

# Myotonic dystrophy: molecular windows on a complex etiology

Zeljka Korade-Mirnics, Paul Babitzke<sup>1</sup> and Eric Hoffman\*

Department of Molecular Genetics and Biochemistry, BSTW1211, School of Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA and <sup>1</sup>Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA

Received November 4, 1997; Revised and Accepted December 17, 1997

## ABSTRACT

**Myotonic dystrophy (DM) is the most common form of adult onset muscular dystrophy, with an incidence of ~1 in 8500 adults. DM is caused by an expanded number of trinucleotide repeats in the 3'-untranslated region (UTR) of a cAMP-dependent protein kinase (DM protein kinase, DMPK). Although a large number of transgenic animals have been generated with different gene constructions and knock-outs, none of them faithfully recapitulates the multisystemic and often severe phenotype seen in human patients. The transgenic data suggest that myotonic dystrophy is not caused simply by a biochemical deficiency or abnormality in the DM kinase gene product. Emerging studies suggest that two novel pathogenetic mechanisms may play a role in the disease: the expanded repeats appear to cause haploinsufficiency of a neighboring homeobox gene and also abnormal DMPK RNA appears to have a detrimental effect on RNA homeostasis. The complex, multisystemic phenotype may reflect an underlying multifaceted molecular pathophysiology: the facial dysmorphism may be due to pattern defects caused by haploinsufficiency of the homeobox gene, while the muscle disease and endocrine abnormalities may be due to both altered RNA metabolism and deficiency of the cAMP DMPK protein.**

## THE CLINICAL AND GENETIC FEATURES OF MYOTONIC DYSTROPHY

Myotonic dystrophy (DM) is a dominantly inherited disorder which shows a very wide range of presentations and progressions (1), with an incidence of ~1 in 8500 adults (2). The eponym derives from the adult onset form, which typically presents with distal muscle weakness (dystrophy) and impaired relaxation of muscle (myotonia) after 20 years of age. The disease is progressive and often leads to significant disability. Characteristic facial changes are also common: low-set ears, a hatchet chin and drooping of the lips and eyelids (ptosis). Severe cases of adult-onset DM also show a high incidence of presenile cataracts, testicular atrophy, diabetes, kidney failure and early frontal balding in males. Mental retardation and 'difficult' personalities can also be seen. A second, more severe, congenital form of DM is characterized by general hypotonia and respiratory distress at birth. The congenital disease shows delayed motor development and mental retardation. In both forms the symptoms

and severity vary greatly among family members and between generations.

Genetic linkage experiments localized the DM gene locus to chromosome 19q13.3 (3). The DM gene is ~14 kb and encodes 2.3 kb of mRNA with 15 exons and a protein (cAMP-dependent kinase) of 624 amino acids (4,5). Surprisingly, no mutations were found within the coding region of the DM gene in DM patients. Instead, the mutation responsible for the DM disease was found in the 3'-untranslated region (3'-UTR) (6–9). The normal DM gene contains 5–37 CTG repeats in the 3'-UTR, while all DM patients have the repeat expanded from 50 to several thousand CTG trinucleotides. It appears that most patients have received the mutant gene from a common ancestor (founder effect). Despite this low initial mutation rate, the mutant gene shows dramatic changes between generations of families (meiotic or germline instability) and also dramatic changes in different cells of any single patient (mitotic or somatic instability) (10–12). On average the severity of DM symptoms correlates with the number of repeats in a patient (7,13).

The molecular mechanisms leading to triplet repeat instability have been the focus of intense study (14–17). Progressive triplet number changes have been reported only for the CTG (CAG), CGG (GCC) and GAA (CTT) sequences (18). Recent studies of replication of repeat sequences in yeast suggest that a replication fork is formed and that the lagging strand is appropriately single-stranded, however, repeating triplets autohybridize and form 'hairpin' structures (17–19). DNA polymerase can misread the hairpin (Okazaki fragments would not be bound to the hairpin structure), leaving a decreased number of triplet repeats on the newly synthesized strand after replication is complete. The same looping and hairpin formation appears to play a crucial role in triplet expansion: once the Okazaki fragments are synthesized on the lagging strand, the fragment containing the triplet repeats would form a hairpin structure with itself or loop back onto the template strand. This would trigger *de novo* replication of the already replicated triplets from the template strand. Endonuclease would not be able to repair this duplication of the Okazaki fragment, rather, after DNA ligation a corresponding repeat expansion would be added to the template strand, resulting in an expansion of the triplet repeat region.

## TRANSGENIC MOUSE MODELS OF THE HUMAN DISEASE

To test if the triplet repeat alone showed meiotic and mitotic instability in mice, Monckton *et al.* (20) created transgenic mice using only the 3'-UTR with 162 CTG repeats. The construct did

\*To whom correspondence should be addressed. Tel: +1 412 648 9024; Fax: +1 412 624 1401; Email: eric@hoffman.mgen.pitt.edu

not include the DM gene itself, but only a small region of its 3'-end. In mice containing a single copy of the transgene transmission through the female germline showed mild meiotic instability, with a tendency for reduction in the number of repeats (+2 to -7), while in male transmission there was a tendency to increase the number of repeating CTG triplets (0 to +7). Somatic/mitotic instability was even less evident, with the number of repeats in different tissues showing little variability ( $\pm 5$ ). In addition, these mice did not show the DM phenotype. Although this model does not attempt to address the pathophysiology of DM, this study showed that the CTG repeats themselves can be intrinsically unstable during meiosis.

The meiotic and mitotic instabilities observed in these transgenic mice were much less dramatic than those seen in human patients. The mice had a pathological repeat of 165, while most human patients have 500 or more repeats. The extent to which repeat changes are observed may simply be correlated with the length of the initial repeat. It is important to note here that the tissue which shows the most dramatic cellular abnormalities in DM patients, namely muscle, is post-mitotic. Yet it has been well established that DM patients show a striking increase in length of the repeat in their muscle as they age (10,21). Thus there must be a mechanism for repeat expansion in non-replicating post-mitotic tissues such as muscle. Here it is interesting to note the sharing of some subunits of the replication machinery responsible for error repair with the transcriptional machinery. This suggests that there may be transcription-coupled DNA repair and that it is this process that is responsible for the dramatic expansions seen in post-mitotic cells (22). This issue has not yet been addressed; all studies to date have looked at repeat instability during replication.

In an attempt to better understand the pathophysiology and symptomatology of DM itself, Reddy *et al.* (23) created DM protein kinase (DMPK) knock-out mice. Surprisingly, the heterozygote animals showed none of the symptoms characteristic of the DM phenotype, suggesting that the dominantly inherited human disease is not due to loss-of-function of the gene (haploinsufficiency). Homozygous mice (complete loss-of-function), while normal up to 9 months, did eventually show some histological evidence of degeneration and fibrosis in muscle, with some subsequent muscle weakness. However, the characteristic plethora of symptoms associated with the DM phenotype, such as myotonia, cataracts and loss of fertility, did not develop despite complete loss-of-function of DMPK. Jansen *et al.* (24) approached the same question using an elaborate experimental design: in addition to knocking out the endogenous mouse DM gene, they simultaneously overexpressed the normal human DM gene. These knock-out/overexpressor mice with multiple copies of the human DM gene also did not develop the characteristic human DM phenotype. However, mice with >20 copies of the human DM gene showed signs of cardiac myopathy and reduced lifespan, but no changes in the striated musculature.

Most recently, Gourdon *et al.* (25) made transgenic mice using the complete 45 kb human genomic locus encompassing a defective DM gene (55 CTG repeats) and two surrounding genes, DMR-N9 and DMAHP (homeobox gene). In this study only 6.8% of descendants showed any increase in the number of CTG repeats (meiotic instability). The mice had 1-4 transgene copies and in most cases only one of the copies showed any CTG expansion, and then it was very minor (-1 to +6, most common +1). Both germinal and somatic instability were noted, but there

was no parental bias as suggested in the Monckton study (20). No evidence of DM phenotype was seen in the mice.

Although DM transgenic animal models have provided interesting data, they have uniformly failed to reproduce the pathognomonic clinical features of the human disease, a finding which suggests that they may not be reproducing the molecular pathology. They were able, in some cases, to reproduce meiotic and mitotic instability, but to a very minor extent relative to the dramatic germline and tissue changes in repeat size seen in patients (10-12,21). There have been hints of muscle disease, but again not nearly to the extent seen in human patients (20,25). Are these differences simply a consequence of the differing physiology of mouse and human or do they reflect incorrect construction of mouse models (none are exact molecular replicas of the human DM mutations). It is important to examine the human disease, both with regard to inheritance patterns and models of molecular pathogenesis, and determine if the mouse data argues for or against any of the possible models.

## DOMINANTLY INHERITED LOSS-OF-FUNCTION

The DM gene codes for DMPK, a member of the cAMP-dependent serine-threonine subfamily of protein kinases (5). Its catalytic domain shares ~45% identity with the *Drosophila warts* gene, which encodes a protein that is a tumor suppressor (26). Although once thought to be tightly membrane bound, new studies have shown that DMPK is more loosely associated with the membrane and can form high molecular weight disulfide-linked complexes (26). The substrates of DMPK and its specific actions *in vivo* are not known. Experiments *in vitro* suggest that DMPK phosphorylates muscle-specific voltage gated Na<sup>+</sup> channels and may regulate the excitability of muscle cells (27). In addition, Timchenko *et al.* showed that DMPK phosphorylates the  $\beta$ -subunit of the dihydropyridine receptor voltage-dependent Ca<sup>2+</sup> release channel (28).

DM patients are heterozygous for the expansion mutation, with one normal and one mutant DMPK gene. If the expansion mutation destroys the ability of the mutant gene to produce functional DMPK then the disease should fall under the rubric of 'haploinsufficiency'. According to the haploinsufficiency model there is not enough protein produced by the normal allele to prevent biochemical abnormalities (50% is not enough). In this model loss of protein production from the mutant allele could be because the expanded CTG repeats in the 3'-UTR prevent poly(A) tailing or interfere with mRNA and/or protein synthesis. Consistent with this model, most studies have shown a negative effect of the repeat on gene transcription and translation (29-31; Fig. 1). Moreover, studies of patient muscle have shown a decrease in the amounts of DM mRNA and DMPK protein (30-33). However, there are some major problems with haploinsufficiency as a pathogenetic model. First, one would expect *complete* loss-of-function to show a much more severe phenotype than haploinsufficiency. On the contrary, knock-out mice show only a very mild phenotype (23). Moreover, a human patient homozygous for DM has been reported and this patient shows the typical congenital form of the disease (34). A second major problem is more theoretical in nature: DM is one of the most clinically variable inherited disorders known; can cells be so exquisitely sensitive to the levels of DMPK protein? If the most severe disease is due to having 50% of the normal DMPK level (0% from the mutant allele and 50% from the normal allele) and

adult onset is due to having 70% of the normal protein level (20% from the mutant allele and 50% from the normal allele), then this would be the first disorder in any animal showing such dramatic differences in phenotype due to relatively subtle changes in protein levels. However, it is possible that haploinsufficiency does not apply to the DM gene product alone, but also neighboring genes through effects on chromatin structure (see next section).

A final problem with the haploinsufficiency model is that most studies of DMPK mRNA and protein in patient muscle have shown far less than 50% of normal levels (32,33). This data suggests that the mutant gene has some effect on the normal allele (see section on RNA metabolism). In conclusion, data from a variety of sources and experimental systems suggest that the abnormal CTG repeat does *not* cause the disease through simple loss-of-function of the mutant allele.

### DISRUPTION OF NEIGHBORING GENES (CONTIGUOUS GENE SYNDROME)

Expanded CTG repeats have an ability to interfere with the local chromatin structure: Wang and Griffith (35) showed that 75–130 CTG repeats cause an increase in nucleosome assembly in *Xenopus borealis*. In both *in vivo* (35) and *in vitro* (36) experiments nucleosome stability was dependent on the number of repeats, a finding readily explained by hairpin structure formation and autohybridization of the CTG repeats (37). The unusual nucleosome helix structure could also account for the greater than expected electrophoretic mobility of DNA segments containing long stretches of CTG and CGG repeats (38).

This ability of expanded CTG repeats to interfere with the local chromatin configuration could affect expression of both DMPK and neighboring genes. The enhancer and the first exon of the adjacent more centromeric DMAHP homeobox gene are in close proximity to exon 15 of the DM gene, making them a vulnerable target for regional chromatin changes (39–41). The expanded number of CTG repeats could affect transcription from the DMAHP homeobox gene. In fact, recent experiments showed that the level of DMAHP mRNA in fibroblasts and myoblasts from DM patients is significantly lower than in normal control cells (41,42; Fig. 1). Quantification of the mRNA transcribed from the mutant DMAHP allele (DM associated allele with Cac8I polymorphism) showed that DMAHP mRNA is reduced to 10–30% compared with the normal allele (41) or in some cases transcription is completely inactivated (42). Many of the disorders caused by mutations of transcription factors (homeobox, paired box) cause disease through haploinsufficiency (43–45). Fifty percent of the corresponding transcription factor is not enough to fully complete pattern formation and tissue development, leading to dysmorphology. Thus haploinsufficiency of DMAHP caused by the repeat expansion could be the cause of the facial dysmorphology characteristic of DM patients.

Gene 59, upstream of the DM gene, is another candidate to be adversely affected by the CTG repeat expansion (5,46). The mouse homolog of this gene, DMR-N9, is strongly expressed in the brain, heart, liver, testis and kidney, but not detected in skeletal muscle (47). If affected by the chromatin changes, altered transcription from gene 59 could account for some of the clinical features observed in DM patients: mental retardation, reduced fertility and kidney failure.

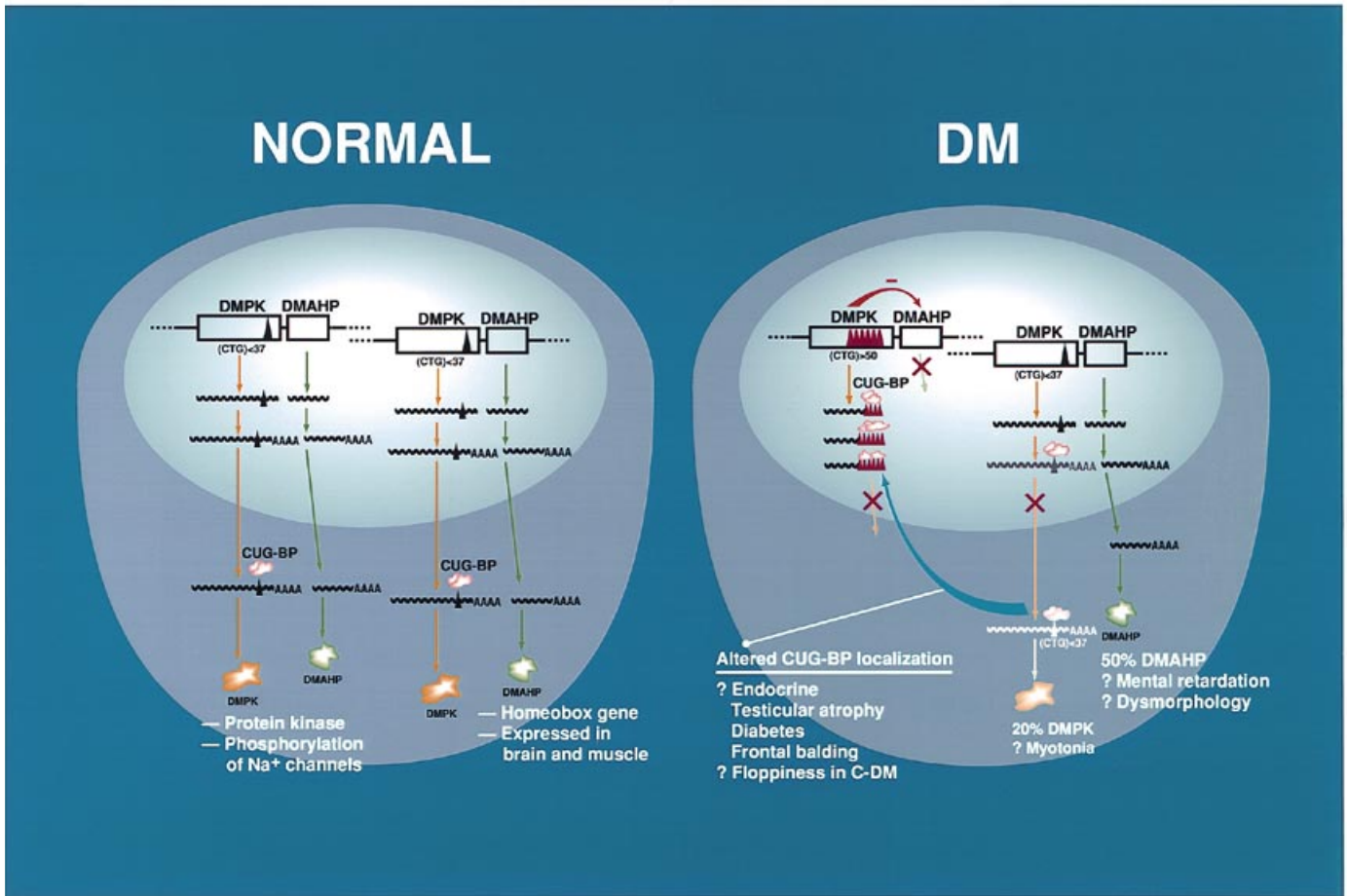
A model where chromatin changes alter the function of several contiguous genes could account for both the congenital and the adult onset phenotypes. In congenital DM the greatly expanded CTG repeats could interfere with many contiguous genes, while the smaller expansions in adult onset DM would have more restricted effects. As homeobox genes play a primary role throughout development and regulate cascades of events, DMAHP deficiency could have serious consequences for a developing organism, leading to a diverse, though specific, symptomatology.

### DEFECTS IN RNA METABOLISM (TRANS-DOMINANT RNA DISORDER)

The pathogenetic mechanisms described above assume that the pathological repeat expansion has a negative (loss-of-function) effect on transcription of the DM gene. On the contrary, a number of recent studies have shown that the mutant allele is in fact transcribed and accumulates to levels in total RNA pools which are similar to the normal allele in patient muscle biopsies (48–50). However, the same patients show a significant reduction in DMPK poly(A)<sup>+</sup> RNA levels (48,49). This suggests that the expanded CTG repeat interferes with normal splicing or polyadenylation. While this RNA processing defect should still lead to loss-of-function of the mutant allele (haploinsufficiency), one study showed that RNA processing of the *normal* DMPK gene was also severely depressed, leading to levels of poly(A)<sup>+</sup> RNA in patient biopsies that were just 20–25% of the levels of that seen in muscles from patients with unrelated disorders (48; Fig. 1). This observation led to the hypothesis that RNA metabolism may be altered in a more global manner, with DMPK RNA containing the expansion mutation having a detrimental effect on processing of many different RNAs (48). Consistent with this hypothesis, abnormal levels of insulin receptor mRNA were observed in patient muscle (51), which could explain the increased insulin resistance seen in many patients.

A series of recent studies using different experimental approaches have provided support to the RNA metabolism defect hypothesis. It is known that RNA–protein interactions are critical for correct processing and trafficking of RNA molecules within cells (52). Some mRNAs have documented localization signals within the 3′-UTR, but the proteins that bind to them are largely unknown (for a review see 52). A number of studies strongly suggest that abnormal RNA–protein interactions play a role in DM pathogenesis: DMPK mRNA in DM patient fibroblasts and myofibers was abnormally accumulated in foci within the nucleus (five to hundreds of copies), while the control cells had a cytoplasmic perinuclear localization of the transcripts (50,53,54). The foci of DMPK RNA contained non-poly(A)<sup>+</sup> RNA and contained the pathological repeat (53), in complete agreement with the RNA studies of Wang *et al.* and Krahe *et al.* cited above (48,49). Two recent papers by Timchenko *et al.* have begun to offer molecular mechanisms for the abnormal RNA metabolism in DM (55,56). Using the band shift assay Timchenko *et al.* isolated a protein, CUG-BP, from HeLa cells that binds to synthetic (CUG)<sub>8</sub> RNA oligonucleotides. Careful analysis of this protein revealed two novel isoforms of a heterogeneous nuclear ribonucleoprotein (hnRNP) of the type called hNab50. The two isoforms were dubbed CUG-BP1 (49 kDa) and CUG-BP2 (51 kDa). hnRNP proteins are typically present within the nucleus, where they bind to poly(A)<sup>+</sup> RNAs and are involved in nuclear RNA processing





**Figure 1.** Processing of DMPK and DMAHP genes within normal and DM cells. In the normal cell, the two DMPK genes are both transcribed, the RNAs processed in the nucleus and transported to the cytoplasm where they are bound by specific RNA binding proteins (CUG-BP), and then translated into the DM protein kinase (DMPK). DMPK has been shown to phosphorylate muscle-specific, voltage-gated sodium channels, and in this way may regulate the excitability of the muscle cell. Centromeric to the DMPK gene is the DMAHP gene. The DMAHP gene contains a homeobox domain and it is expressed in brain and muscle during development and through childhood. The presence of the homeobox domain suggests a possible role of this gene in early pattern establishment. In cells from myotonic dystrophy patients (DM), the DMPK gene is transcribed from both normal and mutant alleles but the transcripts are largely retained within the nucleus. Transcripts from the mutant DMPK allele are retained within the nucleus because the expanded number of CUG repeats bind and sequester mRNA transport proteins (CUG-BP). These proteins are normally cytoplasmic (see 'normal'), but in patient cells CUG-BP is largely localized within the nucleus, where it may interfere with the processing of other types of mRNAs. This *trans*-effect on RNA metabolism may lead to different symptoms such as endocrine defects or floppiness in congenital DM. In addition to retaining the cytoplasmic and nuclear transport proteins within the nucleus, the expanded numbers of CTG repeats within the 3' untranslated region of the DMPK gene interferes with the transcription of the neighbouring DMAHP gene by masking its enhancer that lies within the CTG island. This leads to decreased amount of DMAHP transcript and decreased amount of DMAHP protein. Haploinsufficiency of the DMAHP protein might be responsible for the signs of dysmorphology and mental retardation in DM patients. Finally, the altered RNA metabolism of both the mutant and normal DMPK RNA transcripts leads to deficiency of the DMPK protein (<20%), which could lead to changes in phosphorylation of proteins, and symptoms such as myotonia.

(57,58). Photocrosslinking studies showed that CUG-BP1 and CUG-BP2 were able to preferentially bind the 3'-UTR of poly(A)<sup>+</sup> DMPK mRNA, suggesting that these hnRNPs might be transcript specific or interact with only a subset of mRNAs. Further band shift assay studies showed that CUG-BP1 protein is predominantly in the cytoplasmic fraction and CUG-BP2 is predominantly in the nuclear fraction within normal lymphoblasts. However, in DM lymphoblast lines the amount of CUG-BP1 declined in the cytoplasmic fraction, while CUG-BP2 shifted localization from the cytoplasmic to the nuclear fraction. Although the function of CUG-BP proteins is not clear at present, it is likely that they could function both in nuclei and the

cytoplasm. For example, they could be involved both in splicing of mRNAs containing (CUG)<sub>n</sub> repeats and in their transport.

The only protein other than the CUG-BPs that has been shown to bind RNA triplet repeats is the *trp* RNA binding attenuation protein (TRAP) of *Bacillus subtilis* and its close relative *Bacillus pumilus* (for a review see 59). Tryptophan-activated TRAP regulates expression of the *B. subtilis trp* operon at both the transcriptional and translational levels by binding to 11 (G/U)AG repeats present in the untranslated *trp* operon leader. In addition, TRAP regulates translation of the unlinked *trpG* gene by binding to nine triplet repeats. However, the trinucleotide repeats in the TRAP binding targets are not contiguous, as is the case for

CUG-BP/hNab50, rather the repeats are usually separated by 2 or 3 nt. The *B. subtilis* TRAP complex is composed of 11 identical 75 amino acid subunits arranged in a single ring. This consists of 11 seven-stranded antiparallel  $\beta$ -sheets, with each sheet composed of four  $\beta$ -strands from one subunit and three  $\beta$ -strands from the adjacent subunit. This novel protein structure has been termed the  $\beta$ -wheel (60).

Recent evidence indicates that the TRAP RNA binding domain consists of a novel KKR motif (61). In the crystal structure these three residues are aligned near the outer edge of TRAP. Eleven clusters of this KKR motif, containing Lys37 from one subunit and Lys56 and Arg58 from the adjacent subunit, fall on a single line encircling the TRAP oligopeptide. These findings suggest that the mechanism of RNA binding involves interaction of one KKR motif with one (G/U)AG repeat, with the RNA wrapping around the periphery of the TRAP complex. The amino acid sequence of CUG-BP/hNab50 was found to be related to a family of RNA binding proteins which possess three RNA binding domains (55). It is interesting to note that each of the RNA binding domains in CUG-BP/hNab50 contains at least one KK, KR or RK motif in the primary sequence. It is tempting to speculate that these amino acid residues may play a role in binding to the CUG repeats in the DMPK mRNA.

In the light of these studies it is not hard to envision a mechanism whereby the mutant DM allele leads to the symptoms of the disease (Fig. 1). The 3'-UTR expanded CTG repeats are transcribed, but not processed appropriately, whereupon the transcripts accumulate in the nucleus. In the nucleus the expanded RNA molecules bind excessive numbers of transport proteins, sequestering them from their normal location in the cytoplasm. The abnormal RNA-protein complexes would not be able to leave the nucleus, thereby possibly sequestering other mRNAs in the nucleus as well. This would prevent translation of a number of mRNA species, including normal DMPK mRNA from the normal allele, leading to loss of DMPK, insulin receptor and other proteins. At the same time excessive binding to the extended CTG repeats would create a shortage of transport proteins required for cytoplasmic localization of many RNAs. This could account for the discrepancy seen between normal total RNA and decreased poly(A)<sup>+</sup> RNA levels and would explain the dominant effect of the mutant gene. This model could also explain the multisystemic nature of the disease and the dramatic clinical variability. Different cells and tissues would be affected differently, depending on which subset of RNAs was involved in each cell type. Moreover, the extent to which proteins and RNAs are abnormally sequestered likely depends on the length of the CTG expansion, hence the correlation of CTG expansion with clinical severity.

## CONCLUSION

The concept of repeating triplets causing maladies in humans is not an isolated phenomenon. Triplet repeats within the coding regions of the Huntington (CAG), spinal and bulbar muscular atrophy (CGG) and spinocerebellar ataxia type 1 (CAG) genes interrupt the respective protein (62). In these cases the abnormal protein product containing expanded glutamine tracts causes cellular dysfunction/cell death, resulting in the clinical symptoms of the diseases. However, DM is unique among all inherited disorders in that a mutation in the 3'-UTR of an RNA causes a dominant disease. In DM, unlike the rest of the triplet disease disorders, the 3'-UTR region of the gene expands with CTG

repeats leading to the disorder. Although we know more about this disease than ever, many aspects of the pathophysiology remain elusive. We believe that the full pathophysiology of DM (and other triplet diseases) is a complex process, probably to some degree involving most of the above-described mechanisms.

## REFERENCES

- Harper, P. and Rudel, R. (1994) In Engel, A. and Franzini-Armstrong, C. (eds), *Myology*, 2nd Edn. McGraw-Hill, New York, NY, pp. 1192-1219.
- Harper, P. (1989) *Myotonic Dystrophy*, 2nd Edn. W.B. Saunders Co., Philadelphia, PA.
- Harley, H., Brook, J., Floyd, J., Rundle, S., Crow, S., Walsh, K., Thibault, M.-C., Harper, P. and Shaw, D. (1991) *Am. J. Hum. Genet.*, **49**, 68-75.
- Mahadevan, M.S., Amemiya, C., Jansen, G., Sabourin, L., Baird, S., Neville, C.E., Wormskamp, N., Segers, B., Batzer, M., Lamerdin, J., de Jong, P., Wieringa, B. and Komeluk, R.G. (1993) *Hum. Mol. Genet.*, **2**, 299-304.
- Shaw, D.J., McCurrach, M., Rundle, S.A., Harley, H.G., Crow, S.R., Sohn, R., Thirion, J.-P., Hamsheer, M.G., Buckler, A.J., Harper, P.S., Housman, D.E. and Brook, J.D. (1993) *Genomics*, **18**, 673-679.
- Brook, J. et al. (1992) *Cell*, **68**, 799-808.
- Buxton, J., Shelbourne, P., Davies, J., Jones, C., Tongeren, T., Aslanidis, C., John, P., Jansen, G., Anvret, M., Riley, B., Williamson, R. and Johnson, K. (1992) *Nature (Lond.)*, **355**, 547-548.
- Fu, Y. et al. (1992) *Science*, **255**, 1256-1258.
- Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutter, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., Leblong, S., Earle-MacDonald, J., Hohn, P., Wieringa, B. and Komeluk, R. (1992) *Science*, **255**, 1253-1255.
- Anvret, M., Ahlberg, G., Grandell, U., Hedberg, B., Johnson, K. and Edstrom, L. (1993) *Hum. Mol. Genet.*, **2**, 1397-1400.
- Jansen, G., Willems, P., Coerwinkel, M., Nillesen, W., Smeets, H., Vits, L., Howeler, C., Brunner, H. and Wieringa, B. (1994) *Am. J. Hum. Genet.*, **54**, 575-585.
- Martorell, L., Martinez, J., Carey, N., Johnson, K. and Baiget, M. (1995) *J. Med. Genet.*, **32**, 593-596.
- Harley, H., Rundle, S., MacMillan, J., Myring, J., Brook, J., Crow, S., Reardon, W., Fenton, I., Shaw, D. and Harper, P. (1993) *Am. J. Hum. Genet.*, **52**, 1164-1174.
- Kang, S., Jaworski, A., Ohshima, K. and Wells, R. (1995) *Nature Genet.*, **10**, 213-218.
- Kang, S., Ohshima, K., Shimizu, M., Amirhaeri, S. and Wells, R. (1995) *J. Biol. Chem.*, **270**, 27014-27021.
- Wells, R. (1996) *J. Biol. Chem.*, **271**, 2875-2878.
- Tishkoff, D., Filosi, N., Gaida, G. and Kolodner, R. (1997) *Cell*, **88**, 253-263.
- Gordenin, D., Kunkel, T. and Resnick, M. (1997) *Nature Genet.*, **16**, 116-118.
- Kunkel, T., Resnick, M. and Gordenin, D. (1997) *Cell*, **88**, 155-158.
- Monckton, D., Coolbaugh, M., Ashizawa, K., Siciliano, M. and Caskey, C. (1997) *Nature Genet.*, **15**, 193-196.
- Thornton, C.A., Johnson, K. and Moxley, R.T. (1994) *Annals Neurol.*, **35**, 104-107.
- Seroz, T., Hwang, J., Moncollin, V. and Egly, J. (1995) *Curr. Opin. Genet. Dev.*, **5**, 217-221.
- Reddy, S., Smith, D., Rich, M., Leferovich, J., Reilly, P., Davis, B., Tran, K., Rayburn, H., Bronson, R., Cros, D., Balice-Gordon, R. and Housman, D. (1996) *Nature Genet.*, **13**, 325-335.
- Jansen, G. et al. (1996) *Nature Genet.*, **13**, 316-324.
- Gourdon, G., Radvanyi, F., Lia, A.-S., Duros, C., Blanche, M., Abitbol, M., Junien, C. and Hofmann-Radvanyi, H. (1997) *Nature Genet.*, **15**, 190-192.
- Waring, J.D., Haq, R., Tamai, K., Sabourin, L.A., Ikeda, J.-E. and Komeluk, R.G. (1996) *J. Biol. Chem.*, **271**, 15187-15193.
- Chahine, M. and George, A. (1997) *FEBS Lett.*, **412**, 621-624.
- Timchenko, L., Nastainczyk, W., Schneider, T., Patel, B., Hofmann, F. and Caskey, C.T. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 5366-5370.
- Parsons, M.A., Izban, M.G. and Sinden, R.R. (1997) *Genes Dev.*, submitted for publication.
- Fu, Y.-H., Friedman, D.L., Richards, S., Pearlman, J.A., Gibbs, R.A., Pizzuti, A., Ashizawa, T., Perryman, H.B., Scarleto, G., Fenwick, R.G. and Caskey, C.T. (1993) *Science*, **260**, 235-238.
- Carango, P., Nobler, J.E., Marks, H.G. and Funanage, V.L. (1993) *Genomics*, **18**, 340-348.
- Hofmann-Radvanyi, H., Lavedan, C., Rabes, S., Savoy, D., Duros, C., Johnson, K. and Junien, C. (1993) *Hum. Mol. Genet.*, **2**, 1263-1266.

- 33 Novelli,G., Gennarelli,M., Zelano,G., Pizzuti,A., Fattorini,C., Caskey,C. and Dallapiccola,B. (1993) *Biochem. Mol. Biol. Int.*, **29**, 291-297.
- 34 Martorell,L., Illa,I., Rosell,J., Benitez,J., Sedano,M. and Baiget,M. (1996) *J. Med. Genet.*, **33**, 783-785.
- 35 Wang,Y.-H. and Griffith,J. (1995) *Genomics*, **25**, 570-573.
- 36 Wang,Y.-H., Amirhaeri,S., Kang,S., Wells,R. and Griffith,J. (1994) *Science*, **265**, 1709-1712.
- 37 Pearson,C. and Sinden,R. (1996) *Biochemistry*, **35**, 5041-5053.
- 38 Chastain,P.D., Eichler,E.E., Kang,S., Nelson,D.L., Levene,S.D. and Sinden,R.R. (1995) *Biochemistry*, **34**, 16125-16131.
- 39 Otten,A. and Tapscott,S. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 5465-5469.
- 40 Boucher,C.A. *et al.* (1995) *Hum. Mol. Genet.*, **4**, 1919-1925.
- 41 Klesert,T., Otten,A., Bird,T. and Tapscott,S. (1997) *Nature Genet.*, **16**, 402-406.
- 42 Thornton,C., Wymer,J., Simmons,Z., McClain,C. and Moxley,R. (1997) *Nature Genet.*, **16**, 407-409.
- 43 Zlotogora,J., Lerer,I., Bar-David,S., Ergaz,Z. and Abeliovich,D. (1995) *Am. J. Hum. Genet.*, **56**, 1173-1178.
- 44 Tassabehji,M., Newton,V. and Read,A. (1994) *Nature Genet.*, **8**, 251-255.
- 45 Steingrimsson,E., Moore,K., Lamoreux,M., Ferre-D'Amare,A., Burley,S., Sanders Zimring,D., Skow,L., Hodgkinson,C., Arnheiter,H., Copeland,N. and Jenkins,N. (1994) *Nature Genet.*, **8**, 256-263.
- 46 Jansen,G., Mahadevan,M., Amemiya,C., Wormskamp,N., Segers,B., Hendriks,W., O'Hoy,K., Baird,S., Sabourin,L., Lennon,G., Jap,P., Iles,D., Coerwinkel,M., Hofker,M., Carrano,A., de Jong,P., Korneluk,R. and Wieringa,B. (1992) *Nature Genet.*, **1**, 261-266.
- 47 Jansen,G., Bachner,D., Coerwinkel,M., Wormskamp,N., Hameister,H. and Wieringa,B. (1995) *Hum. Mol. Genet.*, **4**, 843-852.
- 48 Wang,J., Pegoraro,E., Menegazzo,E., Gennarelli,M., Hoop,R., Angelini,C. and Hoffman,E. (1995) *Hum. Mol. Genet.*, **4**, 599-606.
- 49 Krahe,R., Ashizawa,T., Abbruzzese,C., Roeder,E., Carango,P., Giacanelli,M., Funanage,V.L. and Siciliano,M.J. (1995) *Genomics*, **28**, 1-14.
- 50 Hamshe,M., Newman,E., Alwazzan,D., Athwal,D. and Brook,J. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 7394-7399.
- 51 Morrone,A., Pegoraro,E., Angelini,C., Zammarchi,E., Marconi,G. and Hoffman,E. (1997) *J. Clin. Invest.*, **99**, 1691-1698.
- 52 St Johnston,D. (1995) *Cell*, **81**, 161-170.
- 53 Taneja,K.L., McCurrach,M., Schalling,M., Housman,D. and Singer,R.H. (1995) *J. Cell Biol.*, **128**, 995-1002.
- 54 Davis,B., McCurrach,M., Taneja,K., Singer,R. and Housman,D. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 7388-7393.
- 55 Timchenko,L.T., Miller,J.W., Timchenko,N.A., DeVore,D.R., Datar,K.V., Lin,L., Roberts,R., Caskey,C.T. and Swanson,M.S. (1996) *Nucleic Acids Res.*, **24**, 4407-4414.
- 56 Timchenko,L.T., Timchenko,N.A., Caskey,C.T. and Roberts,R. (1996) *Hum. Mol. Genet.*, **5**, 115-121.
- 57 Wilson,S., Datar,K., Paddy,M., Swedlow,J. and Swanson,M. (1994) *J. Cell Biol.*, **127**, 1173-1184.
- 58 Anderson,J., Wilson,S., Datar,K. and Swanson,M. (1993) *Mol. Cell Biol.*, **13**, 2730-2741.
- 59 Babitzke,P. (1997) *Mol. Microbiol.*, **26**, 1-9.
- 60 Antson,A., Otridge,J., Brzozowski,A., Dodson,E., Dodson,G., Wilson,K., Smith,T., Yang,M., Kurecki,T. and Gollnick,P. (1995) *Nature*, **374**, 693-700.
- 61 Yang,M., Chen,X.-P., Militello,K., Hoffman,R., Fernandez,B., Baumann,C. and Gollnick,P. (1997) *J. Mol. Biol.*, **270**, 696-710.
- 62 Hoffman,E.R. (1992) *Curr. Biol.*, **2**, 309-311.