National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

*Your file Votre référence*

*Our file Notre référence*

## NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

## AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

# THE MYOTONIC DYSTROPHY KINASE 3'UNTRANSLATED REGION AND ITS EFFECT ON GENE EXPRESSION

A thesis submitted to the School of Graduate Studies at the University of Ottawa in partial fulfillment of the requirements for the degree of Master of Science, Department of Microbiology and Immunology, Faculty of Medicine.

By Caroline W.-Y. Ang

ISBN 0-612-07834-5

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

## ABSTRACT

Myotonic dystrophy (DM) is a dominantly inherited neuromuscular disease known to be the most commonly occurring form of inherited muscular dystrophy in adults. The mutation responsible for DM has been found to be an expansion of a CTG trinucleotide repeat in the 3'untranslated region (3'UTR) of a gene encoding a putative serine-threonine kinase. A positive correlation between repeat size and disease severity has been observed providing a molecular explanation for genetic anticipation associated with the transmission of DM.

A mutation occurring outside the coding region in the 3'UTR of the myotonic dystrophy kinase (DMK) suggests the formation of a defective gene product may not be involved in the disease process. Instead, the effect of the repeat expansion may be directed toward gene regulation or expression. In this study, investigation of the possibility of cellular factors capable of directly binding DNA or RNA revealed the presence of at least two factors in cytosolic extract. Mobility shift assays were done using an RNA probe comprising a portion of the 3'UTR containing the repeat region and approximately 150bp of upstream and downstream flanking sequence. Probes with repeat sizes of 13 repeats, 45 repeats, and 90 repeats did not appear to vary greatly in gel mobility complex size. Binding might therefore not directly involve the repeat but may in fact be specific to the surrounding sequences. Binding factors sharing similar mobilities were found in cytosolic extracts derived from congenital skeletal muscle and cardiac tissue, as well as mouse liver. Since the 3'UTR sequence appears able to bind proteins from a diverse range of tissues, the interacting factors may be commonly found in the cytoplasm of many cells.

Half-life studies of endogenously expressed DMK mRNA from TE32 rhabdomyosarcoma cells suggest the transcript may have a half-life of over seven hours. Transcription was arrested using actinomycin D and total RNA extracted at various time points for Northern analysis. A probe to detect the ß-actin transcript was used as a control

for long-term stability. The two signals remained comparatively constant throughout the seven-hour assay period suggesting DMK mRNA may be relatively long-lived in the cells at fairly constant levels.

Finally, clones using the CAT reporter gene (Gorman et al., 1982) with the DMK 3'UTR containing repeat sizes from 5 to 90 repeats were constructed. Transient transfection of the constructs into TE32 cells and assay for CAT gene expression revealed increased CAT activity correlating with increasing repeat size. Addition of the DMK 3'UTR alone with a wildtype number of repeats (five and eleven repeats) was enough to boost CAT activity compared to the same construct with a 3'UTR consisting of only a polyadenylation signal. When the size of the repeat was increased to mutation range (45 and 90 repeats), CAT activity was further increased. As well, deletion of portions of the 3'UTR sequence, either up- or downstream of the repeat region abrogated any CAT activity from construct containing these variants. The DMK 3'UTR would therefore appear to have an expression enhancing effect that becomes more pronounced as repeat number increases. The complete inactivity of the deletion clones suggests the repeat sequence must be presented in the context of the full 3'UTR to impose any regulatory control.

Taken together, these data suggest a role for the DMK 3'UTR in the regulation of gene expression. mRNA half-life studies of endogenous DMK revealed the transcript to be quite stable having a half-life of at least seven hours. Factors binding to the transcript may be involved in regulating stability. Cytosolic factor(s) detected by gel-shift assay was found to bind probes comprising the repeat and surrounding regions. Whether regulation occurs at the transcriptional or post-transcriptional stage is a subject for future investigation. Nevertheless, the autosomal dominant inheritance pattern of DM and the observation of genetic anticipation suggest a gain-of-function mutation model with increasing repeat size leading to greater gene activity and therefore greater disease

severity. This mechanism may prove to be an involved process, possibly requiring controls such as regulation of mRNA stability and mRNA binding factors.

## ACKNOWLEDGMENTS

I thank Dr. Robert G. Korneluk for his guidance and unflagging support of this work. Helpful and critical discussion provided by Monica Narang, Luc Sabourin and Dr. Mani Mahadevan is greatly appreciated. I also thank Stephen Baird for providing computer expertise, and all my colleagues for their generous encouragement and tolerance. This work was supported by operating grants for R.G.K. from the Muscular Dystrophy Associations of Canada and the United States of America, The Medical Research Council of Canada, and the Canadian National Centres of Excellence Genetic Disease Network.

# TABLE OF CONTENTS

## LIST OF FIGURES

# CHAPTER I- INTRODUCTION TO MYOTONIC DYSTROPHY

## General Introduction

The degenerative neuromuscular disease myotonic dystrophy (DM) is the most common inherited form of adult muscular dystrophy with an incidence of approximately 1 in 8000 worldwide. The autosomal dominantly inherited mutation has been discovered to be expansion of a CTG trinucleotide repeat sequence mapping to the 3'-untranslated region (3'UTR) of a gene coding for a serine-threonine kinase (the myotonic dystrophy kinase, or DMK).

Expanding trinucleotide repeats have been characterized in several autosomal dominant genetic disease genes isolated within the last few years. X-linked spinal and bulbar muscular atrophy (SBMA or Kennedy's disease) (La Spada et al., 1991), spinocerebellar ataxia type 1 (SCA1) (Orr et al., 1993; Banfi et al., 1994), Huntington's disease (HD) (The Huntington's Disease Collaborative Research Group, 1993), dentarubral-pallidoluysian atrophy (DRPLA) (Koide et al., 1994) and its variant Haw River Syndrome (Burke et al., 1994), and Machado-Joseph disease (MJD) (Kawagucki et al., 1994) are all progressively neurodegenerative and with the exception of X-linked SBMA are all autosomal dominant. The mutant alleles characteristically display an expanding CAG repeat in the coding region near the amino terminus. *FMR-1*, the gene responsible for fragile X disease mental retardation (FRAXA) was found to contain an unstable CCG repeat in the 5'untranslated region (Oberlé et al., 1991; Verkerk et al., 1991; Fu et al., 1991). In affected families, all diseases including DM show increasing repeat size in successive generations. The expansion coincides with increased disease severity indicative of the phenomena of genetic anticipation. The mechanism of mutation for the CAG expansion is postulated to be alteration of protein structure or function resulting from an elongated polyglutamine tract. The CCG expansion leads to hypermethylation thereby shutting off gene expression from the mutant allele (Oberlé et al., 1991; Knight et al., 1993). The DMK

gene is the only one of the expanding trinucleotide repeat diseases discovered so far to contain the mutation in the 3'UTR. A number of molecular mechanisms could explain the effect of a trinucleotide repeat expansion in the 3'untranslated region of a gene. For example, RNA splicing might be affected giving rise to alternate protein isoforms (Whiting et al., 1995), or mRNA steady-state levels could be altered (Sabourin et al., 1993). Another possibility could involve the mutation imposing a direct influence on gene expression. How this might occur will be discussed on the following pages.

**Pathology of DM**

The well characterized, highly variable clinical phenotype of DM involves several systems. In addition to myotonia (an inability to relax contracted muscle) and progressive muscle weakness and wasting, cardiac conduction and smooth muscle defects, cataracts, hypersomnia, mental deterioration, abnormal glucose response, premature male pattern baldness and testicular tubular atrophy are also observed (Harper, 1989). Severity of disease varies from individual to individual. The mildest form can escape diagnosis and manifests itself in middle or old age by cataract formation and little, if any muscle defects. A severe congenital disorder clinically distinct from the adult form displays hypotonia, facial displegia, neonatal respiratory distress, mental retardation and poor neonatal viability. Interestingly, the congenital form is only seen with maternal transmission.

**Cloning of the DM gene and the mutation**

Since no gross chromosomal defects could be detected in individuals with DM using cytogenetic techniques, molecular biology methods had to be employed to uncover the location of the gene. Efforts to clone the gene mapped it to the long arm of chromosome 19 (Davies et al., 1983) and more specifically to region 19q13.3 after physical (Korneluk et al., 1989; Smeets et al., 1990; Brook et al., 1991a, 1991b) and genetic analysis (Brunner et al., 1989; Johnson et al., 1990; Harley et al., 1991; Tsilfidis et al., 1991). Southern blot analysis

identified in normal individuals the presence of an 8.5 or 9.5 kb HindIII or a 9.0 or 10.0 kb EcoRI insertion polymorphism when probed with genomic fragments mapping within a 10 kb Eco RI genomic band contained in 19q13.3 (Harley et al., 1992a; Buxton et al., 1992; Brook et al., 1992). In affected individuals, a larger allele of variable size was detected and found by genetic phasing to be derived from the 10 kb EcoRI allele. Upon isolation of cDNAs and sequence analysis, the expansion was found to comprise a CTG trinucleotide repeat mapping between the stop codon and the polyadenylation tail, therefore within the 3'untranslated region of the gene (Mahadevan et al., 1992; Brook et al., 1992; Fu et al., 1992). Approximately 70% of DM individuals display definite expansion by Southern analysis while the other 30% show detectable increases in allele size only by PCR assay which can amplify the repeat region specifically. As the expansion reaches sizes of several hundred repeats, mutant alleles probed on Southern blots appear less distinct compared to normal and less expanded alleles. CTG-repeats are therefore probably heterogeneous and unstable in somatic cells of individuals with large expansions. The CTG repeat amplification has been detected in at least 99% of all reported DM cases.

Repeat size in the normal population is highly polymorphic ranging from 5 to 35 repeats with 5 repeats and 13 repeats being the most prevalent alleles (35% and 19%, respectively) (Mahadevan et al., 1992). Progressive allelic expansion has been shown to occur in successive affected generations correlating with increasing disease severity (Hunter et al., 1992; Tsilfidis et al., 1992; Harley et al., 1992b) giving a molecular basis for the phenomena of genetic anticipation thought to exist in DM (Höweler et al., 1989). Indeed, two cases report reduction to wildtype range of an expanded allele upon transmission from a clearly affected father to offspring displaying no evident DM symptoms (Brunner et al., 1993; O'Hoy et al., 1993). Such evidence lends further support to the hypothesis that repeat amplification is associated with disease phenotype. In most cases of noncongenital DM, expansions of a few hundred repeats are observed while expansions of around 2000 repeats are often seen in congenital cases. Mothers of congenital offspring also display a higher

degree of expansion compared to mothers of less affected individuals indicating large maternal amplifications may place offspring at higher risk for congenital DM (Tsilfidis et al., 1992). Individuals possessing approximately 50 to 80 repeats are said to carry the protomutation indicating they display a very mild or asymptomatic form of DM (Barceló et al., 1993). The protomutation can be passed with little or no amplification from generation to generation until, by an as yet undetermined mechanism, stability is lost and intergenerational expansion occurs. Repeat instability has been positively correlated to repeat size with protomutations in the upper threshold being more prone to amplification. Repeat sizes of 80 or more almost always display an expanded allele transmission.

In genes where the triplets fall within the coding region, an approximate doubling of 40 to 100 repeats is observed in the affected population compared to the normal 6 to 37 range. In the case of *DMK* and *FMR-1* which normally have 5 to 37 repeats located in the 3'- and 5'-untranslated regions, respectively, mutation size can range from 50 repeats up into the thousands. Thus, there would appear to be a limiting effect on expansion size imposed by it being translated. The cause of expansion of these dynamic trinucleotide repeats has been the subject of much speculation. Both *cis*-acting elements and *trans*-acting factors have been suggested as well as a slippage-mediated model that accounts for both small and large amplifications (reviewed by Richards and Sutherland, 1994). Recombination would not appear to be a factor in repeat expansion, at least in the case of DM, since a uniform recombination frequency was observed over a 1.5 Mb region which included the DM locus (Shutler et al., 1993).

Haplotype analysis suggests DM exhibits a strong founder effect since extended analysis over a distance of 1.3 Mb reveals three major haplotypes representing 45.6% of DM chromosomes (Tsilfidis et al., 1994). High resolution genetic analysis focusing on a 30 kb region encompassing the DM gene demonstrated the presence of only one haplotype on all DM chromosomes studied (Neville et al., 1994). It is possible that the haplotype background against which the mutation occurred might actually predispose the CTG repeat

to amplify. The reduced reproductive fitness of individuals with congenital DM should have resulted in a gradual decline in disease incidence, but instead, DM is apparently being maintained in the population. The inability of protomutation carriers to be detected until diagnosis of a more severely affected relative could account for a larger pool of those predisposed to DM than is currently documented. Meiotic drive may be a possible mechanism for maintenance of larger repeat alleles in the population. Preferential transmission of the larger of two normal alleles has been observed both maternally and paternally particularly with CTG n>19 (Carrey et al., 1994; Shaw et al., 1994). However, the normal and affected populations appear to be mutually exclusive since transmission of a protomutation allele derived from a normal allele has never been observed. How the jump might be made from normal to mutation repeat range is unclear.

**Gene structure**

Sequence analysis has revealed the gene responsible for DM contains 15 exons spanning a distance of approximately 14 kb (Mahadevan et al., 1993). Genomic DNA and cDNA comparison showed the gene to be transcribed in the telomeric to centromeric direction with the CTG repeat sequence transcribed on the coding strand. Northern blot analysis indicates a 3.0-3.3 kb mRNA transcript most strongly expressed in heart, skeletal muscle to a lesser degree, weakly in lung and bladder (both of which are comprised of smooth muscle), and brain (Brook et al., 1992; Jansen et al., 1992). Homology searches have revealed the amino acid sequence encoded by exons 2 to 8 shares up to 40% identity with the catalytic core domain of the serine-threonine protein kinase family that includes the S6 specific kinases, cAMP-dependent kinases and protein kinase C. Weaker homology to an -helical coiled-coil motif similar to that found in myosin and myofibrilar or filamentous proteins in exons 9 to 12, and a hydrophobic stretch of amino acids possibly representing a transmembranous region encoded by exon 15 were also detected. Transcripts from cardiac tissue lack exons 13 and 14. This change results in an open reading frame shift and the use

of a termination codon located 5 bp from the start of exon 15 effectively removing the carboxy-terminal hydrophobic region (Mahadevan et al., 1993). Thus far, three potentially protein-encoding mRNA isoforms have been detected in humans (Jansen et al., 1992). In addition to the full-length and the cardiac-specific isoform mentioned above, a cDNA clone was isolated from human brain frontal cortex and found to be missing the last five codons of exon 8 as the result of the use of a cryptic 5' splice site.

The presence of another gene, designated DMR-N9, located just upstream of DMK has been found as a 3 kb transcript expressed in brain and testes (Jansen et al., 1992). No significant amino acid sequence homology to previously reported proteins has been demonstrated. Like DMK, it is transcribed from telomere to centromere and the 3'UTR of some partial cDNAs isolated has been mapped to just a few hundred bases telomeric to the kinase domain. Whether DMR-N9 has any effect on the expression of DMK or in the manifestation of the disease is not known. Little is also known about the promoter of DMK of which part could potentially lie within DMR-N9. No consensus CAAT or TATA box sequences have been detected in the region upstream of the DMK start codon although a number of putative transcription factor binding sites are present. Interestingly, four E-box sequences, as found in many muscle gene-specific promoters and enhancers, are located in the first intron of DMK and have been shown to increase transcription in a location- and orientation-specific manner in conjunction with sequences extending 5' of the start site (Storbeck et al., 1994).

The amino acid sequence of the longest DMK isoform predicts a protein of approximately 69 kDa. Polyclonal rabbit antisera raised against a glutathione-S-transferase (GST) fusion protein encoding exons 11 to 15 of DMK recognizes 74kDa and 82kDa protein isoforms in Western blots containing human and skeletal muscle, ependyma and choroid plexus. In brain, smaller species have also been detected ranging in size from 45kDa to 50 kDa (Whiting et al., 1995). DMK protein generated by a recombinant baculovirus system expressing either full-length or a truncated isoform missing sequence

from exons 13 to 15 was also detected with the antisera at sizes comparable to their predicted molecular weight (Waring et al., 1994). Immunohistochemical studies have localised DMK to the postsynaptic region of neuromuscular junctions in both human and rodent skeletal muscle as well as in the cytoplasm of skeletal and smooth muscle (van der Ven et al., 1993; Whiting et al., 1995). DMK was also detected in intercalated disks of rat quadricep muscle. No gross histological differences were observed in muscle from normal compared to affected individuals. The distribution of DMK protein in skeletal muscle sections from adult and congenital patients, compared to non-DM controls, as well as the level of protein also did not appear to differ significantly (Sabourin et al., 1995).

**The 3'UTR and regulation of gene expression**

The location of the expansion mutation of DM occurring in the 3'UTR raises the question of a potential role for the sequence in DMK gene expression. Regulatory control regions located in the 3'UTR have been identified in several genes. c-*myc* and c-*fos*, two proto-oncogenes belonging to the immediate-early gene (IEG) class (Brewer and Ross, 1988; Kabnick and Housman, 1988), as well as the lymphokine granulocyte-macrophage colony-stimulating factor (GM-CSF) (Shaw and Kamen, 1986), are all transiently expressed and possess an AU-rich element (ARE) in their 3'UTRs which targets the mRNA for rapid decay. ARE sequences are also in the 3'UTRs of the *pim*-1 proto-oncogene (Wingett et al., 1991), and the urokinase-type plasminogen activator (uPA) gene (Nanbu et al., 1994). Insertion of the ARE consensus sequence into the 3'UTR of the rabbit B-globin gene resulted in a reduction in mRNA half-life from greater than 17 hours, as seen normally, to less than 30 minutes (Shaw and Kamen, 1986). Rapid degradation of c-*fos* and c-*myc* mRNA occurs in two steps: shortening and removal of the poly(A) tail followed by degradation of the remainder of the transcript (Brewer and Ross, 1988; Shyu et al., 1991). The ARE appears to be needed only for the second stage. Proteins have been identified that bind ARE sequences (Vakalopoulou et al., 1991; Brewer, 1991; Bohjanen et al., 1991; You

et al., 1992; Bickel et al., 1992; Savant-Bhonsale and Cleveland, 1992; Zhang et al., 1993) and some have been shown to be tightly associated with mRNA degradation. In the case of GM-CSF, ribosome binding, or ongoing translation of the mRNA is known to be needed for ARE-mediated degradation together with the presence of a >20S divalent cation-independent complex (Savant-Bhonsale and Cleveland, 1992; Aharon and Schneider, 1993). Another protein, AU-A known to bind AREs and other U-rich RNA sequences has been reported to shuttle from the nucleus to the cytoplasm when cells are transcriptionally inhibited possibly to prevent degradation of AU-rich transcripts (Katz et al., 1994).

Postranscriptional regulation is also involved in expression of the transferrin receptor gene mRNA. The transferrin receptor binds iron when iron is bound to transferrin. The complex is endocytosed and iron is released to the cytosol where it is used in cellular processes or sequestered by ferritin. Present in the 3'UTR are five stem-loop structures known as iron-responsive elements (IREs) (Müllner and Kühn, 1988) that are bound by the cytoplasmic IRE-binding protein (IRE-BP) when iron levels are low (Müllner et al., 1989). Binding of IRE-BP stabilizes transferrin mRNA and allows increased translation. A second unidentified element aside from the IREs is required for actual mRNA destabilization and it is thought binding of the IREs by IRE-BP blocks the association of degradation factors with this sequence (reviewed by Sachs, 1993).

In the different mRNA isoforms encoding the PR264/SC35 general splicing factor required for the early steps of splice site selection, only the 3'UTR sequence differs (Sureau and Perbal, 1994). Generated by alternative splicing, the various isoforms possess different half-lives with the shortest 3'UTR being the most stable. Although no ARE sequence was detected, two regions showing high conservation between human and chicken may contain novel elements involved in RNA turnover since both sequences were present in the mRNA with the shortest half-life.

In addition to mRNA degradation, the 3'UTR has been shown to contain elements responsible for controlling translation. Among the proteins required for sperm maturation in

mice are the protamines which are small basic proteins that take the place of histones and testis basic proteins during sperm nuclear condensation. Using a human growth hormone reporter gene system, Braun et al., (1989) discovered a sequence in the 3'UTR of the mouse protamine 1 gene (mPrm-1) capable of conferring translational suppression of mRNA in haploid round spermatids until maturation to the elongation phase. This temporal effect provided by the mPrm-1 3'UTR also influenced subcellular localization with early expression occurring in the acrosome while appropriate later expression was observed intracellularly, not acrosomally. In the related mammalian protamine 2 (Prm-2) mRNA, the highly conserved Y and H elements in the 3'UTR were found to specifically bind an 18 kDa phosphoprotein when translationally repressed (Kwon and Hecht, 1991; Kwon and Hecht, 1993). Similarly, the 3'UTR of tumor necrosis factor (TNF), a cytokine produced predominantly by macrophages, functions as a translational suppressor of mRNA until the time of macrophage stimulation (Han et al., 1991).

3'UTRs have also been found to be involved in gene regulation via an *in trans* pathway. The 3'UTRs from the muscle structural genes tropomyosin, troponin I, and -cardiac actin were able to initiate differentiation in a differentiation-defective mutant myoblast cell-line (NMU2) and suppress proliferation in 10T1/2 fibroblasts (Rastinejad and Blau, 1993). Additional work revealed expression of a 0.2Kb region from the $\alpha$-tropomyosin 3'UTR with no open reading frame sequence was sufficient to inhibit tumor formation by NMU2 cells upon injection into adult mouse muscle (Rastinejad et al., 1993).

# CHAPTER II - FUNCTION OF THE DMK 3'UTR

## INTRODUCTION

### Effect of the expansion mutation in DM

The switch from proliferation to differentiation is tightly regulated in muscle cells. Ectopic expression of the MyoD family of basic helix-loop-helix (bHLH) proteins, which includes MyoD, Myf-5, MRF4 and myogenin, can arrest proliferation and induce nonmuscle cells to transcribe skeletal muscle genes (reviewed by Olson and Klein, 1994). Mouse C2C12 myoblasts, which endogenously express the murine form of DMK, undergo terminal differentiation to myotubes when exposed to serum deprivation or upon reaching confluence. In experiments using C2C12 cells, Sabourin et al. (1994) demonstrated over-expression of a full-length human DMK cDNA containing 5 CTG repeats blocked cell terminal differentiation. Under-expression of DMK, achieved by the introduction of antisense DMK oligonucleotides, conversely appeared to enhance myotube fusion suggesting DMK may be involved in myogenesis. The degree of DMK expression in normal and affected cases has been the subject of some controversy. Increased steady-state levels of mRNA in congenital compared to normal tissue have been shown, suggesting elevated DMK levels may be responsible for the disease phenotype (Sabourin et al., 1993).

On the other hand, contrary results have also been reported. Fu et al., (1993) have reported reduced mRNA expression from the mutant allele in adult DM cases, while Funanage et al., (1994) have published similar results for congenital DM. Equal expression of hnRNA from both alleles in adult and congenital DM but reduced mutant mRNA levels due to aberrant processing (Krahe et al., 1994), and no mutant allele expression accompanied by increased normal allele expression (Radvanyi et al., 1994) have also been observed.

Transgenic mouse studies have been undertaken by some groups and preliminary results with DM knockout mice show no detectable phenotypic abnormalities even when bred to homozygosity (D. Housman, Nature Genetics meeting, April 1994; B. Wieringa, personal communication). Since the complete absence of a functional DMK gene does not appear to affect development, it may serve a redundant function much like in the case of *MyoD* where abrogation of *MyoD* expression in mice is compensated by increased *Myf-5* mRNA levels (Rudnicki et al., 1992). An overexpressing mouse model has also been generated by random integration of tandem copies of a genomic fragment containing the human DMK gene (B. Wieringa and M. Narang, personal communication). Mice containing the largest number of transgene copies (up to 30) display evidence of cardiac muscle necrosis but no skeletal muscle defects. The cardiac muscle defect seems to be directly due to the effect of the transgene and not to the disruption at the site of integration since it was observed in two different mouse lines each with a different point of integration. Therefore, although some have suggested that the highly variable disease phenotype observed in DM stems from reduced levels or reduced activity of DMK from the mutant allele and/or normal allele, a more plausible explanation is a gain of function mutation could be responsible for the dominant inheritance pattern of DM. This hypothesis is supported by *in vitro* and *in vivo* over-expression studies conducted, respectively, with C2C12 myoblasts and transgenic mice.

**Studying the DMK 3'UTR**

The main genetic alteration in mutant *DMK* being an expanding CTG repeat in the 3'UTR implies that over-expression seems the most probable mechanism of mutation in DM. One can speculate that the 3'UTR, including the repeat sequence, is somehow involved in gene regulation. To investigate this hypothesis, I have undertaken to characterize the 3'UTR of *DMK* and analyze its effect on gene expression using a variety of methods. DNA-protein gel mobility shift assays were performed using both double- and

single-stranded probes containing 13 repeats (wildtype) or 63 repeats (protomutation). No specific DNA-protein bands were seen with either double-stranded or single-stranded probes. When mobility shifts using RNA transcripts comprising the CTG repeat and surrounding regions were assayed, we did detect shifted complexes. Protein extracts isolated from a number of different tissues of congenital individuals were found to form complexes with transcripts generated from clones containing 13 repeats (wildtype), 35 repeats(wildtype), and 80 repeats (protomutation). The complexes did not appear to differ with increasing repeat size and were seen in only cytosolic protein extracts, particularly with those from skeletal muscle. These results suggest protein binding specifically to the DNA repeat region may not occur physiologically, but do support the presence of a cytosolic protein(s) capable of binding DMK 3'UTR mRNA.

To study the behavior of the mRNA in the cell, mRNA half-life studies were performed using actinomycin D to arrest transcription in human TE32 rhabdomyosarcoma cells, which express DMK endogenously. Preliminary experiments demonstrated DMK mRNA to have a half-life of several hours. Whether half-life is altered with an increasing repeat expansion or whether binding proteins are involved has yet to be determined.

To study the effect of the DMK 3'UTR on heterologous gene expression, a reporter gene system was constructed. The constitutive CMV promoter was used to drive the chloramphenicol acetyl transferase (CAT) gene joined to the complete DMK 3'UTR containing 5 repeats, 11 repeats, 45 repeats, and 90 repeats. CAT is originally derived from the E. coli transposable element Tn9 that confers antibiotic resistance to chloramphenicol. Similar studies using the 3'UTRs of GM-CSF, c-*myc*, and c-*fos*, fused to a *neo* marker gene were done by Schuler and Cole (1988). Neither CAT nor *neo* is endogenous to mammalian cells making detection easily distinguishable from regular cellular activities. In addition, CAT has the added advantage of a well-documented, and efficient protocol for measuring expression levels by CAT activity (Gorman et al., 1982). CAT fusions have been used to investigate putative regulatory control regions in several genes including the promoter and

3'UTR of the mouse TNF gene (Han et al., 1991; Beutler and Brown, 1991). I have shown that upon transient transfection into TE32 cells, a three-fold increase in CAT activity is observed from constructs containing a wildtype 3'UTR (5 and 11 repeats) compared to a control construct containing only a poly(A) signal. Reporter constructs with repeats in the protomutation (45 repeats) and mutation (90 repeats) range showed a greater than 10-fold increase over control CAT activity. As well, deletion mutants were constructed missing regions upstream and downstream of the repeat. No CAT activity was evident with any of the deletion clones. Together, the data indicate the repeat sequence, and both up- and downstream regions of the DMK 3'UTR contain elements capable of regulating gene activity and conferring increased gene expression.

## MATERIALS AND METHODS

### Preparation of extracts from tissue and cultured cells

For DNA-binding assays, nuclear proteins were extracted from congenital DM diaphragm according to the method as outlined by Dignam et al. (1983). S100 extract for RNA-binding assays was obtained from human normal and congenital DM brain, congenital DM diaphragm, congenital DM heart, and mouse liver. Tissue was first frozen solid in liquid nitrogen, then ground by mortar and pestle to a fine dust. All further homogenizations were done using a microtube-size Teflon hand-held homogenizer.

### Construction of plasmids

pTA-13 (13 CTG repeats), pTA-45 (45 CTG repeats) and pTA-90 (90 CTG repeats) were generated by subcloning the SmaI/HindIII fragment found in the DMK 3'UTR (gift from M. Narang) into pBluescript SK$^-$. These clones were subsequently used for generating labeled transcripts in RNA-protein binding assays.

The pCAT-Basic vector (Promega) was digested with Sau3AI yielding a 791bp fragment containing the entire chloramphenicol acetyl transferase (CAT) open reading frame (ORF). The CAT ORF was ligated into the BamHI site of the pcDNA3 vector (Invitrogen) and the construct named pcDNA3-CAT (fig. 6a).

The 3'UTR of DMK was PCR amplified from a cosmid (11 repeats) using primer 1137 which started 22bp upstream of the stop codon and primer 1140 that ended 34bp downstream of the polyadenylation signal. The amplified product was TA-cloned into the pCRII vector (Invitrogen) (orientation unknown) then digested with BstXI resulting in a fragment 809bp in size (fig. 6a). This fragment was ligated into the BstXI site of pcDNA3-CAT generating pcDNA3-CAT-11rpt 3'UTR. Deletion clones were generated in a similar manner to give pcDNA3-CAT-Del1 (primers 1137 and 1139), pcDNA3-CAT-Del2 (primers

1137 and 410), pcDNA3-CAT-Del3 (primers 409 and 1140), and pcDNA3-CAT-Del4 (primers 1138 and 1140) (fig. 6b).

primer 409 5' GAAGGGTCCTTGTAGCCGGGAAT

primer 410 5' AGAAATGGTCTGTGATCCC

primer 1137 5' CCCAGGAGCCGCCCGCGCTCCCTGAACCC

primer 1138 5' GGGGGATCACAGACCATTTCT

primer 1139 5' CATTCCCGGCTACAAGGACCCTTC

primer 1140 5' GTGGAGTCCAGAGCTTTGGGCAGATGGAGG

The 5 repeat, 45 repeat, and 90 repeat clones were constructed by subcloning the 1.6Kb KpnI/XbaI fragment (containing CAT-11rpt 3'UTR) from pcDNA3-CAT-11rpt 3'UTR into pBluescriptSK- (Stratagene) digested with XbaI/KpnI. pBS-CAT-11rpt 3'UTR was digested with SmaI/HindIII, thereby removing the repeat, and replaced with SmaI/HindIII fragments from a clone with 5 repeats, and the pTA clones containing 45 repeats and 90 repeats. The new constructs were digested with KpnI/XbaI and ligated back into pcDNA3 KpnI/XbaI to generate pcDNA3-CAT-5rpt 3'UTR, -45rpt 3'UTR, and -90rpt 3'UTR (fig. 6a).

**DNA-binding assays**

The 13-repeat and 63-repeat fragments were generated by polymerase chain reaction (PCR) using primers 20 nucleotides long flanking the repeat region. The fragments were end-labeled by incubating at 37°C for one hour: 100ng of DNA with 10µl gamma$^{32}$P-ATP (100µCi), 3µl One-Phor-All buffer (Pharmacia), 2µl (2 units) T4 polynucleotide kinase (Gibco-BRL) and $H_2O$ TO 30µl. The reaction was stopped with 20mM EDTA and extracted with an equal volume of phenol/chloroform. DNA was precipitated with 10mM $MgCl_2$, 3M ammonium acetate and 2.5-fold volume of absolute ethanol, then resuspended in TE and precipitated again. The final pellet was resuspended in TE to 15,000 cpm/µl. Flanking probes with no repeat sequence were obtained by digestion of labeled 13-repeat

probe with Fnu4HI. Immediately after labeling, the T4 kinase was denatured by incubating at 75°C for 10 minutes. Exactly 10 units of Fnu4HI were added and the digest incubated at 37°C for 1 hour. DNA precipitation was as explained above.

Binding reactions contained 2µl 10x bandshift buffer (250mM HEPES (pH 7.9), 50mM $MgCl_2$, 340mM KCl),1.5µl poly dI/dC (1 µg/µl), 6ul BSA (1 µg/µl), 15 000cpm DNA probe, 7.5µg nuclear protein extract and $H_2O$ to 20µl. The components were left on ice for 15 minutes, then 4µl loading dye was added (20% w/v Ficoll 400, 0.1% bromophenol blue) and the mixture run on a 4% non-denaturing acrylamide gel. The binding assay was conducted using both annealed and denatured probe. The probes were denatured by boiling probe and $H_2O$ for 5 minutes, then quickly putting on ice. Remaining reactants were then added. As a control, a dsDNA Sp1 binding site was included comprised of the sequence 5'-GGCCTTCCTGGGGGCGGGGCCTTAGCTGCA-3'. The sense and antisense sequences were synthesized separately and annealed by incubation at 96°C for 3 minutes, then slow cooling to 60°C where it remained for another 15 minutes before cooling down to 23°C.

**RNA-binding assays**

The pTA-13, -45, and -90 plasmids were linearized with HindIII and transcribed from the T7 promoter using the Transprobe T kit from Pharmacia. Transcripts consisted of 150 to 175 bases of non-repeat flanking sequence on each side of the repeats. 40µg of cytosolic (S100) and nuclear protein extracts were incubated with probe (50 000 cpm per reaction) for 30 minutes at 0°C with buffer (20mM HEPES (pH7.9), 20% (v/v) glycerol, 0.1 KCl, 0.2mM EDTA) to 15-20µl. One unit of RNase T1 was added for 10 minutes, followed by heparin (5mg/ml final), both at 0°C. Reaction mixtures were loaded onto 4% nondenaturing polyacrylamide gels (acrylamide/bisacrylamide=19:1), electrophoresed at 180V, 4°C, then dried and autoradiographed. Where a no protein control was done, protein buffer D, as used by Dignam et al. (1983), was substituted. The binding reactions for the

competition assay were the same with the addition of cold transcript in ratios of (hot:cold) 1:1, 1:3, and 1:6.

**RNA half-life studies**

TE32 rhabdomyosarcoma cells were plated on 100mm culture dishes and grown in standard -minimal essential media (MEM) + 10% fetal calf serum (FCS). For actinomycin D-induced transcription arrest, TE32 cells grown to 60% confluence were treated with actinomycin D (1 mg/ml in absolute ethanol stock; Boeringher Mannheim) at a final concentration of 5 μg/ml. Total cellular RNA was isolated from 0 hours to 7 hours post-actinomycin D treatment. RNA was extracted according to the method outlined by Birnboim (1988). Exactly 10μg of RNA per sample in a volume of 10μl was denatured by mixing with 5μl formamide and 5μl formaldehyde/phosphate (100μl 37% formaldehyde, 6.6μl 1M sodium phosphate (pH 6.8) and 94μl formamide) and incubated at 55°C for 30 minutes. The samples were electrophoresed in a vertical 1.2% agarose gel 1.5mm thick containing 20mM sodium phosphate (pH6.8), 2mM EDTA and 1M formaldehyde. Gels were blotted with (+)-nylon membrane (Biotrans). Filters were probed with a 530bp PstI fragment from DMK corresponding to the 5' region of the gene. A probe specific to ß-actin was used to provide a comparison control for long mRNA half-life.

**CAT Assays**

Cells were split approximately 16 hours before transfection into 60mm tissue culture plates at a density of $3x10^5$ cells/plate (20-30% confluence) in α-(MEM)+10% (FCS). Two hours before transfection, new media was added. DNA was transiently transfected into TE32 cells by calcium phosphate precipitation with duplicate plates. Co-transfection was with a ß-galactosidase plasmid driven by the Rous sarcoma virus (RSV) promoter to standardize transfection efficiency. Transfections were done in duplicate for each pcDNA3-CAT-3'UTR vector. 5μg each of CAT-3'UTR DMK plasmid and control ß-galactosidase

plasmid was added to 250mM $CaCl_2$ and $H_2O$ up to 500ul. The calcium/DNA solution was added dropwise to 500µl of 2x HEPES buffered saline (HeBS), pH 7.05 (16.4 g NaCl, 11.9 g HEPES acid, 0.21 g $Na_2HPO_4$ for 1 liter, pH with NaOH) while air was bubbled through the 2x HeBS. The mixture was allowed to precipitate at room temperature for 45 minutes, then 500ul was added to each 60mm plate. After incubation for 4 to 6 hours, the plates were rinsed twice with 1x phosphate buffered saline (PBS) and fresh media added. Cells were allowed to grow for 72 hours, then washed three times with 1x PBS, harvested by scraping and pelleted in 1.5ml microfuge tubes. Cells could be stored at -80°C at this point for a few weeks until time to assay. Cell extracts were obtained by freeze/thaw lysis. Cell pellets were resuspended in 100µl 0.25M Tris-HCl pH 7.8 and frozen on dry ice for two minutes. The suspension was then thawed at 37°C for five minutes, vortexed, and placed on dry ice again. The freeze/thaw cycle was repeated a third time and the cellular debris pelleted.

For the CAT assay (modified from the Promega protocol), 25µl of cell extract supernatant was mixed with 90µl of 0.25M Tris-HCl pH 7.8 and incubated at 60°C for 10 minutes to eliminate any acetylase activity. 5µl of 5mg/ml n-buturyl CoA (Sigma), and 5µl $^{14}C$- chloramphenicol (Amersham) were then added and the reaction allowed to proceed at 37°C for one hour. Each reaction was extracted with 300µl mixed xylenes (BDH) and the upper organic phase transferred to a fresh tube. The xylene phase was back-extracted with 100µl 0.25M Tris-HCl pH 7.8 then 200µl of upper xylene phase was added to 5 ml scintillation fluid (Scintiverse II, Fischer) and counted.

# RESULTS

## DNA-protein complexes specific to the repeat do not form

Demonstration of proteins binding the repeat or flanking 3'UTR regions of DMK in DNA form could indicate regulatory control at the transcriptional level. To study this possibility, DNA fragments primers were generated by PCR using primers 409 and 410 which encompass the repeat sequence and 20 nucleotides of up- and downstream flanking sequence. Clones containing fragments with 13 repeats and 63 repeats were amplified. As a control for protein binding, a 30bp DNA fragment with one consensus Sp1-binding site was also used as described by Dynan et al., (1986). The binding assay was conducted using both annealed and boiled-denatured probe to study double-stranded and single-stranded binding. Nuclear protein extract isolated from diaphragm muscle of a congenital DM patient was used in the binding assay. Shifted complexes were observed with the Sp1 control oligo, and with the denatured fragment containing 13-repeats. No bandshifts were observed with the 63-repeat fragment either in annealed or denatured form (data not shown). Probes representing the flanking regions of the repeat, i.e. the primer sequences, were generated by restriction digestion of the 13-repeat fragment with the restriction enzyme Fnu4H1 which cuts within the repeat sequence to assay for the possibility of binding factors specific to the non-repeats regions. No shift was seen with either annealed or denatured flanking fragments when compared to the full-length, denatured 13-repeat fragment (data not shown). A bandshift was consistently observed with the denatured 13-repeat fragment. However, when a competition assay was performed using cold 13-repeat fragment, cold Sp1 oligo, and poly dI:dC DNA as competitors at ratios of 1:1, 1:2, and 1:4 only cold poly dI:dC DNA was observed to effectively compete with probe. It is possible that not enough specific cold repeat DNA was used to detect competition effectively. However, since the protein-DNA complex involving denatured repeat-containing probe could be competed with

non-specific, single-stranded DNA, the DNA-binding factor probably did not have high affinity for the repeat.

**RNA-specific factors are present in cytosolic protein extracts**

Lack of protein binding to DNA does not exclude the possibility of factors present that are capable of associating with RNA. Such factors would exert control at the post-transcriptional level, perhaps via regulation of RNA stability, cellular transport, or translation. To investigate, RNA probes comprising the SmaI/HindIII fragment of the 3'UTR were generated by *in vitro* transcription from linearized plasmids containing 13-repeats (pTA-13), 45-repeats (pTA-45), or 90-repeats (pTA-90). In addition to the repeat sequence, the labeled transcripts had approximately 200 bases of flanking sequence on each side. Cytosolic (S100) protein extracts isolated from congenital DM diaphragm muscle (i.e. skeletal muscle), and heart were incubated with RNA probe. Congenital muscle tissue was used in an attempt to relate the results more closely with the disease. However, since no normal human controls were used, the role of the reported binding factor(s) in either normal function or disease pathology is uncertain. A strong band shift was observed for all three probes with congenital diaphragm S100 extract in addition to two weaker bands running at faster mobility (fig. 1). The smallest complex was also seen with congenital heart S100 extract although the larger complexes present with congenital diaphragm were absent. Tissue-specific differences between skeletal and cardiac muscle could account for the discrepancies in bandshift pattern. Upon substituting congenital heart S100 extract with nuclear extract from congenital diaphragm, no shifts were observed with the nuclear extract. As a positive control, congenital diaphragm S100 extract was used to test for reproducible RNA-protein binding (fig. 2). Surprisingly, the largest complex seen with congenital diaphragm S100 extract was also present in mouse liver S100 extract (fig. 3) running at the same mobility. DMK is expressed at low levels in human liver in the full-length form and the strong bandshift detected in such high quantity in both congenital skeletal muscle, and

**Fig.1 RNA mobility shift assay with congenital diaphragm and congenital heart S100 extracts.** Strong band shifts are observed for all three probes with congenital diaphragm S100 extract in addition to two weaker bands running at faster mobility. The smallest complex is also seen with congenital heart S100 extract. Lanes a, e, and i are probe only controls. Lanes b, f, and j are probe/no protein controls. S100 extracts from congenital diaphragm (lanes c, g, and k) and congenital heart (lanes d, h, and l) were mixed with labeled transcripts containing 13 repeats (lanes a to d), 45 repeats (lanes e to h) and 90 repeats (lanes i to l).

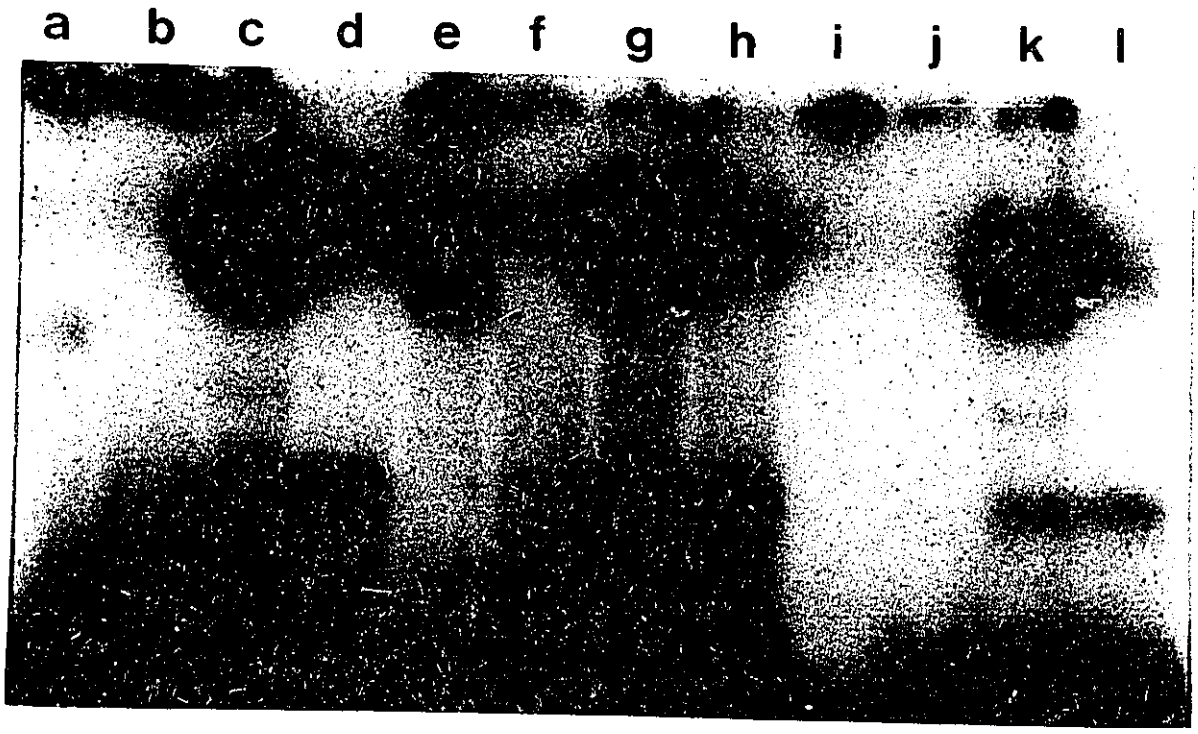

a b c d e f g h i j k l

**Fig. 2 RNA mobility shift assay with S100 and nuclear extract from congenital diaphragm.** The RNA-binding factors seen with congenital diaphragm S100 extract are absent with congenital diaphragm nuclear extract. Binding factors therefore appear to be cytosol-specific. Lanes a, e, and i are probe only controls. Lanes b, f, and j are probe/no protein controls. S100 extract from congenital diaphragm (lanes c, g, and k) and nuclear extract from congenital diaphragm (lanes d, h, and l) were mixed with labeled transcripts containing 13 repeats (lanes a to d), 45 repeats (lanes e to h) and 90 repeats (lanes i to l).

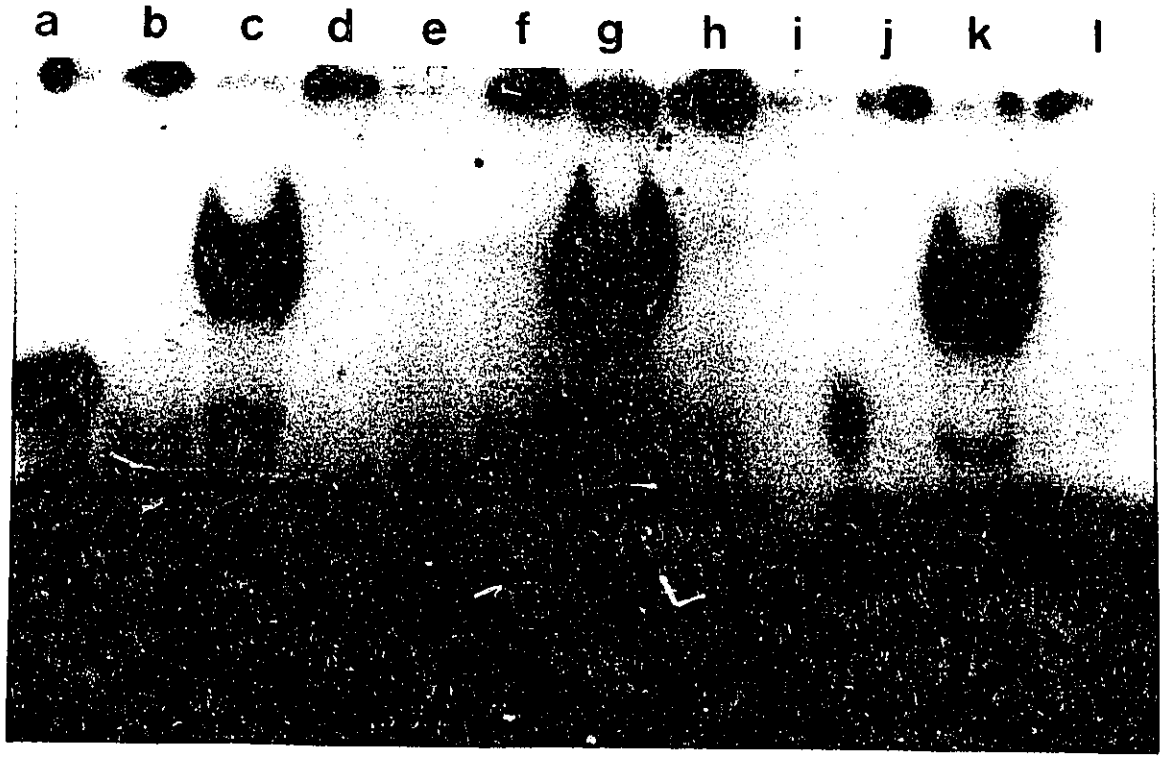
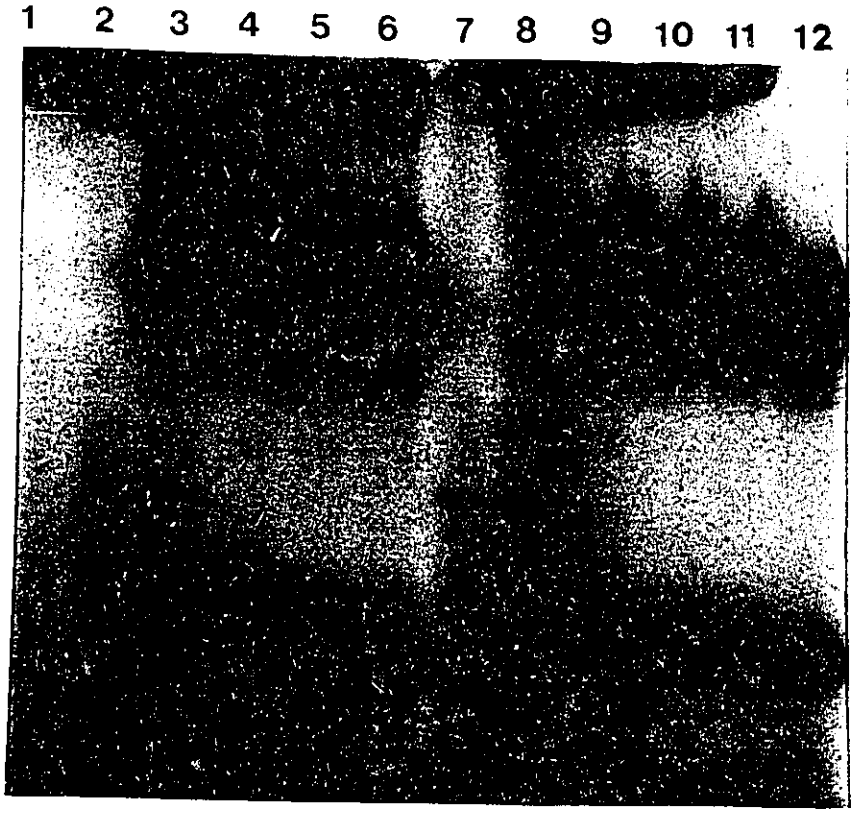

a
b
c
d
e
f
g
h
i
j
k
l

**Fig. 3 RNA mobility shift assay using congenital diaphragm and mouse liver S100 extracts.** The largest RNA-binding factor seen with congenital diaphragm S100 also appear to be present in mouse liver S100 extract. Lanes 1, and 5 are probe only controls and lanes 2, and 6 are probe /no protein controls. S100 extracts from congenital diaphragm (lanes 3, and 7) and mouse liver (lanes 4, and 8) were mixed with labeled transcripts containing 45 repeats (lanes 1 to 4) and 90 repeats (lanes 5 to 8).


1
2
3
4
5
6
7
8

**Fig. 4 Competition assay to determine RNA-binding specificity.** A competition assay using S100 extract from mouse liver mixed with labeled transcripts containing 13 repeat or 35 repeats was mixed with an increasing amount of cold transcript added in ratios of (hot:cold) 1:0. 1:1, 1:3, and 1:6. No reduction in intensity is observed with increasing amounts of cold RNA, though a noticeable decrease in complex size seems to correlate with added with cold transcript. Probes with 13 repeats (lanes 1 to 6) and 45 repeats (lanes 7 to 12) were mixed with mouse liver S100 extract (lanes 3 to 6, and 9 to 12). Lanes 1 and 7 are probe only controls. Lanes 2 and 8 are probe/no protein controls. Lanes 3 and 9, 4 and 10, 5 and 11, and 6 and 12 have ratios of 1:0, 1:1, 1:3, and 1:6, hot:cold transcript, respectively.


1 2 3 4 5 6 7 8 9 10 11 12

liver could represent a ubiquitous RNA-binding protein with specificity for the repeat and/or flanking sequences.

A competition assay using S100 extract from mouse liver mixed with labeled transcripts containing 13 repeats or 45 repeats was done. Increasing amounts of the same cold transcript were added in ratios of (hot:cold) 1:1, 1:3, and 1:6. The two shifted complexes fail to be reduced in intensity with increasing amounts of cold RNA (fig.4). However, a noticeable decrease in complex size seemed to correlate with increasing cold transcript. The change in band mobility may be an actual phenomena due to competition, or may be caused by an increase in reaction volume as increased amounts of competitor were added. Binding of the murine-derived factors may be so strong that cold competitor used at six times excess may not have been sufficient to detectably reduce the signal. In experiments performed by others, concentrations of competitor as high as 1000 times in excess have been required to fully deplete any signal. Whether the human factors interact as strongly has to be tested.

**DMK mRNA has a half-life of several hours**

Establishment of the half-life value of DMK mRNA would be the basis of further studies involving mRNA stability. Once the half-life of wild-type DMK is determined, the effect on half-life of increases in repeat size can be compared. If the assumed turnover mechanism is not involved then no change in half-life should be observed and the possibility of increased transcription can be considered. The rhabdomyosarcoma cell line TE32 expresses endogenously the DMK gene with a wildtype-size repeat at levels detectable by Northern blot analysis. To study the half-life of the endogenous mRNA as a precursor to experiments using clones with larger repeat expansions, TE32 cells were incubated with actinomycin D to halt transcriptional processes. A northern blot containing total cellular RNA sampled at 0 hours, 0.5 hours, 1 hour, 2 hours, 4 hours, and 7 hours post-actinomycin D treatment was probed with a 530pb PstI fragment from the 5' region of DMK. Also used

as probe was a control plasmid containing ß-actin which has a half-life of approximately 17 hours and is expressed at very high levels. DMK mRNA was present at high levels, though at a lower concentration than ß-actin. A signal was detectable as long as long as seven hours post-treatment (fig. 5). The signal may be starting to decrease compared to the control at the seven hour mark, but only slightly. Endogenous mRNA transcribed by TE32 cells would therefore appear to be about as stable as ß-actin mRNA up to seven hours indicating a stable transcript.

**CAT assays reveal a regulatory function of the 3'UTR**

To establish whether the 3'UTR of DMK possesses regulatory activity, clones driven by the CMV promoter and containing the CAT open reading frame and complete DMK-3'UTR with repeat sizes of 5 repeats, 11 repeats, 45 repeats or 90 repeats were constructed (figs. 6a and 6b). The clones were transiently transfected in duplicate into TE32 human rhabdomyosarcoma-derived cells by calcium phosphate precipitation. TE32 cells express DMK endogenously and at significant levels and probably have the necessary factors required for proper gene regulation. A plasmid containing the reporter gene ß-galactosidase was co-transfected to standardize transfection efficiency. After 72 hours, cells were harvested and extracts obtained by freeze/thaw lysis. CAT activity of the extracts was assayed by buturylation of $^{14}$C-chloramphenicol followed by xylene extraction and liquid scintillation counting (LSC).

In a preliminary experiment to measure activity of the constructs, the construct with the 11 repeat DMK-3'UTR demonstrated an approximately two-fold higher activity over a control construct that had only a polyadenylation signal as its 3'UTR (fig. 7). As well, another CAT vector using the Rous sarcoma virus promoter (pRSV-CAT) expressed CAT at a much lower level showing the CMV-driven vectors to be a functional and more sensitive reporter system.

To identify specific regulatory regions within the 3'UTR, deletion clones lacking the regions upstream and downstream of the repeat were generated. The four clones constructed were derived from pcDNA3-CAT-11rpt 3'UTR and contained sequences upstream of the repeat (Del1); upstream of and including the repeat (Del2); downstream of and including the repeat (Del3); and downstream of the repeat (Del4) as illustrated in fig. 1. The control BGH polyadenylation signal was included in all clones. All the deletion clones displayed only background CAT activity even when the repeat sequence was present. In comparison, the pcDNA3-CAT-11rpt 3'UTR and pcDNA3-CAT clones expressed CAT at the same two-fold higher ratio level seen in previous trials (fig.8). An incomplete DMK-3'UTR therefore gives rise to a non-functional reporter construct indicating the sequences deleted to be essential for gene expression. This experiment also demonstrates inclusion of the repeat sequence in the deleted 3'UTRs is not enough to permit reporter gene expression. Inclusion of the flanking sequences would therefore appear necessary to generate activity by acting in conjunction with the repeat sequence.

Since the mutation observed in DM is expansion of the repeat sequence, another assay was performed to study the effect of increasing repeat number on CAT expression. Using clones that were identical except for repeat size, up to 90 repeats were incorporated. Unfortunately, attempts to generate constructs with more than 90 repeats resulted in highly unstable clones so only the effects of a minimally symptomatic mutation could be observed. An approximately three-fold increase in CAT activity was observed from constructs containing a wildtype 3'UTR (5 and 11 repeats) compared to the polyA-only control construct. Reporter constructs with repeats in the protomutation (45 repeats) and mutation (90 repeats) range showed a greater than 10-fold increase over control CAT activity. Increasing repeat size would therefore appear to have a positive effect on gene expression. However, since results with the deletion clones show the 3'UTR must be present in its entirety to function as an activation element, the enhanced expression imposed by an

expanded number of repeats is probably detectable only when in the context of the other 3'UTR sequences.

**Fig. 5 Analysis of mRNA half-life of DMK.** TE32 rhabdomyosarcoma cells were transcriptionally inhibited to measure the half-life of the DMK transcript. TE32 cells were treated with actinomycin D (5μg/ml final concentration) and mRNA levels were assayed by Northern blot analysis. As a control for the amount of total RNA loaded per lane, ß-actin was also probed. After a period of seven hours, DMK mRNA levels may just be beginning to diminish in comparison with ß-actin levels which remain constant. Time points sampled post-treatment: a=0 hrs, b=0.5 hrs, c=1 hr, d=2hrs, e=4hrs, and f=7hrs.

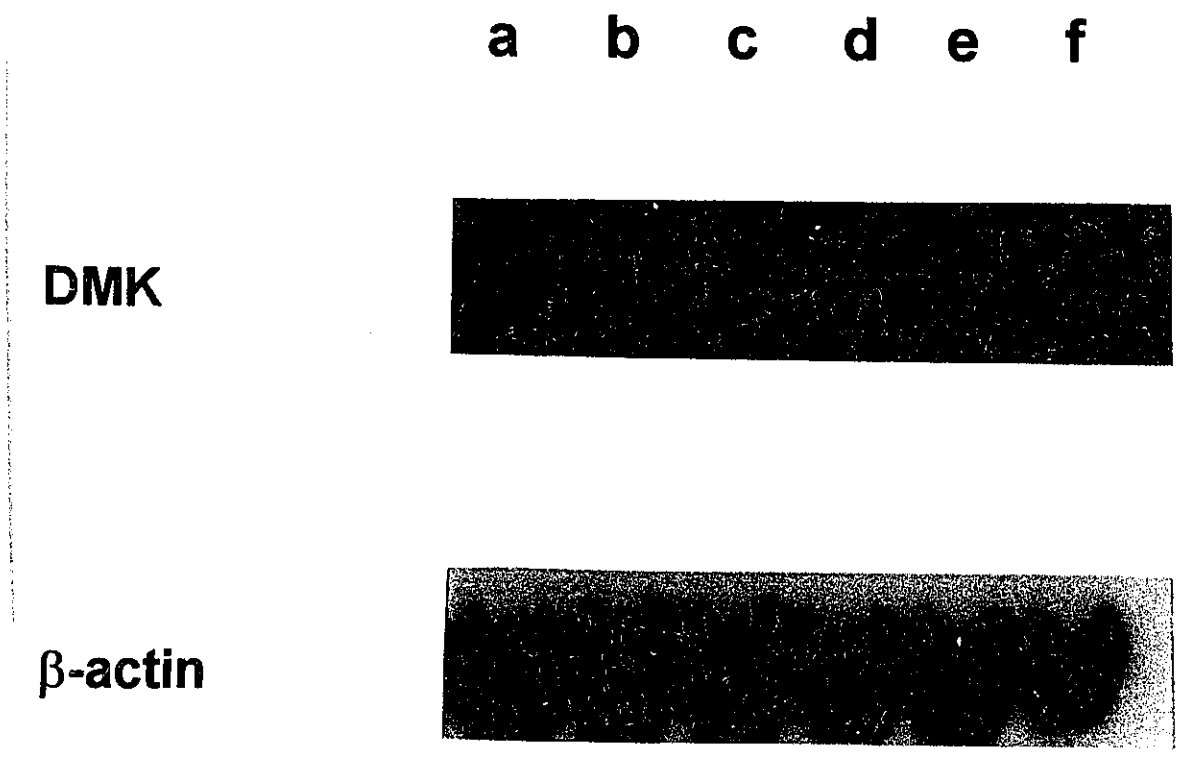
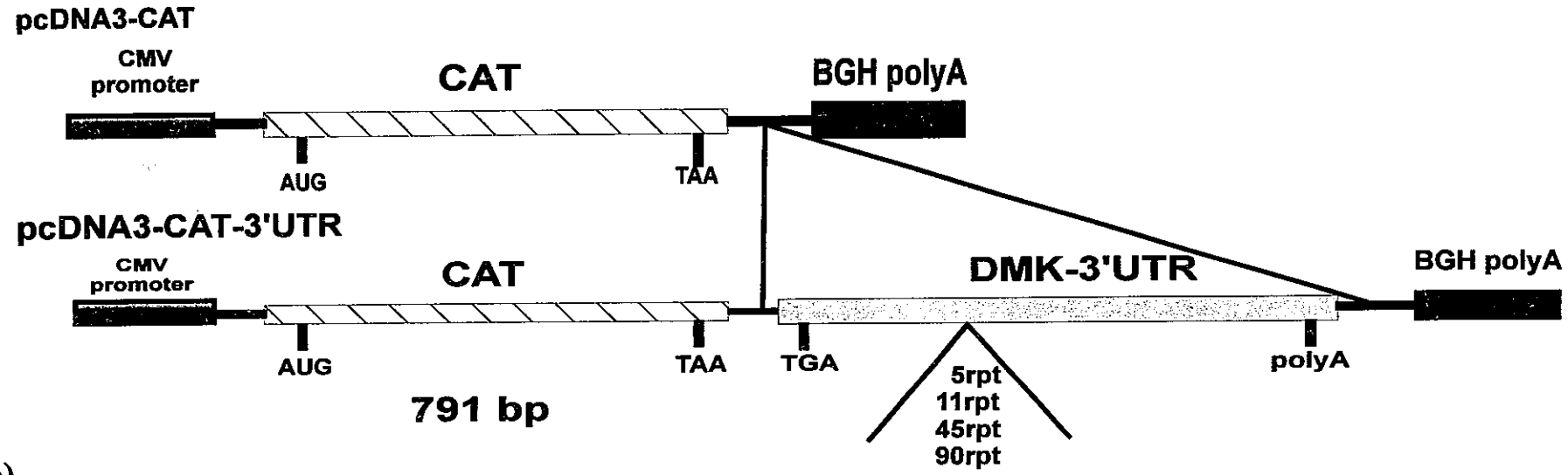

a b c d e f
DMK
β-actin

**Fig. 6 Construction of plasmids with the DMK 3'UTR.** a) The CAT ORF (hatched region) contained in a Sau3AI fragment was ligated into the BamHI site of the pcDNA3 vector to generate pcDNA3-CAT. PCR was used to amplify the DMK 3'UTR (gray bar) from a cosmid with 11 repeats. The PCR product was cloned into the pCRII vector, then subcloned into pcDNA3-CAT via the BstXI site. The 11 repeats were replaced with 5 repeats, 45 repeats, or 90 repeats to generate the pcDNA3-CAT-3'UTR series of clones. b) Deletion clones were generated by PCR specific for the regions up- or downstream of the repeat again using the cosmid with 11 repeats.

a)


pcDNA3-CAT
CMV promoter
CAT
BGH polyA
AUG
TAA
pcDNA3-CAT-3'UTR
CMV promoter
CAT
DMK-3'UTR
BGH polyA
AUG
TAA
TGA
polyA
791 bp
5rpt
11rpt
45rpt
90rpt

b)

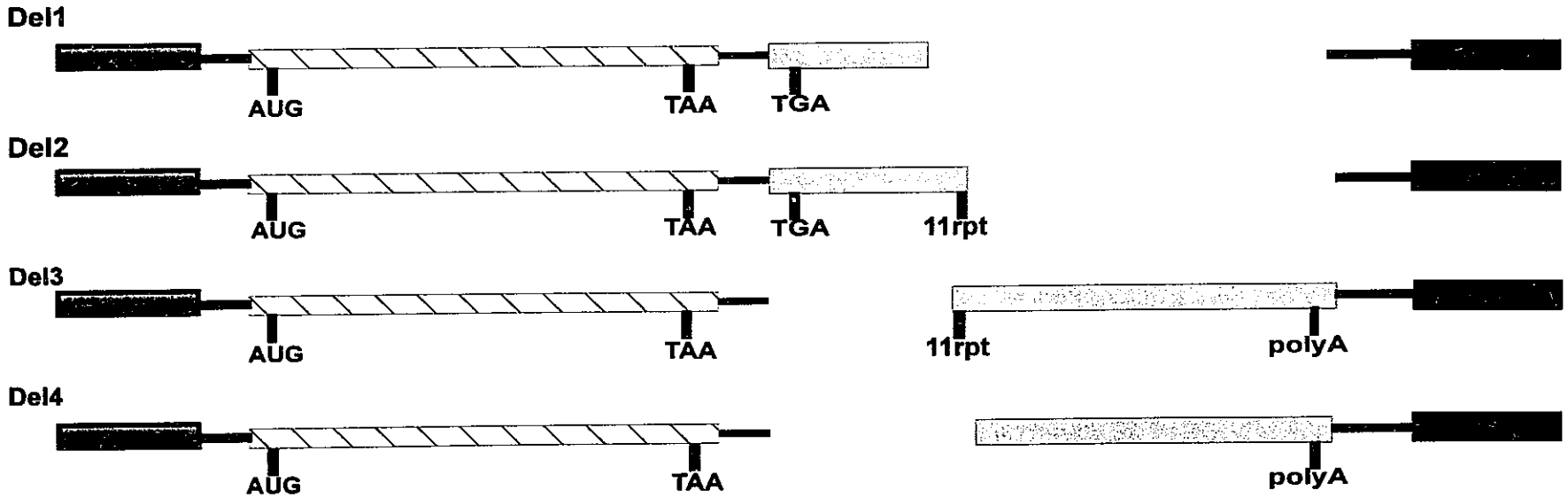

Del1
AUG
TAA
TGA
Del2
AUG
TAA
TGA
11rpt
Del3
AUG
TAA
11rpt
polyA
Del4
AUG
TAA
polyA

**Fig. 7 Comparison of CAT activity of constructs with and without the DMK 3'UTR.** Addition of a DMK-3'UTR with 11 repeats to a CAT reporter plasmid resulted in an increase in CAT activity compared to a construct containing only a polyadenylation signal. Activity was approximately two-fold higher for pcDNA3-CAT-11rpt 3'UTR over pcDNA3-CAT. Comparison with another CAT vector using the Rous sarcoma virus promoter (pRSV-CAT) revealed the CMV-containing constructs to possess at least twice the CAT activity. Left to right: no reporter gene plasmid; pRSV-CAT reporter plasmid; pcDNA3-CAT reporter plasmid; pcDNA3-CAT-11rpt 3'UTR.

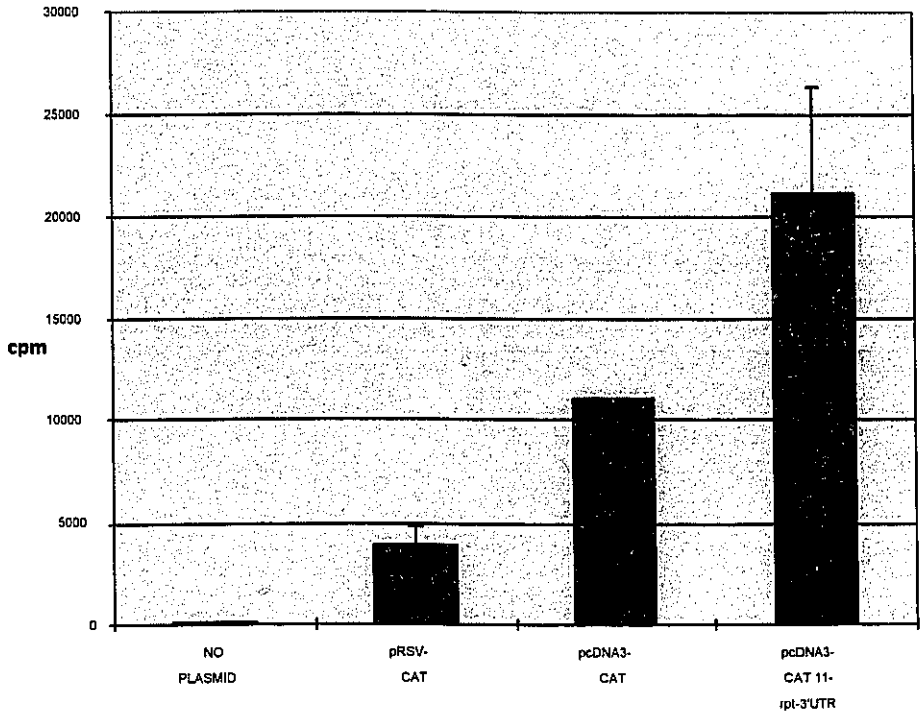

30000
25000
20000
15000
cpm
10000
5000
0
NO PLASMID
pRSV-CAT
pcDNA3-CAT
pcDNA3-CAT 11-rpl-3'UTR

**Fig. 8 Comparison of CAT activity of deletion mutant constructs.** Deletion analysis of the regions upstream and downstream of the repeat region in the DMK 3'UTR was undertaken to identify potential regulatory regions. Results show all deletion clones to have little or no CAT activity despite the presence of the control poly(A) signal. Left to right: pcDNA3-CAT; pcDNA3-CAT-11rpt 3'UTR; pcDNA3-CAT-Del1; pcDNA3-CAT-Del2; pcDNA3-CAT-Del3; pcDNA3-CAT-Del4.

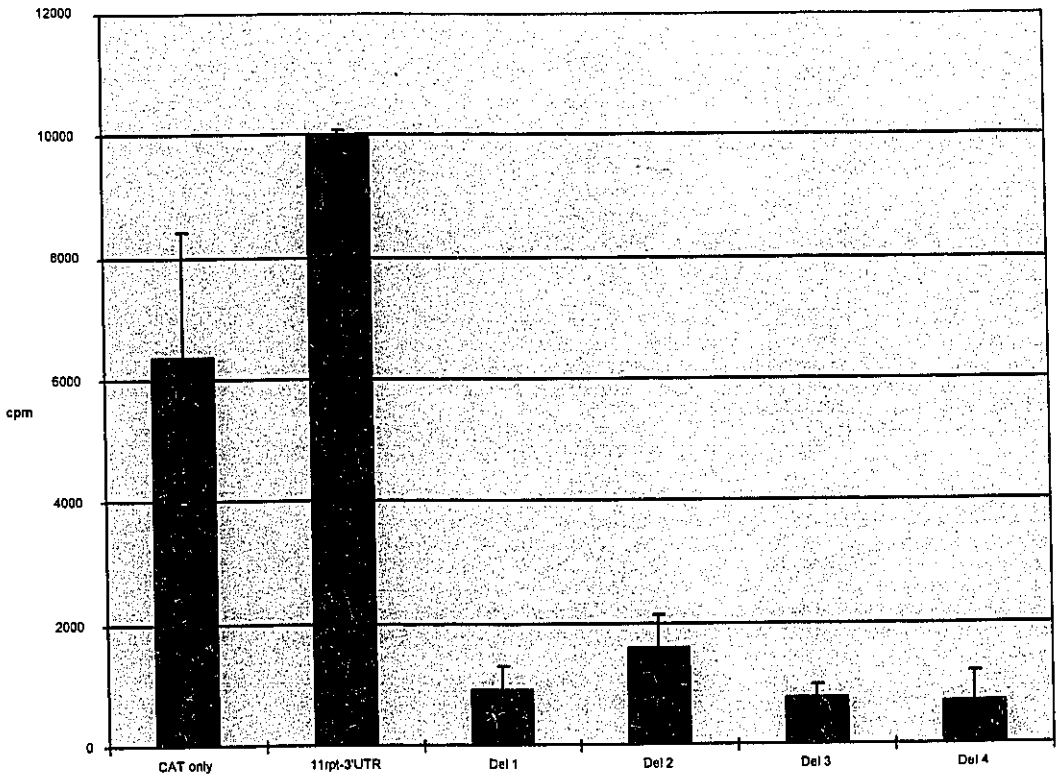
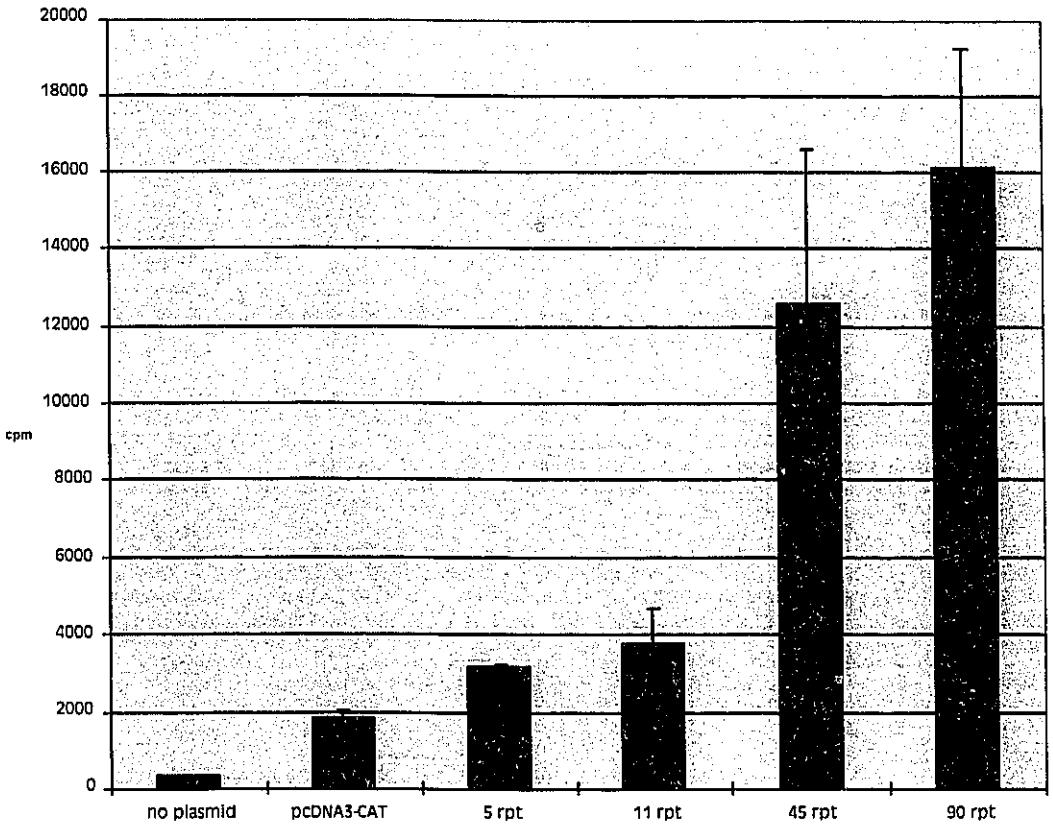

12000
10000
8000
6000
4000
2000
0
cpm
CAT only
11rpt-3'UTR
Del 1
Del 2
Del 3
Del 4

**Fig. 9 Comparison of CAT activity of DMK 3'UTR constructs containing an increasing number of repeats.** Chimeric DMK 3'UTR-CAT constructs were made containing $[CTG]_n$ sequences ranging from 5 repeats up to 90 repeats. An approximately three-fold increase in CAT activity from constructs containing a wildtype 3'UTR (5 and 11 repeats) was observed compared to a control construct containing only a poly(A) signal. Constructs with repeats in the protomutation (45 repeats) and mutation (90 repeats) range showed a greater than 10-fold increase over control CAT activity. Left to right: No reporter gene plasmid; pcDNA3-CAT; pcDNA3-CAT-5rpt 3'UTR; pcDNA3-CAT-11rpt 3'UTR; pcDNA3-CAT-45rpt 3'UTR; pcDNA3-CAT-90rpt 3'UTR.


20000
18000
16000
14000
12000
10000
cpm
8000
6000
4000
2000
0
no plasmid
pcDNA3-CAT
5 rpt
11 rpt
45 rpt
90 rpt

## DISCUSSION

The nature of the DM mutation being expansion of a CTG repeat in the 3'UTR of the gene has provided a unique opportunity to study a potentially novel mechanism of gene regulation. While DM does share many characteristics with the other trinucleotide repeat diseases, namely, genetic anticipation correlating with increased expansion, and autosomal dominant inheritance (with the exception of the X-linked disorders), the mechanism of mutation is not as immediately obvious. The mutation is most likely not protein-related since the expansion occurs outside the coding region. Instead, a change in the expression of DMK effected by the expansion is a logical model. Experiments were undertaken to establish the involvement of the DMK 3'UTR in gene expression.

Protein binding experiments using both DNA and RNA encompassing the repeat and the surrounding regions identified three possible RNA binding factors present in the cytoplasm. A transcript containing the CTG repeat in the 3'UTR of the DMK gene appears able to form RNA-protein complexes together with congenital skeletal (diaphragm) cytosolic protein extract. One of the binding factors also seems to be present in congenital heart cytosolic extracts. No RNA-protein complexes are seen to form using nuclear extract from congenital skeletal muscle suggesting proteins specific for the transcripts are present only in the cytoplasm. The largest RNA-binding factor is also found in mouse liver S100 extract possibly indicating binding of a common factor to the 3'UTR sequences. Competition experiments to establish specificity of binding were preliminary and require further study.

Half-life studies of DMK mRNA carried out using actinomycin D to inhibit transcription in TE32 rhabdomyosarcoma cells show that DMK mRNA has a half-life of several hours. After seven hours post-actinomycin D treatment, the DMK signal, as detected by northern blot analysis, is beginning to diminish. Such a long mRNA half-life may be indicative of a high, or at least persistent level of DMK protein in cells. Whether

repeat expansion alters the stability of the DMK transcript will have to be investigated in future studies.

To directly observe the ability of the 3'UTR to regulate gene expression, the sequence was cloned downstream of the CAT reporter gene driven by a CMV promoter. Substitution of the DMK 3'UTR for an ordinary polyadenylation signal results in an increase in CAT activity suggesting an element(s) may be present in the sequence that is able to boost basal level gene expression. Large deletions of the DMK 3'UTR inhibit expression causing a reduction in CAT activity to background levels. A complete 3'UTR would therefore appear necessary to function in an expression-enhancing manner. To attempt to more closely represent the disease situation, reporter constructs were made to contain repeat sequences from wildtype to mutation range, the upper limit being 90 repeats. An increase in repeat size was found to correlate with increasing CAT activity. Expansion of the repeat may therefore influence gene expression in a positive manner resulting in increasing DMK levels as repeat number escalates. However, work by Sabourin et al., (1995), seems to suggest the protein levels of DMK do not differ greatly in muscle tissue derived from disease individuals compared to non-DM controls. The 3'UTR-specific $[CTG]_n$ expansion may therefore not exert a direct influence on the DMK gene, but may work, *in trans* perhaps by titrating protein factors that are able to bind the 3'UTR mRNA. Loss of these factors in the cellular milieu could result in impairment of cellular metabolic functions giving rise to the disease phenotype.

# CHAPTER III - CONCLUSIONS

To investigate a potential role of the DMK 3'UTR (including the $[CTG]_n$ expansion mutation) in gene regulation, characterization studies were carried out. Specifically, RNA-protein gel mobility shift assays, mRNA half-life studies and a CAT reporter gene system were developed to try and analyze its behavior.

CAT-DMK 3'UTR chimeric reporter constructs containing an increasing number of repeats correlated with an increase in CAT activity. Whether the increase in CAT activity occurs at the transcriptional or post-transcriptional levels is not known. mRNA half-life studies using the CAT constructs could be done comparing stability of clones with a wildtype-size repeat to mutant range expansions to investigate whether disruption of normal message turnover occurs. The factors observed binding RNA may also be involved in mRNA stability. To determine if RNA-specific proteins are involved in regulation of half-life, cyclohexamide can be used to pretreat cells for arrest of protein synthesis. A change in half-life value would suggest RNA-protein interactions are necessary for controlling stability of the DMK message.

Taken together, the data seems to point to a regulatory role for the DMK 3'UTR in gene expression. Since RNA-binding factors are present in cytoplasmic extracts, the repeat and the 176bp and 150bp of up- and downstream flanking sequence, respectively, may function by sequestering proteins thus maintaining these factors at a specific cellular concentration. The presence of a high affinity factor in mouse liver S100 extract, which expresses DMK at low levels only, suggests the 3'UTR may bind ubiquitous factors. No change in complex size was observed with transcripts that contained larger repeat expansions suggesting proteins may not be binding the repeat itself but perhaps a secondary structure of which the repeat sequence is a part. Repeat expansion may alter sequence binding affinity causing an imbalance in protein concentration in the cytoplasm. Thus, the disease pathology may involve other genes or factors that interact with the DMK 3'UTR in

addition, or perhaps exclusion to DMK itself. Experiments using the CAT reporter constructs showed an upward trend in CAT activity following an increasing number of repeats from wildtype to mutation range. Such behavior is characteristic of a gain of function mutation which would explain the dominant inheritance associated with DM. If these results can be extrapolated to the disease situation, then increased levels of DMK protein can be expected as the repeat continues to expand. Keeping in mind DMK is a serine/threonine kinase and that the mutation may correspond to hundreds or thousands of repeats in affected individuals, the action of this protein may become highly perturbed in the disease state. However, until more is known about the biological function and interactions of DMK, the relationship between aberrant kinase activity and disease pathology remains unclear.

Comparing human and mouse sequence, a corrupted version of five imperfect repeats based on the core CXG sequence is observed in the mouse DMK gene (Jansen, et al., 1992). More significantly, regions of high homology both up- and downstream of the repeat are present in the human and mouse 3'UTR. In CAT assays, constructs containing large deletions of the 3'UTR that include the regions of high homology show complete abrogation of CAT gene activity. Further deletion mapping of the 3'UTR and mutational analysis should be done in future work to specifically define any functional element(s). Targeting the conserved regions for analysis should provide evidence as to whether they serve a regulatory purpose. Perhaps the two sequences share a similar secondary structure which becomes distorted when repeat size extends beyond a certain point. If the repeats only become significant upon reaching obtrusive proportions, the function of this particular sequence under normal conditions comes into question. It would be interesting to perform the same CAT reporter gene experiments substituting the DMK 3'UTR with sequence free of repeats to determine if a minimal trinucleotide sequence is necessary for proper function.

Since these results are based solely on the effect of the 3'UTR with no other gene sequences present, they could be very much out of context with respect to *in vivo* gene

regulation. Other regulatory elements may be present in the 5'UTR or coding region. Intronic sequences may also prove to be important as they have for the promoter (Storbeck et al., 1994). All data collected for the mRNA half-life and CAT-DMK 3'UTR transfection studies have used TE32 cells as the preferred cell line due to the fact DMK is expressed at significant levels. However, it must be remembered that all results are obtained against a tumor cell background and may not perfectly reflect the *in vivo* situation. To explore this avenue more fully, other cell lines will also have to be tested to determine if the data reported is cell-type specific or occurs only with TE32. Ultimately, the same CAT constructs, or similar using the ß-galactosidase reporter gene could be integrated into the genomes of transgenic mice to study their behavior *in vivo*.

The discovery of a trinucleotide expansion in the 3'UTR of the DM gene occurring in over 99% of all affected individuals has led to speculation about the possible role of this region in gene regulation. I have succeeded in demonstrating an expression control function of the 3'UTR of DMK with possible RNA-protein involvement. I have also shown that while an expanding repeat sequence is capable of imposing increased activity from a reporter gene construct, it must be together with the complete 3'UTR. This result has important implications in terms of the pathology of the disease, namely, that the molecular behavior can translate phenotypically into a gain of function mutation. Experiments to investigate *in vivo* behavior of the 3'UTR sequence will have to be conducted in the future to further test this hypothesis.

## References

Aharon, T., and R.J. Schneider. 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. Mol. Cell. Bio. **13**: 1971-1980.

Ang, C.W.Y., L.A. Sabourin, J.M Barceló, M.A. Narang, and R.G. Korneluk. 1993. Gel-shift analysis of RNA-specific proteins able to bind the CTG-repeat region of the myotonic dystrophy kinase gene. Am. J. of Hum. Genet. **53**: Supple. A663.

Ang, C.W.Y., L.A. Sabourin, M.A. Narang, and R.G. Korneluk. 1994. The myotonic dystrophy kinase 3'-untranslated region and its effect on gene expression. Am. J. of Hum. Genet. **55**: Supple. A730.

Banfi, S., A. Servadio, M.-y. Chung, T.J. Kwiatkowski Jr., A.E. McCall, L.A. Duvick, Y. Shen, E.J. Roth, H.T. Orr, and H.Y. Zoghbi. 1994. Identification and characterization of the gene causing type 1 spinocerebellar ataxia. Nat. Genet. **7**: 513-

Barceló, J.M., M.S. Mahadevan, C. Tsilfidis, A.E. MacKenzie, and R.G. Korneluk. 1993. Intergenerational stability of the myotonic dystrophy protomutation. Hum. Molec. Genet. **2**: 705-709.

Beutler, B., and T. Brown. 1991. A CAT reporter construct allows ultrasensitive estimation of TNF synthesis and suggests that the TNF gene has been silenced in non-macrophage cell lines. J. Clin. Invest. **87**: 1336-1344.

Bickel, M., Y. Iwai, D.H. Pluznik, and R.B. Cohen. 1992. Binding of sequence-specific proteins to the adenosine- plus uridine-rich sequences of the murine granulocyte/macrophage colony-stimulating factor mRNA. Proc. Natl. Acad. Sci. USA **89**: 10001-10005.

Birnboim, H.C. 1988. Rapid extraction of high molecular weight RNA from cultured cells and granulocytes for Northern analysis. Nuc. Acid. Res. **16**: 1487-1497.

Bohjanen, P.R., B. Petryniak, C.H. June, C.B. Thompson, and T. Lindsten. 1991. An inducible cytoplasmic factor (AU-B) binds selectively to AUUUA multimers in the 3' untranslated region of lymphokine mRNA. Mol. Cell. Bio. **11**: 3288-3295.

Braun, R.E., J.J. Peschon, R.R. Behringer, R.L. Brinster, and R.D. Palmiter. 1989. Protamine 3'-untranslated sequences regulate temporal translational control and subcellular localization of growth hormone in spermatids of transgenic mice. Genes & Dev. **3**: 793-802.

Brewer, G., and J. Ross. 1988. Poly(A) shortening and degradation of the 3' A+U -rich sequences of human c-*myc* mRNA in a cell-free system. Mol. Cell. Bio. **8**: 1697-1708.

Brewer, G. 1991. An A+U-rich element RNA-binding factor regulates c-*myc* mRNA stability in vitro. Mol. Cell. Bio. **11**: 2460-2466.

Brook, J.D., S.L. Knight, S.H. Roberts, H.G. Harley, K.V. Walsh, S.A. Rundle, K. Freyne, M.C. Koch, N.D. Epstein, B. Wieringa, H. Smeets, K. Haddinghan, M.J. Siciliano, D.H. Palmer, J.S. Miles, C.R. Wolf, C. Fonatsch, and D.J. Shaw. 1991a. The physical map of chromosome arm 19q: some new assignments, confirmations and re-assessments. Hum. Genet. **87**: 65-72.

Brook, J.D., H.G. Harley, K.V. Walsh, S.A. Rundle, M.J. Siciliano, P.S. Harper, and D.J. Shaw. 1991b. Identification of new DNA markers close to the myotonic dystrophy locus. J. Med. Genet. **28**: 84-88.

Brook, J.D., M.E. McCurrach, H.G. Harley, A.J. Buckler, D. Church, H. Aburatani, K Hunter, V.P. Stanton, J.-P. Thirion, T. Hudson, R. Sohn, B. Zemelman, R.G. Snell, S.A. Rundle, S. Crow, J. Davies, P. Shelbourne, J. Buxton, C. Jones, V. Juvonen, K. Johnson, P.S. Harper, D.J. Shaw, and D.E. Housman. 1992. Mólecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3'end of a transcript encoding a protein kinase family member. Cell **68**: 799-808.

Brunner, H.G., H. Smeets, H.M.M. Lamberman, M. Coerwinkel-Driessen, B.A. Van Oost, B. Wieringa, and H.-H. Roper. 1989. A multipoint linkage map around the locus for myotonic dystrophy on chromosome 19. Genomics **5**: 589-595.

Brunner, H.G., G. Jansen, W. Nillesen, M.R. Nelen, C.E.M. de Die, C.J. Höweler, B.A. van Oost, B. Wieringa, H.-H. Ropers, and H.J.M. Smeets. 1993. Reverse mutation in myotonic dystrophy (brief report). New Eng. J. Med. **238**: 476-480.

Burke, J.R., M.S. Wingfield, K.E. Lewis, A.D. Roses, J.E. Lee, C. Hulette, M.A. Pericak-Vance, and J.M. Vance. 1994. The Haw River syndrome: dentatorubropallidoluysian atrophy (DRPLA) in an African-American family. Nat. Genet. **7**: 521-

Buxton, J., P. Shelbourne, J. Davies, C. Jones, T. Van Tongeren, C. Aslanidis, P. de Jong, G. Jansen, M. Anvret, R. Williamson, and K. Johnson. 1992. Detecton of an unstable fragment of DNA specific to individuals with myotonic dystrophy. Nature **355**: 547-548.

Carey, N., K. Johnson, P. Nokelainen, L. Peltonen, M.-L. Svontaus, V. Juvonen, M. Anvret, U. Grandell, K. Chotai, E. Robertson, H. Middleton-Price, and S. Malcolm. 1994. Meiotic drive at the myotonic dystrophy locus? Nat. Genet. **6**: 117-118.

Davies, J.D., J. Jackson, R. Williamson, P.S. Harper, S. Ball, M. Sarfarazi, L. Meredith, and G. Fey. 1983. Linkage analysis of myotonic dystrophy and sequences on chromosome 19 using a cloned complement 3 gene probe. J. Med. Genet. **21**: 259-263.

Dynan, W.S., S. Sazer, R. Tijan, and R.T. Schimke. 1986. Transcription factor Sp1 recognizes a DNA sequence in the mouse dihydrofolate reductase promoter. Nature **319**: 246-248.

Fu, Y.-H., D.P.A. Kuhl, A. Pizzuti, M. Piertti, J.S. Sutcliffe, S. Tichards, A.J.M.H. Verderk, J.J.A. Holden, R.G. Fenwick, Jr., S.T. Warren, B.A. Oostra, D.L. Nelson, and C.T. Caskey. 1991. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. Cell **67**: 1047-1058.

Fu, Y.-H., A. Pizzuti, R.G. Fenwick, Jr., J. King, S. Rajnarayan, P.W. Dunne, J. Dubel, G.A. Nasser, T. Ashizawa, P. de Jong, B. Wieringa, R. Korneluk, M.B. Perryman, H.F. Epstein, and C.T. Caskey. 1992. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science **255**: 1256-

Fu, Y.-H., D.L. Friedman, S. Richards, J.A. Pearlman, R.A. Gibbs, A. Pizzuti, T. Ashizawa, M.B. Perryman, G. Scarlato, R.G. Fenwick, Jr., and C.T. Caskey. 1993. Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. Science **260**: 235-238.

Funanage, V.L., P. Carango, P.M. Moses. and H.G. Marks. 1994. Reduced expression of the normal *DMPK* allele in congenital DM patient. Am. J. Hum. Genet. **55**: Suppl. A1282.

Gorman, C.M., L.F. Moffat, and B.H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Bio. **2**: 1044-1051.

Han, J., G. Juez, and B. Beutler. 1991. Interactive effects of the tumor necrosis factor promoter and 3'-untranslated region. J. of Immu. **146**: 1843-1848.

Harley, H.G., J.D. Brook, S.A. Rundle, S. Crow, W. Reardon, A.J. Buckler, P.S. Harper, D.E. Housman, and D.J. Shaw. 1991. Localisation of the myotonic dystrophy locus to 19q 13.2-19q 13.3 and its relationship to twelve polymorphic loci on 19q. Hum Genet. **87**: 73-80.

Harley, H.G., J.D. Brook, S.A. Rundle, S. Crow, W. Reardon, A.J. Buckler, P.S. Harper, D.E. Housman, and D.J. Shaw. 1992a. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature **355**: 545-546.

Harley, H.G., S.A. Runde, W. Reardon, J. Myring, S. Crow, J.D. Brook, P.S. Harper, and D.J. Shaw. 1992b. Unstable DNA sequence in myotonic dystrophy. Lancet **339**: 1125-1128.

Harper, P.S. 1989a. Myotonic dystrophy - the clinical picture. p. 13-36. In J. Walton (ed.) *Myotonic Dystrophy*, 2nd ed. London: W.B. Saunders Co.

Harper, P.S. 1989b. Myotonic dystrophy in infancy and childhood. p. 187-225. In J. Walton (ed.) *Myotonic Dystrophy*, 2nd ed. London: W.B. Saunders Co.

Harper, P.S. 1989c. Muscle pathology in myotonic dystrophy. p. 227-230. In J. Walton (ed.) *Myotonic Dystrophy*, 2nd ed. London: W.B. Saunders Co.

Höweler, C.J., H.F.M. Busch, J.P.M. Geraedts, M.F. Niermeijer, and A. Staal. 1989. Anticipation in myotonic dystrophy: fact or fiction. Brain **112**: 779-797.

Hunter, A., C. Tsilfidis, G. Mettler, P. Jacob, M. Mahadevan, L. Surh, and R.G. Korneluk. 1992. The correlation of age of onset with CTG trinucleotide repeat amplification in myotonic dystrophy. J. Med. Genet. **29**: 774-779.

The Huntington's disease collaborative Research Group. 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell **72**: 971-983.

Jansen, G., M. Mahadevan, G. Amemiya, N. Wornskamp, B. Segers, W. Hendriks, K. O'Hoy, S. Baird, L. Sabourin, G. Lennon, P.L. Jap, D. Iles, M. Coerwindel, M. Hofker, A.V. Carrano, P.J. de Jong, R.G. Korneluk, and B. Wieringa. 1992. Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. Nat. Genet. **1**: 261-266.

Johnson, K., P. Shelbourne, J. Davies, J. Buxton, E. Minno, M.J. Siciliano, L.L. Bachinski, H. Harley, S. Rundle, T. Miki, H. Brunner, and R, Williamson. 1990. A new polymorphic probe which defines the region of chromosome 19 containing the myotonic dystrophy locus. Am. J. Hum. Genet. **46**: 1073-1082.

Katz, D.A., N.G. Theodorakis, D.W. Cleveland, T. Lindsten, and C.B. Thompson. 1994. AU-A, an RNA-binding activity distinct from hnRNP A1, is selective for AUUUA repeats and shuttles between the nucleus and the cytoplasm. Nucl. Acids. Res. **22**: 238-246.

Kawaguchi, Y., T. Okamoto, M. Taniwaki, M. Aizawa, M. Inoue, S. Katayama, H, Kawakami, S. Nadamura, M. Nishimura, I. Akiguchi, J. Kimura, S. Narumiya, and A. Kakizuka. 1994. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. Nat. Genet. **8**: 221-228.

Kabnick, K.S., and D.E. Housman. 1988. Determinants that contribute to cytoplasmic stability of human c-*fos* and B-Globin mRNAs are located at several sites in each mRNA. Mol. Cell. Bio. **8**: 3244-3250.

Koide, R., T. Ikeuchi, O. Onodera, H. Tanada, S. Igarashi, K. Endo, H. Takahashi, R. Kondo, A. Ishikawa, T. Hayashi, M. Saito, A. Tomoda, T. Miike, H. Naito, F. Ikuta, and S. Tsuji. 1994. Unstable expansion of CAG repeat in hereditary dentatorubral--pallidoluysan atrophy (DRPLA). Nat. Genet. **6**: 9-13.

Korneluk, R.G., A.E. MacKenzie, Y. Nakamura, I. Dubé, P. Jacob, and A.G.W. Hunter. 1989. A reordering of human chromosome 19 long-arm DNA markers and identification of markers flanking the myotonic dystrophy locus. Genomics **5**: 596-604.

Kwon, Y.K., and N.B. Hecht. 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. Proc. Natl. Acad. Sci. USA **88**: 3584-3588.

Kwon, Y.K., and N.B. Hecht. 1993. Binding of a phosphoprotein to the 3' untranslated region of the mouse protamine 2 mRNA temporally represses its translation. Mol. Cell.Bio. **13**: 6547-6557.

La Spada, A.R., E.M. Wilson, D.B. Lubahn, A.E. Harding, and K.H. Fischbeck. 1991. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Science **352**: 77-

Mahadevan, M., C. Tsilfidis, L. Sabourin, G. Shutler, C. Amemiya, G. Jansen, C. Neville, M. Narang, J, Barceló, K. O'Hoy, S. Leblond, J. Earle-MacDonald, P.J. de Jong, B. Wieringa, and R.G. Korneluk. 1992. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. Science **255**: 1253-1255.

Mahadevan, M.S., C. Amemiya, G. Jansen, L. Sabourin, S. Baird, C.D. Neville, N. Wormskamp, B. Seges, M. Batzer, J. Lamerdin, P. de Jong, B. Wieringa, and R.G. Korneluk. 1993. Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene. Hum. Molec. Genet. **2**: 299-304.

Müllner, E.W., and L.C. Kühn. 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.

Müllner, E.W., B. Neupert, and L.C. Kühn. 1989. A specific mRNA binding factor regulates the iron-dependent stability of cytoplasmic transferrin receptor mRNA. Cell **58**: 373-382.

Nanbu, R., P.-A. Menoud, and Y. Nagamine. 1994. Multiple instability-regulating sites in the 3' untranslated region of the urokinase-type plasminogen activator mRNA. Mol. Cell. Bio. **14**: 4920-4928.

Neville, C.E., M.S. Mahadevan, J.M. Barceló, and R.G. Korneluk. 1994. High resolution genetic analysis suggests one ancestral predisposing haplotype for the origin of the myotonic dystrophy mutation. Hum. Molec. Genet. **3**: 45-51.

O'Hoy, K.L., C. Tsilfidis, M.S. Mahadevan, C.E. Neville, J. Barceló, A.G.R. Hunter, and R.G. Korneluk. Reduction in size of the myotonic dystrophy trinucleotide repeat mutation during transmission. Science **259**: 809-812.

Olson, E.N., and W.H. Klein. 1994. bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. Genes & Dev. **8**: 1-8.

Orberlé, I., F. Rousseau, D. Heitz, C. Kretz, D. Devys, A. Hanauer, J. Boué, M.F. Berthea, and J.L. Mandel. 1991. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science **252**: 1097-1102.

Orr, H.T., M.-y. Chung, S. Banfi, T.J. Kwiatowski Jr., A. Servadio, A.L. Beaudet, A.E. McCall, L.A. Suvick, L.P.W. Ranum, and H.Y. Zoghbi. 1993. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. Nat. Genet. **4**: 221-

Puymirat, J., P. Etongue-Mayer, R. Faure, J.P. Bouchard, R. Laframboise, M.C. Thibault. 1994. The myotonic dystrophy gene product is a tyrosine/serine kinase: decreased levels of M-PK activity are associated with the congenital and adult forms of myotonic dystrophy. Am. J. Hum. Genet. **55**: Suppl. A83.

Radvanyi, H.H., A. Laurent, G. Gourdon, M. Vidaud, and C. Junien. 1994. Myotonic dystrophy: increased expression of the normal allele in CDM infants muscle. Am. J. Hum. Genet. **55**: Suppl. A1387.

Rastinejad, F., and H.M. Blau. 1993. Genetic complementation reveals a novel regulatory role for 3' untranslated regions in growth and differentiation. Cell **72**: 903-917.

Rastinejad, F., M.J. Conboy, T.A. Rando, and H.M. Blau. 1993. Tumor suppression by RNA from the 3' untranslated region of a-tropomyosin. Cell **75**: 1107-1117.

Richards, R.I., and G.R. Sutherland. 1994. Simple repeat DNA is not replicated simply. Nat. Genet. **6**: 114-116.

Rudnicki, M.A., T. Braun, S. Hinuma, and R. Jaenisch. 1992. Inactivation of *MyoD* in mice leads to up-regulation of the myogenic HLH gene *Myf-5* and results in apparently normal muscle development. Cell **71**: 383-390.

Sabourin, L.A., M.S. Mahadevan, M. Narang, D.S.C. Lee, L.C. Surh, and R.G. Korneluk. 1993. Effect of the myotonic dystrophy (DM) mutation on mRNA levels of the DM gene. Nat. Genet. **4**: 233-238.

Sabourin, L.A., M. Narang, C. Tsilfidis, C.J. Storbeck, L. Podolsky, and R.G. Korneluk. 1994. Effect of the myotonic dystrophy kinase (DMK) on the terminal differentiation of myoblasts in culture. Am. J. Hum. Genet. **55**: Suppl. A4.

Sabourin, L.A., M. Narang, E. Whiting, J.D. Waring, K. Tamai, and R.G. Korneluk. 1995. Expression of the myotonic dystrophy kinase (DMK) in muscle tissues and primary cell cultures from DM patients. Submitted.

Sachs, A.B. 1993. Messenger RNA degradation in eukaryotes. Cell **74**: 413-421.

Savant-Bhonsale, S., and D.W. Cleveland. 1992. Evidence for instability of mRNAs containing AUUUA motifs mediated through translation-dependent assembly of a >20S degradation complex. Genes & Dev. **6**: 1927-1939.

Schuler, G.D., and M.D. Cole. 1988. GM-CSF and oncogene mRNA stabilities are independently regulated in *trans* in a mouse monocytic tumor. Cell **55**: 1115-1122.

Shaw, A.M., R.A. Barnetson, M.F. Phillips, P.S. Harper, and H.G. Harley. 1994. Evidence for meiotic drive at the myotonic dystrophy locus. Am. J. Hum. Genet. **55**: Suppl. A1413.

Shaw, G., and R. Kamen. 1986. A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell **46**: 659-667.

Shyu, A.-B., J.G. Belasco, and M.E. Greenberg. 1991. Two distinct destabilizing elements in the c-*fos* message trigger deadenylation as a first step in rapid mRNA decay. Genes & Dev. **5**: 221-231.

Shutler, G.G., A.E. MacKenzie, and R.G. Korneluk. 1994. The 1.5-Mb region spanning the myotonic dystrophy locus shows uniform recombination frequency. Am. J. Hum. Genet. **54**: 104-113.

Siomi, H., M.C. Siomi, R.L. Nussbaum. and G. Dreyfuss. 1993. The protein product of the fragile X gene, *FMR1*, has characteristics of an RNA-binding protein. Cell **74**: 291-298.

Smeets, H., L. Bachinski, M. Coerwinkel, J. Schepens, J. Hoelimakers, M. Duin, K.H. van Grzeschik, C.A. Weber, P. De Jan, M.J. Sicilian, and B. Wieringa. 1990. A long range restriction map of the human chromosome 19q13 region: close physical linkage between cumm and the ERCC1 and ERCC2 genes. Am. J. Hum. Genet. **46**: 492-502.

Storbeck, C.J., L. Sabourin, S. Baird, and R.G. Korneluk. 1994. *In vitro* mapping of myotonic dystrophy (DM) gene promoter. Am. J. Hum. Genet. **55**: Suppl. A796.

Sureau, A., and B. Perbal. 1994. Several mRNAs with variable 3' untranslated regions and different stability encode the human PR264/SC35 splicing factor. Proc. Natl. Acad. Sci. USA **91**: 932-936.

Tsilfidis, C., A.E. MacKenzie, G. Shutler, S. Lablond, J. Bailly, K. Johnson, R. Williamson, J. Siegel-Bartelt, and R.G. Korneluk. 1991. D19S51 is closely linked with and maps distal to the myotonic dystrophy locus on 19q. Am. J. Hum. Genet. **49**: 961-965.

Tsilfidis, C., A.E. MacKenzie, G. Mettler, J. Barceló, and R.G. Korneluk. 1992. Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. Nat. Genet. **1**: 192-195.

Tsilfidis, C., E.J. Whiting, A.E. MacKenzie, R.G. Korneluk. 1994. Convergent myotonic dystrophy (DM) haplotypes on 19q13.3: potential inconsistencies in human disease gene localization. Am. J. Hum. Genet. **55**: Suppl. A1195.

Vakalopoulou, E., J. Schaack, and T. Shenk. 1991. A 32-kilodalton protein binds to AU-rich domains in the 3' untranslated regions of rapidly degraded mRNAs. Mol. Cell. Bio. **11**: 3355-3364.

van der Ven, P.F.M., G. Jansen, T.H.M.S.M. van Kuppevelt, M.B. Perryman, M. Lupa, P.W. Dunne, H.J. ter Laak, P.H.K. Jap, J.H. Veerkamp, H.F. Epstein, and B. Wieringa. 1993. Myotonic dystrophy kinase is a component of neuromuscular junctions. Hum. Molec. Genet. **2**: 1889-1894.

Verkerk, A.J.M.H., M. Pieretti, J.S. Sutcliffe, Y.-H. Fu, D.P.A. Kuhl, M.R. Victoria, F. Zhang, B.E. Eussen, G.-J. B. van Ommen, L.A.J. Blonden, G.J. Riggins, J.L. Chastain, C.B. Kunst, H. Galjaard, C.T. Caskey, D.L. Nelson, B.A. Oostra, and S.T. Warren. 1991. Identification of a gene (*FMR-1*) containing a CCG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell **65**: 905-914.

Waring, J.D., R. Haq, M.A. Narang, K. Tamai, M.S. Mahadevan, J. Ikeda, and R.G. Korneluk. 1994. Characterization of myotonic dystrophy kinase in heterologous expression systems. Am. J. Hum. Genet. **55**: Suppl. A799.

Whiting, E.J., J.D. Waring, K. Tamai, M.J. Somerville, M. Hincke, W.A. Staines, J.-E Ikeda, and R.G. Korneluk. 1995. Characterization of myotonic dystrophy kinase (DMK) protein in human and rodent muscle and central nervous tissue. Hum. Molec. Genet. In press.

Wingett, D., R. Reeves, and N.S. Magnuson. 1991. Stability changes in *pim-1* proto-oncogene mRNA after mitogen stimulation of normal lymphocytes. J. of Immu. **147**: 3653-3659.

You, Y., C.-Y.A. Chen, and A.-B. Shyu. 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. Mol. Cell. Bio. **12**: 2931-2940.

Zhang, W., B.J. Wagner, K. Ehrenman, A.W. Schaefer, C.T. DeMaria, D. Crater, K. DeHaven, L. Long, and G. Brewer. 1993. Purification, characterization, and cDNA cloning of an AU-rich element RNA-binding protein, AUF1. Mol. Cell. Bio. **13**: 7652-7665.