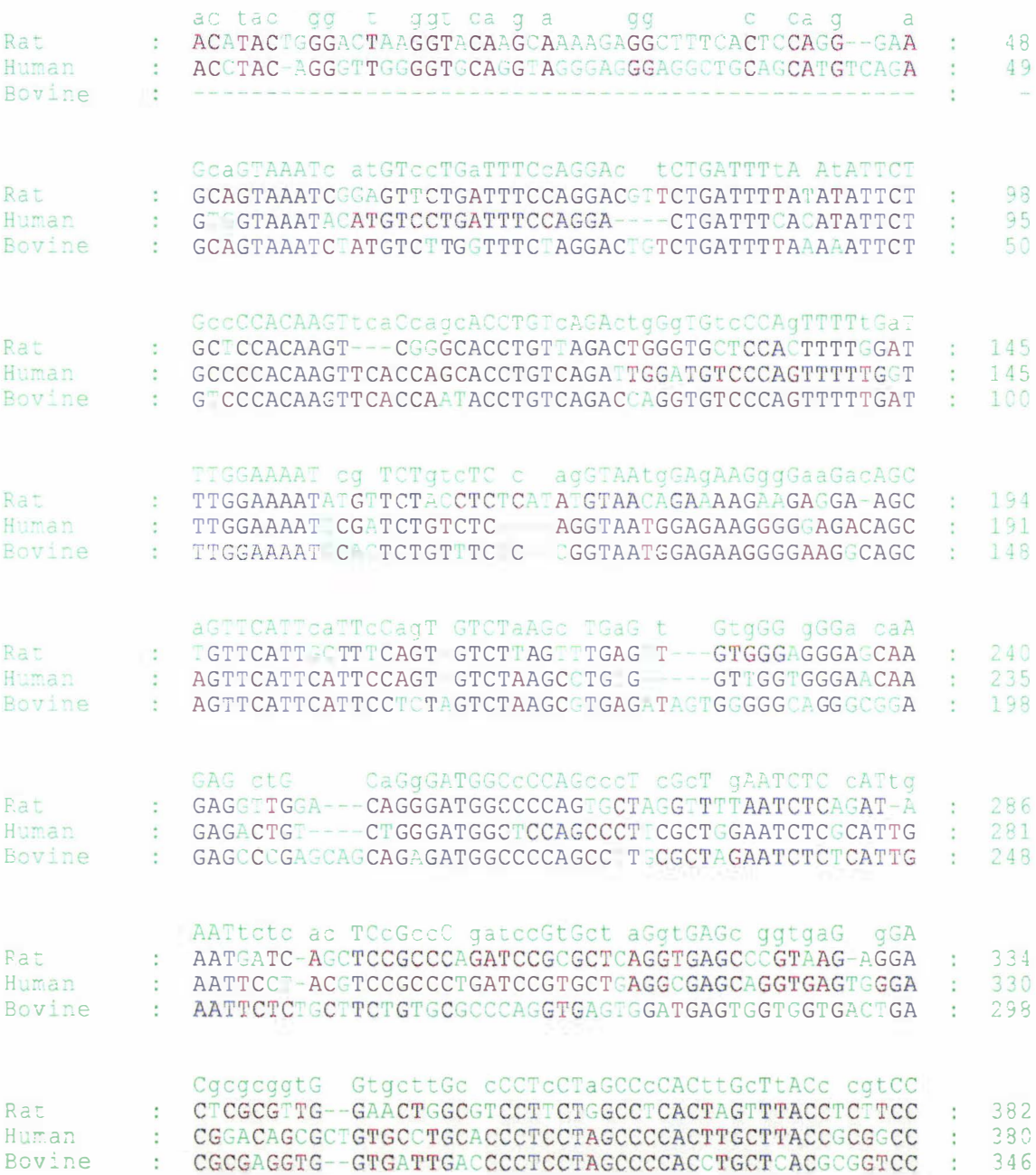
03201 GGCTCGCGCT GTGTGCCGAT GGGGGCGGGG AGAAGCCCGC CAGCGCCCGG

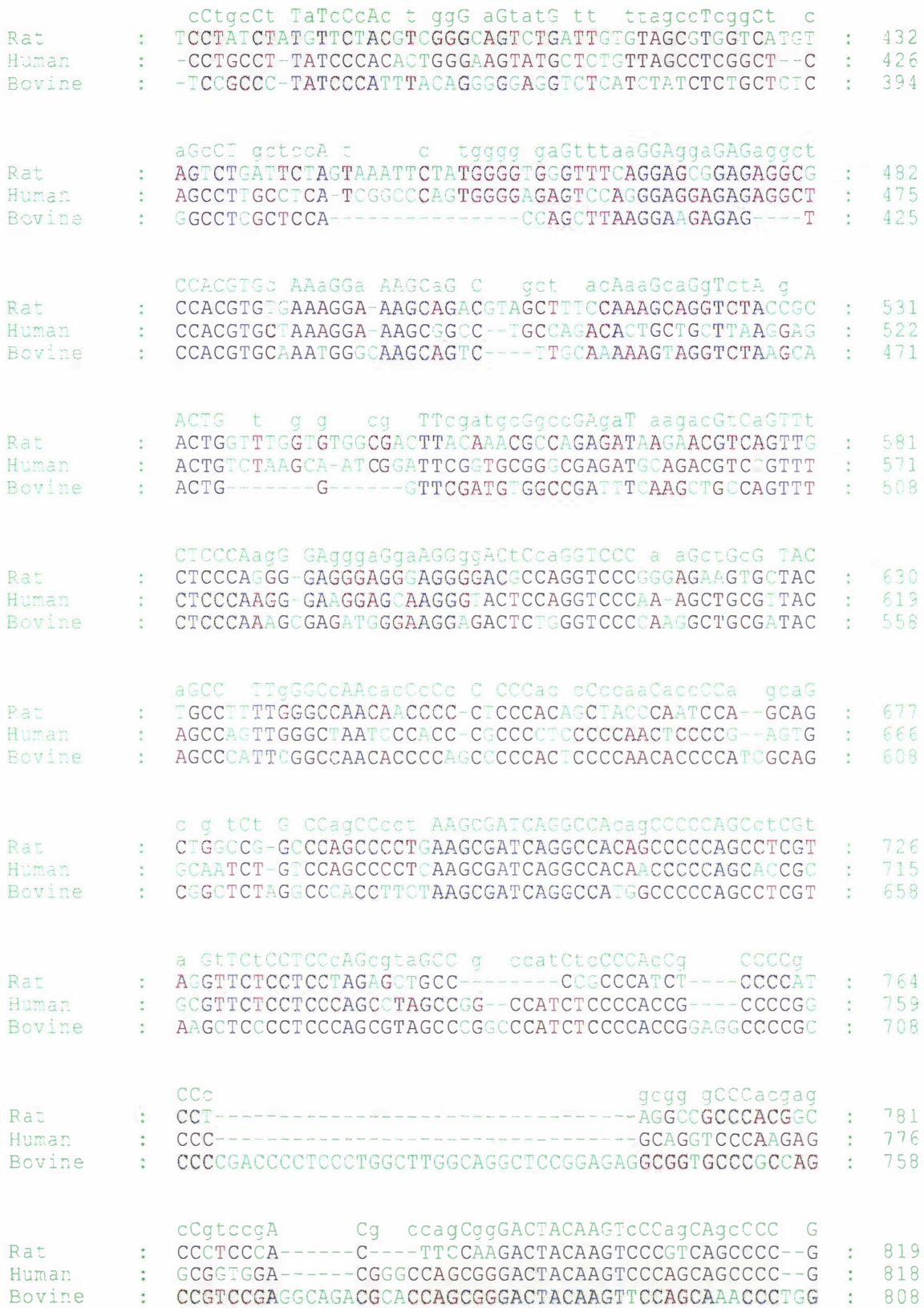
···---------(12.6797) GCF S02197
---------- -(14.8265) Sp1 Q00158
---------- --(12.8961) V$SP1_Q6 M00196
========(12.6797) Sp1 S02492
---------(12.6797) HiNF-C,Sp1 S01946, S02489
========(12.6797) AP-1,Sp1 S01287
--------- -(12.6797) Sp1 S01230, S02139
========= =(12.5708) Sp1 M00008
========(12.6797) Sp1 S01251
--------- -(12.6797) MAZ,Sp1,TBP,TFIID,Sp1 S02473, S02481, S02483
--------- ---(12.3580) Sp1 Q00161
==========(12.6797) Sp1 S02478
==========(14.2647) Sp1 S02480
-------- --(12.6797) GCF S02443

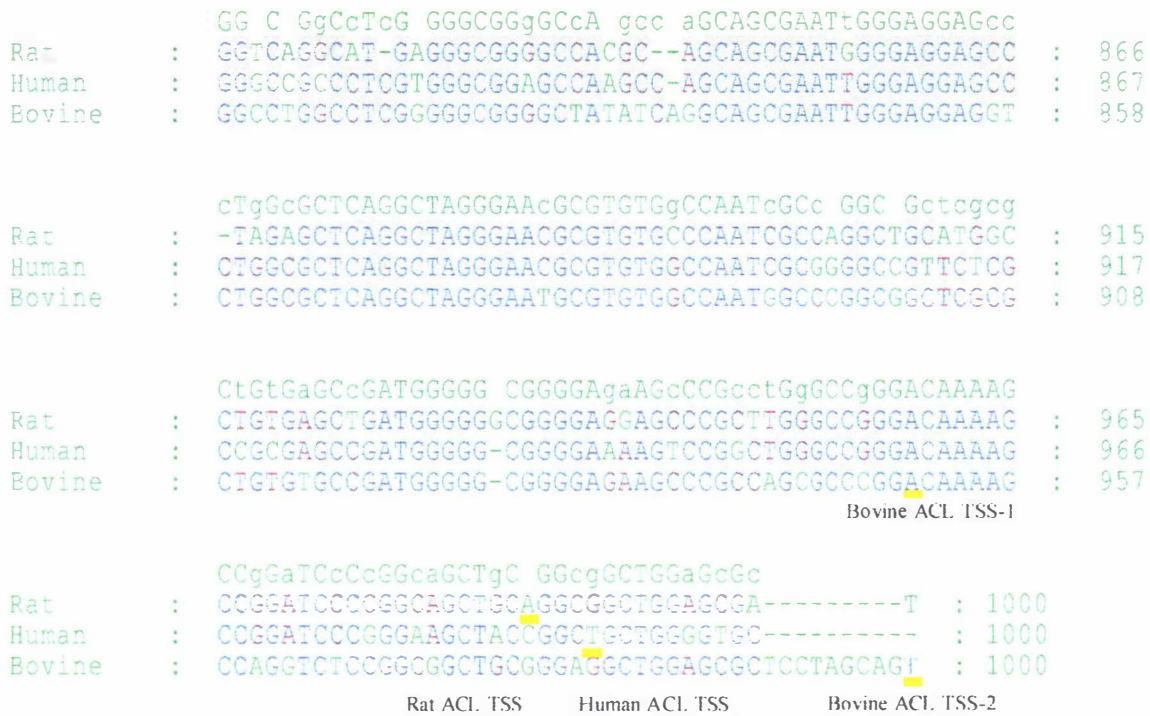
--- --------(14.2647) DI S02536
------- ----(14.2647) AP-2 S00530
---------···(12.6797) GCF S03087
03251 ACAAAAGCCA GGTCTCCGGC GGCTGCGGGA GGCTGGAGCG CTCCTAGCAG
···-(12.6797) GCF S03087
------- -----(12.6594) repressor of CAR1 expression Q00155
= ========== =(12.4346) V$AP2_Q6 M00189
---···(14.2647) Sp1 S01050
03301 TGGGCGGTGA CTTGGGTTCT GTAGACTGAA CGCCGCGCAT AGACTTTCTC
···--------(14.2647) Sp1 S01050
------ -----(14.2647) ssDBP-1,ssDBP-2 S01143
03351 GGAGAGGTGA GTGGCCGACG GCTGTTTTCG CGAGATGGAC CCCACCTGGT
----- -----(12.6797) NF-kappaB.Sp1 S02466
----- ----------···(15.7386) LF-A1 S03376
03401 CCAGGAGCTT CCTGCTTCCC TCCCCCAACC CGGAACCGCC TCCTTCGAAG
···--(15.7386) LF-A1 S03376
-------- -------(14.5897) Elk-1 M00007
------- ---(12.6797) TIN-1 S00557
----- ----(12.6797) c-Ets-1_68 S02074
===== =====(12.3425) c-Ets-1_68 Q00026
---- ------(12.6797) Sp1 S02489
=== =======(14.2647) H4TF-1 S02487
-- --------(12.6797) MPBF,NF-1 S01158
03451 CCAGCTTGGG GCTCACGCGC GCTCCCTGCA CCTGCTTGGA GTTTTGGGGG
===== ========(14.7480) I$SN_01 M00060
---- ---------(12.1367) Sn M00044

Appendix 3

Comparison of the Promoters From the bovine, Human and Rat ATP-Citrate Lyase genes.




cCtgcCt TaTcCcAc t ggG aGtatG tt ttagccTcggCt c
Rat : TCCTATCTATGTTCTACGTCGGGCAGTCTGATTGTGTAGCGTGGTCATGT : 432
Human : -CCTGCCT-TATCCCACACTGGGAAGTATGCTCTGTTAGCCTCGGCT--C : 426
Bovine : -TCCGCCC-TATCCCATTTACAGGGGAGGTCTCATCTATCTCTGCTCTC : 394
aGcCT gctccA t c tgggg gaGtttaaGGAggaGAGaggct
Rat : AGTCTGATTCTAGTAAATTCTATGGGGTGGGTTTCAGGAGCGGAGAGGCG : 482
Human : AGCCTTGCCTCA-TCGGCCCAGTGGGGAGAGTCCAGGGAGGAGAGAGGCT : 475
Bovine : GGCCTCGCTCCA---------------CCAGCTTAAGGAAGAGAG----T : 425
CCACGTGc AAaGGa AAGCaG C gct acAaaGcaGgTctA g
Rat : CCACGTGTGAAAGGA-AAGCAGACGTAGCTTTCCAAAGCAGGTCTACCGC : 531
Human : CCACGTGCTAAAGGA-AAGCGGCC--TGCCAGACACTGCTGCTTAAGGAG : 522
Bovine : CCACGTGCAAATGGGCAAGCAGTC----TTGCAAAAAGTAGGTCTAAGCA : 471
ACTG t g g cg TTcgatgcGgccGAgaT aagacGtCaGTTt
Rat : ACTGGTTTGGTGTGGCGACTTACAAACGCCAGAGATAAGAACGTCAGTTG : 581
Human : ACTGTCTAAGCA-ATCGGATTCGGTGCGGGCGAGATGCAGACGTCCGTTT : 571
Bovine : ACTG-------G------GTTCGATGTGGCCGATTTCAAGCTGCCAGTTT : 508
CTCCCAagG GAgggaGgaAGGggACtCcaGGTCCC a aGctGcG TAC
Rat : CTCCCAGGG-GAGGGAGGGAGGGGACGCCAGGTCCCGGGAGAAGTGCTAC : 630
Human : CTCCCAAGG-GAAGGAGCAAGGGTACTCCAGGTCCCAA-AGCTGCGTTAC : 619
Bovine : CTCCCAAAGCGAGATGGGAAGGAGACTCTGGGTCCCCAAGGCTGCGATAC : 558
aGCC TTgGGCcAAcacCcCc C CCCac cCccaaCaccCCa gcaG
Rat : TGCCTTTTGGGCCAACAACCCC-CTCCCACAGCTACCCAATCCA--GCAG : 677
Human : AGCCAGTTGGGCTAATCCCACC-CGCCCCTCCCCCAACTCCCCG--AGTG : 666
Bovine : AGCCCATTCGGCCAACACCCCAGCCCCCACTCCCCAACACCCCATCGCAG : 608
c g tCt G CCagCCcct AAGCGATCAGGCCAcagCCCCCAGCctCGt
Rat : CTGGCCG-GCCCAGCCCCTGAAGCGATCAGGCCACAGCCCCCAGCCTCGT : 726
Human : GCAATCT-GTCCAGCCCCTCAAGCGATCAGGCCACAACCCCCAGCACCGC : 715
Bovine : CGGCTCTAGGCCCACCTTCTAAGCGATCAGGCCATGGCCCCCAGCCTCGT : 658
a GtTCtCCTCCcAGcgtaGCC g ccatCtcCCCAcCg CCCCg
Rat : AGGTTCTCCTCCTAGAGCTGCC--------CCGCCCATCT----CCCCAT : 764
Human : GCGTTCTCCTCCCAGCCTAGCCGG--CCATCTCCCCCACCG----CCCCGG : 759
Bovine : AAGCTCCCCTCCCAGCGTAGCCCGGCCCATCTCCCCACCGGAGGCCCCGC : 708
CCc gcgg gCCCacgag
Rat : CCT----------AGGCCGCCCACGGC : 781
Human : CCC----------GCAGGTCCCAAGAG : 776
Bovine : CCCCGACCCCTCCCTGGCTTGGCAGGCTCCGGAGAGGCGGTGCCCGCCAG : 758
cCgtccgA Cg ccagCggGACTACAAGTcCCagCAgcCCC G
Rat : CCCTCCCA------C----TTCCAAGACTACAAGTCCCGTCAGCCCC--G : 819
Human : GCGGTGGA------CGGGCCAGCGGGACTACAAGTCCCAGCAGCCCC--G : 818
Bovine : CCGTCCGAGGCAGACGCACCAGCGGGACTACAAGTTCCAGCAAACCCTGG : 808



The identity of the three sequences together is 63%. The transcription start sites (TSS) for each gene are underlined.