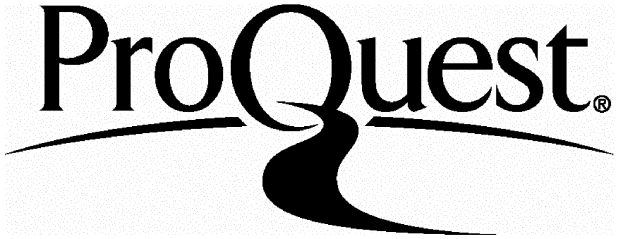
The Isolation and Function of the 3'Untranslated Region of The Myosin Heavy Chain Genes of Skeletal Muscle

Arpna Kiri BSc, MSc

A Dissertation submitted to the Board of Studies of the University of London for the Degree of Doctor of Philosophy

April 2000



ABSTRACT

In recent years the 3'untranslated regions (3'UTRs) of eukaryotic mRNAs have emerged as a repository of signals that determine mRNA localization, stability, translation, and cytoplasmic polyadenylation. Furthermore, these regions bind trans-acting factors to form specific complexes that control gene expression at the post-transcriptional level. This thesis describes the investigations of the functional role of the 3'untranslated region of four murine skeletal myosin heavy chain (MyHC) mRNAs. 3' RACE was used firstly to isolate and clone the slow type 1, and fast 2a, 2b, and 2x MyHC 3'UTRs. Sequence analysis revealed not only the isotypic conservation of the 3'UTRs but also two conserved motifs in the fast 3'UTRs which could be involved in regulation of gene expression or act as putative binding sites for proteins.

The results of transfection experiments with hybrid CAT/MyHC 3'UTRs plasmids showed that, after taking into account the effect of deleting the CAT 3'UTR, the MyHC 3'UTRs had no major effects on translation and stability of reporter mRNA in transfected C2C12 cells.

RNA-protein interactions of the MyHC 3'UTRs were investigated by bandshift analyses. Results demonstrated that the MyHC 3'UTR-protein complexes were muscle-specific. Competition assays to determine sequence specificity showed that binding could be competed out by unlabelled MyHC 3'UTR and polyadenylic acid but not rat growth hormone 3'UTR. Deletion analysis suggests that secondary structure is important for protein binding. Purification of the protein from shifted complexes identified a 33-kDa protein, aldolase A, as a MyHC 3'UTR-binding protein.

These results and the role of aldolase A in the post-transcriptional regulation of MyHC are discussed.

ACKNOWLEDGEMENTS

First and foremost I would like to thank Professor Goldspink for giving me the opportunity to study for a PhD. He has been very patient and understanding during the course of my studies and has been there to encourage me and give me confidence. I would also like to thank Dr Godfrina McKoy, who not only provided a good role model for hard work and self-belief, but also encouraged me and helped me with certain aspects of this PhD project. I must also thank Dr Nick Totty for his assistance with the protein sequencing.

I am grateful to all my colleagues past and present in the Department of Anatomy for their friendship and support. Thanks also to Ann Chaplin, our secretary, for the administrative help and, more importantly, for her friendship.

Most of all I would like to thank my family for their support and encouragement throughout the course of my studies.

This PhD was funded by a grant from the Wellcome Trust.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ABBREVIATIONS

μM	micromolar
3' UTR	3' untranslated region
A	adenosine
ATP	adenosine triphosphate
bp	base pairs
BSA	bovine serum albumin
CAT	chloramphenicol acetyl transferase
c.p.m	counts per minute
cDNA	complementary DNA
CIAP	calf intestinal alkaline phosphatase
CPSF	cleavage nd polyadenylation specificty factor
CstF	cleavage stimulation factor
CTP	Cytosine triphosphate
DEAE	diethyl aminoethyl
DEPC	diethyl polycarbonate
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulphoxide
DNase	deoxyribonuclease
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
GTP	guanosine triphosphate
HEPES	4-(2hydroxyethyl)-1-piperazineethanesulphonic acid
IPTG	isopropyl-β-D-thiogalactosidase
kb	kilobases
kDA	kiloDalton
M	molar

MALDI	**matrix-assisted laser desorption**
mM	**millimolar**
MMu-LV	**Maloney murine Leukemia virus**
mRNA	**messenger RNA**
MyHc	**myosin heavy chain**
nm	**nanometers**
nM	**nanomolar**
nt	**nucleotide**
OD	**optical density**
ONPG	***o*-nitrophenyl-β-glactopyranoside**
PABP	**poly A-binding protein**
PAGE	**polyacrylamide gel electrophoresis**
PBS	**phosphate-buffered saline**
PCR	**polymerase chain reaction**
PMSF	**phenylmethylsulphonyl fluoride**
PNK	**polynucleotide kinase**
poly A	**polyadenylic acid**
RACE	**rapid amplification of cDNA ends**
RNase	**ribonuclease**
rNTP	**ribonucleotide triphosphate**
RT-PCR	**reverse transcription-polymerase chain reaction**
SDS	**sodium dodecyl sulphate**
SDS-PAGE	**SDS-polyacrylamide gel electrophoresis**
snRNP	**small nuclear ribonucleoprotein particle**
SSC	**standard saline citrate**
SV40	**simian virus 40**
TBE	**Tris, borate, EDTA**
TE	**Tris-HCl, EDTA**
TEMED	**tetraethylmethlenediamine**
TMPD	**tetramethylpentadecane**
TTP	**thymidine triphosphate**

tRNA	**transfer RNA**
UV	**Ultra violet**
X-gal	**5-bromo-4-chloro-3-indolyl-β-D-galactoside**

Chapter 1

General Introduction

1.1. Introduction

During eukaryotic cell differentiation qualitative and quantitative changes occurring in cytoplasmic mRNA result in similar changes in protein composition within the cell. Differential processing of the same genetic information ultimately determines the distinct differentiated phenotype of each cell (Darnell, 1979; Darnell, 1982; Lewin, 1975a; Lewin 1975b). In order to achieve this, specific regulation of gene expression is required. Gene expression is a complex process involving transcritpional, post-transcriptional, translational and post-translational events, which culminate in the production and accumulation of protein product (figure 1.1).

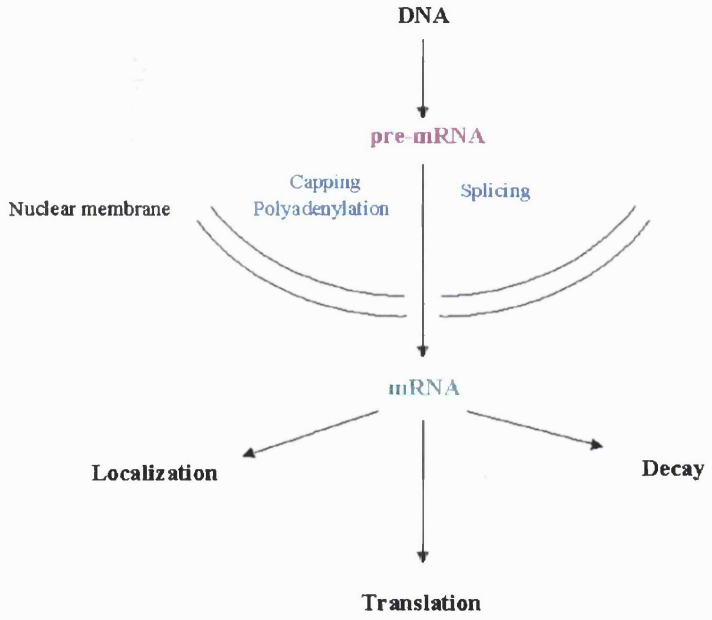


Figure 1.1. Schematic diagram showing the steps involved in transcription of DNA into RNA and translation into protein. The initial product of transcription is pre-mRNA. The pre-mRNA is capped, spliced, cleaved and polyadenylated before being transported out of the nucleus. Once in the cytoplasm the mature mRNA may be translated into protein or degraded. Some mRNAs may be localized.

Initially it was thought that this specific regulation was controlled most frequently at the level of transcription and translation with translational control

resulting from specific RNA-protein interactions at or near the 5' end of the mRNA. However, it is now clear that there are several mechanisms of regulation at the post-transcriptional level. This thesis is concerned with post-transcriptional regulation of gene expression specifically by the 3' untranslated region of mRNAs.

The primary transcript (pre-mRNA) is not the template that the translational complex scans in order to produce the corresponding protein. Pre-mRNAs undergo several modifications (capping, methylation, polyadenylation, splicing) within the nucleus prior to becoming mature mRNAs. The mature mRNAs are then exported through the nuclear membrane. The mature mRNA, 5' to 3', consists of a 5' cap structure followed by an untranslated region, the coding region, another untranslated region and finally a poly A tail (figure 1.2).

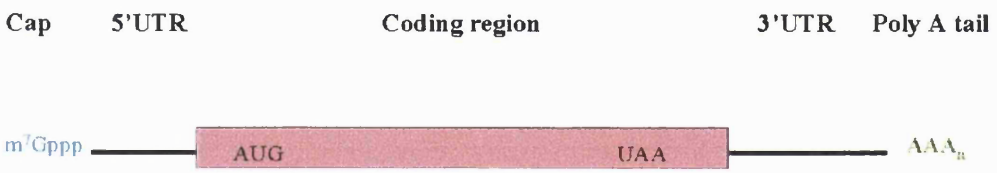


Figure 1.2. Schematic representation of mature eukaryotic mRNA. The mRNA has a cap structure and poly A tail at its respective 5' and 3' ends. The coding region which is translated into to protein (boxed) is flanked by 5' and 3' untranslated regions.

The 5' untranslated regions (5'UTRs) of mRNAs are usually short, whereas the 3'UTRs are very variable, and can be very long. The average size of the 5'UTR as studied by comparisons in various taxonomic groups is 200 nucleotides are one and a half to three times shorter than corresponding 3'UTRs (Pesole *et al*, 1997).

Comparisons of nucleotide sequences from different vertebrate species by Duret *et al*, (1993) revealed the existence of highly conserved regions with more than 70% similarity in the non-coding regions of several genes. Although there was no significant homology between the non-coding regions of most orthologous genes, highly conserved regions were found to be situated predominantly within the 3'UTR, indicating a more selective pressure to preserve the sequence of the 3' non-coding region. Thus, there is more constraint on the mature mRNA rather than on the pre-mRNA or DNA inferring that the highly conserved regions are involved in post-transcriptional mechanisms. Yaffe *et al*, 1985 were the first to show that portions of the 3' untranslated regions (3'UTRs) of skeletal muscle alpha-actin and cytoplasmic beta-actin had been preserved since the divergence of mammals and birds, and that the conservation was specific for each actin isotype. From their analyses Duret *et al*, (1993) have established that highly conserved regions within the 3'UTR do not code either for proteins or any known structural RNAs such ribosomal and transfer RNAs. Many of the genes that contain highly conserved regions within the 3'UTR belong to multigene families where they code for isoforms that are similar at the protein level, but the non-coding regions are specific for each isotype. The MyHC is encoded by a multigene family whose isoforms are very similar at the protein level, but whose 3'UTRs are highly divergent. The 3'UTRs of many genes are now known to be involved in post-transcriptional mechanisms of regulation. From work published to date, this region of mRNA has been shown to act as a repository of signals controlling polyadenylation, determining mRNA localization, regulating mRNA stability, as well as of signals controlling translation (see reviews Ross, 1995; St.Johnston, 1995; Sonenberg, 1994). In many cases regulation by the 3'UTRs involves protein-RNA interactions (see review Standart & Jackson, 1994).

This thesis is concerned with the role of the 3'UTR of the myosin heavy chain mRNAs. Myosin heavy chain (MyHC) is the major contractile protein in muscle. The mechanisms of MyHC gene expression regulation have been postulated to occur at both transcriptional (Ordahl & Caplan, 1976) and post-transcritpional levels (Leibovitch *et al*, 1978; Affara *et al*, 1980; Medford *et al*, 1983), and thus it is conceivable that some post-transcriptional regulation may occur through the 3'UTR.

In order to understand how the 3'UTR is involved in regulating gene expression it is necessary to explain how RNA is processed at the 3'end. This introduction will present the current understanding of 3' end processing of mRNA with a focus on how the untranslated regions are involved in regulating gene expression. Emphasis is placed on the known functional roles of the 3'UTR in the regulation of gene expression. In the latter part of this introduction the structure and function of the myosin heavy chain is described.

1.2. The 3'untranslated region.

1.2.1. 3' end processing of RNA

Early sequence analysis experiments on a few easily purified mRNAs revealed that these eukaryotic mRNAs have a polyadenylic acid sequence at their 3' ends (Lim & Canellakis, 1970; Edmonds *et al*, 1971; Lee *et al*, 1971). These poly A tails are not encoded in the DNA, but are added to the mRNA post-transcriptionally (Birnboim *et al*, 1973). Subsequently it was shown that the vast majority of eukaryotic mRNAs possess a 3' polyadenylate tail that is added in a post-transcriptional reaction. The only eukaryotic mRNAs known to lack a poly A tail are those encoding the major histones (Adesnik & Darnell, 1972; Adesnik *et al*, 1972; Greenberg & Perry, 1972a; Greenberg & Perry, 1972b). Analyses of viral mRNA synthesis in infected cells first demonstrated that transcription proceeds past the polyadenylation site, suggesting that the 3' end receiving the poly A tail is generated by nuclease action and not transcription termination (Nevins & Darnell, 1978; Ford & Hsu, 1978). In eukaryotic cells, polyadenylation occurs in the nucleus and involves endonucleolytic cleavage of the primary transcript at the site of poly A addition, followed by addition of the poly A tail (Moore & Sharp, 1984; Moore & Sharp 1985; Wahle, 1992; Wahle & Keller, 1992; Keller, 1995). The cleavage of the primary transcripts is directed by specific RNA sequences upstream and downstream of the site of poly A addition. They are the hexanucleotide sequence AAUAAA, the nucleotide to which the poly A is added (normally an adenosine), and sequences downstream of this nucleotide. The AAUAAA hexamer, known as the poly A signal, is typically located 10-35 bases upstream of the polyadenylation site. The sequence is highly conserved (Proudfoot & Brownlee, 1976) and is important for 3'end formation. Systemic mutagenesis of this sequence *in vitro* has shown that variations from the AAUAAA sequence greatly reduce 3'end formation and addition of poly A at the poly A site. Furthermore, in RNAs that extend beyond the cleavage site, cleavage is prevented (Montell *et al*, 1983; Sheets *et al*, 1990; Wickens & Stephenson, 1984). However, there are some RNAs that do not contain a recognizable AAUAAA element. The majority of these RNAs involve alternative

polyadenylation. For example, of the four major mRNAs derived from the mouse dihydrofolate reductase gene, one lacks the AAUAAA element as defined above, and two others have unusual variants (Setzer *et al*, 1980; Hook & Kellems, 1988).

In addition to the poly A signal, downstream elements are also required for 3' end cleavage (McDevitt *et al*, 1984; Gil & Proudfoot, 1984; Sadofsky & Alwine, 1984). They are more diffuse than the poly A signal but are generally rich in U or G and U residues (McLauchlan *et al*,1985). Furthermore, the position of these elements with respect to AAUAAA is precise, and an abnormal separation of the two elements abolishes 3' end formation (McDevitt *et al*, 1986). Gil & Proudfoot, (1987) demonstrated that the GU- and U-rich sequence elements were both required for efficient rabbit β mRNA processing. If only one element was present the efficiency of processing was greatly diminished, and the level of 3' end formation was also decreased when the distance between the two elements was expanded. Therefore, the GU- and U-rich elements function synergistically in the 3' end formation of globin mRNA.

The endonucleolytic cleavage and addition of poly A are tightly coupled and involve a cleavage and polyadenylation specificity factor (CPSF) (Whale & Keller, 1992; Jenny *et al*, 1994) and a cleavage stimulation factor (CstF). The CPSF binds specifically to the AAUAAA signal. CstF binds to the GU/U rich elements (MacDonald *et al*, 1994). These two factors act synergistically; the simultaneous binding of CPSF is required for efficient binding of CstF, and the interaction of CPSF with the RNA is stabilized by CstF. Cleavage factors present within the 3' processing complex carry out the endonucleolytic cleavage step. Poly A polymerase, which catalyzes the synthesis of the poly A tail, is also required for the cleavage of most pre-mRNAs. The downstream fragment is degraded, CPSF and poly A polymerase remain bound to the upstream cleaved fragment and add a short tract of about ten adenosine residues. This activates poly A polymerase, which by itself has a low affinity for RNA and is unable to recognize the pre-mRNA substrate specifically, to elongate the poly A tail (Bienroth *et al*, 1993). The polyadenylation complex is joined by another component, poly A binding protein (PABP). This protein binds to the poly a tail in the polyadenylation complex (Wahle & Keller,

1992) and stimulates a rapid burst of processive synthesis of a poly A tail of approximately 250 nucleotides. This elongation requires the simultaneous presence of pre-mRNA that has a wild type poly A signal, CPSF, poly A polymerase and PABP in a quaternary complex that transiently stabilizes the binding of the poly A polymerase to the RNA 3' end (Bienroth *et al*, 1993). In the absence of CPSF, poly A polymerase will non-specifically polyadenylate an RNA substrate with a 3'OH, but in the presence of CPSF, poly A polymerase activity is stimulated to specifically polyadenylate those RNA's containing an intact AAUAAA sequence. Therefore, only a complex formed from all three proteins is competent for the processive synthesis of a full-length poly A tail. The growing poly A tail is covered by additional molecules of PABP and the length of the poly A tail is presumed to be controlled by interruption of the protein-protein interactions between CPSF, poly A polymerase and PABP, where the 3' end of the poly A tail is no longer sequestered in the protein-RNA complex, when the correct length has been reached (Wahle, 1995). There is some evidence to suggest that PABP also functions to protect mRNA from rapid degradation. When de-proteinized β-globin mRNA was incubated in a cell-free mRNA decay reaction mixture with polysomes, the half-life is greater than 60 minutes. However, if the polysomes were depleted of PABP, the mRNA was rapidly degraded (Bernstein *et al*, 1989). Addition of purified PABP to the PABP-depleted reaction mixtures restabilized the mRNA. Therefore the PABP complex formation with the poly A tail acts to protect polyadenylated mRNA from degradation.

Since its discovery, several hypotheses have been put forward as to the function of the poly A tail. Among the candidate functions were mRNA transport from the nucleus, the control of mRNA stability (Bernstein & Ross, 1989), compartmentalization of mRNA in the cytoplasm, and even a mechanism for the control of translation (Brawerman, 1981). It is likely that the poly A tail has multiple functions that affect nuclear processing of pre-mRNA, transport to the cytoplasm and translation, as well as mRNA stability. It seems to be required for some of the splicing events in RNA processing. In the cytoplasm the poly A tail is a target for the degradation process. Complete loss of the poly A tail is probably followed by

1.2.2. The 3'UTR and control of translation.

The regulation of gene expression at the level of translation is of critical importance during early embryonic development and during cell growth and differentiation. In recent years a vast number of works has led to a growing list of mRNAs that contain sequence motifs within their 3'UTR that are responsible for their positive and negative translational regulation. The translation of some mRNAs is determined by the length of the poly A tail, which in turn is specified by sequences present in the 3'UTR. The translational regulation of such mRNAs and other mRNAs mediated by sequence elements that reside in the 3'UTR are given below.

1.2.2.1. Translation dependent on polyadenylation.

Polyadenylation can be a target for controlling gene expression. For example the human U1 small nuclear ribonucleoprotein (SnRNP)-specific protein U1A autoregulates its production by binding its own pre-mRNA and inhibiting polyadenylation (Gunderson *et al*, 1994). The U1 SnRNP particle functions in splicing of pre-mRNA (reviewed by Green, 1991; Guthrie, 1991). U1A binds to the second hairpin loop of U1 SnRNA with a high specificity (Scherly *et al*, 1989). It also binds with high affinity in a specific manner to an RNA sequence in the 3'UTR of the U1A pre-mRNA (Boelens *et al*, 1993). This region which is conserved in vertebrates, consists of two short sequences, one a perfect match and the other a single mis-match to the sequence AUUGCAC found in the second hairpin loop of U1 SnRNA. These sequences are necessary for specific binding of two U1A protein molecules. The bound U1A protein inhibits polyadenylation of the U1A pre-mRNA through specific interaction with poly A polymerase and this in turn, down-regulates U1A mRNA accumulation.

There are many examples where polyadenylation of a particular mRNA is tightly correlated with activation of its translation. This is seen in a wide range of organisms from Xenopus to mouse. The majority of these examples relate to developmental systems, although it is not particular just to development. Early work with clam oocytes demonstrated a correlation between the length of the poly A tail of

a given mRNA and its translation. mRNAs that underwent elongation of their poly A tail after fertilization were recruited onto polysomes, whereas those that were deadenylated were no longer present on polysomes (Rosenthal *et al*, 1983, Rosenthal & Ruderman, 1987). In particular, the mRNAs for ribonucleotide reductase and cyclins A and B are stored untranslated in abundance in oocytes with short poly A tails. Regulation by the 3'UTR was demonstrated by *in vitro* translation experiments. The translation of these mRNAs remained tightly repressed when extracts from oocytes were added to rabbit reticulocyte lysates. However, specific translation of each mRNA was achieved when the relevant anti-sense 3'UTR was added. This activation was not accompanied by polyadenylation. Various lengths of anti-sense 3'UTR mapped the critical region for repression of translation to a centrally located sequence of 134 nucleotides within the 3'UTR, termed the masking region (Standart *et al*, 1990). An 82-kDa protein from oocytes binds to this element in both the ribonucleotide reductase and cyclin A mRNAs. This protein is phosphorylated after fertilization, and this probably results in polyadenylation and translational activation *in vivo*, although it has been shown by *in vitro* anti-sense experiments that polyadenylation is not a prerequisite for translation.

The mRNAs synthesized during oogenesis can be grouped into two broad classes. One group is typified by actin and ribosomal protein mRNAs which retain fairly long poly A tails and are translated efficiently during oogenesis. But shortly after nuclear envelope breakdown at maturation their poly A tails are shortened, often completely, and their translation ceases. This is thought to occur through a default pathway, for which no specific *cis*-acting signals are required. The timing of deadenylation could be explained by mixing of nuclear and cytoplasmic components, since it has been shown that two components, one cytoplasmic and the other nuclear, are necessary for deadenylation of these mRNAs *in vitro*. While enucleated oocytes did not deadenylate either endogenous or micro-injected RNA upon maturation, whole cells extracts from both immature and mature oocytes deadenylated exogenous RNA to the same extent. Deadenylation activity was not detected in isolated nuclear or cytoplasmic extracts obtained from immature oocytes, but is reconstituted when these fractions are combined *in vitro*. Therefore, the factors

isolated nuclear or cytoplasmic extracts obtained from immature oocytes, but is reconstituted when these fractions are combined *in vitro*. Therefore, the factors required for deadenylation activity were present in immature oocytes, but poly A removal was prevented by sequestration of one or more of these components in the nucleus. Therefore, maturation specific deadenylation of maternal mRNAs occurred upon the release of nuclear factors into the cytoplasm at germinal vesicle breakdown (Varnum *et al*, 1992).

The second group of mRNAs requires signals that reside in the 3'UTR. These mRNAs are stored in the cytoplasm of the developing oocyte as masked untranslated transcripts with short poly A tails. This involves shortening of the long poly A tail originally added in the nucleus. At maturation or following fertilization, these mRNAs undergo polyadenylation and translation activation. In Xenopus and mouse oocytes, it has been shown directly that the length of the poly A tail can control translation of maternal mRNAs (McGrew *et al*, 1989; Vassalli *et al*, 1989; McGrew & Richter, 1990; Paris & Richter, 1990). The majority of mRNAs undergo deadenylation and the total poly A content declines by about 50% (Sagata *et al*, 1980). These mRNAs become dormant, having a poly A tail that is too short for efficient translation. During meiotic maturation, or at specific times after fertilization these mRNAs are mobilized onto polysomes. One such mRNA is G10, which is recruited for translation during oocyte maturation. The poly A tail of this mRNA is elongated from 90 to 200 residues during maturation coincidentally with its translation. The cytoplasmic polyadenylation involves addition of A residues to the pre-existing short oligo A tail, and the elongation of the poly A tail is then sufficient to trigger translation. The tissue-type plasminogen activator (tPA) mRNA is stored in the cytoplasm of fully-grown primary mouse oocytes in a stable and untranslated form. The dormancy of this mRNA is associated with a short poly A tail, and poly A tail elongation controls tPA mRNA translational activation during meiotic maturation (Huarte *et al*, 1992). This stage-specific lengthening of the poly A tail has been found to require two types of motif in the 3'UTR: the AAUAAA nuclear polyadenylation signal, and a U-rich motif known as the cytoplasmic polyadenylation element (CPE) in the Xenopus system, or adenylation control

al, 1997) and the polyadenylation that occurs subsequent to maturation. As with nuclear processing of primary transcripts, if the AAUAAA motif is mutated, cytoplasmic polyadenylation is prevented suggesting that the process of nuclear and cytoplasmic polyadenylation may be mechanistically similar (Fox *et al*, 1989; Huarte *et al*, 1992).

The nucleotide sequences UUUUUAU and UUUUAAU have been identified as CPEs in Xenopus oocytes (McGrew *et al*, 1989; Fox *et al* 1989). However, in another maternal mRNA G10 the CPE differs in sequence and length; UUUUUUAUAAAG (Paris & Richter, 1990). Therefore, there is no simple unique sequence that is required since related but different sequences have been found to be present in different mRNAs (see table1). These variations in the CPE may dictate the subtleties with respect to the timing and degree of polyadenylation, since at meiotic maturation different mRNAs will not be adenylated at the same time or to the same degree. In the mouse, the sequence element AUUUUAAU is the equivalent of the CPE identified in Xenopus oocytes. The rat tPA mRNA, which is also translated during meiotic maturation (Huarte *et al*, 1985) contains this sequence about 100 bases upstream of the AAUAAA sequence (Feng *et al*, 1990). The hypoxanthine-guanine phosphoribosyltransferase (HPRT) mRNA in the mouse is also specifically polyadenylated during meiotic maturation (Paynton *et al*, 1988). It contains an AUUUUAAU sequence about 200 nucleotides upstream of AAUAAA in addition to a UUUUAAAU sequence close to the AAUAAA that has been suggested to function as the CPE (Paris & Richter, 1990). Whereas the CPE sequences, which are critical for polyadenylation during meiotic maturation, are also responsible for deadenylation in primary oocytes, the AAUAAA sequence is not required for deadenylation.

The above examples show that, since the same region of the 3'UTR directs cytoplasmic deadenylation and readenylation, it functions as an adenylation control element (ACE), acting in *cis* to control the length of the poly A tail of an mRNA, and therefore time of its translation. Table 1.1. shows some of 3'UTR regulatory motifs identified that have been proven to regulate polyadenylation. Although these elements are generally composed of A and U residues, they vary in their sequence

and distance form the poly A signal, reflecting the subtle differences in polyadenylation of different mRNAs.

Table 1.1. mRNAs subject to translational control dependent on polyadenylation and 3'UTR motifs (after Standart & Jackson, 1994)

Cell type, condition and mRNA *species*	*Regulatory 3'UTR motif*	*Comments*
Xenopus oocyte maturation		
G10 mRNA	**UUUUUUAUAAAG**guguAAUAAA	CPEs proven to regulate maturation-dependent polyadenylation and translation
B4 mRNA	**UUUUUAAU**guuuaauucuauaAAUAAA	
D7 mRNA	**UUUUUAU**cacaAAUAAA	CPE proven to regulate polyadenylation: **evidence in respect of translation**
Xenopus early embryogenesis		
C12 mRNA	$(U)_{>11}$ ◄— 31 nt —► AAUAAA	Shown to regulate polyadenylation, and presumed to be the element that regulates translation
Spisula oocyte maturation		
ribonucleotide reductase (RR) mRNA	~ 135 nt motif in 5'-proximal half of 3'UTR	Motifs shown to regulate translation at least *in vitro*, since anti-sense RNAs activate translation. Relationship to CPEs unknown
cyclin A mRNA	~ 130 nt motif centrally located in 3'-UTR	
Mouse oocyte maturation		
tissue plasminogen activator (tPA) mRNA	AUUUUAAU & AUUUUA (two copies of each motif)	ACE motifs located in terminal 20% of 3'UTR. 60-125 nt upstream of 3-end. Proven to regulate polyadenylation. Involved in translational regulation
HPRT mRNA	AUUUUAAU & UUUUAAAU	Sequences though to function as ACE/CPE

CPE. cytoplasmic polyadenylation element. ACE, adenylation control element.

1.2.2.2. Translational control independent of polyadenylational changes.

In the above section, it was shown how translation of some mRNAs are determined by the lengths of their poly A tails, which are specified by signals within the 3'UTRs of the transcripts. There are now many examples known of mRNAs that are translationally regulated by signals within their 3'UTR independent of polyadenylation changes. Regulation of gene expression at the level of translation is of critical importance during early embryonic development, during cell growth and differentiation. It is also important in cells that are subjected to nutritional changes and some forms of stress. For example, the generation of body pattern during development requires temporal and spatial regulation of gene expression. In organisms such as *Drosophila*, *Caenorhabditis*, and *Xenopus*, proteins synthesized from maternally provided mRNAs control events in embryogenesis. Because these mRNAs are present in oocytes prior to fertilization, their expression during development needs to be controlled by mechanisms other than transcription. The majority of maternal mRNAs are stored in a translationally silent state until after fertilization, when they become active and are translated to produce proteins necessary for the newly formed embryo (Curtis *et al*, 1995; St.Johnston, 1995). The fibroblast growth factor-1 (FGF-1) receptor mRNA is an example of a translationally repressed mRNA in *Xenopus* whose translation is inhibited by an element within the 3'UTR. The mRNA encoding FGF-1 receptor, which is necessary for normal development, is stored as a stable, untranslated transcript in immature oocytes. Its translation is activated at meiotic maturation. Robbie *et al*, (1995) have identified a 180-nucleotide *cis*-acting element within the 3'UTR of FGF mRNA that inhibited the translation of chimeric transcripts without changing the stability. The inhibition of translation was demonstrated to occur independently of polyadenylational changes, since the deletion of the poly A tract or polyadenylation signal sequences did not affect translational inhibition and the activation of the chimeric transcripts did not require lengthening of the poly A tail. A reversal of translational inhibition was seen at meiotic maturation, without poly A tail lengthening. The translational inhibition was shown to be mediated by a 43-kDa oocyte cytoplasmic protein that bound specifically to the translation inhibitory element and Y box proteins, which are

thought to bind non-specifically to dormant mRNAs (see below and Murray *et al*, 1992).

Both mRNA localization and translational regulation are important in restriction of specific key molecules in the *Drosophila* embryo. Anterior-posterior polarity of the *Drosophila* embryo is initiated during oogenesis via differential maternal mRNA localization. Proteins from these mRNAs are expressed in gradients, and specify cell fates along the anterior-posterior axis of the embryo. A gradient of Bicoid (Bcd) protein starting from the anterior pole, where *bicoid* mRNA is localized, controls development of the head and thoracic structures (Dreiver & Nusslein-Volhard, 1988a). *Nanos* mRNA is localized to the posterior pole, from which a gradient of Nanos (Nos) protein controls the development of the abdomen (Barker *et al*, 1992; Gavis & Lehman, 1992). Although Bcd acts mainly as a transcription factor controlling the activation of the segmentation gene *hunchback* (Struhl *et al*, 1989), it has been shown to inhibit translation of *caudal* mRNA in the anterior of the embryo, where it forms a gradient emanating from evenly distributed RNA in the oocyte, in the opposite direction to Bcd (Rivera-Pomar, 1996). In the absence of a Bcd gradient, Caudal protein becomes evenly distributed throughout the embryo. Therefore, Bcd controls the region-specific translation of *caudal* mRNA through a Bcd-binding region of *caudal* mRNA.

Nos also acts by inhibiting the translation of the transcriptional repressor *hunchback* mRNA, ultimately leading to the degradation of this mRNA and allowing the expression of genes required for formation of the abdomen. (Wharton & Struhl, 1991). It also represses the translation of *Bcd* mRNA, which is coupled with deadenylation but not degradation of the mRNA. Repression requires a short sequence motif known as the Nanos response element (NRE) present as one copy in the 3'UTR of the *Bcd* mRNA and two copies in the *hunchback* mRNA. The NREs are both necessary and sufficient to confer Nos-dependent regulation. Nos therefore, acts to control Hunchback expression (and hence abdominal pattern) as a function of its concentration-dependent interaction with NREs.

The expression of Nos itself depends on the localization of the *nos* mRNA at the posterior pole of the embryo, (Wang *et al*, 1994). In oocytes from

females mutant for *nos* there is no localization of *nos* mRNA (Lehman & Nusslein-Volhard, 1991) and the embryos resulting from these oocytes lack abdominal segments. Thus, spatial regulation of *nanos* is achieved by a combination of localization and localization-dependent translation. The loss of *nos* activity in embryos defective in *nos* mRNA localization has been shown to be a result of the translational repression of the unlocalized mRNA (Gavis & Lehman, 1994). This translational repression and the localization of *nos* mRNA are mediated through sequences within the 3'UTR (Gavis & Lehman, 1994; Gavis *et al*, 1996). The signal for translational repression of unlocalized *nos* mRNA is a discrete 90 nucleotide element which is conserved between related species (Gavis *et al*, 1996). It lies within the localization signal but functions independently of it.

In the nematode *C. Elegans*, there are two sexes; males and hermaphrodites. Hermaphrodites produce sperm initially and then switch to producing oocytes. Both spermatogenesis and switch to oogenesis are regulated by the 3'UTRs of two sex-determining genes. The onset of spermatogenesis is controlled by the 3'UTR of the *tra-2* gene (Goodwin *et al*, 1993), and the switch to oogenesis is regulated by the 3'UTR of the *fem-3* gene (Ahringer & Kimble, 1991). Inhibition of *tra-2* translation is necessary to allow spermatogenesis. The inhibition of translation has been mapped by the inclusion of these elements in a reporter RNA to a perfect direct repeat of 28 nucleotides within the 3'UTR of the *tra* mRNA, known as the direct repeat elements (DREs). The DREs inhibit translation of *tra-2* mRNA by specifically binding a protein called direct repeat factor (DRF) (Goodwin *et al*, 1993). Similarly, repression of *fem-3* is required to allow the switch from spermatogenesis to oogenesis. Mutation analysis of a 5-nucleotide motif centrally located in the 3'UTR showed that it was responsible for translational regulation of *fem-3*. A protein, *fem-3* binding factor (FBF), which binds specifically to the 3'UTR and is necessary for determining sexual fate, has been identified (Zhang *et al*. 1997).

Translational regulation is not limited to oocytes and early embryonic development. For example during red blood cell differentiation, the mRNA coding for erythroid 15-lipoxygenase (LOX) is synthesized in the early stages of mammalian erythropoiesis, but is only activated for translation in peripheral reticulocytes. LOX

catalyzes the dioxygenation of polyenoic fatty acids, and is unique in its ability to attack intact phospholipids. It is thought to be the main factor responsible for the breakdown of internal membranes such as those in mitochondria, when reticulocytes mature into erythrocytes (Rapoport & Schewe, 1986). LOX mRNA is transcribed at the bone marrow cell stage, but is stored untranslated and expression is only initiated prior to maturation of reticulocytes into erythrocytes (Thiele *et al*, 1982). The 3'UTR of the LOX mRNA was observed to contain 10 tandem repeats of a 19 nucleotide pyrimidine-rich motif, consensus CCCCA/GCCCUCUUCCCCAAG (Fleming *et al*, 1989; Hunt, 1989). A 48 kDa protein that specifically binds to the 3'UTR repeat region was identified by gel retardation assays and UV cross-linking. The protein, purified by affinity chromatography from rabbit reticulocytes, was shown to specifically inhibit the translation of mRNAs containing the LOX mRNA 3'UTR sequence. The inhibition is specifically *cis*-acting as demonstrated by the addition of two species of mRNA, one containing the 3'UTR repeat region and the other lacking it to *in vitro* translation assays (Ostareck-Lederer *et al*, 1994). The inhibition of translation was proved to be due to RNA-protein interaction and not RNA degradation. Since mRNAs used in experiments contained neither a poly A tail nor a poly A signal, the translational control of LOX mRNA was independent of changes in polyadenylation.

Similarly, during spermatogenesis in the mouse, the transcription of mRNAs coding for protamine, an arginine-rich chromosome protein, occurs shortly after completion of meiosis and stops about a week later. During this time the mRNAs are stored, polyadenylated with some 160 A residues, in the cytoplasm as translationally inert ribonucleoprotein particles (Kleene *et al*, 1984). Activation of translation in elongated spermatids is coupled with a shortening of the poly A tail to about 30 residues, and since this occurs slightly after translational activation, deadenylation is thought unlikely to be a primary cause for activation. Transgenic studies have demonstrated that the temporal expression of the mouse protamine 1 mRNA in the testis is determined by the 3'UTR (Braun *et al*, 1989). A 62-nucleotide motif that is necessary and sufficient for temporal translational regulation has been identified in this region (Braun *et al*, 1990). Work published by Kwon & Hecht, (1991) identified

two highly conserved regions in the protamine 2 3'UTR termed Y and H boxes that interact with protein. These sequence elements were found to be highly conserved among post-meiotic translationally regulated nuclear proteins of the mammalian testis. An 18-kDa protein from the testis cytoplasmic extract was found to bind specifically to the mouse protamine 2 mRNA, conferring translational repression. This protein is phosphorylated at the stages when protamine mRNAs are masked, and dephosphorylated in elongating spermatids. Phosphorylation of the 18-kDa protein is necessary for it to bind at the Y box (Kwon & Hecht, 1993). In contrast, Fajardo *et al*, (1994) have identified cytoplasmic proteins of 53 and 55 kDa that bind specifically to 22 and 20 nucleotide regions within the 3'UTRs of protamine 1 and protamine 2 mRNAs respectively. However, these binding sites did not fall into the 62-nucleotide region necessary for translational repression, suggesting that the binding activity observed here serves some other purpose in spermatid differentiation.

Another example how the 3'UTR acts to repress translation is seen in the muscle transcription factor *mef2a* gene. The myocyte enhancer factor, MEF2, family of transcription factors are important for the activation of muscle-specific transcription in skeletal, cardiac and smooth muscle cells (Olson *et al*, 1995). One of these factors MEF2A seems to be controlled at the level of translation (Suzuki *et al*, 1995), and although the *mef2a* transcripts are expressed in a wide range of tissues, the protein is more restricted (Yu *et al*, 1992). The 3'UTR of *mef2a* when fused to the chloramphenicol acetyltransferase reporter (CAT) gene inhibited CAT expression (Black *et al*, 1997). The inhibition of expression was shown to be a result of translational repression as demonstrated by RNase protection assays, which showed that the steady-state level of CAT mRNA was not affected by the fusion of the *mef2a* 3'UTR. The translational repression, which was mapped to an internal fragment of about 400 nucleotides, is regulated during muscle cell differentiation, when inhibition caused by the 3'UTR is relaxed. Therefore, the *mef2a* 3'UTR functions as a *cis*-acting translational repressor, and this repression may contribute to the tissue-restricted expression of MEF2A.

The examples of translational regulation by the 3'UTR given so far describe the repression of translation. The 3'UTR can influence translation in a positive manner as seen in the human heat-shock protein 70 (HSP70) gene. When cells are under stress they synthesize a family of proteins called heat shock proteins (HSP). They mediate tolerance to thermal and chemical injury and therefore the regulation of HSPs is critical to survival. Although evidence suggests that expression of HSP70 is regulated at the level of transcription (Wu *et al*, 1986), it was postulated that post-transcriptional mechanisms also existed to regulate production of HSP70 protein. The injured cell has the ability to identify and preferentially translate HSP70 mRNA to the exclusion of all other messages (Mizzen & Welch, 1988; Peterson & Lindquist, 1989). When the 3'UTR of HSP70 mRNA, as a mechanism of post-transcriptional regulation, was investigated it was found that the 3'UTR when fused to the CAT reporter gene up-regulated its expression in response to heat stress (Moseley *et al*, 1993). HSP70 message increased in response to heat stress and therefore, it is likely that the 3'UTR of HSP70 mRNA contains a regulatory element to stabilize the mRNA at increased temperatures, and preferentially translate HSP70 mRNA.

The 3'UTR of the human p53 mRNA has been shown to participate in both translational repression and activation. In the blast cells of patients with acute myelogenous leukaemia, the level of p53 protein does not correlate with levels of p53 mRNA, indicating a mechanism of translational control. In addition, when two cell lines with similar levels of p53 but different levels of p53 mRNA were examined, there was preferential association of p53 mRNA with large polysomes in the cell line with less p53 RNA. It was subsequently shown that the p53 3'UTR functions in *cis* to repress translation (Fu *et al*, 1996). Furthermore, when human cells were gamma-irradiated, p53 protein accumulated. However, there was no parallel increase in p53 mRNA. Instead, the transcripts became increasingly associated with polysomes, suggesting post-transcriptional regulation. In studies with a chimeric reporter construct containing p53 3'UTR the translation of this hybrid mRNA was repressed *in vivo*, but after gamma-irradiation, reporter activity was

elevated (Fu & Benchimol, 1997). This demonstrated that the 3'UTR could regulate gene expression in both a positive and negative manner in the same gene.

The examples described here and summarized in Table 1.2. are just a few of many which demonstrate the post-transcriptional regulation of gene expression at the level of translation that are directed by sequence elements within the 3'UTR. Although many of the 3'UTR elements act, independent of the state polyadenylation of the mRNA, to repress translation, there are examples where such elements promote translation, often in response to stimuli.

Table 1.2 Some mRNAs subject to translational control dependent on 3'UTR motifs **(after Standart & Jackson, 1994)**

Cell type, condition and mRNA *species*	*Regulatory 3'UTR motif*	*Comments*
Mouse spermatogenesis		
protamine- I mRNA	62 nt motif at extreme 3' end of 3'-UTR	Shown to confer appropriate temporal regulation of expression of chimaeric transgene
C. elegans hermaphrodite germ line sex determination		
fem-3 mRNA	**UCUUGU**	Centrally located within 3'-UTR in each mRNA species. Shown by analysis of mutations to the responsible for translational regulation
tra-2 mRNA	**(NNUAUUUAAUUUCUUAUCUACUCAUAUCUANN)$_2$**	
Drosophila early embryogenesis		
hunchback mRNA *bicoid* mRNA	**GUUGUNNNNNAUUGUA**	NRE (nanos response element): 2 copies of motif in *hunchback* mRNA, one copy in *bicoid* mRNA. Proven to be responsible for translational regulation.
Erythroid cell differentiation		
15 -lipoxygenase mRNA	**CCCRGCCCUCUUCCCAAG (consensus)**	10 copies in rabbit erythroid LOX mRNA, 4 copies in human RNA, located in 5'-half of 3'UTR. Binds 48 kDa protein which specifically represses translation of LOX mRNA *in vitro* -
Muscle cell differentiation		
mef2a mRNA	~400 nt sequence centrally located within 3'UTR	*cis*-acting translational repressor, which may contribute to tissue-restricted expression of MEF2Afunction
Human cells		
heat shock protein 70 mRNA	Heat responsive element in 3'UTR?	Regulation of HSP70 by 3'UTR through stabilization and translation
p53 mRNA	Repressor element	3'UTR shown to repress translation, increased translational activity in response to gamma-irradiation.

1.2.3. Determinants of mRNA stability.

The stability of mRNAs in cells varies widely. Some mRNAs are short-lived whereas others are stable for many hours. Therefore mRNA stability plays an important part in regulating gene expression. The primary and secondary structure, translational rate, and intracellular localization are some of the variables that can influence mRNA stability. In addition, mRNA stability can be determined by several other factors, which include *cis*-acting sequence elements, *trans*-acting regulatory factors and hormones and growth factors (see figure 1.3). Different specific features of the 3'UTR determine mRNA stability

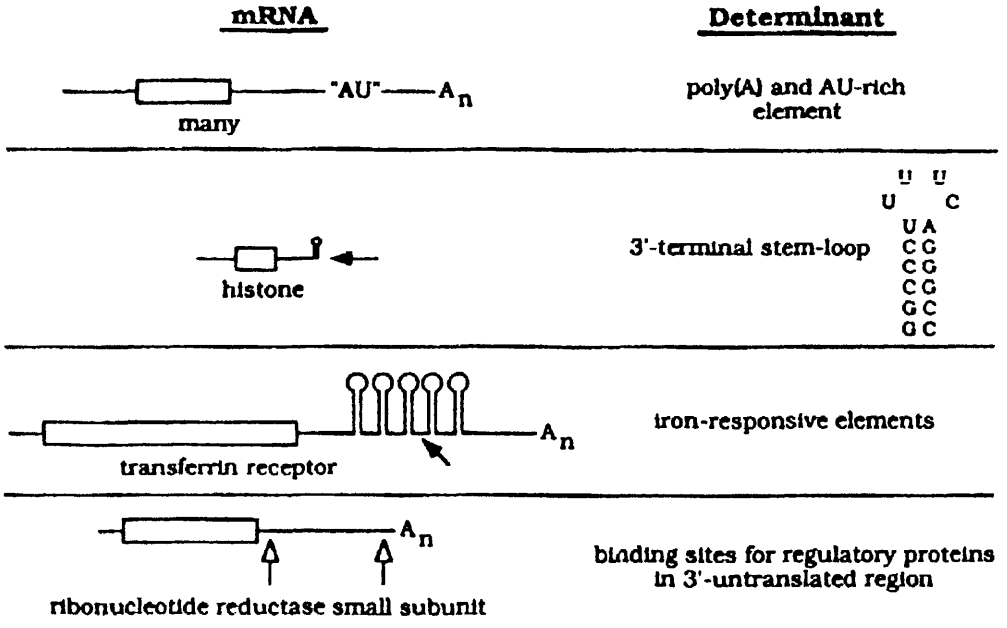


Figure 1.3. Schematic representation of some 3'UTR mRNA stability determinants. "AU" denotes AU-rich region, solid arrows indicate determinants that are also cleavage sites. Open arrows indicate location of determinants. Boxes represent coding region and straight lines untranslated regions (adapted from Ross, 1995).

1.2.3.1. The Poly A tail.

Of the *cis*-acting sequences that determine stability, the poly A tail probably has some role in protecting the mRNA from degradation. Two observations implicate the poly A tail in protecting mRNAs to a certain extent. 1) Deadenylation is the first step in mRNA degradation of many RNAs (see reviews by Decker & Parker, 1994; Beelman & Parker, 1995). In transcriptional pulse-chase experiments using a regulatable promoter, the mammalian c-*fos* mRNA and several yeast mRNAs do not decay until their poly A tails have been shortened (Shyu *et al*, 1989; Decker & Parker, 1993). 2) A complex between poly A and PABP protects mRNAs from rapid degradation *in vitro*. Poly A mRNA substrates are rapidly degraded when incubated with extracts lacking PABP, but stabilized when excess exogenous PABP is added. In contrast mRNAs lacking a poly A tail are unstable whether PABP poly is present or not (Bernstein *et al*, 1989). However, these observations do not mean that deadenylation automatically triggers mRNA decay, as there are several deadenylated mRNAs such as that for actin, which are quite stable in cells (Geoghegan & McCoy, 1986).

1.2.3.2. The histone mRNA 3'end.

Most of the works published to date on mRNA stability have identified elements within the 3'UTR that signal mRNA decay. In many cases the use of chimeric constructs and transfection experiments have revealed how the 3'UTR can function as an instability element independent of the rest of the mRNA. Some of these elements have been shown to be sites for protein binding. In histone mRNAs, the signals for processing and stability are found in a stem-loop structure in their 3'UTR. In the cell cycle, histone mRNAs are transcribed rapidly during the S phase. The half-life of the transcripts during this phase is 40 minutes. At the end of the S phase, the transcription rate drops. The efficiency of pre-mRNA processing decreases, and the half-life of cytoplasmic mRNA drops to 10 minutes (Alterman *et al*, 1984; Heintz *et al*, 1983; Morris *et al*, 1991). When the last 30 nucleotides of histone mRNA were added to the end of globin mRNA in a chimeric construct, the globin mRNA was regulated post-transcriptionally in the same manner as the wild-

type histone mRNA (Luscher *et al*, 1983). It was shown subsequently that a stem-loop structure, which is present in all histone mRNAs that are cell-cycle regulated, and adjacent sequences are necessary for mRNA processing and stability (Pandey & Marzluff, 1987; Pandey *et al*, 1994).

1.2.3.3. The iron–responsive element.

The transferrin receptor functions to import iron into cells. The levels of the receptor mRNA are modulated by cellular iron content via iron-response-elements (IREs) within the 5'UTR and 3'UTR of the transferrin receptor mRNA (reviewed by Thiel, 1990; Casey *et al*, 1988; Casey *et al*, 1989). The IRE is 23-27 base pair stem with a mismatched C and a 6-nucleotide loop (Müllner & Kühn, 1988). The IRE in the 3'UTR functions by binding a 90-kDa monomeric protein. This protein, termed iron regulatory protein (IRP), binds to the IRE when intracellular iron levels are low, and stabilizes the transferrin receptor mRNA by protecting it from degradation, resulting in increased synthesis of transferrin receptor. In situations where intracellular iron is abundant, there is no IRE-iron regulatory protein complex formation and the mRNA is relatively unstable. Therefore the IREs regulate mRNA half-life and translation of the transferrin receptor, in response to iron content of the cell.

1.2.3.4. AU-rich and U-rich elements.

The most common sequence elements within 3'UTRs to be linked with mRNA instability are the AU-rich elements (AUREs). It was found that many short-lived mRNAs such as those for cytokines (e.g. GM-CSF, tumour necrosis factor) and oncogenes (e.g. c-*fos*, c-*myc*) contained single or multiple copies of the pentamer AUUUA and/or U rich regions within their 3'UTR (Shaw & Kamen, 1986; Caput *et al*, 1986). In transfection studies, when the AURE from the unstable mRNA coding for GM-CSF was placed within the 3'UTR of the stable β-globin mRNA or fused to a CAT reporter gene, the chimeric transcript was rapidly degraded (Shaw & Kamen, 1986; Iwai *et al*, 1991). Interestingly, treatment with phorbol esters stabilized some short-lived mRNAs. The sequence that mediates phorbol ester stabilization of GM-

CSF mRNA has been identified as a 60 nucleotide region upstream of the AU-rich region (Iwai *et al*, 1991).

Within the c-*fos* 3'UTR there is a region of approximately 40 to 50 nucleotides long which contains several AUUUA motifs, and another region which contains an approximately 20 nucleotide U-rich sequence (Chen *et al*, 1993). In experiments where the AUUUA motifs were mutated, the deadenylation rate decreased slightly, but the transcript were stabilized by at least five-fold. In comparison, deletion of the U-rich region only resulted in decreased deadenylation, but only a two-fold increase in stability (Chen & Shyu, 1994). Therefore, in c-*fos* mRNA the AUUUA element facilitates degradation of the mRNA, and the U-rich region promotes removal of the poly A tail and enhances the destabilizing function of the AUUUA motifs (You *et al*, 1992). Similarly, although Jones & Cole (1987) showed that deletions of the c-*myc* 3'UTR containing the AUUUA motif resulted in stabilization of the c-*myc* transcripts, work published by Alberta *et al* (1994) demonstrated that the AUUUA motif was not the essential component of c-*myc* destabilizing activity. Chimeric β-globin constructs mapped the destabilizing activity of the c-*myc* 3'UTR to three redundant instability elements, one of which was adjacent to, but distinct from, the AUUUA pentamer. This region consists of a 39 nucleotide U-rich region which, in chimeric constructs, co-operates with adjacent sequences to destabilize the β-globin mRNA. The importance of the U-rich region is emphasized by the finding that other short-lived mRNAs have about 40-50% U residues in the region flanking the AUUUA motif. In comparison, stable mRNAs that contain AUUUA motifs contain only about 25% U residues in the flanking region (Alberta *et al*, 1994).

The mechanisms whereby AU rich elements influence mRNA stability are probably complex, with processes differing in various mRNAs. The sequence UUAUUUA(U/A)(U/A) or UUAUUUAUU repeated several times and placed within the 3'UTR of a stable mRNA has been found to be the most effective destabilization element of related AUREs tested (Lagnado *et al*, 1994). Interleukin-3 mRNA, which is unstable in a mast cell line cultured in low calcium, contains eight AUUUA motifs within its 3'UTR. The mRNA becomes stable if more than two of these pentamers

are deleted. Interleukin-1a mRNA, which is has also been shown to be unstable in a rabbit reticulocyte system, has four closely spaced AUUUA motifs. If the motifs are separated the mRNA is rendered stable (Gorospe & Baglioni, 1994). In some mRNAs the AUREs function as destabilizing elements only when the message is being translated, with the mRNA being stable if it is not being translated (Aharon & Schneider, 1993).

Although there is much evidence to suggest that the AUUUA elements influence mRNA stability, it is unlikely that a five base element alone is sufficient to act as a destabilizing element in every case. In fact mRNAs such β-globin and murine osteonectin, which are stable, contain an AUUUA motif in their 3'UTR. In some mRNAs 3'UTR AU-rich sequences act to suppress translation. The AU-rich element in the human interferon-beta mRNA is a potent suppressor of its translation (Kruys *et al*, 1987; Kruys *et al*, 1988; Kruys *et al*, 1989).In *in vitro* experiments, the association of the AU-rich sequence with the poly A tail reduces the translational efficiency (Grafi *et al*, 1993). The tumour necrosis factor (TNF) 3'UTR contains a UUAUUUA element. This motif and flanking sequences in chimeric CAT reporter constructs was found to suppress the translation of the hybrid mRNA (Han *et al*, 1990). Furthermore, the TNF 3'UTR has been shown to contain an endotoxin response element, which activates translational de-repression of TNF 3'UTR (Han *et al*, 1990).

1.2.3.5. Factors that affect mRNA stability.

The intracellular half-lives of many mRNAs are determined by various factors. The stability of some mRNAs can be influenced by factors such *trans*-acting proteins, hormones, growth factors and even oxygen tension. These factors act to either stabilize or destabilize the mRNA and in many cases this process is mediated through the 3'UTR. For example, cell activation with phorbol ester (Shaw & Kamen, 1986), or calcium ionophore (Wodnar-Filipowicz & Moroni, 1990) treatment has been shown to stabilize AU-rich mRNAs. Malter and Hong (1991) have demonstrated that AU binding protein activity is up-regulated by 12-*O*-tetradecanoylphorbal-13-acetate (TPA) and/or ionophore treatment, and propose that

AU-binding factor (AUBF) binding may mediate TPA and ionophore mediated labile message stabilization. Stephens *et al* (1992) have demonstrated that the AU binding factor from 3T3-L1 preadipocytes not only binds to the AU-rich region of glucose transporter (GLUT-1) mRNA 3'UTR, but that this activity is up-regulated by TPA, tumour necrosis factor α (TNFα) and cAMP, all of which stabilize the GLUT-1 mRNA. It is thought that AUBF mediates GLUT-1 mRNA stabilization by occupying the AU rich region in the 3'UTR.

The cytochrome P-450 family of mono-oxygenases catalyzes the oxidation of various lipophilic compounds. The expression of one of these enzymes, CYP2a5, is enhanced by pyrazole, which increases the half-life of the CYP2a5 mRNA and causes elongation of its poly A tail (Aida & Negishi, 1991;Hahnemann *et al*, 1992). A 44 kDa protein, the activity of which is induced by pyrazole, binds specifically to the CYP2a5 3'UTR mRNA. Although the 3'UTR contains a single AUUUA motif, the inducible protein forms a complex with the 3'UTR independently of this sequence. The finding that the 44 kDa protein is found in nuclei and polyribosomes suggests that this protein is involved in nuclear and cytoplasmic mRNA turnover and/or binding and stabilization of the CYP2a5 mRNA to the polyribosomes (Geneste *et al*, 1996).

The stability of the mRNA for tyrosine hydroxylase (TH), an enzyme involved in the biosynthesis of catecholamine neurotransmitters, and that of vascular endothelial growth factor (VEGF), is increased by reduced oxygen tension (hypoxia) (Czyzyk-Krzeska *et al*, 1997; Levy *et al*, 1997). Increased stability of the TH mRNA during hypoxia is a result of the enhanced binding of a 50-kDa cytoplasmic protein (hypoxia inducible protein, HIP) to a pyrimidine-rich region within the TH mRNA 3'UTR. Mutational analysis has identified the optimal protein-binding site to be a (U/C)(C/U)CCCU motif (Czyzyk-Krzeska *et al*, 1997; Paulding *et al*, 1999). The 3'UTR of the VEGF mRNA contains five hypoxia-inducible RNA-protein binding sites, which mediate mRNA stability (Levy *et al*, 1997). The binding activity of erythropoietin mRNA binding proteins is also up-regulated by hypoxia (Rondon *et al*, 1991). These cytosolic proteins bind specifically to a 120-nucleotide fragment

within the erythropoietin mRNA, and have been proposed to participate in the erythropoietin mRNA turnover.

Mammalian ribonucleotide reductase R1 mRNA stability under normal and TPA-stimulated conditions provides a contrasting example of the involvement of 3'UTR and protein interactions. The expression of the ribonucleotide reductase gene R1 is elevated by increased stability in Balb/c 3T3 fibroblasts treated with TPA. Chen *et al*, (1993) have demonstrated the existence of a 49 nucleotide TPA-responsive element with the R1 mRNA 3'UTR. The octamer CAAACUUC has been defined as the sequence within the 49 nucleotides region that selectively binds a 57 kDa protein in unstimulated cells (Amara *et al*, 1995). However this binding activity is abolished in TPA treated cells. Therefore although the 49-nucleotide region may control TPA induced message stabilization, the protein binding activity observed is probably involved in mRNA degradation. Down–regulation of protein binding protects R1 mRNA by inhibiting RNA degradation.

1.2.3.6. AURE binding proteins.

Some *trans*-acting factors, namely RNA-binding proteins, appear to stabilize some RNAs to which they are bound while others act as destabilizers. A number of proteins, termed AU or UR binding proteins/factors (AUBP/Fs, URBP/Fs), have been identified that bind with high affinity to mRNAs containing AU-rich and U rich elements. For example, four U-rich sequence binding cytoplasmic proteins that interact with a 20 nucleotide U-rich sequence within the c-*fos* ARE have been identified (You *et al*, 1992). Bandshift analyses and competition experiments have showed that the proteins form specific complexes with the c-*fos* AURE and recognize the U rich region rather than AUUUA motifs. The function of the URBPs has been suggested to be an involvement in the first step of c-*fos* AURE-mediated mRNA degradation, the removal of the poly A tail. Three cytoplasmic proteins, 33, 39 and 42 kDa in size, that bind in a sequence-specific manner to the AU-rich region in the GM-CSF 3'UTR have also been detected (Bickel *et al*, 1992). The 39 nucleotide U-rich region of the c-*myc* 3'UTR that has been implicated in mRNA destabilization has been shown to specifically bind a 59-kDa factor contained in

cytoplasmic extracts of Balb/c3T3 cells. Two other proteins of approximately 33 and 37 kDa were also found to be present in the complex. Brewer (1991) has demonstrated that proteins, which accelerated c-myc mRNA degradation in a cell-free decay system, bound specifically to both the c-*myc* and GM-CSF 3'UTR. These proteins are phosphorylated and can be found in complex with other polypeptides (Zhang *et al*, 1993). The 37-kDa protein, AUF1, has been cloned and is found to be localized in both the nucleus and the cytoplasm, suggesting a role in both nuclear and cytoplasmic mRNA turnover.

Another AU binding protein, 34 kDa AU-A, which is one of three factors found to bind to AUUUA multimers in lymphokine mRNA 3'UTRs (Bohjanen *et al*, 1991; Bohjanen *et al*, 1992) is also found in both the nucleus and cytoplasm. It binds to several other U-rich regions, and the AU-rich region of c-*myc*.

1.2.3.7. The 3'UTR and stable mRNAs.

Whereas it is evident that unstable mRNAs contain instability elements, it is less clear whether stable mRNAs contain stabilizing elements or are stable by default. One of the most well characterized examples of mRNA stability dependent on *cis*-acting determinants is that of the human globin genes. During erythroid differentiation, transcription shuts down midway, and mRNAs coding for the globins need to remain stable for several days to be fully expressed. The stability of globin mRNAs has been estimated to be between 16 and 48 hours depending on the experimental system used (Aviv *et al*, 1976; Lodish & Small, 1976; Krowczynska *et al*, 1985; Ross & Sullivan, 1985). This remarkable stability allows the globin mRNAs to accumulate and exceed to more than 95% of the total cellular mRNA in terminally differentiated erythroid cells. The observations of Krowczynska *et al*, (1985) of a lack of correlation between the tendency to lose poly A and the rate of mRNA decay, and the accumulation of poly A deficient actin mRNA in erythrolukaemia cells treated with DMSO, suggested that stability of the actin mRNAs was not determined solely by the presence of a poly A tail. Initial evidence for a discrete stabilizing element in the α-globin mRNA was found by investigations of a naturally occurring α-globin mutant, in which the translational termination

codon UUA is changed to UCA (Weiss & Liebhaber, 1994). The base substitution allows the ribosome to read through into the 3'UTR. This read-through appears to disrupt sequences or structures within the 3'UTR, leading to destabilization of the α-globin mRNA. Mutational analysis has mapped a *cis*-determining stabilizing element to three cytidine-rich (C-rich) regions within the 3'UTR of the α-globin mRNA (Weiss & Liebhaber, 1995). Mutations in these segments result in translationally independent destabilization of the α-globin mRNA. Wang *et al*, (1995) have detected at least three *trans*-acting factors which interact with the C-rich regions The RNA-protein complex (α-complex) formation is sequence-specific since no such interactions were observed with the β-globin 3'UTR. Furthermore, base substitutions within the C-rich region, which destabilized α-globin mRNA *in vivo* resulted in a parallel disruption of the RNA-protein complex formation. The proteins in the α-complex appear to be ubiquitous since bandshift analysis showed that the α-complex was formed with cytosolic extracts from non-erythroid as well erythroid cells (Wang *et al*, 1995). The lack of tissue- and species-specificity of the proteins forming the α-complex suggests that these proteins are conserved between species and have similar functions in different species. Indeed, the stability of murine α-globin mRNAs is controlled by a parallel mechanism. A CU-rich region within the 3'UTR is the determinant of murine α-globin mRNA stability. One of the protein components of the human α-complex has been identified as a 39-kDa poly C-binding protein (PCBP). While PCBP is essential for complex formation it is not sufficient on its own for α-complex formation. In the mouse, a 48-kDa cytoplasmic poly CU-binding protein has been identified (Wang & Liebhaber, 1996). Interestingly, a protein of the same molecular weight has been found to bind to the LOX mRNA 3'UTR, which contains 10 tandem repeats of a 19-nucleotide pyrimidine-rich sequence (Fleming *et al*, 1989; Hunt, 1989). Although LOX mRNA is synthesized in the early stages of erythropoiesis, translation is only activated several days later in peripheral reticulocytes. In the intervening period, the mRNA must be maintained in a stable and translationally inert state. Ostareck-Leder *et al*, (1994) demonstrated that the interaction of a 48 kDa protein and the pyrimidine-rich sequence was necessary and

sufficient for the specific translational block of LOX and reporter mRNAs *in vitro*. It is highly likely that this protein functions to not only repress translation of LOX mRNA but also stabilize it during the period of repression.

The above examples of how the 3'UTR of mRNAs influence stability are just a few of many which show that post-transcriptional regulation at the level of mRNA stability by the 3'UTR plays an important role in gene expression. From the various observations it can be seen that mRNA stabilization or destabilization occurs through sequence specific interactions with different regulatory proteins, and that this in turn can be influenced by factors such as hormones. Many of the mentioned factors that bind to specific regions within the 3'UTRs of the mRNAs are now being characterized. A better understanding of the mechanisms involved in controlling mRNA stability may have practical value, since it could be used to enhance or reduce the expression of certain genes.

1.2.4. The 3'UTR and mRNA localization.

The cellular morphology and physiology of differentiated cells is a result of the ordered arrangement of specialized proteins. Even in the undifferentiated cell, compartments have particular macromolecular identities composed of polypeptide complexes unique to their sub-cellular function. Although the targeting of membrane-associated proteins is quite well understood (Pelham & Munro, 1993), the mechanisms involved in the localization of cytoplasmic proteins are less clear. It has been generally assumed that once translated, the proteins diffuse to regions of the cell where they are needed. However, simple diffusion would appear to be an inefficient and error-prone way of localizing protein. Another way that proteins can be targeted to discrete sub-cellular regions is by localization of the mRNAs that encode these proteins. In this way specific proteins could be synthesized in the sub-cellular regions where they are required and their translation could be prevented in regions where they are not required. Localized protein synthesis would also seem to be more efficient in that presumably less energy would be required to localize a single mRNA which can be translated many times than to transport many protein

molecules. In addition, localized translation of mRNAs would also allow local translational control.

Evidence to support the intracellular localization of specific mRNAs first came from studies of Ascidian eggs in which actin mRNA was enriched in the myoplasm (Jeffery *et al*, 1983). *In situ* hybridization showed that actin mRNAs and polyadenylated RNA were preferentially localized near the myoplasm while histone RNA was distributed homogeneously. Since then examples of localized transcripts have been found in increasing number of cell types, including fibroblasts, myoblasts, neurons and epithelial cells (see review by Wilhelm & Vale, 1993). Many of the known localized mRNAs are maternal RNAs in *Drosophila* and *Xenopus* eggs. The majority of localized mRNAs studied so far have been found to contain *cis*-acting sequences required for localization within their 3'UTRs. These signals, where precisely mapped, have proved to be fairly long. These relatively large signals probably contain multiple protein binding sites.

1.2.4.1. Localized mRNAs in *Xenopus* oocytes.

In *Xenopus* oocytes, several maternal mRNAs are sorted to either the vegetal pole or the animal pole (Rebagliati, *et al*, 1985). Of these the Vg1 mRNA, which encodes a member of the transforming growth factor β (TGFβ) family, is the best characterized. Vg1 mRNA is localized to the vegetal pole of late-stage oocytes and remains in the vegetal hemisphere of early embryos until after gastrulation (Weeks & Melton, 1987). Localization of Vg1 mRNA appears to be a two-step process, with translocation of the message to the vegetal hemisphere being the first, and anchoring of the message to the cortex being the second (Yisraeli *et al*, 1990). Translocation is thought to involve both microtubules and microfilaments since microtubule de-polymerizing drugs prevent translocation of the mRNA, but do not interfere with already localized mRNA, and microfilament inhibitors cause release of localized message without inhibiting the translocation (Yisraeli *et al*, 1990). By using chimeric constructs linking different regions of the Vg1 mRNA to β-globin reporter sequences, Mowry and Melton (1992) identified a 340-nucleotide sequence within

the 3'UTR that directed the localization of the otherwise non-localized β-globin reporter transcript. Two distinct protein-binding regions separated by 78 nucleotides have been identified in the 3'UTR of the Vg1 mRNA (Schwartz *et al*, 1992). The downstream region corresponds to the previously identified localization signal. A 69-kDa protein binds to this region with high affinity in a specific manner. Competition experiments have established that this protein, although not a general RNA binding-protein, does recognize another *Xenopus* maternal mRNA 3'UTR , TGF*β*-5, which is also a member of the TGF*β* family. Like Vg1, TGF*β*-5 mRNA is tightly localized along the vegetal cortex of middle- and late-stage oocytes. The binding is not, however, inhibited by the animally localized mRNA An2. Although no obvious stretches of homology are found between Vg1 and TGF*β*-5 3'UTRs, they may form similar secondary structures, which may be the primary recognition signal for protein binding and localization.

1.2.4.2. Localized mRNAs in *Drosophila* embryo.

The anterior-posterior axis in *Drosophila* embryos is established by localization of maternal mRNAs encoding proteins that determine morphology (Nusselin-Volhard *et al*, 1987). These proteins are expressed in gradients, and specify cell fates along the anterior-posterior axis of the embryo. The *bicoid* (*bcd*) and *nanos* (*nos*) mRNAs encode the primary determinants of anterior-posterior axis. Bicoid (Bcd) protein controls development of the head and thoracic structures (Dreiver &Nusslein-Volhard, 1988a), while Nanos (Nos) controls the development of the abdomen (Barker *et al*, 1992; Gavis & Lehman, 1992). The generation of Bcd and Nos protein gradients depends on the prior localization of the *bcd* and *nos* mRNAs at the anterior and posterior poles of the embryo, respectively (Dreiver & Nusslein-Volhard, 1988b; Wang *et al*, 1994). The determinant of *bcd* mRNA localization spans a 625-nucleotide portion of the *bcd* 3'UTR which is capable of forming extensive secondary structure (MacDonald & Struhl, 1988). This region was shown to confer anterior localization to a *lacZ/α-tubulin* reporter transcript in transgenic flies. The localization of *bcd* mRNA at different stages of oogenesis is affected by three other

genes, *exuperantia*, *swallow*, and *staufen* (St.Johnston *et al*, 1989; Stephenson *et al*, 1988). Experiments using mutants for these genes and cytoskeletal inhibitors suggest that *exuperantia* and *swallow* function to mediate, stabilize or regulate the association between *bcd* mRNA and microtubules (Pokrywka & Stephenson, 1991). *Staufen* protein appears to function to anchor *bcd* mRNA at the anterior pole.

Wharton & Struhl, 1989, have shown that expression of Nos protein at the anterior of the embryo results in the suppression of normal anterior development and duplication of abdominal structures at the anterior. In oocytes from females mutant for *nos* there is no localization of *nos* mRNA and the embryos resulting from these oocytes lack abdominal segments (Lehman & Nusslein-Volhard, 1991). Therefore, the localization of *nos* RNA is essential for Nos expression and correct pattern formation. The loss of Nos activity in embryos defective in *nos* mRNA localization has been shown to be a result of the translational repression of the unlocalized mRNA (Gavis & Lehman, 1994). The sequence that determines *nos* mRNA localization has been mapped to a large region (780 nucleotides) of the 3'UTR which is composed of partially redundant localization elements. (Gavis *et al*, 1996). Similarly to *bcd* mRNA, the process of *nos* mRNA localization involves several other genes. These include *cappuccino*, *spire*, *staufen*, *oskar*, and *vasa*. *Staufen* protein and *oskar* mRNA are the first molecules to localize at the posterior pole. Since *Staufen* is also necessary for the anterior localization of *bcd* mRNA, it probably functions to anchor the RNAs to structures common to both poles (St. Johnston *et al*, 1991). *Oskar* mRNA directs pole plasm formation (Ephrussi *et al*, 1991) and *Cappuccino* and *Spire* proteins are necessary to localize *oskar* mRNA to the posterior pole. They are presumed to form a cytoskeletal framework for posterior localization. The next protein to be localized is Vasa, and then lastly *nos* mRNA.

Although sequence analysis of *bcd*, *nos*, *vg1* and *an2* mRNA has not revealed any significant homology, Gottlieb (1992) has identified a 9-nucleotide motif YUGUUYCUG, common to the 3'UTR of all of these mRNAs. Deletion of this nonamer caused *bcd* mRNA to partially mislocalize, without affecting mRNA stability. This motif is not within the 625 nucleotides required for *bcd* mRNA localization and might therefore be a general mRNA localization signal.

In the early blastoderm stage of *Drosophila* embryos, segmentation gene transcripts also show differing patterns of localization in the periplasm surrounding the cortical layer nuclei. At the end of this stage, the periplasm of an individual cell is sub-divided into apical and basal periplasm by invagination of cell membranes. Sequences in the 3'UTR of the transcripts for three genes, *fushi tarazu*, *hairy*, and *even-skipped*, direct the mRNAs to the apical periplasm (Davis & Ish-Horowicz, 1991). In the *even-skipped* 3'UTR the last 125 nucleotides were shown to confer apical localization to a α-globin reporter transcript. Deletion of 59 bases at the 5'end of this sequence caused hybrid mRNAs to de-localize.

1.2.4.3. Localization of mRNAs for cytoskeletal proteins.

Analysis of the intracellular distributions of mRNAs coding for the cytoskeletal proteins actin, vimentin and tubulin in chicken myoblasts and fibroblasts revealed that these mRNAs were localized in non-random patterns (Lawrence & Singer 1986). Although polyadenylated RNA was homogeneously distributed throughout the cytoplasm, actin mRNA was found localized in the lamellipodia. Vimentin and tubulin mRNAs, in contrast, were distributed near the nucleus and in the peripheral cytoplasm respectively.

Experiments to investigate the mechanisms involved in actin mRNA localization where protein synthesis inhibitors were used have shown that the peripheral localization occurs independently of translation (Sundell & Singer, 1990). The localization signal was determined by fusing chicken β-actin cDNA sequences to the β-galactosidase reporter gene and expressing the chimeric constructs transiently in chick embryonic fibroblasts. These experiments revealed that a 54-nucleotide sequence termed the zipcode, in the β-actin 3'UTR was sufficient to confer peripheral localization to the normally non-localized reporter gene (Kislauskis *et al*, 1993). This localization was disrupted by anti-sense oligonucleotides directed against the zipcode. While protein concentration and mRNA levels were unaffected, the changes in β-actin mRNA distribution caused the lamellae to collapse and cells to become symmetrical, demonstrating that the localized synthesis of the β-actin protein

is essential for maintaining cell polarity (Kislauskis *et al*, 1994). A 68-kDa protein which binds to the proximal (27-nucleotide) half of the zipcode with high specificity has been identified and shown to contain regions homologous to several RNA-binding proteins (Ross *et al*, 1997). The mechanism of localization of β-actin mRNA is energy dependent, since cordycepin, an inhibitor of ATP production prevented this process (Latham *et al*, 1994). The general mechanism of mRNA sorting is thought to involve the formation of a ribonucleoprotein (RNP) complex followed by translocation of this complex to the correct cellular location and anchoring of the RNP to the cytoskeleton (Wilhelm & Vale, 1993). In the case of β-actin the transport and anchoring steps are suggested to involve the actin cytoskeleton since localization is inhibited by disruptors of the actin cytoskeleton but not by inhibitors of microtubules (Sundell & Singer, 1991). The zipcode binding protein could therefore, play a role in anchoring of the β-actin mRNA to the cytoskeleton.

In developing muscle, vimentin filaments rearrange greatly during early myogenesis (Danowski *et al*, 1992). *In situ* hybridization has revealed that the vimentin mRNA location also changes during muscle development. In young myoblasts vimentin mRNA displays a bipolar distribution. This changes to perinuclear in elongated myoblasts and diffuse in developing myotubes (Cripe *et al*, 1993). The differing patterns of distribution suggest that mRNA in one location carries out a function that mRNA elsewhere in the cells cannot and that the location may be important in maintaining differentiated cytoskeletal structures.

Although the 3'UTR of vimentin mRNA has not yet shown to be important for localization, a 61-114-nucleotide region within it is required to specifically bind a 46-kDa protein (Zehner *et al*, 1997). This region is highly conserved from *Xenopus* to man. Therefore the 3'UTR of vimentin mRNA and associated proteins may be important for vimentin mRNA localization.

1.2.5. Other functions of the 3'UTR.

Another functional aspect of the 3'UTR was observed for the muscle structural proteins troponin I, tropomyosin and cardiac actin. The 3'UTRs of these transcripts have been shown by genetic complementation to permit differentiation in a mutant myoblast cell line, which is defective in differentiation (Rastinejad & Blau 1993). The 3'UTR of tropomyosin was most effective in promoting the expression of differentiation-specific genes in myogenic cell lines and inhibiting growth in fibroblasts, suggesting that the 3'UTRs of troponin I, tropomyosin and cardiac actin are *trans*-acting regulators in a feedback loop that can inhibit proliferation and promote differentiation. The 3'UTR of tropomyosin has additionally been shown to suppress tumour growth in a mutant myogenic cell line which gives rise to tumours in mice (Rastinejad *et al*, 1993). In an extension of these observations L'Ecuyer *et al* (1995), demonstrated that chicken embryonic fibroblasts transfected with a plasmid encoding the 3'UTR of skeletal muscle tropomyosin were induced to express skeletal tropomyosin. These cells became spindle-shaped, fused and expressed titin, a marker of striated muscle differentiation. The skeletal muscle tropomyosin and titin organized in sarcomeric arrays. Transfection of osteoblasts with the same plasmid induced them to become spindle shaped but no fusion or organization of these proteins was observed. Therefore skeletal muscle tropomyosin 3'UTR induces trans-differentiation into a striated muscle phenotype in a cell-type specific manner.

The many examples given above demonstrate the importance of the 3'UTR in the post-transcriptional regulation of gene expression. It can be seen that no single function can be attributed to the 3'UTR as, depending on the mRNA type, this region acts as a repository of signals determining mRNA stability, mRNA translation, mRNA localization and control of polyadenylation. In embryonic development the 3'UTR of some mRNAs contains multiple signals that determine localization and translational repression.

1.3. The structure and function of the Myosin Heavy Chain.

Myosin is a highly conserved ubiquitous protein found in all eukaryotic cells where it functions as a molecular motor to generate the force for cellular movements. Sarcomeric myosin is the major protein involved in contraction. of skeletal and cardiac muscles and is composed of two heavy chain subunits and four light chain subunits, which form a hexameric molecule. The heavy chain subunits contain ATPase activity and provide the energy requirements that drive the contractile process. They exist as a number of isoforms, which are temporally and developmentally regulated and define muscle fibre phenotype. An outline of the function and properties of muscle and the myosin heavy chains are given below.

1.3.1. Skeletal muscle ultrastructure.

Adult skeletal muscles are composed of bundles of myofibres, which run from tendon to tendon and provide the characteristic contractile activity of skeletal muscle. Each fibre is a single multinucleated, membrane-bound cell. When observed under the microscope they have a characteristic striated appearance. The fibres contain numerous myofibrils, which are composed of repeated contractile units called sarcomeres. The organized arrangement of the sarcomeres gives rise to the characteristic striated appearance of skeletal muscle. Although the ultra-structure and molecular composition of the sarcomeres is similar among different muscle types, there is a high degree of molecular variability due to the existence of multiple forms of myofibrillar components. The sarcomere is a complex structure composed of numerous proteins arranged in a precise manner (figure1.4.). It is bordered on either side by a line known as the Z disk, and each Z disk bisects a lighter band termed the I band, since it is isotropic in polarized light. At the centre of the sarcomere is a dark band called the A band (anisotropic in polarized light). The A band is bisected by the H zone which itself contains a denser region called the M line. Within the A band is an array of thick and thin filaments. The main constituent of the thick filaments is myosin. Additional proteins are associated with myosin including C- and H-proteins, and titin. The thin filaments are composed of actin, which also has associated

proteins, of which troponin and tropomyosin are the most important in the regulation of contraction. The main constituent of the Z disks is α-actinin. The sliding action of actin and myosin relative to one another without change in length of either filament is responsible for muscle contraction at the level of the sarcomere (Huxley & Hanson, 1954; Huxley & Niedergerke, 1954). They are responsible for converting chemical energy into mechanical force during muscle contraction. The sliding action of the thick and thin filaments is facilitated by projections present on the myosin molecules, known as cross-bridges. These cross-bridges act in a cyclical manner, attaching to actin, changing conformation, which pulls on actin, detaching, and then re-attaching further along the filament to repeat the cycle. The conformational changes of the cross-bridges on actin provide the propulsive force between filaments, with each cycle being powered by the hydrolysis of ATP.

Since this thesis is concerned with the myosin heavy chain genes, the introduction concentrates on myosin. The other components of the sarcomere are discussed elsewhere (Craig, 1994, in Myology, volume 1; Schiaffino & Reggiani, 1996).

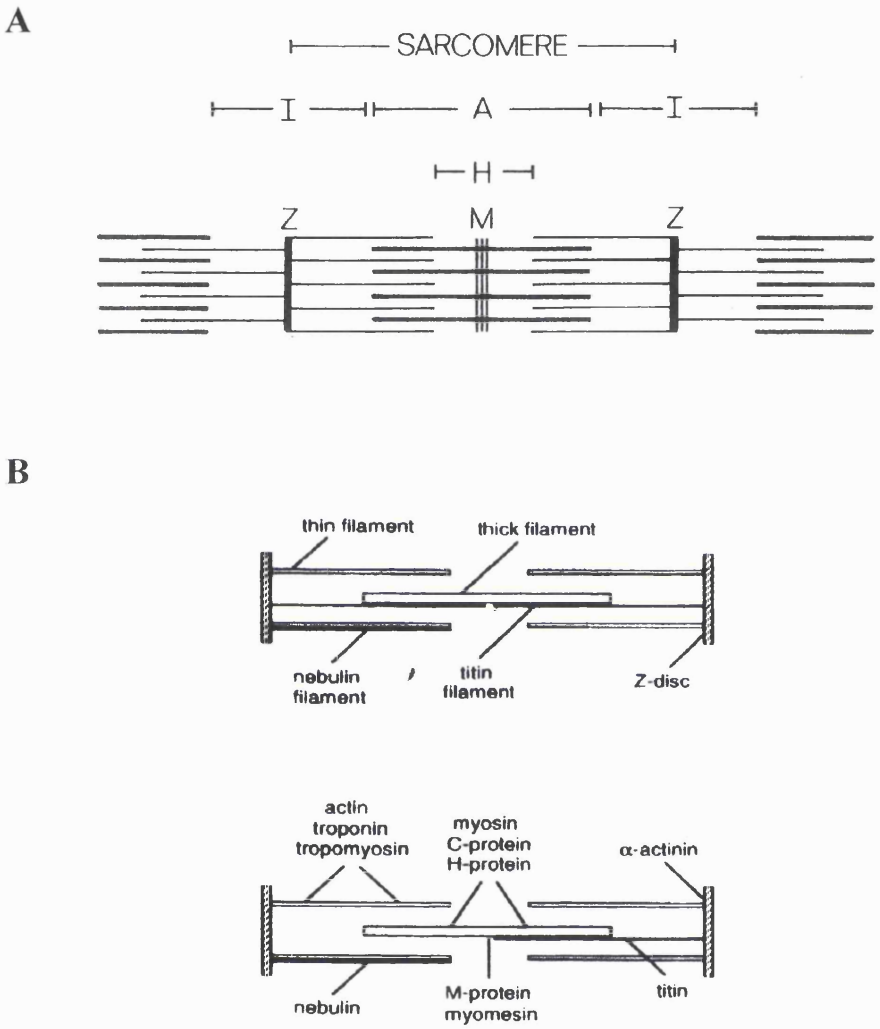


Figure 1.4. Schematic diagrams showing the banding pattern and structure of the sarcomere (A), and the arrangement of the main components of the within the sarcomere (after Schiaffino & Regianni, 1996).

1.3.2. Muscle fibre types.

Skeletal muscles are composed of arrays of different types of fibre, and this in turn is responsible for the unique contractile properties of the muscle as a whole, and gives rise to different muscle types. In adult muscle, fibre types have been classified according to physiological, biochemical and histochemical properties. In physiological terms, muscle fibres are classed as fast or slow depending on their speed of contraction. The sarcomeres of fast and slow muscle fibres differ intrinsically in their speed of shortening by 2 to 3-fold. The biochemical basis for the difference in speed results in part from the differential expression of myofibrillar proteins, and in particular the myosin heavy chain (MyHC), (Bandman, 1992; Moss *et al*, 1995; Schiaffino *et al*, 1988). The MyHCs are largely responsible for the ATPase activity and hence the contraction rate of individual fibres (Bottinelli *et al*, 1991; Reiser *et al* 1985a; Reiser *et al* 1985b). Therefore, the MyHC content of a fibre is the molecular determinant of fibre type. The major fibre types in skeletal muscle include one slow (type I), and three fast types (IIA, IIB, IIX) (Gauthier, 1994 in Myology voll; Pette & Staron, 1990.) Each contains a different MyHC isoform encoded by a distinct gene (Schiaffino *et al*, 1989; Termin *et al*, 1989; Saez & Leinwand, 1986). Slow type I muscle fibres contain MyHC isoforms which have a low ATPase activity, whilst fast type II fibres contain MyHC isoforms with high ATPase activities (Barany 1967; Reiser *et al*,1985). In addition to fibres containing a single MyHC isoform, there are hybrid fibres, which co-express different MyHCs, resulting in a spectrum of MyHC combinations (Schiaffino & Reggiani, 1994.).

1.3.3. Muscle fibre-type diversity.

Muscle development is characterized by the asynchronous differentiation of successive fibre generations. Skeletal muscle fibres are formed by temporally discrete rounds of myoblast fusion to form multinucleated myotubes that mature into myofibres (reviewed by Stockdale, 1992; Gunning & Hardeman, 1991). In rodent and avian limb muscles, the first round of myoblast fusion gives rise to primary myotubes, which begin to form before the cleavage of the muscle mass into distinct muscles. After the primary myotubes have formed, A second round of fusion is

initiated to form secondary myotubes (Condon *et al*, 1990). In humans a third round of myoblast fusion can give rise to what have been called tertiary myotubes (Draeger *et al*, 1987). Primary and secondary myotubes differ by the MyHCs they express. In primary myotubes, the first MyHC gene expressed encodes the embryonic (MyHC-emb), isoform. Subsequently a neonatal (MyHC-neo) isoform, and then the slow isoform MyHC-β are observed in all primary myotubes except those in the soleus muscle. The mutually exclusive continued expression of either MyHC-β or MyHC-neo determines a fast or slow fibre-type. The fast lineage fibres undergo a step to determine the type of adult fast MyHC they will express at a later stage. Secondary fibres express MyHC-emb and MyHC-neo simultaneously at first, and then a decision is made whether to express MyHC-β. MyHC-β-expressing fibres are fated to become slow. Those fibres that are restricted to expressing MyHC-emb and MyHC-neo are fated to become fast fibres and subsequently express adult fast isoforms MyHC-2A, MyHC-2B, and MyHC-2X.

The transcripts of the three fast MyHC isoforms show a distinct pattern of distribution in different muscles and fibres form the earliest stages of development (DeNardi *et al*, 1993). The adult fast isoforms are initially co-expressed with MyHC-emb and MyHC-neo. These developmental isoforms then disappear, with the time course of down-regulation varying according to fibre type and muscle type. The isoform profile of skeletal muscle fibres undergoes further changes with maturation and ageing. For example, in the rat soleus muscle, fast fibres expressing MyHC-2A are progressively transformed into fibres expressing exclusively MyHC-β throughout post-natal development (Butler-Browne & Whalen, 1984). Similarly, fibres in fast muscles of the mouse expressing MyHC-β also disappear with age, and in the rat a decrease in MyHC-2B correlates with an increase in MyHC-2X in ageing fast skeletal muscles (Whalen *et al*, 1984; Larsson *et al*, 1993).

Additional factors such as hormonal signals and innervation affect fibre-type diversity and MyHC isoform transitions during development (see review by Gunning & Hardeman, 1991). In particular thyroid hormones appear to accelerate the normal maturation process. In the absence of thyroid hormone, the transition from MyHC-

neo to adult fast MyHC is greatly reduced, but can be accelerated by administration of thyroid hormone (Butler-Browne *et al*, 1984; Gambke *et al*, 1983). This effect is independent of innervation, since it is also observed in denervated muscles (Russell *et al*, 1988).

Studies by Buller *et al*, 1960, showed that switching a nerve from a slow-twitch muscle to innervate a fast-twitch muscle resulted in the transformation of the muscle to a slow phenotype, and conversely a nerve from a fast-twitch muscle transformed a slow-twitch muscle to fast phenotype. This is partly explained by the fact that fast and slow nerves have different firing patterns (Pette & Vrbova, 1985). These results showed that innervation was important in determining muscle fibre type. Neural influences also influence muscle fibre growth and differentiation during development. In the absence of innervation, secondary fibres and some primary fibres disappear, and some slow fibres lose their slow myosin. However it has been observed that fibre type diversification can occur in the absence of neural influences. In rat embryos, fibre types are formed even when treatment with neurotoxins to destroy innervation before invasion of the muscle by motor axons, is carried out (Condon *et al*, 1990). In the mouse, removal of motoneurons during the fetal stages does not interfere with the transition from expression of the developmental MyHCs to the adult MyHC-2B (Weydert *et al*, 1987).

The expression of the MyHC genes can be modulated in the adult stages by electrical stimulation and mechanical factors in addition to innervation and thyroid hormone. For example the fast MyHC-2X which is not normally found in the rat soleus muscle can be induced in this muscle by either thyroid hormone, high frequency electrical stimulation, or mechanical unloading after hindlimb suspension (Ausoni *et al*, 1990; Campione *et al*, 1993; Izumo *et al*, 1986). The regulatory mechanisms responsible for fibre-type specific expression are partly modulated by transcription factors.

1.3.4. Myosin.

The myosin molecule is composed of two myosin heavy chain (MyHC) subunits (approximately 220 kDa) and two pairs of non-identical myosin light chain

(MLC) subunits (17-23 kDa), referred to as the essential light chains and regulatory light chains (Lowey, 1994 in Myology). The MyHC provides both the motor and filament forming functions of the myosin molecule. The heavy chain subunits form a structure with two domains (figure 1.5.). The globular amino terminal head domain corresponds to the motor domain and contains the ATP-binding site and the actin-binding site. The long α-helical, coiled-coil carboxyl rod domain contains the filament forming properties of myosin. The light chains are associated with the neck region of myosin (Katoh & Lowey, 1989; Rayment *et al*, 1993). The MLCs also exist in multiple isoforms that are differentially expressed in different fibre types.

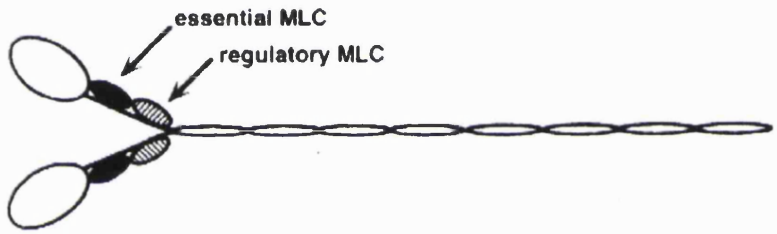


Figure 1.5. A schematic representation of the structure of the myosin molecule. Myosin consists of two myosin heavy chain subunits, two essential myosin light chains, and two regulatory myosin light chains. The heavy chains have a globular head domain, which contains ATP- and actin-binding sites. The light chains are bound to the neck region (from Schiaffino & Reggiani, 1996)

1.3.5. The myosin light chain isoforms.

Five major myosin essential light chains have been identified in mammalian striated muscle (Barton & Buckingham, 1985). Fast skeletal muscle fibres contain the isoforms MLC-1fast and MLC-3fast, whereas slow skeletal muscle fibres contain MLC-1slow/ventricular (also called MLC-1slow-b) and MLC-1slow-a. MLC-1slow-b is also expressed in ventricular myocardium (Hailstones & Gunning, 1990). A developmental isoform MLC-1emb/atrial also exists. This isoform, as its name suggests, is expressed in developing skeletal muscle and the atrial myocardium. The

suggests, is expressed in developing skeletal muscle and the atrial myocardium. The essential MLCs are encoded by distinct genes, apart from MLC-1fast and MLC-3fast, which are transcribed by alternative utilization of transcription initiation sites in a single gene, (Periasamy *et al*, 1984). There are four known isoforms of regulatory MLC. MLC-2fast and MLC-2slow are expressed in skeletal muscle with MLC-2slow being additionally expressed in the ventricular myocardium. MLC-2atrial is present in the atrial myocardium. The distinct isoform MLC-2m has been detected in mandibular muscles of carnivores (Rowlerson *et al*, 1981).

1.3.6. The myosin heavy chain isoforms.

The myosin heavy chains are encoded by a multigene family, whose members exhibit temporal and tissue-specific expression (Bandman, 1985; Mahdavi *et al*, 1987). The MyHC gene family is comprised of several genes, which code for distinct isoforms. There are two MyHC isoforms expressed in cardiac muscle, MyHCβ/slow and MyHCα. Chromosomally they are located closely linked on chromosome 14 in both human and mouse (Saez *et al*, 1987; Weydert *et al*, 1985; Gulick *et al*, 1991). The genes encoding these isoforms are differentially regulated during expression and have differing patterns of distribution. MyHC-α is the predominant isoform in adult ventricular and atrial myocardium of small mammals (Lompre *et al*, 1984). Although its expression is almost exclusively restricted to cardiac muscle, low levels of MyHC-α expression are seen in extraocular and masseter muscles (d'Albis *et al*, 1993; Rushbrook *et al*, 1994). MyHC-β is abundant in the ventricular myocardium of all mammals during foetal development, and is the predominant MyHC isoform in the ventricular myocardium of adult large mammals. In addition MyHC-β is also expressed in slow (type I) skeletal muscle fibres. (Lompre *et al*, 1984) have shown that this isoform is identical to that expressed in the ventricular myocardium.

The skeletal MyHC genes are clustered on chromosome 17 in humans and chromosome 11 in the mouse (Leinwand *et al*, 1983; Weydert *et al*, 1985). The major MyHCs expressed in skeletal muscle are the developmental isoforms

embryonic MyHC (MyHC-emb) and neonatal MyHC (MyHC-neo), slow type 1/β-cardiac (MyHC-β/slow) and fast types MyHC-2A, MyHC-2B and MyHC-2X, although no MyHC-2B isoform has been shown to exist in human skeletal muscle (Ennion *et al,* 1995). The developmental MyHC isoforms, MyHC-emb and MyHC-neo, were first identified at the protein level in the rat (Whalen *et al*, 1981) and then at the mRNA level in other species as well (see review by Emerson & Bernstein, 1987). Although these isoforms are expressed predominantly at the embryonic and neonatal stages of development, and transiently during regeneration (Whalen *et al*, 1981; Periasamy *et al*, 1984; Weydert *et al*, 1987) they are found in some fibres of extraocular and masseter muscle (Sartore *et al*, 1987; Butler-Browne *et al*, 1988).

1.3.7. The myosin heavy chain genes.

The MyHCs are conserved proteins and this is reflected in the structural homology of the genes that encode the distinct isoforms in a number of species. MyHC genes that have been sequenced show similar sizes at the level of DNA and mRNA. The MyHC gene is approximately 24 kb and the transcripts are about 6 kb. The coding regions of the genes are highly homologous, with corresponding isoforms in different species being more homologous than different isoforms within a species. For example the rat and human embryonic MyHCs share 74, 90 and 80% identity in their 5' untranslated, coding and 3' untranslated regions respectively (Stedman *et al,* 1990). The 5' and 3' untranslated region, in contrast show more sequence divergence between the transcripts of various isoforms. Hence, these regions have been used to generate isoform specific probes for studying the distribution and expression of specific MyHC isoforms (McKoy *et al*, 1998). However, it is of interest to note that the 3' untranslated regions in particular, show remarkable sequence similarities between the same isoform across species (DeNardi *et al*, 1993; Saez & Leinwand 1986), suggesting a functional role. Furthermore common sequence motifs are found in the 3'UTRs of some isoforms which are found in the corresponding isoforms from different species (Weydert *et al,* 1983; Kavinsky *et al*, 1983; Saez & Leinwand 1986; DeNardi *et al*, 1993).

1992; Buckingham, 1996; Rudnicki & Jaenisch, 1995). Following proliferation, myoblasts withdraw from the cell cycle and start to differentiate to form multinucleated myofibres. This is accompanied by the activation and expression of muscle-specific genes. The MyoD family of transcription factors are responsible for inducing skeletal myogenesis, and have the ability to convert non-muscle cells into muscle cells. They have a sequence homology for the basic-helix-loop-helix (bHLH) regulatory motif. The basic domain allows binding to DNA and the helix-loop-helix domain allows the formation of heterodimers with other transcription factors such as E12 (see review by Olson, 1992; Dias *et al*, 1994;. Weintraub *et al*, 1991). These MyoD transcription factors bind a CANNTG sequence motif called the E-box present in the promoters and enhancers of muscle genes. They are expressed at different periods during development, with Myf-5 the first to be expressed in the somites, and MRF4 being expressed at the late foetal stage of development. Myf-5 and MyoD are more homologous to each other than the others an are expressed before differentiation. Myogenin and MRF4 show more homology to each other and are expressed upon differentiation (see review by Megeney & Rudnicki, 1995). While Myf-5 expression is diminished, and MyoD and myogenin are reduced during development, MRF4 is still expressed at high levels in both fast and slow adult muscles (Voytik *et al*, 1993). It has been suggested that MyoD and myogenin are responsible for fibre-type specific gene expression since myogenin transcripts are more abundant in slow type I fibres and MyoD transcripts are expressed at higher levels in fast type II fibres. In addition a corresponding change in MyoD and myogenin is observed when fibre-type transformations are induced by thyroid hormone or cross-innervation. (Hughes *et al*, 1993; Voytik *et al*, 1993).

1.3.8. Post-transcriptional regulation of MyHC gene expression.

The formation of skeletal muscle fibres occurs in two stages, with the formation of primary and then secondary myotubes. Once they are formed myofibres undergo a process of specialization, giving rise to individual fibre types in the adult. During this process of development the different MyHC genes are individually transcriptionally regulated and expressed giving rise to the different phenotypes. The

During this process of development the different MyHC genes are individually transcriptionally regulated and expressed giving rise to the different phenotypes. The developmental transitions are reflected in relative changes in mRNA levels. Although the regulation of MyHC gene expression occurs mostly at the level of transcription, it has been postulated to occur at the post-transcriptional level as well. Studies by Lyons *et al*, (1990) on the expression of myosin genes in developing skeletal muscle in the mouse embryo demonstrated that perinatal MyHC transcripts were detectable at 10.5 days post coitum (d.p.c). Vivarelli *et al,* (1988), however showed that the perinatal MyHC protein was not detectable in 10-day somites or 13-day limb buds. Furthermore, Harris *et al* (1989) showed that the first detectable accumulation of the perinatal MyHC protein occurred at 16 d.p.c. Therefore, these observations and those of Roy *et al*, 1984 suggest post-transcriptional controls of gene expression for myosin genes during development, through changes in stability and/or rate of translation. The mechanisms that regulate the accumulation of the MyHC mRNAs differ. Cox *et al*, 1991 have demonstrated that MyHC-emb, MyHC-neo and MyHC-2B showed independent transcriptional and post-transcriptional regulation during differentiation of a mouse muscle cell line. Although there were high levels of MyHC-emb gene transcription, the levels of corresponding cytoplasmic mRNA were relatively low. During differentiation, despite low levels of transcription of the MyHC-neo gene, cytoplasmic mRNA accumulated to levels greater than that for MyHC-emb, indicating different mechanisms involved in regulating accumulation of these mRNAs. Similarly MyHC-2B was shown to accumulate in older myotubes. Studies by Medford *et al*,(1983)attributed an increase in the effective stability of MyHC mRNA during myogenesis in experiments with a rat skeletal muscle cell line to withdrawal of myoblasts from the cell cycle prior to fusion. The large number of MyHC transcripts are observed in chick myoblasts prior to cell fusion have been shown to be stored mainly in the form of messenger RNA protein complexes (mRNPs) (Dym *et al*, 1979). These transcripts are transferred to the polysomal fraction in differentiating myoblasts (Doetschman *et al*, 1980).

1.4. Aims.

The MyHC isoforms are a major determinant of the contractile and metabolic characteristic of skeletal muscle fibres. As discussed earlier, they are expressed in a developmentally and temporally regulated manner. Furthermore, muscle fibres are capable of adapting their phenotype in response to factors such as neuronal activity, stretch, electrical stimulation, and thyroid hormones by altered patterns of MyHC gene expression. Although the regulation of expression the MyHC genes is thought to occur mainly at the level of transcription, experimental data has indicated that regulation of MyHC gene expression also occurs at the post-transcriptional level. The aim of this study was to investigate the role of the MyHC 3' untranslated region in post-transcriptional regulation. The first part of this introduction has highlighted the various functional roles attributed to the 3'UTRs of different genes. Observations by several workers that this region of the MyHC mRNAs contain conserved sequence motifs further suggests some functional role for this region in MyHC gene expression. Mouse skeletal muscle was used as a source of tissue for the isolation of MyHC 3' untranslated region sequences. These were then used to investigate functional aspects of the MyHC 3'UTR.

Chapter 2

Isolation and Sequence Analysis of Myosin Heavy Chain 3' Untranslated Regions

2.1. Introduction.

The Polymerase Chain Reaction (PCR) was originally developed to amplify defined regions of DNA *in vitro* from minute amounts of starting material, allowing the analysis of low abundance gene sequences (Saiki *et al*, 1985; Mullis. *et al*, 1986; Mullis and Faloona, 1987). Since the introduction of the original PCR, the basic protocol has been further developed and adapted, and can be applied to study mRNA (Innis *et al*, 1990). PCR can be used in conjunction with reverse transcription (RT-PCR) to amplify specific region of mRNA. The mRNA is first transcribed into single stranded complimentary DNA (first strand cDNA) by primer extension using an oligonucleotide specifically designed to copy the mRNA, or oligo-dT to prime at the poly A tail, or by using random hexamers (random priming). If the two ends of the mRNA sequence to be amplified are known, specific primers can be used to amplify from the first strand cDNA. Where only the sequence at one end is known the RACE (Rapid Amplification of cDNA Ends) protocol can be applied. (Frohman *et al*,s 1988). The RACE method allows the amplification of nucleic acid sequences from an mRNA template between a specific site within the transcript and unknown sequences at either the 5' or 3' end.

To amplify sequences at the 5' end (5' RACE), a homopolymer tail is added to the 3' end of the first strand cDNA using terminal transferase, or alternatively, specific linkers can be added with T4 RNA ligase (Dumas Milne Edwards *et al*, 1991). PCR is then performed with a gene specific primer and a primer that anneals to the homopolymer tail or specific linker sequence.

The methods for amplification of 3'ends (3' RACE) takes advantage of the fact that most mRNAs have a poly A tail. The reverse transcription step uses a chimeric primer composed of an oligo-dT sequence at the 3' end and a specific sequence at the 5' end. Amplification by PCR from the first strand cDNA is carried out with a primer whose sequence is the same at the 5' sequence of the chimeric primer and a gene specific primer.

The results obtained from RACE depend not only on the conditions of on PCR, but also on the integrity of the RNA, which will ultimately affect the quality of

the cDNA synthesised. Therefore, it is essential to isolate reasonably pure undegraded RNA by effective disruption of tissue, denaturation of nucleoprotein complexes, inactivation of endogenous ribonucleases (RNases), and removal of contaminating DNA. Of these the inactivation of endogenous RNases, which are released from membrane bound organelles upon cell disruption, is the most critical.

The aim of the work described in this chapter was to employ the 3'RACE protocol to amplify the 3' sequences of the different myosin heavy chain mRNAs from mouse skeletal muscle as a first step to studying the role of the myosin heavy chain 3' untranslated regions in post-transcriptional regulation of gene expression. The gene specific primers used in 3'RACE to amplify MyHC 3' ends were originally designed from the C terminal coding region of the MyHC genes which is well conserved between different isoforms within a species and between different species. Amplified products were cloned and characterized by sequence analysis.

2.2. Materials and methods.

2.2.1. Source of tissue.

The soleus, gastrocnemius, tibialis anterior and diaphragm muscles were excised from adult C57BL/10 black mice and frozen immediately in liquid nitrogen to prevent degradation of RNA by endogenous RNases.

2.2.2. Total RNA isolation.

The single step method of RNA isolation using acid guanidium thiocyanate-phenol-chloroform after Chomczynski and Sacchi (1987) was used to isolate total RNA. Precautions were taken to prevent the inadvertent introduction of, and contamination by, RNases by wearing disposable gloves at all times. All glassware was baked at 200° C overnight. Eppendorf tubes and pipette tip were autoclaved and baked overnight at 80° C. Solutions, apart from those containing primary amines, were treated with diethylpyrocarbonate (DEPC) 0.1% v/v.

Frozen samples of mouse muscle, weighing approximately 40-90 mg, were crushed and homogenised in 1ml of denaturing solution containing 4 M guanidium thiocyanate, 25 mM sodium citrate pH 7.0, 0.5% sarcosyl, and 0.1 M 2-mercaptoethanol, in a 1.5ml Eppendorf tube using a Kontes® pellet pestle. Then 33µl of 3 M sodium acetate pH 4.8 (DEPC treated), 1ml of phenol and 0.2 ml chloroform-isoamyl alcohol (24:1) were sequentially added to the homogenate, with thorough mixing after the addition of each reagent. The final suspension was vortexed vigorously three times for 10 seconds, and cooled on ice for 15 minutes.

The samples were then centrifuged at 14000 rpm for 30 minutes at 4°C. After centrifugation the upper aqueous phase, which contains the RNA was removed, taking care not to disturb the interphase (which contains DNA) and transferred to a fresh Eppendorf tube. Two volumes of ice-cold absolute ethanol were added to the samples, which were then placed at -20° C overnight to precipitate the RNA.

The precipitated RNA was centrifuged at 14000 rpm for 30 min. at 4° C. The supernatant was discarded and the resultant RNA pellet dissolved in 200µl of denaturing solution. To determine RNA purity and concentration, an aliquot of this

was precipitated with two volumes absolute ethanol, washed with 70% ethanol 3 times and resuspended in an appropriate volume of DEPC water. The OD at 260 nm and 280 nm was then measured. The rest of the RNA was stored as a precipitate at -70°C by adding 400µl of absolute ethanol.

2.2.3. Oligonucleotides for 3'RACE PCR.

The oligonucleotides RoRi-dT_{17} and Ro were synthesised according to Harvey *et al* (1991). RoRi-dT_{17}, which consists of 17 thymidine residues with an adapter sequence at the 5' end, was used to prime first strand cDNA synthesis. The Ro primer corresponds to the first 25 bases of RoRi-dT_{17}, and was used in PCR together with the either of the gene specific primers FG2Exn40 (Ennion *et al*, 1995), or RnExn40 (McKoy *et al*, 1998) in order to amplify the 3' ends of the myosin heavy chains. These gene specific primers were designed from conserved sequences in exon 40 of the MyHC genes.

RoRi-dT_{17}

5' TCGATGGTCGACGCATGCGGATCCAAAGCTTGAATTCGAGCTTTTTTTTTTTTTTT 3'

Ro

5' ATCGATGGTCGACGCATGCGGATCC 3'

FG2Exn40

5' AGGAAGGTCCAGCACGAACTGGAGG 3'

RnExn40

5' CGCAAGGTGCAGCACGAGCTGGAT/C 3'

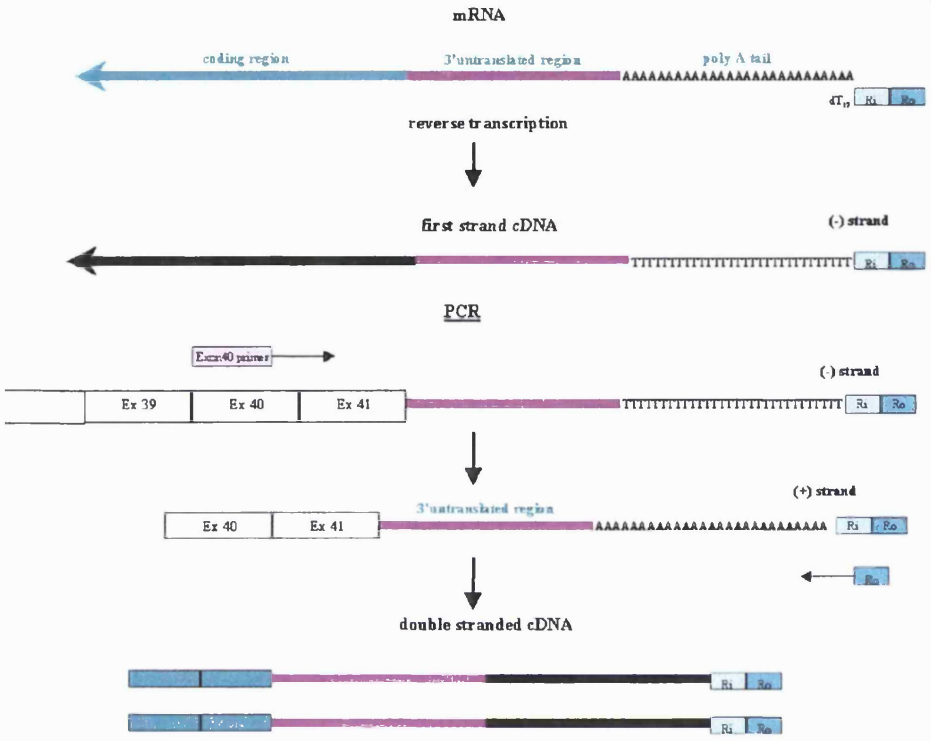


Figure 2.1. Schematic diagram showing the 3'RACE protocol employed to amplify the 3' sequences of MyHC mRNAs. Total RNA was extracted form mouse skeletal muscle and used in a reverse transcription reaction to synthesize first strand cDNA. Amplification of 3' sequences was performed using gene specific primers. The location of the gene specific primers in relation to MyHC mRNA sequence are indicated

2.2.4. First strand cDNA synthesis.

To prepare the RNA for cDNA synthesis the appropriate volume of precipitate was centrifuged at 14000 rpm for 10 minutes at 4° C. After removing the supernatant the RNA pellet was washed three times in ice-cold 70% ethanol and following removal of all traces of supernatant, the pellet was briefly vacuum dried. The RNA was dissolved in 7µl of RNase-free water, heated at 70° C for 10 min. and rapidly cooled on ice to eliminate secondary structure prior to cDNA synthesis.

The cDNA synthesis reactions were assembled on ice in a 20 µl volume containing buffer (final concentration 10 mM Tris-HCl pH 8.3, 15 mM KCl, 0.6 mM MgCl), 10 mM DTT, 0.5 mM each dGTP, dATP, dTTP, dCTP (Boerhinger), 66 µM R_OR_i-dT_{17} primer, 40 units RNase inhibitor (Promega), 4µg total RNA, and 200 units M-MLV reverse transcriptase (GIBCO BRL) per µg RNA. Control reactions were set up as above but omitting the addition of reverse transcriptase. The reactions were incubated at 37° C overnight and subsequently stored at -20° C.

2.2.5. RACE-PCR.

Double-stranded cDNA was amplified from the first strand cDNA reaction in a PCR using either FG2Exn40 or RnExn40 primers. Reactions were performed in a 100µl volume containing reaction buffer (final concentrations 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 0.2 mM each of dATP, dTTP, dCTP, dGTP Boerhinger, 25µM each Ro and exon 40 primers and 10µl first strand cDNA. The reaction components were gently mixed and briefly centrifuged. 100µl of mineral oil were overlaid on top of the reaction mixture to reduce evaporation during the PCR. The samples were heat denatured at 94° C for 5 minutes and then 2.5 units of Taq polymerase (Boerhinger) were added. Amplification was then carried out by 32 cycles of denaturation at 94° C for 1 minute followed by annealing at 55° C for 1 minute, followed by elongation at 72° C for 2 minutes. A final elongation step at 72° C for 10 minutes was included to ensure all PCR products had 3' A overhangs. PCR products were analysed by electrophoresis of an aliquot of the reaction on a 1% agarose gel containing ethidium bromide (10 mg/ml).

2.2.6. Cloning of PCR products.

Amplification products can be efficiently cloned into specifically designed vectors using the TA cloning method (Mead *et al*, 1991). Since Taq DNA polymerase has a non template-dependent activity which adds a single adenosine residue to the 3' ends of PCR products (Clark, 1988), it is possible to ligate PCR products into linearised vectors that have a single thymidine residue added to their 3' ends. The TA cloning system (Invitrogen) was used to clone 3'RACE products obtained from the amplification reactions. PCR products were electrophoresed on a 1% agarose gel and bands of interest excised and purified using Wizard PCR purification kit (Promega).7µl of purified PCR product was ligated to 50 ng of pCR™II vector in a 10µl volume, in the presence of ligation buffer, (final concentrations 6 mM Tris-Cl, pH7.5, 6 mM $MgCl_2$, 5 mM NaCl, 0.1mg/ml BSA, 7 mM β-ME, 0.1 mM ATP, 2 mM DTT, 1 mM spermidine), and 4 units of T4 DNA ligase (Invitrogen). Reactions were incubated at 14° C overnight.

Transformation of INVαF' One Shot™ competent cells (Invitrogen) was carried out as follows; 2µl of 0.5M beta-mercaptoethanol was added to each vial containing 50µl of competent cells with gentle mixing by stirring. Then 2µl of ligation reaction were pipetted directly into the cells. The vials were incubated on ice for 30 minutes and then heatshocked for 30 seconds at 42°C, followed by incubation on ice for 2 minutes. 450µl of room temperature SOC medium was added to each vial, and they were then incubated in a rotary shaker for 1 hour at 37°C. Cells were spread on LB agar plates containing 50µg/ml ampicillin which had been spread with 40µl X-gal (40mg/ml) The plates were incubated overnight at 37° C.

Positive transformants were initially identified by blue/white colony selection (Yanisch-Perron *et al*, 1985). Small-scale plasmid preparations were then carried out from overnight cultures using the Wizard Minipreps DNA purification System (Promega). The presence of PCR product insert was verified by restriction digestion of plasmid DNA and electrophoresis on a 1% agarose gel containing ethidium bromide.

2.2.7. Sequencing of positive plasmid clones.

Double stranded sequencing of clones using the dideoxy chain termination method (Sanger *et al*, 1977) was carried out to identify the PCR products. 3-5 μg Plasmid DNA was denatured using the alkaline-denaturation method (Lim & Pene, 1988) by addition of 1/10 volume 2 M NaOH, 2 mM EDTA and incubation at 37°C for 30 minutes. 1/10 volume 3M NaAcetate, pH 4.8, was added to stop the reaction and the DNA was precipitated with two volumes of ethanol at -70°C for 15 minutes. After washing the pellet in 70% ethanol, and briefly air drying, the DNA was resuspended in 7μl distilled water. Sequencing reactions were performed using the Sequenase Version 2.0 kit (Amersham). T7 and Sp6 primers were used for sequencing. 2 pmol of primer was annealed to denatured plasmid by mixing them together in reaction buffer (final concentrations 40 mM Tris-HCl, pH 7.5, 20 mM $MgCl_2$, 50 mM NaCl) in a 10μl volume, heating for 2 minutes at 65° C and allowing to cool slowly to less than 35° C over 15-30 minutes. While cooling tubes were filled with 2.5μl of each termination mix (80μM dGTP, 80μM dATP, 80μM dCTP, 80μM dTTP, 50 mM NaCl, and 8μM of either ddGTP, or ddATP, or ddCTP, or ddTTP). These were pre-warmed in a 37° C water-bath. The annealed DNA mixture was chilled on ice and then to it was added 1μl of 0.1 M DTT, labelling mix (1.5 μM each of dGTP, dCTP, and dTTP). 0.5μl of [^{35}S]dATP and 2μl Sequenase polymerase (diluted 1:8 with 10 mM Tris-HCl, pH 7.5, 5 mM DTT, 0.5 mg/ml BSA buffer). The reaction was mixed gently and incubated at room temperature for 2-5 minutes. 3.5μl of the labelling mix was added to each of the pre-warmed termination tubes (G, A, T and C), mixed and incubated at 37° C for 5 minutes. Reactions were stopped with the addition of 4μl of stop solution (95% formamide, 20 mM EDTA, 0.05% Bromophenol Blue, 0.05% Xylene Cyanol FF. Samples were heated to 75° C for 2 minutes prior to loading on to a 6% sequencing gel and electrophoresis in 1x TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA). After electrophoresis, gels were soaked in 5% acetic acid, 15% methanol for 15 minutes to remove urea prior to vacuum drying and subsequent overnight exposure to autoradiography film.

2.2.8. Computer-based analyses.

Sequence analyses were performed using BLAST searches on the EMBL database. Protein and nucleotide sequence alignments were generated using the CLUSTAL programme (Higgins & Sharp, 1988) on PCGENE. Secondary structure predictions were obtained using the *mfold* programme (Zuker, 1989; Jaeger *et al*, 1989). This programme predicts the secondary structure of RNAs based on the minimum free energy of folding. Default settings were used for the predictions with the paramters set as linear sequence, 1M NaCl ionic conditions and a folding temperature of 37° C.

2.3. Results.

2.3.1. PCR Amplification.

Total RNA was extracted from mouse soleus, gastrocnemius, tibialis anterior and diaphragm muscles of seven week old C57BL/10 black mice. First strand cDNA was synthesised from these RNAs and used in PCR to amplify the 3' ends of myosin transcripts. The results of the amplification are shown in figure 2.2. The PCR included controls for each muscle type cDNA where the addition of reverse transcriptase was omitted in the first strand cDNA synthesis as well as a no template control. Absence of bands in the controls demonstrated that PCR products were amplified from cDNA and not from any contaminating genomic DNA (no reverse transcriptase control) or external contaminants (no template control).

Panel A shows the products amplified using the FG2Ex40 and Ro primers. The size of the PCR products observed were approximately 300 base pairs for all muscles samples. This is the expected size of the PCR products for the 3' regions of the MyHC, as calculated from previously published data (Ennion *et al*, 1995; McKoy *et al*, 1998). These bands were gel purified and cloned for analysis by sequencing

The RnEx40 and Ro primers amplified products that were also about 300 base pairs (Panel B). However, another band of approximately 500 base pairs was additionally amplified from diaphragm cDNA. Increasing the annealing temperature did not eliminate the amplification. Both the 500 and 300 bp bands were agarose gel purified and cloned for sequence analysis.

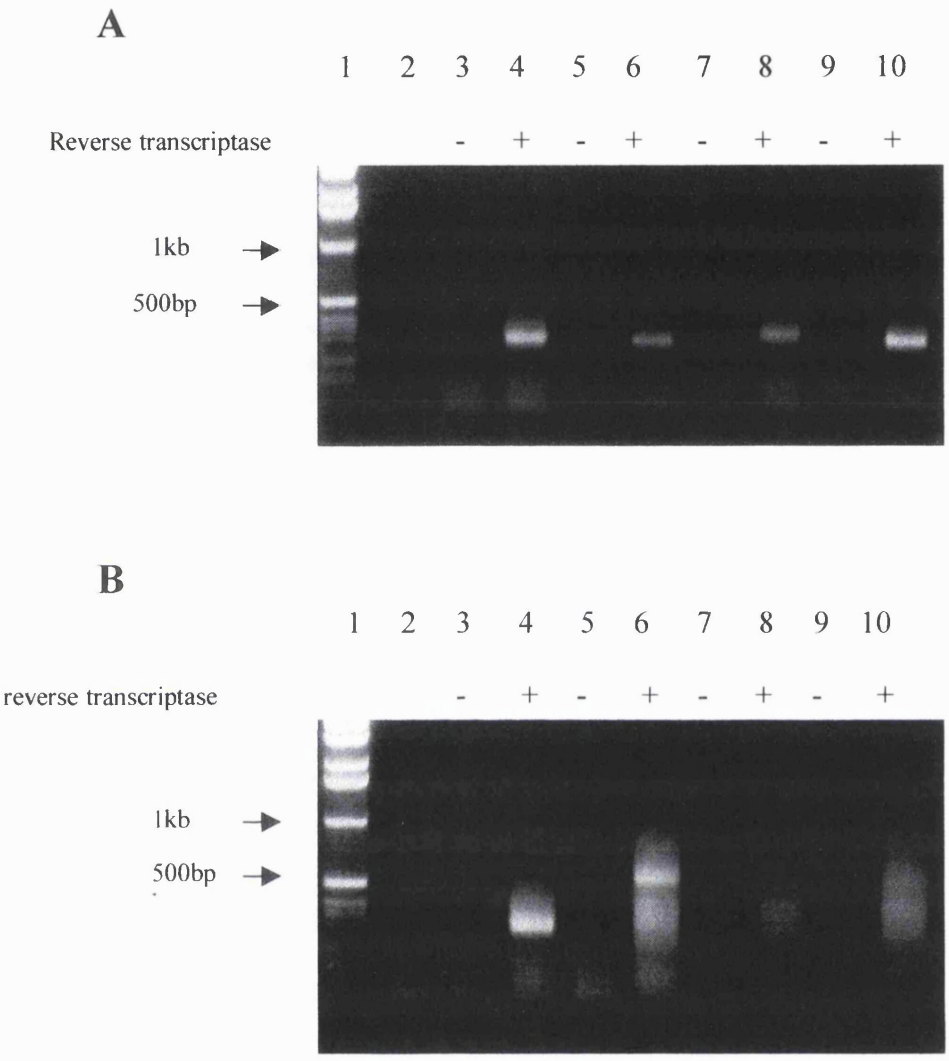


Figure 2.2. PCR amplification of myosin heavy chain cDNA. Total RNA was extracted from four different muscles and converted to cDNA. Amplification was carried out with 32 cycles of denaturation at 94°C for 1 minute, annealing at 55°C for 1 minute, and elongation at 72°C for 2 minutes. Panel A shows PCR products obtained using FG2Exn40 and Ro primer: Lane 1 marker; lane 2 no template control; lanes 3,5,7,9 are no reverse transcriptase controls; lanes 4,6,8,10 amplification from soleus, tibialis anterior, diaphragm and gastrocnemious cDNA respectively. Panel B: amplification using RnExn40 and Ro primer. PCR cycles as for FG2Ex40. Lane 1: marker; lane 2 no template control; lanes 3,5,7,9 no reverse transcriptase control; lanes 4,6,8,10 PCR products from soleus, diaphragm, gastrocnemious and tibialis anterior cDNA.

2.3.2. Sequence analyses.

Sequencing of the cloned 300bp PCR bands indicated that several myosin heavy chains had been specifically amplified. The 500bp band amplified in the Ro/RnEx40 reaction was identified as apolipoprotein A-I. This band could not be eliminated by increasing the annealing temperature, indicating that the RnExn40 primer amplified non-specifically.

Figure 2.3. summarises the different myosin heavy chain isoform 3' ends isolated from the muscles indicated. They were identified according to sequence comparisons and homologies to existing sequences for myosin heavy chain available in the EMBL database. In total four different MyHc isoforms were isolated and identified. More than one isoform was generally isolated from each muscle. As expected no embryonic or neonatal MyHC isoforms were amplified from adult muscle by either of the gene specific primers.

Muscle	Myosin Heavy Chain Isoform
Diaphragm	**SlowType 1/β-cardiac** **Fast 2a, 2x**
Gastrocnemious	**Fast 2a, 2b, 2x**
Soleus	**SlowType 1/β-cardiac** **Fast 2a, 2x**
Tibialis anterior	**Fast 2a, 2b, 2x**

Figure 2.3. Summary of the MyHC isoforms isolated and cloned from the indicated mouse skeletal muscles.

The 3' nucleotide sequences for the slow type 1/ βcardiac, fast 2a, 2b and 2x isoforms together with the deduced amino acid sequences are shown in figures 2.4. 2.5. 2.6. and 2.7. respectively. The sequences shown are the consensus of at least five sequences for each isoform obtained from different subclones from different PCR reactions. These sequences correspond to the last two exons and the entire 3' untranslated regions for each isoform. The length of the 3' untranslated regions for each isoform are 105, 125, 105 and 104 bases for slow type 1, 2a, 2b, and 2x

respectively. It is interesting to note that the poly A signal for MyHc 2x differs from the typical AAUAAA observed in most eukaryotic mRNAs.

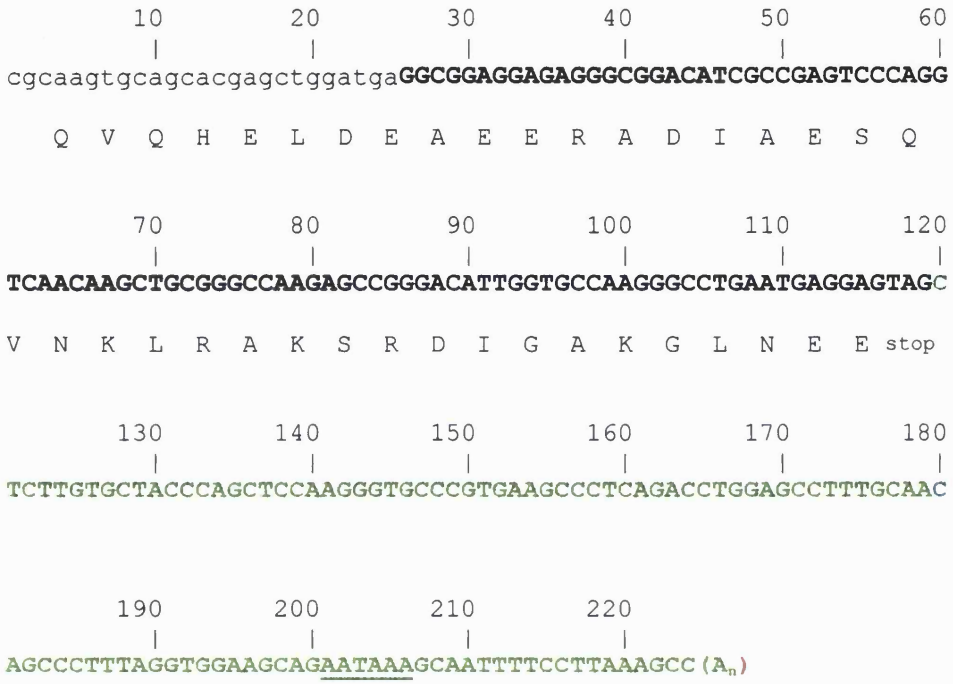


Figure 2.4. Sequence for 3' end of myosin heavy chain identified as slow type I /beta cardiac. Nucleotides in lower case correspond to RbEx40 primer. Amino acids are represented by the single letter code.

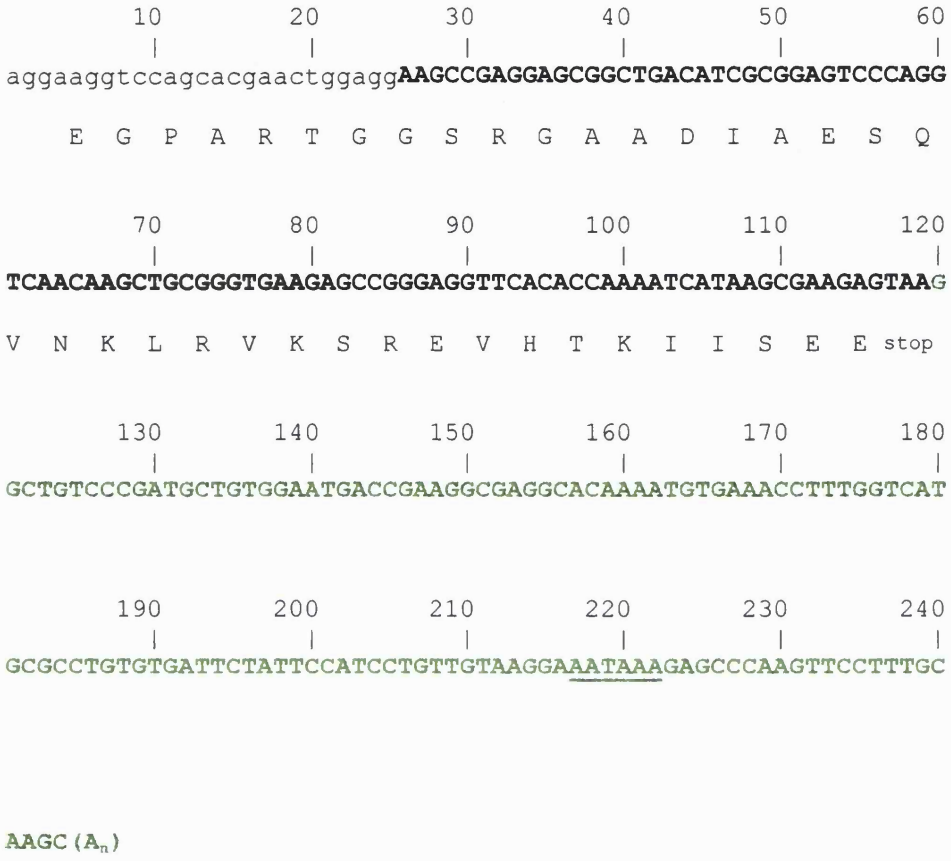


Figure 2.5. Nucleotide and deduced amino acid sequence for 3' end of myosin heavy chain 2a. FG2Ex40 primer is indicated in lower case. Amino acids are represented by single letter code.

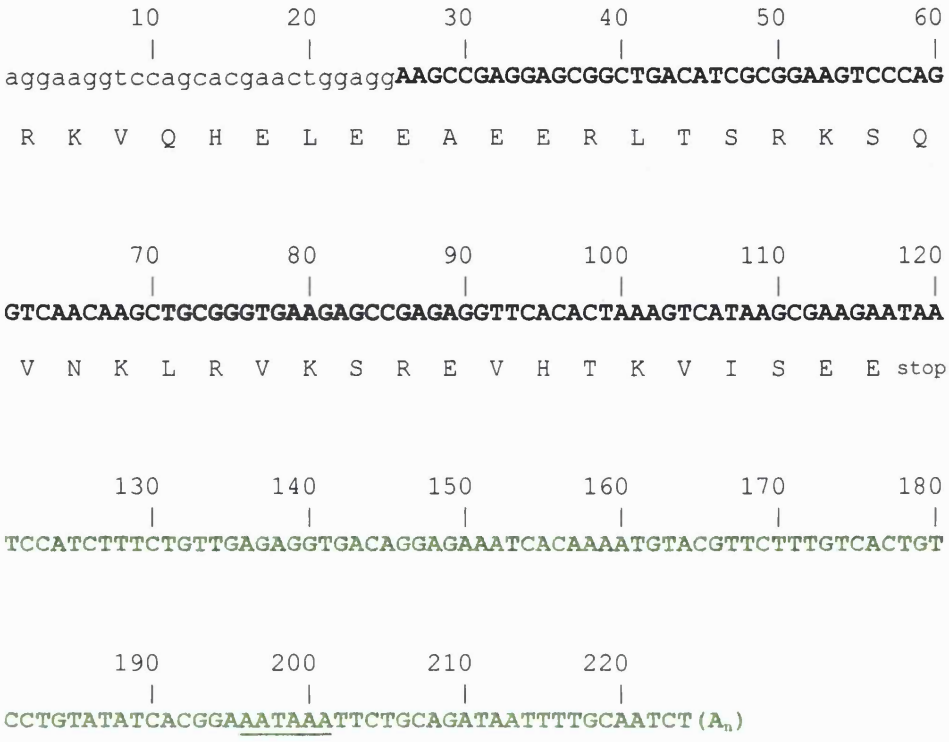


Figure 2.6. Nucleotide and amino acid sequence for 3' end of myosin heavy chain 2b. FG2Ex40 primer is indicated in lower case, and deduced amino acid sequence is represented by single letter code.

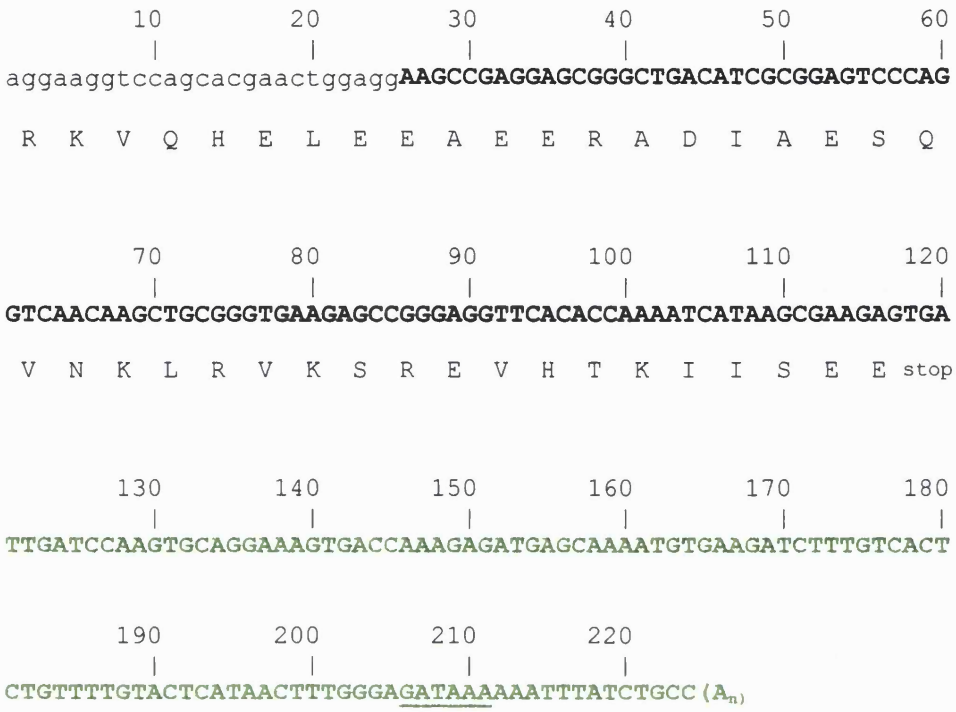


Figure 2.7. Nucleotide and dedudced amino acid sequence for myosin heavy chain 3' end identified as 2x, amplified using FG2Ex40 primer (lower case).

The CLUSTAL DNA programme was used to produce alignments of the amino acid and nucleotide sequence data obtained for the mouse MyHC genes.

type I	AEERADIAESQVNKLR**A**KSR**DIGA**K**GLN**EE
2a	**SRGAA**DIAESQVNKLRVKSREVHTKIISEE
2b	AEER**LTSRK**SQVNKLRVKSREVHTK**V**ISEE
2x	AEERADIAESQVNKLRVKSREVHTKIISEE

Figure. 2.8. Alignment of amino acid sequence for MyHC 3' coding region. The single letter code is used to represent the amino acids, bold letters indicate non-conserved amino acids.

Figure 2.8. demonstrates the conserved nature of the myosin heavy chains. It can be seen that the predicted amino acid sequence varies only slightly between the different proteins. The nucleotide sequences coding for myosin heavy chains are highly homologous across the four genes isolated here (figure 2.9.) The 3' untranslated regions, however are somewhat divergent. The percentage homology ranges from 55% to 71%. Noticeably, when the fast isoforms were aligned, two regions of homology stood out within the untranslated sequence (indicated in bold, figure 2.10).

When the mouse myosin heavy chain 3' untranslated sequences were compared with those of other species, it was evident that a particular isoform is more similar to its counterpart in another species than to other isoforms within that species as expected, confirming the identification of the MyHC types encoded by the different genes (see figures 2.11. and 2.12.). Table 2.1. shows the percentage homology between the different MyHC 3'UTRs within a species and across species. Percentage homologies in some cases as high as 99% for comparative isoforms (see figures in bold).

Mperi	53.6																					
Mslow	42.9	55.2																				
M2a	51.2	68.9	71.4																			
M2b	61	64.2	69.5	61.9																		
M2x	53.3	61.5	70.5	67.6	55.2																	
Remb	**90.2**	58.2	47.6	56.6	54.9	52.5																
Rperi	50.4	**88.4**	57.1	71	70.3	64.5	55.7															
Rslow	53.3	55.2	**81**	73.3	71.4	68.6	53.3	57.1														
R2a	54.1	63.1	55.2	**88.5**	68.6	65.7	56.6	66.4	55.2													
R2b	63.1	68.9	53.4	68.9	**89.3**	58.3	62.1	54.4	43.7	68.9												
R2x	54.4	66	45.6	68.9	61.2	**93.2**	57.3	64.1	42.7	70.9	56.3											
Rbemb	**65.6**	56.6	49.5	53.3	56.2	57.7	**63.9**	54.9	54.3	55.7	52.4	56.3										
Rbperi	52.5	**66.7**	51.4	60.8	70.5	64.4	50	**65**	51.4	65	67	59.2	47.5									
Rbslow	49.6	50.4	**73.3**	52.9	51.4	51	49.6	49.6	**68.6**	47.9	54.4	46.6	46.3	40.8								
Rb2a	55.2	57.5	55.2	**77.6**	66.7	71.2	58.2	58.2	50.5	**77.1**	68	73.8	54.9	65.8	44.6							
Rb2b	61.8	61.8	55.9	64.7	**80.4**	65.7	60.8	60.8	56.9	70.6	**73.5**	64.7	57.8	68.6	52	68.6						
Rb2x	57.7	64.4	53.8	70.2	60.6	**86.5**	60.6	66.3	46.2	72.1	61.2	**88.4**	55.8	64.4	43.3	75	67.6					
Hemb	**80**	57.6	42.9	48.8	66.7	58.1	**78.7**	52	55.2	57.4	65	66	**70.5**	56.7	47.1	62.4	63.7	64.4				
Hperi	56.8	**71.3**	49.5	61.6	65.7	59	51.6	**68.2**	48.6	60.7	63.1	60.2	50.8	**83.3**	43.8	63.6	69.6	68.3	58.4			
Hslow	51.3	46	**77.1**	63.7	68.1	69	45.1	46	**74.3**	53.1	53.4	49.5	45.1	44.2	**77.9**	54	56.9	48.1	50.4	50.4		
H2a	55.2	58.6	50.5	**76.8**	69.5	73.3	59	57.8	44.8	**73**	68.9	71.8	54.1	66.7	46.3	**85.2**	68.6	72.1	58.4	66.4	49.6	
H2x	53.6	55.4	52.4	76.8	70.5	**70.5**	59	56.9	46.7	69.7	68	**71.8**	55.7	68.3	47.9	85.4	68.6	**71.2**	57.6	66.7	50.4	99
	Memb	Mperi	Mslow	M2a	M2b	M2x	Remb	Rperi	Rslow	R2a	R2b	R2x	Rbemb	Rbperi	Rbslow	Rb2a	Rb2	Rb2x	Hemb	Hperi	Hbeta	H2a

Table 2.1. Percentage homology of 3' untranslated regions of skeletal myosin heavy chain genes from mouse (M), rat (R), rabbit (Rb), and human (H). embryonic (emb) and perinatal (peri) sequences for mouse, rat, rabbit and human, were obtained from EMBL under accession numbers M38129, M12289, X05004, K02111, Y13203, Y13204, X13100, and Y00821 respectively

```
type I   cgcaagtgc-agcacgagctggatgAGGCGGAGGAGAGGGCGGACATCGC    49
2a       aggaaggtccagcacgaactggAGGAAGCCGAGGAGCGG-CTGACATCGC    49
2x       aggaaggtccagcacgaactggAGGAAGCCGAGGAGCGGGCTGACATCGC    50
2b       aggaaggtccagcacgaactggAGGAAGCCGAGGAGCGG-CTGACATCGC    49
          * ***..* *******.*****.**.** ****** ** *.********

type I   CGA-GTCCCAGGTCAACAAGCTGCGGGCCAAGAGCCGGGACATTGGTGCC    98
2a       GGA-GTCCCAGGTCAACAAGCTGCGGGTGAAGAGCCGGGAGGTTCACACC    98
2x       GGA-GTCCCAGGTCAACAAGCTGCGGGTGAAGAGCCGGGAGGTTCACACC    99
2b       GGAAGTCCCAGGTCAACAAGCTGCGGGTGAAGAGCCGAGAGGTTCACACT    99
          ** ********************** ********.** .** . .*

type I   AAGGGCCTGAATGAGGAGTA--GCT---CTTGTGCTACCCAGC-TCCAAG   142
2a       AAAATCATAAGCGAAGAGTAAGGCTGTCCCGATGCTGTGGAATGACCGAA   148
2x       AAAATCATAAGCGAAGAGTGATT-GATCCAAGTGCAGGAAAGTGACCAAA   148
2b       AAAGTCATAAGCGAAGAATAAT-CCATCTTTCTGTTGAGAGGTGACAGGA   148
         **...* *.*. **.**.*.          . ** ..   ..  .* ...

type I   GG---TGCCC-----GTGAAGCCCTCAGACCTGGAGCCTTTG-----CAA   179
2a       GGCGAGGCACAAAATGTGAAACCTTTGGTCATGC-GCCTGTGTGATTCTA   197
2x       GA-GATGAGCAAAATGTGAAGATCTTTGTCA--C-TC-TGTTTTGTAC--   191
2b       GA--AATCACAAAATGT-ACGTTCTTTGTCACTG-TCCTGTATATCAC--   192
         *.   ..  *     ** * .   * .*.*     .* *.*.     *

type I   --CAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTTTCCTTA-AAGCC     224
2a       TTCCATCCTGTTGTAAGGAAATAAAGAGCCCAAGTTCCTTTGCAAGC-     244
2x       -TCATAACTTTGGGAGATAAAAAATTTATC----------TGC---C-     224
2b       ----------------GGAAATAAATTCTGCAGATAATTTTGCAATCT     224
                        ..*. ..*...            *.    *
```

Figure 2.9. Alignment of 3' coding sequences and untranslated regions of mouse MyHc type I, 2a, 2b, and 2x mRNAs. Character to show that a position in the alignment is perfectly conserved (*). Character to show that a position is well conserved: (.) Gaps indicated by (-) were inserted to maintain alignments. Lower case letters refer to primer sequence. Stop codons are shown in bold and the poly A signals are underlined.

```
2a      aggaaggtccagcacgaactggaggAAGCCGAGGAGCGG-CTGACATCGC    49
2x      aggaaggtccagcacgaactggAGGAAGCCGAGGAGCGGGCTGACATCGC    50
2b      aggaaggtccagcacgaactggAGGAAGCCGAGGAGCGG-CTGACATCGC    49
        ************************************ **********

2a      GGA-GTCCCAGGTCAACAAGCTGCGGGTGAAGAGCCGGGAGGTTCACACC    98
2x      GGA-GTCCCAGGTCAACAAGCTGCGGGTGAAGAGCCGGGAGGTTCACACC    99
2b      GGAAGTCCCAGGTCAACAAGCTGCGGGTGAAGAGCCGAGAGGTTCACACT    99
        *** ********************************.***********

2a      AAAATCATAAGCGAAGAGTAAGGCTGTCCCGATGCTGTGGAATGACCGAA   148
2x      AAAATCATAAGCGAAGAGTGATT-GATCCAAGTGCAGGAAAGTGACCAAA   148
2b      AAAGTCATAAGCGAAGAATAAT-CCATCTTTCTGTTGAGAGGTGACAGGA   148
        ***.*************.*.*.   .**  . ** .*.....**** ..*

2a      GGCGAGGCACAAAATGTGAAACCTTTGGTCATGCGCCTGTGTGATTCTAT   198
2x      GA-GATGAGCAAAATGTGAAGATCTTTGTCA--CTC-TGTTTTGTAC---   191
2b      GA--AATCACAAAATGT-ACGTTCTTTGTCACTGTCCTGTATATCAC---   192
        *.  *.. .******** * .   **.****   .* ***.*.. .*

2a      TCCATCCTGTTGTAAGGAAATAAAGAGCCCAAGTTCCTTTGCAAGC-    244
2x      TCATAACTTTGGGAGATAAAAAATTTATC----------TGC---C-    224
2b      ---------------GGAAATAAATTCTGCAGATAATTTTGCAATCT    224
                     ..***.**...             ***   *
```

Figure 2.10. Alignment of 3' end of the fast isoforms. Primer sequence is in lower case. Stop codons indicated in bold case. Poly A signal sequences are underlined. The highly conserved nature of the coding region is very evident. Two regions of homology in the untranslated region are highlighted in bold.

Type I MyHc

```
Mouse      CT---------CTTGTGCTACCCAGCTCCAAGGGTG---CCCGTGAAGCC    38
Rat        AT---------CTTGCTCTACCCAACCCTAAGGATG---CCTGTGAAGCC    38
Rabbit     CCTGGTGACGCCTTGATCCGCCCAGCCCTGAGGACGACGCCAGTGAAGTC    50
Human      CTTTGCCACATCTTGATCTGCTCAGCCCTGGAGGTG---CCAGCAAAGCC    47
                     **** .* .* **.* * ...*. *   ** * .*** *

Mouse      CTCAGACCTGGAGC---CTTTGCAACAGC-CCTTTAGGTGGAAGCAGAAT    84
Rat        CTGAGACCTGGAGC---CTTTG-AAAAGCACCTTCAGGCAGAAACACAAT    84
Rabbit     CCTTGTCTGGGAGCTCACATAGCAGCAGCCCCTT-GGGAAGAAGCAGAAT    99
Human      CCATG--CTGGAGC---CTGTGTAACAGCTCCTT-GGGAGGAAGCAGAAT    91
           *  .*   .*****   *...* *. *** **** .** .***.** ***

Mouse      AAAGCAATTTTCCTTAAAGCC-                                105
Rat        AAAGCAATTTTCCTTCAAGCC-                                105
Rabbit     AAATCAGTTTTCCTCGAAGCTG                                121
Human      AAAGCAATTTTCCTTGAAGCCG                                113
           ***.**.*******  ****
```

2a MyHc

```
Mouse      GGCT-GTCCCGATGCTGTGGAATGACCGAAGGCGAGGCACAAAATGTGAA    49
Rat        GGCA-GCTCTGATGCTGTAGAATGACCGAAGA-AAGGCACAAAATGTGAA    48
Rabbit     GCCAAGTCCTGAGGCTGTGGAATGTCCAAAGAGA-GGCACAAAATGTGAA    49
Human      TC-ATGTCCTGATGCCATGGAATGACTGAAGACA-GGCACAAAATGTGAC    48
           .  . *  * **.** .*.*****.* .***. . **************

Mouse      ACCTTTGGTCATGCGCCTGTGTGATTC-----TATTCCATCCTGTTG-TA    93
Rat        GCCTTTGGTCATGCCCCCATGTGATTC-----TATTTAATCCTATTG-TA    92
Rabbit     ATCTTTGGTTATTTCCCTCTGTAATTACTGTAAAT-CTACCCTACTGCAA    98
Human      ATCTTTGGTCATTTCCCTCTGTAATTATTGTGTATTCTACCCTGTTGCAA    98
. ******* **.  **  ***.***      .**   * ***. ** .*

Mouse      AGGAAATAAAGAGCCCAAGTTCCTTTGCAAGC----                 125
Rat        AGGAAATAAAGAGCCCAAGTTC--TTGCAAGC----                 122
Rabbit     AGGAAATAAAGAGCATAGGGTGTTTTGCAAACAATC                 134
Human      AGGAAATAAAG--CATAGGGTAGTTTGCAAAC----                 128
           ***********  *  *.*.*   ******.*
```

Figure 2.11. Nucleotide sequence alignments of the untranslated regions of mouse, rat, rabbit and human slow type I and fast 2a myosin heavy chain genes. The mouse type I and 2a 3' coding and UTR sequences have been submitted to the EMBL database and are available under the accession numbers AJ223362 and AJ002521 respectively. Sequences for rat , rabbit and human for type I and 2a were obtained under the accession numbers K01463, Y13205, M30605, X72589, Y13201, and Z32858 respectively. The poly A signals are underlined. Sequence motifs as highlighted in figures 2.8. and 2.9. are indicated in bold for the 2a alignment.

2b MyHc

```
Mouse      TCCA--TCTTTCTGTTGAGAGGTGACAGGAGAAAT-CACA-AAATGTGAC    46
Publish.   TCCA--TCTTTCTGTTGAGAGGTGACAGGAGAAAT-CACA-AAATGTGAC    46
Rat        CTCAATTCCTTCTGTTGAAAGGTGACAGAAGAAAT-CACACAAATGTGAC    49
Rabbit     TCCA-TTCTAT-TGCTAAAAGGTGACCAAAGAAATGCACA-AAATGTGAA    47
             **  ** .* ** *.*.******* ..****** **** ********

Mouse      GTTCTTTGTCACTGTCCTGTATATCACGGAAATAAATTCTGCAGATAATT    96
Publish.   GTTCTTTGTCACTGTCCTGTATATCACGGAAATAAATTCTGCAGATAATT    96
Rat        GTTCTTTGTCACTGTCCTGTATATCAAGGAAATAAAAGCTGCAGATAATT    99
Rabbit     GTTCTTTGTCACTACCTTGTGCATCAAGTAAATAAAATCTGCA----ATT    93
           *************. * ***. **** *.*******..*****    ***

Mouse      TTGCAATCT                                            105
Publish.   TTGCAATCT                                            105
Rat        TTGC-----                                            103
Rabbit     TTGCAATCT                                            102
           ****
```

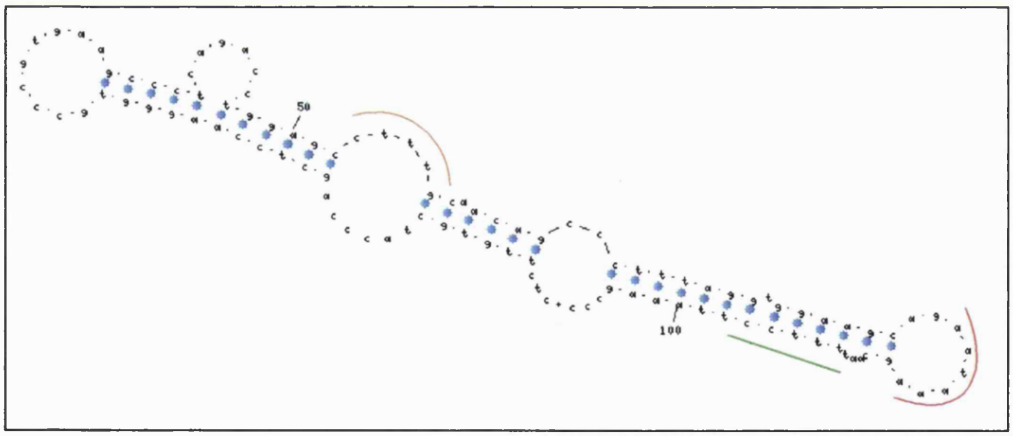
2x MyHc

```
Mouse      TTGA-TCCAAGTGCAGGAAAGTGACCAAAGAGATGAGCAAAATGTGAAGA    49
Rat        TCGA-TCCAAA-GCAGGAAAGTGACCAAAGAGATGAGCAAAATGTGAAGA    48
Rabbit     TCCA-TCCACATGCTGGAAAGTGACCAAAGAAATGCACAAAATGTGAAAC    49
Human      TCATGTCCTGATGCCATGGAATGACTGAAGACAGGCACAAAATGTGACA-    49
           *  . ***. . ** ....*.**** .**** *.* .********** .

Mouse      TCTTTG------TCACTCTGTTTTG-----------TACTCA-TAACTTT    81
Rat        TCTTTG------TCACTCCATTTTG-----------TACTTA-CGACTTT    80
Rabbit     TCTTTG-------CACTCTGTTTTG-----------TACTTATCGATTTT    81
Human      TCTTTGGTCATTTCCCTCTGTAATTATTGTGTATTCTACCCTGTTGCAAA    99
           ******       * *** .*..*.           ***  .  .. ...

Mouse      GGGAG------ATAAAAAATTTATCTGCC--                      104
Rat        GGGAG------ATAAAAAATTTATCTGCC--                      103
Rabbit     GGGAG------ATAAAAAATTTATCTGCC--                      104
Human      GGAAATAAAGCATAGGGTAGTTTGCAAACAA                      130
           **.*.      ***....*.**..*.. *
```

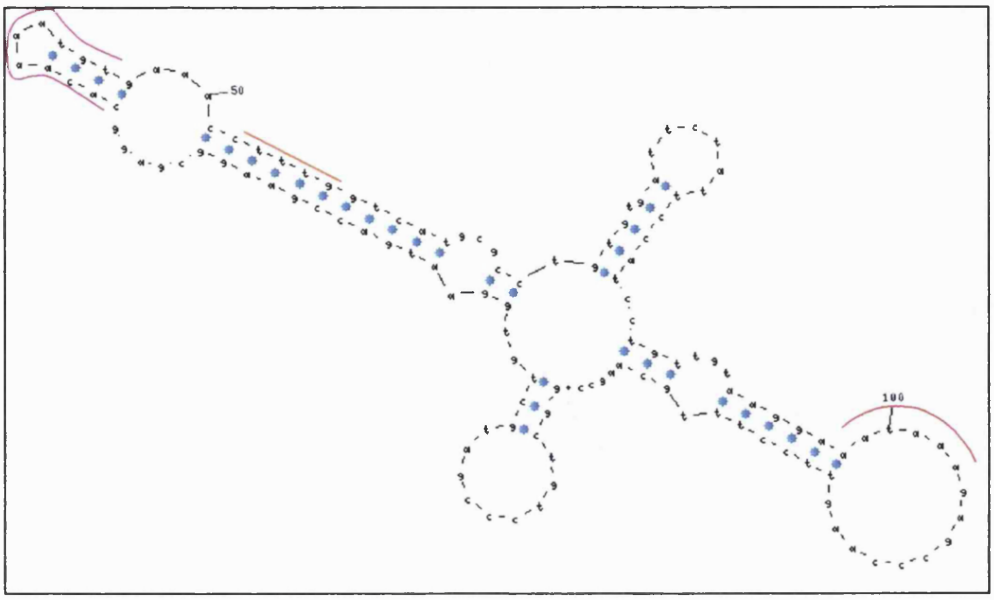
Figure 2.12. Alignment of 3' untranslated region nucleotide sequences for mouse, rat rabbit, and human myosin heavy chains. The mouse 3' coding and UTR sequences for 2b and 2x have been submitted to the EMBL database under the accession numbers AJ223361 and AJ002522 respectively. Rat, rabbit, and the previously described mouse 2b (publish.) sequences were obtained under accession numbers X72590, Y13200, and K00988 respectively. The rat, rabbit and human sequences were obtained under accession numbers X72591, Y13202, and S73840 respectively. The poly A signals are underlined and conserved sequence elements highlighted in bold.

2.3.3. Secondary structure predictions.

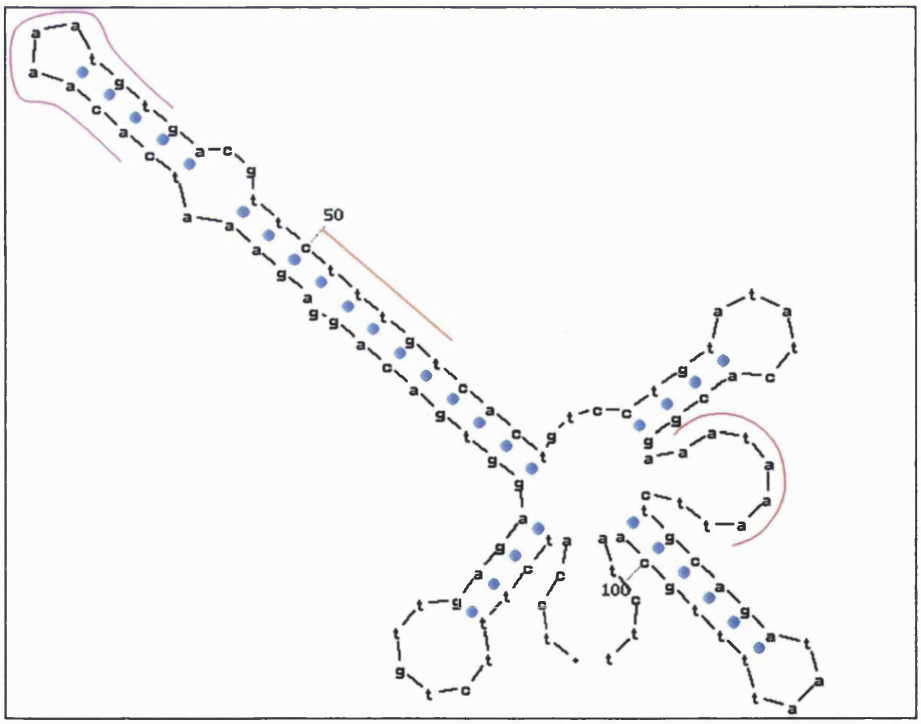
RNA molecules have various structures, which allow diverse mechanisms for protein recognition. RNA-binding proteins generally target single-stranded regions within secondary structure domains of the RNA. The secondary structures of the mouse MyHC 3'UTRs were predicted using the *mfold* programme (Zuker, 1989; Jaeger *et al*, 1990). Secondary structures are predicted by attempting to minimize the free energy of the RNA molecule by maximizing the number of favourable base pairing interactions. The free energy associated with base pairing contributes mainly to the stability of RNA secondary structure. This free energy can be well approximated from stacking energies that depend on neighbouring base pairs as well as a single base pair. The secondary structures of the type 1, 2a, 2b, and 2x 3'UTRs are shown below. The positions of the conserved CACAAAATGT (purple line) and CTTTG motifs (orange line) are indicated. The green line in the type 1 structure shows a motif conserved across species within this isoform. The position of the poly A signal is also highlighted by a red line.



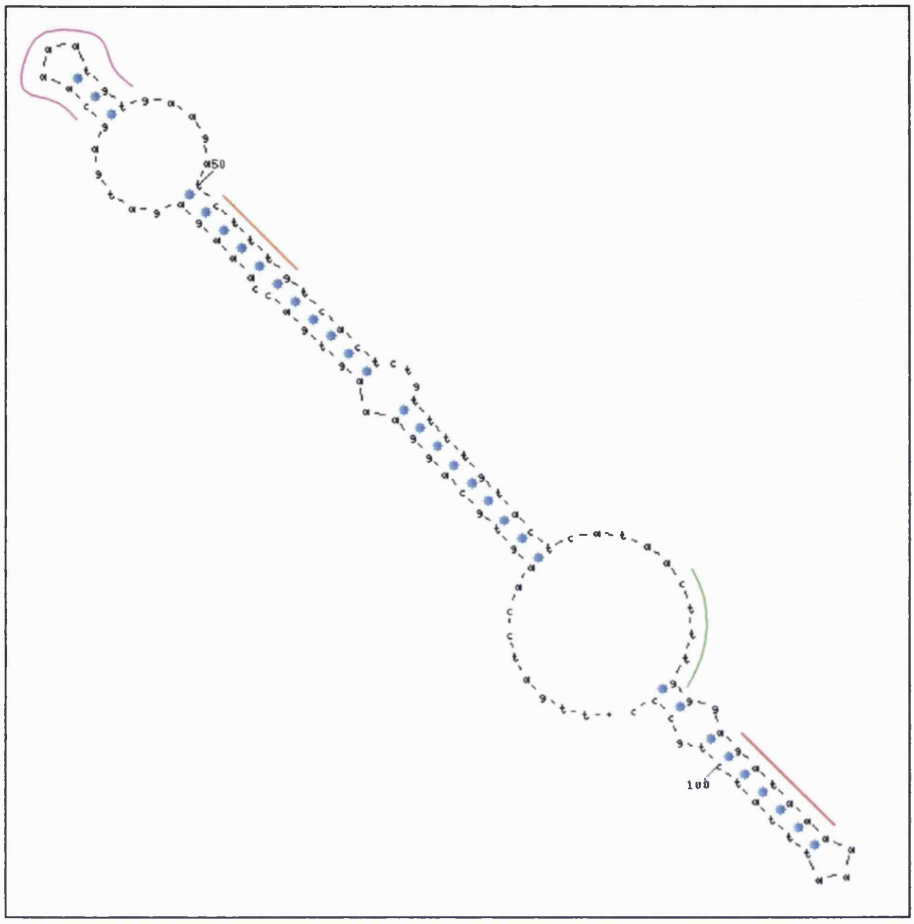
MyHC type 1 3'UTR . Free energy = -23.3 kcal/mol.



MyHC 2a 3'UTR. Free energy = -31.7 kcal/mol. .



MyHC 2b 3'UTR. Free energy = -27.5 kcal/mol.



MyHC 2x 3'UTR. Free energy = -26 kcal/mol.

2.4. Discussion.

The aim of the experimental work described in this chapter was to isolate the 3' untranslated regions of the mouse skeletal myosin heavy chains. This was achieved through employment of reverse transcription coupled with PCR, in a technique known as RACE. The technique allowed the amplification of the closely related MyHC genes by utilizing a specifically designed oligo-dT primer, which reduced the frequency of false amplification products from contaminating DNA during PCR, and the gene specific primers FG2Exn40 and RnExn40. These primers correspond to a highly conserved region within exon 40 of the MyHC gene coding sequence of fish white and rabbit alpha cardiac muscle respectively (Ennion *et al*, 1995 McKoy *et al*, 1998). Single bands were amplified in RACE PCR using FG2Exn40 primer. The RnExn40 primer showed some non-specific amplification since there was another prominent band from the PCR that proved to be apolipoprotein A-I

By comparison with previously described MyHc genes from various species, four myosin heavy chain genes were identified as having been cloned; slow type I, fast 2a, fast 2b, and fast 2x. These are the genes expressed in adult skeletal muscle. The FG2Ex40 primer amplified only the fast MyHC genes and therefore RnEx40, which was designed specifically from slow type I MyHC sequences, was used to amplify this isoform from the different muscle samples. The MyHc amplified from a particular muscle is dependent on the predominant gene(s) expressed in that muscle. From figure 2.3. it can be seen that the diaphragm and soleus express both slow and fast MyHCs, whereas only fast isoforms could be amplified from the gastrocnemius and tibialis anterior, reflecting the MyHC composition of these muscles in mice.

The slow type I MyHC 3' cloned form skeletal muscle sequence is identical to the beta cardiac isoform expressed in murine heart (Sanchez *et al*, 1991). The 2a and 2x sequences amplified from muscle mRNA are similar to two linked mouse myosin heavy chain genes, MHC2A and MdMs, isolated from mouse genomic library by Parker-Thornburg *et al*, 1992. Although MHC2A was identified as MyHC 2a, the identity of MdMs was only postulated to correspond to MyHC 2x. The 2a and 2x

sequences amplified by RACE PCR here are virtually identical to those published by Parker-Thornburg *et al*, 1992, with some minor sequence differences that may be due to allelic variations between strains or minor sequence errors (see appendix). The 2b sequence is almost identical to that described by Weydert *et al*, 1983, whose 3' sequence was incomplete (see appendix). It is clear from the 3' sequences derived in this study and those from other species that for a given isoform, there is less variation across the species than there is for different MyHCs within the same species. The presence of conserved regions suggests that the 3'end of each type of MyHC message serves a specific function.

The MyHC 2x transcript contains an atypical poly A signal, GAUAAA. The poly A signal is important for 3' end formation, where primary transcripts are cleaved at a specific point and a tail of adenosine residues are added to the free 3' end (Wahle & Keller, 1992; Keller, 1995). The point of cleavage is determined not only by the poly A signal, which is situated approximately 10-30 bases upstream of the cleavage site (Proudfoot & Brownlee, 1976), but also by downstream GU and or U rich sequences (McLauchlan *et al*, 1985). Mutation analysis has shown that base changes at every position in typical AAUAAA greatly reduce cleavage and polyadenylation *in vitro* (Sheets *et al*, 1990). This reduction varies greatly according to the position of the changed base, with some mutations being tolerated relatively well. Although variants of AAUAAA do occur naturally, they are not very common. The variant AUUAAA is found in 10% of poly A sites, and the two less common variants AAUACA and AAUUAA are found in approximately 2% of poly A sites. The last two variations have been shown to result in less efficient processing than the normal AAUAAA sequence (Wickens & Stephenson, 1984). The poly A signal for the mouse 2x mRNA is assumed to be the sequence GAUAAA since it is situated 15 bases upstream of the polyadenylation site, the normally expected position for the poly A signal. This poly A signal is not unique to the mouse MyHC 2x sequence, but is also observed in rat and rabbit MyHC 2x 3'UTR. The human MyHC 2x transcript however contains a typical poly A signal (figure 2.12.). A mutation of AAUAAA to GAUAAA in *in vitro* processing assays has been shown to result in a low efficiency of cleavage and even lower efficiency of polyadenylation (Sheets *et al*, 1990). It is

possible that the 2x transcript could be similarly inefficiently processed, although other sequences in the primary transcript could compensate for the variant poly A signal to allow 3' end formation and subsequent expression.

It is well established that the coding regions of the myosin heavy chain genes are highly conserved (Saez & Leinwand, 1986; Stedman et al, 1990; Moore *et al*, 1993). Results obtained from sequencing of the mouse myosin heavy chains demonstrated this high homology. In contrast, the 3' untranslated regions of the MyHC transcripts are rather different between the various isoforms. These regions can be used therefore, not only to identify a particular isoform, but also to generate specific probes for studying expression and distribution patterns (Weydert *et al*, 1983; Parker-Thornburg *et al*, 1992; McKoy *et al*, 1998). However, it is noteworthy that the 3' UTRs of the same myosin heavy chain gene across species show a degree of homology. Table 2.1. shows the degree of homology between the skeletal myosin heavy chain gene 3' untranslated regions in mice, rats, rabbits and humans. The homology of each isoform to all the others is indicated as a percentage. Since different myosin heavy chain genes are expressed in different tissues, perhaps this conservation could be due to a functional role of 3' untranslated regions in tissue specific expression. Two short blocks of sequence appear to be well conserved in the mouse fast isoforms, CACAAAATGTGA and CTTTGTCA. These motifs appear to be common to all fast isoforms in all species analysed here, occurring at approximately in the same position in the 3' untranslated region. The CACAAAATGTGA motif is not found in the type 1 sequence from any of the species analysed here. This could be explained by the fact that the slow type 1 MyHC is encoded on a separate chromosome, whereas the fast MyHCs are tandemly linked on one chromosome and probably arose form gene duplication events. The consensus sequence for the two motifs can be read as CAAAATGTGA and CTTTG. The CTTTG motif is present more than once within the 2x 3' UTR, and is also observed in the mouse and rat and human, but not rabbit type 1 3'UTR. This position of this motif within the human, mouse and rat type 1 differs, beginning at base 1 in the human type 1 3'UTR and at base 53 in the mouse and rat type 1 3'UTRs. The main conserved sequence seen within the type 1 3'UTR in different species is the

motif TTTTCCTT. A database search of the CAAAATGTGA motif shows that this sequence is not unique to the myosin heavy chains. It is also present in the 3' UTR of a Xenopus laevis LIM class homeodomain protein gene (Taira *et al*, 1992) and heat shock transcription factor 33 mRNA from soybean (90% homologous). Furthermore, the LIM homeodomain 3' UTR also contains a CTTTG motif.

The *mfold* programme was used to predict the secondary structures of the MyHC 3'UTRs. Although only one structure is shown for each type of MyHC 3'UTR, the programme produced several foldings for each 3'UTR, with similar free energies of folding. The secondary structures shown are those with the most negative free energy of folding since the more negative the free energy, the more stable the secondary structure. Prediction of the secondary structures for the mouse MyHC 3'UTRs revealed stem-loop structures for all four 3'UTRs with internal loops and bulges which could be targets for protein binding. RNAs have various structures that allow diverse mechanisms for protein recognition. Unlike DNA binding proteins, where sequence specific recognition occurs by insertion into the major groove of the DNA helix, RNA binding proteins are thought to not usually target fully double stranded tracts (Mattaj, 1993). RNA binding proteins generally target single stranded regions within secondary structure domains. These are hairpin loops (stem-loops), internal loops and bulges where functional groups on the bases may be more readily available for sequence specific recognition. Helical stems are formed by Watson-Crick base pairing. These can be interrupted by bulges where one or more nucleotides on one strand of the stem structure are not base-paired, while all the nucleotides on the opposite strand are base-paired. The unpaired bases can be stacked within the helix, or bulged outside of the helix and represent targets for site-specific recognition of RNA molecules. In addition, they may also form tertiary contacts between distant regions of the RNA secondary structure since bulges bend RNA helices significantly, depending on the number and identity of bulged nucleotides. Internal loops are formed when two double-stranded helices are separated on each strand by several non-Watson-Crick-paired nucleotides. Stem-loop structures occur when the phosphodiester backbone folds back on itself to form a double helical tract (stem) leaving unpaired nucleotides (loop). This is the dominant feature of the 2a and

2x 3'UTRs, present also within the type 1 and 2b 3'UTR secondary structures and is a potential binding site for RNA-binding proteins. The conserved motif CACAAAATGT is positioned within such a stem loop structure and could be a sequence specific recognition site for protein binding. The CTTTG motif on the other hand, resides within the stem structure of the slow type 1, 2a, 2b and 2x 3'UTRs. There is a second CTTTG motif present in the 2x 3'UTR that forms part of a loop. The type 1 conserved motif TTTTCCTT forms a stem structure. These 'stem' motifs could play a part in positioning the protein correctly through non-sequence-specific contacts with the RNA backbone. The position of the poly A signal in the predicted secondary structures is within a loop for type 1, 2a and 2b 3'UTRs, but for the 2x which has the unusual GAUAAA poly A signal, this motif is situated in a stem region. Upstream of this stem is an internal loop formed partly by the bases CATAAC. These bases could possibly behave like a poly A signal to which the cleavage and polyadenylation specificity factor binds.

Although the secondary structures can be computed from an RNA sequence and theoretical free energy minimization under specified conditions, the computed structures may not fully mimic *in vivo* situations, and the RNAs themselves may require helper molecules in order to fold into biologically active structures. Sequence comparisons to similar molecules across phylogeny and chemical and enzymatic probing are necessary to confirm secondary structure predictions. Comparisons of a particular sequence across species highlighting conserved secondary structures would indicate *in vivo* function, and chemical and enzyme analyses would allow stem-loop structures and helical stem structures to be mapped. The secondary structures of several 3'UTRs which contain motifs for post-transcriptional regulation have been elucidated and are available for comparison on the website http://bio-www.ba.cnr.it:8000/BioWWW. The MyHC 3'UTRs have no structures or motifs common with sequences in this database which include histone, IRE, CPE, AUREs *nanos* and other 3'UTRs.

Over recent years, the 3' UTRs of various genes have been assigned various functional roles. These include localization of mRNAs (Singer, 1992; St. Johnston, 1995; Hesketh, 1996), translational regulation (Curtis *et al*, 1995; Jackson &

Standart, 1990) and mRNA stability (Ross, 1995; Jacobson and Peltz, 1996). Often these functions act in conjunction to regulate the expression of genes. In all cases sequence elements in 3' UTR have been identified that are responsible for such functionality. For example. the MEF2A 3'UTR functions as a *cis*-acting translation repressor (Black *et al*, 1997). Kislauskis *et al*, 1994, showed that a 'zipcode' in the β-cytoplasmic actin 3'UTR was responsible for the intracellular localization of chimeric mRNA. An example of the role of the 3'UTR in mRNA stability has been demonstrated by Kiledjian *et al*, 1995 in globin mRNA. AU-rich elements have been found to be common to mRNAs that are unstable (Caput *et al*, 1986; Shaw and Kamen, 1986; Han *et al*, 1990). No such motifs are observed in the MyHc 3' UTRs. Studies by Medford *et al*, 1983 have demonstrated an increase in the effective stability of MyHC mRNA during myogenesis in experiments with a rat skeletal muscle cell line in response to withdrawal of myoblasts from the cell cycle prior to fusion. Observations by others have suggested that changes in stability and/or the rate of translation account for accumulation of MyHC mRNA without expression of the corresponding protein during development (Lyons *et al*, 1990; Vivarelli *et al*, 1988; Roy *et al*, 1984). Hence the MyHC 3'UTRs could influence stability of the message during muscle cell differentiation as it is the case for 3'UTRs of several other transcripts (see introduction). Furthermore, it has been shown that many 3' UTRs bind cytoplasmic proteins in electrophoretic mobility shift assays, with the secondary structure of the 3'UTR playing an important role in the binding of proteins. The *mfold* programme was used to predict the secondary structures of the MyHC 3'UTR mRNA sequences. From the predicted structures it can be seen that the MyHC 3'UTRs form similar secondary structures. The loop and bulge structures may allow for specific recognition since they expose the RNA backbone to interaction with proteins. Similar secondary structures are predicted for the MyHC 3'UTRs from other species (not shown). The fact that the position of the sequence motif CACAAAATGT element is conserved could be of significance for secondary structure formation, and subsequent protein binding.

Chapter 3

The Effect of the Myosin Heavy Chain 3'UTRs on Reporter Gene Expression and mRNA Stability

3.1. Introduction.

The study of eukaryotic gene regulation and expression has been greatly advanced by the ability to introduce DNA or RNA into eukaryotic cells. The expression of exogenous genes in mammalian cells can be used to analyse transcriptional regulatory sequences, the mechanisms involved in RNA processing, intracellular sorting and post-translational processing of secreted proteins. Physical and chemical methods can be used to introduce nucleic acids into cells in a process that is referred to as transfection. The DNA of interest is usually introduced into cells through vectors. Many different types of vectors have been used to introduce DNA in to cultured mammalian cells. By far the most common method is the use of plasmid DNA expression vectors. These vectors contain sequences that include a set of restriction sites to facilitate insertion of foreign DNA and a gene coding for antibiotic resistance allowing selection of plasmids containing bacterial colonies. The transcriptional unit within the vector comprises of a promoter element (such as the SV40 early promoter, Rous sarcoma virus [RSV], or the human cytomegalovirus [CMV] promoter) to drive expression of inserted genes, enhancer elements to increase efficiency of expression, and polyadenylation signals.

Initial transfection studies by Vaheri & Pagano (1965), and Graham & van der Eb (1973), used DEAE-dextran and calcium phosphate-mediated transfection techniques. Both of these methods produce a chemical environment that results in the DNA attaching to the cell surface, and the DNA is then endocytosed. Since the early experiments with DNA transfer, transfection technology has advanced greatly with the advent of molecular biology. Cloning of plasmid DNA and ease of DNA manipulation has allowed the development of additional, more efficient, methods of transfection, namely electroporation (Wong & Neumann, 1982), direct microinjection (Capecchi, 1980), biolistic particle delivery (Ye *et al*, 1990), and lipid-mediated transfer (Felgner *et al*, 1987) of nucleic acids into cells. The first three of these are physical methods. Electroporation uses an electric field to open up pores in the cell, and the DNA diffuses through the pores (Shigekawa & Dower, 1988; Shigekawa & Dower, 1989). Although this technique can be used with virtually all

cell types, a drawback of this protocol is the substantial cell death, thus requiring large numbers of cells. The microinjection technique involves using a fine needle to deliver the DNA directly into the nucleus or cytoplasm of the cell. Since this method is labour-intensive, it is not appropriate for studies requiring large numbers of cells. The biolistic particle method of gene delivery relies on high velocity delivery of nucleic acids on microprojectiles to recipient cells. Although this method has a high efficiency of transfection in culture and *in vivo*, its use is prohibited by cost.

Lipofection or lipid–mediated transfer is a chemical method that uses synthetic cationic and neutral lipids to transfect cells. The cationic portion of the lipid molecule associates with the negatively charged nucleic acid, resulting in the formation of a lipid/nucleic acid complex. The overall positive charge of the complex mediates its association with the negatively charged cell membrane and uptake of the DNA may occur either by endocytosis (Friend *et al*, 1996) or by fusion with the plasma membrane (Bangham, 1992). This method of transfection is highly efficient, and has a greater reproducibility. However, the lipid formulation can be toxic to cells at high concentrations. Therefore, optimization of the amount of lipid and nucleic acid necessary for efficient transfection is required depending on the type of cells to be transfected and the type of lipid formulation that is used.

When cells are transfected, although the DNA will enter a large proportion of the cells in culture, only a small number of these cells will actually take up the DNA into the nucleus. Such DNA is expressed transiently for a few days. In an even smaller proportion of cells, the DNA is integrated into a chromosome and may be stably expressed in future generations of this cell population. These stable transfectants can be identified and maintained by using a selectable marker either in a recombinant plasmid or on a separate plasmid that is co-transfected with the plasmid DNA that is being studied. The selectable marker usually encodes resistance to a lethal drug. Cells that survive the drug treatment expand into clonal groups that can be individually selected and propogated. One drawback of stable transfection is that it can take several weeks and often longer to get through the process of drug-selection, isolation and propagation of stably transfect clonal cell lines.

The most commonly used expression vectors used for studying regulation of gene expression are those which contain reporter genes coding for enzymes such as chloramphenicol acetyl transferase, beta-galactosidase, firefly luciferase and alkaline phosphatase. Putative regulatory sequences or their modified derivatives are joined to the reporter gene, which is not expressed or is present at very low levels in the cells of interest. Following transfection cells can be assayed for reporter mRNA and/or reporter protein. This method has led to the identification of promoter, enhancer and silencer elements in many eukaryotic genes, and is useful for studying mRNA processing, translation and recombination events (Groskreutz & Schenborn, 1997).

The reporter system was applied in this chapter to study the 3'UTR of the MyHC genes. The 3'UTR of some genes have been shown to be involved in translation control (reviewed by Jackson & Standart, 1990). One mechanism of translational control is thought to occur through localization and cytoplasmic polyadenylation of transcripts. This is observed in many developmentally regulated mRNAs (Huarte *et al*, 1992; Gavis & Lehman, 1994). Specific sequences within the 3'UTR of other mRNAs have been shown to repress translation. For example, the 3'UTR of human p53 mRNA has been shown to contain a negative regulatory element, which is capable of repressing its translation (Fu *et al*, 1996). In *C. elegans*, the *tra*-2 sex-determining gene is regulated at the translational level by two direct repeat elements located within its 3'UTR which are required for repression of translation (Goodwin *et al*, 1993). In some mRNAs, the 3'UTR has been shown to upregulate transcription in response to physiological and biochemical factors (Moseley *et al*, 1993; Fu & Benchimol, 1997).

Cytoplasmic MyHC mRNA has been shown to accumulate in myoblasts, and then levels decrease prior to myosin synthesis and cell fusion (Robbins & Heywood, 1978; Dym *et al*, 1979). The accumulation of MyHC mRNA without expression of the corresponding protein until later has been attributed to changes in stability or translation of the mRNA (Medford *et al*, 1983; Roy *et al*, 1984). In chapter 2, sequence analyses of the 3'UTRs of the skeletal MyHC genes revealed the isoform specific conservation of this region, suggesting a functional role. It is possible that the 3'UTRs of the MyHC genes play a role in translation control. In order to

investigate this aspect of the MyHC 3'UTRs, hybrid reporter expression vectors were constructed. The CAT reporter system was used for the analyses. The MyHC 3'UTRs were spliced on to the end of the CAT gene coding region and the hybrid constructs were transfected into myoblasts by lipofection. The effects of MyHC 3'UTRs were determined by assaying cells lysates for CAT activity.

3.2. Materials and methods.

3.2.1. Amplification of Myosin heavy chain 3'UTRs.

From the sequences for the mouse myosin heavy chain 3' coding and untranslated regions obtained in chapter 2, primers were designed for use in PCR to amplify only the 3' untranslated region, inclusive of poly A signals but not the stop codon, of slow type 1/β, fast 2a, 2b, and 2x myosin heavy chain mRNAs. Each pair of primers were designed such that the forward primer contained a *Hind* III restriction site and the reverse primer contained an *Xba* I site, to facilitate subcloning. The forward (F) and reverse (R) primers used for each myosin 3' UTR are shown below.

BETAF	**5'**	**CCGAAGCTTCTCTTGTGCTA**	**3'**
BETAR	**5'**	**CGGTCTAGAGGCTTTAAGGA**	**3'**
2AF	**5'**	**AATAAGCTTGGCTGTCCCGATGC**	**3'**
2AR	**5'**	**GCGTCTAGAGCTTGCAAAGGAAC**	**3'**
2BF	**5'**	**CCGTGAAGCTTTCCATCTTTCTG**	**3'**
2BR	**5'**	**CGGCCGTCTAGAAGATTGCAAAA**	**3'**
2XF	**5'**	**GGCCAAGCTTTTGATCCAAGTG**	**3'**
2XR	**5'**	**CGCGCTCTAGAGGCAGATAAAT**	**3'**

The 3'ends amplified in chapter 2 were used as templates for PCR. The pCRII plasmids containing the cloned MyHC 3'ends were digested with *EcoR* I and electrophoresed on an agarose gel. Bands corresponding to the MyHC 3' ends were gel purified and 5 μl used in a PCR reaction as described in chapter 2 section 2.2.5. with the appropriate pairs of primers. PCR products were analysed by agarose gel electrophoresis and correct size bands were gel purified. The PCR products were

cloned into pCR™II vector as described in chapter 2, section 2.2.6 and sequenced to verify correct amplification.

3.2.2. Plasmid constructs.

The reporter gene concept using the bacterial chloramphenicol acetyltransferase (CAT) gene and CAT assay system was developed by Gorman *et al*, (1982). The control plasmid pSVLCAT was constructed by inserting the gene coding for chloramphenicol acetyltransferase from the plasmid pCM7 (Pharmacia) into the vector pSVL (Pharmacia). pSVL is designed for high-level transient expression in eukaryotic cells, and contains a multiple cloning site into which genes can be inserted. The expression of these genes is driven by the Simian virus 40 (SV40) late promoter (Gorman *et al,* 1982; Sprague *et al*, 1983; Templeton & Eckhart, 1984). The plasmid also contains processing signals for splicing and polyadenylation of transcripts (see figure 3.1) (Crowley *et al*, 1983).

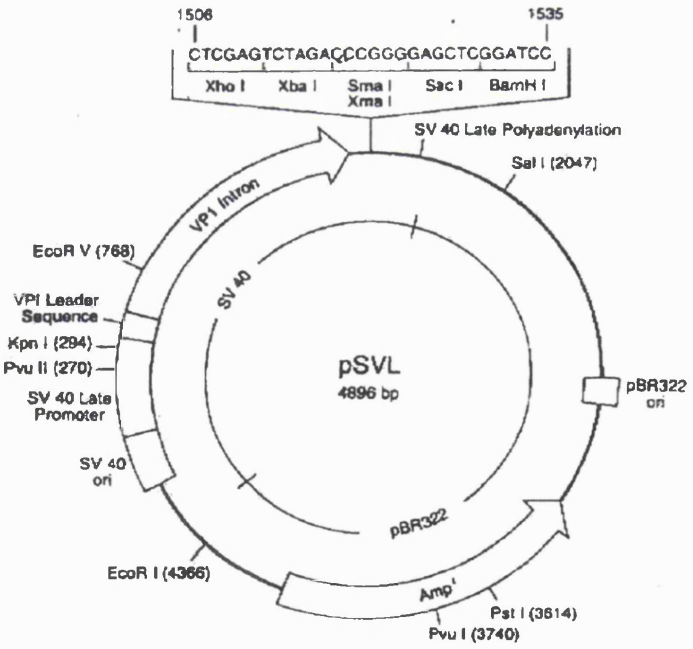
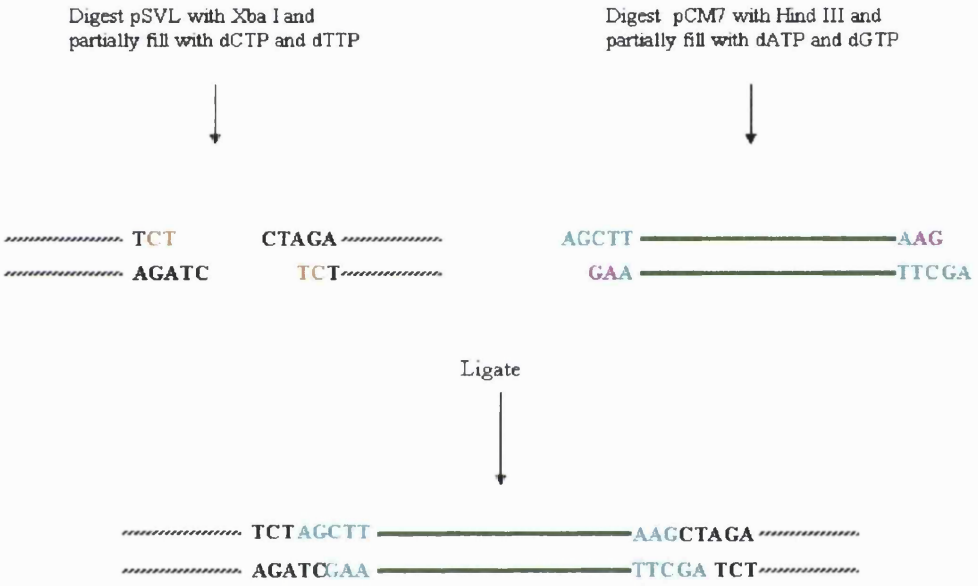


Figure 3.1. A map of the plasmid pSVL (adapted from Pharmacia) which was used to produce chimeric CAT-MyHC 3'UTR constructs for transfection studies. The plasmid contains SV40 promoter and polyadenylation signal sequences for expression of genes inserted into the multiple cloning site.

Since there were no compatible restriction sites between the CAT gene and the multiple cloning site within pSVL for ligation, a partial fill-in strategy was employed to generate compatible cohesive ends. The recessed 3' termini were filled using the polymerase activity of the Klenow fragment of *E.Coli* DNA polymerase I in the presence of the appropriate dNTPs.



5µg of plasmid pCM7 was digested with *Hind* III at 37° C for 2 hours, in a 20µl volume. Then 1µl of each of dATP and dGTP at 2 mM each and 1µl of Klenow enzyme (5u/µl, Promega) were added to the digest. The reaction was incubated for a further 30 minutes at 37° C and then the Klenow DNA polymerase enzyme was inactivated by heating at 75° C for 10 minutes. This reaction was electrophoresed on a 1% agarose gel and the fragment corresponding to the CAT gene was excised and purified from the gel.

The plasmid pSVL was treated in a similar manner. 2µg of the plasmid DNA was digested with *Xba* I restriction enzyme reaction for 2 hours at 37° C. A small aliquot of the digest was checked for linearization on an agarose gel. Then 1µl of

each of dCTP and dTTP at 2 mM each, and 1µl of Klenow DNA polymerase were added and the reaction incubated for 30 minutes at 37° C. After heat inactivation of the Klenow as described above, the plasmid DNA was cleaned by phenol/chloroform extraction and then precipitated with 1/10 volume 3 M sodium acetate, pH 4.8, and 2 volumes of ice-cold 100% ethanol. The plasmid was then resuspended in TE buffer at a concentration of 100 ng/µl. Ligation of CAT and pSVL was performed by mixing 200 ng of linearized pSVL with 6µl of purified CAT fragment, ligase buffer, and 4 units of T4 DNA ligase in a 10 µl reaction. The ligation was incubated at 14° C overnight.

Transformation of competent DH5α cells (GIBCO BRL) was carried out by adding 2µl of the ligation to 50µl of DH5α cells and incubating on ice for 30 minutes. The cells were heat-shocked at 37° C for 30 seconds and then incubated on ice for 2 minutes. 450µl of LB medium was added and the cells were then incubated with shaking at 37° C for 1 hour. They were plated out on LB agar plates containing ampicillin (100 µg/ml) and incubated overnight at 37° C.

Colonies were picked and used to inoculate medium for growth of cultures. Plasmid preparations from overnight cultures were digested with restriction *Xho* I and *Sac* I enzymes to screen for CAT insert. Positive plasmid preparations were tested for orientation of the insert by double digestion with *Xho* I and *Eco*R I.

The plasmid pSVLCAT▲ contains the CAT gene minus its 3'UTR. This plasmid was constructed by digesting pCM7 with *Hin*d III and *Sca* I, which removed part of the coding region and the 3'UTR of the CAT gene (see figure 3.2). The coding region was then repaired with the use of sense and antisense oligonucleotides incorporating *Hin*d III site.

Sense oligonucleotide

Hind S 5' ACTGCGATGAGTGGCAGGGCGGGGCGTAAA 3'

Antisense oligonucleotide

Hind A 5' AGCTTTTACGCCCCGCCCTGCCACTCATCGCAGT 3'

The oligonucleotides were designed such that once annealed the double stand was blunt at the 5' end and the 3' end had a cohesive end, compatible with a *Hind* III restricted end. This eliminated the need to digest the repaired CAT DNA further prior to cloning into the vector pGEM4Z. 20μg of pCM7 was digested with *Hind* III and *Sca* I restriction enzymes in a double digest for 4 hours at 37° C. The CAT *Hind* III / *Sca* I fragment was gel purified after electrophoresis on an agarose gel.

To repair the coding region, 100 pmoles (~1μg) of each of the oligonucleotides were annealed together by incubating at 90° C for 5 minutes in the presence of 10 mM Tris-HCl, pH 8.0, 2 mM $MgCl_2$, 50 mM NaCl, 1 mM EDTA, and allowing to cool slowly to room temperature. The annealed oligonucleotides were extracted with phenol/chloroform and precipitated with 2 volumes of ice-cold absolute ethanol. After resuspension in TE buffer, pH 8.0, the annealed oligonucleotides were phosphorylated. Buffer consisting of 70 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 0.5 mM ATP and 10 units of T4 polynucleotide kinase (Promega), were added to the annealed oligonucleotides. The reaction was incubated at 37° C for 30 minutes and then the polynucleotide kinase was heat-inactivated by incubating the reaction at 70° C for 10 minutes. The reaction was extracted with phenol/chloroform and precipitated with ethanol.

To ligate the phosphorylated, annealed oligonucleotides to the CAT *Hind* III / *Sca* I fragment, approximately 1μg of oligonucleotides were mixed with 1μg of CAT fragment, ligase buffer and 4 units of T4 DNA ligase. The ligation was performed at 14° C overnight.

The repaired CAT fragment was then inserted into pGEM4Z (Promega) as follows. pGEM4Z was digested with *Hind* III for 4 hours at 37° C, and then heated at 75° C for 10 minutes to inactivate the restriction enzyme. After cooling to room temperature, the linearized plasmid was dephosphorylated to enhance ligation to insert and reduce self-ligation. 0.01 units of calf intestinal alkaline phosphatase (CIAP) per pmol of DNA ends and dephosphorylation buffer (50 mM Tris-HCl, pH 8.5, 0.1 mM EDTA) were added to the digested DNA. The reaction was incubated at 37° C for 30 minutes, and then phenol/chloroform extracted and ethanol precipitated. *Hind* III digested pGEM4Z was mixed, in a 1:5 molar ratio of vector to insert, with

repaired CAT DNA, ligase buffer and 4 units of T4 DNA ligase and incubated at 14° C overnight.

2μl of the ligation reaction were used to transform competent DH5α cells. Cells were plated on LB agar plates containing ampicillin (100μg/ml), X-gal, and IPTG, allowing blue/white selection. Plasmid preparations from positive colonies were digested with *Hind* III to verify presence of repaired CAT insert. Plasmids with correct inserts were sequenced to ensure that the sequence of the CAT gene had been repaired correctly.

To construct the plasmid pSVLCAT▲, a two-base pair fill-in strategy as described for the construction of plasmid pSVLCAT was employed to clone the CAT gene minus its 3'UTR into pSVL. pGEM4ZCAT▲ plasmid DNA was digested with *Hind* III and then partially filled-in with dATP and dGTP using Klenow enzyme After inactivation of the Klenow enzyme, the reaction was electrophoresed on a 1% agarose gel and the fragment corresponding to CAT▲ was excised and gel purified. Ligation of CAT▲ DNA to pSVL plasmid, which had been digested with *Xba* I and filled in with dCTP and dTTP, and transformations, were carried out as described above for pSVLCAT. The orientation of CAT▲ within pSVL was determined by double digestion with *Xho* I and *Eco*R I enzymes of plasmid preparations that contained inserts.

The plasmids pSVLCAT▲β, pSVLCAT▲2a, pSVLCAT▲2b, and pSVLCAT▲2x were constructed by placing the MyHC β, 2a, 2b, and 2x 3'UTRs into the *Sma* I cloning site of pSVLCAT▲. Digestion with *Sma* I produces blunt ends, therefore the MyHC 3'UTRs had to be blunted in order to clone them into the *Sma* I site of pSVLCAT▲. The pCRII plasmids containing MyHC 3'UTRs were digested with *Hind* III and *Xba* I enzymes for 4 hours at 37° C. Then 2μl of a dNTP solution containing dATP, dCTP, dGTP, and dTTP at 2 mM each and 5 units of Klenow enzyme were added to the reaction to fill in the ends. The reactions were incubated for a further 30 minutes at 37° C and then heated to inactive the Klenow enzyme. The digests were electrophoresed on a 1.2% agarose gel, and bands corresponding to the 3'UTRs were excised and purified from the gel. The plasmid

pSVLCAT▲ was digested with *Sma* I enzyme for 2 hours at 30° C. The digested plasmid was dephosphorylated by adding 1 unit of CIAP per pmol of DNA ends with dephosphorylation buffer and incubating at 55° C for 1 hour. The plasmid DNA was then extracted with phenol/chloroform and ethanol precipitated. Ligations of pSVLCAT▲ and blunted MyHC 3'UTRs were set up as follows. 200ng of plasmid were mixed with 6µl of 3'UTR, ligase buffer, and 400 units of T4 DNA ligase (New England Biolabs), and incubated at room temperature overnight. 2µl of the ligation reactions were used to transform competent DH5α cells. Cells were plated on LB agar containing ampicillin (50µg/ml) and incubated at 37° C overnight.

Plasmid DNA preparations from overnight cultures of colonies were digested with *Xho* I and *Sac* I and run on a gel together with pSVLCAT▲ which had also been digested with *Xho* I and *Sac* I to determine which clones contained 3'UTR inserts. Positive clones containing MyHC 3'UTRs were identified by comparing the size of the CAT▲ plus 3'UTR band with CAT▲. Positive clones were sequenced to verify the presence of 3'UTR insert and determine the orientation.

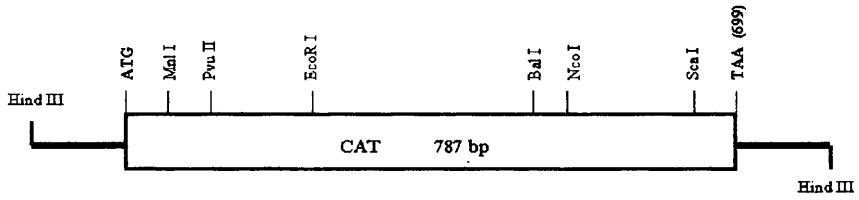


Figure 3.2. Schematic diagram of the CAT gene, indicating the positions of the start and stop codons and some restriction sites.

Figure 3.3. shows a schematic diagram of the CAT constructs made The SV40 late promoter drives CAT expression from the ATG within the CAT gene. The stop codon is provided from within the CAT gene, and downstream splicing and polyadenylation signals for mRNA processing are provided by sequences within pSVL plasmid (see figure 3.1.).

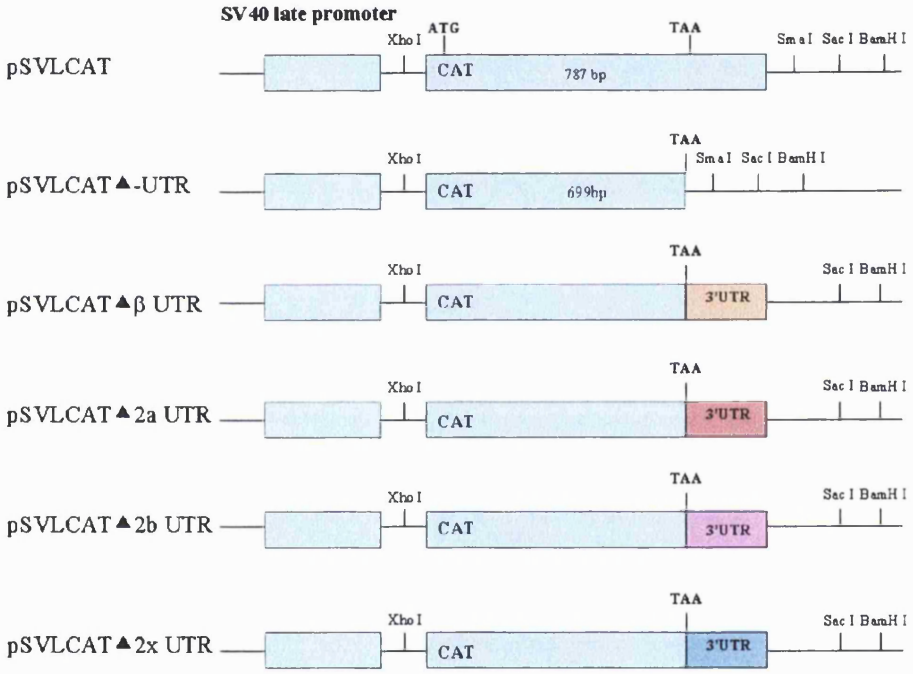


Figure 3.3. Schematic diagram showing the CAT constructs used to transfect C2C12 myoblasts. To make pSVLCAT, the CAT gene was excised form pCM7 and cloned into the *Xba* I site in pSVL. The CAT 3'UTR was removed and replaced with MyHC 3'UTRs. The entire MyHC 3'UTR, including the poly A signal was used to produce hybrid constructs. The expression of CAT and hybrid CAT-MyHC 3'UTR constructs is driven by the SV40 late promoter, and transcripts are spliced and polyadenylated using the SV40 processing signal within pSVL.

3.2.3. Large scale preparation of plasmid DNA for transfection.

Plasmids for use in transfection studies were prepared from DH5α bacterial cultures selectively grown in the presence of 50μg/ml ampicillin. The bacteria were lysed and plasmid DNA purified according to the protocol from QIAGEN megaprep purification kit. The DNA was resuspended in TE buffer, pH 8.0 and the concentration and purity determined by measuring the OD at 260 and 280 nm.

3.2.4. Cell culture.

The C2C12 myoblast cell line (Blau *et al*, 1983), a subclone of the C2 line isolated from mouse skeletal muscle (Yaffe & Saxel, 1977) was used for studying the effects of MyHC 3'UTRs on CAT gene expression and mRNA stability. Cells were seeded onto plates and maintained as proliferating myoblasts in growth medium containing Dulbecco's modified Eagle's medium (DMEM) with 4.5g/l glutamax (GIBCO BRL) supplemented with 10% (vol./vol.) foetal calf serum and ampicillin and gentamycin (0.1 mg/ml each, SIGMA). Cultures were incubated in a humidified atmosphere of 95% air and 5% CO_2 at 37° C. Cells were passaged when 80% confluent by rinsing twice with PBS and incubating for 2 minutes in 1 x trypsin/EDTA solution. Cells were counted and re-seeded at the appropriate density.

3.2.5. Optimization of transfection conditions.

To obtain the highest efficiency of transfection and maintain low toxicity, optimization of the transfection conditions was required. The amount of plasmid, lipid, and cell numbers for optimal transfection were determined using pSVLCAT plasmid and LipofectAMINE (GIBCO BRL). LipofectAMINE is 3:1 (w/w) liposome suspension of the polycationic lipid 2,3-dioleyloxy-N-[2(sperminecarboxamido)ethyl]-N,N-dimethyl-1-propanaminium trifluoroacetate (DOSPA) and the neutral lipid dioleyl phosphatidylethanolamine (DOPE) in water.

For optimization of transfection conditions, myoblasts were seeded initially at a density of 150,000 cells per dish in 60mm diameter tissues culture dishes and incubated in growth medium overnight. The following day plasmid DNA and lipid complexes were prepared for transfecting into the cells. pSVLCAT plasmid (2.5-10

μg) and LipofectAMINE reagent were diluted separately in 150μl of serum-free medium (OPTI-MEM®, GIBCO BRL). The diluted lipid and plasmid DNA were then combined and incubated at room temperature for 45 minutes to allow DNA-liposome complexes to form. While complexes were forming, the cells were rinsed twice with serum-free medium. Then for each transfection 2.7 ml of serum-free medium was added to the DNA-lipid complex. After gentle mixing, the diluted complex was added to the monolayer of cells. Cells were incubated with the lipid-DNA for 6 hours at 37° C in a CO_2 incubator. The transfection mixture was then removed and replaced with DMEM supplemented with 5% horse serum, and ampicillin and gentamycin (0.1mg/ml each) to induce differentiation.

3.2.6. Transfection of 3'UTR constructs.

To evaluate the effects of MyHC 3'UTR on CAT reporter expression, the plasmid constructs pSVL, pSVLCAT, pSVLCAT▲, pSVLCAT▲β, pSVLCAT▲2a, pSVLCAT▲2b, and pSVLCAT▲2x were transfected into C2C12 myoblasts. The plasmid pCHII0 (Hall *et al,* 1983) which contains the beta-galactosidase reporter gene expressed from the SV40 early promoter was co-transfected with each 3'UTR plasmid as a measure of transfection efficiency (Herbomel *et al*, 1984). 1 x 10^5 C2C12 cells were seeded onto 60-mm culture dishes, and incubated in growth medium overnight. The following day 5μg of pCHII0 and 5μg of 3'UTR plasmid DNA were diluted in 150μl of OPTI-MEM. 40μg of LipofectAMINE was diluted in 150 μl of OPTI-MEM. The lipid and DNA were then combined and incubated at room temperature for 45 minutes to allow complex formation. The lipid-DNA complex was then diluted by addition of 2.7 mls of OPTI-MEM. The diluted lipid-DNA was overlaid on to the C2C12 cells, which had been rinsed twice with OPTI-MEM. The cells were incubated for 6 hours before the transfection mixture was removed and replaced with differentiation medium. Transfections were carried out in triplicate for each plasmid.

3.2.7. Harvesting of transfected cells.

Cells were harvested 60 hours post-transfection and assayed for protein concentration, beta-galactosidase activity, and CAT activity. Cells were rinsed with PBS twice, and then scraped off the dishes using a cover slip in 100µl of 250 mM Tris-HCl, pH 7.5. Extracts were derived from the cells by passaging them repeatedly through a 25-gauge needle and syringe (Lee *et al*, 1988). The disrupted cells were centrifuged at 14,000 rpm, 4° C for 10 minutes to pellet cell debris. The supernatant was carefully removed and stored at –20° C.

3.2.8. Protein Assays.

Cell extracts were assayed for protein concentration using the Bicinchoninic acid (BCA) assay (Smith *et al*, 1985). The assay consists of mixing 1 volume of sample (standard or unknown) with 20 volumes of standard working reagent. The standard working regent was prepared by mixing 50 volumes of BCA and 1 volume of 4% $CuSO_4$ $5H_2O$ To quantify protein concentration in the cell extracts, 5µl of each sample was made up to 50µl with distilled water. Bovine serum albumin (BSA) was used to prepare a set of standards of known concentration in a total volume of 50µl. 1 ml of working reagent was added to each standard and to each sample. The reactions were vortexed to mix and then incubated at 60° C for 30 minutes. The reactions were allowed to cool to room temperature and the absorbance of each sample was measured at 562 nm. Readings from BSA standards were used to produce a standard curve from which the protein concentrations of the cell extracts were determined.

3.2.9. Beta-galactosidase assays.

The beta-galactosidase assay kit from Stratagene was used to determine enzyme activity from cell lysates of cells co-transfected with pCHII0 and 3'UTR constructs. 50µl of extract were pipetted into an Eppendorf tube and double distilled water added to make the volume up to 100µl. 900µl of buffer A (100 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 50 mM β-mercaptoethanol) was added and the lysates were then incubated at 37° C for 5 minutes. Then 200µl of ONPG (4mg/ml in 100

mM NaH_2PO_4) was added. The reactions were vortexed for 5 seconds, and incubated at 37° C for 3 hours. The reactions were stopped by addition of 500µl of 1 M Na_2CO_3. The optical density was then read at 420 nm.

3.2.10. CAT assays.

Extracts were assayed for CAT activity using a phase extraction protocol (Seed & Sheen, 1988). Cell extracts were heated to 65° C for 10 minutes and then cooled to room temperature to inactivate endogenous deacetylases (Crabb & Dixon, 1987). CAT assays were performed by mixing 20µl of extract with 4µl of ^{14}C-labelled chloramphenicol (0.05µCi/µl), 5µl butyryl-Coenzyme A (5 mg/ml) and 5µl of 2 M Tris-HCl, pH.8.0 in a final volume of 50µl. Reactions were incubated for 90 minutes at 37° C. Commercially available CAT enzyme (Promega) was used as a standard to ensure that CAT activities from lysates were within the linear range. It was diluted to known concentrations and reactions performed as above. 200µl of a 2:1 (v/v) mixture of tetramethylpentadecane (TMPD) and mixed xylenes were then added to the reactions and standards. The reactions were vortexed vigorously, and then centrifuged for 10 minutes. The top phase was removed into a scintillation vial. 4 ml of scintillation fluid was added and the samples were counted in a scintillation counter.

3.2.11. Actinomycin D studies.

For studies of RNA stability, C2C12 cells were transfected as before, but the amount of lipid and DNA scaled up to allow the transfection of cells seeded at a density of 1 x 10^6 onto 150 mm dishes. 60 hours after transfection, actinomycin D (5µg/ml in DMSO) was added to each dish to arrest transcription. RNA was extracted from cells at 0, 6 and 12 hours after the addition of actinomycin D.

3.2.12. RNA extraction.

RNA was extracted from cells using the method of Chomczynski & Sacchi, 1987. Cells were rinsed twice with PBS. After removal of excess PBS, cells were lysed in the dishes with the addition of 5 ml of denaturing solution. The lysed cells

were removed into 15ml falcon tubes and then 165µl of sodium acetate, pH 4.8, 5ml of phenol, and 1ml of chloroform/isoamylalcohol were added sequentially with mixing in between. The samples were vortexed for 10 seconds three times and incubated on ice for 30 minutes. The samples were then centrifuged at 6000 rpm, 4° C, for 30 minutes. The supernatant containing the RNA was carefully removed to a clean 15ml falcon tube, and 2 volumes of ice-cold 100% ethanol were added. The RNA was precipitated at –20° C overnight. Samples were centrifuged at 6000rpm, 4° C, for 30 minutes to pellet the RNA. The supernatant was discarded, and the RNA pellet was resuspended in 500µl of denaturing solution. The resuspended RNA was transferred to an Eppendorf tubes and re-precipitated with 1 volume of isopropanol, at -20° C overnight. The precipitated RNA was then pelleted by centrifugation at 14,000rpm, 4° C for 1 hour. The RNA pellet was resuspended in 30µl of nuclease-free water. The concentration and purity of the RNA was determined by measuring the optical density at 260 and 280 nm.

3.2.13. Electrophoresis of RNA.

To analyse the RNA extracted from the actinomycin D treated cells, total RNA was separated under denaturing conditions using formaldehyde agarose gel electrophoresis (Lehrach *et al*, 1977). A 1.2% formaldehyde gel was prepared by boiling 3.6 g of agarose in 231 ml of DEPC-treated water to dissolve. After cooling to 55° C 9 ml of 37% formaldehyde and 60 ml of 5x MOPS buffer (0.1 M MOPS, pH 7.0, 40 mM sodium acetate, 5 mM EDTA, pH 8.0) were added. The gel was poured and allowed to set. RNA samples were prepared for electrophoresis as follows. 20µg of total RNA was mixed with 4µl of 5x MOPS buffer, 6µl of formaldehyde, and 20µl of formamide. Samples were incubated at 65° C for 15 minutes and then chilled on ice. After centrifugation briefly to collect all of the fluid at the bottom of the microfuge tubes, 4µl of gel loading buffer (50% glycerol, 1 mM EDTA pH 8.0, 0.25% bromophenol blue, 0.25% xylene cyanol FF) was added to each sample. Samples were then loaded onto the gel, which had been pre-run for 5 minutes at 5V/cm, and electrophoresed in 1x MOPS buffer overnight at 40V.

3.2.14. Northern blotting.

After electrophoresis the gel was briefly stained in ethidium bromide for visualization purposes. The gel was destained and rinsed in DEPC-treated water to remove excess formaldehyde prior to transfer of the RNA to a nylon membrane. A capillary blot was set up essentially as described in Sambrook *et al.* A piece of Whatman 3MM paper was placed on an upturned gel casting tray to form a support equal to or greater than the size of the gel. The support was placed in a large plastic tray and 20x SSC added until the level of the liquid reached almost to the top of the support. Once the 3MM paper was thoroughly wet, any air bubbles present were smoothed out using a sterile pipette. The gel was placed on top of the support in an inverted position, making sure there were no air bubbles between it and the 3MM paper. The gel was surrounded, but not covered, with Saran Wrap film to prevent liquid flowing directly to paper towels placed on top of the gel. A piece of Hybond N+ membrane (Amersham) cut to the exact size was placed on top of the gel, smoothing out air bubbles as before. Then 2 pieces of 3MM paper cut to the same size as the gel were wetted in 2x SSC and placed on top of the Hybond membrane. After smoothing, a stack of paper towels was placed on top. A glass plate was put on top of the towels and weighed down with a weight. The transfer was allowed to proceed overnight.

Following transfer, the blotting set-up was dismantled and the membrane was carefully removed. After marking the positions of the wells, the membrane was soaked in 6x SSC for 5 minutes. The membrane was removed from the 6x SSC and dried at room temperature for 30 minutes. The dried filter was then placed between two pieces of 3 MM paper and baked for 2 hours at 80° C. The baked filter was stored at 4° C until required.

3.2.15. Labelling of probes and hybridization.

Specific RNAs were detected on the membrane using DNA probes labelled with ^{32}P. The Prime-a-Gene® Labelling System from Promega was used to label all probes. This kit, which uses a mixture of random hexamers to prime DNA synthesis

in vitro from any linear double-stranded DNA template, is based on the method developed by Feinberg & Vogelstein, 1983, 1984.

The c-myc and CAT probes were used for Northern analysis. The CAT probe corresponds to the CAT minus its 3'UTR used in constructs for transfection. The c-myc cDNA probe was kindly provided by Dr John Hesketh in the form of a plasmid, pT7-2. The DNA templates were prepared by digesting with the appropriate enzymes to yield gene specific fragments: pGEM4ZCAT▲ was digested with *Hind* III, and pT7-2 (c-myc) with *Xho* I. Digests were electrophoresed on an agarose gel and the bands corresponding to CAT and c-myc fragments were gel purified. The concentrations of the purified fragments were determined by measuring the OD at 260 nm.

Labelled probes were synthesized as follows. The double-stranded DNA template was denatured by heating at 100° C for 2 minutes and rapidly chilling on ice. The labelling reaction was assembled on ice by mixing 25ng of denatured DNA template with labelling buffer (50 mM Tris-HCl, pH 8.0, 5 mM $MgCl_2$, 2 mM DTT, 200 mM HEPES, pH 6.6, and 26 A_{260} units/ml of random hexadeoxyribonucleotides.), 20 μM each of dATP, dTTP, dGTP, 20μg nuclease-free BSA, 50 μCi of [α-^{32}P]dCTP (3000 Ci/mmol), and 5 units of DNA polymerase I Klenow fragment, in a final volume of 50 μl. The labelling reactions were allowed to proceed by incubating at room temperature for 2 hours. They were then heated at 100° C for 2 minutes to stop the reaction and subsequently chilled on ice. EDTA was added to a final concentration of 20 mM and the labelled probe was purified from unincorporated labelled dCTP by size exclusion chromatography using Sephadex® G-50 spin columns (Pharmacia).

Hybridizations with labelled probes were performed in hybridization bottles. The membrane was incubated with pre-hybridization solution consisting of 10% dextran sulphate, 1 M NaCl, 50% formamide, 10x Denhardt's reagent (0.2% BSA, 0.2% polyvinylpyrrolidone, 0.2% ficoll), 50 mM Tris-HCl, pH 7.6, 0.1% $Na_4P_2O_7$. 10 H_2O, 1% SDS, and 100μg/ml of denatured salmon sperm DNA for 3-4 hours at 42° C in a HYBAID oven. The pre-hybridization solution was then poured out and replaced with hybridization solution (same as pre-hybridization solution) containing

1 x 10^6 cpm/ml labelled probe. The membrane was incubated with the probe at 42° C overnight. After hybridization the blot was transferred to a tray for washing. The blot was washed in 6x SSC/0.1 % DS for 30 minutes, followed by two 15 minute washes in 0.5x SSC/0.1 % SDS. The blot was wrapped in Saran Wrap and placed with x-ray film and intensifying screens at -70° C for autoradiography. The blot was stripped by pouring a boiling solution of 0.5% SDS over the membrane and allowing to cool to room temperature with gentle agitation.

3.2.16. Statistical analyses

Statistical significance analyses on the data obtained from the transfection optimization experiments and between CAT activities from chimeric CAT/3'UTR constructs performed using the SIGMASTAT programme.

3.3. Results.

3.3.1. PCR amplification of MyHC 3'UTRs.

In order to study any potential role that the skeletal MyHC 3'UTRs may have in regulation of gene expression at the level of translation and mRNA stability, the 3'UTRs only were amplified from the MyHC 3'ends that were previously amplified in RACE-PCR (see chapter 1, section 2.2.5). Since the 3'UTRs are somewhat divergent in their sequences, it was necessary to design specific pairs of primers for each 3'UTR to be amplified. To facilitate subsequent cloning and manipulations of the DNA, restriction sites were incorporated into the primers. The 3'ends were firstly excised from the pCRII plasmids, into which they were originally cloned, and purified before using as a template for PCR. This was carried out to avoid problems of non-specific amplification from plasmid DNA. Figure 3.4. shows the amplification product for each 3'UTR. The expected size of PCR products are 123, 143, 128 and 125 base pairs for slow type I, 2a, 2b, and 2x 3'UTRs respectively. From figure 3.4 it can be seen that the 3'UTRs amplified to approximate the correct size, although to determine the size accurately, the PCR products need to be electrophoresed on a polyacrylamide gel. The PCR products were cloned into pCRII and the identity of each clone confirmed by sequencing.

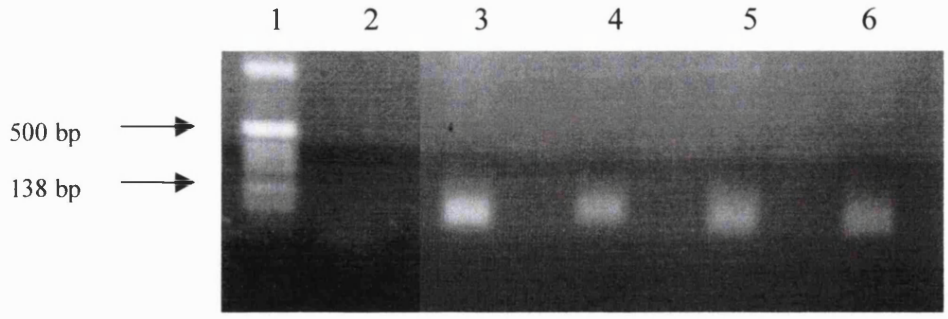


Figure 3.4. PCR amplification of myosin heavy chain 3'UTRs. Amplification was carried out with 32 cycles of denaturation at 94°C for 1 minute, annealing at 55°C for 1 minute, and elongation at 72°C for 2 minutes using isoform specific primers. Lane 1 marker; lane 2 no template control; lanes 3, 4, 5, and 6, slow type I, 2a, 2b, and 2x 3'UTR amplification products

3.3.2. Plasmid constructs.

The plasmid pSVL was used as the basis of all the constructs used for transfection studies. It is designed for high-level transient expression in eukaryotic cells, and genes inserted into the multiple cloning site are expressed from the first ATG codon using the SV40 late promoter. The plasmid pCM7 was used as a source of the CAT reporter gene. It was excised from pCM7 and inserted into pSVL. Figure 3.5. shows all of the CAT constructs digested with *Xho* I and *Sac* I enzymes. The double restriction enzyme digestion was used as a means of testing for insertion of CAT, CAT▲, and CAT▲ plus MyHC 3'UTRs. The difference in size of CAT and CAT▲ can be clearly seen.

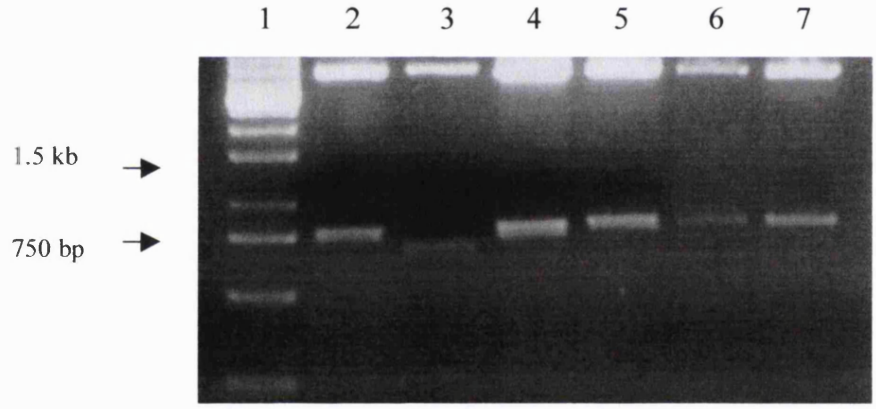


Figure 3.5. CAT and CAT hybrid constructs, digested with *Xho* I and *Sac* I restriction enzymes. Lane 1 marker; lane 2 pSVLCAT; lane 3 pSVLCAT▲; lane 4 pSVLCAT▲β; lane 5 pSVLCAT▲2a; lane 6 pSVLCAT▲2b; lane 7 pSVLCAT▲2x.

3.3.3. Optimization of transfection assays.

When transfecting cells, whatever method is used, it is usually necessary to optimize several parameters, such as ratio of transfection reagent to DNA and amount of DNA, in order to maximize transfection efficiency (Trivedi & Dickson, 1995). The efficiency of transfection and expression can vary as a function of cell density and amount of lipid and DNA used. Therefore, optimization is particularly important in transient transfection especially where multiple constructs are to be compared. Lipofection with LipofectAMINE, a liposome suspension of polycationic and neutral lipid, was employed for the studies carried out in this chapter. Although high levels of transfection are generally achieved with small amounts of DNA and LipofectAMINE, the optimal efficiency varies with cell and plasmid type. The amount of positive charge contributed by the cationic component of LipofectAMINE has to equal or exceed the amount of negative charge contributed by the phosphates on the plasmid DNA, resulting in a net neutral or positive charge. The optimal amount of DNA depends on the type of DNA and the target cell line. Therefore, initial experiments were carried out to establish the optimal lipid to DNA ratio to maximize transfection efficiency. 2.5, 5 and 10μg of plasmid pSVLCAT were mixed with LipofectAMINE in ratios of 1:2, 1:3, 1:4, and 1:5, and transfected into C2C12 cells which had been seeded at 1.5 x 10^5 cells per 60 mm dish. Lysates from harvested cells were then tested for CAT activity, using a quantitative assay. The results of the effect of lipid to DNA ratio on CAT activity are shown in figure 3.6. At a DNA to lipid to ratio of 1:2, very little CAT activity was observed, even with 10μg of DNA. Increasing the DNA to lipid ratio results in increased activity, with a peak observed at a ratio of 1:4. However, this activity decreases with a ratio of 1:5. The peak ratio of 1:4 using 2.5μg of DNA is similar to the levels of activity observed at ratios of 1:3, 1:4, and 1:5 when 5μg of DNA is used to transfect. The greatest CAT activity was seen when 10μg of plasmid DNA was complexed with lipid in a ratio of 1:4, which was statistcially significant compared to other DNA lipid combinations ($p<0.001$, see appendix). Although this ratio requires quite a high amount of lipid, very little cell toxicity was observed with this amount of LipofectAMINE since incubations with lipid/DNA complexes did not exceed 6 hours . The optimal cell

density for transfection, using 10µg of DNA complexed with 40µg of lipid was then determined. Cells were seeded at densities of $1x10^5$, $2x10^5$, $3x10^5$, and $4x10^5$ per dish and transfected. Figure 3.7. shows the results of the optimization of cell numbers for optimal transfection conditions. The highest CAT activity wasoserved at a cell density of 100,000 cells per 60 mm dish. However, the differences between the means of the CAT acitvities at different cell densities did not show any statistical differences (see appendix)

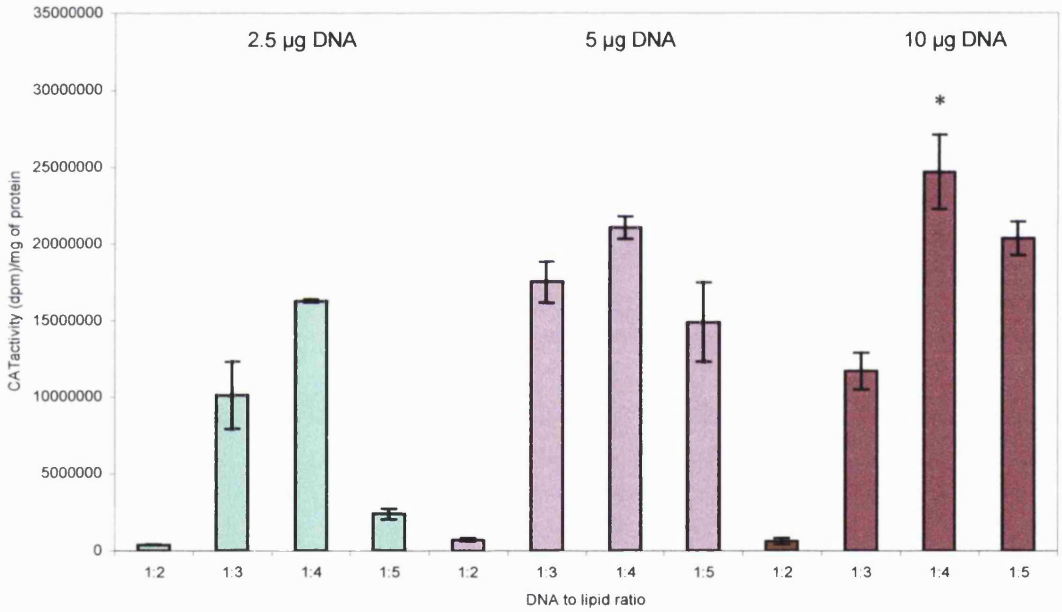


Figure 3.6. Graph showing the effects of DNA to lipid ratio on CAT activity. C2C12 cells were transfected with 2.5, 5 and 10µg of pSVLCAT plasmid DNA complexed with LipofectAMINE at the indicated ratios of DNA to lipid. The standard error is given for each mean (n=3), and the level of significance between the means is denoted as * =$p<0.001$

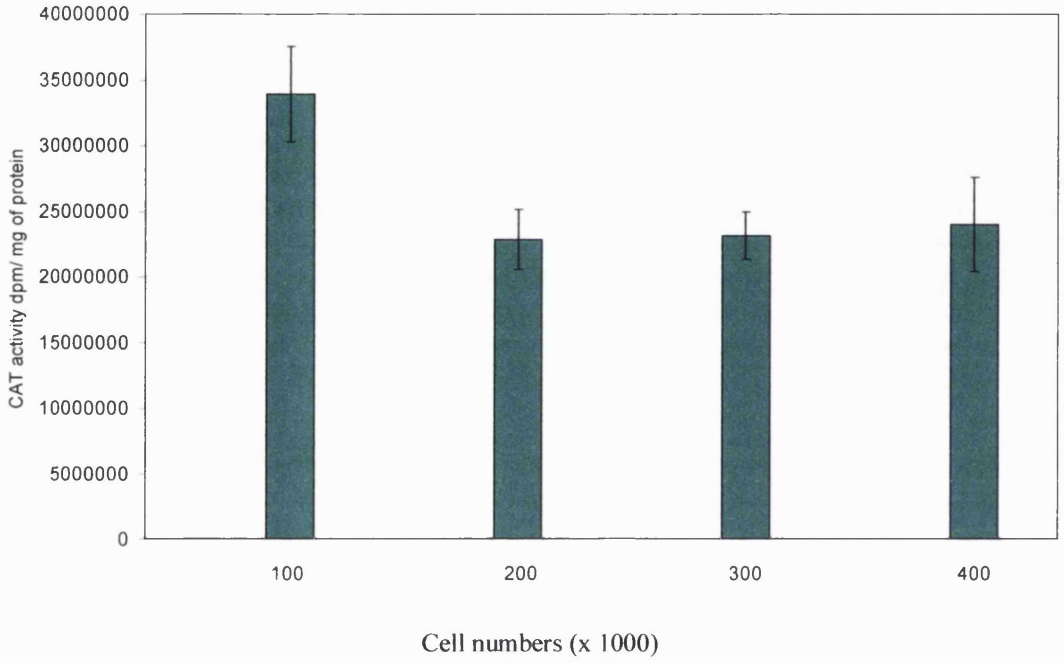


Figure 3.7. Graph showing the effects cell density on transfection efficiency. 10μg of pSVLCAT mixed with 40μg of lipid were used to transfect myoblasts seeded at the indicated cell densities. The standard errors of the means are shown (n=3). Statistical analysis (ANOVA) showed that there was significance between the means.

3.3.4. Transfections with 3'UTR constructs.

C2C12 cells were transfected with CAT and hybrid CAT▲/MyHC 3'UTR constructs to assess the effects of the 3'UTRs on CAT expression. 5μg of each plasmid and 5μg of pCHII0 were complexed with 40μg of LipofectAMINE, and used to transfect C2C12 myoblasts, seeded at $1x10^5$ cells per dish. pCHII0 was co-transfected with each construct as a measure of transfection efficiency and for purposes of normalization between transfected constructs.

Transfection of each plasmid was performed in triplicate, that is three dishes per construct in each experiment. Harvested cell extracts were assayed for protein, beta-galactosidase activity and CAT activity. Figure 3.8 and table 3.1 shows the average CAT activities of the different constructs from three separate transfection

(CAT minus its 3'UTR) and pSVLCAT▲/MyHC 3'UTR constructs. pSVLCAT▲β and pSVLCAT▲2x showed increased activity compared to pSVLCAT▲, but the activities of pSVLCAT▲2a and 2b were similar to pSVLCAT▲. One way ANOVA indicated a statistically significant difference between the mean values of the CAT activities of the different plasmids ($P<0.001$). Pairwise multiple comparisons were performed using the Fisher LSD Method, Student-Newman-Keuls Method and Tukey Test (SIGMASTAT). Significant differences between the means of pSVLCAT and pSVLCAT▲β ($p<0.001$), and between pSVLCAT and pSVLCAT▲2a, pSVLCAT▲2b and pSVLCAT▲2x ($p<0.005$) were indicated by the Fisher LSD Method and Student-Newman-Keuls Methods. There was significant difference in the CAT activities between the means of pSVLCAT▲β and pSVLCAT▲2a ($p<0.05$) and between pSVLCAT and pSVLCAT▲ ($p<0.001$). However with the Tukey Test, which is more stringent, the only significant differences were between pSVLCAT and pSVLCAT▲β ($p<0.001$), pSVLCAT and pSVLCAT▲2x ($p<0.005$), pSVLCAT▲β and pSVLCAT▲2a ($p<0.05$), and pSVLCAT and pSVLCAT▲ ($p<0.001$). The results of individual transfection experiments (not shown) exhibited variable levels of CAT expression from the different constructs between sets of experiments, with pSVLCAT only showing consistently low CAT activity between experiments. One way ANOVA on results of individual experiments showed significance in the differences in the mean values between plasmids in some experiments but not others.

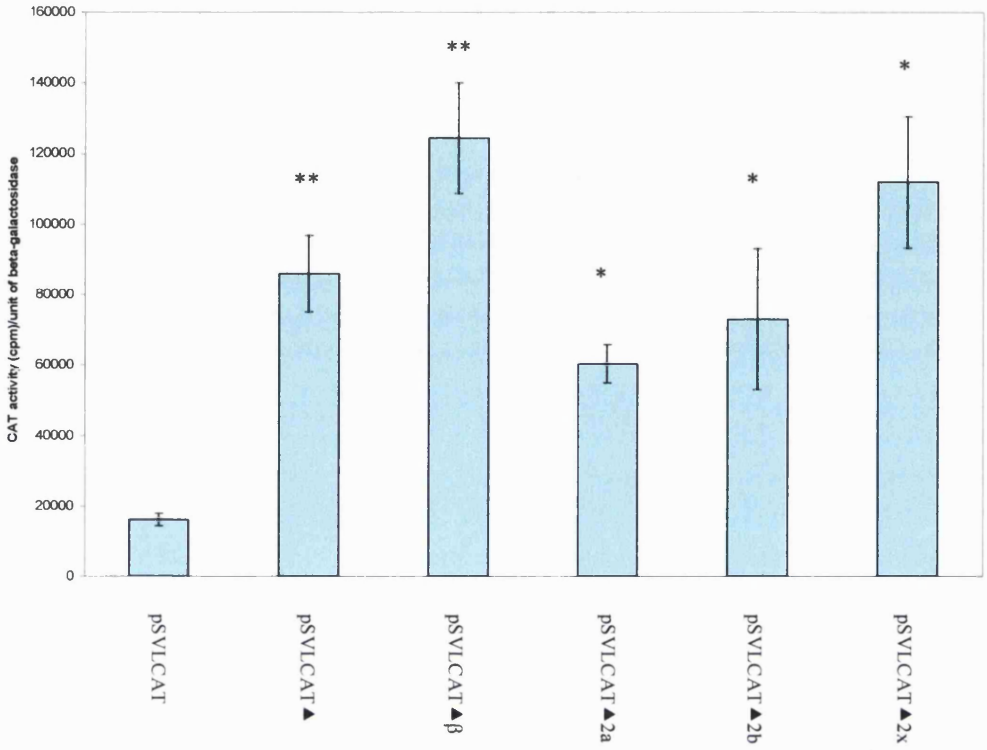


Figure 3.8. The effects of the MyHC 3'UTRs on CAT activity. Cells were transfected with CAT plasmids. Plasmid pCH110 was co-transfected with each construct for normalization purposes. Cells extracts were assayed for CAT activity, beta-galactosidase activity and protein concentration. The CAT activity of each construct is expressed normalized for beta-galactosidase activity and protein concentration. Results are the means of three separate experiments where the transfections were performed in triplicate (n=9). Significant differences between the means of pSVLCAT (control) and pSVLCAT▲β ($p<0.001$)**, and between pSVLCAT and pSVLCAT▲2a, pSVLCAT▲2b and pSVLCAT▲2x ($p<0.005$)* (Fisher LSD Method and Student-Newman-Keuls Methods) are indicated. Significant difference in the CAT activities between the means pSVLCAT and pSVLCAT▲ ($p<0.001$)** were also observed.

SUMMARY

Groups	*Count*	*Average*	*± SEM*
pSVLCAT	9	16072	± 1736
pSVLCAT▲	9	85877	± 10824
pSVLCAT'▲β	9	124277	± 15658
pSVLCAT▲2a	9	60343	± 5487
pSVLCAT'▲2b	9	72998	± 20063
pSVLCAT'▲2x	9	111772	± 18640

Table 3.1. Summary of the mean CAT activities obtained from each CAT construct.

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	6.7627E+10	5	1.3525E+10	7.87230529	0.00001777	2.408513
Within Groups	8.2469E+10	48	1718101914			
Total	1.501E+11	53				

Table 3.2. One way analysis of variance of the data in table 3.1. Results indicate statistical significance within the data group.

3.3.5. Northern analysis of Actinomycin D treated transfections.

Transfected cells were treated with actinomycin D in order to assess the stability of all of the CAT constructs. Total RNA was extracted from transfected cells at 0, 6 and 12 hours after the addition of actinomycin D. The RNA was electrophoresed on a denaturing gel and subsequently transferred to a filter for hybridization with a probe for CAT mRNA. Figure 3.9. shows the result of the actinomycin D studies. Panel A demonstrates the results of hybridization with a probe for CAT mRNA. At 0 hour of actinomycin D treatment there was very little mRNA for pSVLCAT, which correlated with the low CAT activities observed in protein extracts. The levels of pSVLCAT▲ and pSVLCAT▲/MyHC 3'UTR mRNAs in comparison were relatively high at 0 hours of treatment. No pSVLCAT mRNA was seen at 6 hours or 12 hours after treatment. pSVLCAT▲ and pSVLCAT▲/MyHC 3'UTR mRNAs were still present, although the level of pSVLCAT▲ mRNA decreased compared to 0 hours. The amount of pSVLCAT▲/MyHC 3'UTR mRNAs at 6 hours appeared to be greater than pSVLCAT▲, but were still present at similar levels at 12 hours. Panel B shows the results of hybridization with a probe for c-myc. This probe was used as an endogenous control since c-myc has a very short half-life and decays rapidly. At 6 hours after actinomycin treatment, no c-myc mRNA was present, an indication of its instability. Panel C shows the electrophoresis of the total RNA.

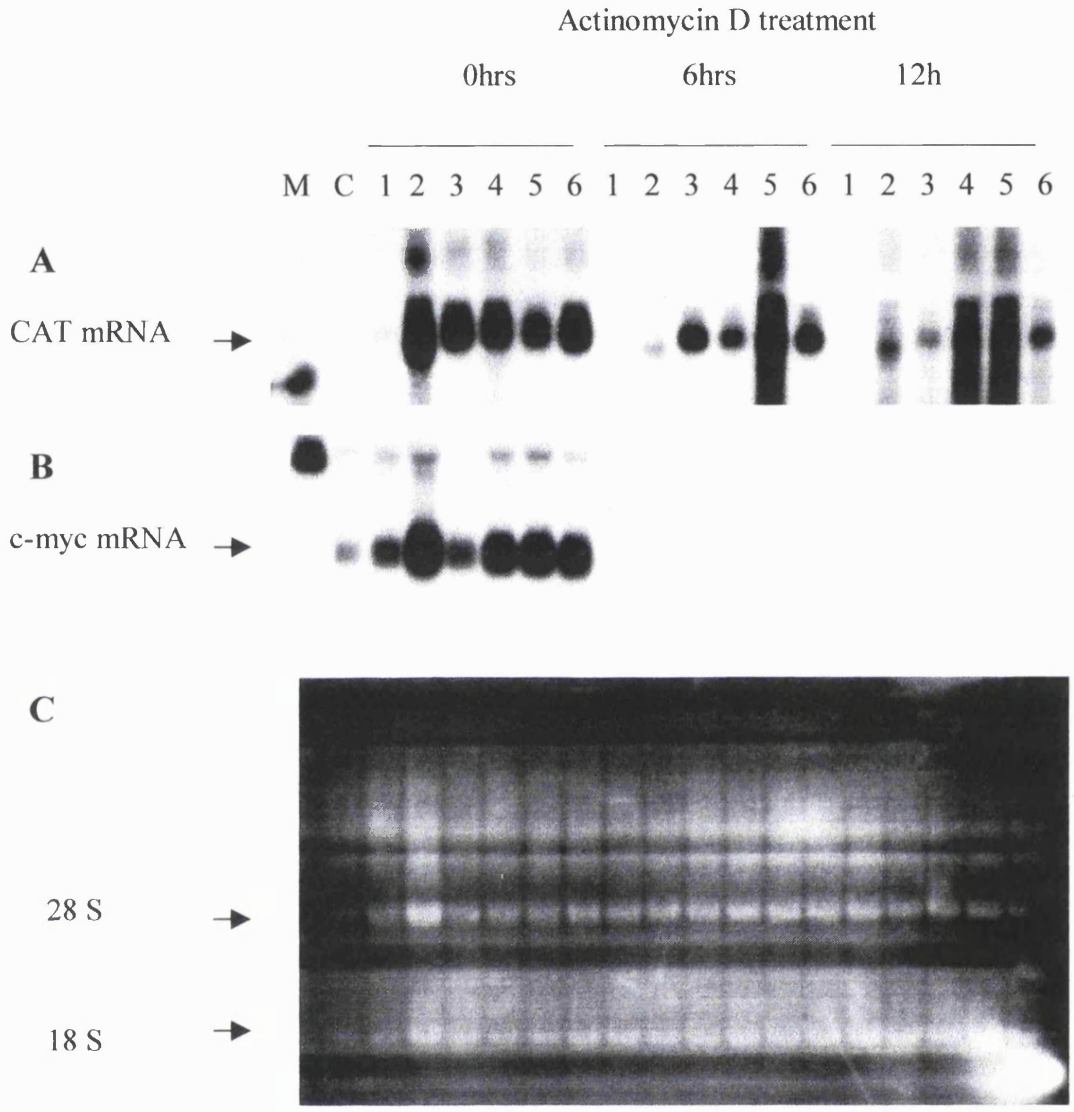


Figure 3.9. Northern analyses of total RNA extracted from cells transfected with CAT constructs. The figure shows the level of CAT transcripts (panel A) and c-myc mRNA (panel B). M =RNA marker; C =control RNA from non- transfected cells; lane 1 pSVLCAT; lane 2 pSVLCAT▲; lane 3 pSVLCAT▲β; lane 4 pSVLCAT▲2a; lane 5 pSVLCAT▲2b; lane 6 pSVLCAT▲2x. The length of actinomycin D treatment is indicated at the top.

3.4. Discussion.

Eukaryotic gene expression is a multi-step process of transcriptional, post-transcriptional, and translational events. In eukaryotic cell differentiation, there are qualitative and quantitative changes in the cytoplasmic mRNA population, therefore regulation must occur at the transcriptional and post-transcriptional level. 3'UTRs can repress or activate translation by modulating the length of the poly A tail, which in turn regulates the rate of translational initiation (Salles *et al*, 1994). They can also enhance or repress translation by association with cellular factors. It is possible that certain factors bind sequences within the 3'UTR, which in turn may promote interaction with the 5' end of the transcript. Since regulation of eukaryotic gene transcription depends on interaction of *cis*-acting regulatory sequences with *trans*-acting factors, the reporter gene system provides a useful strategy for identifying these sequences. Functional studies of mRNA stability and expression have been carried out using reporter constructs to measure increase or decrease in transcriptional activity, stability and translation of reporter genes. The use of reporter constructs overcomes the problems of differentiating between exogenous and endogenous gene products. In this chapter the CAT reporter system was employed to test the ability of the MyHC 3'UTRs to regulate transcription and expression of the CAT gene. In order to do this, the CAT 3'UTR was removed and replaced with those of the MyHC genes. These constructs were then transiently transfected into C2C12 myoblasts and the activity of the CAT gene measured by a phase extraction assay. C2C12 myoblasts express multiple myosin heavy chain isoforms. At the myoblast stage, they express mainly embryonic MyHC, and as they differentiate into multinucleated myotubes, they also express perinatal, and slow MyHC proteins (Silberstein *et al*, 1986; Weydert *et al*, 1987). In well-differentiated older myotubes fast 2b mRNA is also expressed, although corresponding protein could be not be detected (Weydert *et al*, 1987). Unlike the situation *in vivo*, there is no transition between the mRNAs, since the embryonic mRNA remains the major transcript, although this could be due to the non-homogeneity of cells. There is, however, a sequential accumulation of the perinatal and fast mRNAs. Changes in stability and/or

the rate of translation could account for accumulation of MyHC mRNA without expression of the corresponding protein during development. (Lyons *et al*, 1990; Vivarelli *et al,* 1988; Roy *et al*, 1984).

CAT/MyHC 3'UTR constructs were transfected into C2C12 myoblasts, which were allowed to differentiate. The level of CAT protein expression from cells transfected with each construct were expressed in terms of CAT activity normalized for protein concentration and beta-galactosidase activity from the co-transfected plasmid to allow for differences in transfection efficiencies between constructs. The transfected plasmids were studied at the post-transcriptional and translational level by measuring the mRNA levels and CAT activity from each construct. The results shown in figure 3.8 represent the mean CAT activities from three separate sets of transfections in which transfections were performed in triplicate. Although these results would indicate that the MyHC 3'UTRs increase CAT activity, the fact that pSVLCAT▲ also showed greater CAT activity than pSVLCAT calls into question the effects of the presence of the MyHC 3'UTRs on CAT expression. The increased levels of CAT activity in the constructs containing MyHC 3'UTRs could therefore be due to the removal of the CAT 3'UTR rather than the presence of a MyHC 3'UTR. Furthermore, the results of individual transfection experiments gave variable levels of CAT activity for each construct, with only pSVLCAT producing a consistent level of CAT activity. It must be noted that in the optimization experiments, a second plasmid was not co-transfected, hence the low expression could be due to the presence of a second plasmid. In experiments where transfected cells were allowed to differentiate in culture for a longer time (up to 10 days), it was not possible to gauge very well the level of CAT expression in lysates prepared from these cells as there was a loss of both CAT activity and beta-galactosidase activity. Therefore, the change in reporter gene expression in response to differentiation could not be determined.

There are several possible reasons for the variability between sets of transfections. There were numerous experimental steps performed before CAT activities for each construct were obtained. Therefore, experimental error could be one explanation for the variations observed. Although the transfections were

Additionally, when comparing the ability of different sequences to regulate expression of a reporter gene following transfection of cells, it was necessary to control for potential differences in transfection efficiencies. This can be done by either normalizing assays for protein concentration and conducting the experiment several times with at least two different preparations of each DNA, or by including a second plasmid, used as an internal control. This plasmid codes for a different reporter gene, and the product of this gene is used to normalize for transfection. In the transfections carried out in this chapter, the plasmid pCHII0, which codes for beta-galactosidase was co-transfected with CAT plasmids. Transfecting with beta-galactosidase plasmids has been suggested to be unreliable due to its variable sensitivity depending on isolation and storage conditions (Lu, 1992). It has also been suggested that there can be changes in beta-galactosidase activity when co-transfected with reporter constructs because of promoter cross competition (Farr & Roman, 1991), and presence of specific DNA sequences (Bergeron *et al*, 1995). Therefore, it is possible that hybrid CAT/MyHC 3'UTRs constructs were affecting the level of beta-galactosidase expression and do not necessarily reflect differences in transfection efficiency. This theory could be applied in reverse; that is, the presence of beta-galactosidase could affect CAT plasmids. Another point to be taken into consideration is that different promoters can use the same general transcription factors, and bias towards one promoter could effect availability of these transcription factors. Variation effects could also be cell specific; expression of beta-galactosidase has been shown to be dependent on promoter and promoter strength in a particular cell type (Ray & Gage, 1992). DNA purity and topology also effect efficiency of transfection. RNA contamination can reduce transfection efficiency. This can be discounted as a factor in the transfections performed here since QIAGEN purified plasmids which are relatively pure and RNA free were used. It has been suggested that supercoiled DNA is more efficiently transfected than linear DNA, and that the supercoiling of plasmid DNA affects the ability of transcription factors to activate a given promoter once bound. (Weintraub *et al*, 1986). Work published by Martin, (1990), which demonstrated variable results from CAT transfection experiments, showed that the variability could not be explained by differences in quality of DNA,

showed that the variability could not be explained by differences in quality of DNA, growth states of the cells used, culture medium employed, or the time at which transfected cells were harvested for CAT assays.

The stability of transcripts from the CAT constructs was investigated by extracting RNA from cells transfected with the CAT constructs after transcription had been arrested by the addition of actinomycin D, and analysed by Northern blotting. Total RNA was extracted from cells at three time points after actinomycin D had been added to the cells. The probe for CAT mRNA showed that the transcript from pSVLCAT was not very abundant compared to pSVLCAT▲ and pSVLCAT▲/MyHC 3'UTR transcripts at 0 hours of treatment. The high levels of pSVLCAT▲ mRNA and protein expression were unexpected, since the half-life of CAT mRNA has been postulated to be approximately 6 hours. Although the strategy for studying the MyHC 3'UTRs has been employed previously, (Han *et al,* 1990; Iwai *et al*, 1991; Moseley *et al*, 1993) the effect of removing the CAT 3'UTR has not been fully investigated and no data has been presented. Therefore, it is difficult to say whether the differences in CAT activities between pSVLCAT and pSVLCAT▲ plasmid presented here are due to sequence elements within the CAT 3'UTR that might effect transcript stability, and hence expression of the gene, or experimental error.

The abundance of the transcripts correlate with the CAT activities observed from cell lysates and although the levels of these transcripts appear to decrease with time, they are still detectable. Although it would seem that the presence of a MyHC 3'UTR stabilizes CAT mRNA, the greater abundance of pSVLCAT▲ mRNA and the fact that at 12 hours pSVLCAT▲ mRNA seems to be more abundant than at 6 hours, calls into question the reliability of transient transfection studies for studying mRNA stability. Ideally, actinomycin D studies and RNA half-life determination need to be performed on stably transfected cells that are expressing uniformly.

The control of mRNA stability and translation is increasingly being recognized as an important mode of regulating gene expression in eukaryotes, and specific sequence elements within the 3'UTR have been identified which are responsible for modulating this. From the experimental data presented in this chapter it is difficult to

obtain conclusive data that supports a role for the MyHC 3'UTRs in regulating mRNA stability and translation. The MyHC mRNAs are relatively stable transcripts with a half-life of 55-60 hours but whether this stability is mediated through the 3'UTR remains unclear.

Chapter 4

RNA Protein Interactions of Myosin Heavy Chain 3' Untranslated Regions

4.1. Introduction.

RNA-protein interactions are involved in most of the steps of gene expression, including splicing, nucleocytoplasmic transport, translation and degradation of mRNA and its precursors. These usually involve specific recognition of sequences and structural elements in mRNA molecules by different proteins. Some of the RNA-protein interactions are intrinsic to the control processes, whereas others are involved in regulating these processes in response to environmental or developmental influences. Soon after synthesis, the RNA undergoes a series of interactions with various proteins. Early RNA-protein interactions include the capping of the 5' terminal nucleotide and the 3' end formed by pre-determined cleavage sites. Pre-mRNA processing involves spliceosomal proteins and heteronuclear ribonucleoproteins (hnRNPs). The mature mRNA is then transported through the nuclear pores in to the cytoplasm accompanied by various other proteins (see reviews by Guthrie C, 1991; Rosbash M & Singer RH, 1993). At the level of translation, many protein factors are required for translational initiation, elongation and termination. mRNA-binding proteins may interact with stable secondary structures within the 5' untranslated region, such as stem-loop structures, to inhibit translational, or alternatively nucleotide sequences in these structured regions may allow binding of specific mRNA-binding proteins (reviewed by McCarthy & Kollmus, 1995).

The best characterized RNA-protein interactions at the 3' end of the transcript are those involved in 3' end formation and polyadenylation (see chapter 1), where the AAUAAA hexamer binds the cleavage and polyadenylation specificity factor and the cleavage stimulation factor to bring about the cleavage of the 3' end. Polyadenylation then occurs with the involvement of another factor, the poly A-binding protein. Other sequence elements situated within the 3'UTR which are known to regulate gene expression by controlling mRNA stability, mRNA translation, and message localization (see chapter 1) have also been shown to bind proteins through electromobility shift assays. For example, the U1 SnRNP-specific protein U1A autoregulates its production by binding to the 3'UTR of its own pre-mRNA (Boelens

et al, 1993). In some developmental systems, masking proteins 'silence' maternal mRNAs until translation of these transcripts is needed at a particular stage in development. The repression of translation of the ribonucleotide reductase and cyclin mRNAs in S*pisula*, mediated via the 3'UTR, requires the binding of a protein factor. Similarly, RNA-protein interactions within their 3'UTR is required for the inhibition of translation of the mammalian LOX mRNA and mouse protamine mRNAs (Kwon & Hecht, 1991; Ostareck-Lederer *et al*, 1994). RNA-binding proteins are also important in the control of mRNA stability mediated through the 3'UTR. A number of factors have been shown to bind AU-rich and U-rich elements within the UTR that are implicated in the degradation of mRNA (see chapter 1). Conversely, an RNA-binding protein interacts with the C-rich regions of the α-globin 3'UTR to stabilize the transcript (Wang *et al*, 1995). The mRNA localization signals contained within the 3'UTRs of transcripts such as those of *Xenopus* oocyte Vg1, chicken β-actin and vimentin mRNAs have also been show to bind proteins (Schwartz *et al*, 1992; Ross *et al*, 1997; Danowski *et* al, 1992).

Although the function of the MyHC 3'UTRs is not as yet clearly defined, evidence for RNA-protein interactions between specific regions within the 3'UTRs of many mRNAs, would suggest that they contain protein recognition sequences. Furthermore, in chapter 2, sequence alignments of the MyHC 3'UTRs highlighted the iso-specific conservation of this region across species and in particular a short sequence motif CAAAATGTGA present in the fast isoforms. This sequence motif could be a recognition site for protein binding. Predicted secondary structure analyses showed that this sequence motif lies within a loop region, which is thought to be one of the single-stranded targets for RNA-binding proteins. Therefore, the aim of the work described in this chapter was to investigate RNA-protein interactions of the MyHC 3'UTRs through electrophoretic mobility shift assays (EMSA). The basis of the EMSA, also known as bandshift or gel retardation assay, is the change in the electrophoretic mobility of a nucleic acid molecule upon binding to a protein or other molecule. Labelled nucleic acid, containing sequences of interest, is incubated with protein extract, and then electrophoresed on a polyacrylamide gel. If the nucleic acid binds one or more protein or other molecules, the migration of the nucleic acid

through the gel is retarded. Although the EMSA is a technique that has been widely used for identifying DNA-binding proteins (reviewed by Dent & Latchman, 1993), its use for studying RNA-binding proteins is more difficult. Whereas DNA can be readily synthesized and labelled, particular care is necessary when dealing with RNA to prevent contamination by RNases. For RNA bandshift analyses, the RNA sequence to be investigated for protein-binding capabilities is usually synthesized by 'run-off' transcription from template DNA, using either T7 or T3, or Sp6 bacteriophage RNA polymerases (Krieg and Melton, 1987; Melton et al, 1984; Milligan et al, 1987). These enzymes have a high specificity for their respective promoters, allowing the synthesis of large amounts of RNA of any desired sequence, including cellular RNAs. Inclusion of a radiolabelled nucleotide allows internal labelling of the RNA to produce a probe that can be readily detected by autoradiography. Several plasmids are commercially available that contain multiple cloning sites flanked by one or more phage promoter, allowing any DNA cloned downstream of the promoter to act as template for transcription. The MyHC 3'UTRs were cloned into a vector suitable for transcription, and cRNA probes were synthesized by run-off transcription. The probes were used in electrophoretic mobility assays with protein extracts derived from mouse tissues and assayed for protein binding. Competition assays were also performed to determine the specificity of protein binding.

4.2. Materials and methods.

4.2.1. Protein extraction.

Protein extracts were prepared as described by Dignam et al, 1983 with the following modifications. Liver, kidney, brain, heart, and muscle tissue were dissected from C57BL/10 black adult mice and minced in 4 mls (50-100µl/mg of tissue) of ice-cold buffer A (10 mM HEPES (pH 7.9), 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT), supplemented with the protease inhibitors aprotonin at 0.5µg/ml, pepstatin at 0.5µg/ml, leupeptin at 0.5µg/ml and PMSF at 1µg/ml. The tissue was homogenised on ice using a teflon pestle and glass tube. Nuclei, unbroken cells and connective tissue were removed from the lysate by centrifugation at 5000 rpm for 15 minutes at 4^0 C in a Sorvall SS34 rotor. The resulting pellet was retained and kept on ice and the supernatant (the cytoplasmic extract) was mixed with 0.25 volumes of extraction buffer B (250 mM HEPES, pH 7.9, 750 mM KCL, 17.5 mM $MgCl_2$, 5 mM DTT, 0.5 M EDTA, 50% glycerol), and centrifuged for 1 hour at 100,000 g in a Beckman SW 41 Ti rotor at 4^0 C. The high speed supernatant (S100/cytoplasmic extract) was carefully removed, flash frozen in liquid nitrogen and stored at $^-70^0$ C.

The "nuclear" pellet was washed briefly in buffer A, and then resuspended in 0.4 ml buffer C (420 mM NaCl, 20 mM HEPES, pH7.9, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM DTT, 25% glycerol), supplemented with protease inhibitors. After homogenization on ice, the extract was placed on a rocker for 30 minutes at 4° C. The homogenate was then centrifuged for 30 minutes in a Sorvall SS34 rotor at 15000 rpm, 4° C. The supernatant (nuclear extract) was dialyzed against 50 volumes of buffer D (20 mM HEPES, pH 7.9, 75 mM KCl, 2 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 10% glycerol) overnight at 4° C. The dialysed extracts were frozen in liquid nitrogen and stored in aliquots at -70° C.

4.2.2. Measurement of protein concentrations.

Concentrations of protein extracts were measured using the bicinchoninic acid micro assay (Smith et al, 1985) as described in Chapter 3 section 3.2.8.

4.2.3. SDS PAGE of protein extracts.

Protein extracts were analysed by vertical SDS-polyacrylamide gel electrophoresis, using a discontinuous gel system. Approximately 20μg of protein were subjected to electrophoresis through a 5% stacking gel and a 10% resolving gel (acrylamide to bisacrylamide ratio 19:1) in the manner of Laemmli (1970). The stacking gel contained 125 mM Tris-HCl, pH 6.8 and 0.1% SDS. The resolving gel contained a final concentration of 375 mM Tris-HCl, pH 8.8 and 0.1% SDS. Loading buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 2% β-mercaptoethanol and 0.001% bromophenol blue dye) was added to each sample in a ratio of 1:1. The samples were boiled in water for 3 minutes to dissociate the proteins and then loaded onto a 1.5 mm thick gel. Electrophoresis was carried out in Tris-glycine buffer (25 mM Tris, 192 mM glycine, pH 8.3, 0.1% SDS) for 5hrs at 200v using the Protean II gel system (Biorad).

After electrophoresis was complete the gel plates were carefully separated and the gel placed in a dish for proteins staining with Coomassie Blue (0.1% w/v brilliant blue G, 25% v/v methanol, 5% v/v acetic acid, SIGMA). 5 gel volumes of Coomassie Blue were poured over the gel and left to stain with gentle agitation overnight. The gel was destained in 30% methanol, 10% acetic acid.

4.2.4. Plasmid constructs for *in vitro* transcription.

For the purpose of cRNA synthesis by *in vitro* transcription, the MyHC 3' UTRs were excised from the plasmid pCRII into which they had previously been cloned (see chapter 3), and subcloned into pGEM4Z (Promega). T7 and Sp6 promoter regions that allow in vitro transcription flank the multiple cloning site in this plasmid vector. pGEM4Z was prepared for subcloning as follows: Plasmid DNA was digested with *Hin*d III and *Xba* I, at 37° C for 2 hours. The digested plasmid was purified by phenol/chloroform extraction and precipitated with 1/10 volume 3 M sodium acetate, pH 4.8 and 2 volumes of ice-cold absolute ethanol at -20° C. Each of the plasmids pCRIIβ3'UTR, pCRII2a3'UTR, pCRII2b3'UTR and pCRII2x3'UTR was digested with *Hin*d III and *Xba* I restriction enzymes. The digestions were

subsequently electrophoresed on a 1.2 % ethidium bromide agarose gel. Bands corresponding to the 3' UTRs were gel purified using Wizard PCR Preps kit (Promega), and ligated to *Hind* III and *Xba* I digested pGEM4Z. Digested pGEM4Z was mixed with purified 3'UTR in the presence of ligase buffer and units of T4 DNA ligase (New England Biolabs. Ligation reactions were incubated at 14° C overnight and used to transform competent DH5α cells. Positive clones were selected by blue/white colony screening and grown as small-scale cultures for plasmid preparations. The resulting plasmid DNAs were digested with *Hind* III and *Xba* I and then sequenced to verify the presence and orientation of 3'UTR.

4.2.5. PCR amplification and cloning of rat growth hormone 3' UTR.

To study the specificity of the myosin heavy chain 3'UTR RNA-protein interactions, a non-myosin 3'UTR was used for comparison. PCR was used to amplify the rat growth hormone (rGH) 3'UTR for this purpose. A forward primer incorporating a *Hind* III restriction sequence (rGH1) and reverse primer with an *Xba* I restriction site (rGH2) were designed from the rat growth hormone cDNA sequence (Seeburg et al, 1977).

Forward primer

rGH1 5' GCGAAGCTTGCACACACTGGT 3'

Reverse primer

rGH2 5' CGGTCTAGAGGACAAAGTGTA 3'

Amplification of the rat growth hormone 3'UTR was performed using the plasmid pC3E-rGH (provided by Gavin MacColl, Royal Free Hospital), which contains the rat growth hormone cDNA, as follows. In a 100μl volume, reaction buffer (10 mM TrisHCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl), 0.2 mM each of dATP, dTTP, dCTP, dGTP, 25μM each of rGH1 and rGH2 primers and 1μl (1μg) of pCE-rGH plasmid DNA were gently mixed and briefly centrifuged. 100μl of mineral

oil were overlaid on top of the reaction to reduce evaporation during the PCR. The reaction was heat denatured at 94° C for 5 minutes and then 2.5 units of Taq polymerase (Boerhinger) were added. Amplification was then carried out by 32 cycles of denaturation at 94° C for 45 seconds followed by annealing at 55° C for 45 seconds, followed by elongation at 72° C for 1 minute. A final elongation step at 72° C for 10 minutes was included to ensure all PCR products had 3' A overhangs. A control reaction, omitting template DNA, was included. PCR products were analysed by electrophoresis of an aliquot on a 1% agarose gel containing ethidium bromide. The band corresponding to the rGH 3'UTR was excised and purified using the Wizard PCR purification kit (Promega). The gel purified PCR product was ligated to 50 ng of pUAg vector (R&D Systems) in a 10µl volume with 3 units of T4 DNA ligase in ligation buffer (20 mM Tris-HCl, pH7.6, 5 mM $MgCl_2$, 0.5 mM ATP, 5 mM DTT). Reactions were incubated at 14° C overnight.

Transformation of INVαF' competent cells (Invitrogen) was carried out as described in chapter 2 section 2.2.6. Plasmid DNA prepared from overnight cultures of positive clones was digested with restriction enzymes to confirm presence of insert. These plasmids were then sequenced to confirm identity of the rGH 3'UTR insert.

4.2.6. Synthesis of radiolabelled probes.

T7 phage polymerase was used to synthesize RNA transcripts from DNA plasmid templates. RNase and DNase free tips and tubes were used throughout and solutions were DEPC-treated where necessary. 3µg of the plasmids pGEM4Z-β, pGEM4Z-2a, pGEM4Z-2b, and pGEM4Z-2x were linearized with either 40 units of *BamH* I for synthesis of sense cRNA or with 40 units of *Hind* III for synthesis of antisense cRNA, in a 100µl volume. pGEM4Z-rGH was linearized with *EcoR* I. Linearized plasmids were cleaned by phenol/chloroform extraction followed by clean-up with PCR prep resin (Promega). The DNA was eluted from the PCR preps column in RNase-free TE buffer, pH 8.0. The recovery of DNA was checked by agarose gel electrophoresis.

Using the Maxiscript™ In Vitro Transcription Kit (Ambion) radiolabelled RNA probes were synthesized as follows: Reactions were performed in a 30µl volume using 1µg linearized plasmid template in the presence of 40 mM Tris-HCl (pH7.9), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, 0.5 mM each of ATP, GTP, UTP, 3µM CTP, 50 µCi [α-^{32}P]CTP (800 Ci/mmol, 10 mCi/ml) NEN Dupont), 40 units RNase inhibitor (Promega), and 20 units of either T7 to generate sense probes or Sp6 polymerase to make antisense probes. Reactions were incubated at 37° C for 2 hours. DNA templates were then removed by the addition of 2 units of RNase-free DNase I to the reactions, and incubation for a further 1 hour at 37° C. The transcription reactions were then extracted with acid (pH 4.5) phenol/chlorofom/isoamylalcohol 25:24:1). A 1µl aliquot from each reaction was removed for determination of percentage incorporation of radio-label and specific activity.

Radiolabelled RNA markers were synthesized using RNA Century Marker Template Set (Ambion) which comprises of 5 linearized plasmids for use as templates for *in vitro* transcription with T7 polymerase. The marker synthesis reaction was performed in a 20µl volume using 0.5µg of Century Marker Template in the presence of 40 mM Tris-HCl (pH7.9), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, 0.5 mM each of ATP, GTP, UTP, CTP, 30 µCi [α-^{32}P]CTP (800 Ci/mmol, 10 mCi/ml) NEN Dupont), 40 units RNase inhibitor (Promega), and 20 units of T7. Reactions were incubated at 37° C for 2 hours. DNA templates were then removed by the addition of 2 units of RNase-free DNase I to the reaction and incubation for a further hour at 37° C. Transcribed RNA markers were purified from unincorporated labelled radionucleotide by precipitation at -20° C with 1/10 volume 5 M ammonium acetate and two volumes of ice cold ethanol.

Full-length transcripts were purified by polyacrylamide gel electrophoresis, using the Protean II vertical gel system (Biorad). The electrophoresis apparatus was treated to ensure it was RNase-free and gel plates, spacers and combs were made RNase-free with the use of RNaseZap (Ambion). An equal volume of gel loading buffer (95% formamide, 0.5 mM EDTA pH 8.0, 0.025% SDS, 0.025% xylene cyanol, 0.025% bromophenol blue) was added to each reaction, and the tubes were

heated at 95° C for 5 minutes. Heated reactions were then loaded together with labelled RNA markers onto a 0.75 mm thick 5% polyacrylamide, 8 M urea gel and run at 250 V for 2 hours in 1x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA). Full-length products were visualised by autoradiography, excised from the gel and placed in RNase-free tubes. The labelled RNAs were eluted from the gel slice overnight in 500 µl of elution buffer (0.5 M ammonium acetate, 1 mM EDTA, 0.2% SDS) at 37° C. The tubes were centrifuged at high speed to pellet the acrylamide and the eluate carefully removed into fresh tubes. RNAs were precipitated with 0.2 M NaCl, 40 µg yeast tRNA, and 2.5 volumes of absolute ice-cold ethanol at -20° C overnight.

The precipitated RNAs were pelleted by centrifugation at 14000 rpm, 4° C, for 1 hour, then washed in cold 70% ethanol (DEPC), and resuspended in 30µl of nuclease free water after briefly air-drying.

4.2.7. Determination of transcript yield and specific activity.

The percentage incorporation of radio-label and specific activity of each of the labelled transcripts was determined by adsorption to DE-81 filters (Whatman). These are positively charged filters and hence strongly adsorb and retain nucleic acids and are particularly good for oligonucleotides that are too small to be precipitated by trichloroacetic acid (TCA). Unincorporated nucleotides adsorb less well and are removed by extensive washing in buffer. The 1µl aliquot of each of the transcription reactions retained in section 4.2.6. were diluted by addition of 9µl of RNase-free water. 5µl was spotted onto two filters and allowed to dry by evaporation. One filter from each pair was transferred to a scintillation vial for counting the total radioactivity, and the other was transferred to a beaker containing 0.5 M Na_2HPO_4, pH 7.0. The filters were swirled in the buffer for 2 minutes and then transferred to a fresh beaker of phosphate buffer. The washing was repeated two more times, after which the filters were briefly immersed in 70% ethanol (DEPC), and then air-dried. The filters were then inserted into scintillation vials for counting and the percentage incorporation and specific activity calculated (see appendix for calculations).

4.2.8. Large scale synthesis of unlabelled RNA.

Large quantities of unlabelled β, 2a, 2b, 2x, and rGH 3'UTR RNA were synthesized for use in competition assays using the Ribomax System™ (Promega). The kit uses an enzyme mix, which contains T7 polymerase at a high concentration, inorganic pyrophosphatase, and RNase inhibitor to produce high yields of RNA transcript. Inorganic pyrophosphate precipitates at high rNTP concentrations upon incorporation of the rNTPS and inhibits the polymerase, hence the inclusion of inorganic pyrophosphatase enzyme in the mix.

Transcription reactions were carried out in 100µl volume using 10 µg of linearized plasmid in the presence of 80 mM HEPES-KOH, pH 7.5, 24 mM $MgCl_2$, 2 mM spermidine, 40 mM DTT, 7.5 mM each of rATP, rGTP, rUTP, rCTP, and 800 units of T7 polymerase. The reaction components were gently mixed and incubated at 37° C for 4 hours. Following transcription the DNA templates were removed by incubation for 1 hour at 37° C with the 10 units of RNase-free DNase I. The RNA was extracted with 1 volume of acid (pH 4.5) phenol /chloroform/isoamyl alcohol (25:24:1), and then 1 volume of chloroform/isoamyl alcohol (24:1). Unincorporated nucleotides were removed by size exclusion chromatography using Sephadex® G50 columns (Pharmacia). The column-purified RNAs were then precipitated at -20° C with 1/10 volume 3 M sodium acetate, pH 5.2, and 1 volume of isopropanol. The RNA was pelleted by centrifugation at 14000 rpm for 30 minutes at 4° C. The pellets were washed with ice cold 70% ethanol (DEPC) and briefly air-dried. The RNAs were resuspended in 50µl of nuclease free water and stored at -70° C until required. The concentration and purity of the RNAs were determined by measuring the OD at 260 nm and 280 nm.

4.2.9. RNA-protein binding reactions and electrophoresis of complexes.

All solutions were treated with DEPC, or made with DEPC-treated water and all equipment and glassware to be used were treated with RNaseZap (Ambion). RNA bandshifts were performed as described by Fajardo et al, 1994, with the following modifications. *In vitro* transcribed [^{32}P]CTP-labelled RNA probes were heated at 70° C for 5 minutes and then allowed to renature by cooling slowly to room temperature. Renatured RNA probe (~1x10^5 cpm per reaction) was incubated, in a 15µl reaction volume, with 10-20µg of protein extract in binding buffer containing 20 mM HEPES (pH 7.9), 3 mM $MgCl_2$, 40 mM KCl, 2 mM DTT, and 5% glycerol, in the presence of 200 ng/µl yeast tRNA. The reactions were incubated at room temperature for 30 minutes and subsequently treated with 1 unit of RNase T1 (GIBCO BRL) for 20 minutes at room temperature, followed by 6µg/ml (final concentration) of heparin (SIGMA) for 10 minutes at room temperature. 5µl of 50% glycerol were added to each reaction. Samples were then subjected to electrophoresis through a 1.5 mm thick non-denaturing 5% polyacrylamide gel (acrylamide/bisacrylamide ratio 65:1) in 0.5x TBE. Electrophoresis was carried using Protean II vertical gel system (Biorad), at 250V for 2 hours at room temperature. The gel was then carefully transferred to Whatman paper, and dried at 80° C on a vacuum gel drier (Biorad). Dried gels were exposed to X-ray film overnight.

4.2.10. Competition assays.

To determine the specificity of RNA-protein complexes, competition assays were carried out using unlabelled RNA probes, or the homoribonucleotide polymer poly A as competitor. Cold competitor was pre-incubated with protein extract for 20 minutes at room temperature. Labelled probe was then added and binding assays carried out as described in section 4.2.9.

4.3. Results.

4.3.1. SDS-PAGE analysis of protein extracts.

Figure 4.1. shows SDS-PAGE analysis of the proteins extracted from mouse heart, kidney, liver, brain, and skeletal muscle tissues. The extracts were analysed by SDS-PAGE to ensure that the proteins had not degraded during the process of extraction from the tissues prior to use in electromobility shift assays.

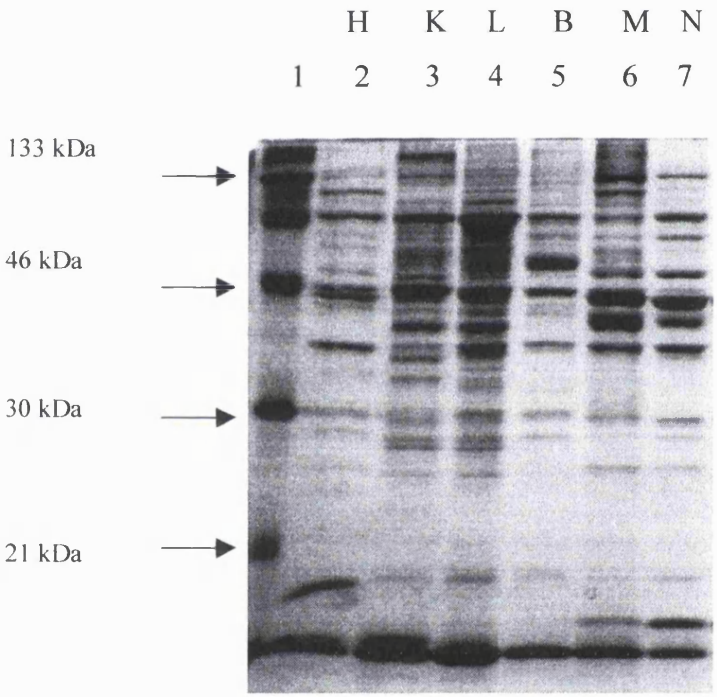


Figure 4.1. SDS-PAGE of proteins extracted from mouse tissues. The proteins were electrophoresed on a 10% polyacrylamide gel and stained with Coomassie blue. Lane 1 protein marker; lanes 2, 3, 4, 5, and 6, protein extracts from heart (H), kidney (K), liver (L), brain (B) tissue and skeletal muscle (M). Lane 7 nuclear extract from skeletal muscle tissue.

4.3.2. PCR amplification of rat growth hormone 3'UTR.

The rat growth hormone 3'UTR was used as a comparison in bandshift assays as measure of specificity of RNA-protein interactions between the MyHC 3'UTRs and muscle protein extracts. The size of the rat growth hormone PCR product was 94 base pairs. Figure 4.2 shows the rat growth hormone 3'UTR sequence. The sequence has 40-50% homology with each MyHC 3'UTR if aligned in pairwise, but only 14% similarity in sequence to the MyHC 3'UTRs overall.

5' cgaagcttGCACACACTGGTGTCTCTGCGGCACTCCCCCG

TTACCCCCCTGTACTCTGGCAACTGCCACCCCTACACT

TTGTCCtctagacccg 3'

Figure 4.2. Sequence of the rat growth hormone amplified from cloned cDNA. The growth hormone sequence is shown in bold type. Primers used in PCR are underlined with lowercase letters referring to primer linker sequence.

4.3.3. Synthesis of 3'UTR transcripts.

Figure 4.3. shows the 'run-off' transcripts synthesized from DNA templates. The size of the full length transcripts inclusive of vector sequence and restriction site sequences from primers was calculated to be 124 bp for type I, 144 bp for 2a, 124 bp for 2b, 123 bp for 2x, and 103 bp for rGH. For the labelled probes it can be seen that as well as full-length transcripts, there were shorter transcripts, necessitating the purification of the correct size products from polyacrylamide gel. The presence of smaller RNA products could be attributed to premature termination due to the presence of sequences with in the DNA template that resemble termination sequences recognized by the phage RNA polymerase, or the lower purity of the template DNA which can also affect the termination rate of the polymerase. The yield of radiolabelled transcripts was generally between 100 and 300 ng, and labelling was routinely performed to a specific activity of 1-5x 10^8 cpm/μg. Cold

probes were purified by column chromatography in order to remove the high concentration of unincorporated nucleotides. The yield of unlabelled transcripts was generally in the range of 300-700ng/µl.

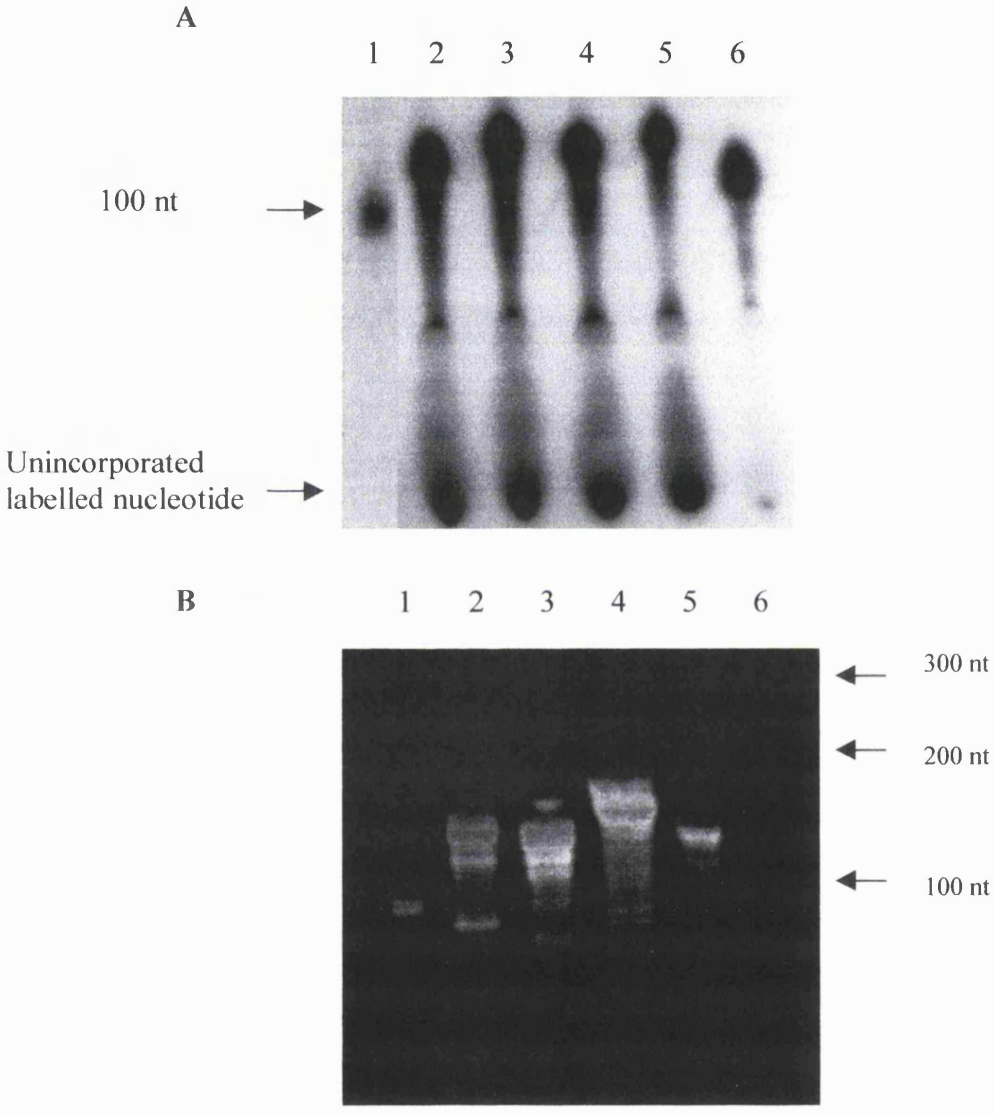


Figure 4.3. Run-off transcripts of MyHC and rat growth hormone 3'UTRs. Panel A shows labelled transcripts: Lane 1 marker (nt stands for nucleotides); lane 2 MyHC slow type I (beta), lane 3 MyHC 2a; lane 4 MyHC 2b; lane 5 MyHC 2x; lane 6 rat growth hormone. Panel B shows the unlabelled transcripts electrophoresed on a 5% polyacrylamide gel stained with ethidium bromide: Lane 1 rat growth hormone; lane 2 MyHC 2x; lane 3 MyHC 2b; lane 4 MyHC 2a; lane 5 MyHC beta; lane 6 marker.

4.3.4. Optimization of bandshift assay conditions and complex formation.

In analysing RNA-protein interactions, it was important to determine both the specificity and the stability of a given complex. Therefore, it was necessary to initially optimize the conditions for binding and eliminate non-specific binding. Cytoplasmic and nuclear extracts in particular have relatively high concentrations of non-specific binding proteins such as heterogenous nuclear ribonucleoprotein particles (hnRNPs). These bind readily to small RNAs giving rise to non-specific complexes that cause smearing in gel retardation assays. The non-specific binding can be minimized by use of a competitor such as tRNA. Yeast tRNA was included at 1000-fold excess in all the binding reactions. Non-specific complex formation was further reduced by the addition of heparin and RNase T1 enzyme. The results of experiments to establish the conditions for optimal binding of protein to RNA are illustrated with radiolabelled 2a 3'UTR probe. Figure 4.4. shows the retardation of probe with increasing amounts of protein. 5, 10, 20, 30 and 40μg of protein were incubated with the probe at room temperature for 30 minutes. It can be seen that increased binding is observed as the amount of protein incubated with the probe is increased. An increased amount of smearing was seen with 30 and 40 μg of protein. In subsequent bandshift assays, approximately 20μg of protein was used as good retardation is seen with this amount of protein and it allowed the total reaction volume to be kept manageable for gel electrophoresis.

Figure 4.5. shows the results of an RNase T1 digestion titration of the probe alone. This was carried out to determine how much RNase T1 was necessary to digest away probe. Complete digestion of the probe was observed with 1 unit of RNase T1 and therefore this amount of RNase T1 enzyme was used to digest away any probe not protected by bound protein in optimized bandshift assays.

Heparin was added to the binding reactions as a competitor to destabilize non-specific interactions between the RNA and protein. A heparin titration was carried out initially to determine the optimal amount required to remove non-specific binding (figure 4.6.) Addition of more than 100 ng heparin disrupted binding of protein to probe and therefore less binding and retardation was observed at heparin concentrations greater than 100 ng. Furthermore, the addition of heparin also showed

the presence of additional complexes. 100 ng of heparin was used in optimized bandshift assays to analyse MyHC 3'UTR-protein interactions.

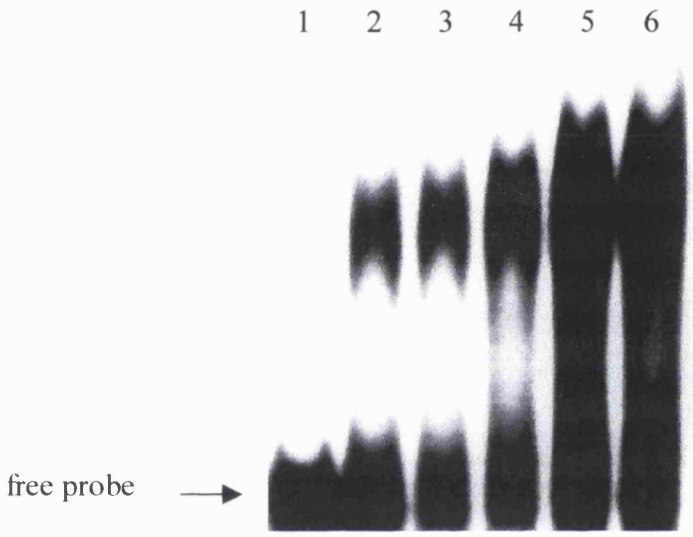


Figure 4.4. Bandshift assay with MyHC 2a 3'UTR probe, showing the formation of RNA-protein complexes with muscle cytoplasmic extract. ^{32}P labelled transcript was incubated with different amounts of extract at room temperature and analysed by electrophoresis in non-denaturing polyacrylamide gel: Lane 1 free probe, lanes 2, 3, 4, 5, and 6 probe incubated with 5, 10, 20, 30, and 40 μg of protein, respectively.

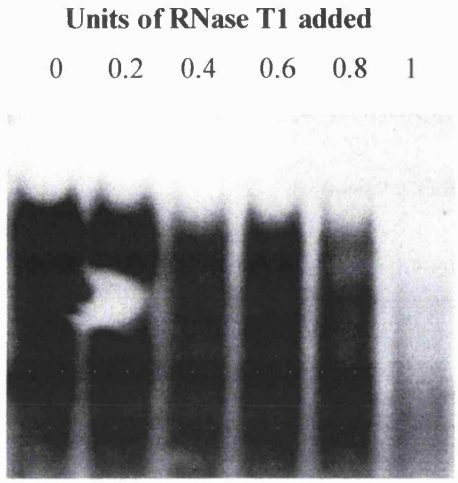


Figure 4.5. The effects of the addition of RNase T1 to probe alone. Reactions were incubated at room temperature for 20 minutes in binding buffer and RNase T1 enzyme as indicated. Reactions were then electrophoresed in a non-denaturing polyacrylamide gel.

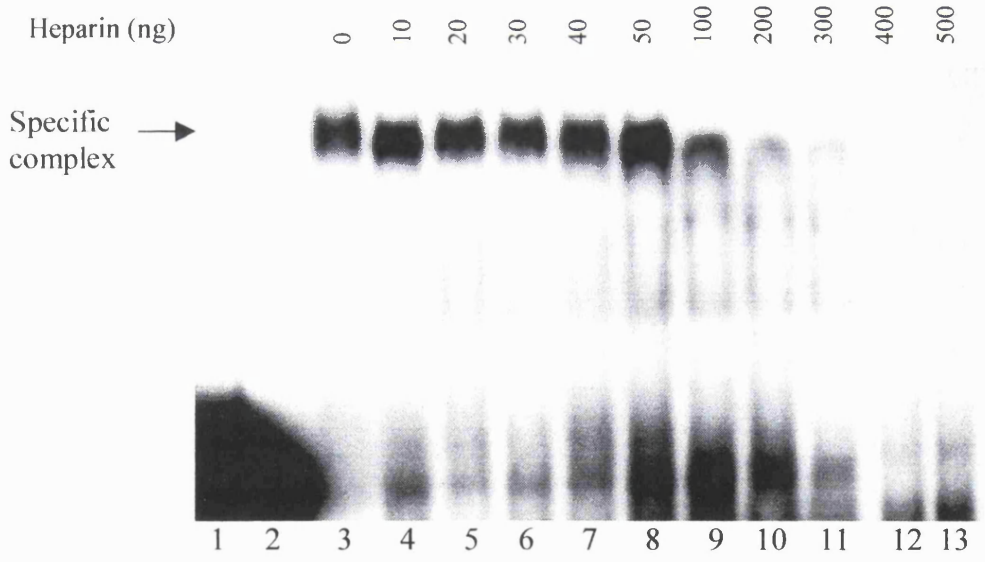


Figure 4.6. The effect of heparin on the RNA-protein complex formation. 20 µg of protein was incubated with probe as described in section 4.2.9. 1 unit of RNase T1 was added to each reaction except the reaction containing free probe. Heparin was then added in increasing amounts as a non-specific competitor. Lane 1 free probe; lane 2 probe plus RNase T1; lanes 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, and 13 probe plus extract, with addition of 1 unit of RNase T1 enzyme and heparin as indicated.

4.3.5. Bandshift assays with muscle protein extracts.

Under the conditions defined by experiments to determine optimum binding, bandshift assays were performed with slow type 1 (β-cardiac), 2a, 2b, 2x and rGH 3'UTR probes. The results are shown in figure 4.7. Free probe was electrophoresed with binding reactions in each assay in order to detect shifted complexes. It can be seen that incubation of probe with protein resulted in some smearing of the retarded complex through the gel, particularly for the type 1 and 2x probes, even when tRNA was included as a non-specific competitor. This is probably due to non-specific RNA-protein interactions. Small, fainter bands were observed that migrated faster, in addition to a larger, slower migrating complex. These smaller bands and the smearing were not diminished by the addition of 100ng of heparin, but were virtually eliminated by addition of RNase T1. The addition of heparin and RNase T1 gave similar results to that observed with the addition of RNase T1 alone, although it can be seen that for MyHC 2a, and to a lesser extent for MyHC 2x, that addition of heparin and RNase T1 changes the mobility of the complex compared to addition of RNase T1 alone. It is not possible to distinguish from these results whether the complex is a single large band or more than one complex migrating closely together. The type of complexes observed for the MyHC probes was not seen with radiolabelled rGH 3'UTR probe, although two shifted complexes, one with a similar mobility, and the other a faster mobility, than the MyHC 3'UTR complexes were detected. This would suggest that the complexes detected with MyHC 3'UTR probes are specific.

A binding reaction with extracts pre-treated with proteinase K confirmed that the retarded complexes consisted of RNA and protein (see figure 4.8). No complexes were observed when probes were incubated with extracts treated with proteinase K.

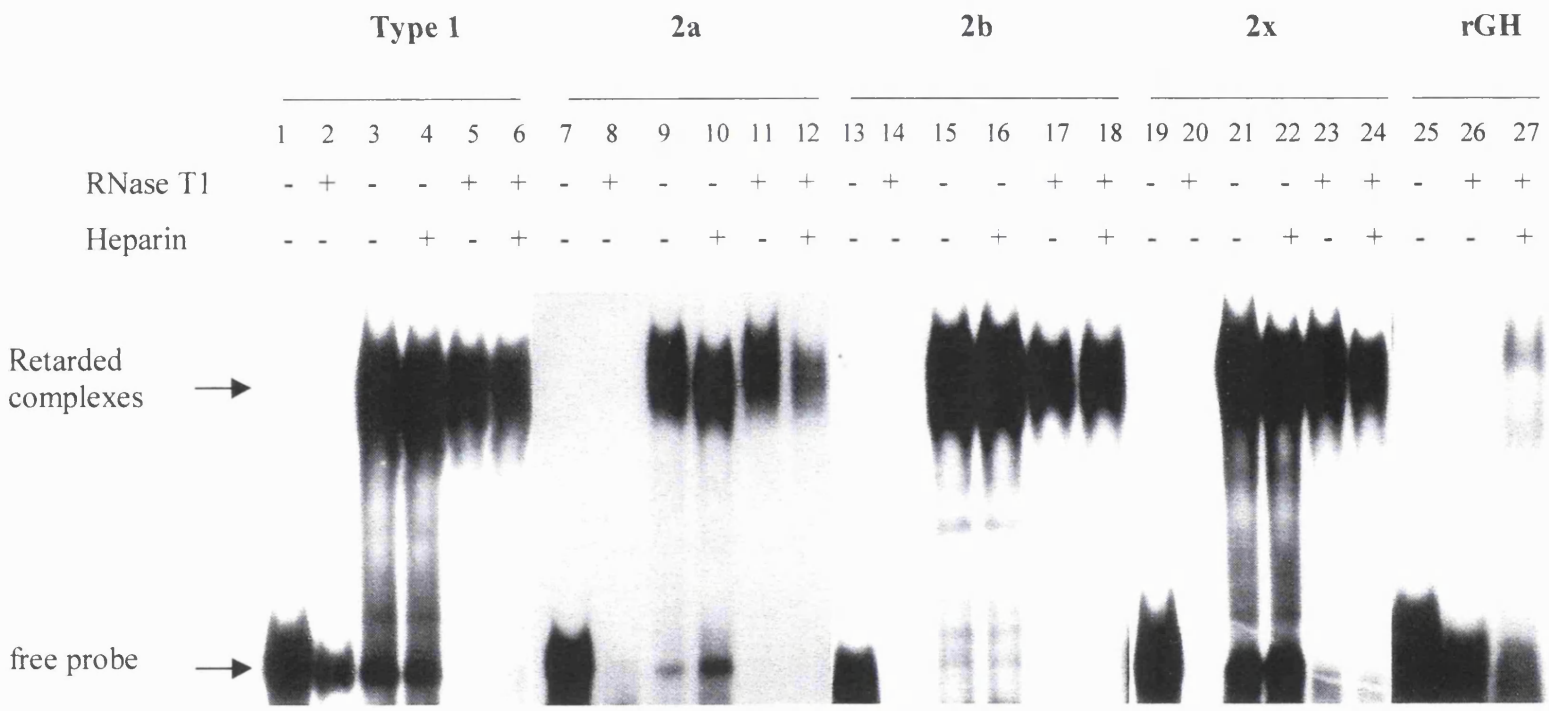


Figure 4.7. Formation of specific complexes between Type 1, 2a, 2b, 2x MyHC 3'UTR RNA probes. Lanes 1, 7, 13, 19, and 25 contain free probes. RNA samples in lanes 2, 8, 14, 20, and 26 were treated with RNase T1 alone at room temperature for 20 minutes. All other lanes, probes ($1x10^5$ cpm) were incubated with 20 µg of extract in binding buffer, in the presence of 3 µg yeast tRNA. Lanes 3, 9, 15, and 21 contain probe with extract. Lanes 4, 10, 16, and 22 contain samples that were treated with 100 ng of heparin for 10 minutes at room temperature. Samples in lanes 5, 11, 17, and 23 were treated with RNase T1 (0.1 unit for 20 minutes at room temperature). Samples in lanes 6, 12, 18, 24 and 27 were treated sequentially with 0.1 unit of RNase T1 for 20 minutes and 100 ng of heparin for 10 minutes at room temperature.

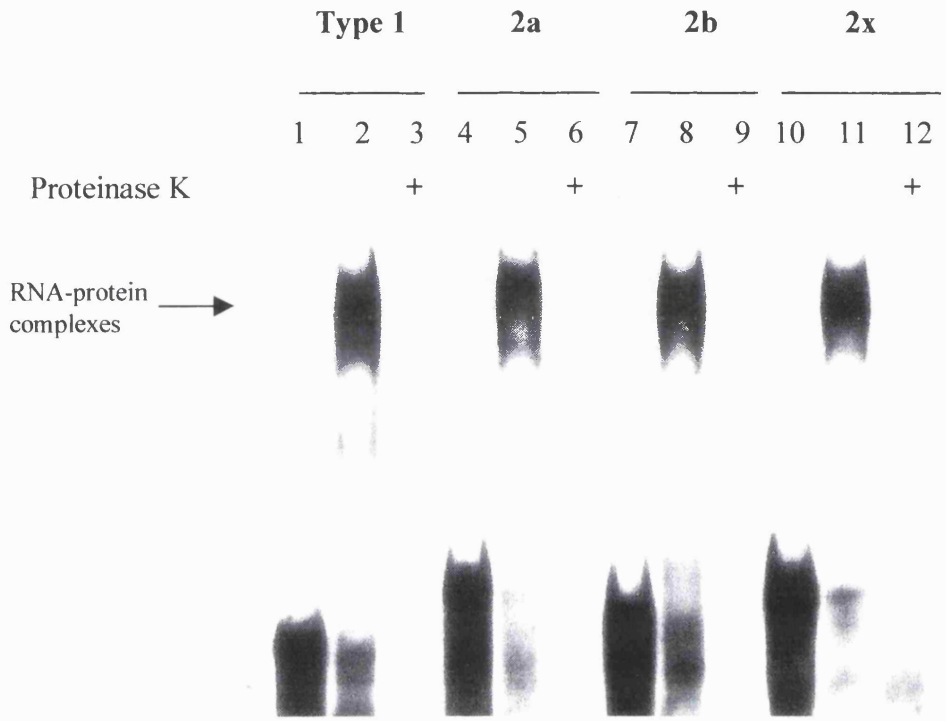


Figure 4.8. Bandshift analyses with pre-treated protein extracts. Lanes 1, 4, 7 and 10 contain free probe. Lanes 2, 5, 8, and 11 probes were incubated with protein and subsequently treated with RNase T1 and heparin. Lanes 3, 6, 9 and 12 probes incubated with protein extracts pre-treated with proteinase K, final concentration 50 µg/ml at 37° C for 15 minutes.

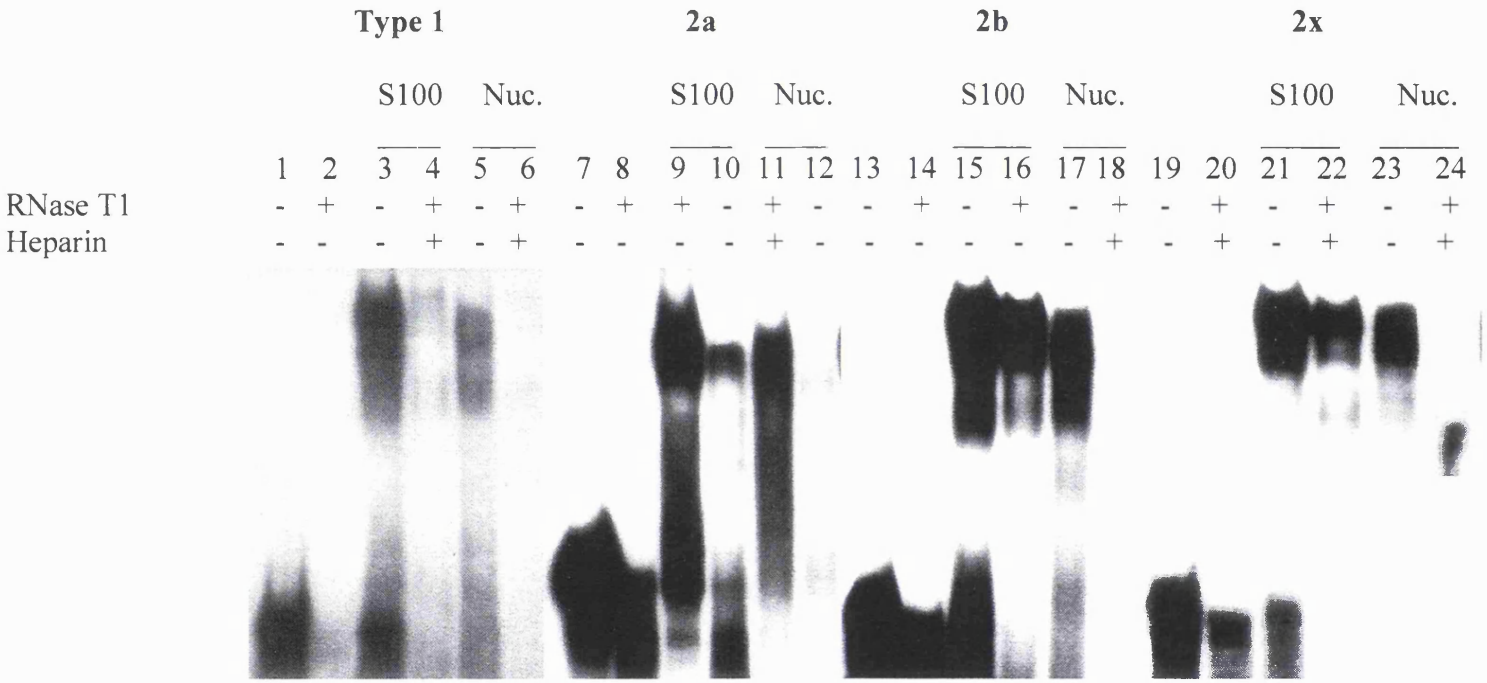


Figure 4.9. Bandshift assay with S100 (cytoplasmic) and nuclear protein extracts from muscle. Probes ($1x10^5$ cpm) were incubated with 20 µg of muscle cytoplasmic or nuclear extract, and subsequently treated with RNase T1 and heparin as indicated. Lanes 1, 7, 13, and 19 contain free probe. Lanes 2, 8,14, and 20 contain probe treated with RNase T1 alone. Nuc. refers to nuclear extract.

4.3.6. Myosin heavy chain 3'UTRs bind muscle cytoplasmic protein.

To characterize the nature of the protein bound to the MyHC 3'UTRs, bandshift assays were performed with S100 and nuclear muscle protein extracts. Figure 4.9. shows the results of these bandshift assays. Although shifted complexes were detected for all of the MyHC 3'UTR probes when incubated with nuclear extracts, addition of heparin and RNase T1 eliminated binding for type 1 and 2b 3'UTRs. For 2a and 2x 3'UTR probes faint bands could still be detected but the migration pattern was different to that observed with S100 extract. From these results, it would seem that MyHC 3'UTRs form complexes with cytoplasmic protein.

4.3.7. Tissue specificity of RNA-protein complex formation.

The tissue specificity of the MyHC 3'UTR RNA-protein complex formation was investigated by performing bandshift assays with protein extracts prepared from mouse cardiac muscle, brain, liver, and kidney tissue in addition to skeletal muscle. The results of these experiments are shown in figures 4.10 and 4.11. Reactions with the MyHC 3'UTR probes were performed with each type of extract alone and with the addition of heparin and RNase T1 to show that any absence of binding in the RNase T1 reactions was not due to degradation of the probe from endogenous RNases but from the lack of specific binding. For the type 1 3'UTR, although some retardation of probe was observed with cardiac, liver and brain extracts, this was hardly detectable when heparin and RNase T1 was added to the binding reactions. Furthermore, these bands showed a different mobility to that seen with muscle extracts. No binding was observed with extracts from kidney. With the 2a 3'UTR retarded complexes with different mobilities to that of muscle extract were detected with cardiac, kidney, liver and brain extracts when incubated with the probe in the absence of heparin and RNase T1. However these complexes were not detectable when binding reactions were performed with the addition of heparin and RNase T1. A similar pattern of complex formation was detectable for the 2b and 2x 3'UTR probes when they were incubated with cardiac, kidney, liver and brain extracts without and with addition of heparin and RNase T1 (figure 4.11). These results suggest that MyHC 3'UTRs form a specific complex with a skeletal muscle protein,

since no binding was detected with extracts from cardiac muscle, kidney, liver or brain.

4.3.8. Competition assays

To confirm that the complexes formed were specific to the MyHCs, competition assays were performed. The results of these are shown in figure 4.12 and 4.13. Muscle protein extracts were pre-incubated with increasing excess of unlabelled 3'UTR, rGH 3'UTR, or polyadenylic acid. For the slow type 1 3'UTR, unlabelled type 1 3'UTR reduced binding of protein to labelled probe at 100-fold excess, with 1000-fold excess competing out labelled probe. Polyadenylic acid (poly A) also competed out binding of labelled type 1 probe at 100-fold excess, but unlabelled rGH 3'UTR did not compete for binding even at 2000-fold excess (figure 4.12). The 2a, 2b and 2x probes were competed out by unlabelled 3'UTR at greater than 1000-, 100- and 1000-fold excess of corresponding unlabelled probe, respectively. The 2a, 2b, and 2x MyHC 3'UTR protein complexes were also competed out by poly A at greater than 2000-fold excess for the 2a probe, and greater than 1000-fold excess for 2b and 2x probes. However, as for the type 1 probe unlabelled rGH 3'UTR up to 2000 fold excess did not compete for protein (figure 4.12 and 4.13.).

Figure 4.14 shows the results of a cross-competition experiment where unlabelled type 1, 2b, and 2x MyHC 3'UTR cRNAs were pre-incubated with muscle protein extracts prior to addition of labelled MyHC 2a probe. It can be seen that MyHC 2a 3'UTR-protein complex formation was competed out by type 1 3' UTR at 2000-fold excess and to a lesser extent by 2b and 2x 3'UTRs at this fold excess.

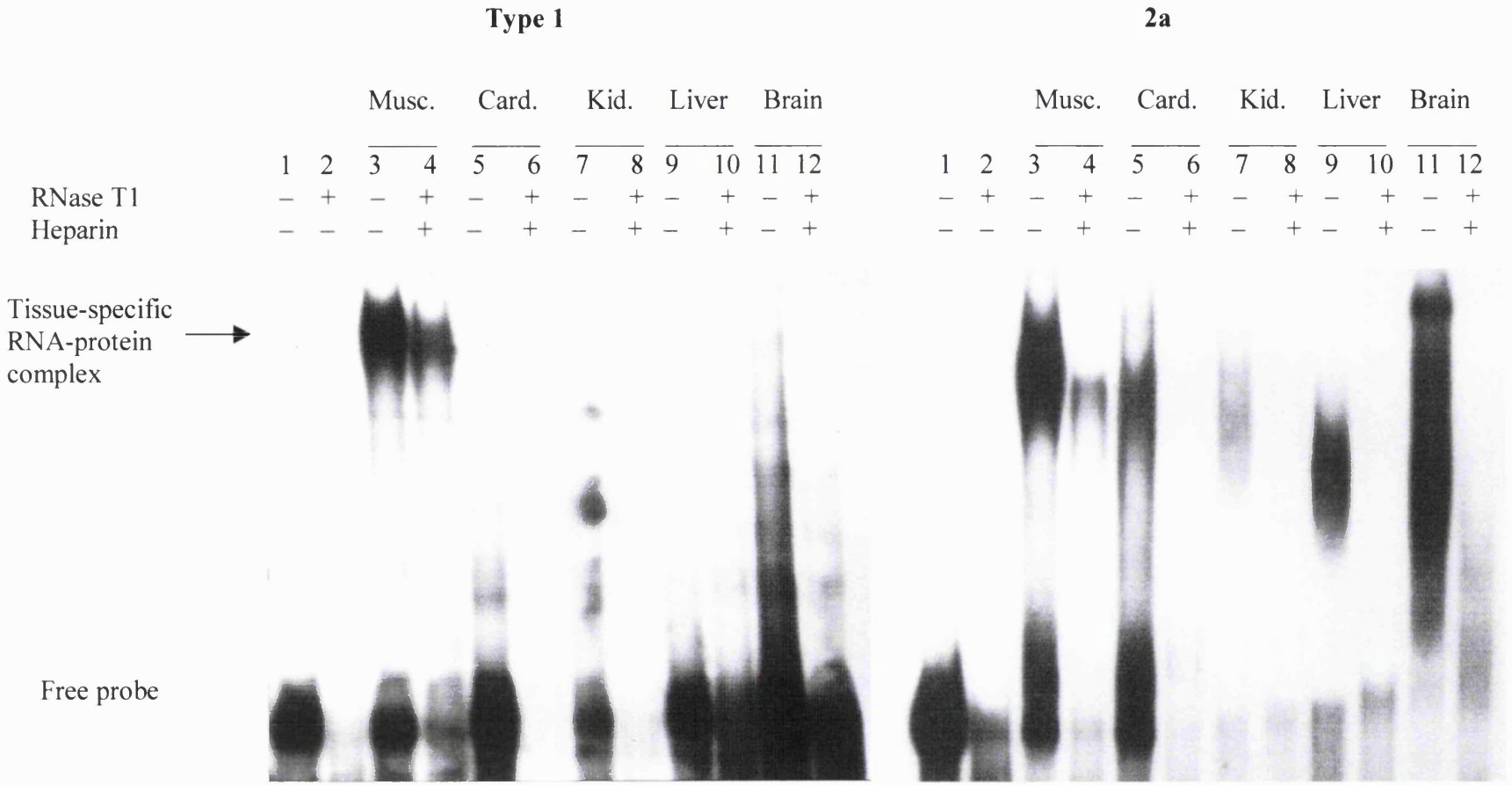


Figure 4.10. RNA bandshift showing MyHC slow type 1 and 2a 3'UTR binding activity with extracts prepared from skeletal muscle (Musc.), cardiac (Card.), kidney (Kid.), liver and brain tissue.

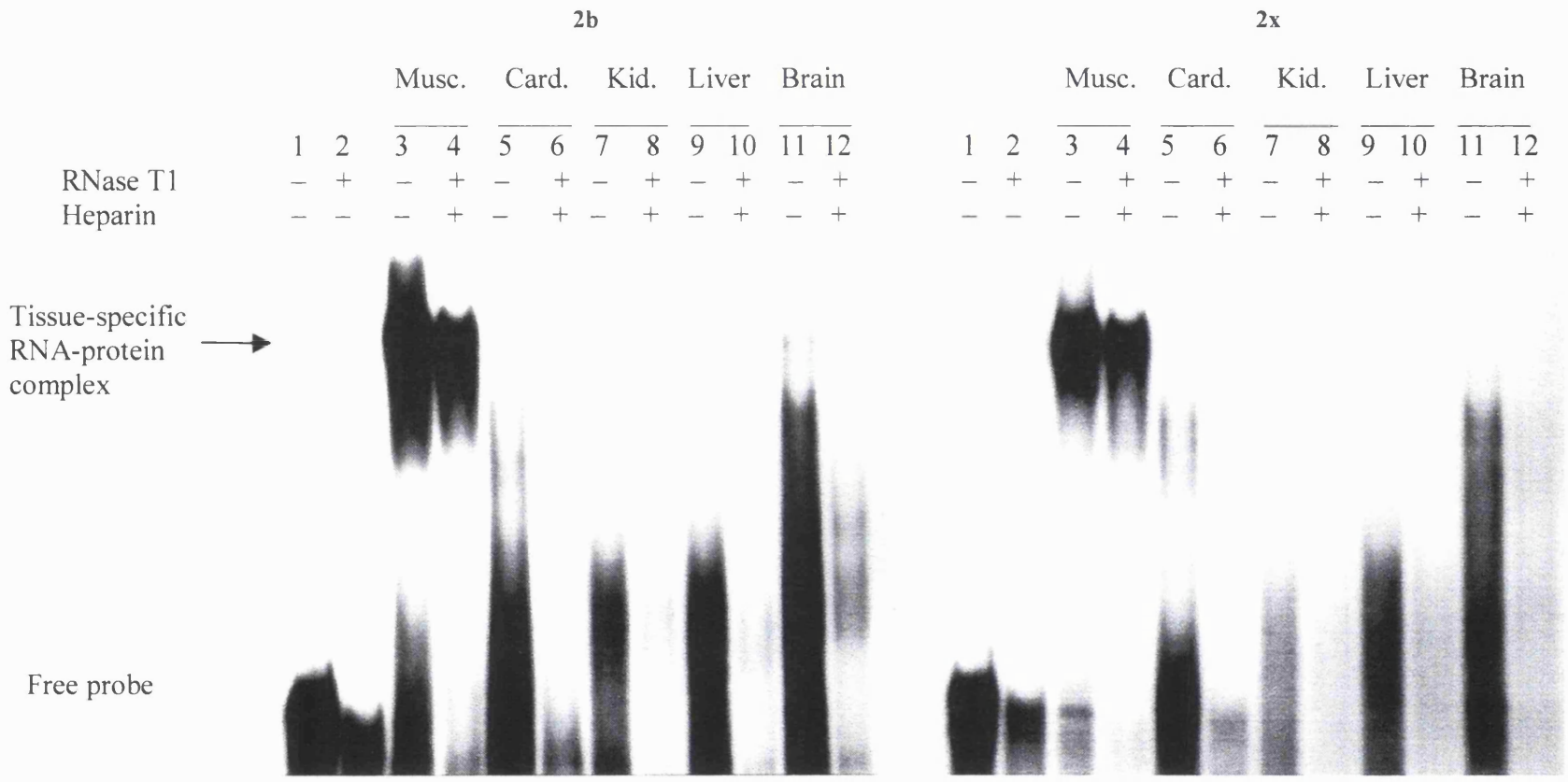


Figure 4.11. RNA bandshift showing MyHC 2b and 2x 3'UTR binding activity with extracts prepared from skeletal muscle (Musc.), cardiac (Card.), kidney (Kid.), liver and brain tissue.

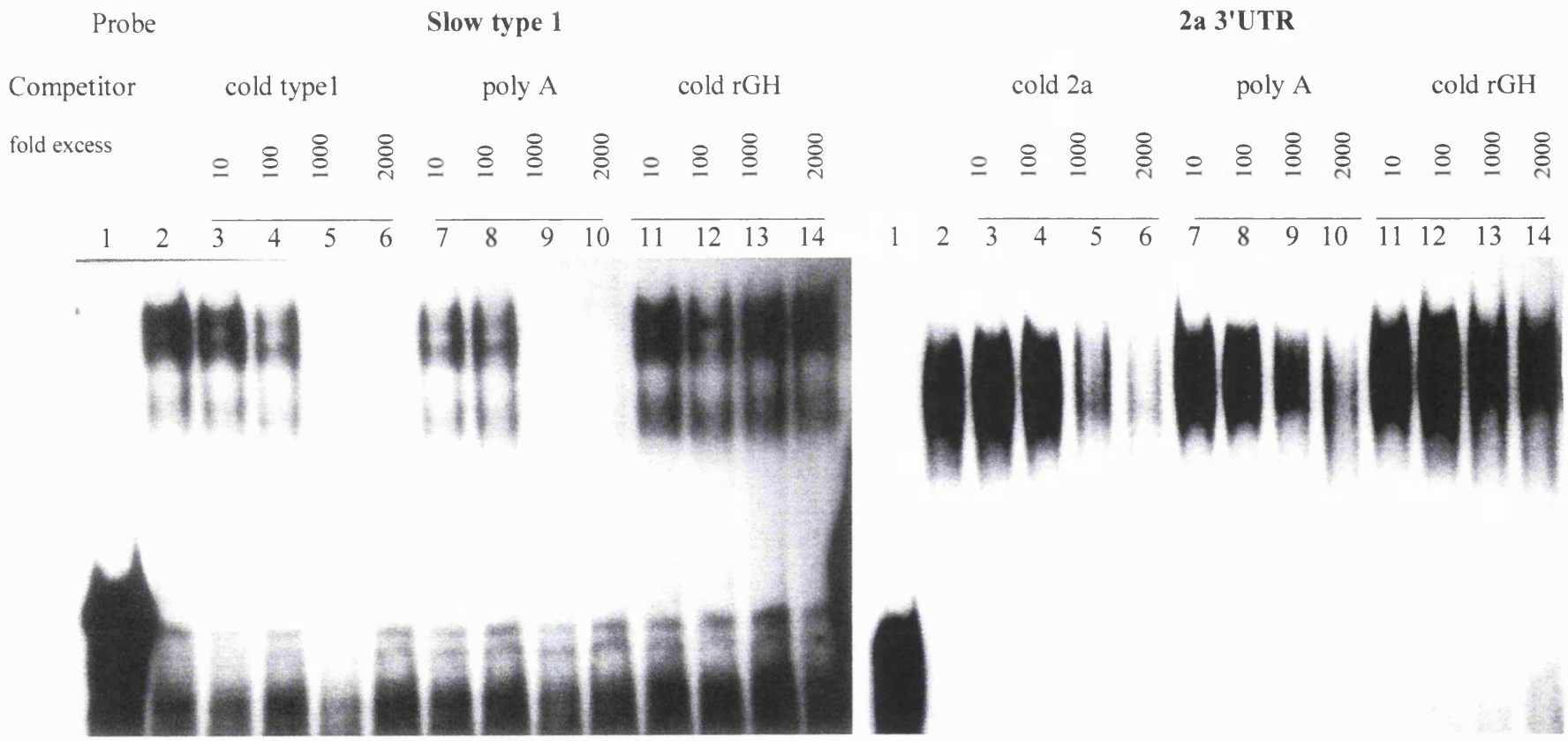


Figure 4.12. Competition assay using specific and non-specific competitors for slow type 1 and fast 2a MyHC 3'UTR probes. Protein extracts were incubated with competitors for 20 minutes at room temperature prior to addition of labelled probe (~ 0.3 ng) and subsequent treatment with RNase T1 as described in the methods. Lane 1 contains free probe; lane 2 probe plus extract; lanes 3, 4, 5, 6 pre-incubation with 10-, 100-, 1000-, and 2000-fold excess of unlabelled RNA probe. Lanes 7, 8, 9, and 10 pre-incubated with 10-, 100-, 1000-, and 2000-fold excess of poly A. Lanes 11, 12, 13, 14 contain protein pre-incubated with 10-, 100-, 1000-, and 2000-fold excess of unlabelled rGH as competitor.

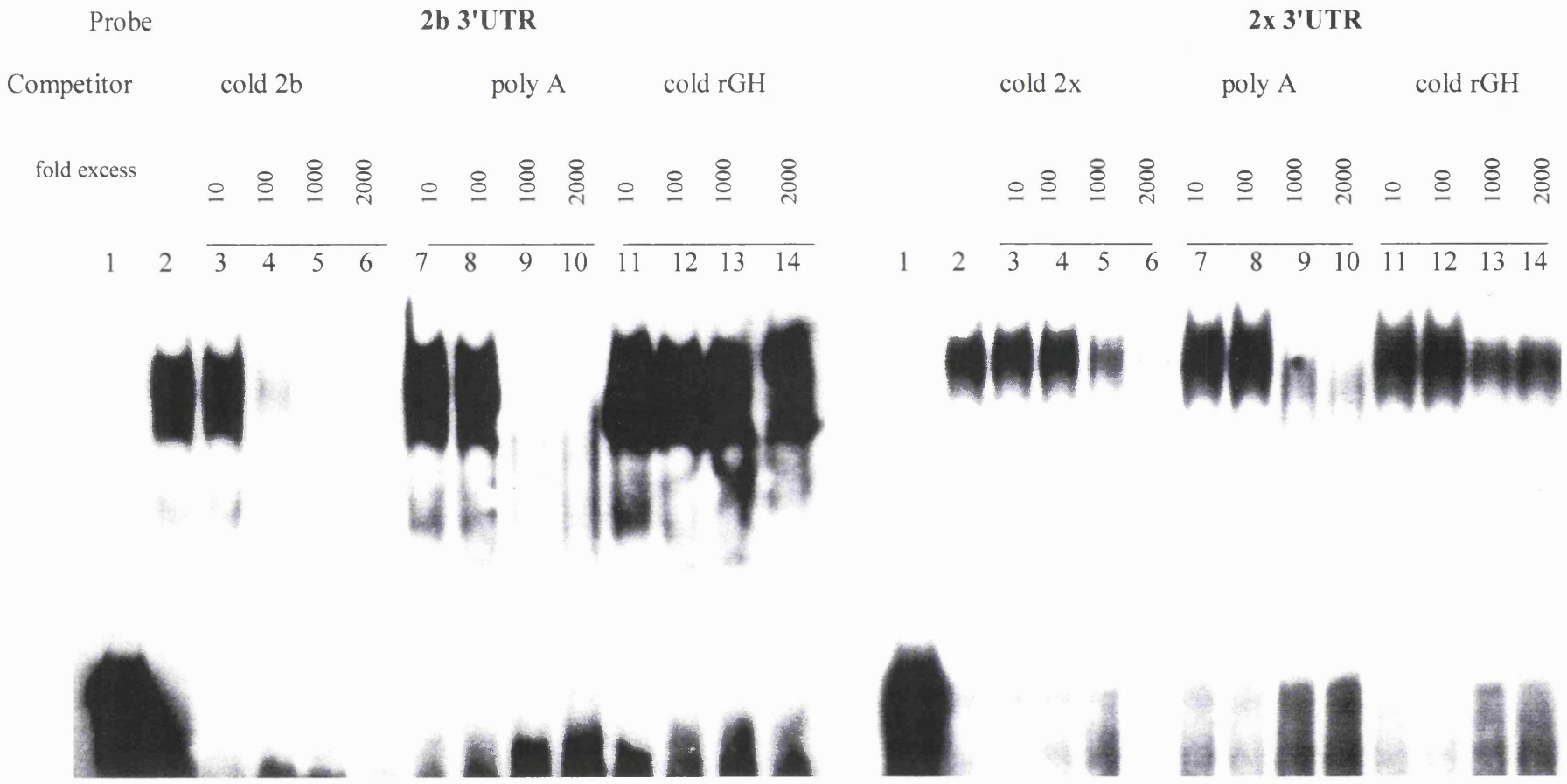


Figure 4.13. Competition assay using specific and non-specific competitors for 2b and 2x MyHC 3'UTR probes. Protein extracts were incubated with competitors for 20 minutes at room temperature prior to addition of labelled probe (~ 0.3 ng) and subsequent treatment with RNase T1 as described in the methods. Lane 1 contains free probe; lane 2 probe plus extract; lanes 3, 4, 5, 6 pre-incubation with 10-, 100-, 1000-, and 2000-fold excess of unlabelled RNA probe. Lanes 7, 8, 9, and 10 pre-incubated with 10-, 100-, 1000-, and 2000-fold excess of poly A. Lanes 11, 12, 13, 14 contain protein pre-incubated with 10-, 100-, 1000-, and 2000-fold excess o unlabelled rGH as competitor.

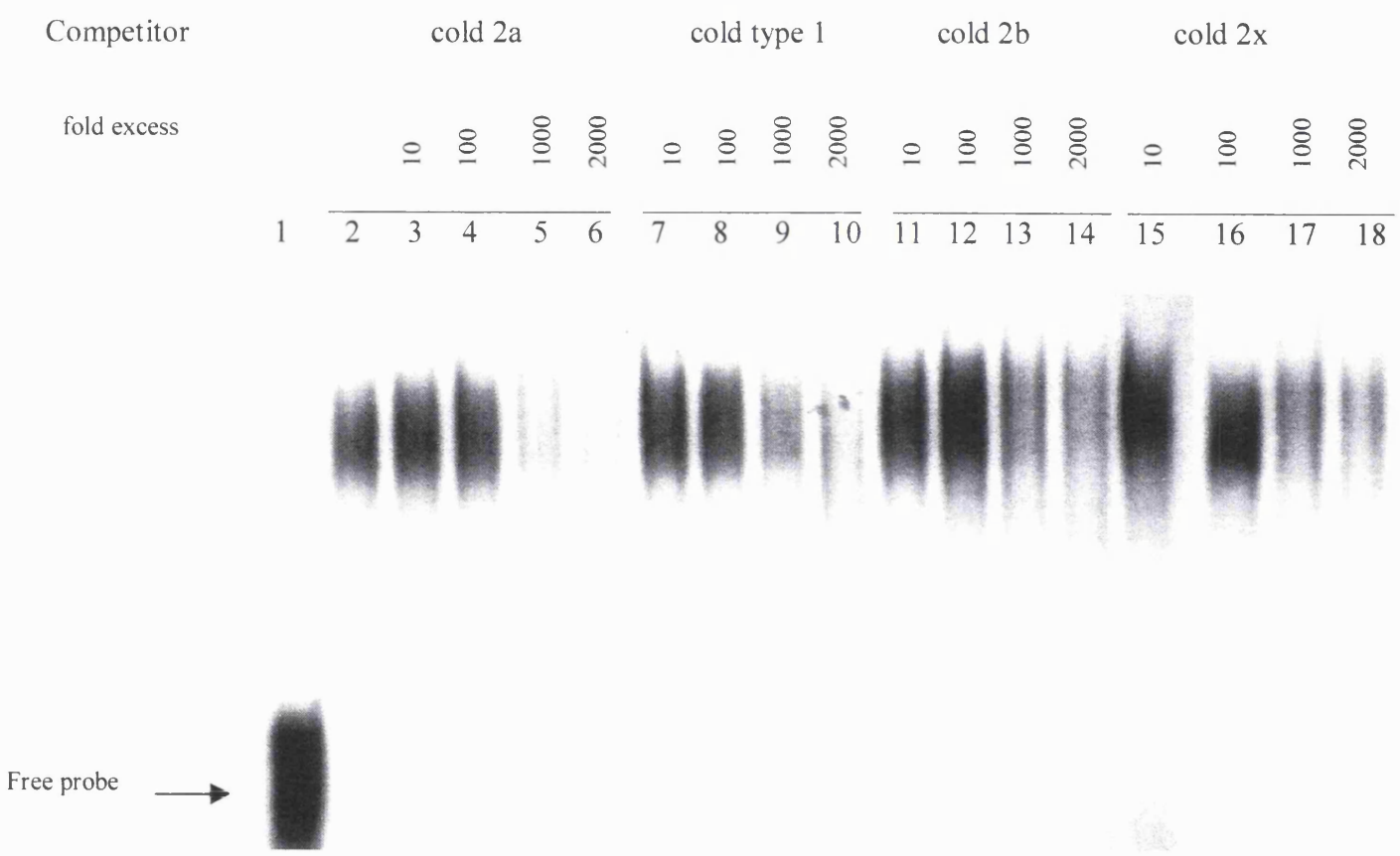


Figure 4.14. Cross-competition analyses using cold MyHC 3'UTRs. Protein extracts were pre-incubated with the indicated cold 3'UTRs for 20 minutes at room temperature before additon of labelled MyHC 2a 3'UTR probe (~0.3ng) and subsequent treatment with heparin and RNase T1. Lane 1 contains free probe; lane 2 probe plus extract; lanes 3, 4, 5, 6 pre-incubation with 10-, 100-, 1000-, and 2000-fold excess of unlabelled 2a 3'UTR. Lanes 7, 8, 9, and 10 pre-incubated with 10-, 100-, 1000-, and 2000-fold excess of unlabelled type 1 3'UTR. Lanes 11, 12, 13, 14 contain protein pre-incubated with 10-, 100-, 1000-, and 2000-fold excess of unlabelled 2b 3'UTR. Lanes 15, 16, 17, and 18 contain protein incubated with unlabelled 2x 3'UTR.

4.4. Discussion.

In this chapter, the binding of muscle protein to the slow type 1, fast 2a, 2b, and 2x MyHC 3'UTRs was investigated by electrophoretic mobility assays. Although the transfection studies with hybrid reporter constructs in chapter 3 did not conclusively demonstrate a regulatory function for the MyHC 3'UTRs in gene expression and mRNA stability, this did not preclude RNA-protein interactions. There are many illustrations of RNA-protein interactions between the 3'UTRs and specific proteins, and several cytoplasmic proteins have been shown to interact with the 3'UTRs of different muscle genes. For example, Ross et al, 1997, have demonstrated the binding of several proteins to the β-actin mRNA zipcode, which functions to localize β-actin mRNA. Similarly, the 3'UTR of vimentin mRNA, which is also involved in localization, also binds a protein (Zehner et al, 1997). Therefore it was hypothesized that the MyHC 3'UTRs were potential targets for RNA-protein interactions. RNA-protein complexes were detected with type 1, fast 2a, 2b, and 2x MyHC 3'UTR probes. Initial bandshift assays were performed to define the optimal binding conditions. The optimal conditions for binding were similar to published protocols (see Fajardo *et al*, 1994; Maatta & Pettinen, 1993). The main difference appeared to be in the concentration of heparin added to remove non-specific protein binding. Whereas most protocols have used 5mg/ml heparin, more than 6μg/ml abolished protein binding to the MyHC 3'UTRs. This might suggest that the complexes formed are not as stable as those exposed to higher concentrations of heparin, although 6μg/ml heparin was the optimal concentration required to detect mouse catalase gene, *CAS*-1, 3'UTR RNA-protein complexes (Reimer & Singh, 1996).

The use of rat growth hormone 3'UTR as a comparison in bandshifts might be thought to be a random choice and using another muscle RNA 3'UTR might be considered to have been more appropriate. However, it can be argued that another muscle 3'UTR could have formed a similar complex to the MyHC 3'UTRs, and the aim here was to demonstrate that the pattern of complex formation was specific to the MyHC 3'UTRs. The pattern of complex formation in initial bandshift

experiments suggests that the RNA-protein complexes formed with MyHC 3'UTR probes are specific since the complexes formed with the rat growth hormone 3'UTR probe differ in the banding pattern and extent of binding. The specificity of the MyHC 3'UTR RNA-protein interactions was further illustrated in competition assays where protein binding was eliminated by excess unlabelled MyHC 3'UTR, but not unlabelled rat growth hormone 3'UTR even when present at 2000-fold excess. Polyadenylic acid was, however, able to compete with labelled 3'UTRs for protein. This suggests that RNA-protein interactions either involve an A rich region within the 3'UTR, or that the binding protein can also recognize and bind poly A sequences. Competition with other polynucleotides such as poly C and poly U were not performed.

The conserved motif CAAAATGTGA has four A residues within it and could be a target for protein binding, as suggested by competition with poly A. However, the absolute requirement for this motif is questionable since complex formation was detected with slow type 1 3'UTR which does not contain such a motif. Furthermore, complexes were also detected with antisense cRNA probes (results not shown). It would be intriguing to see if poly U is able to compete with antisense 3'UTRs probes for protein. In a preliminary cross-competition assay unlabelled type 1, 2b and 2x 3'UTRs were found to be able to compete with labelled 2a 3'UTR for protein. Whereas unlabelled 2a competed out labelled 2a probe at over 1000-fold excess, unlabelled type 1, 2b and 2x were required at a minimum 2000-fold to compete out labelled 2a probe. Since some cross-competition is observed, it is likely that RNA-protein interactions between the MyHC 3'UTRs and muscle extracts depend on structural elements within the secondary structure rather than just specific sequences, and the conserved CAAAATGTGA element may not even be a target for protein binding. Computer-based predictions of secondary structures showed that the MyHC 3'UTRs formed stem-loop structures (see chapter 2, section 2.3.3.). The bulges and loops could be potential targets for muscle protein binding (Roy *et al*, 1990; Calnan *et al*, 1991). Although the CAAAATGTGA motif lies within a loop region, it may be involved more in protein recognition by inducing bends within helices in the stem regions, which is thought to be a mechanism for recognition of RNA by proteins.

The degree of bending depends on the number of bulged nucleotides and the identity of the unpaired bases. This bending could help to correctly position the protein relative to other factors. It has been suggested that the absence of rigid stem-loop structures favours the recognition of unpaired nucleotides by proteins, while stem regions may help to position the protein correctly by non-sequence specific contacts with the backbone.

The tissue specificity of RNA-protein complexes was illustrated by bandshift assays performed with protein extracts from, cardiac muscle, kidney, liver and brain tissue. Since no complex formation was detected with extracts form tissues other than skeletal muscle, it can be concluded that the protein(s) interacting with the MyHC 3'UTRs are skeletal muscle specific. It is worth noting that the slow type 1 3'UTR whose sequence is identical to the 3'UTR of beta-cardiac MyHC expressed in the heart did not form a complex with cardiac extract as might have been expected. Therefore, the RNA-protein interactions of the slow type 1 3'UTR are skeletal muscle specific under the conditions used in bandshift analyses. From results of bandshifts performed with skeletal muscle S100 extract, which is essentially the cytoplasmic fraction, and what was assumed to be the nuclear extract, it would appear that the MyHC 3'UTRs binds cytoplasmic extracts. However, possible leakage of nuclear proteins during the homogenization of muscle tissue cannot be disregarded which could lead to absence of protein binding. Some faint bands could be detected in reactions with nuclear extracts (see figure 4.9.).

The RNA-protein complexes were resolved by non-denaturing polyacrylamide gel electrophoresis, and therefore, it was not possible to determine the molecular weight of the binding protein(s) from the bandshift gels. The methods used to analyse the MyHC 3'UTR-protein interactions further are described in the following chapter.

In conclusion, the adult skeletal MyHC 3'UTRs interact in a tissue specific manner with a cytoplasmic protein factor expressed in skeletal muscle. Although the binding activity is specific for the MyHC 3'UTRs, this factor can also bind polyadenylic acid.

Chapter 5

Characterization of MyHC 3'UTR mRNA-Protein Interactions

5.1. Introduction.

Interactions involving mRNA and protein are critical in the post-transcriptional regulation of many eukaryotic genes such as in splicing (see review by Keene & Query, 1991), mRNA stability(Jacobson & Peltz, 1996). Moreover, proteins have been shown to interact with regulatory elements within the 3'UTR of many eukaryotic mRNAs. The mRNA-protein interactions have been shown to be dependent on nucleotide sequence and secondary structure. For example specific proteins have been shown to bind to AU-rich regions in the 3'UTR of many cytokine mRNAs, which are responsible for their rapid degradation (Bickel *et al*, 1992; Stephens et al, 1992; Brewer, 1991; Shaw & Kamen, 1986). The general mechanism here is thought to involve some kind of destabilization, leading to increased susceptibility to exonuclease digestion (Alberta *et al*, 1994). Although the role for 3'UTR mRNA sequence elements and interacting proteins is recognized in several systems, the molecular mechanisms behind their involvement in post-transcriptional regulation remains poorly understood. Progress is being made in this direction with the characterization of the 3'UTR-binding proteins. The majority have been characterized by determining their molecular weight using ultraviolet cross-linking (UV cross-linking). Proteins associated with RNA and DNA can be covalently linked by ultraviolet irradiation (Pelle & Murphy, 1993), and the cross-linked reactions electrophoresed on SDS polyacrylamide gels to elucidate the molecular weight of bound proteins. Cross-linking experiments have been used effectively to determine the molecular weights of several proteins associated with the 3'UTR. For example Ross et al, 1997, have characterized and cloned a 68 kDa protein that binds to the β-actin 3'UTR zipcode which has been shown to be responsible for localization of chimeric mRNA. Other proteins characterized by molecular weight include the AU-rich binding proteins and the poly CU-binding protein associated with the 3'UTR of α-globin mRNA. From similarities in the molecular weights of some of these proteins, it has been suggested that they may have more than one function. For example the 48 kDa cytoplasmic poly CU-binding protein was found to stabilize the α-globin mRNA. A 48 kDa protein has also been found to interact with LOX mRNA

3'UTR to repress translation. Therefore, the same protein may serve not only to stabilize the mRNA, but also to repress translation.

In the previous chapter, it was shown that the slow type 1, fast 2a, 2b and 2x skeletal MyHC 3'UTRs bound cytoplasmic protein prepared from skeletal muscle. The work described in this chapter aimed to characterize the protein binding to the MyHC 3'UTRs. The characterization of RNA protein interactions is more difficult than that for DNA binding proteins, but the interactions can be analysed by measuring the effect of protein modification on binding or, vice versa, looking at the effect of modifying the RNA on protein binding. Since the identity of the protein is unknown the RNA was modified by sequential deletion. Bandshift assays were performed with deleted RNAs and the protein characterized by UV cross-linking.

5.2. Materials and methods.

5.2.1. Elution of bound probe.

In order to determine the whether full-length MyHC 3'UTR was required for binding muscle cytoplasmic protein, the bound probes were eluted from the shifted protein/probe complex. This was carried out by performing a binding reaction and running a preparative bandshift gel. After the gel had finished running, the plates were carefully dismantled and the gel and plate were wrapped in Saran wrap. The gel was placed in an X-ray cassette with film and exposed for 2 hours. The developed film was aligned with the gel to identify the shifted complexes. These were excised from the gel and placed in RNase-free tubes. 700µl of gel elution buffer (0.5 M ammonium acetate, 1 mM EDTA, 0.2% SDS) were added to each gel slice, and they were incubated at 37° C overnight. The tubes were centrifuged at high speed to pellet the acrylamide and the eluate carefully removed into fresh tubes. Eluted RNA probes were precipitated with 0.2 M NaCl, 40 µg yeast tRNA, and 2.5 volumes of ice-cold absolute ethanol at -20° C overnight.

The precipitated RNAs were pelleted by centrifugation at 14000 rpm, 4° C, for 1 hour, then washed in cold 70% ethanol (DEPC). The probes were briefly air-dried and resuspended in RNase-free water. Both eluted probes and original full-length probes were then electrophoresed on a polyacrylamide gel. An equal volume of gel loading buffer (95% formamide, 0.5 mM EDTA pH 8.0, 0.025% SDS, 0.025% xylene cyanol, 0.025% bromophenol blue) was added to each sample, and the tubes were heated at 95° C for 5 minutes. Heated reactions were loaded, together with labelled RNA markers, onto a 0.75 mm thick 5% polyacrylamide, 8 M urea gel and run at 250 V for 2 hours in 1x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA). The gel was transferred to Whatman paper, dried at 80° C on a vacuum gel drier and exposed to X-ray film overnight.

5.2.2. Generation of nested deletions

In order to determine the minimum sequence required for formation of RNA-protein complex, the MyHC 3'UTRs were sequentially deleted and used to produce

probes by run-off transcription. The probes were then used in bandshift assays. A double-stranded nested deletion kit from Pharmacia was used to generate a set of deletions for the MyHC 3' UTRs. The kit uses the enzyme *exonuclease* III (*Exo* III) which catalyzes the stepwise removal of 5' mononucleotides from the recessed or blunt 3'-hydroxyl termini of double stranded DNA in a controlled digestion reaction (Weiss, 1976). Protruding 3' ends are resistant to *exonuclease* activity (Rogers and Weiss, 1980), allowing the construction of unidirectional deletions without the need for further subcloning of target DNA (Henikoff, 1984; Henikoff, 1987).

To create unidirectional deletions of the plasmids pGEM4Z -β, pGEM4Z-2a, pGEM4Z-2b, and pGEM4Z-2x, they were digested with *Xba* I restriction enzyme to generate a nuclease sensitive 3' end and with Sac I to generate a protruding nuclease resistant end in a double digestion reaction (see figure 5.1).

10 μg of plasmid DNA was digested with *Xba* I and *Sac* I in restriction buffer consisting of 25 mM Tris-acetate, pH 7.8, 100 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, in a total volume of 100 μl such that the final concentration of the plasmid DNA was 0.1 μg/μl. Digests were incubated at 37° C overnight. An aliquot of each of the reactions was electrophoresed on an agarose gel to check the digestion. After digestion was complete, the DNA was heated for 10 minutes at 70° C to inactivate the enzymes.

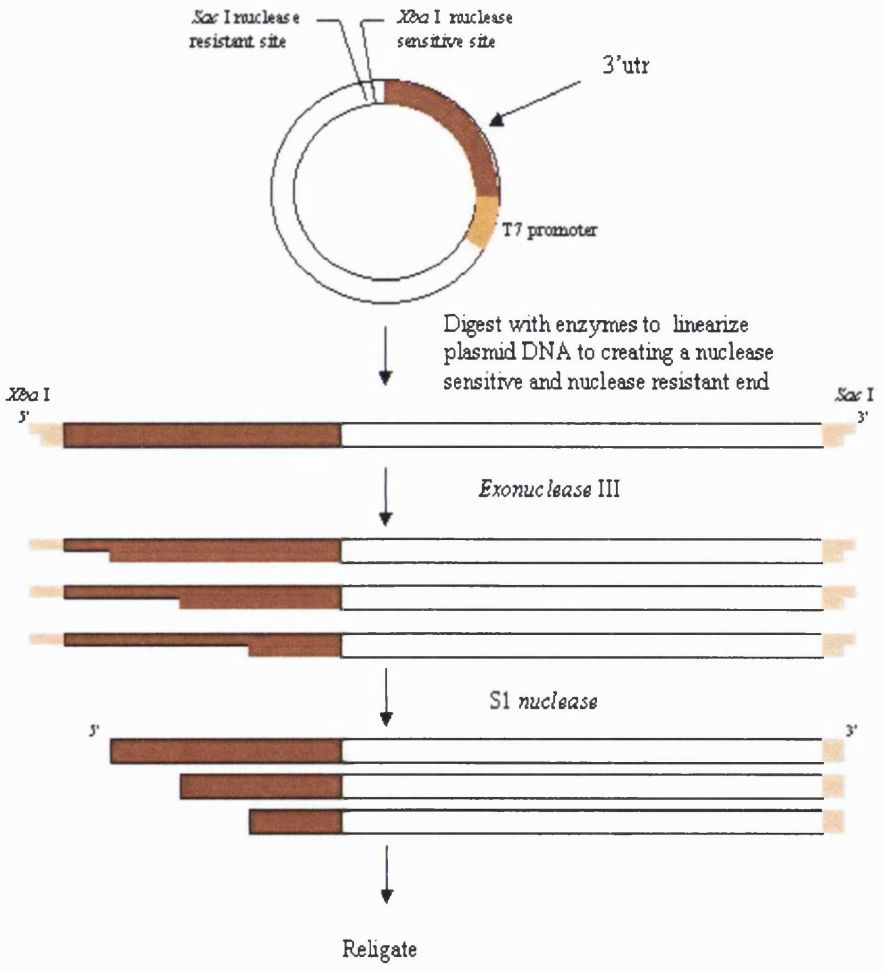


Figure 5.1. Schematic diagram showing the steps involved in the generation of nested deletions.

An S1 nuclease/S1 buffer mixture was prepared by mixing 33 µl of S1 buffer (150 mM potassium acetate, pH 4.6, 1.25 M NaCl, 5 mM $ZnSO_4$, 25% glycerol) with 66 µl distilled water and 1 µl S1 nuclease. 3 µl of this mix was pipetted into each of the 20 Eppendorf tubes, which were placed on ice ready for transfer of timed samples from digestion reactions. 24µl of 2x *Exo* III buffer was prepared by mixing 8µl of Exo III buffer (400 mM Tris-HCl, pH 8, 4 mM $MgCl_2$) with 4µl of 0.3 M NaCl, and 12µl of distilled water. 20µl of this 2x Exo III buffer was mixed with 20µl (2µg) of the double-digested DNA. The mixture was equilibrated at 30° C for 2-3 minutes. A 2µl aliquot was removed for the time-point zero control, and placed in the first of the 20 time point tubes containing 3µl of the S1 nuclease/S1 buffer, mixing well. This tube was kept to one side on ice. 1µl of *exonuclease* III was added to the DNA and mixed gently. The reaction was incubated at 30° C and then every 30 seconds, a 2µl sample was removed. This was mixed immediately and thoroughly with 3µl of the S1 nuclease/buffer mix previously prepared. All of these tubes were kept on ice until all the timed samples had been removed from the *exonuclease* reaction. These were then incubated simultaneously at room temperature for 30 minutes. 1µl of S1 stop solution (303 mM Tris base, 50 mM EDTA) was subsequently added to each sample, which were then incubated at 65° C for 10 minutes.

Each of the timed samples was analysed by electrophoresing 3µl of the reaction on an agarose gel. The remaining 3µl were recircularized by ligation. A bulk ligation mix was prepared as follows; 85µl 5x ligase mix (T4 ligase, 0.13 units/µl, 300 mM Tris-HCl, pH7.6, 50 mM $MgCl_2$, 5 mM ATP, 0.5 mM spermidine, 25 mM DTT, 50% glycerol) 85µl 25% PEG, and 195µl distilled water. 17µl of this mix was added to the 3µl left from the timed samples. They were gently mixed and incubated for 2 hours at room temperature.

After analysis of the extent of deletion by gel electrophoresis, the ligations that showed the required extent of deletion were picked for transformation. 50µl of competent DHα cells was added to these ligated samples. After mixing gently, the transformation reactions were placed on ice for 30 minutes. Then the reactions were heat shocked at 37° C for 30 seconds, followed by 2 minutes chilling on ice. 450µl of LB medium was added and the reactions incubated with shaking for 1 hour at 37° C.

100µl of the transformation reaction was plated out on LB agar plates containing 50µg/ml ampicillin, which were then incubated at 37° C overnight. The following day several colonies were picked and grown in small scale cultures for plasmid preparation using Wizard Mini Prep Kit (Promega). Plasmid preparations were analysed by restriction digestion with *Hin*d III and *Eco*R I enzymes which flanked the insert. Clones that contained inserts were further analysed for the extent of deletion by dideoxy chain termination sequencing.

5.2.3. Synthesis of deleted probes and electrophoresis of complexes.

RNA probes corresponding to deleted MyHC 3'UTR sequences were synthesized by transcription using T7 polymerase as described in chapter 4 section 4.2.6. The transcription reactions were electrophoresed on a 5% polyacrylamide, 8M urea gel. Full-length products were visualised by autoradiography, excised from the gel and placed in RNase-free tubes. The labelled RNAs were eluted from the gel slice overnight in 500 µl of elution buffer (0.5 M ammonium acetate, 1 mM EDTA, 0.2% SDS) at 37° C. The tubes were centrifuged at high speed to pellet the acrylamide and the eluate carefully removed into fresh tubes. RNAs were precipitated with 0.2 M NaCl, 40 µg yeast tRNA, and 2.5 volumes of pure ice-cold ethanol at -20° C overnight. The precipitated RNAs were pelleted by centrifugation at 14000 rpm, 4° C, for 1 hour, then washed in cold 70% ethanol, and resuspended in 30µl of nuclease free water after briefly air-drying.

Electrophoretic mobility shift assays were carried out with the deleted MyHC probes as follows. Labelled RNA deletion probes were heated at 70° C for 5 minutes and then allowed to renature by cooling slowly to room temperature. Renatured RNA probe (~$1x10^5$ cpm per reaction) was incubated, in a 15µl reaction volume, with 10-20µg of cytoplasmic muscle protein extract in binding buffer containing 20 mM HEPES (pH 7.9), 3 mM $MgCl_2$, 40 mM KCl, 2 mM DTT, and 5% glycerol, in the presence of 200 ng/µl yeast tRNA. The reactions were incubated at room temperature for 30 minutes and subsequently treated with 1 unit of RNase T1 (GIBCO BRL) for 20 minutes at room temperature, followed by 6µg/ml heparin (SIGMA) for 10 minutes at room temperature. 5µl of 50% glycerol were added to

each reaction. Samples were then subjected to electrophoresis through a 1.5 mm thick non-denaturing 5% polyacrylamide gel (acrylamide/bisacrylamide ratio 65:1) in 0.5x TBE. Electrophoresis was carried at 250V for 2 hours at room temperature. The gel was then carefully transferred to Whatman paper, and dried at 80° C on a vacuum gel drier. The dried gel was exposed to X-ray film overnight.

5.2.4. UV cross-linking.

UV cross-linking was used to try to characterize the protein(s) bound to the MyHC 3'UTRs. Radiolabelled RNA probe was incubated with protein extracts as previously described. Immediately following incubation, the samples were transferred to a 96 well microtitre plate kept on ice. They were then placed 5 cm below the light source in UV Stratalinker 1800 (Stratagene) and exposed to UV light for increasing lengths of time from 0 to 30 minutes in 10 minute increments. The reactions were then incubated at room temperature for 20 minutes after the addition of RNase T1 (2 units) and then for a further 10 minutes following the addition of heparin (6μg/ml). An equal volume of Laemmli buffer (62.5 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 2% β-mercaptoethanol and 0.001% bromophenol blue dye) was added to each sample, and after boiling for 5 minutes they were subjected to electrophoresis through a 5% stacking gel containing 125 mM Tris-HCl, pH 6.8 and 0.1% SDS, and a 10% resolving gel containing 375 mM Tris-HCl, pH 8.8 and 0.1% SDS (acrylamide to bisacrylamide ratio 19:1). Electrophoresis was carried out in Tris-glycine buffer (25 mM Tris, 192 mM glycine, pH 8.3, 0.1% SDS) for 4 hours at 250V using the Protean II vertical gel system. SDS-PAGE protein markers (Amersham) were electrophoresed alongside for determination molecular weights of cross-linked protein. After electrophoresis, the gel was transferred to Whatman paper and dried. The gel was exposed to X-ray film overnight.

5.2.5. Elution and SDS-PAGE of bound protein.

In order to identify the protein(s) bound to the MyHC 3'UTRs, a preparative bandshift assay was performed and the protein from shifted RNA-protein complexes purified. RNA probe ($1x10^5$ cpm) was incubated with 20-30μg of cytoplasmic

muscle protein extract in a 30µl reaction, and incubated at room temperature for 30 minutes. The reaction was subsequently treated with RNase T1 (2 units) and 6µg/ml heparin as previously described. Reactions were electrophoresed on a 5% non-denaturing polyacrylamide gel for 3 hours at 200 V. After the gel had finished running, the plates were carefully dismantled and the gel and plate were wrapped in Saran wrap. The gel was placed in an X-ray cassette with film and exposed for 2 hours at 4° C. The developed film was aligned with the gel to identify complexed probes and protein, which were excised and placed in clean tubes. Gel slices were incubated at 37° C for 4 hours in 1ml of modified protein elution buffer consisting of 1% Triton-X100, 20 mM HEPES pH 7.6, 1 mM EDTA, 100 mM NaCl, 2 mM DTT, 20µg/ml BSA, 0.1 mM PMSF (Holcik & Liebhaber, 1997). The tubes were then centrifuged at high speed to pellet the acrylamide and the eluate carefully removed into fresh tubes. Eluted protein was precipitated with acetone by adding the eluate drop-wise to ice-cold 80% acetone and placing at -20° C overnight. The tubes were centrifuged at 14000 rpm, 4° C for 1 hour. The resulting pellet was washed with 1 ml of 50% acetone and briefly air-dried. The protein pellets were resuspended in 20µl of Laemmli buffer and boiled for 5 minutes. The heated samples were then loaded onto and SDS-polyacrylamide gel consisting of a 5% stacking gel and a 10% resolving gel as described above. Electrophoresis was carried out for 3 hrs at 100v using the ATTO mini gel system (GRI). After electrophoresis was complete the gel was silver stained as follows to visualize protein bands. The gel was fixed for 1 hour in 40% methanol/10% acetic acid, then washed twice, 30 minutes each time, in 10% ethanol/5% acetic acid. After oxidization for 10 minutes in 0.2% $K_2Cr_2O_7$, the gel was washed 3 times, 10 minutes for each wash, in distilled water. Silver staining was performed for 30 minutes in a 0.1 % (w/v) solution of $AgNO_3$. The gel was washed for 2 minutes in distilled water and developed in 3% (w/v) Na_2CO_3, 0.5% formalin solution. The reaction was stopped by transferring the gel to a 1% acetic acid solution. For protein sequencing, eluted proteins were electrophoresed on SDS-PAGE gel and stained with Coomassie blue (SIGMA). Stained bands were excised and stored at -20° C until protein mapping was performed.

5.2.6. Protein mapping.

Protein mapping reactions were performed at the Ludwig Institute by Dr Nick Totty. Protein bands excised from Coomassie stained SDS-PAGE gels were destained and dehydrated with 50% acetonitrile, 25 mM ammonium bicarbonate and digested overnight with modified trypsin (Promega) in 25 mM bicarbonate. The peptides were further extracted 2 more times with 505 acetonitrile and 55 trifluoroacetic acid. The supernatants were pooled and dried using a speedvac drier. 100µl of HPLC grade water was added and then the peptides were dried again to remove volatile salts. Samples were resuspended in 5µl of water and 0.5µl aliquots were used for MALDI analysis. This was performed using a Perseptive Biosystems Voyager Elite XL instrument with 2,5-dihyroxybenzoic acid as sample matrix. Internal calibration was performed using two well-characterized trypsin autolysis peptide ions, present in all of the samples. Monotypic peptide ions were used to search the Protein Prospector Database (University of California) using the MS-Fit programme.

5.3. Results.

5.3.1. Nearly full-length probe binds to muscle protein.

MyHC 3'UTR probes that were bound to protein and retarded in electrophoretic mobility assays were purified by excising the gel slices containing the shifted complex and eluting from the gel. The eluted probes were subsequently electrophoresed together with full-length MyHC 3'UTR probes on a polyacrylamide gel. Figure 5.2. shows the results of the electrophoresis. The eluted probes were of similar length to the original full-length 3'UTR probes, indicating that the protein bound to virtually the entire 3'UTR. A reason for this could be the secondary structures formed by the MyHC 3'UTRs. They form stem loop structures and the double stranded stem regions would be protected from RNase T1 which cleaves single stranded RNA 5' to G residues. Although the MyHC 3'UTR sequences are very G-rich the majority of these residues lie within stem regions of the secondary structures (see chapter 2). and G residues situated in looped single- stranded regions could be protected by protein, which is thought to bind to these regions.

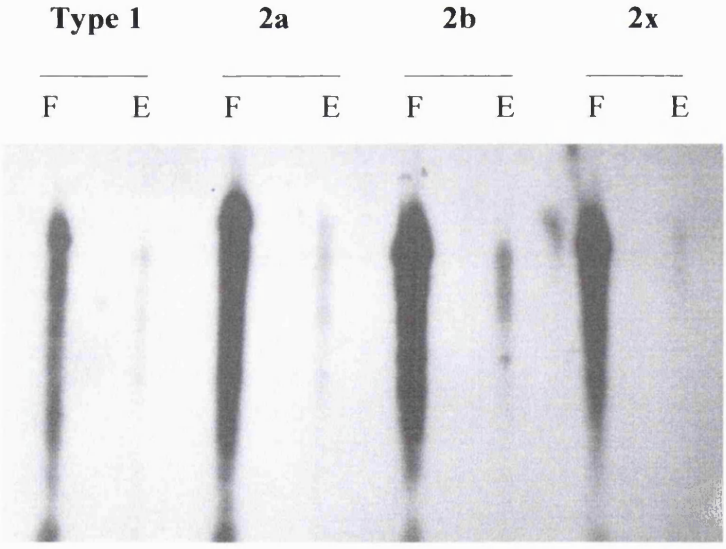


Figure 5.2. Electrophoresis of full-length MyHC 3'UTR probes and probes bound to muscle cytoplasmic protein. The probes were eluted from gel slices containing RNA-protein complexes excised from a bandshift gel. F indicates full-length original probe, and E probe eluted from complex.

5.3.2. RNA bandshift with deleted probes.

To more accurately map the sequence requirements for MyHC 3'UTR mRNA-protein interaction, a set of deletions were produced from pGEM4Z-3'UTR plasmids. Deletions from the 3' end of the UTR were carried out using *exonuclease* III enzyme. ^{32}P-labelled probes were then generated from the deletions by *in vitro* transcription, and were used in gel shift assays.

A set of nested 3'UTR deletions, deleted from the 3' end, was successfully produced for the 2a MyHC 3'UTR. Although 3'UTR deletions were obtained for the slow type 1, 2b and 2x 3'UTRs as determined by agarose gel electrophoresis, sequence analysis indicated that clones from these deletions did not encompass the 3'UTR or were not deleted enough. The small size of the 3'UTRs (~100-125 bp), and the speed of enzymatic reaction, requires the *exonuclease* digestion reactions to be controlled such that a range of deletion can be produced within the 3'UTR. The rates of *exonuclease* III digestion are affected by temperature and salt concentration, with digestion more rapid the higher the temperature and the lower the salt concentration (Hoheisel, 1993; Henikoff, 1987). Reactions performed at 30° C and NaCl at 75 mM were expected to digest about 90 bases per minute. However, the precise rate of *exonuclease* III digestion varies quite significantly from one DNA sequence to the next (Linxweiler & Horz, 1982) and this was apparent when subclones were sequenced. In the digestions performed here aliquots were removed every 30 seconds. The reactions were analysed by agarose gel electrophoresis. Figure 5.3. illustrates the results of the deletion reaction. It was observed that at the last time point (8.5 minutes from start of digestion) that the plasmid had been deleted by almost a kilobase in length. Subsequent sequence analysis of clones showed that after time point 4, that is 2 minutes into the digestion reaction, the 3'UTR had been virtually digested away. With the MyHC 2a 3'UTR, seven nested deletions of 92 (2ADEL5), 83 (2ADEL1E), 72 (2ADEL11), 66 (2ADEL1C), 59 (2ADEL2D), 46 (2ADEL2G), and 28 bp (2ADEL9) in length were produced. The extent of deletion between these fragments varies from 33 bases to as little as six bases. The rate of digestion was rapid during the first few minutes of *exonuclease* activity and this is reflected by the large deletion observed for 2ADEL5 from the full-length 125 bp 2a 3'UTR. Figure 5.4 shows the labelled probes synthesized from the deletions of the 2a

3'UTR. Sequences of the full-length and deleted MyHC 2a 3'UTRs are illustrated in figure 5.6. with conserved motifs highlighted in red.

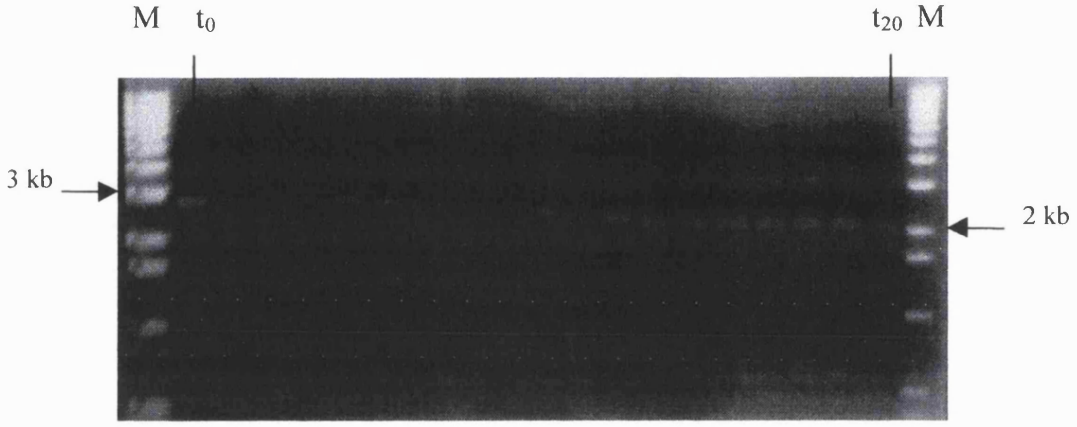


Figure 5.3.. MyHC 2a 3'UTR was digested with *exonuclease* III to produce a set of deletions. They were analysed by 1% agarose gel electrophoresis. M stands for 1 kb ladder. Aliquots were removed from the digestion reaction every 30 seconds. Transformants were produced from ligation of plasmid digested up to time point 4. Clones were sequenced to determine extent of deletion.

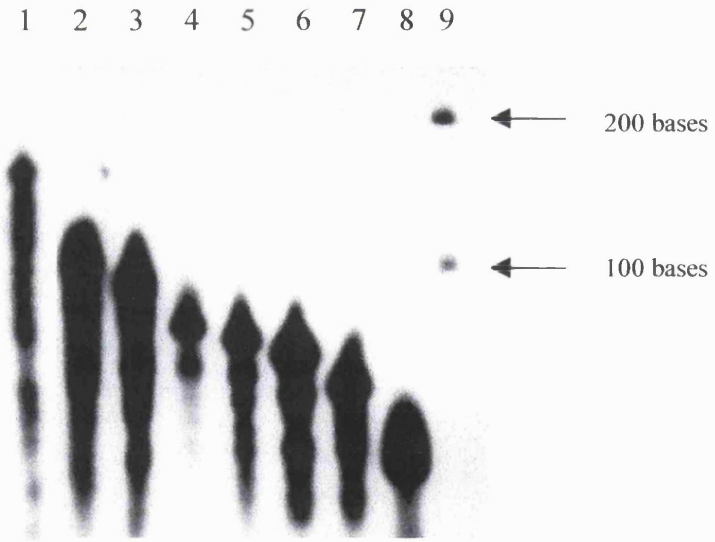


Figure 5.4. Radiolabelled probes synthesized by *in vitro* transcription wild type and deleted MyHC 3'UTR. Seven deletions of 2a MyHC 3'UTR were obtained. 1) 2A full length, 125 nt (nucleotides); 2) 2ADEL5 92 nt; 3) 2ADEL1E 83 nt; 4) 2ADEL11 72 nt; 5) 2ADEL1C 66 nt; 6) 2ADEL2D 59 nt; 7) 2ADEL2G 46 nt; 8) 2ADEL9 28 nt; 9) radiolabelled RNA marker

Figure 5.5. shows RNA-protein complex formation detected with deleted probes. Deletions 2ADEL5 and 2ADEL1E 92 and 83 bp in length respectively form a complex as observed with full-length 2a 3'UTR. 2ADEL5 contains no poly A signal, indicating that the complex is not formed by a protein binding to this region. Loss of complex formation was seen with probes 2ADEL11 and 2ADEL1C, although a faint band could still be detected, suggesting that at least 83 bases of the 2a MyHC 3'UTR are required for protein binding. Both of these probes still contain the CACAAAATGT and CTTTG motifs. A shifted band was detected with deletions 2ADEL2D and 2ADEL2G, which are 59 and 46 nucleotides long, respectively. This is a single band, whereas the complexes observed with 3'UTR of 83 bp or greater are more diffuse and could be two closely migrating complexes. The predicted secondary structures of the deleted 2a 3'UTRs and their energies of folding are shown in figure 5.7. The deletions still form stem loop structures but the energies of folding were less negative, which could mean that they are less stable.

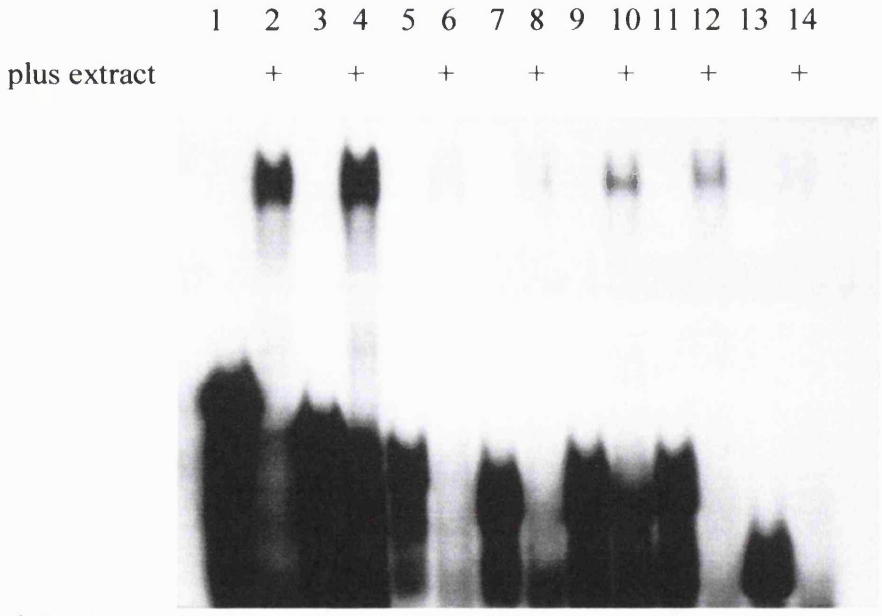


Figure 5.5. EMSA with deletion probes of MyHC 2a 3'UTR. Odd lanes contain free probe and even lanes correspond to probe incubated with protein extract. Lanes 1 and 2, 2ADEL5; lanes 3 and 4, 2ADEL1E; lanes 5 and 6, 2ADEL11; lanes 7 and 8, 2ADEL1C; lanes 9 and 10, 2ADEL2D; lanes 11 and 12, 2ADEL2G; lanes 13 and 14, 2ADEL9.

Full length 2a MyHC 3'UTR sequence

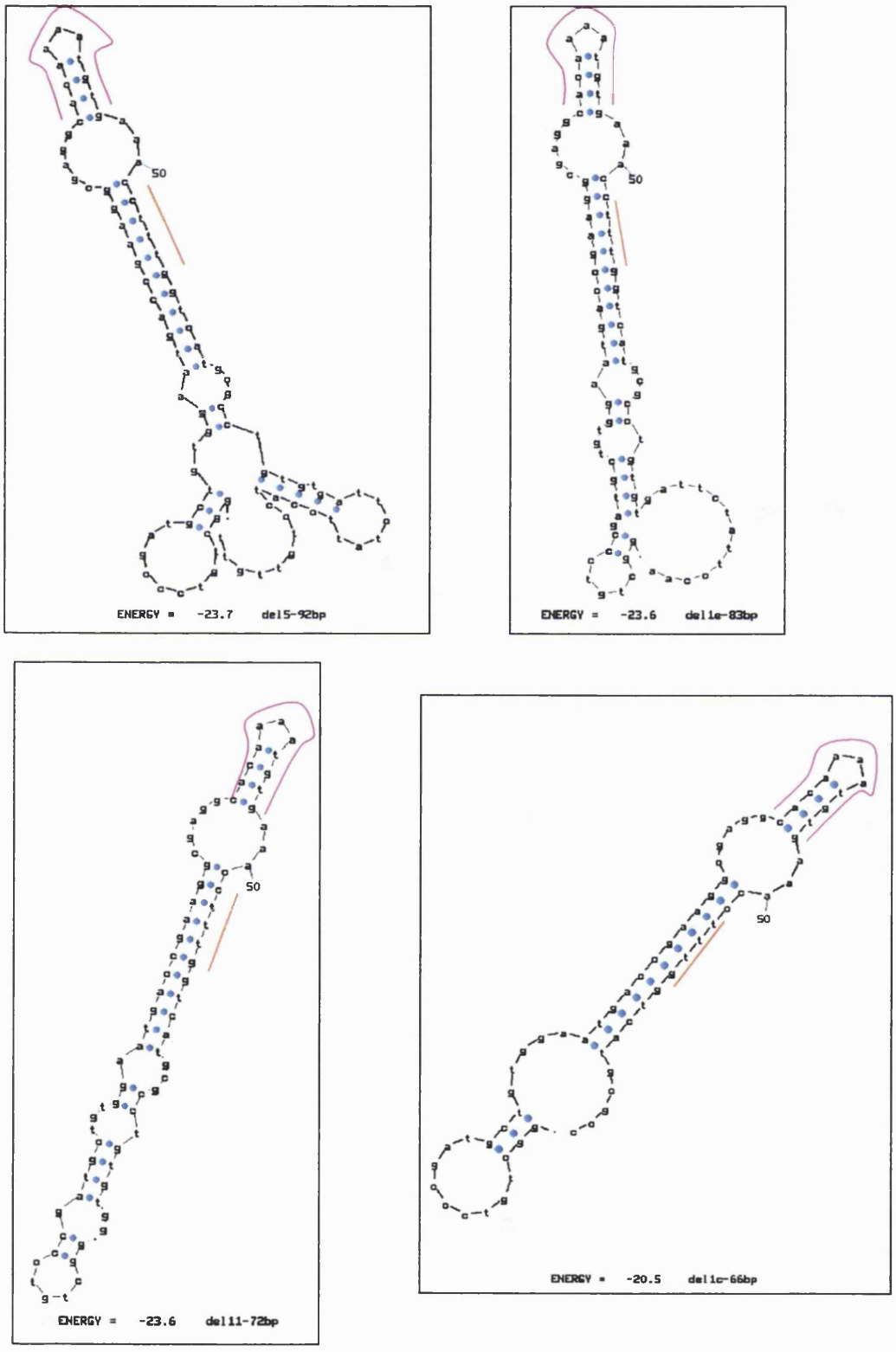
1) GGCTGTCCCGATGCTGTGGAATGACCGAAGGCGAGGCACAAAATGTGAAACCTTTGGTCATGCGCCTGTGTGATTCTATTCCATCCTGTTGTAA
GGAAATAAAAAGCCCAAGTTCCTTTGCAAGC **(125 bp)** +++

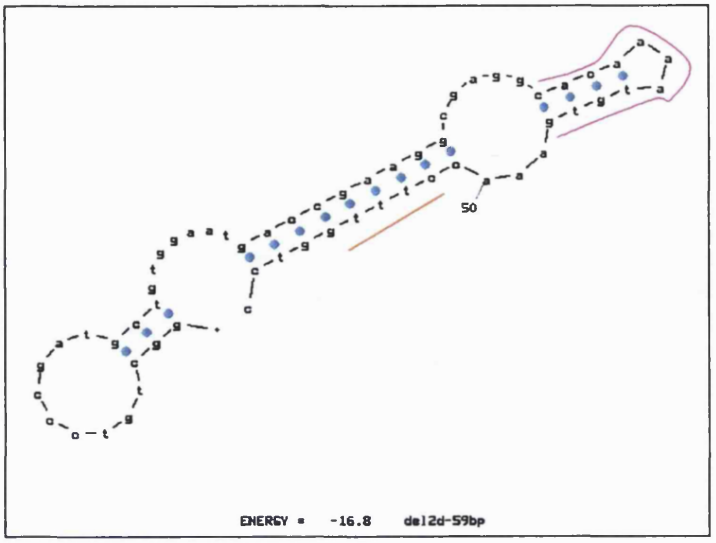
Deletions of 2a MyHC 3'UTR produced and used in gel shift assays.

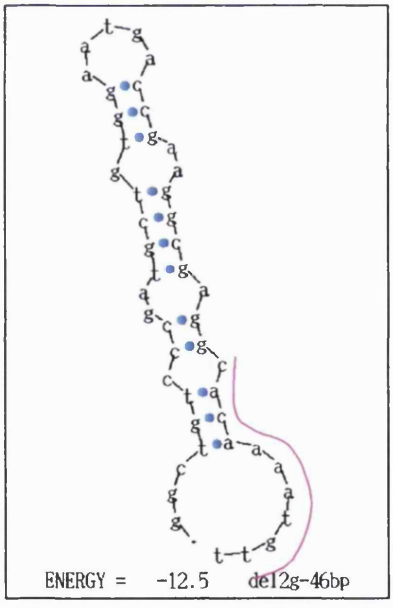
2)	**2ADEL5**	GGCTGTCCCGATGCTGTGGAATGACCGAAGGCGAGGCACAAAATGTGAAACCTTTGGTCATGCGCCTGTGTGATTCTATTCCATCCTGTTGT	**(92 bp)**	+++
3)	**2ADEL1E**	GGCTGTCCCGATGCTGTGGAATGACCGAAGGCGAGGCACAAAATGTGAAACCTTTGGTCATGCGCCTGTGTGATTCTATTCCA	**(83 bp)**	+++
4)	**2ADEL11**	GGCTGTCCCGATGCTGTGGAATGACCGAAGGCGAGGCACAAAATGTGAAACCTTTGGTCATGCGCCTGTGTG	**(72 bp)**	+/
5)	**2ADEL1C**	GGCTGTCCCGATGCTGTGGAATGACCGAAGGCGAGGCACAAAATGTGAAACCTTTGGTCATGCGCC	**(66 bp)**	+/-
6)	**2ADEL2D**	GGCTGTCCCGATGCTGTGGAATGACCGAAGGCGAGGCACAAAATGTGAAACCTTTGGTC	**(59 bp)**	++
7)	**2ADEL2G**	GGCTGTCCCGATGCTGTGGAATGACCGAAGGCGAGGCACAAAATGT	**(46 bp)**	++
8)	**2ADEL9**	GGCTGTCCCGATGCTGTGGAATGACCGA	**(28 bp)**	+/-

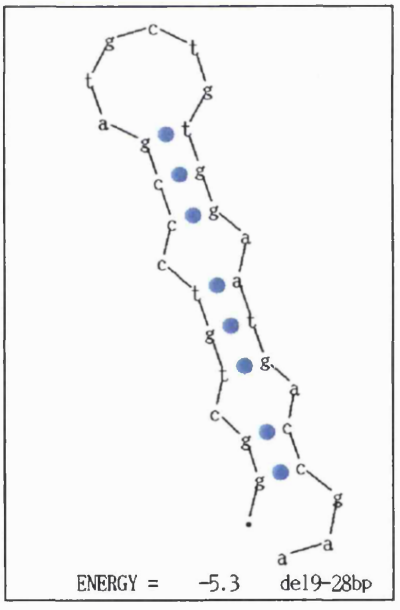
Figure 5.6. Sequences of full-length and deleted MyHC 2a 3'UTRs. The lengths of the sequences and protein binding capacities are indicated.

Figure 5.7. Predicted secondary structures for deleted MyHC 2a 3'UTRs. CACAAAATGT and CTTTG motifs are highlighted by purple and orange lines respectively. Energies of folding (kcal/mol) are indicated




50
ENERGY = -16.8 del2d-59bp


ENERGY = -12.5 del2g-46bp


ENERGY = -5.3 del9-28bp

5.3.4. UV cross-linking analysis.

Ultraviolet irradiation induces covalent cross-linking of DNA or RNA to tightly associated proteins (Dreyfuss et al, 1984; Sommerville, 1992; Mayrand et al, 1981). These complexes have been shown to be resistant to heat, detergent and alkali (Pashev, 1991). Therefore, following UV irradiation, the complexes can be analysed by SDS-PAGE. Initially, UV cross-linking experiments were carried out in order to determine the length of time required for exposure to UV light. MyHC 2a 3'UTR probe was exposed to UV for increasing lengths of time. The reactions were electrophoresed on a SDS gel. This approach produced a lot of background (results not shown) and it was hard to distinguish whether signals were from probe cross-linked to protein. In order to eliminate background, different cross-linking approaches were tried with all four MyHC 3'UTR probes. Attempts to cross-link RNA-protein complexes within excised gel slices after running binding reaction on a native gel were unsuccessful as this required elution of cross-linked complexes from the gel which proved to be difficult. Experiments where the cross-linked reactions were first electrophoresed on a native gel and then excised and treated with RNase T1 prior to running on SDS-PAGE was also unsuccessful, largely due to the protein being retained in the gel slice. An alternative strategy was employed, where binding reactions were exposed to UV light, treated with RNase T1 and heparin and then electrophoresed on a native polyacrylamide gel. Gel slices containing irradiated complexes were excised and incubated in protein elution buffer. The eluted proteins were precipitated with acetone, and subjected to electrophoresis on a SDS polyacrylamide gel. These results are shown in figure 5.8. The first three lanes correspond to MyHC 2a 3'UTR probe and protein binding reaction exposed to UV light for increasing lengths of time starting with 0 minutes. No signal was expected to be seen for the non-irradiated sample, as no cross-linking should have occurred and RNA probe would not be expected to be associated with labelled probe. However, from the results it can be seen that the results of non-irradiation are similar to that of exposure to UV light, even at longer exposures. Six bands were detected cross-linked to MyHC 2a 3'UTR. The three smaller bands were also detected cross-linked to MyHC 2b and 2x 3'UTRs. These reactions did not display the extra bands

as seen with 2a 3'UTR. Three bands for the type 1 3'UTR migrated differently to those of the fast MyHC 3'UTRs with approximate molecular weights of 27, 31 and 34 kDa, suggesting that the proteins complexed to the slow type 1 MyHC 3'UTRs were different. The strongest band detected for 2a, 2b and 2x 3'UTRs was approximately 30 kDa in size.

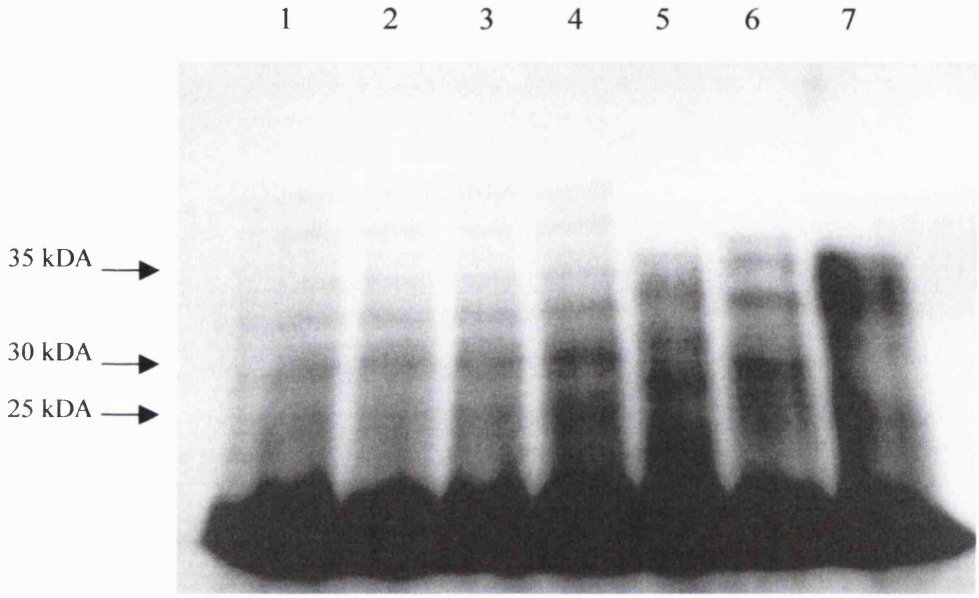


Figure 5.8. Autoradiograph of SDS-PAGE of UV cross-linked MyHC 3'UTR RNA-protein complexes. Lanes 1, 2, 3, and 4 2a probe-protein exposed to UV for 0, 10, 20, and 30 minutes. Lanes 5, 6, and 7, type 1, 2b and 2x probe-protein complexes irradiated for 30 minutes.

5.3.5. Protein elution and analysis.

Acetone precipitation was used to purify the protein(s) bound to the MyHC 3'UTRs. Gel slices containing RNA-protein complexes were excised from a preparative EMSA gel and proteins eluted in buffer. The proteins were then precipitated with acetone and separated on SDS-PAGE. The proteins were visualized by silver staining and the results of this can be seen in figure 5.9. Four distinctive protein bands were detected eluted from complexes formed with type 1, 2a, and 2x 3'UTRs. and three protein bands were seen for 2b 3'UTR. The bands were of about

60, 46, and 30, and 25 kDa in size. Coomassie stained protein bands for 2a 3'UTR complex were excised from the gel and mapped by tryptic digestion. The 60 kDa band was identified as bovine serum albumin. The 46 and 25 kDa proteins were identified as truncated forms of bovine serum albumin, indicating that BSA present in the elution buffer that was used to harvest the proteins from the complex was co-precipitated. The 30 kDa band was identified as aldolase A. Although the proteins bands for type 1, 2b, and 2x 3'UTR complexes were not sequenced, they are probably also aldolase A, since the same banding pattern was observed.

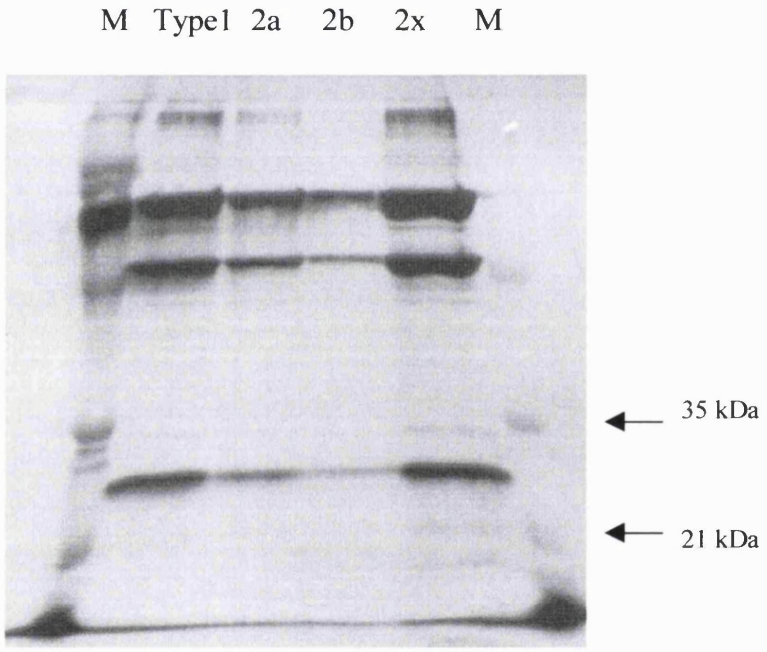


Figure 5.9. SDS-PAGE of proteins eluted from MyHC 3'UTR RNA-protein complexes eluted from EMSA gel, and precipitated with acetone. The gel was silver stained to visualize protein bands. Protein bands were excised from a Coomassie stained gel for protein identification by tryptic digestion.

5.3.5. EMSA with commercial aldolase.

A confirmatory EMSA was performed with commercially available muscle aldolase A enzyme. Rabbit muscle aldolase (SIGMA) was incubated with each of the four MyHC 3'UTR probes, treated with RNase T1 and heparin and electrophoresed

5.3.5. EMSA with commercial aldolase.

A confirmatory EMSA was performed with commercially available muscle aldolase A enzyme. Rabbit muscle aldolase (SIGMA) was incubated with each of the four MyHC 3'UTR probes, treated with RNase T1 and heparin and electrophoresed as described for bandshift assays with muscle protein extract. The results of this are shown in figure 5.10. All four MyHC 3'UTRs bound the commercial aldolase A. The rGH 3'UTR, which was used as a control in bandshift assays also formed a complex with the aldolase A. One reason for this could be the high concentration and purity of the aldolase A compared to that in muscle extracts, and hence the stronger signal form complexes, which were observed even with 2μg of commercial aldolase A (results not shown).

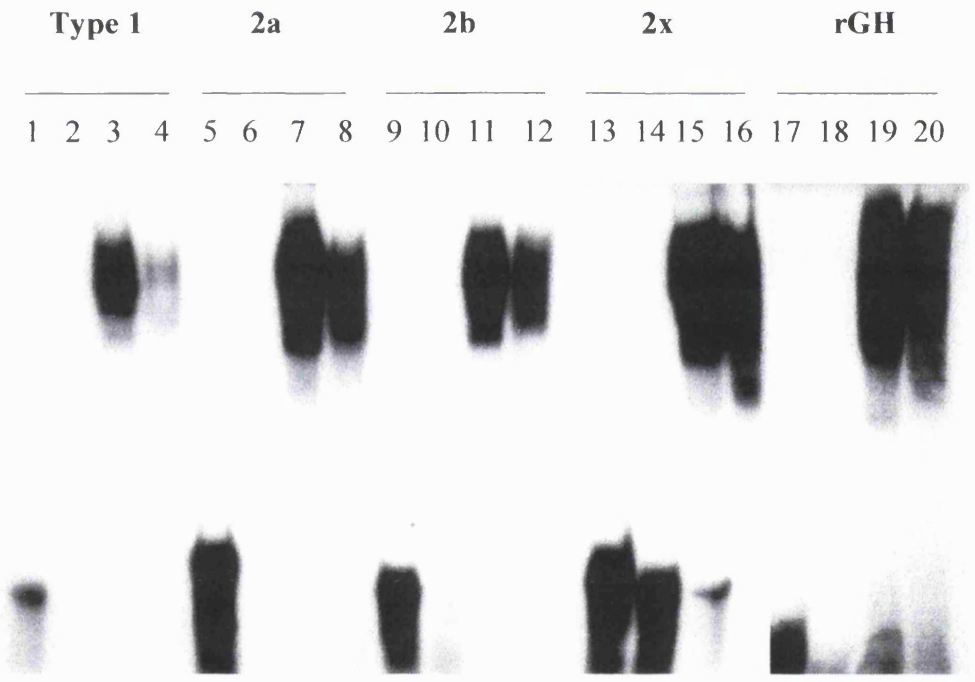


Figure 5.10. Bandshift assay with commercial rabbit skeletal muscle aldolase A. Aldolase was incubated with full-length MyHC 3'UTR probes. Lanes 1, 5, 9, 13 and 17 contain free probe. Lanes 2, 6, 10, 14 and 18, probe digested with RNase T1. Lanes 3, 7, 11, 15 and 19 probe incubated with aldolase A alone. Lanes 4, 8, 12, 16 and 20, probe incubated with Aldolase A and RNase T1.

5.4. Discussion.

The regulation of gene expression from transcription to translation is based mainly on the recognition of DNA and RNA sequences by proteins. There are several steps between initial synthesis of primary RNA transcript and production of a protein, which offer opportunities for subtle regulation of gene expression. The mechanism of RNA binding proteins as a means of processing mRNAs and regulating is well documented (Neupert et al, 1990; Rothenberger et al, 1990; Gorlach *et al*, 1993; Burd & Dreyfuss, 1994). However, among the various levels at which gene expression is regulated, the emphasis has been on transcriptional control and the 5' untranslated region. Now that the 3' untranslated region has become a focus for an additional mechanism of controlling gene expression, the RNA protein interactions of these regions are being extensively studied. The majority of proteins binding to mRNA 3'UTRs have been characterized by bandshift assays and UV cross-linking, and additionally by affinity purification experiments. In this chapter, the protein binding to the MyHC 3'UTR and sequence requirements were partially characterized. Experiments in which bound probe was eluted from retarded complexes showed that virtually the entire 3'UTR was bound to protein for all four skeletal MyHC 3'UTRs. Nested deletions for MyHC 2a 3'UTR were used to map more precisely the minimum sequence required for protein binding. Deletion analyses showed that the minimum number of bases needed for complex formation as comparable to that of full-length probe was 83 .However, whereas deletions subsequent to 2ADEL11 (72 bp) showed no binding, an anomaly was observed with deletions 2ADEL2D and 2ADEL2G, which are 59 and 46 nucleotides long, respectively. The bands detected with these deletions were more compact compared to the longer probes. Aldolase A was identified as the protein binding to the MyHC 2a 3'UTR after elution from bandshift gels and purification by SDS-PAGE. There is a possibility that other proteins that may have been present in the complex could have been lost during this process of elution.

Although from these results it would appear that the majority of the 3'UTR is interacting with a protein, it is possible that the protein is in fact only in contact with certain nucleotides, but has a requirement for specific secondary structure for protein

binding. Sequential deletion of the 2a 3'UTR resulted in changes in the predicted secondary structure of the probes, whose free energies of folding were less negative, and therefore less stable. The conserved motif is present in all but the largest deletion. Sequence comparisons in chapter 2 showed that the motifs CACAAAATGTG and CTTTG were conserved within the fast MyHC 3'UTRs across species, and hence it was hypothesized that these sequences could be protein binding sites. However, from deletion bandshift analyses, it was observed that these elements were not essential for complex formation since protein binding was decreased with deletions 2ADEL11(72 bases) and 2ADEL1C (66 bases), both of which contain these sequences. Furthermore, protein-binding seemed to increase with subsequent deletions, 2ADEL2D (59 bases) and 2ADEL2G (46 bases). These complexes are presumably formed by interaction with the same protein that binds to the full-length 2A 3'UTR since the mobility is the same. Looking at the predicted secondary structures of the full-length 3'UTR and the deleted 3'UTRs, it can be seen that the stem-loop structure formed by the CACAAAATGTG and CTTTG motifs does not change with respect to position and base pairing structure, except within the two smallest probes, where the CACAAAATGTG motif forms a bigger loop at the end of the RNA, and .is deleted totally in the smallest probe. A decrease in complex formation is first observed with 2ADEL11, in which the sequence ATTCTATTCCA has been deleted. These bases form a loop in the full-length secondary structure and in the deletions 2ADEL5 (92 bases) and 2ADEL1E (83 bases) which could be necessary for protein binding. The complex formation with 2ADEL2D and 2ADEL2G could be due to the protein recognizing an alternative stem-loop structure. From these observatons it can be concluded that secondary structure is plays an important part in complex formation and the conserved sequences within the 3'UTR may function to bring about the correct base-parings to allow RNA-protein interactions.

UV cross-linking experiments performed to identify protein-RNA complexes by size on SDS-PAGE proved to be difficult in that a high degree of background was observed. Several different approaches were tried, with elution of protein from irradiated reactions excised from native polyacrylamide gel proving to give the best

results, although there was still a lot of background. Several bands were seen, ranging from approximately 25 to 35 kDa cross-linked to the 3'UTR probes, with additional bands observed for the 2a probe. Although the UV cross-linking would appear to indicate that more than one protein binds to the MyHC 3'UTRs, the results of the protein elution experiments in section 5.3.5. contradict this by showing a single band of about 30 kDa. An explanation for this again could be the loss of proteins during the elution and purification process.

Coomassie stained bands sequenced by tryptic digestion identified the 30 kDa protein band eluted and purified from shifted complexes as aldolase A. Aldolase A is a glycolytic enzyme which converts fructose 1,6-bisphosphate and fructose 1 phosphate to dihydroxyacetone phosphate. It exists in several forms with aldolase A generally thought to be the muscle specific isoform, aldolase B the liver isoform and aldolase C expressed mostly in the brain. In the rat, two aldolase A mRNA species have been detected (Schweighoffer et al, 1986). One is expressed in non-muscle, fetal and adult slow-twitch muscle, the other is specific to adult fast-twitch muscle. These two isoforms differ by their 5' ends.

Although the binding of aldolase A to the MyHC 2a 3'UTR was unexpected, the binding of glycolytic enzymes to nucleic acids is not in itself a totally new concept. There are several reports which show that glycolytic enzymes function as DNA binding proteins (Ronai, 1993; Grosse et al, 1986). These include aldolase, GAPDH, PGK and LDH. GAPDH, in particular, has been assigned new roles in DNA repair, translational control of gene expression and in DNA replication (reviewed by Sirover 1996). Moreover, GAPDH has been reported to bind to not only single-stranded DNA but also RNA, poly A and poly(dA-dT) (Perucho et al, 1980). It is worth noting that in competition assays with MyHC 3'UTR in chapter 4, it was shown that binding of protein to the MyHC 3 UTRs could be competed out by poly A. Extensive studies on GAPDH have shown that it binds specifically to RNA from diverse sources such as transfer RNA, hepatitis A virus 5'UTR and AU-rich elements in the 3'UTR of several genes (Singh & Green, 1993; Nagy & Rigby, 1995; McGowan & Pekala, 1996; Schultz et al, 1996). It has been suggested that GAPDH may participate in RNA export since it discriminated between wild type tRNA and

tRNA mutants defective in nuclear export. The observation that GAPDH is present in both the nucleus and cytoplasm of HeLa cells raises the possibility that it may act as a shuttle between compartments. From experiments described in chapter 4 of this thesis, it was observed that the MyHC 3'UTRs bound cytoplasmic protein, but it was not clear whether they bound protein extracted form muscle nuclei. Therefore, a role for aldolase as a shuttle between the nucleus and the cytoplasm cannot be discounted.

In cell culture, aldolase has been shown to exist in a soluble as well as structure-bound form (Minaschek, et al, 1992). The soluble form is distributed homogeneously throughout the cytoplasm, excluding the nucleus and vesicles. The structure-bound form is associated with the actin cytoskeleton. A considerable amount of aldolase A is also located around the preinuclear region. Wiseman *et al* (1997), have reported the existence of a localization signal within the 3'UTR of the type 1 MyHC mRNA, with perinuclear localization of hybrid mRNA containing the MyHC 3'UTR. From this it could be hypothesized that aldolase A is somehow involved in the localization of myosin heavy chain mRNA 3'UTR.

Although this chapter has identified aldolase as 3'UTR binding protein that interacts with MyHC 2a 3'UTR, its interaction with type 1, 2b and 2x MyHC 3'UTRs still needs to be investigated with deletion and UV cross-linking analyses.

Chapter 6

General Discussion

The aim of the work described in this thesis was to investigate the function of the 3'UTR of the myosin heavy chain mRNAs. The 3'RACE method was successfully employed to amplify four adult mouse skeletal myosin heavy mRNA 3' ends corresponding to the slow type 1(β-cardiac-like) and fast types 2a, 2b, and 2x. Although partial sequences had been published for the fast 2a and 2b types, the work in this thesis has described the complete 3'UTR sequences for these MyHC mRNAs. Sequence analyses showed that, although these mouse MyHC 3'UTRs were isoform specific, as expected they showed more sequence similarity to homologous isoforms from other species, with sequence identities typically greater than 63%. Two sequence elements CACAAAATGT and CTTTG were found to be highly conserved within the mouse fast 2a, 2b and 2x MyHC mRNA 3'UTRs. The conservation of these motifs in other species as well suggested that they could be elements involved in regulation of gene expression or putative binding sites for proteins. This hypothesis was further supported the fact that studies by Cox *et al*, (1991) and Roy *et al*, (1984) suggested that differences observed in the transcription and cytoplasmic content of corresponding MyHC mRNAs, and the lack of corresponding MyHC protein were a result of changes in message stability and/or rate of translation. The hypothesis was tested in this thesis initially by transfection experiments with chimeric CAT reporter constructs to analyse the effects of the MyHC 3'UTRs on translation and message stability. The results from these experiments, while not definitive, showed that the MyHC 3'UTRs, after taking into account the effect of deleting the CAT 3'UTR, have no major effects on translation and mRNA stability in transfected C2C12 cells. The sensitivity and reproducibility of the experimental methods was insufficient to detect any subtle effects. Consistent with this finding is that the intrinsic stability of MyHC mRNA in cultured rat L6 myoblasts is approximately 55-60 hours. The half-life of CAT mRNA in comparison is about 6 hours. Therefore, increased stability of hybrid CAT-MyHC 3'UTR transcripts would have been expected after arrest of transcription with actinomycin D. The only major difference observed was an increase in CAT mRNA following deletion of the CAT 3'UTR, but whether this is a transcriptional or stability effect is not clear. A transient transfection system contributed to the variation in transfection results. Perhaps

differences in transitional rates and mRNA stabilities might have been observed in response to prolonged differentiation. However, this was difficult to assess, as a certain amount of cell death was observed and reporter activity was lost over longer periods in extracts from these cells. Argument against this hypothesis is provided by the finding of Medford *et al*, (1983) that the intrinsic stability of MyHC mRNA is the same in both myoblasts and myotubes, indicating that the intrinsic stability of the MyHC mRNA molecule does not change during myogenesis, and that MyHC protein synthesis is probably regulated by cytoplasmic mRNA concentration (Medford et al, 1983). Of course, one way of testing the MyHC 3'UTRs for stability elements would be to splice them on to the end of the coding region of a reporter gene that is rendered unstable by the removal of its 3'UTR and performing stable transfections. It would also be interesting to see if the expression and stability of reporter transcripts change in response to treatments, for example with hormones.

The RNA bandshift experiments described in chapter 4 demonstrated that the MyHC 3'UTRs bound protein in a tissue-specific manner. The RNA-protein interaction was detected specifically with protein present only in skeletal muscle as shown by bandshift analyses with extracts from different tissues. This protein was considered to cytoplasmic, as the complex formation with this fraction was stronger than with the nuclear fraction. However, faint complex bands were observed with nuclear protein extracts, and therefore it cannot be stated that the MyHC 3'UTR-binding factor is a cytoplasmic protein. An intriguing observation in the binding assays was that the slow MyHC 3'UTR, although identical to that of cardiac type, did not bind protein from cardiac muscle extract. Hence it was concluded that the binding factor is either specific to skeletal muscle and not present in cardiac muscle, or is expressed at very low levels in this tissue. Competition assays clearly showed that the RNA-protein interactions were specific for the MyHC 3'UTRs since unlabelled rat growth hormone could not compete out protein binding, However polyadenylic acid was able to compete for binding, suggesting an A rich binding site. Bandshift analysis with deleted MyHC 2a 3'UTR was not able to clarify whether the conserved elements CACAAAATGT and CTTTG found in the fast 3'UTRs were essential for protein binding. The results from these experiments showed that protein

>@■■■■■■«ÆÆ■@»ã¼âª ¼<¼Þ-|¼'Þ±Ä»Ä■@»ã▒Ã ¼~ª ½¼|■¼Ÿ-■½|âþ±Í¾»Ä■'±ãÞ{ Á¼>ª? ¼xþ

binding was diminished with deletions which still contained the conserved motifs, suggesting that secondary structure and not just sequence plays an important part in protein binding. Of the predicted secondary structures the 2a and 2b 3'UTRs are the most similar and the slow type 1 and 2x are more similar to each other.

From experiments carried out in chapter 5, a 30 kDa protein binding to MyHC 2a 3'UTR was identified as aldolase A. This is a glycolytic enzyme which exists in several isoforms. These isoforms are expressed in a tissue-specific manner, with aldolase A expressed predominantly in the muscle. Although aldolase A mRNA has been shown to be expressed in several tissues, including cardiac and liver in the rat, three different aldolase mRNAs (I, II and III) were detected. These differ in their 5'UTR length and sequence. Aldolase mRNA I (which has the shortest 5' region) was found to be expressed almost exclusively in skeletal muscle, and nearly completely repressed in other tissues where the mRNAs II and III replace mRNA I (Mukai *et al*, 1986). Furthermore, they type II and III mRNAs are found in fetal liver, but not adult liver. Therefore, aldolase A is developmentally regulated.

Although the proteins binding to the slow type 1, 2b and 2x MyHC 3'UTRs were not identified by protein sequencing, gel analysis of the purified bound protein suggested that these 3'UTRs also bound aldolase A. Bandshift assays with commercial aldolase A confirmed that all four of the MyHC 3'UTRs bound aldolase A. However, since the concentration of the purified commercial aldolase A was greater, the sequence specificity could not be confirmed with the rat growth hormone 3'UTR. Despite the unexpected result of finding that a glycolytic enzyme bound to MyHC 3'UTR RNA, this is not the first instance where a glycolytic enzyme has been found to bind to RNA. The role of GAPDH as a protein that binds not only RNA, but also poly A and single stranded DNA has been well-documented (Perucho et al, 1977). GAPDH has been implicated in a number of cellular activities that are unrelated to glycolysis. Apart from binding to single stranded nucleic acids, the involvement of GAPDH in protein phosphorylation, DNA repair and interaction with microtubules among others, has also been published (Huitorel & Pantaloni, 1985; Allen et al, 1987; Kawamoto & Caswell, 1986). The role of aldolase A in binding the MyHC 3'UTRs is as yet unclear. Similarly to GAPDH, aldolase A can

function in a manner other than glycolytic. Mejean et al, (1989) have shown that aldolase A exists in both the fluid and solid phases of the cytoplasm, suggesting roles in addition to glycolysis, in organization of the cytoplasm and cell motility. Aldolase A has been found to interact with certain DNA sequences in mouse SEWA sarcoma cells, and has been shown to be present in the nuclei of these cells (Ronai et al, 1992). Notably, like GAPDH, aldolase A also probably binds to polyadenylic acid as suggested by competition assays with cold poly A (chapter 4).

In mature muscle a functional role for aldolase A could be to provide a method of attachment to microtubules which are thought to be involved in translocation. Glycolytic enzymes have been shown to interact with microtubules (Walsh *et al* 1989). In muscle, the nuclei in fibres are surrounded by microtubules in the sub-sarcolemmal region and moreover, microtubules are more abundant in the end portion of muscle fibres. (Boudriau et al, 1993; Kano et al, 1991). Perhonen *et al* (1998), have shown that microtubules are necessary for dispersal of alpha-cardiac MyHC mRNA in neonatal rat cardiomyocytes. When contractile activity of the cardiomyocytes was arrested, the MyHC mRNA aggregated in the perinuclear region. The 3'UTR of alpha cardiac MyHC mRNA was necessary to produce the perinuclear localization in contraction arrested myocytes (Goldpsink *et al*, 1997). After resumption of contractile activity, the mRNA distribution returned to normal. However, if microtubules were depolymerzied prior to resuming contractile activity, the perinuclear localization of the mRNA was maintained. These myocytes still produced new myofibrils, suggesting that myofibrillogenesis may occur independently of mRNA localization.

MyHC isoform switching occurs during development, altered muscle activity, and in response to hormones. This requires the exchange of one type of myosin molecule with another better suited the functional demands placed on the muscle. Therefore, during muscle development and growth and during isoform transition, targeting of newly synthesized sarcomeric proteins is essential (Russell et al, 1992). Since mRNA transcripts act as a template for translation into protein, their localization would allow specific proteins to be synthesized in sub-cellular regions where they are required. Localized protein synthesis would be an efficient way of

targeting proteins to the correct site, as presumably less energy would be required to transport a single mRNA that can be translated several times than to transport many protein molecules. In addition, localization opens up the possibility of local translation control. Non homogenous distribution of mRNA transcripts has been observed in *Xenopus* oocytes (Yisraeli & Melton, 1988), cultured fibroblasts (Sundell & Singer, 1990) and developing neurons (Garner et al, 1988). In cultured chicken muscle vimentin, desmin, and vinculin mRNAs localize to the costameres where their protein products are concentrated (Morris & Foulton, 1994). In several asymmetric cell types (fibroblasts, endothelial cells and myoblasts), β-actin mRNA is localized near the leading edge of the cell, (Lawrence & Singer, 1986; Hill et al, 1994; Hoock et al, 1991; Hill & Gunning 1993). The sequence elements required for β-actin mRNA sorting are contained within its 3'UTR (Kislauskis et al, 1993; Kislauskis et al, 1994). The mechanism of β-actin localization involves the binding of several proteins in a complex to the 3'UTR. A 68 kDa protein, ZBP-1, which binds specifically to the β-actin 3'UTR contains an RNA binding domain which has strong homology to RNP1 and 2 motifs of the RNA recognition motif (Burd & Dreyffuss, 1994). ZBP-1 is thought to function as complex with other proteins regulated by a signal transduction event to localize β-actin mRNA to wards the signalling event, providing actin protein for motility (Ross *et al*, 1997).

The targeting of newly synthesized myosin could, in theory be achieved in three ways. By targeting of the protein itself, co-translational assembly due to targeting signals in the nascent protein or RNA, or by targeting of specific mRNAs for local synthesis of protein close to site of function. Larsen et al, (1969 nature) and Horne & Hesketh (1990) both proposed that myosin was synthesized close to or in association with myofibrils. Furthermore MyHC mRNA has been reported to be present in the myofibrillar cytoplasm (Hesketh et al, 1991). Russell *et al*, (1992) have shown that myosin translation from endogenous mRNA occurs in the inter-myofibrillar spaces along the periphery of the myofibrils. The myosin mRNA is probably anchored on the cytoskeleton. In situ hybridization has revealed that the distribution of MyHC mRNA coincides with the cytoskeleton, perinuclear region, and subsarcolemma (Russell & Dix, 1992). It has been suggested that association of

the MyHC mRNA with the cytoskeleton would retard diffusion of the mRNA and concentrate it around the nuclei and in regions of rapid growth and repair. This theory is supported by the finding that, when muscles are stretched, higher levels of MyHC mRNA are found at the myotendinous junction (Dix & Eisenberg, 1990). Therefore, the myotendinous junction is a site for mRNA accumulation which allows regional protein synthesis and myofibril assembly. Furthermore, Wiseman et al, 1997 have reported evidence of a localization signal in the slow skeletal MyHC 3'UTR which targeted a β-globin-MyHC 3'UTR hybrid mRNA to the perinuclear cytoplasm in myoblasts and early myotubes. These observations suggest a non-glycolytic functional role for aldolase A in perhaps MyHC mRNA translocation to regions of myosin synthesis, or localization and/or anchoring, since microtubules are more abundant in the end portion of muscle fibres and glycolytic enzymes have been shown to interact with microtubules (Walsh *et al* 1989; (Kano et al, 1991; Boudriau et al, 1993). When mRNA is transported form the nucleus to the cytoplasm, it is assembled into a messenger ribonucleoprotein particle (mRNP). The proteins associated with the mRNA could bind to both the mRNA and the cytoskeleton, accounting for the intracellular distribution of MyHC mRNA. Therefore the role of aldolase A could be a mechanism for attachment to the cytoskeleton as well as cytoplasmic transport of MyHC mRNA to regions of rapid growth such as the myotendinous junction. Binding of aldolase A to the MyHC 3'UTR and cytoskeletal association could in theory stabilize the mRNA and/or enhancing translation. Aldolase A could further provide a link between metabolic adaptation and structural adaptation. In response to exercise, muscle fibres produce higher levels of glycolytic as well as oxidative enzymes. This is particularly true during the conversion of type 2B and 2X fibres into type 2A and then to type 1.

As yet there is no data to suggest that MyHC 2a, 2b, or 2x 3'UTRs are involved in localization of mRNA. However, Shoemaker et al, (1999), have used *in situ* hybridization to localize slow type 1, fast 2a 2b, and 2x MyHC mRNAs in rat skeletal muscle fibres Two distinct patterns of sub-cellular localization were observed. Type 1 and type 2a MyHC mRNAs were located preferentially in the muscle periphery, whereas, type 2b and type 2x mRNAs showed homogenous

distribution across the fibre section. The slow type 1 3'UTR and 2a 3'UTRs which localize to the periphery share the 11 base motif UGUGAAGCCCU, whilst types 2b and 2x have the common motif UCUUUGUCACU. These motifs are in contrast to those highlighted in this thesis where myosin heavy chain 3'UTR sequences from various species were aligned. It was shown in chapter 2 that the motif CTTTG was common to slow type 1 and fast 2a, 2b, and 2x 3'UTRs, and the 10 base motif CACAAAATGT was commonly shared by types 2a, 2b and 2x but not slow type 1 3'UTRs. The motifs pointed out by Shoemaker et al do not appear to be as well conserved as the motifs highlighted in this thesis across species (see figure 6.1.). Even when comparing the shared sequences between type 1 and 2a, and type 2b and type 2x, in species other than the rat there are differences in bases. However the data presented provides evidence for isoform specific MyHC mRNA localization.

```
type I    AAGGGCCTGAATGAGGAGTA--GCT---CTTGTGCTACCCAGC-TCCAAG
2a        AAAATCATAAGCGAAGAGTAAGGCTGTCCCGATGCTGTGGAATGACCGAA
2x        AAAATCATAAGCGAAGAGTGATT-GATCCAAGTGCAGGAAAGTGACCAAA
2b        AAAGTCATAAGCGAAGAATAAT-CCATCTTTCTGTTGAGAGGTGACAGGA
          **...* *.*. **.**.*.          . ** ..   ..  .* ...

type I    GG---TGCCC-----GTGAAGCCCTCAGACCTGGAGCCTTTG-----CAA
2a        GGCGAGGCACAAAATGTGAAACCTTTGGTCATGC-GCCTGTGTGATTCTA
2x        GA-GATGAGCAAAATGTGAAGATCTTTGTCA--C-TC-TGTTTTGTAC--
2b        GA--AATCACAAAATGT-ACGTTCTTTGTCACTG-TCCTGTATATCAC--
          *.   ..  *     ** * .   * .*.*     .* *.*.     *

type I    --CAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTTTCCTTA-AAGCC
2a        TTCCATCCTGTTGTAAGGAAATAAAGAGCCCAAGTTCCTTTGCAAGC-
2x        -TCATAACTTTGGGAGATAAAAAATTTATC----------TGC---C-
2b        ----------------GGAAATAAATTCTGCAGATAATTTTGCAATCT
                        ..*. ..*...            *.    *
```

Figure 6.1. Alignment of mouse skeletal types 1, 2a, 2b and 2x MyHC 3'UTR sequences. The stop codons are indicated by boldface type and asterisks indicate absolutely conserved sequences. Conserved motifs as determined by sequence alignments in this thesis are highlighted by boldface maroon type, and boxed text denotes conserved motifs between type 1 and fast 2a, and fast types 2b and 2x as demonstrated by Shoemaker et al, 1999.

The introduction to this thesis described work published providing evidence for the functional roles of the 3'UTR of many mRNAs. Specific sequence elements within the 3'UTR can influence message stability, control cytoplasmic polyadenylation, regulate translation and are responsible for the localization of transcripts. In almost all of the 3'UTRs studied, the binding protein to specific sequences has been detected. It is clear that the 3'UTR and association of RNA-binding plays an important part in regulating gene expression. It is apparent that the 3'UTR does not have one specific function, and in some instances has more then one function, such as in the *Drosophila* embryo, where the 3'UTR of *nos* mRNA mediates its translational repression and localization during development. The mechanisms behind the control of expression by the 3'UTRs are, although as yet unclear, under continuous investigation. Several hypotheses have been put forward as to how the 3'UTRs regulate expression. In the control of gene expression by translation dependent on the polyadenylation state of the mRNA, the presence of cytoplasmic polyadenylation control elements (CPEs/ACEs) is well established. Several questions arise in terms of how translation is activated by cytoplasmic polyadenylation. Does polyadenylation activate translation? Or does translational activation cause polyadenylation? Or is the correlation of polyadenylation and activation of translation triggered by an earlier event, where neither is dependent on each other. One example where polyadenylation is the prime cause of translation activation is that of the tissue plasminogen activator (tPA) mRNA which is stored in a stable untranslated form with a short poly A tail in immature oocytes (Strickland *et al*, 1988). During oocyte maturation, the poly A tail lengthens and a burst of tPA synthesis is observed. When the 3' terminal 455 nucleotides of tPA mRNA were micro-injected into primary oocytes or attached to a reporter gene, maturation dependent polyadenylation and activation of translation were observed (Vassalli *et al*, 1989). Both polyadenylation and translation were blocked by the addition of cordycepin to the 3' end of the chimeric construct. In addition, reporter constructs without the tPA 3'UTR polyadenylated *in vitro* prior to micro-injection were also translated Therefore, the poly A tail is a major determinant of translatability in this case. The role of phosphorylation in RNA-protein interactions between the ACE of

tPA in mouse oocytes was demonstrated by the binding of an 85 kDa protein in primary oocytes. This protein was phosphorylated after maturation.

In *Xenopus*, the situation is less clear as illustrated by G10 mRNA. This mRNA has a short poly A tail in immature oocytes, which is elongated following progesterone treatment. However, only 60%-70% of this polyadenylated mRNA is associated with polysomes implying that, although polyadenylation may be required for activation of translation, it is not sufficient (McGrew *et al*, 1989). In experiments where *in vitro* polyadenylated G10 mRNA was micro-injected, the transcripts were not mobilized into polysomes in immature oocytes, but upon progesterone-induced maturation, the poly A tail was further lengthened, and approximately 30% of the mRNA was associated with polysomes. Whereas maturation is not a requirement for activation of translation of *in vitro* polyadenylated tPA mRNA in the mouse oocytes or B4 mRNA in *Xenopus* oocytes, maturation is necessary in for G10 mRNA and another *Xenopus* mRNA C12 (Simon *et al,* 1992). Therefore, it has been suggested that in the latter, it is the process of extension of the poly A tail rather than the length of the poly A tail that causes transitional activation (McGrew *et al,* 1989).

The general assumption is that binding of specific proteins to *cis*-acting 3'UTR motifs such the CPE/AC motifs is central to the mechanisms of control of polyadenylation and/or translation. Two modes of regulatory RNA-protein interactions can be visualized. First, that the protein acts as repressor, where its binding to the 3'UTR blocks translation initiation as observed with LOX mRNA, or blocks polyadenylation by occluding the CPE or AAUAAA motif. At the appropriate stage, phosphorylation or dephosphorylation might then, reduce the affinity of this protein for the RNA, activating translation and/or polyadenylation. Secondly, the protein could be a positive activator, with the phosphorylation state again, activating its binding to the 3'UTR at the correct stage. This in turn would promote transitional activation and/or polyadenylation. It is likely that proteins that are involved in polyadenylation and transitional activation fall into the second category. For example, a 58-kDa protein in *Xenopus* oocytes, which specifically cross-linked to the CPE in B4 mRNA, was found to be phosphorylated at maturation (Paris *et al*, 1991). Analysis of the cDNA sequence of this protein showed an RNA recognition motif

and a consensus Cdc2 kinase phosphorylation site. *In vitro* experiments have demonstrated that both polyadenylation of the mRNA and phosphorylation of the protein could be induced in oocyte extracts by the addition of cyclin B and depletion of extracts of Cdc2 kinase prevented both events (Richter, 1991). Another CPE binding protein, CPEB, purified from *Xenopus* oocytes was found to contain two RNA recognition motifs and within this region there was a 62% homology with an oocyte-specific RNA-binding form *Drosophila* (Hake & Richter, 1994). The 62-kDa CPEB, which was detected in oocytes but not in embryos beyond the gastrula stage, was phosphorylated doing oocyte maturation, correlating with the induction of polyadenylation. The requirement of this protein for polyadenylation was demonstrated by depletion of CPEB from oocyte extracts which were incapable of adenylating exogenous RNA.

Phosphorylation is also implicated in 3'UTR RNA-protein binding regulating translation independently of polyadenylation. For example, the temporal repression of mouse protamine 2 mRNA involves the binding of an 18 kDa protein to elements within its 3'UTR. This developmentally regulated protein requires phosphorylation to bind RNA in order to repress translation. At a later stage in male gamete differentiation, binding of the 18 kDa protein is no longer observed, probably due to dephosphorylation, thus allowing the protamine mRNA to be translated (Kwon & Hecht, 1993).

Inhibition of translation by the 3'UTR-protein interactions presumably involves some form of interaction with the 5'end of the mRNA (see figure 6.2.). One model for the mechanism of inhibition could entail the 3'UTR binding protein occluding the 5'-cap structure either directly through an affinity of the 3'UTR binding protein for the cap (figure 6.2. A), or indirectly through protein-protein interactions with another unidentified protein (figure 6.2. B). Alternatively the binding of a specific protein to the 3'UTR could act as a signal for other proteins to bind to the mRNA, resulting in a messenger ribonucleoprotein particle (mRNP) which is inaccessible for translation (figure 6.2. C).

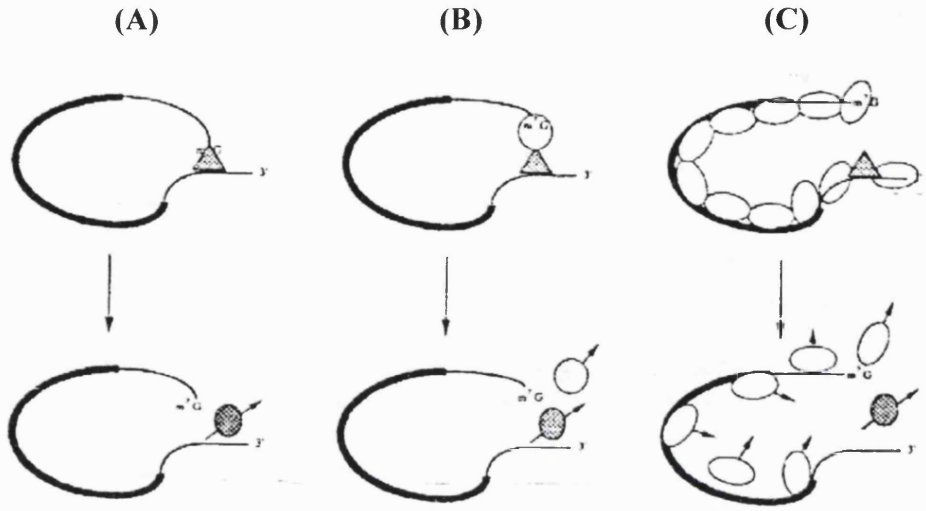


Figure 6.2. Models for inhibition of transitional initiation by binding of specific proteins to 3'UTR motif. Thickened line represents coding region of the mRNA; the specific 3'UTR binding protein is shown in its bound state as a stippled triangle, which is presumed to undergo a conformational change (symbolized by the change to an oval shape) that reduces its binding affinity and allows translation. Open ovals indicate other non-specific RNA-binding proteins. (A) Occlusion of the 5'-cap structure through an affinity of the 3'UTR binding protein for the cap. (B) Indirect h protein-protein interactions with another unidentified protein. (C) Binding of a specific protein to the 3'UTR acting as a signal for other proteins to bind to the mRNA, resulting in a messenger ribonucleoprotein particle (mRNP) (adapted from Jackson, 1994).

RNA-protein interactions are also important in the regulation of mRNA stability. The rapid decay of mRNAs coding for cytokines, lymphokines and oncogenes has been attributed to AU and U rich *cis*-elements located within their 3'UTRs. These elements bind proteins, which influence the turnover of such mRNAs. As described in the introduction, the removal of the poly A tail, the first critical step in the AU-rich element mediated decay of *c-fos* mRNA, is thought to be caused by the formation of a complex between a protein and a 20 nucleotide U-rich region within the AU-rich element in the 3'UTR (Shyu *et al*, 1991). Such a complex could bring about mRNA decay by perhaps destabilizing the poly A binding protein-poly A tail complex, which is a major determinant of mRNA stability (Bernstein *et al*, 1989), making the poly A tail more vulnerable to exonuclease attack from the 3'end. Alternatively, the protein complex could activate or recruit a poly A-specific exonuclease. Conversely the binding of proteins to the AUUUA motif found in the

3'UTRs of labile mRNAs of lymphokines and cytokines acts to stabilize the message. One such protein, the 44 kDa AU-binding factor (AUBF) has been shown to bind several lymphokine, cytokine and oncogene mRNAs through the AUUUA motif in the 3'UTR. (Gillis & Malter, 1990). Phosphorylation is required for binding activity, which is up-regulated by treatment with tetradecanoyl phorbol acetate (TPA) or calcium ionophore, which are known to stabilize cytokine and lymphokine messages. These agonists activate protein kinase C, thus the mechanism of stabilization of labile mRNAs upon TPA/ionophore treatment involves activation of AUBF and binding to the AUUUA motifs, delaying the turnover of the transcript. The binding of such factors could block the recognition or degradation.

A 34 kDa AU-rich binding protein, AU-A, which is one of three factors found to bind to AUUUA multimers in lymphokine mRNA 3'UTRs (Bohjanen *et al*, 1991; Bohjanen *et al*, 1992) is found in both the nucleus and cytoplasm. It binds to several other U-rich regions, and the AU-rich region of c-*myc*. It has been shown that AU-A shuttles between the cytoplasm and the nucleus, suggesting that it is a part of a ribonucleoprotein complex that is involved in nucleo-cytoplasmic transport of mRNA. It has been proposed that exchange of other RNA-binding proteins for AU-A on AUUUA sequences in the cytoplasm could be involved in the regulation of mRNA decay (Katz *et al*, 1994).

The stability of the human α-globin mRNA is determined by the formation of an RNA-protein complex (α -complex) with polypyrimidine-rich tracks within the α-globin mRNA 3'UTR. The α-complex is thought to provide a protective mechanism in conjunction with the poly A binding protein, since sequestration of alphaCP1 and alphaCP2, two of the poly C binding proteins enhanced deadenylation and decay of the mRNA. PABP was found to essential for the stability of the mRNA *in vitro*, since rapid deadenylation resulted in its depletion (Wang *et al*, 1999). Therefore this suggest interaction between the α -complex and PABP to stabilize the globin mRNA. AlphaCP1 was found, in addition, to be able to specifically bind to poly C regions within the 3'UTR of mRNAs encoding the transmembrane receptor protein, TAPA-1, and the mitochondrial cytochrome c oxidase subunit II enzyme, cox II (Trifillis *et al*, 1999). An interesting finding in the study of the α-complex was that among the other

proteins which are an essential requirement for alphaCP1 and alphaCP2 to bind to the poly C region was AUF1, previously identified as an AU-rich binding factor (Kiledjian *et al*, 1997) Thus, the α -complex is probably a dynamic rather than fixed complex.

Evidently, signal transduction and phosphorylation events play a key role in 3'UTR-mediated post-transcriptional regulation. From the work published to date it would appear that the 3'UTR-protein interactions are dynamic and probably involve several proteins, some common to several 3'UTRs, acting in conjunction to regulate expression overall. In some cases such as that of mRNA stability, abetter understanding of the general and specific mechanisms through which mRNA stability is modulated raises the potential of novel therapeutic applications to common genetic diseases by altering the stability of target mRNAs.

Future experiments.

Although the work in this thesis has provided evidence for the binding of aldolase A to the 3'UTR of MyHC mRNA, the precise function of the MyHC 3'UTRs and the association of aldolase A needs to be further investigated. The main questions that need to be answered are whether the MyHC 3'UTRs confer localization, whether aldolase A functions to localize or anchor the MyHC mRNAs and whether phosphorylation is involved. The localization studies will involve transfection of reporter-3'UTR chimeric plasmid constructs into myoblasts and *in situ* hybridisation with probes to look at the localization and distribution of reporter/3'UTR mRNAs. The RNA-protein interactions of the MyHC 3'UTRs will be further analysed through poly U, poly C and oligo competition assays and phosphorylation studies. RNase footprinting will also be performed to identify the protein-binding sequence of each 3'UTR. Hopefully, such analyses will provide answers that will lead to further insights in the regulation of the MyHC and implications for muscle structure and function.

APPENDIX

I. Bacterial strains and media.

LB Medium (Luria-Bertani Medium)

Per Litre: 10g tryptone
5g yeast extract
10g NaCl

SOC Medium

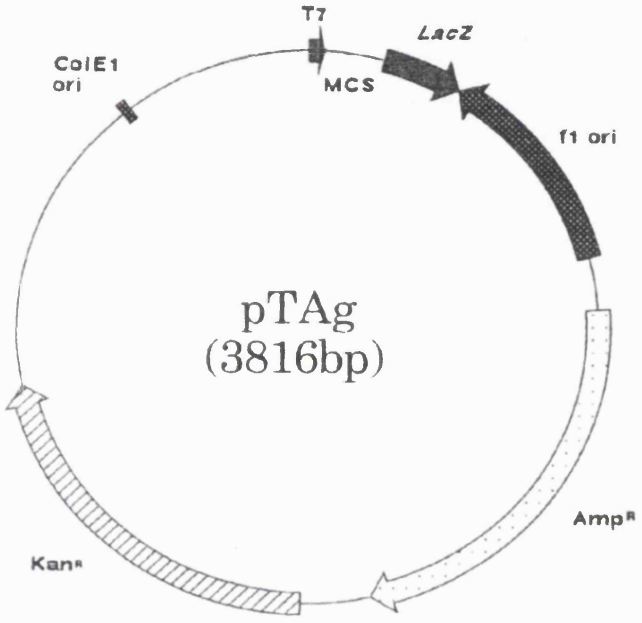
Per Litre: 20g tryptone
5g yeast extract
0.5g NaCl
2.5 mM KCl
10 mM $MgCl_2$
20 mM glucose

Strain	Genotype
JM109	*endA*1, *recA*1, *gyrA*96, *thi*, *hsd*R17 (r_k^-, m_k^+), *relA*1, *supE*44, Δ(*lac-proAB*), [F', *traD*36, *proAB*, *lacI*^{q}ZΔM15]
DH5α	Φ80d*lacZ*ΔM15, *recA*1, *endA*1, *gyrA*96, *thi*-1, *hsd*R17 (r_k^-, m_k^+), *supE*44, *relA*1, *deoR*, Δ(*lacZYA-argF*)U169
INVαF'	F' *endA*1, *recA*1, *hsd*R17 (r_k^-, m_k^+), *supE*44, *thi*-1, *gyrA*96, *relA*1, Φ80d*lacZ*ΔM15, Δ(*lacZYA-argF*)U169λ$^-$

Symbol	Description	Effect
*deo*R	Regulatory gene mutation allowing constitutive expression of genes for deoxyribose synthesis	Allows uptake of large plasmids.
*endA*1	Endonuclease mutation	Improves quality of plasmid DNA isolations.
*gyrA*96	DNA gyrase mutation	Confers resistance to nalidixic acid.
*hsd*R17 (r_k^-, m_k^+)	Restriction minus, modification positive	Allows cloning without cleavage of transformed DNA by endogenous restriction endonucleases. DNA prepared from this strain can be used to transform r_k^+ *E.coli* strains.
lacI^q	Over production of the *lac* repressor protein	Leads to high levels of the *lac* repressor protein, inhibiting transcription from the *lac* promoter.
*lac*ZΔM15	Partial deletion of β-D-galactosidase gene	Allows complementation of β-D-galactosidase activity by α-complementation sequence in vector. Allows blue/white selection of recombinant colonies when plated on X-Gal
proAB	Mutations in proline metabolism	Requires proline for growth in minimal media.
*recA*1	Mutation in recombination	Prevents recombination of introduced DNA with host DNA, ensuring stability of inserts.
relA	Relaxed phenotype; mutation eliminating stringent factor	Allows RNA synthesis in the absence of protein synthesis
supE	Suppressor mutation	Suppresses amber (UAG) mutations
thi-1	Mutation in thiamine metabolism	Thiamine required for growth in minimal media.
*traD*36	Transfer factor mutation	Prevents trnasfer of F' episome

II. Plasmids used in cloning.

Map of the plasmid pTAG,,used for cloning of PCR products with the LigATor system (R&D Systems)



pTAg SEQ 5' primer — T7 Promoter ——>

GCT ATG ACC ATG ATT ACG CCA AGC TCT AAT ACG ACT CAC TAT AGG GAA AGC
CGA TAC TGG TTC TAA TGC GGT TCG AGA TTA TGC TGA GTG ATA TCC CTT TCG

Pst I — Insert

Kpn I — Sph I — Mlu I — SnaB I — BamH I — EcoR I

TCG GTA CCA CGC ATG CTG CAG ACG CGT TAC GTA TCG GAT CCA GAA TTC GTG ATT A
AGC CAT GGT GCG TAC GAC GTC TGC GCA ATG CAT AGC CTA GGT CTT AAG CAC TA TT

Sal I
Acc I — Avr II — Nhe I — BstX I — Age I
Hinc II — Hind III — Xho I — Sty I — Xba I — Pml I — EcoO109 I

TCT GTC GAC AAG CTT CTC GAG CCT AGG CTA GCT CTA GAC CAC ACG TGT GGG GGC
AGA CAG CTG TTC GAA GAG CTC GGA TCC GAT CGA GAT CTG GTG TGC ACA CCC CCG

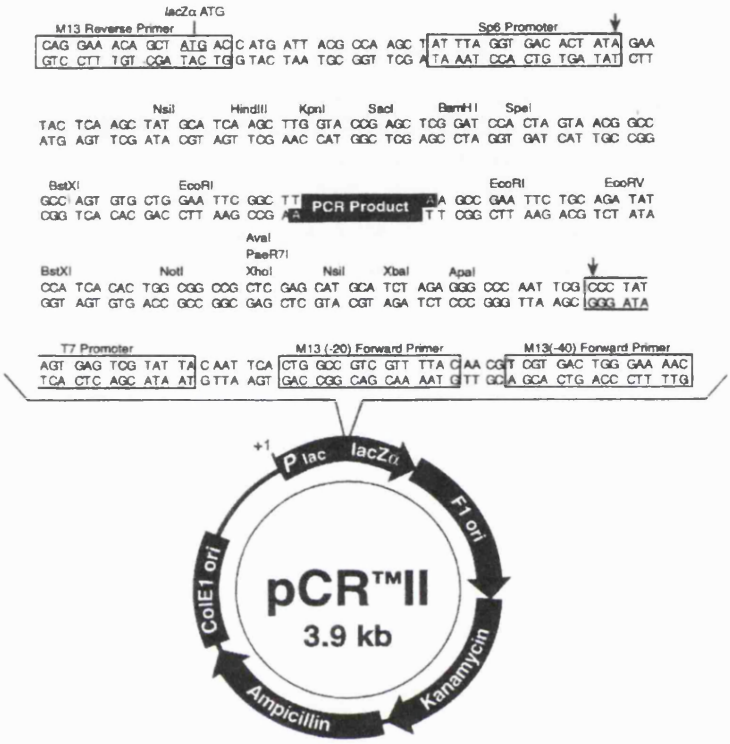
Not I
Sac I — Eag I

CCG AGC TCG CGG CCG CAC AAT TCA CTG GCC GTC GTT TTA CAA
GGC TCG AGC GCC GGC GTG TTA AGT GAC CGG CAG CAA AAT GTT

pTAg SEQ 3' primer

Map of pCR™II

The map of the linearized vector, pCR™II, is shown below. The sequence of the multiple cloning site is shown with a PCR product inserted by TA Cloning®. Arrows indicate the start of transcription for Sp6 and T7 RNA polymerases, respectively.

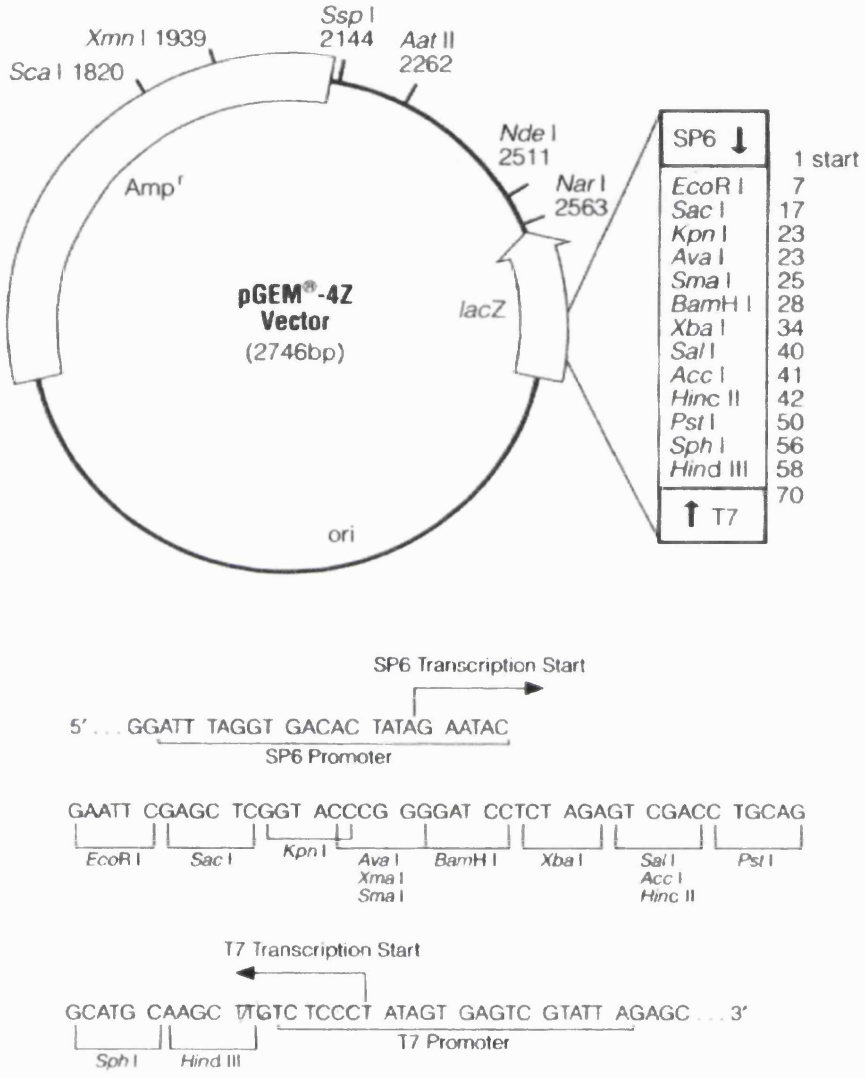


Features and Functions of pCR™II

The table below describes the features of pCR™II.

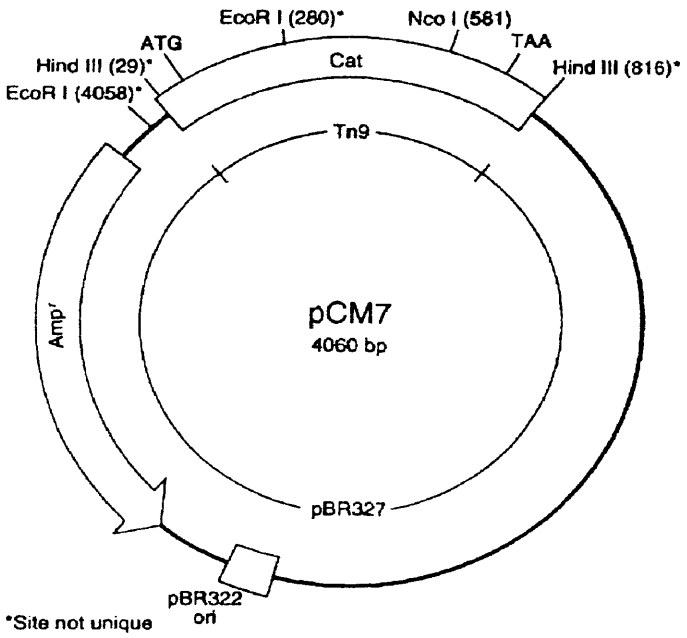
Feature	Function
lac promoter	For bacterial expression of the *lacZα* fragment for α-complementation (blue-white screening).
lacZα fragment	Encodes the first 146 amino acids of β-galactosidase. Complementation in *trans* with the Ω fragment gives active β-galactosidase for blue-white screening.
Kanamycin resistance gene	Selection and maintenance in *E. coli*, especially useful when cloning products amplified from ampicillin-resistant plasmids.
Ampicillin resistance gene	Selection and maintenance in *E. coli*.
ColE1 origin	Replication, maintenance, and high copy number in *E. coli*.
T7 and Sp6 promoters and priming sites	*In vivo* or *in vitro* transcription of sense and antisense RNA. Sequencing of insert.
M13 Forward (-20 and -40) and M13 Reverse Priming Sites	Sequencing of insert.
f1 origin	Rescue of sense strand for mutagenesis and single-strand sequencing.

(Adapted from Original TA Cloning Kit manual, Invitrogen).

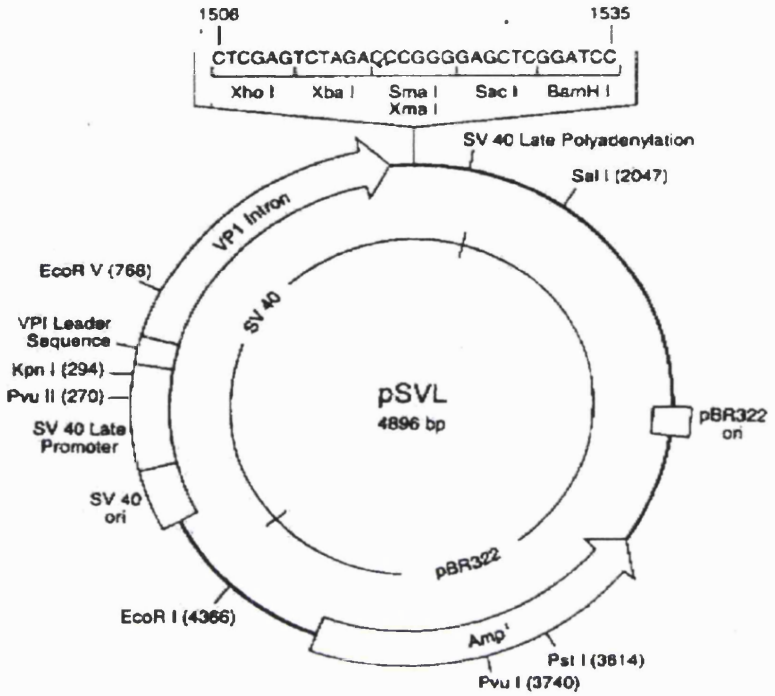


Map of pGEM4Z vector promoter and multiple cloning site sequence. Sequence shown corresponds to RNA synthesized by SP6 RNA polymerase and is complimentary to RNA synthesized by T7 RNA polymerase (Adapted from Protocols and Applications Guide, Promega).

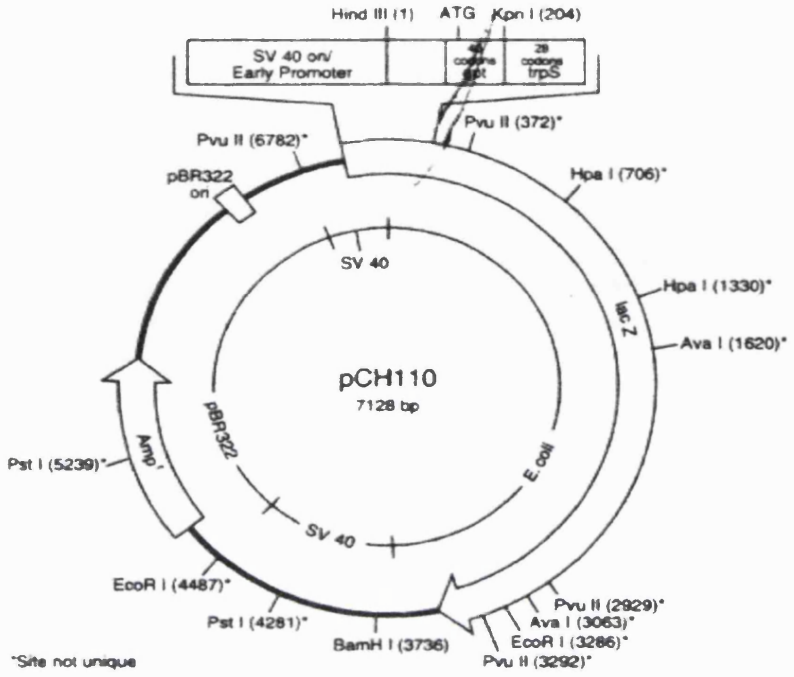
III. Plasmids used for chimeric constructs and transfections.



Map of plasmid pCM7, which contains the complete polypeptide coding sequence of chloramphenicol acetyl transferase gene. The CAT cartridge can be excised from the plasmid with *Hind* III enzyme (from Pharmacia).



Map of plasmid pSVL, which is designed for high-level transient expression in eukaryotic cells. Genes cloned into the multiple cloning site are expressed from the first ATG codon using the SV40 late promoter (Sprague *et al*, 1983; Templeton & Eckhart, 1984). Transcripts are spliced and polyadenylated using the SV40 VP1 processing signals.



Map of plasmid pCH110 assay vector for monitoring and normalizing expression in eukaryotic cells. It contains a functional *lacZ* gene which is expressed from the SV40 early promoter in eukaryotes (Hall *et al*, 1983). When co-transfected with a second plasmid, pCH110 provides an internal marker to monitor and normalize transient and stable expression between experiments (Herbomel *et al*, 1984).

IVPublished sequences for MyHC 2a and 2x 3'ends

MdMs (2x)

CGGGTGAAGAGCCGGGAGTTCACACCAAAATCATAAGCGAAGGAGTGAtt gatccaagtgcaggaaagtgaccaaagagatgagcaaaatgtgaagatctttgtcactctgttttgtactcataactttggga g**aataaa**aaatttatctgccaagagctctgtcagtctcttccttttcccttccatggtggccgatcatttgtacatcaggagag gaaaaagaaagacc

2A

CGGGTGAAGAGCCGGGAGGTTCACACCAAATCATAAGCGAAGAGTAAggc tgtcccgatgctgtggaatgaccgaaggcgaggcacaaaatgtgaaacctttggtcatgcgcctgtgtgattctattccatc ctgttgtaagga**ataaa**gagcccaagttcctttgcaagcaatcaccttggttggcctctctttgtcccttcctta

Sequences for the MyHC 2A and 2X (MdMs) 3' noncoding regions isolated and identified from genomic library by Parker-Thornburg *et al* (1992). The poly A signals are highlighted in bold and the underlined sequences correspond to 3'UTR sequences as identified in chapter 2.

V. Formulae for calculation of specific activity and yield of radiolabelled probes.

The specific activity and yield of radiolabelled probes was calculated using the following formulae.

% incorporation = $\frac{\text{incorporated cpm}}{\text{total cpm}} \times 100$

Total cpm incorporated

= incorporated cpm x dilution factor x $\frac{\text{reaction volume}}{\text{volume counted}}$

Total amount of RNA synthesized. This is determined by the amount of limiting rNTP present in the reaction, the maximum theoretical yield and the percent incorporation.

nmol of labelled rNTP = $\frac{\mu\text{Ci rNTP in reaction}}{\text{isotope concentration } (\mu\text{Ci/nmol})}$

nmol of limiting cold rNTP

= µl of limiting cold rNTP x 100µM rNTP x $\frac{10^3\text{nmol}}{1\mu\text{mol}} \times \frac{1\text{L}}{10^6\mu\text{l}}$

total nmol of limiting rNTP

= nmol of labelled rNTP + nmol of limiting cold rNTP

maximum theoretical RNA yield

= total nmol of limiting rNTP x 4 (rNTPs) x $\frac{330\text{ng (rNTP)}}{\text{nmol}}$

total µg of RNA syntesized

= % incorporation x. maximum theoretical RNA yield

VI. One Way Analysis of Variance

Statistical data for transfection optimization experiment results:
DNA/lipid ratios.

One Way Analysis of Variance
Normality Test: Failed (P = 0.003)
Equal Variance Test: Passed (P = 0.221)

Group Name	N	Missing	Mean	Std Dev	SEM
2.5 1/2	3	0	391526.667	44012.446	25410.597
2.5 1/3	3	0	10103005.000	3784241.401	2184832.792
2.5 1/4	3	0	16288516.000	160139.944	92456.840
2.5 1/5	3	0	2387919.000	599901.722	346353.421
5 1/2	3	0	707547.667	181435.367	104751.758
5 1/3	3	0	17513760.000	2286615.195	1320177.898
5 1/4	3	0	21042694.000	1271451.159	734072.669
5 1/5	3	0	14906213.333	4476109.006	2584282.739
10 1/2	3	0	612175.000	336087.357	194040.126
10 1/3	3	0	11695601.667	2112648.445	1219738.149
10 1/4	3	0	24686768.333	4163830.103	2403988.431
10 1/5	3	0	20352210.000	1896235.222	1094791.916

Source of Variation	DF	SS	MS	F	P
Between Groups	11	2.571E+015	2.337E+014	41.773	<0.001
Residual	24	1.343E+014	5.594E+012		
Total	35	2.705E+015			

The differences in the mean values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (p<0.001).
Power of performed test with alpha = 0.050:1.000

All Pairwise Multiple Comparison Procedures (Student-Newman-Keuls Method):
Comparisons for factor:

Comparison	Diff of Means	p	q	P	P<O.050
10 1/4 vs. 2.5 1/2	24295241.667	12	17.791	<0.001	Yes
10 1/4 vs. 10 1/2	24074593.333	11	17.630	<0.001	Yes
10 1/4 vs. 5 1/2	23979220.667	10	17.560	<0.001	Yes
10 1/4 vs. 2.5 1/5	22298849.333	9	16.329	<0.001	Yes
10 1/4 vs. 2.5 1/3	14583763.333	8	10.680	<0.001	Yes
10 1/4 vs. 10 1/3	12991166.667	7	9.513	<0.001	Yes
10 1/4 vs. 5 1/5	9780555.000	6	7.162	<0.001	Yes
10 1/4 vs. 2.5 1/4	8398252.333	5	6.150	0.002	Yes
10 1/4. Vs. 5 1/3	7173008.333	4	5.253	0.006	Yes
10 1/4 vs. 10 1/5	4334558.333	3	3.174	0.084	No
10 1/4 vs. 5 1/4	3644074.333	2	2.669	0.071	Do Not Test
5 1/4 vs. 2.5 1/2	20651167.333	11	15.123	<0.001	Yes
5 1/4 vs. 10 1/3	20430519.000	10	14.961	<0.001	Yes
5 1/4 vs. 5 1/29	20335146.333	9	14.891	<0.001	Yes
5 1/4 vs. 2.5 1/5	18654775.000	8	13.661	<0.001	Yes
5 1/4 vs. 2.5 1/3	10939689.000	7	8.011	<0.001	Yes
5 1/4 . 10 1/3	9347092.333	6	6.845	<0.001	Yes
5 1/4 vs. 5 1/5	6136480.667	5	4.494	0.030	Yes
5 1/4 vs. 2.5 1/4	4754178.000	4	3.481	0.092	No
5 1/4 vs. 5 1/3	3528934.000	3	2.584	0.182	Do Not Test
5 1/4 vs. 10 1/5	690484.000	2	0.506	0.724	Do Not Test
10 1/5 vs. 2.5 1/2	19960683.333	10	14.617	<0.001	Yes
10 1/5 vs. 10 1/2	19740035.000	9	14.456	<0.001	Yes
10 1/5 vs. 5 1/2	19644662.333	8	14.386	<0.001	Yes
10 1/5 vs. 2.5 1/5	17964291.000	7	13.155	<0.001	Yes

10 1/5 vs. 2.5 1/3	10249205.000	6	7.506	<0.001	Yes
10 1/5 vs. 10 1/3	8656608.333	5	6.339	0.001	Yes
10 1/5 vs. 5 1/5	5445996.667	4	3.988	0.044	Yes
10 1/5 vs. 2.5 1/4	4063694.000	3	2.976	0.110	Do Not Test
10 1/5 vs. 5 1/3	2838450.000	2	2.079	0.155	Do Not Test
5 1/3 vs. 2.5 1/2	17122233.333	9	12.539	<0.001	Yes
5 1/3 vs. 10 1/2	16901585.000	8	12.377	<0.001	Yes
5 1/3 vs. 5 1/2	16806212.333	7	12.307	<0.001	Yes
5 1/3 vs. 2.5 1/5	15125841.000	6	11.077	<0.001	Yes
5 1/3 vs. 2.5 1/3	7410755.000	5	5.427	0.007	Yes
5 1/3 vs. 10 1/3	5818158.333	4	4.261	0.029	Yes
5 1/3 vs. 5 1/5	2607546.667	3	1.910	0.382	No
5 1/3 vs. 2.5 1/4	1225244.000	2	0.897	0.532	Do Not Test
2.5 1/4 vs. 2.5 1/2	15896989.333	8	11.641	<0.001	Yes
2.5 1/4 vs. 10 1/2	15676341.000	7	11.480	<0.001	Yes
2.5 1/4 vs. 5 1/2	15580968.333	6	11.4t0	<0.001	Yes
2.5 1/4 vs. 2.5 1/5	13900597.000	5	10.179	<0.001	Yes
2.5 1/4 2.5 1/3 6	6185511.000	4	4.530	0.019	Yes
2.5 1/4 vs. 10 1/3	4592914.333	3	3.363	0.064	No
2.5 1/4 vs. 5 1/5	1382302.667	2	1.012	0.481	Do Not Test
5 1/5 vs. 2.5 1/2	14514686.667	7	10.629	<0.001	Yes
5 1/5 vs. 10 1/2	14294038.333	6	10.468	<0.001	Yes
5 1/5 vs. 5 1/2	14198665.667	5	10.398	<0.001	Yes
5 1/5 vs. 2.5 1/5	12518294.333	4	9.167	<0.001	Yes
5 1/5 vs. 2.5 1/3	4803208.333	3	3.517	0.051	No
5 1/5 vs.10 1/3	3210611.667	2	2.351	0.110	Do Not Test
10 1/3 vs. 2.5 1/2	11304075.000	6	8.278	<0.001	Yes
10 1/3 vs. 10 1/3	11083426.667	5	8.116	<0.001	Yes
10 1/3 vs. 5 1/2	10988054.000	4	8.047	<0.001	Yes
10 1/3 vs. 2.5 1/5	9307682.667	3	6.816	<0.001	Yes
10 1/3 vs. 2.5 1/3	1592596.667	2	1.166	0.418	Do Not Test
5 1/3 vs. Cot S	9711478.333	5	7.112	<0.001	Yes
5 1/3 vs. 10 1/2	9490830.000	4	6.950	<0.001	Yes
5 1/3 vs. 5 1/2	9395457.333	3	6.880	<0.001	Yes
5 1/3 vs. 2.5 1/5	7715086.000	2	5.650	<0.001	Yes
2.5 1/5 vs. 2.5 1/2	1996392.333	4	1.462	0.732	No
2.5 1/5 vs. 10 1/2	1775744.000	3	1.300	0.634	Do Not Test
2.5 1/5 vs. 5 1/2	1680371.333	2	1.231	0.393	Do Not Test
5 1/2 vs. 2.5 1/2	316021.000	3	0.231	0.985	Do Not Test
5 1/2 vs. 10 1/2	95372.667	2	0.0698	0.961	Do Not Test
10 1/2 vs. 2.5 1/2	220648.333	2	0.162	0.910	Do Not Test

A result of "Do Not Test" occurs for a comparison when no significant difference is found between two means that enclose that comparison. For example, if you had four means sorted in order, and found no difference between means 4 vs. 2, then you would not test 4 vs. 3 and 3 vs. 2, but still test 4 vs. I and 3 vs. 1(4 vs. 3 and 3 vs. 2 are enclosed by 4 vs. 2: 4 3 2 1). Note that not testing the enclosed means is a procedural rule, and a result of Do Not Test should be treated as if there is no significant difference between the means, even though one may appear to exist.

Statistical data for transfection optimization experiment results:
Cell number optimization.
Normality Test: Passed (P> 0.200)
Equal Variance Test: Passed (P = 0.685)

Group Name	N	Missing	Mean	Std Dev	SEM
100	3	0	33921507.333	6291026.550	3632125.872
200	3	0	22865614.333	3996236.273	2307228.088
300	3	0	23143445.333	3137442.885	1811403.494
400	3	0	24006414.333	6212770.327	3586944.621

Source of Variation	DF	SS	MS	F	P
Between Groups	3	2.541E+014	8.471E+013	3.258	0.081
Residual	8	2.080E+014	2.600E+0 13		
Total	11	4.621E+014			

The differences in the mean values among the treatment groups are not great enough to exclude the possibility that the difference is due to random sampling variability; there is not a statistically significant difference (P = 0.081).
Power of performed test with alpha = 0.050: 0.374

Statistical data for transfection experiment results

Normality Test: Failed (P = 0.001)
Equal Variance Test: Failed (P = 0.009)

Group Name	N	Missing	Mean	Std Dev	SEM
pSVLCAT	9	0	16072.333	5206.891	1735.630
pSVLCAT	9	0	85877.000	32471.766	10823.922
pSVLCA□β	9	0	124276.889	46974.340	15658.113
pSVLCAT□2a	9	0	60343.000	16460.150	5486.717
pSVLCAT□2b	9	0	72998.333	60187.854	20062.618
pSVLCAT'□2x	9	0	111772.222	55919.417	18639.806

Source of Variation	DF	SS	MS	F	P
Between Groups	5	67627113894.148	13525422778.830	7.87	<0.001
Residual	48	82468891852.444	1718101913.593		
Total	53	150096005746.593			

The differences m the mean values among the treatment groups are greater than would be expected by chance; there is a statistically significant difference (P <0.001).
Power of performed test with alpha = 0.050: 0.998

All Pairwise Multiple Comparison Procedures (Fisher LSD Method):
Comparisons for factor:

Comparison	Diff of Means	LSD(alpha=O.050)	P	D
pSVLCAT□β vs. pSVLCAT	108204.556	39287.229	<0.001	Yes
pSVLCAT□β vs. pSVLCAT□2a	63933.889	39287.229	0.002	Yes
pSVLCAT□β vs. pSVLCAT□2b	51278.556	39287.229	0.012	Yes
pSVLCAT□β vs. pSVLCAT□	38399.889	39287.229	0.055	No
pSVLCAT□β vs. pSVLCAT□2x	12504.667	39287.229	0.525	Do Not Test
pSVLCAT□2x vs. pSVLCAT	95699.889	39287.229	<0.001	Yes

pSVLCAT□2x vs. pSVLCAT□2a	51429.222	39287.229	0.011	Yes
pSVLCAT□2x vs. pSVLCAT□2b	38773.889	39287.229	0.053	No
pSVLCAT□2x vs. pSVLCAT□	25895.222	39287.229	0.191	Do Not Test
pSVLCAT□ vs. pSVLCAT	69804.667	39287.229	<0.001	Yes
pSVLCAT□ vs. pSVLCAT□2a	5534.000	39287.229	0.198	No
pSVLCAT□ vs. pSVLCAT□2b	12878.667	39287.229	0.513	Do Not Test
pSVLCAT□2b vs. pSVLCAT	56926.000	39287.229	0.005	Yes
pSVLCAT□2b vs. pSVLCAT□2a	12655.333	39287.229	0.520	Do Not Test
pSVLCAT□2a vs. pSVLCAT	44270.667	39287.229	0.028	Yes

A result of "Do Not Test" occurs for a comparison when no significant difference is found between two means that enclose that comparison. For example, if you had four means sorted in order, and found no difference between means 4 vs. 2, then would not test 4 vs. 3 and 3 vs. 2, but still test 4 vs. 1 and 3 vs. 1(4 vs. 3 and 3 vs. 2 are enclosed by 4 vs. 2: 4 3 2 1). Note not testing the enclosed means is a procedural rule, and a result of Do Not Test should be treated as if there is no significant difference between the means, even though one may appear to exist.

All Pairwise Multiple Comparison Procedures (Student-Newman-Keuls Method):
Comparisons for factor:

Comparison	Diff of Means	p	q	P	P<0.050
pSVLCAT□β vs. pSVLCAT	108204.556	6	7.831	<0.001	Yes
pSVLCAT□β vs. pSVLCAT□2a	63933.889	5	4.627	0.016	Yes
pSVLCAT□β vs. pSVLCAT□2b	51278.556	4	3.711	0.055	No
pSVLCAT□β vs. pSVLCAT□'	38399.889	3	2.779	0.132	Do Not Test
pSVLCAT□β vs. pSVLCAT□2x	12504.667	2	0.905	0.525	Do Not Test
pSVLCAT□2x vs. pSVLCAT	95699.889	5	6.926	<0.001	Yes
pSVLCAT□2x vs. pSVLCAT□2a	51429.222	4	3.722	0.054	No
pSVLCAT□2x vs. pSVLCAT□2b	38773.889	3	2.806	0.127	Do Not Test
pSVLCAT□2x vs. pSVLCAT□	25895.222	2	1.874	0.191	Do Not Test
pSVLCAT□ vs. pSVLCAT	69804.667	4	5.052	0.005	Yes
pSVL CAT□ vs. pSVLCAT□2a	25534.000	3	1.848	0.398	Do Not Test
pSVLCAT'□vs. pSVLCAT□2b	12878.667	2	0.932	0.513	Do Not Test
pSVLCAT□2b vs. pSVLCAT	56926.000	3	4.120	0.015	Yes
pSVLCAT□2b vs. pSVLCAT□2a	12655.333	2	0.916	0.520	Do Not Test
pSVLCAT□2a vs. pSVLCAT	44270.667	2	3.204	0.028	Yes

All Pairwise Multiple Comparison Procedures (Tukey Test):
Comparisons for factor:

Comparison	Diff of Means	p	q	P	P<0.050
pSVLCAT□β vs. pSVLCAT	108204.556	6	7.831	<0.001	Yes
pSVLCAT□β vs. pSVLCAT□2a	63933.889	6	4.627	0.023	Yes
pSVLCAT□β vs. pSVLCAT□2b	51278.556	6	3.711	0.111	No
pSVLCAT□β vs. pSVLCAT□'	38399.889	6	2.779	0.377	Do Not Test
pSVLCAT□β vs. pSVLCAT□2x	12504.667	6	0.905	0.987	Do Not Test
pSVLCAT□2x vs. pSVLCAT	95699.889	6	6.926	<0.001	Yes

pSVLCAT□2x vs. pSVLCAT□2a	51429.222	6	3.722	0.109	No
pSVLCAT□2x vs. pSVLCAT□2b	38773.889	6	2.806	0.366	Do Not Test
pSVLCAT□2x vs. pSVLCAT□	25895.222	6	1.874	0.770	Do Not Test
pSVLCAT□ vs. pSVLCAT	69804.667	6	5.052	0.010	Yes
pSVL CAT□ vs. pSVLCAT□2a	25534.000	6	1.848	0.780	Do Not Test
pSVLCAT'□vs. pSVLCAT□2b	12878.667	6	0.932	0.986	Do Not Test
pSVLCAT□2b vs. pSVLCAT	56926.000	6	4.120	0.057	No
pSVLCAT□2b vs. pSVLCAT□2a	12655.333	6	0.916	0.987	Do Not Test
pSVLCAT□2a vs. pSVLCAT	44270.667	6	3.204	0.228	Do Not Test

VII. Publications and presentations.

Papers

Myosin regulatory elements as vectors for gene transfer by intramuscular injection. Skarli, M; Kiri, A; Vrbova, G; Lee, CA; Goldspink, G. Gene Therapy, 1998. Vol. 5, (4), p 514-520.

Feeding induced stimulation of skeletal muscle protein synthesis following chronic partial starvation includes changes in gene transcription. Svanberg, E., Kiri, A., Goldspink, G., Lundholm, K. Accepted for publication by the Journal of Clinical Investigations.

In preparation: RNA-protein interactons of the skeletal myosin heavy chain 3' untranslated regions.

Abstracts

Mechanisms behind feeding induced stimulation of skeletal muscle protein synthesis following undernutrition. Svanberg, E., Kiri, A., Goldspink, G., Lundholm, K. European Society for Parenteral and Enteral Nutrition meeting September 1998

Does the 3' untranslated region of the myosin heavy chain mRNA control its translation?. Kiri, A., Skarli, M., Goldspink, G. Poster presentation, Keystone Symposium, 1996.

Muscle-specific promoters in gene-therapy of haemophilia and muscular dystrophy. Skarli, M.; Kiri, A.; Vrbova, G., Goldspink, G. Journal of Cellular Biochemistry 1995 S21A SIA, p 383. Meeting abstract, Keystone Symposium,

REFERENCES

ADESNIK, M., SALDITT, M., THOMAS, W. & DARNELL, J.E. (1972). Evidence that all messenger RNA molecules (except histone messenger RNA) contain Poly (A) sequences and that the Poly(A) has a nuclear function. *Journal of Molecular Biology* **71**, 21-30.

ADESNIK, M. & DARNELL, J.E. (1972). Biogenesis and characterization of histone messenger RNA in HeLa cells. *Journal of Molecular Biology* **67**, 397-406.

AFFARA, N.A., ROBERT, B., JACQUET, M., BUCKINGHAM, M.E. & GROS, F. (1980). Changes in gene expression during myogenic differentiation. I. Regulation of messenger RNA sequences expressed during myotube formation. *Journal of Molecular Biology* **140**, 441-458.

AFFARA, N.A., DAUBAS, P., WEYDERT, A. & GROS, F. (1980). Changes in gene expression during myogenic differentiation. II. Identification of the proteins encoded by myotube-specific complementary DNA sequences. *Journal of Molecular Biology* **140**, 459-470.

AHARON, T. & SCHNEIDER, R.J. (1993). Selective destabilization of short-lived mRNAs with the granulocyte-macrophage colony-stimulating factor AU-rich 3' noncoding region is mediated by a cotranslational mechanism. *Molecular & Cellular Biology* **13**, 1971-1980.

AHRINGER, J. & KIMBLE, J. (1991). Control of the sperm-oocyte switch in Caenorhabditis elegans hermaphrodites by the fem-3 3' untranslated region. *Nature* **349**, 346-348.

AIDA, K. & NEGISHI, M. (1991). Posttranscriptional regulation of coumarin 7-hydroxylase induction by xenobiotics in mouse liver: mRNA stabilization by pyrazole. *Biochemistry* **30**, 8041-8045.

ALBERTA, J.A., RUNDELL, K. & STILES, C.D. (1994). Identification of an activity that interacts with the 3'-untranslated region of c-myc mRNA and the role of its target sequence in mediating rapid mRNA degradation. *Journal of Biological Chemistry* **269**, 4532-4538.

ALLEN, R.W., TRACH, K.A., HOCH, J.A., KAWAMOTO, R.M. & CASWELL, A.H. (1987). Identification of the 37-kDa protein displaying a variable interaction with the erythroid cell membrane as glyceraldehyde-3-phosphate dehydrogenase.
Autophosphorylation of glyceraldehydephosphate dehydrogenase and phosphorylation of protein from skeletal muscle microsomes. *Journal of Biological Chemistry* **262**, 649-653.

ALTERMAN, R.B., GANGULY, S., SCHULZE, D.H., MARZLUFF, W.F., SCHILDKRAUT, C.L. & SKOULTCHI, A.I. (1984). Cell cycle regulation of mouse H3 histone mRNA metabolism. *Molecular & Cellular Biology* **4**, 123-132.

AMARA, F.M., CHEN, F.Y. & WRIGHT, J.A. (1995). Defining a novel cis element in the 3'-untranslated region of mammalian ribonucleotide reductase component R2 mRNA: role in transforming growth factor-beta 1 induced mRNA stabilization. *Nucleic Acids Research* **23**, 1461-1467.

AUSONI, S., GORZA, L., SCHIAFFINO, S., GUNDERSEN, K. & LOMO, T. (1990). Expression of myosin heavy chain isoforms in stimulated fast and slow rat muscles. *Journal of Neuroscience* **10**, 153-160.

AVIV, H., VOLOCH, Z., BASTOS, R. & LEVY, S. (1976). Biosynthesis and stability of globin mRNA in cultured erythroleukemic Friend cells. *Cell* **8**, 495-503.

BANDMAN, E. (1985). Myosin isoenzyme transitions in muscle development, maturation, and disease. [Review] [256 refs]. *International Review of Cytology* **97**, 97-131.

BANDMAN, E. (1992). Contractile protein isoforms in muscle development. [Review] [135 refs]. *Developmental Biology (Orlando)* **154**, 273-283.

BANGHAM, A.D. (1992). Liposomes: realizing their promise. [Review] [12 refs]. *Hospital Practice (Office Edition)* **27**, 51-56.

BARANY, M. (1967). ATPase activity of myosin correlated with speed of muscle shortening. *Journal of General Physiology* **50**, Suppl-218

BARKER, D.D., WANG, C., MOORE, J., DICKINSON, L.K. & LEHMANN, R. (1992). Pumilio is essential for function but not for distribution of the Drosophila abdominal determinant Nanos. *Genes & Development* **6**, 2312-2326.

BARTON, P.J. & BUCKINGHAM, M.E. (1985). The myosin alkali light chain proteins and their genes. [Review] [161 refs]. *Biochemical Journal* **231**, 249-261.

BEELMAN, C.A. & PARKER, R. (1995). Degradation of mRNA in eukaryotes. [Review] [50 refs]. *Cell* **81**, 179-183.

BERGERON, D., BARBEAU, B., LEGER, C. & RASSART, E. (1995). Experimental bias in the evaluation of the cellular transient expression in DNA co-transfection experiments. *Cellular & Molecular Biology Research* **41**, 155-159.

BERNSTEIN, P. & ROSS, J. (1989). Poly(A), poly(A) binding protein and the regulation of mRNA stability. [Review] [32 refs]. *Trends in Biochemical Sciences* **14**, 373-377.

BERNSTEIN, P., PELTZ, S.W. & ROSS, J. (1989). The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. *Molecular & Cellular Biology* **9**, 659-670.

BICKEL, M., IWAI, Y., PLUZNIK, D.H. & COHEN, R.B. (1992). Binding of sequence-specific proteins to the adenosine- plus uridine-rich sequences of the murine granulocyte/macrophage colony-stimulating factor mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 10001-10005.

BIENROTH, S., KELLER, W. & WAHLE, E. (1993). Assembly of a processive messenger RNA polyadenylation complex. *EMBO Journal* **12**, 585-594.

BIRNBOIM, H.C., MITCHEL, R.E. & STRAUS, N.A. (1973). Analysis of long pyrimidine polynucleotides in HeLa cell nuclear DNA: absence of polydeoxythymidylate. *Proceedings of the National Academy of Sciences of the United States of America* **70**, 2189-2192.

BLACK, B.L., LU, J. & OLSON, E.N. (1997). The MEF2A 3' untranslated region functions as a cis-acting translational repressor. *Molecular & Cellular Biology* **17**, 2756-2763.

BLAU, H.M., WEBSTER, S.G. & PAVLATH, G.K. (1983). Defective myoblasts identified in Duschenne Muscular Dystrophy. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 4856-4860.

BOELENS, W.C., JANSEN, E.J., VAN, V.W., STRIPECKE, R., MATTAJ, I.W. & GUNDERSON, S.I. (1993). The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell* **72**, 881-892.

BOHJANEN, P.R., PETRYNIAK, B., JUNE, C.H., THOMPSON, C.B. & LINDSTEN, T. (1991). An inducible cytoplasmic factor (AU-B) binds selectively to AUUUA multimers in the 3' untranslated region of lymphokine mRNA. *Molecular & Cellular Biology* **11**, 3288-3295.

BOHJANEN, P.R., PETRYNIAK, B., JUNE, C.H., THOMPSON, C.B. & LINDSTEN, T. (1992). AU RNA-binding factors differ in their binding specificities and affinities. *Journal of Biological Chemistry* **267**, 6302-6309.

BOTTINELLI, R., SCHIAFFINO, S. & REGGIANI, C. (1991). Force-velocity relations and myosin heavy chain isoform compositions of skinned fibres from rat skeletal muscle. *Journal of Physiology* **437**, 655-672.

BOUDRIAU, S., VINCENT, M., COTE, C.H. & ROGERS, P.A. (1993). Cytoskeletal structure of skeletal muscle: identification of an intricate exosarcomeric microtubule lattice in slow- and fast-twitch muscle fibers. *Journal of Histochemistry & Cytochemistry* **41**, 1013-1021.

BRAUN, R.E., PESCHON, J.J., BEHRINGER, R.R., BRINSTER, R.L. & PALMITER, R.D. (1989). Protamine 3'-untranslated sequences regulate temporal translational control and subcellular localization of growth hormone in spermatids of transgenic mice. *Genes & Development* **3**, 793-802.

BRAUN, R.E. (1990). Temporal translational regulation of the protamine 1 gene during mouse spermatogenesis. *Enzyme* **44**, 120-128.

BRAWERMAN, G. (1981). The Role of the poly(A) sequence in mammalian messenger RNA. [Review] [92 refs]. *CRC Critical Reviews in Biochemistry* **10**, 1-38.

BREWER, G. (1991). An A + U-rich element RNA-binding factor regulates c-myc mRNA stability in vitro. *Molecular & Cellular Biology* **11**, 2460-2466.

BUCKINGHAM, M. (1996). Skeletal muscle development and the role of the myogenic regulatory factors. [Review] [27 refs]. *Biochemical Society Transactions* **24**, 506-509.

BULLER, A.J., ECCLES, J.C. & ECCLES, R.M. (1960). Interaction Between Motorneurones and Muscles in Respect of the Characteristic Speeds of Their Contraction. *Journal of Physiology* **150**, 417-439.

BURD, C.G. & DREYFUSS, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. [Review] [97 refs]. *Science* **265**, 615-621.

BUTLER-BROWNE, G.S., HERLICOVIEZ, D. & WHALEN, R.G. (1984). Effects of hypothyroidism on myosin isozyme transitions in developing rat muscle. *FEBS Letters* **166**, 71-75.

BUTLER-BROWNE, G.S. & WHALEN, R.G. (1984). Myosin isozyme transitions occurring during the postnatal development of the rat soleus muscle. *Developmental Biology (Orlando)* **102**, 324-334.

BUTLER-BROWNE, G.S., ERIKSSON, P.O., LAURENT, C. & THORNELL, L.E. (1988). Adult human masseter muscle fibers express myosin isozymes characteristic of development. *Muscle & Nerve* **11**, 610-620.

CALNAN, B.J., TIDOR, B., BIANCALANA, S., HUDSON, D. & FRANKEL, A.D. (1991). Arginine-mediated RNA recognition: the arginine fork [published erratum appears in Science 1992 Feb 7;255(5045):665]. *Science* **252**, 1167-1171.

CAMPIONE, M., AUSONI, S., GUEZENNEC, C.Y. & SCHIAFFINO, S. (1993). Myosin and troponin changes in rat soleus muscle after hindlimb suspension. *Journal of Applied Physiology* **74**, 1156-1160.

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

CAPUT, D., BEUTLER, B., HARTOG, K., THAYER, R., BROWN-SHIMER, S. & CERAMI, A. (1986). Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 1670-1674.

CASEY, J.L., HENTZE, M.W., KOELLER, D.M., CAUGHMAN, S.W., ROUAULT, T.A., KLAUSNER, RD & HARFORD, J.B. (1988). Iron-responsive elements: regulatory RNA sequences that control mRNA levels and translation. *Science* **240**, 924-928.

CASEY, J.L., KOELLER, D.M., RAMIN, V.C., KLAUSNER, R.D. & HARFORD, J.B. (1989). Iron regulation of transferrin receptor mRNA levels requires iron-responsive elements and a rapid turnover determinant in the 3' untranslated region of the mRNA. *EMBO Journal* **8**, 3693-3699.

CHEN, C.Y. & SHYU, A.B. (1994). Selective degradation of early-response-gene mRNAs: functional analyses of sequence features of the AU-rich elements. *Molecular & Cellular Biology* **14**, 8471-8482.

CHEN, F.Y., AMARA, F.M. & WRIGHT, J.A. (1993). Mammalian ribonucleotide reductase R1 mRNA stability under normal and phorbol ester stimulating conditions: involvement of a cis-trans interaction at the 3' untranslated region. *EMBO Journal* **12**, 3977-3986.

CHOMCZYNSKI, P., SACCHI, N., (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* **162**, 156-159.

CLARK, J.M. (1988). Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Research* **16**, 9677-9686.

CONDON, K., SILBERSTEIN, L., BLAU, H.M. & THOMPSON, W.J. (1990). Development of muscle fiber types in the prenatal rat hindlimb. *Developmental Biology (Orlando)* **138**, 256-274.

COX, R.D., WEYDERT, A., BARLOW, D. & BUCKINGHAM, M.E. (1991). Three linked myosin heavy chain genes clustered within 370 kb of each other show independent transcriptional and post-transcriptional regulation during differentiation of a mouse muscle cell line. *Developmental Biology (Orlando)* **143**, 36-43.

COX, R.D. & BUCKINGHAM, M.E. (1992). Actin and myosin genes are transcriptionally regulated during mouse skeletal muscle development. *Developmental Biology (Orlando)* **149**, 228-234.

CRABB, D.W. & DIXON, J.E. (1987). A method for increasing the sensitivity of chloramphenicol acetyltransferase assays in extracts of transfected cultured cells. *Analytical Biochemistry* **163**, 88-92.

CRAIG, R. (1994). The structure of the contractile filaments. In *Myology*, eds. ENGEL, A.G. & BARKER, D.D., pp. 73-123. McGraw-Hill, New York.

CRAWFORD, D.R. & RICHTER, J.D. (1987). An RNA-binding protein from Xenopus oocytes is associated with specific message sequences. *Development* **101**, 741-749.

CRIPE, L., MORRIS, E. & FULTON, A.B. (1993). Vimentin mRNA location changes during muscle development. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 2724-2728.

CROWLEY, C.W., LIU, C.C. & LEVINSON, A.D. (1983). Plasmid-directed synthesis of hepatitis B surface antigen in monkey cells. *Molecular & Cellular Biology* **3**, 44-55.

CURTIS, D., LEHMANN, R. & ZAMORE, P.D. (1995). Translational regulation in development. [Review] [70 refs]. *Cell* **81**, 171-178.

CZYZYK-KRZESKA, M.F., PAULDING, W.R., BERESH, J.E. & KROLL, S.L. (1997). Post-transcriptional regulation of tyrosine hydroxylase gene expression by oxygen in PC12 cells. *Kidney International* **51**, 585-590.

D'ALBIS, A., ANGER, M. & LOMPRE, A.M. (1993). Rabbit masseter expresses the cardiac alpha myosin heavy chain gene. Evidence from mRNA sequence analysis. *FEBS Letters* **324**, 178-180.

DANOWSKI, B.A., IMANAKA-YOSHIDA, K., SANGER, J.M. & SANGER, J.W. (1992). Costameres are sites of force transmission to the substratum in adult rat cardiomyocytes. *Journal of Cell Biology* **118**, 1411-1420.

DARNELL, J.E., Jr. (1979). Transcription units for mRNA production in eukaryotic cells and their DNA viruses. *Progress in Nucleic Acid Research & Molecular Biology* **22**, 327-353.

DARNELL, J.E., Jr. (1982). Variety in the level of gene control in eukaryotic cells. [Review] [151 refs]. *Nature* **297**, 365-371.

DAVIS, I. & ISH-HOROWICZ, D. (1991). Apical localization of pair-rule transcripts requires 3' sequences and limits protein diffusion in the Drosophila blastoderm embryo. *Cell* **67**, 927-940.

DECKER, C.J. & PARKER, R. (1993). A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes & Development* 7, 1632-1643.

DECKER, C.J. & PARKER, R. (1994). Mechanisms of mRNA degradation in eukaryotes. [Review] [46 refs]. *Trends in Biochemical Sciences* **19**, 336-340.

DENARDI, C., AUSONI, S., MORETTI, P., GORZA, L., VELLECA, M., BUCKINGHAM, M. & SCHIAFFINO, S. (1993). Type 2X-myosin heavy chain is coded by a muscle fiber type-specific and developmentally regulated gene. *Journal of Cell Biology* **123**, 823-835.

DENARDI, C., AUSONI, S., MORETTI, P., GORZA, L., VELLECA, M., BUCKINGHAM, M. & SCHIAFFINO, S. (1993). Type 2X-myosin heavy chain is coded by a muscle fiber type-specific and developmentally regulated gene. *Journal of Cell Biology* **123**, 823-835.

DENT, C.L. & LATCHMAN, D. (1993). *Transcription Factors: A Practical Approach*. IRL Press, Oxford, UK.

DIAS, P., DILLING, M. & HOUGHTON, P. (1994). The molecular basis of skeletal muscle differentiation. [Review] [79 refs]. *Seminars in Diagnostic Pathology* **11**, 3-14.

DIGNAM, J.D., LEBOVITZ, R.M. & ROEDER, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Research* **11**, 1475-1489.

DIX, D. J., EISENBERG, B. R. 1990. Myosin mRNA accumulation and myofibrillogenesis at the myotendinous junction of stretched muscle fibers. *Journal of Cell Biology* **111**, 1885-1894

DOETSCHMAN, T.C., DYM, H.P., SIEGEL, E.J. & HEYWOOD, S.M. (1980). Myoblast stored myosin heavy chain transcripts are precursors to the myotube polysomal myosin heavy chain mRNAs. *Differentiation* **16**, 149-162.

DRAEGER, A., WEEDS, A.G. & FITZSIMONS, R.B. (1987). Primary, secondary and tertiary myotubes in developing skeletal muscle: a new approach to the analysis of human myogenesis. *Journal of the Neurological Sciences* **81**, 19-43.

DREYFUSS, G., CHOI, Y.D. & ADAM, S.A. (1984). Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. *Molecular & Cellular Biology* **4**, 1104-1114.

DREYFUSS, G., MATUNIS, M.J., PINOL-ROMA, S. & BURD, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. [Review] [207 refs]. *Annual Review of Biochemistry* **62**, 289-321.

DRIEVER, W. & NUSSLEIN-VOLHARD, C. (1988). A gradient of bicoid protein in Drosophila embryos. *Cell* **54** , 83-93.

DRIEVER, W. & NUSSLEIN-VOLHARD, C. (1988). The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell* **54**, 95-104.

DUMAS MILNE EDWARDS JB, VALDENAIRE, O. & MALLET, J. (1997). Anchoring a defined sequence to the 5' ends of mRNAs. The bolt to clone rare full-length mRNAs. *Methods in Molecular Biology* **67**, 261-278.

DURET, L., DORKELD, F. & GAUTIER, C. (1993). Strong conservation of non-coding sequences during vertebrates evolution: potential involvement in post-transcriptional regulation of gene expression. *Nucleic Acids Research* **21**, 2315-2322.

DYM, H.P., KENNEDY, D.S. & HEYWOOD, S.M. (1979). Sub-cellular distribution of the cytoplasmic myosin heavy chain mRNA during myogenesis. *Differentiation* **12**, 145-155.

ECUYER, T.J., TOMPACH, P.C., MORRIS, E. & FULTON, A.B. (1995). Transdifferentiation of chicken embryonic cells into muscle cells by the 3' untranslated region of muscle tropomyosin. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 7520-7524.

EDMONDS, M., VAUGHAN, M.H.J. & NAKAZATO, H. (1971). Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship. *Proceedings of the National Academy of Sciences of the United States of America* **68**, 1336-1340.

EMERSON, C.P.J. & BERNSTEIN, S.I. (1987). Molecular genetics of myosin. [Review] [232 refs]. *Annual Review of Biochemistry* **56**, 695-726.

ENNION, S., SANT, SARGEANT, A.J., YOUNG, A. & GOLDSPINK, G. (1995). Characterization of human skeletal muscle fibres according to the myosin heavy chains they express. *Journal of Muscle Research & Cell Motility* **16**, 35-43.

ENNION, S., GAUVRY, L., BUTTERWORTH, P., GOLDSPINK, G., (1995). Small-diameter white myotomal muscle fibres associated with growth hyperplasia in the carp (Cyprinus carpio) express a distinct myosin heavy chain gene. *Journal of Experimental Biology* **198**, 1603-1611.

EPHRUSSI, A., DICKINSON, L.K. & LEHMANN, R. (1991). Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* **66**, 37-50.

FAJARDO, M.A., BUTNER, K.A., LEE, K. & BRAUN, R.E. (1994). Germ cell-specific proteins interact with the 3' untranslated regions of Prm-1 and Prm-2 mRNA. *Developmental Biology (Orlando)* **166**, 643-653.

FARR, A. & ROMAN, A. (1992). A pitfall of using a second plasmid to determine transfection efficiency. *Nucleic Acids Research* **20**, 920

FEINBERG, A.P. & VOGELSTEIN, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**, 6-13.

FEINBERG, A.P. & VOGELSTEIN, B. (1984). "A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity". Addendum. *Analytical Biochemistry* **137**, 266-267.

FELGNER, P.L., GADEK, T.R., HOLM, M., ROMAN, R., CHAN, H.W., WENZ, M., NORTHROP, J.P., RINGOLD, G.M. & DANIELSEN, M. (1987). Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 7413-7417.

FENG, P., OHLSSON, M. & NY, T. (1990). The structure of the TATA-less rat tissue-type plasminogen activator gene. Species-specific sequence divergences in the promoter predict differences in regulation of gene expression. *Journal of Biological Chemistry* **265**, 2022-2027.

FLEMING, J., THIELE, B.J., CHESTER, J., O'PREY, J., JANETZKI, S., AITKEN, A., ANTON, I.A., RAPOPORT, S.M. & HARRISON, P.R. (1989). The complete sequence of the rabbit erythroid cell-specific 15-lipoxygenase mRNA: comparison of the predicted amino acid sequence of the erythrocyte lipoxygenase with other lipoxygenases. *Gene* **79**, 181-188.

FORD, J.P. & HSU, M.T. (1978). Transcription pattern of in vivo-labeled late simian virus 40 RNA: equimolar transcription beyond the mRNA 3' terminus. *Journal of Virology* **28**, 795-801.

FOX, C.A., SHEETS, M.D. & WICKENS, M.P. (1989). Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. *Genes & Development* **3**, 2151-2162.

FRIEND, D.S., PAPAHADJOPOULOS, D. & DEBS, R.J. (1996). Endocytosis and intracellular processing accompanying transfection mediated by cationic liposomes. *Biochimica et Biophysica Acta* **1278**, 41-50.

FROHMAN, M.A., DUSH, M.K. & MARTIN, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 8998-9002.

FU, L., MINDEN, M.D. & BENCHIMOL, S. (1996). Translational regulation of human p53 gene expression. *EMBO Journal* **15**, 4392-4401.

FU, L. & BENCHIMOL, S. (1997). Participation of the human p53 3'UTR in translational repression and activation following gamma-irradiation. *EMBO Journal* **16**, 4117-4125.

GAMBKE, B., LYONS, G.E., HASELGROVE, J., KELLY, A.M. & RUBINSTEIN, N.A. (1983). Thyroidal and neural control of myosin transitions during development of rat fast and slow muscles. *FEBS Letters* **156**, 335-339.

GAUTHIER (1994). Skeletal muscle fibre types. In *Myology*, eds. ENGEL, A.G. & BANKER, B.G., pp. 255-283. McGraw-Hill, New York.

GAVIS, E.R. & LEHMANN, R. (1992). Localization of nanos RNA controls embryonic polarity. *Cell* **71**, 301-313.

GAVIS, E.R. & LEHMANN, R. (1994). Translational regulation of nanos by RNA localization. *Nature* **369**, 315-318.

GAVIS, E.R., CURTIS, D. & LEHMANN, R. (1996). Identification of cis-acting sequences that control nanos RNA localization. *Developmental Biology (Orlando)* **176**, 36-50.

GAVIS, E.R., LUNSFORD, L., BERGSTEN, S.E. & LEHMANN, R. (1996). A conserved 90 nucleotide element mediates translational repression of nanos RNA. *Development* **122**, 2791-2800.

GENESTE, O., RAFFALLI, F. & LANG, M.A. (1996). Identification and characterization of a 44 kDa protein that binds specifically to the 3'-untranslated region of CYP2a5 mRNA: inducibility, subcellular distribution and possible role in mRNA stabilization. *Biochemical Journal* **313**, 1029-1037.

GEOGHEGAN, T.E. & MCCOY, L. (1986). Biogenesis and cell cycle relationship of poly(A)- actin mRNA in mouse ascites cells. *Experimental Cell Research* **162**, 175-182.

GIL, A. & PROUDFOOT, N.J. (1984). A sequence downstream of AAUAAA is required for rabbit beta-globin mRNA 3'-end formation. *Nature* **312**, 473-474.

GIL, A., PROUDFOOT, N.J., JENNY, A., HAURI, H.P. & KELLER, W. (1987). Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation. Characterization of cleavage and polyadenylation specificity factor and cloning of its 100-kilodalton subunit. *Cell* **49**, 399-406.

GILLIS, P. & MALTER, J.S. (1991). The adenosine-uridine binding factor recognizes the AU-rich elements of cytokine, lymphokine, and oncogene mRNAs. *Journal of Biological Chemistry* **266**, 3172-3177.

GOLDSPINK, P., SHARP, W. & RUSSELL, B. (1997). Localization of cardiac (alpha)-myosin heavy chain mRNA is regulated by its 3' untranslated region via mechanical activity and translational block. *Journal of Cell Science* **110**, 2969-2978.

GOODWIN, E.B., OKKEMA, P.G., EVANS, T.C. & KIMBLE, J. (1993). Translational regulation of tra-2 by its 3' untranslated region controls sexual identity in C. elegans. *Cell* **75**, 329-339.

GORLACH, M., BURD, C.G., PORTMAN, D.S. & DREYFUSS, G. (1993). The hnRNP proteins. [Review] [70 refs]. *Molecular Biology Reports* **18**, 73-78.

GORMAN, C.M., MOFFAT, L.F. & HOWARD, B.H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Molecular & Cellular Biology* **2**, 1044-1051.

GOROSPE, M. & BAGLIONI, C. (1994). Degradation of unstable interleukin-1 alpha mRNA in a rabbit reticulocyte cell-free system. Localization of an instability determinant to a cluster of AUUUA motifs. *Journal of Biological Chemistry* **269**, 11845-11851.

GOTTLIEB, E. (1992). The 3' untranslated region of localized maternal messages contains a conserved motif involved in mRNA localization. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 7164-7168.

GRAFI, G., SELA, I. & GALILI, G. (1993). Translational regulation of human beta interferon mRNA: association of the 3' AU-rich sequence with the poly(A) tail reduces translation efficiency in vitro. *Molecular & Cellular Biology* **13**, 3487-3493.

GRAHAM, F.L. & EB AJ VAN DER. (1973). Transformation of rat cells by DNA of human adenovirus 5. *Virology* **54**, 536-539.

GREEN, M.R. (1991). Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. [Review] [221 refs]. *Annual Review of Cell Biology* 7, 559-599.

GREENBERG, J.R. & PERRY, R.P. (1972). The isolation and characterization of steady-state labeled messenger RNA from L-cells. *Biochimica et Biophysica Acta* **287**, 361-366.

GREENBERG, J.R. & PERRY, R.P. (1972). Relative occurrence of polyadenylic acid sequences in messenger and heterogeneous nuclear RNA of L cells as determined by poly (U)-hydroxylapatite chromatography. *Journal of Molecular Biology* **72**, 91-98.

GROSKREUTZ, D. & SCHENBORN, E.T. (1997). Reporter systems. [Review] [88 refs]. *Methods in Molecular Biology* **63**, 11-30.

GROSSE, F., NASHEUER, H.P., SCHOLTISSEK, S. & SCHOMBURG, U. (1986). Lactate dehydrogenase and glyceraldehyde-phosphate dehydrogenase are single-stranded DNA-binding proteins that affect the DNA-polymerase-alpha-primase complex. *European Journal of Biochemistry* **160**, 459-467.

GULICK, J., SUBRAMANIAM, A., NEUMANN, J. & ROBBINS, J. (1991). Isolation and characterization of the mouse cardiac myosin heavy chain genes. *Journal of Biological Chemistry* **266**, 9180-9185.

GUNDERSON, S.I., BEYER, K., MARTIN, G., KELLER, W., BOELENS, W.C. & MATTAJ, L.W. (1994). The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell* **76**, 531-541.

GUNNING, P. & HARDEMAN, E. (1991). Multiple mechanisms regulate muscle fiber diversity. [Review] [54 refs]. *FASEB Journal* **5**, 3064-3070.

GUTHRIE, C. (1991). Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. [Review] [83 refs]. *Science* **253**, 157-163.

HAHNEMANN, B., SALONPAA, P., PASANEN, M., MAENPAA, J., HONKAKOSKI, P., JUVONEN, LANG, M.A., PELKONEN, O. & RAUNIO, H. (1992). Effect of pyrazole, cobalt and phenobarbital on mouse liver cytochrome P-450 2a-4/5 (Cyp2a-4/5) expression. *Biochemical Journal* **286**, 289-294.

HAILSTONES, D.L. & GUNNING, P.W. (1990). Characterization of human myosin light chains 1sa and 3nm: implications for isoform evolution and function. *Molecular & Cellular Biology* **10**, 1095-1104.

HAKE, L.E. & RICHTER, J.D. (1994). CPEB is a specificty factor that mediates cytoplasmic polyadenylation during oocyte maturation. *Cell* **79**, 617-627.

HALL, C.V., JACOB, P.E., RINGOLD, G.M. & LEE, F. (1983). Expression and regulation of Escherichia coli lacZ gene fusions in mammalian cells. *Journal of Molecular & Applied Genetics* **2**, 101-109.

HAN, J., BROWN, T. & BEUTLER, B. (1990). Endotoxin-responsive sequences control cachectin/tumor necrosis factor biosynthesis at the translational level [published erratum appears in J Exp Med 1990 Mar 1;171(3):971-2]. *Journal of Experimental Medicine* **171**, 465-475.

HARRIS, A.J., FITZSIMONS, R.B. & MCEWAN, J.C. (1989). Neural control of the sequence of expression of myosin heavy chain isoforms in foetal mammalian muscles. *Development* **107**, 751-769.

HARVEY, R.J., VREUGDENHIL, E., ZAMAN, S.H., BHANDAL, N.S., USHERWOOD, P.N., BARNARD, EA & DARLISON, M.G. (1991). Sequence of a functional invertebrate GABAA receptor subunit which can form a chimeric receptor with a vertebrate alpha subunit. *EMBO Journal* **10**, 3239-3245.

HEINTZ, N., SIVE, H.L. & ROEDER, R.G. (1983). Regulation of human histone gene expression: kinetics of accumulation and changes in the rate of synthesis and in the half-lives of individual histone mRNAs during the HeLa cell cycle. *Molecular & Cellular Biology* **3**, 539-550.

HENIKOFF, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**, 351-359.

HENIKOFF, S. (1987). Unidirectional digestion with exonuclease III in DNA sequence analysis. *Methods in Enzymology* **155**, 156-165.

HERBOMEL, P., BOURACHOT, B. & YANIV, M. (1984). Two distinct enhancers with different cell specificities coexist in the regulatory region of polyoma. *Cell* **39**, 653-662.

HESKETH, J.E. (1996). Sorting of messenger RNAs in the cytoplasm: mRNA localization and the cytoskeleton. [Review] [123 refs]. *Experimental Cell Research* **225**, 219-236.

HESKETH, J.E. (1996). mRNA targeting: signals in the 3'-untranslated sequences for sorting of some mRNAs. [Review] [55 refs]. *Biochemical Society Transactions* **24**, 521-527.

HESKETH, J., CAMPBELL, G., LOVERIDGE, N., RUSSELL, B., DIX, D. J. (1991).Myosin heavy-chain mRNA is present in both myofibrillar and subsarcolemmal regions of muscle fibresMechanisms for intracellular distribution of mRNA: in situ hybridization studies in muscle. *Biochemical Journal* **279**, 309-310.

HILL, M. A., GUNNING, P. (1993). Beta and gamma actin mRNAs are differentially l ocated within myoblasts. *Journal of Cell Biology* **122**, 825-832.

HILL, M. A., SCHEDLICH, L., GUNNING, P. (1994). Serum-induced signal transduction determines the peripheral location of beta-actin mRNA within the cell. *Journal of Cell Biology* **126**, 1221-1229.

HOHEISEL, J.D. (1993). On the activities of Escherichia coli exonuclease III. *Analytical Biochemistry* **209**, 238-246.

HOLCIK, M. & LIEBHABER, S.A. (1997). Four highly stable eukaryotic mRNAs assemble 3' untranslated region RNA-protein complexes sharing cis and trans components. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 2410-2414.

HOOCK, T. C., NEWCOMB, P. M., HERMAN, I. M. (1991). Beta actin and its mRNA are localized at the plasma membrane and the regions of moving cytoplasm during the cellular response to injury. *Journal of Cell Biology* **112**, 653-664.HIGGINS, D.G. & SHARP, P.M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**, 237-244.

HOOK, A.G. & KELLEMS, R.E. (1988). Localization and sequence analysis of poly(A) sites generating multiple dihydrofolate reductase mRNAs. *Journal of Biological Chemistry* **263**, 2337-2343.

HUARTE, J., BELIN, D. & VASSALLI, J.D. (1985). Plasminogen activator in mouse and rat oocytes: induction during meiotic maturation. *Cell* **43**, 551-558.

HUARTE, J., STUTZ, A., O'CONNELL, M.L., GUBLER, P., BELIN, D., DARROW, A.L., STRICKLAND, S. & VASSALLI, J.D. (1992). Transient translational silencing by reversible mRNA deadenylation. *Cell* **69**, 1021-1030.

HUGHES, S.M., TAYLOR, J.M., TAPSCOTT, S.J., GURLEY, C.M., CARTER, W.J. & PETERSON, C.A. (1993). Selective accumulation of MyoD and myogenin mRNAs in fast and slow adult skeletal muscle is controlled by innervation and hormones. *Development* **118**, 1137-1147.

HUITOREL, P. & PANTALONI, D. (1985). Bundling of microtubules by glyceraldehyde-3-phosphate dehydrogenase and its modulation by ATP. *European Journal of Biochemistry* **150**, 265-269.

HUNT, T. (1989). On the translational control of suicide in red cell development [see comments]. *Trends in Biochemical Sciences* **14**, 393-394.

HUXLEY, A.F. & NIEDERGERKE, R. (1954). Structural changes in muscle during contraction. Interference microscopy of living muscle firbres. *Nature* **173**, 971-973.

HUXLEY, H.E. & HANSON, J. (1954). Changes in the cross-striations of muscle during contraction and stretch an their structural interpretation. *Nature* **173**, 973

INNIS, M. A., GELFAND, M. A, SNINSKY, J. J., AND WHITE, T. J. eds. PCR protocols: A Guide to Methods & Applications. Academic Press, Inc. San Diego,CA. 1990. ()

IWAI, Y., BICKEL, M., PLUZNIK, D.H. & COHEN, R.B. (1991). Identification of sequences within the murine granulocyte-macrophage colony-stimulating factor mRNA 3'-untranslated region that mediate mRNA stabilization induced by mitogen treatment of EL-4 thymoma cells. *Journal of Biological Chemistry* **266**, 17959-17965.

IZUMO, S., NADAL-GINARD, B. & MAHDAVI, V. (1986). All members of the MHC multigene family respond to thyroid hormone in a highly tissue-specific manner. *Science* **231**, 597-600.

JACKSON, R.J. & STANDART, N. (1990). Do the poly A tail and 3' untranlsated region control mRNA translation? *Cell* **62**, 15-24.

JACOBSON, A. & PELTZ, S.W. (1996). Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. [Review] [400 refs]. *Annual Review of Biochemistry* **65**, 693-739.

JAEGER, J.A., TURNER, D.H. & ZUKER, M. (1989). Improved predictions of secondary structures for RNA. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 7706-7710.

JEFFERY, W.R., TOMLINSON, C.R. & BRODEUR, R.D. (1983). Localization of actin messenger RNA during early ascidian development. *Developmental Biology (Orlando)* **99**, 408-417.

JONES, T.R., COLE, M.D., DEPARTMENT OF MOLECULAR BIOLOGY, L.T.L.P. & UNIVERSITY, N.J. (1987). Rapid cytoplasmic turnover of c-myc mRNA: requirement of the 3' untranslated sequences. *Molecular & Cellular Biology* 7, 4513-4521.

KANO, Y., FUJIMAKI, N. & ISHIKAWA, H. (1991). The distribution and arrangement of microtubules in mammalian skeletal muscle fibers. *Cell Structure & Function* **16**, 251-261.

KATOH, T. & LOWEY, S. (1989). Mapping myosin light chains by immunoelectron microscopy. Use of anti-fluorescyl antibodies as structural probes. *Journal of Cell Biology* **109**, 1549-1560.

KATZ, D.A., THEODORAKIS, N.G., CLEVELAND, D.W., LINDSTEN, T. & THOMPSON, C.B. (1994). AU-A, an RNA-binding activity distinct from hnRNP A1, is selective for AUUUA repeats and shuttles between the nucleus and the cytoplasm. *Nucleic Acids Research* **22**, 238-246.

KAVINSKY, C.J., UMEDA, P.K., SINHA, A.M., ELZINGA, M., TONG, S.W., ZAK, R., JAKOVCIC & RABINOWITZ, M. (1983). Cloned mRNA sequences for two types of embryonic myosin heavy chains from chick skeletal muscle. I. DNA and derived amino acid sequence of light meromyosin. *Journal of Biological Chemistry* **258**, 5196-5205.

KAWAMOTO, R. M., CASWELL, A. H. (1986). Autophosphorylation of glyceraldehydephosphate dehydrogenase and phosphorylation of protein from skeletal muscle microsomes. *Biochemistry* **25**, 657-661.

KEENE, J.D.a.Q.C.C. (1991). Nuclear RNA-binding proteins. *Progress in Nucleic Acid Research & Molecular Biology* **41**, 179-202.

KELLER, W. (1995). No end yet to messenger RNA 3' processing:. [Review] [29 refs]. *Cell* **81**, 829-832.

KILEDJIAN, M., WANG, X. & LIEBHABER, S.A. (1995). Identification of two KH domain proteins in the alpha-globin mRNP stability complex. *EMBO Journal* **14**, 4357-4364.

KILEDJIAN, M., DEMARIA, C.T., BREWER, G. & NOVICK, K. (1997). Identification of AUF1 (hetergenous nulcear ribonuicleoprotein D) as a component of the alpha-globin mRNA stability complex. *Molecular & Cellular Biology* **17**, 4870-4876.

KISLAUSKIS, E.H., LI, Z., SINGER, R.H. & TANEJA, K.L. (1993). Isoform-specific 3'-untranslated sequences sort alpha-cardiac and beta-cytoplasmic actin messenger RNAs to different cytoplasmic compartments [published erratum appears in J Cell Biol 1993 Dec;123(6 Pt 2):following 1907]. *Journal of Cell Biology* **123**, 165-172.

KISLAUSKIS, E.H., ZHU, X. & SINGER, R.H. (1994). Sequences responsible for intracellular localization of beta-actin messenger RNA also affect cell phenotype. *Journal of Cell Biology* **127**, 441-451.

KLEENE, K.C., DISTEL, R.J. & HECHT, N.B. (1984). Translational regulation and deadenylation of a protamine mRNA during spermiogenesis in the mouse. *Developmental Biology (Orlando)* **105**, 71-79.

KOZAK, M. (1992). Regulation of translation in eukaryotic systems. *Annual Review of Cell Biology* **8**, 197-225.

KRIEG, P. A., MELTON, D. A. (1987). In vitro RNA synthesis with SP6 RNA polymerase. *Methods in Enzymology* **155** 397-415.

KROWCZYNSKA, A., YENOFSKY, R. & BRAWERMAN, G. (1985). Regulation of messenger RNA stability in mouse erythroleukemia cells. *Journal of Molecular Biology* **181**, 231-239.

KRUYS, V., WATHELET, M., POUPART, P., CONTRERAS, R., FIERS, W., CONTENT, J. & HUEZ, G. (1987). The 3' untranslated region of the human interferon-beta mRNA has an inhibitory effect on translation. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 6030-6034.

KRUYS, V., MARINX, O., SHAW, G., DESCHAMPS, J. & HUEZ, G. (1989). Translational blockade imposed by cytokine-derived UA-rich sequences. *Science* **245**, 852-855.

KRUYS, V.I., WATHELET, M.G. & HUEZ, G.A. (1988). Identification of a translation inhibitory element (TIE) in the 3' untranslated region of the human interferon-beta mRNA. *Gene* **72**, 191-200.

KWON, Y.K. & HECHT, N.B. (1991). Cytoplasmic protein binding to highly conserved sequences in the 3' untranslated region of mouse protamine 2 mRNA, a translationally regulated transcript of male germ cells. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 3584-3588.

KWON, Y.K. & HECHT, N.B. (1993). Binding of a phosphoprotein to the 3' untranslated region of the mouse protamine 2 mRNA temporally represses its translation. *Molecular & Cellular Biology* **13**, 6547-6557.

LAEMMLI, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

LAGNADO, C.A., BROWN, C.Y. & GOODALL, G.J. (1994). AUUUA is not sufficient to promote poly(A) shortening and degradation of an mRNA: the functional sequence within AU-rich elements may be UUAUUUA(U/A)(U/A). *Molecular & Cellular Biology* **14**, 7984-7995.

LARSON, P. F., HUDGSON, P., WALTON, J. N. 1969. Morphological relationship of polyribosomes and myosin filaments in deeloping and regenerating skeletal muscle. *Nature* **222**, 1168-1169.

LARSSON, L., BIRAL, D., CAMPIONE, M. & SCHIAFFINO, S. (1993). An age-related type IIB to IIX myosin heavy chain switching in rat skeletal muscle. *Acta Physiologica Scandinavica* **147**, 227-234.

LATHAM, V.M.J., KISLAUSKIS, E.H., SINGER, R.H. & ROSS, A.F. (1994). Beta-actin mRNA localization is regulated by signal transduction mechanisms. *Journal of Cell Biology* **126**, 1211-1219.

LAWRENCE, J.B. & SINGER, R.H. (1986). Intracellular localization of messenger RNAs for cytoskeletal proteins. *Cell* **45**, 407-415.

LEE, K.A., BINDEREIF, A. & GREEN, M.R. (1988). A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing. *Gene Analysis Techniques* **5**, 22-31.

LEE, S.Y., MENDECKI, J. & BRAWERMAN, G. (1971). A polynucleotide segment rich in adenylic acid in the rapidly-labeled polyribosomal RNA component of mouse sarcoma 180 ascites cells. *Proceedings of the National Academy of Sciences of the United States of America* **68**, 1331-1335.

LEHMANN, R. & NUSSLEIN-VOLHARD, C. (1991). The maternal gene nanos has a central role in posterior pattern formation of the Drosophila embryo. *Development* **112**, 679-691.

LEHRACH, H., DIAMOND, D., WOZNEY, J.M. & BOEDTKER, H. (1977). RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry* **16**, 4743-4751.

LEIBOVITCH, M.P., TICHONICKY, L. & KRUH, J. (1978). Chromatin protein kinases and phosphoproteins during myoblast growth and differentiation. *Biochemical & Biophysical Research Communications* **81**, 623-629.

LEINWAND, L.A., FOURNIER, R.E., NADAL-GINARD, B. & SHOWS, T.B. (1983). Multigene family for sarcomeric myosin heavy chain in mouse and human DNA: localization on a single chromosome. *Science* **221**, 766-769.

LEVY, N.S., GOLDBERG, M.A. & LEVY, A.P. (1997). Sequencing of the human vascular endothelial growth factor (VEGF) 3' untranslated region (UTR): conservation of five hypoxia-inducible RNA-protein binding sites. *Biochimica et Biophysica Acta* **1352**, 167-173.

LEWIN, B. (1975). Units of transcription and translation: sequence components of heterogeneous nuclear RNA and messenger RNA. *Cell* **4**, 77-93.

LEWIN, B. (1975). Units of transcription and translation: the relationship between heterogeneous nuclear RNA and messenger RNA. [Review] [56 refs]. *Cell* **4**, 11-20.

LIM, H.M. & PENE, J.J. (1988). Optimal conditions for supercoil DNA sequencing with the Escherichia coli DNA polymerase I large fragment. *Gene Analysis Techniques* **5**, 32-39.

LIM, L. & CANELLAKIS, E.S. (1970). Adenine-rich polymer associated with rabbit reticulocyte messenger RNA. *Nature* **227**, 710-712.

LINXWEILER, W. & HORZ, W. (1982). Sequence specificity of exonuclease III from E. coli. *Nucleic Acids Research* **10**, 4845-4859.

LODISH, H.F. & SMALL, B. (1976). Different lifetimes of reticulocyte messenger RNA. *Cell* 7, 59-65.

LOMPRE, A.M., NADAL-GINARD, B. & MAHDAVI, V. (1984). Expression of the cardiac ventricular alpha- and beta-myosin heavy chain genes is developmentally and hormonally regulated. *Journal of Biological Chemistry* **259**, 6437-6446.

LORENZ, L.J. & RICHTER, J.D. (1985). A cDNA clone for a polyadenylated RNA-binding protein of Xenopus laevis oocytes hybridizes to four developmentally regulated mRNAs. *Molecular & Cellular Biology* **5**, 2697-2704.

LOWEY, S. (1986). The structure of vertebrate muscle myosin. In *Myology*, ed. ENGEL, A.G.a.B.B.Q., pp. 563-587. McGrawHill, New York.

LU, J.X. (1992). Magnesium chelation inactivates beta-galactosidase in -20 degrees C storage. *Biotechniques* **12**, 177-181.

LUSCHER, B., STAUBER, C., SCHINDLER, R. & SCHUMPERLI, D. (1985). Faithful cell-cycle regulation of a recombinant mouse histone H4 gene is controlled by sequences in the 3'-terminal part of the gene. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 4389-4393.

LYONS, G.E., ONTELL, M., COX, R., SASSOON, D. & BUCKINGHAM, M. (1990). The expression of myosin genes in developing skeletal muscle in the mouse embryo. *Journal of Cell Biology* **111**, 1465-1476.

MAATTA, A. & PENTTINEN, R.P.K. (1993). A fibroblast protien binds the 3'-untranslated region of the pro-alpha1(I) collagen mRNa. *Biochemical Journal* **295**, 691-698.

MACDONALD, C.C., WILUSZ, J. & SHENK, T. (1994). The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Molecular & Cellular Biology* **14**, 6647-6654.

MACDONALD, P.M. & STRUHL, G. (1988). cis-acting sequences responsible for anterior localization of bicoid mRNA in Drosophila embryos. *Nature* **336**, 595-598.

MAHDAVI, V., IZUMO, S. & NADAL-GINARD, B. (1987). Developmental and hormonal regulation of sarcomeric myosin heavy chain gene family. *Circulation Research* **60**, 804-814.

MALTER, J.S. & HONG, Y. (1991). A redox switch and phosphorylation are involved in the post-translational up-regulation of the adenosine-uridine binding factor by phorbol ester and ionophore. *Journal of Biological Chemistry* **266**, 3167-3171.

MARTIN, J.D. (1990). Application of the Two-Phase Partition Assay for Chloramphenicol Acetyl Transferase (CAT) to Transfections with Simian Virus 40-CAT Plasmids. *Analytical Biochemistry* **191**, 242-246.

MATTAJ, I.W. (1993). RNA recognition: a family matter?. [Review] [33 refs]. *Cell* **73**, 837-840.

MAYRAND, S. & PEDERSON, T. (1981). Nuclear ribonucleoprotein particles probed in living cells. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 2208-2212.

MCCARTHY, J.E. & KOLLMUS, H. (1995). Cytoplasmic mRNA-protein interactions in eukaryotic gene expression. [Review] [51 refs]. *Trends in Biochemical Sciences* **20**, 191-197.

MCDEVITT, M.A., IMPERIALE, M.J., ALI, H. & NEVINS, J.R. (1984). Requirement of a downstream sequence for generation of a poly(A) addition site. *Cell* **37**, 993-999.

MCDEVITT, M.A., HART, R.P., WONG, W.W. & NEVINS, J.R. (1986). Sequences capable of restoring poly(A) site function define two distinct downstream elements. *EMBO Journal* **5**, 2907-2913.

MCGOWAN, K. & PEKALA, P.H. (1996). Dehydrogenase binding to the 3'-untranslated region of GLUT1 mRNA. *Biochemical & Biophysical Research Communications* **221**, 42-45.

MCGREW, L.L., DWORKIN-RASTL, E., DWORKIN, M.B. & RICHTER, J.D. (1989). Poly(A) elongation during Xenopus oocyte maturation is required for translational recruitment and is mediated by a short sequence element. *Genes & Development* **3**, 803-815.

MCGREW, L.L., DWORKIN-RASTL, E., DWORKIN, M.B. & RICHTER, J.D. (1989). Poly(A) elongation during Xenopus oocyte maturation is required for translational recruitment and is mediated by a short sequence element. *Genes & Development* **3**, 803-815.

MCGREW, L.L. & RICHTER, J.D. (1990). Translational control by cytoplasmic polyadenylation during Xenopus oocyte maturation: characterization of cis and trans elements and regulation by cyclin/MPF. *EMBO Journal* **9**, 3743-3751.

MCKOY, G., LEGER, M.E., BACOU, F. & GOLDSPINK, G. (1998). Differential expression of myosin heavy chain mRNA and protein isoforms in four functionally diverse rabbit skeletal muscles during pre- and postnatal development. *Developmental Dynamics* **211**, 193-203.

MCLAUCHLAN, J., GAFFNEY, D., WHITTON, J.L. & CLEMENTS, J.B. (1985). The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. *Nucleic Acids Research* **13**, 1347-1368.

MEAD, D.A., PEY, N.K., HERRNSTADT, C., MARCIL, R.A. & SMITH, L.M. (1991). A universal method for the direct cloning of PCR amplified nucleic acid. *Bio/Technology* **9**, 657-663.

MEDFORD, R.M., NGUYEN, H.T. & NADAL-GINARD, B. (1983). Transcriptional and cell cycle-mediated regulation of myosin heavy chain gene expression during muscle cell differentiation. *Journal of Biological Chemistry* **258**, 11063-11073.

MEDFORD, R.M., NGUYEN, H.T. & NADAL-GINARD, B. (1983). Transcriptional and cell cycle-mediated regulation of myosin heavy chain gene expression during muscle cell differentiation. *Journal of Biological Chemistry* **258**, 11063-11073.

MEGENEY, L.A. & RUDNICKI, M.A. (1995). Determination versus differentiation and the MyoD family of transcription factors. [Review] [100 refs]. *Biochemistry & Cell Biology* **73**, 723-732.

MEJEAN, C., PONS, F., BENYAMIN, Y. & ROUSTAN, C. (1989). Antigenic probes locate binding sites for the glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, aldolase and phosphofructokinase on the actin monomer in microfilaments. *Biochemical Journal* **264**, 671-677.

MELTON, D.A., KRIEG, P.A., REBAGLIATI, M.R., MANIATIS, T., ZINN, K. & GREEN, M.R. (1984). Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Research* **12**, 7035-7056.

MILLIGAN, J.F., GROEBE, D.R., WITHERELL, G.W. & UHLENBECK, O.C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Research* **15**, 8783-8798.

MINASCHEK, G., GROSCHEL-STEWART, U., BLUM, S. & BEREITER-HAHN, J. (1992). Microcompartmentation of glycolytic enzymes in cultured cells. *European Journal of Cell Biology* **58**, 418-428.

MIZZEN, L.A. & WELCH, W.J. (1988). Characterization of the thermotolerant cell. I. Effects on protein synthesis activity and the regulation of heat-shock protein 70 expression. *Journal of Cell Biology* **106**, 1105-1116.

MONTELL, C., FISHER, E.F., CARUTHERS, M.H. & BERK, A.J. (1983). Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3' consensus sequence AAUAAA. *Nature* **305**, 600-605.

MOORE, C.L. & SHARP, P.A. (1984). Site-specific polyadenylation in a cell-free reaction. *Cell* **36**, 581-591.

MOORE, C.L. & SHARP, P.A. (1985). Accurate cleavage and polyadenylation of exogenous RNA substrate. *Cell* **41**, 845-855.

MOORE, L.A., TIDYMAN, W.E., ARRIZUBIETA, M.J. & BANDMAN, E. (1993). The evolutionary relationship of avian and mammalian myosin heavy-chain genes. *Journal of Molecular Evolution* **36**, 21-30.

MORRIS, E. J., FULTON, A. B. (1994). Rearrangement of mRNAs for costamere proteins during costamere development in cultured skeletal muscle from chicken. *Journal of Cell Science* **107**, 377-386.

MORRIS, T.D., WEBER, L.A., HICKEY, E., STEIN, G.S. & STEIN, J.L. (1991). Changes in the stability of a human H3 histone mRNA during the HeLa cell cycle. *Molecular & Cellular Biology* **11**, 544-553.

MOSELEY, P.L., WALLEN, E.S., MCCAFFERTY, J.D., FLANAGAN, S. & KERN, J.A. (1993). Heat stress regulates the human 70-kDa heat-shock gene through the 3'-untranslated region. *American Journal of Physiology* **264**, L533-L537

MOSS, R.L., DIFFEE, G.M. & GREASER, M.L. (1995). Contractile properties of skeletal muscle fibers in relation to myofibrillar protein isoforms. [Review] [194 refs]. *Reviews of Physiology Biochemistry & Pharmacology* **126**, 1-63.

MOWRY, K.L. & MELTON, D.A. (1992). Vegetal messenger RNA localization directed by a 340-nt RNA sequence element in Xenopus oocytes. *Science* **255**, 991-994.

MUKAI, T., JOH, K., ARAI, Y., YATSUKI, H. & HORI, K. (1986). Tissue-specific Expression of Rat Aldolase A mRNAs. *Journal of Biological Chemistry* **261**, 3347-3354.

MULLIS, K., FALOONA, F., SCHARF, S., SAIKI, R., HORN, G. & ERLICH, H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology* **51 Pt 1**, 263-273.

MULLIS, K.B. & FALOONA, F.A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* **155**, 335-350.

MULLNER, E.W. & KUHN, L.C. (1988). A stem-loop in the 3' untranslated region mediates iron-dependent regulation of transferrin receptor mRNA stability in the cytoplasm. *Cell* **53**, 815-825.

MURRAY, M.T., SCHILLER, D.L. & FRANKE, W.W. (1992). Sequence analysis of cytoplasmic mRNA-binding proteins of Xenopus oocytes identifies a family of RNA-binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 11-15.

NAGY, E. & RIGBY, W.F. (1995). Glyceraldehyde-3-phosphate dehydrogenase selectively binds AU-rich RNA in the NAD(+)-binding region (Rossmann fold). *Journal of Biological Chemistry* **270**, 2755-2763.

NEUPERT, B., THOMPSON, N.A., MEYER, C. & KUHN, L.C. (1990). A high yield affinity purification method for specific RNA-binding proteins: isolation of the iron regulatory factor from human placenta. *Nucleic Acids Research* **18**, 51-55.

NEVINS, J.R. & DARNELL, J.E. (1978). Groups of adenovirus type 2 mRNA's derived from a large primary transcript: probable nuclear origin and possible common 3' ends. *Journal of Virology* **25**, 811-823.

NUSSLEIN-VOLHARD, C., FROHNHOFER, H.G. & LEHMANN, R. (1987). Determination of anteroposterior polarity in Drosophila. *Science* **238**, 1675-1681.

OLSON, E. (1992). Activation of muscle-specific transcription by myogenic helix-loop-helix proteins. [Review] [38 refs]. *Symposia of the Society for Experimental Biology* **46**, 331-341.

OLSON, E.N. (1993). Regulation of muscle transcription by the MyoD family. The heart of the matter. [Review] [27 refs]. *Circulation Research* **72**, 1-6.

OLSON, E.N., PERRY, M. & SCHULZ, R.A. (1995). Regulation of muscle differentiation by the MEF2 family of MADS box transcription factors. [Review] [102 refs]. *Developmental Biology (Orlando)* **172**, 2-14.

ORDAHL, C.P. & CAPLAN, A.I. (1976). Transcriptional diversity in myogenesis. *Developmental Biology (Orlando)* **54**, 61-72.

OSTARECK-LEDERER, A., OSTARECK, D.H., STANDART, N. & THIELE, B.J. (1994). Translation of 15-lipoxygenase mRNA is inhibited by a protein that binds to a repeated sequence in the 3' untranslated region. *EMBO Journal* **13**, 1476-1481.

PANDEY, N.B. & MARZLUFF, W.F. (1987). The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. *Molecular & Cellular Biology* 7, 4557-4559.

PANDEY, N.B., WILLIAMS, A.S., SUN, J.H., BROWN, V.D., BOND, U. & MARZLUFF, W.F. (1994). Point mutations in the stem-loop at the 3' end of mouse histone mRNA reduce expression by reducing the efficiency of 3' end formation. *Molecular & Cellular Biology* **14**, 1709-1720.

PARIS, J. & RICHTER, J.D. (1990). Maturation-specific polyadenylation and translational control: diversity of cytoplasmic polyadenylation elements, influence of poly(A) tail size, and formation of stable polyadenylation complexes. *Molecular & Cellular Biology* **10** , 5634-5645.

PARIS, J., SWENSON, K., PIWNICA-WORMS, H. & RICHTER, J.D. (1991). Maturation-specific polyadenylation: in vitro activation by p34cdc2 and phosphorylation of a 58-kD CPE-binding protein. *Genes & Development* **5**, 1697-1708.

PARKER-THORNBURG, J., BAUER, B., PALERMO, J. & ROBBINS, J. (1992). Structural and developmental analysis of two linked myosin heavy chain genes. *Developmental Biology (Orlando)* **150** , 99-107.

PASHEV, I.G., DIMITROV, S.I. & ANGELOV, D. (1991). Crosslinking proteins to nucleic acids by ultraviolet laser irradiation. [Review] [19 refs]. *Trends in Biochemical Sciences* **16**, 323-326.

PAULDING, W.R. & CZYZYK-KRZESKA, M.F. (1999). Regulation of tyrosine hydroxylase mRNA stability by protein-binding, pyrimidine-rich sequence in the 3'-untranslated region. *Journal of Biological Chemistry* **274**, 2532-2538.

PAYNTON, B.V., REMPEL, R. & BACHVAROVA, R. (1988). Changes in state of adenylation and time course of degradation of maternal mRNAs during oocyte maturation and early embryonic development in the mouse. *Developmental Biology (Orlando)* **129**, 304-314.

PELHAM, H.R. & MUNRO, S. (1993). Sorting of membrane proteins in the secretory pathway. [Review] [19 refs]. *Cell* **75**, 603-605.

PELLE, R. & MURPHY, N.B. (1993). In vivo UV-cross-linking hybridization: a powerful technique for isolating RNA binding proteins. Application to trypanosome mini-exon derived RNA. *Nucleic Acids Research* **21**, 2453-2458.

PERHONEN, M., SHARP, W.W. & RUSSELL, B. (1998). Microtubules are needed for dispersal of alpha-myosin heavy chain mRNA in rat neonatal cardiac myocytes. *Journal of Molecular & Cellular Cardiology* **30**, 1713-1722.

PERIASAMY, M., WIECZOREK, D.F. & NADAL-GINARD, B. (1984). Characterization of a developmentally regulated perinatal myosin heavy-chain gene expressed in skeletal muscle. *Journal of Biological Chemistry* **259**, 13573-13578.

PERIASAMY, M., STEHELER, E.E., GARFINKEL, L.I., GUBITS, R.M., RUIZ-OPAZO, N. & NADAL-GINARD, B. (1984). Fast skeletal myosin light chains 1 and 3 are produced from a single gene by a combined process of differential RNA transcription and splicing. *Journal of Biological Chemistry* **259**, 13595-13604.

PERUCHO, M., SALAS, J. & SALAS, M.L. (1977). Identification of the mammalian DNA-binding protein P8 as glyceraldehyde-3-phosphate dehydrogenase. *European Journal of Biochemistry* **81** , 557-562.

PERUCHO, M., SALAS, J. & SALAS, M.L. (1980). Study of the interaction of glyceraldehyde-3-phosphate dehydrogenase with DNA. *Biochimica et Biophysica Acta* **606**, 181-195.

PESOLE, G., LIUNI, S., GRILLO, G. & SACCONE, C. (1997). Structural and compositional features of untranslated regions of eukaryotic mRNAs. *Gene* **205**, 95-102.

PETERSEN, R.B. & LINDQUIST, S. (1989). Regulation of HSP70 synthesis by messenger RNA degradation. *Cell Regulation* **1**, 135-149.

PETTE, D. & VRBOVA, G. (1985). Neural control of phenotypic expression in mammalian muscle fibers. [Review] [215 refs]. *Muscle & Nerve* **8**, 676-689.

PETTE, D. & STARON, R.S. (1990). Cellular and molecular diversities of mammalian skeletal muscle fibers. [Review] [648 refs]. *Reviews of Physiology Biochemistry & Pharmacology* **116**, 1-76.

PHILIPSON, L., WALL, R., GLICKMAN, G. & DARNELL, J.E. (1971). Addition of polyadenylate sequences to virus-specific RNA during adenovirus replication. *Proceedings of the National Academy of Sciences of the United States of America* **68**, 2806-2809.

POKRYWKA, N.J. & STEPHENSON, E.C. (1991). Microtubules mediate the localization of bicoid RNA during Drosophila oogenesis. *Development* **113**, 55-66.

POKRYWKA, N.J. (1995). RNA localization and the cytoskeleton in Drosophila oocytes. [Review] [49 refs]. *Current Topics in Developmental Biology* **31**, 139-166.

PROUDFOOT, N.J. & BROWNLEE, G.G. (1976). 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* **263**, 211-214.

RAPOPORT, S.M. & SCHEWE, T. (1986). The maturational breakdown of mitochondria in reticulocytes. [Review] [195 refs]. *Biochimica et Biophysica Acta* **864**, 471-495.

RASTINEJAD, F. & BLAU, H.M. (1993). Genetic complementation reveals a novel regulatory role for 3' untranslated regions in growth and differentiation. *Cell* **72**, 903-917.

RASTINEJAD, F., CONBOY, M.J., RANDO, T.A. & BLAU, H.M. (1993). Tumor suppression by RNA from the 3' untranslated region of alpha-tropomyosin. *Cell* **75**, 1107-1117.

RAY, J., GAGE, F.H., KRIEG, P.A. & MELTON, D.A. (1992). Gene transfer into established and primary fibroblast cell lines: comparison of transfection methods and promoters. In vitro RNA synthesis with SP6 RNA polymerase. *Biotechniques* **13**, 598-603.

RAYMENT, I., RYPNIEWSKI, W.R., SCHMIDT-BASE, K., SMITH, R., TOMCHICK, D.R., BENNING, M.M., WINKELMANN, D.A., WESENBERG, G. & HOLDEN, H.M. (1993). Three-dimensional structure of myosin subfragment-1: a molecular motor [see comments]. *Science* **261**, 50-58.

REBAGLIATI, M.R., WEEKS, D.L., HARVEY, R.P. & MELTON, D.A. (1985). Identification and cloning of localized maternal RNAs from Xenopus eggs. *Cell* **42**, 769-777.

REIMER, D.L. & SINGH, S.M. (1996). Distinct mRNA-binding proteins interacting with short repeat sequences of the 3'UTR may be involved in the post-transcriptional regulation of the mouse catalase gene, Cas-1. *DNA & Cell Biology* **15**, 317-328.

REISER, P.J., MOSS, R.L., GIULIAN, G.G. & GREASER, M.L. (1985). Shortening velocity in single fibers from adult rabbit soleus muscles is correlated with myosin heavy chain composition. *Journal of Biological Chemistry* **260**, 9077-9080.

REISER, P.J., MOSS, R.L., GIULIAN, G.G. & GREASER, M.L. (1985). Shortening velocity and myosin heavy chains of developing rabbit muscle fibers. *Journal of Biological Chemistry* **260**, 14403-14405.

RICHTER, J.D. (1991). Translational control during early development. [Review] [34 refs]. *Bioessays* **13** , 179-183.

RICHTER, K., GOOD, P.J. & DAWID, I.B. (1990). A developmentally regulated, nervous system-specific gene in Xenopus encodes a putative RNA-binding protein. *New Biologist* **2**, 556-565.

RIVERA-POMAR, R., NIESSING, D., SCHMIDT-OTT, U., GEHRING, W.J. & JACKLE, H. (1996). RNA binding and translational suppression by bicoid [see comments]. *Nature* **379**, 746-749.

ROBBIE, E.P., PETERSON, M., AMAYA, E. & MUSCI, T.J. (1995). Temporal regulation of the Xenopus FGF receptor in development: a translation inhibitory element in the 3' untranslated region. *Development* **121**, 1775-1785.

ROBBINS, J. & HEYWOOD, S.M. (1978). Quantification of myosin heavy-chain mRNA during myogenesis. *European Journal of Biochemistry* **82**, 601-608.

ROGERS, S.G. & WEISS, B. (1980). Exonuclease III of Escherichia coli K-12, an AP endonuclease. *Methods in Enzymology* **65**, 201-211.

ROGERS, S.G. & WEISS, B. (1980). Cloning of the exonuclease III gene of Escherichia coli. *Gene* **11**, 187-195.

RONAI, Z., ROBINSON, R., RUTBERG, S., LAZARUS, P. & SARDANA, M. (1992). Aldolase-DNA interactions in a SEWA cell system. *Biochimica et Biophysica Acta* **1130**, 20-28.

RONAI, Z. (1993). Glycolytic enzymes as DNA binding proteins. [Review] [16 refs]. *International Journal of Biochemistry* **25**, 1073-1076.

RONDON, I.J., MACMILLAN, L.A., BECKMAN, B.S., GOLDBERG, M.A., SCHNEIDER, T., BUNN, HF & MALTER, J.S. (1991). Hypoxia up-regulates the activity of a novel erythropoietin mRNA binding protein. *Journal of Biological Chemistry* **266**, 16594-16598.

ROSBASH, M. & SINGER, R.H. (1993). RNA travel: tracks from DNA to cytoplasm. [Review] [27 refs]. *Cell* **75**, 399-401.

ROSENTHAL, E.T., TANSEY, T.R. & RUDERMAN, J.V. (1983). Sequence-specific adenylations and deadenylations accompany changes in the translation of maternal messenger RNA after fertilization of Spisula oocytes. *Journal of Molecular Biology* **166**, 309-327.

ROSENTHAL, E.T. & RUDERMAN, J.V. (1987). Widespread changes in the translation and adenylation of maternal messenger RNAs following fertilization of Spisula oocytes. *Developmental Biology (Orlando)* **121**, 237-246.

ROSS, A.F., OLEYNIKOV, Y., KISLAUSKIS, E.H., TANEJA, K.L. & SINGER, R.H. (1997). Characterization of a beta-actin mRNA zipcode-binding protein. *Molecular & Cellular Biology* **17**, 2158-2165.

ROSS, J. & SULLIVAN, T.D. (1985). Half-lives of beta and gamma globin messenger RNAs and of protein synthetic capacity in cultured human reticulocytes. *Blood* **66**, 1149-1154.

ROSS, J. (1995). mRNA stability in mammalian cells. [Review] [395 refs]. *Microbiological Reviews* **59**, 423-450.

ROTHENBERGER, S., MULLNER, E.W. & KUHN, L.C. (1990). The mRNA-binding protein which controls ferritin and transferrin receptor expression is conserved during evolution. *Nucleic Acids Research* **18**, 1175-1179.

ROWLERSON, A.P.B.M.J.W.R.G.a.W.A.G. (1981). A novel myosin present in cat jaw closing muscles. *Journal of Muscle Research & Cell Motility* **2**, 425-438.

ROY, R.C., SAIDAPET, C., DASGUPTA, S. & SARKAR, S. (1984). Regulation of mRNA translation during chicken myogenesis in ovo. *Experimental Biology and Medicine* **9**, 284-289.

ROY, S., PARKIN, N.T., ROSEN, C., ITOVITCH, J. & SONENBERG, N. (1990). Structural requirements for trans activation of human immunodeficiency virus type 1 long terminal repeat-directed gene expression by tat: importance of base pairing, loop sequence, and bulges in the tat-responsive sequence. *Journal of Virology* **64**, 1402-1406.

RUDNICKI, M.A. & JAENISCH, R. (1995). The MyoD family of transcription factors and skeletal myogenesis. [Review] [66 refs]. *Bioessays* **17**, 203-209.

RUSHBROOK, J.I., WEISS, C., KO, K., FEUERMAN, M.H., CARLETON, S., ING, A. & JACOBY (1994). Identification of alpha-cardiac myosin heavy chain mRNA and protein in extraocular muscle of the adult rabbit. *Journal of Muscle Research & Cell Motility* **15**, 505-515.

RUSSELL, J.E. & LIEBHABER, S.A. (1996). The stability of human beta-globin mRNA is dependent on structural determinants positioned within its 3' untranslated region. *Blood* **87**, 5314-5323.

RUSSELL, B., WENDEROTH, M. P., GOLDSPINK, P. H. 1992. Remodeling of myofibrils: subcellular distribution of myosin heavy chain mRNA and protein. *American Journal of Physiology* **262**, R339-R345.

RUSSELL, S.D., CAMBON, N., NADAL-GINARD, B. & WHALEN, R.G. (1988). Thyroid hormone induces a nerve-independent precocious expression of fast myosin heavy chain mRNA in rat hindlimb skeletal muscle. *Journal of Biological Chemistry* **263**, 6370-6374.

SADOFSKY, M. & ALWINE, J.C. (1984). Sequences on the 3' side of hexanucleotide AAUAAA affect efficiency of cleavage at the polyadenylation site. *Molecular & Cellular Biology* **4**, 1460-1468.

SAEZ, L. & LEINWAND, L.A. (1986). Characterization of diverse forms of myosin heavy chain expressed in adult human skeletal muscle. *Nucleic Acids Research* **14**, 2951-2969.

SAEZ, L.J., GIANOLA, K.M., MCNALLY, E.M., FEGHALI, R., EDDY, R., SHOWS, T.B. & LEINWAND, L.A. (1987). Human cardiac myosin heavy chain genes and their linkage in the genome. *Nucleic Acids Research* **15**, 5443-5459.

SAGATA, N., SHIOKAWA, K. & YAMANA, K. (1980). A study on the steady-state population of poly(A)+RNA during early development of Xenopus laevis. *Developmental Biology (Orlando)* 77, 431-448.

SAIKI, R.K., SCHARF, S., FALOONA, F., MULLIS, K.B., HORN, G.T., ERLICH, H.A. & ARNHEIM, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**, 1350-1354.

SALLES, F.J., LIEBERFARB, M.E., WREDEN, C., GERGEN, J.P. & STRICKLAND, S. (1994). Coordinate initiation of Drosophila development by regulated polyadenylation of maternal messenger RNAs. *Science* **266**, 1996-1999.

SAMBROOK, J., FRITSCH, E. F., AND MANIATIS, T. Molecular Cloning: A Laboratory Manual. Second Edition. Cold Spring Harbor Laboratory Press, Plain View, New York. 1989.

SANCHEZ, A., JONES, W.K., GULICK, J., DOETSCHMAN, T. & ROBBINS, J. (1991). Myosin heavy chain gene expression in mouse embryoid bodies. An in vitro developmental study. *Journal of Biological Chemistry* **266**, 22419-22426.

SANGER, F., NICKLEN, S. & COULSON, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463-5467.

SARTORE, S., MASCARELLO, F., ROWLERSON, A., GORZA, L., AUSONI, S., VIANELLO, M. & SCHIAFFINO, S. (1987). Fibre types in extraocular muscles: a new myosin isoform in the fast fibres. *Journal of Muscle Research & Cell Motility* **8**, 161-172.

SCHERLY, D., BOELENS, W., VAN, V.W., DATHAN, N.A., HAMM, J. & MATTAJ, I.W. (1989). Identification of the RNA binding segment of human U1 A protein and definition of its binding site on U1 snRNA. *EMBO Journal* **8**, 4163-4170.

SCHIAFFINO, S., AUSONI, S., GORZA, L., SAGGIN, L., GUNDERSEN, K. & LOMO, T. (1988). Myosin heavy chain isoforms and velocity of shortening of type 2 skeletal muscle fibres. *Acta Physiologica Scandinavica* **134**, 575-576.

SCHIAFFINO, S., GORZA, L., SARTORE, S., SAGGIN, L., AUSONI, S., VIANELLO, M., GUNDERSEN, K. & LOMO, T. (1989). Three myosin heavy chain isoforms in type 2 skeletal muscle fibres. *Journal of Muscle Research & Cell Motility* **10**, 197-205.

SCHIAFFINO, S. & REGGIANI, C. (1994). Myosin isoforms in mammalian skeletal muscle. [Review] [85 refs]. *Journal of Applied Physiology* **77**, 493-501.

SCHIAFFINO, S. & REGGIANI, C. (1996). Molecular diversity of myofibrillar proteins: gene regulation and functional significance. [Review] [585 refs]. *Physiological Reviews* **76**, 371-423.

SCHULTZ, D.E., HARDIN, C.C. & LEMON, S.M. (1996). Specific interaction of glyceraldehyde 3-phosphate dehydrogenase with the 5'-nontranslated RNA of hepatitis A virus. *Journal of Biological Chemistry* **271**, 14134-14142.

SCHWARTZ, S.P., AISENTHAL, L., ELISHA, Z., OBERMAN, F. & YISRAELI, J.K. (1992). A 69-kDa RNA-binding protein from Xenopus oocytes recognizes a common motif in two vegetally localized maternal mRNAs. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 11895-11899.

SCHWEIGHOFFER, F., MAIRE, P., TUIL, D., GAUTRON, S., DAEGELEN, D., BACHNER, L. & KAHN, A. (1986). In vivo developmental modifications of the expression of genes encoding muscle-specific enzymes in rat. *Journal of Biological Chemistry* **261**, 10271-10276.

SCIOTE, J.J., ROWLERSON, A.M. & CARLSON, D.S. (1995). Myosin expression in the jaw-closing muscles of the domestic cat and American opossum. *Archives of Oral Biology* **40**, 405-413.

SEED, B. & SHEEN, J.Y. (1988). A simple phase-extraction assay for chloramphenicol acyltransferase activity. *Gene* **67**, 271-277.

SETZER, D.R., MCGROGAN, M., NUNBERG, J.H. & SCHIMKE, R.T. (1980). Size heterogeneity in the 3' end of dihydrofolate reductase messenger RNAs in mouse cells. *Cell* **22**, 361-370.

SHAW, G. & KAMEN, R. (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46**, 659-667.

SHEETS, M.D., OGG, S.C. & WICKENS, M.P. (1990). Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic Acids Research* **18**, 5799-5805.

SHEETZ, M.P., CHASAN, R. & SPUDICH, J.A. (1984). ATP-dependent movement of myosin in vitro: characterization of a quantitative assay. *Journal of Cell Biology* **99**, 1867-1871.

SHIGEKAWA, K. & DOWER, W.J. (1988). Electroporation of eukaryotes and prokaryotes: a general approach to the introduction of macromolecules into cells. *Biotechniques* **6**, 742-751.

SHIGEKAWA, K. & DOWER, W.J. (1989). Electroporation: a general approach to the introduction of macromolecules into prokaryotic and eukaryotic cells. [Review] [32 refs]. *Australian Journal of Biotechnology* **3**, 56-62.

SHOEMAKER, S. D., RYAN, A. F., LIEBER, R. L. (1999). Transcript-specific mRNA trafficking based on the distribution of coexpressed myosin isoforms. *Cells Tissues Organs* **165**, 10-15.

SHYU, A.B., GREENBERG, M.E. & BELASCO, J.G. (1989). The c-fos transcript is targeted for rapid decay by two distinct mRNA degradation pathways. *Genes & Development* **3**, 60-72.

SILBERSTEIN, L., WEBSTER, S.G., TRAVIS, M. & BLAU, H.M. (1986). Developmental progression of myosin gene expression in cultured muscle cells. *Cell* **46**, 1075-1081.

SIMON, R., TASSAN, J. P., RICHTER, J. D. (1992). Translational control by poly(A) elongation during Xenopus development: differential repression and enhancement

SINGER, R.H. (1992). The cytoskeleton and mRNA localization. [Review] [41 refs]. *Current Opinion in Cell Biology* **4**, 15-19.

SINGH, R. & GREEN, M.R. (1993). Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase. *Science* **259**, 365-368.

SIROVER, M.A. (1996). Minireview. Emerging new functions of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in mammalian cells. [Review] [28 refs]. *Life Sciences* **58** , 2271-2277.

SIROVER, M.A. (1997). Role of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in normal cell function and in cell pathology. [Review] [46 refs]. *Journal of Cellular Biochemistry* **66**, 133-140.

SIROVER, M.A. (1999). New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase. [Review] [166 refs]. *Biochimica et Biophysica Acta* **1432**, 159-184.

SMITH, P.K., KROHN, R.I., HERMANSON, G.T., MALLIA, A.K., GARTNER, F.H., PROVENZANO, MD, FUJIMOTO, E.K., GOEKE, N.M., OLSON, B.J. & KLENK, D.C. (1985). Measurement of protein using bicinchoninic acid [published erratum appears in Anal Biochem 1987 May 15;163(1):279]. *Analytical Biochemistry* **150**, 76-85.

SOMMERVILLE, J. (1992). RNA-binding proteins: masking proteins revealed. [Review] [22 refs]. *Bioessays* **14**, 337-339.

SONENBERG, N. (1994). mRNA translation: influence of the 5' and 3' untranslated regions. [Review] [51 refs]. *Current Opinion in Genetics & Development* **4**, 310-315.

SPRAGUE, J., CONDRA, J.H., ARNHEITER, H. & LAZZARINI, R.A. (1983). Expression of a recombinant DNA gene coding for the vesicular stomatitis virus nucleocapsid protein. *Journal of Virology* **45**, 773-781.

ST, JHONSTON, D., DRIEVER, W., BERLETH, T., RICHSTEIN, S. & NUSSLEIN-VOLHARD, C. (1989). Multiple steps in the localization of bicoid RNA to the anterior pole of the Drosophila oocyte. *Development* **107 Suppl**, 13-19.

ST, JHONSTON, D., BEUCHLE, D. & NUSSLEIN-VOLHARD, C. (1991). Staufen, a gene required to localize maternal RNAs in the Drosophila egg. *Cell* **66**, 51-63.

ST, JHONSTON, D. (1995). The intracellular localization of messenger RNAs. *Cell* **81**, 161-170.

STANDART, N., DALE, M., STEWART, E. & HUNT, T. (1990). Maternal mRNA from clam oocytes can be specifically unmasked in vitro by antisense RNA complementary to the 3'-untranslated region. *Genes & Development* **4**, 2157-2168.

STANDART, N. & DALE, M. (1993). Regulated polyadenylation of clam maternal mRNAs in vitro. *Developmental Genetics* **14**, 492-499.

STANDART, N. & JACKSON, R.J. (1994). Regulation of translation by specific protein/mRNA interactions. [Review] [92 refs]. *Biochimie* **76**, 867-879.

STANDART, N. & JACKSON, R. (1994). Translational regulation. Y the message is masked?. *Current Biology* **4**, 939-941.

STEDMAN, H.H., ELLER, M., JULLIAN, E.H., FERTELS, S.H., SARKAR, S., SYLVESTER, J.E., KELLY, A.M. & RUBINSTEIN, N.A. (1990). The human embryonic myosin heavy chain. Complete primary structure reveals evolutionary relationships with other developmental isoforms. *Journal of Biological Chemistry* **265**, 3568-3576.

STEPHENS, J.M., CARTER, B.Z., PEKALA, P.H. & MALTER, J.S. (1992). Tumor necrosis factor alpha-induced glucose transporter (GLUT-1) mRNA stabilization in 3T3-L1 preadipocytes. Regulation by the adenosine-uridine binding factor. *Journal of Biological Chemistry* **267** , 8336-8341.

STEPHENSON, E.C., CHAO, Y.C. & FACKENTHAL, J.D. (1988). Molecular analysis of the swallow gene of Drosophila melanogaster. *Genes & Development* **2**, 1655-1665.

STOCKDALE, F.E. (1992). Myogenic cell lineages. [Review] [105 refs]. *Developmental Biology (Orlando)* **154**, 284-298.

STRUHL, G., STRUHL, K. & MACDONALD, P.M. (1989). The gradient morphogen bicoid is a concentration-dependent transcriptional activator. *Cell* **57**, 1259-1273.

STUMP, W.T. & HALL, K.B. (1993). SP6 RNA polymerase efficiently synthesizes RNA from short double-stranded DNA templates. *Nucleic Acids Research* **21**, 5480-5484.

STUTZ, A., HUARTE, J., GUBLER, P., CONNE, B., BELIN, D. & VASSALLI, J.D. (1997). In vivo antisense oligodeoxynucleotide mapping reveals masked regulatory elements in an mRNA dormant in mouse oocytes. *Molecular & Cellular Biology* **17**, 1759-1767.

SUNDELL, C.L. & SINGER, R.H. (1990). Actin mRNA localizes in the absence of protein synthesis. *Journal of Cell Biology* **111**, 2397-2403.

SUNDELL, C.L. & SINGER, R.H. (1991). Requirement of microfilaments in sorting of actin messenger RNA. *Science* **253**, 1275-1277.

SUZUKI, E., GUO, K., KOLMAN, M., YU, Y.T. & WALSH, K. (1995). Serum induction of MEF2/RSRF expression in vascular myocytes is mediated at the level of translation. *Molecular & Cellular Biology* **15**, 3415-3423.

TAIRA, M., JAMRICH, M., GOOD, P.J. & DAWID, I.B. (1992). The LIM domain-containing homeo box gene Xlim-1 is expressed specifically in the organizer region of Xenopus gastrula embryos. *Genes & Development* **6**, 356-366.

TEMPLETON, D. & ECKHART, W. (1984). N-terminal amino acid sequences of the polyoma middle-size T antigen are important for protein kinase activity and cell transformation. *Molecular & Cellular Biology* **4**, 817-821.

TERMIN, A., STARON, R.S. & PETTE, D. (1989). Myosin heavy chain isoforms in histochemically defined fiber types of rat muscle. *Histochemistry* **92**, 453-457.

THEIL, E.C. (1990). Regulation of ferritin and transferrin receptor mRNAs. [Review] [58 refs]. *Journal of Biological Chemistry* **265**, 4771-4774.

THIELE, B.J., ANDREE, H., HOHNE, M. & RAPOPORT, S.M. (1982). Lipoxygenase mRNA in rabbit reticulocytes. Its isolation, characterization and translational repression. *European Journal of Biochemistry* **129**, 133-141.

TRIFILLIS, P., DAY, N. & KILEDJIAN, M. (1999). Finding the right RNA: identification of cellular mRNA substrates for RNA-binding proteins. *Rna* **5**, 1071-1082.

TRIVEDI, R.A. & DICKSON, G. (1995). Liposome-mediated gene transfer into normal and dystrophin-deficient mouse myoblasts. *Journal of Neurochemistry* **64**, 2230-2238.

VAHERI, A. & PAGANO, J.S. (1965). *Science* **175**, 434

VARNUM, S.M., HURNEY, C.A. & WORMINGTON, W.M. (1992). Maturation-specific deadenylation in Xenopus oocytes requires nuclear and cytoplasmic factors. *Developmental Biology (Orlando)* **153**, 283-290.

VASSALLI, J.D., HUARTE, J., BELIN, D., GUBLER, P., VASSALLI, A., O'CONNELL, M.L., PARTON, L.A., RICKLES, R.J. & STRICKLAND, S. (1989). Regulated polyadenylation controls mRNA translation during meiotic maturation of mouse oocytes. *Genes & Development* **3**, 2163-2171.

VIVARELLI, E., BROWN, W.E., WHALEN, R.G. & COSSU, G. (1988). The expression of slow myosin during mammalian somitogenesis and limb bud differentiation. *Journal of Cell Biology* **107**, 2191-2197.

VOYTIK, S.L., PRZYBORSKI, M., BADYLAK, S.F. & KONIECZNY, S.F. (1993). Differential expression of muscle regulatory factor genes in normal and denervated adult rat hindlimb muscles. *Developmental Dynamics* **198**, 214-224.

VRBOVA, G. & PETTE, D. Reanalysis: impulse activity and fiber-type transformation [letter]. *Muscle & Nerve* **10**, 569-Aug

WAHLE, E. & KELLER, W. (1992). The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors. [Review] [190 refs]. *Annual Review of Biochemistry* **61**, 419-440.

WAHLE, E. (1992). The end of the message: 3'-end processing leading to polyadenylated messenger RNA. [Review] [53 refs]. *Bioessays* **14**, 113-118.

WAHLE, E. (1995). Poly(A) tail length control is caused by termination of processive synthesis. *Journal of Biological Chemistry* **270**, 2800-2808.

WALSH, J.L., KEITH, T.J. & KNULL, H.R. (1989). Glycolytic enzyme interactions with tubulin and microtubules. *Biochimica et Biophysica Acta* **999**, 64-70.

WANG, C., DICKINSON, L.K. & LEHMANN, R. (1994). Genetics of nanos localization in Drosophila. *Developmental Dynamics* **199**, 103-115.

WANG, X., KILEDJIAN, M., WEISS, I.M. & LIEBHABER, S.A. (1995). Detection and characterization of a 3' untranslated region ribonucleoprotein complex associated with human alpha-globin mRNA stability [published erratum appears in Mol Cell Biol 1995 Apr;15(4):2331]. *Molecular & Cellular Biology* **15**, 1769-1777.

WANG, X. & LIEBHABER, S.A. (1996). Complementary change in cis determinants and trans factors in the evolution of an mRNP stability complex. *EMBO Journal* **15**, 5040-5051.

WANG, Z., DAY, N., TRIFILLIS, P. & KILEDJIAN, M. (1999). An mRNA stability complex functions with poly(A)-binding protien to stabilize mRNA in vitro. *Molecular & Cellular Biology* **19**, 4552-4560.

WEEKS, D.L. & MELTON, D.A. (1987). A maternal mRNA localized to the vegetal hemisphere in Xenopus eggs codes for a growth factor related to TGF-beta. *Cell* **51**, 861-867.

WEINTRAUB, H., CHENG, P.F. & CONRAD, K. (1986). Expression of transfected DNA depends on DNA topology. *Cell* **46**, 115-122.

WEINTRAUB, H., DAVIS, R., TAPSCOTT, S., THAYER, M., KRAUSE, M., BENEZRA, R., BLACKWELL, T.K., TURNER, D., RUPP, R. & HOLLENBERG, S. (1991). The myoD gene family: nodal point during specification of the muscle cell lineage. [Review] [69 refs]. *Science* **251**, 761-766.

WEISS, B. (1976). Endonuclease II of Escherichia coli is exonuclease III. *Journal of Biological Chemistry* **251**, 1896-1901.

WEISS, B. (1976). Terminal cross-linking of DNA strands by a mixture of exonuclease III of Escherichia coli and polynucleotide ligase of bacteriophage T4. *Journal of Molecular Biology* **103**, 669-673.

WEISS, I.M. & LIEBHABER, S.A. (1994). Erythroid cell-specific determinants of alpha-globin mRNA stability. *Molecular & Cellular Biology* **14**, 8123-8132.

WEISS, I.M. & LIEBHABER, S.A. (1995). Erythroid cell-specific mRNA stability elements in the alpha 2-globin 3' nontranslated region. *Molecular & Cellular Biology* **15**, 2457-2465.

WEYDERT, A., DAUBAS, P., CARAVATTI, M., MINTY, A., BUGAISKY, G., COHEN, A., ROBERT, B. & BUCKINGHAM, M. (1983). Sequential accumulation of mRNAs encoding different myosin heavy chain isoforms during skeletal muscle development in vivo detected with a recombinant plasmid identified as coding for an adult fast myosin heavy chain from mouse skeletal muscle. *Journal of Biological Chemistry* **258**, 13867-13874.

WEYDERT, A., DAUBAS, P., LAZARIDIS, I., BARTON, P., GARNER, I., LEADER, D.P., BONHOMME, F., CATALAN, J., SIMON, D. & GUENET, J.L. (1985). Genes for skeletal muscle myosin heavy chains are clustered and are not located on the same mouse chromosome as a cardiac myosin heavy chain gene. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 7183-7187.

WEYDERT, A., BARTON, P., HARRIS, A.J., PINSET, C. & BUCKINGHAM, M. (1987). Developmental pattern of mouse skeletal myosin heavy chain gene transcripts in vivo and in vitro. *Cell* **49**, 121-129.

WHALEN, R.G., SELL, S.M., BUTLER-BROWNE, G.S., SCHWARTZ, K., BOUVERET, P. & PINSET-HARSTOM, I. (1981). Three myosin heavy-chain isozymes appear sequentially in rat muscle development. *Nature* **292**, 805-809.

WHALEN, R.G., JOHNSTONE, D., BRYERS, P.S., BUTLER-BROWNE, G.S., ECOB, M.S. & JAROS (1984). A developmentally regulated disappearance of slow myosin in fast-type muscles of the mouse. *FEBS Letters* **177**, 51-56.

WHARTON, R.P. & STRUHL, G. (1989). Structure of the Drosophila BicaudalD protein and its role in localizing the the posterior determinant nanos. *Cell* **59**, 881-892.

WHARTON, R.P. & STRUHL, G. (1991). RNA regulatory elements mediate control of Drosophila body pattern by the posterior morphogen nanos. *Cell* **67**, 955-967.

WICKENS, M. & STEPHENSON, P. (1984). Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3' end formation. *Science* **226**, 1045-1051.

WILHELM, J.E. & VALE, R.D. (1993). RNA on the move: the mRNA localization pathway [published erratum appears in J Cell Biol 1993 Dec;123(6 Pt 1):1625]. [Review] [74 refs]. *Journal of Cell Biology* **123**, 269-274.

WISEMAN, J.W., GLOVER, L.A. & HESKETH, J.E. (1997). Evidence for a localization signal in the 3'untranslated region of myosin heavy chain messenger RNA. *Cell Biology International* **21** , 243-248.

WODNAR-FILIPOWICZ, A. & MORONI, C. (1990). Regulation of interleukin 3 mRNA expression in mast cells occurs at the posttranscriptional level and is mediated by calcium ions. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 777-781.

WONG, T.K. & NEUMANN, E. (1982). Electric field mediated gene transfer. *Biochemical & Biophysical Research Communications* **107**, 584-587.

WU, B.J., KINGSTON, R.E. & MORIMOTO, R.I. (1986). Human HSP70 promoter contains at least two distinct regulatory domains. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 629-633.

YAFFE, D. & SAXEL, O. (1977). Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature* **270**, 725-727.

YAFFE, D., NUDEL, U., MAYER, Y. & NEUMAN, S. (1985). Highly conserved sequences in the 3' untranslated region of mRNAs coding for homologous proteins in distantly related species. *Nucleic Acids Research* **13**, 3723-3737.

YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103-119.

YE, G.N., DANIELL, H. & SANFORD, J.C. (1990). Optimization of delivery of foreign DNA into higher-plant chloroplasts. *Plant Molecular Biology* **15**, 809-819.

YISRAELI, J.K., SOKOL, S. & MELTON, D.A. (1990). A two-step model for the localization of maternal mRNA in Xenopus oocytes: involvement of microtubules and microfilaments in the translocation and anchoring of Vg1 mRNA. *Development* **108**, 289-298.

YOU, Y., CHEN, C.Y. & SHYU, A.B. (1992). U-rich sequence-binding proteins (URBPs) interacting with a 20-nucleotide U-rich sequence in the 3' untranslated region of c-fos mRNA may be involved in the first step of c-fos mRNA degradation. *Molecular & Cellular Biology* **12**, 2931-2940.

YU, Y.T., BREITBART, R.E., SMOOT, L.B., LEE, Y., MAHDAVI, V. & NADAL-GINARD, B. (1992). Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. *Genes & Development* **6**, 1783-1798.

ZEHNER, Z.E., SHEPHERD, R.K., GABRYSZUK, J., FU, T.F., AL-ALI, M. & HOLMES, W.M. (1997). RNA-protein interactions within the 3 ' untranslated region of vimentin mRNA. *Nucleic Acids Research* **25**, 3362-3370.

ZHANG, B., GALLEGOS, M., PUOTI, A., DURKIN, E., FIELDS, S., KIMBLE, J., WICKENS & MP (1997). A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line. *Nature* **390**, 477-484.

ZHANG, W., WAGNER, B.J., EHRENMAN, K., SCHAEFER, A.W., DEMARIA, C.T., CRATER, D., DEHAVEN, K., LONG, L. & BREWER, G. (1993). Purification, characterization, and cDNA cloning of an AU-rich element RNA-binding protein, AUF1. *Molecular & Cellular Biology* **13**, 7652-7665.

ZUKER, M. (1989). Computer prediction of RNA structure. *Methods in Enzymology* **180**, 262-288.