increase k_{cat} [M. L. Bender, G. E. Clement, C. R Gunter, F. J. Kezdy, J. Am. Chem. Soc. 86, 3697 (1964)]. Alaninamide increased k_{cat} equally for the hydrolysis of both sucAAPF-SBzl and sucAAPF-AMC by Ch, which indicates that deacylation is rate determining for both compounds. Alaninamide also increased k_{cat} for the hydrolysis of sucAAPF-SBzl by D189S and mutant Tr \rightarrow Ch[S1+L1+L2], which confirms that deacylation is the rate determining step in these reactions (L. Hedstrom, unpublished experiments).

- 20. In the elastase reaction, increasing peptide length also increases acylation in preference to binding or deacylation [R. Ć. Thompson and E. R. Blout, Proc. Natl. Acad. Sci. U.S.A. 67, 1734 (1970)].
- L. Brady et al., Nature 343, 767 (1990); F. K.
 Winkler, A. D'Arcy, W. Hunziker, *ibid.*, p. 771; M.
 M. G. M. Thunnissen et al., *ibid.* 347, 689 (1990); J. D. Schrag, Y. Li, S. Wu, M. Cygler, *ibid.* **351**, 761 (1991); J.-P. Wery *et al.*, *ibid.* **352**, 79 (1991); J. L. Sussman et al., Science 253, 872 (1991)
- 22. Acetylcholinesterase hydrolyzes acetylcholine 105fold faster than the analogous amide (k_{cat}/K_m) , which is consistent with the relative reactivity of esters and amides [D. E. Moore and G. P. Hess, Biochemistry 14, 2386 (1975)]. In contrast, Ch hydrolyzes sucAAPF-SBzl only sixfold faster than sucAAPF-pNA.
- 23. R. A. Blevins and A. Tulinsky, J. Biol. Chem. 260, 4264 (1985).
- 24. P. G. Jones and A. J. Kirby, J. Chem. Soc. Chem. 24. F. G. Jones and R. J. Kirby, J. Chem. Sol. Chem. Sol. Chem. Commun. 1979, 288 (1979).
 25. We found that Tr, D189S, and Tr→Ch[S1+L1]
- +L2] were also expressed in yeast culture medium by fusing the trypsinogen coding sequences to the α-factor leader sequence as described for carboxy-peptidase A1 [M. A. Phillips, R. Fletterick, W. J. Rutter, J. Biol. Chem. 265, 20692 (1990)]. The final expression vector pYT contains the inducible alcohol dehydrogenase–glyceraldehydephosphate dehydrogenase (ADH-GAPDH) promoter and regulatory regions, the GAPDH transcription terminator, and amp, ura3, and leu2d markers. Medium was isolated from a 2-liter culture by centrifugation and was concentrated on a phenyl-Sepharose column. Trypsinogen was activated and purified (26). Typically, 30 mg of purified Tr was obtained. Trypsins isolated from the yeast expression system have kinetic constants identical to those of enzymes isolated from the Escherichia coli system.
- 26. Rat trypsinogen II mutants were expressed in the periplasm of E. coli and purified with (NH4)2SO4 precipitation and anion exchange chromatography [L. Gráf et al., Biochemistry 26, 2616 (1987)]. The trypsinogens were activated by enterokinase treatment, and Tr was isolated by affinity chromatogra-phy on SBTI Sepharose (Sigma Chemical Co.). Typically, 1 mg of purified mutant Tr was obtained. The reported activities are not likely the result of misfolded protein or contaminating activities. First, all of the mutants bound quantitatively to SBTI Sepharose when activated by enterokinase. The mutant trypsins were stable at 4°C for several months. The activity of mutant Tr→Ch[S1+L1+L2] is observed in several independent preparations of Tr→Ch[S1+L1+L2] and not in any other mutant Tr preparation. Mutants D189S and Tr→Ch-[S1+L1+L2] were expressed in both *E. coli* and yeast (25), with identical steady-state kinetic constants for the hydrolysis of sucAAPF-AMC and benzoyltyrosylethyl ester. The k_{cat} value for ester hydrolysis is equivalent for Ch, Tr, D189S, and $Tr \rightarrow Ch[S1+L1+L2]$, which indicates that the activity is unlikely to represent contaminating proteases and that the preparations are fully active
- 27. M. Zimmerman, B. Åshe, E. C. Yurewicz, G. Patel, Anal. Biochem. 78, 47 (1977).
- L. Gráf, I. Boldogh, L. Szilagyi, W. J. Rutter, in Protein-Structure Function, Z. H. Zaidi, A. Abbasi, M. Smith, Eds. (TWEL Publishers, Karachi, Pakistan, 1990), pp. 49-65. 29. K. Brady and R. H. Abeles, Biochemistry 29, 7608
- (1990).
- 30. W. Bode and P. Schwager, J. Mol. Biol. 98, 693 (1975)
- 31. R. Huber et al., ibid. 89, 73 (1974).
- 32. T. E. Ferrin, C. C. Huang, L. E. Jarvis, R. Lang-

- ridge, J. Mol. Graph. 6, 13 (1988). 33. J. R. Brown and B. S. Hartley, Biochem. J. 101, 214 (1966).
- S. D. Pinsky, K. S. LaForge, V. Luc, G. Scheele, Proc. Natl. Acad. Sci. U.S.A. 80, 7486 (1983).
- K. Titani, L. H. Ericsson, H. Neurath, K. A. Walsh, 35. Biochemistry 14, 1358 (1975). 36. C. S. Craik et al., J. Biol. Chem. 259, 14255
- (1984). 37. This paper is dedicated to R. H. Abeles on the

occasion of his 65th birthday. This work was suported by NIH grant DK21344. We thank J. Cross, M. Phillips, and M. Garabedian for comments on the manuscript. This work is part of an ongoing collaborative project between the W. J. Rutter laboratory at the University of California at San Francisco and the L. Gráf laboratory at Loránd Eötvös University

23 October 1991: accepted 13 January 1992

Myotonic Dystrophy Mutation: An Unstable CTG Repeat in the 3' Untranslated Region of the Gene

MANI MAHADEVAN, CATHERINE TSILFIDIS, LUC SABOURIN, GARY SHUTLER, CHRIS AMEMIYA, GERT JANSEN, CATHERINE NEVILLE, Monica Narang, Juana Barceló, Kim O'Hoy, Suzanne Leblond, JANE EARLE-MACDONALD, PIETER J. DE JONG, BÉ WIERINGA, **ROBERT G. KORNELUK***

Myotonic dystrophy (DM) is the most common inherited neuromuscular disease in adults, with a global incidence of 1 in 8000 individuals. DM is an autosomal dominant, multisystemic disorder characterized primarily by myotonia and progressive muscle weakness. Genomic and complementary DNA probes that map to a 10-kilobase Eco RI genomic fragment from human chromosome 19q13.3 have been used to detect a variable length polymorphism in individuals with DM. Increases in the size of the allele in patients with DM are now shown to be due to an increased number of trinucleotide CTG repeats in the 3' untranslated region of a DM candidate gene. An increase in the severity of the disease in successive generations (genetic anticipation) is accompanied by an increase in the number of trinucleotide repeats. Nearly all cases of DM (98 percent or 253 of 258 individuals) displayed expansion of the CTG repeat region. These results suggest that DM is primarily caused by mutations that generate an amplification of a specific CTG repeat.

YOTONIC DYSTROPHY (DM) IS an autosomal dominant disorder characterized primarily by myotonia and progressive muscle weakness, although central nervous system, cardiovascular, and ocular manifestations also frequently occur (1). The disease shows a marked variability in expression, ranging from a severe congenital form that is frequently fatal just after birth to an asymptomatic condition associated with normal longevity. Furthermore, the progression of DM in affected families may exhibit genetic anticipation, an increase in the severity of the disease in

successive generations. (2).

We have cloned the essential region of human chromosome 19q13.3 that contains the DM locus (3-5). We isolated genomic and cDNA probes that map to this region and that detect an unstable, 10-kb Eco RI genomic fragment in individuals affected with DM. This variable length polymorphism is similar to the instability at the fragile X locus (6-8). The physical map location and genetic characteristics of the variable length polymorphism are consistent with it being the cause of DM. We now further characterize the variable length polymorphism segment to determine the molecular basis of the unstable genomic region.

We identified, using Southern (DNA) blot analysis, two different DNA polymorphisms mapping within the 10-kb Eco RI genomic fragment (Fig. 1). The polymorphisms are detected by two genomic probes (pGB2.6 and pGP1.5) that map to this Eco RI fragment. In normal individuals, these probes detect an 8.5- or 9.5-kb Hind III or a 9.0- or 10.0-kb Eco RI insertion polymorphism. This variation is due to a 1-kb insertion within a 150-bp Pst I-Bg1 I fragment (Fig. 1). The second DNA polymorphism maps to a 1.5-kb Bam HI fragment located 4.0 kb centromeric to the insertion poly-

M. Mahadevan, C. Tsilfidis, L. Sabourin, G. Shutler, C. Neville, M. Narang, J. Barceló, K. O'Hoy, Department of Microbiology and Immunology, University of Otta-wa, Ottawa, Ontario, Canada K1H 8L1. C. Amemiya, and P. J. de Jong, Lawrence Livermore National Laboratories, Livermore, CA 94550. G. Jansen, Department of Human Genetics, University

G. Jansen, Department of Human Genetics, University of Nijmegen, 6500 HB, Nijmegen, the Netherlands. S. Leblond and J. Earle-Macdonald, Research Institute,

Children's Hospital of Eastern Ontario, Ottawa, Canada K1H 8L1.

B. Wieringa, Department of Cell Biology and Histology, University of Nijmegen, 6500 HB, Nijmegen, the Netherlands

R. G. Korneluk, Department of Microbiology and Immunology, University of Ottawa, Ottawa, Ontario, Can-ada K1H 8L1 and Research Institute, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada KÎH 8L1.

^{*}To whom correspondence should be addressed.

morphism and is associated exclusively with chromosomes from individuals with DM. DNA sequence analysis of this genomic fragment revealed a trinucleotide CTG repeat located approximately 300 bp distal to the centromeric Bam HI site (Fig. 1). Chromosome 19 cosmids F18894 and Y100263 (5) contained 20 and 11 CTG repeats, respectively.

We have used the 1.5-kb Bam HI frag-

Fig. 1. (A) Restriction map of the genomic segment containing the variable length polymorphism detected at the DM locus. The map was constructed by the analysis of subclones derived from the cosmid Y100263, a genomic clone previously described (5). The normal insertion polymorphism and the variable length polymorphism (triangles 1 and 2, respectively) are detected by two sequence probes, pGB2.6 and pGP1.5. The orientation of the map is given from telegram (TFP) given from telomere (TER) to centromere (CEN), and the position of relevant restriction sites are given: B, Bam HI; Bg, Bg1 I; E, Eco RI; H, Hind III; and P, Pst I. (B) Nucleotide sequence of the 3' region of a cDNA clone. The clone was isolated from a human heart cDNA library (Stratagene, Inc.) with the 1.5-kb Bam HI fragment containing the

ment, as well as other genomic probes from the DM region, to isolate human and mouse cDNAs from brain, muscle, and heart libraries. Sequence analysis shows that some of the isolated cDNAs contain the variable CTG repeat (Fig. 1). Multiple stop codons are found in all three reading frames 5' to the trinucleotide repeat. Furthermore, a polyadenylation signal and polyadenylated



by the entire 875-bp sequence of the 3'-most exon. The nucleotide position designated by the number I corresponds to the first base of the last exon. A consensus polyadenylation signal sequence is underlined. Relevant restriction enzymes are shown in italics above their recognition sequences. Primers used for sequencing and PCR amplification are underlined and numerically designated. Indicated in bold letters are the five CTG repeats located between primers 409 and 410 found in this cDNA. The genomic sequence derived from cosmid clones F18894 and Y100263 is identical to the sequence of the last exon of the cDNA, except there are 20 and 11 CTG repeats, respectively. Plasmid DNA was prepared by a routine alkaline lysis method. Dideoxy chain termination reactions with vector primers and internal primers were performed with fluorescent dye-labeled dideoxynucleotides, according to manufacturer's specifications (Applied Biosystems), and subsequently run on an ABI 373A automated sequencer. DNA sequencing was also performed with radioactive nucleotides to resolve sequence ambiguities

Fig. 2. Southern blot analysis showing varying degrees of allelic expansion in DM-affected individuals. Eco RI allele sizes of 9 kb and 10 kb are found in the normal population (lanes 1, 2, and 6). Genetic phasing of all our DM families reveals that the 10-kb allele is the expanding allele in DM individuals. Classification of expansion is as follows: E0, no expansion; E1, expansion of 0 to 1.5 kb; E2, expansion of 1.5 to 3.0 kb; E3, expansion of 3.0 to 4.5 kb; E4, expansion of 4.5 to 6.0 kb or more. Approximately 30% of DM individuals (affected status verified by DNA typing) show no



allelic expansion on the Southern blot level (lanes 8 and 11). Varying degrees of expansion are seen in other DM individuals, ranging from E1 (lane 9) to E2 (lanes 3 and 10) to E3 (lanes 5 and 7) to E4 (lane 4). N refers to normal individuals. Genomic DNA (5 µg) was digested with Eco RI, separated by electrophoresis on 0.8% agarose gels, and transferred to nylon membranes. Southerns were probed with the Bam HI-Eco RI fragment of pGB2.6, washed in saline sodium citrate [0.2× standard saline citrate (SSC)] with 0.1% sodium dodecyl sulfate (SDS) at 65°C, and exposed to x-ray film for 1 to 4 days.

tail are present in these clones approximately 500 bp downstream of the repeat region. Therefore, the CTG repeat is contained within the 3' untranslated region of this DM candidate gene.

We investigated the relation of both DNA polymorphisms to DM by Southern blot analysis. Genomic DNA from a total of 280 normal and 258 affected individuals from 98 DM families were used in the study. The insertion polymorphism in normal individuals can be detected with a variety of restriction enzymes (Fig. 1). In all of our affected families, the DM chromosome segregated with the larger insertion allele, which has a frequency of 0.60 in the normal population. This strong linkage disequilibrium suggests that there exists a limited number of DM mutations. Alternatively, the larger 10-kb Eco RI allele may be somehow predisposed to DM mutation.

Increases in the size of the variable length polymorphism (5, 9, 10) up to 6.0 kb and greater were detected (Fig. 2). Nearly 70% of individuals with DM (180 of 258) show distinct increases in the size of the variable length polymorphism by Southern blot analysis. Furthermore, almost all DM families (96 of 98) have at least one individual with expanded DM alleles. Normal individuals (280 analyzed) do not exhibit allelic expansion.

Expanded alleles often have a blurred appearance on Southern blots (Fig. 2). This suggests a somatic cell heterogeneity in the size of the expanded alleles similar to that seen in the fragile X mutation (6-8, 11). Furthermore, preliminary results suggest there is a tendency for the DM variable length polymorphism alleles to increase in size in successive generations of a family and that the greater the expansion of these variable length polymorphism alleles, the greater the clinical severity of the disease. Therefore, genetic anticipation may, in fact, be a real phenomenon in DM (2) and may not simply be due to ascertainment bias, as originally suggested (12).

To determine whether an increased number of CTG repeats is responsible for allelic expansion, primers for the polymerase chain reaction (PCR) flanking the CTG repeat region were derived from the DNA sequence in Fig. 1. All primer combinations produce similar patterns of amplification in normal individuals. A substantial variability in the size of normal alleles can be seen with agarose gel electrophoresis and ethidium bromide staining. Polyacrylamide gel electrophoresis (Fig. 3) revealed that the sizes of normal alleles corresponded to variability in the number of CTG repeats, ranging from 5 to 30 (124 normal chromosomes analyzed). CTG repeat numbers of 5 and 13 are the most common in the normal population

(frequencies of 35% and 19%, respectively). The overall heterozygosity among normal individuals is 81%.

PCR amplification of the CTG repeat region of genomic DNA from individuals affected with DM revealed that although the unaffected alleles are readily amplified, the mutant alleles are not usually visible by ethidium bromide staining and agarose gel electro-



Flg. 3. Size distribution of CTG repeats in normal individuals. Analysis of ³²P-labeled PCR products by denaturing acrylamide electrophoresis shows repeat lengths ranging from 5 to 28 repeats. The most common repeats in the normal population (5 repeats, 35%; 13 repeats, 19%) are designated by the arrows. Primer 409 (1 μ g) was radioac-tively labeled with γ -[³²P]adenosine triphosphate. CTG repeats were amplified with genomic DNA from various normal individuals in standard PCR reactions containing 50 ng of both primer 409 and 410, and 10 ng of radiolabeled primer 409. Amplified products were resolved by gel electrophoresis on 8% polyacrylamide. Dried gels were exposed to Kodak X-AR5 film and autoradiographed at room temperature for 17 hours. Sequencing reactions generated from a control template were used as molecular size standards.

Fig. 4. Southern blot analysis of PCRamplified genomic DNA from normal (N) and DM-affected (E) individuals. Southerns were probed with a (CTG)10 oligonucleotide. PCR-amplified DNA from DM individuals shows a distinct smearing of the hybridization signal, presumably due to heterogeneity of the expanded allele. In contrast, the oligonucleotide probe hybridized only to the normal alleles in unaffected individuals. Individuals who show no expansion (E0) at the Southern



blot level show clear expansion with the PCR-based assay (lane 10). In some PCR-amplified samples from DM individuals, the hybridization smear produced by the labeled (CTG)10 oligonucleotide is faint and could be mistaken for that of an unaffected individual (lane 13). However, examination of Southern blot analysis of Eco RI- or Hind III-digested genomic DNA probed with pGB2.6 or pGP1.5 reveals the presence of a greatly expanded allele (E4) in this individual. Presumably, a greatly increased number of CTG repeats is refractory to PCR amplification under the conditions used. Thus, appropriate molecular diagnosis of DM should include both Southern blot analysis of genomic DNA and the hybridization analysis of PCR-amplified DNA. Genomic DNA (2 µg) was PCR-amplified with primers 406 and 409 by a standard protocol and 30 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1.5 min. Amplified products were separated by electrophoresis on 1% agarose gels, transferred onto nylon membranes, and probed with a labeled (CTG)10 oligonucleotide. Membranes were washed in 6× SSC with 0.1% SDS at 40°C and exposed to x-ray film overnight.

phoresis. In order to determine whether the mutant alleles were in fact amplified, a synthetic (CTG)10 oligonucleotide was used to probe a Southern blot of PCR products. Genomic DNA amplified by PCR from individuals with DM showed a distinct smearing of the hybridization signal above the normal alleles, presumably due to heterogeneity of the expanded allele. In contrast, the oligonucleotide probe hybridized only to the normal alleles in unaffected individuals (Fig. 4).

As discussed above, nearly 70% of individuals with DM show distinct expansion of the variable region by Southern blot analysis of genomic DNA. However, most of the remaining 30% that do not show allelic expansion by Southern blotting show some increase in the sizes of the mutant allele by the PCR-based oligonucleotide assay. Only a few of our DM families (2 of 98) are negative by both types of analysis. These families may represent unique mutations of the DM gene that do not involve an increase in the number of CTG repeats.

Occasionally, in DM individuals a faint higher molecular size band detectable by ethidium bromide and agarose gel electrophoresis is amplified by PCR. This occurs only with DNA from DM individuals that show minimal or no visible expansion of the 10-kb Eco RI or 9.5-kb Hind III allele. These bands produce an intense hybridization signal when bound to the $(CTG)_{10}$ oligonucleotide (Fig. 4). Selective PCR amplification of these bands followed by direct DNA sequence analysis showed no differences in the DNA sequence flanking the CTG repeat (from primer 407 to primer 409) (Fig. 1) in comparison to normal individuals. Only the number of CTG repeats appeared to vary between normal and DM alleles, although this was difficult to

determine unequivocally because of severe compression problems in the sequencing gels, a problem similarly encountered in the DNA sequence analysis of CGG repeat in the fragile X mutation (11). However, the compression problems were typically evident after about ten trinucleotide repeat units were sequenced from either strand of the mutant allele. These sequencing data, combined with the observation that strong hybridization smears were produced by probing with the (CTG)₁₀ oligonucleotide, suggests that the difference in this region between the DM and normal alleles is due to the variable number of CTG repeats.

Our results suggest that DM is caused primarily by a mutation or mutations that increase the size of a genomic fragment due to the amplification of a CTG repeat. This is similar to the trinucleotide repeat mutation in fragile X (8, 11) and in X-linked spinal and bulbar muscular atrophy (13). Nearly all individuals with DM (98%) in our population display an increased number of CTG repeats, which can be detected by Southern blot analysis of genomic DNA or hybridization analysis of a PCR-based assay or both. We have demonstrated that these trinucleotide repeats are located within the last exon of the DM gene. This gene has recently been suggested to encode a putative serine-threonine protein kinase, as deduced from the DNA sequence analysis of human (14) and homologous mouse (15) cDNAs. In light of the surprising result that the DM mutation is found in the 3' noncoding region of the DM candidate cDNA, it will be interesting to examine the role of this defect in the pathogenesis of myotonic dystrophy.

- 1. P. S. Harper, Myotonic Dystrophy (Saunders, Phila-delphia, ed. 2, 1989).
- 2.
- C. J. Höweler et al., Brain 112, 779 (1989). G. Shutler et al., Genomics, in press. 3.
- 4. G. Jansen et al., ibid., in press
- C. Aslanidis et al., Nature 355, 548 (1992). 5.
- I. Oberlé et al., Science 252, 1097 (1991)

- S. Yu et al., ibid., p. 1179.
 A. J. M. H. Verkerk et al., Cell 65, 905 (1991).
 J. Buxton et al., Nature 355, 547 (1992).
 H. G. Harley et al., ibid., p. 545.
 E. J. Kremer et al., Science 252, 1711 (1991).
 J. L. Charley et al., Science 252, 1711 (1991).

- L. S. Penrose, Ann. Eugen. 14, 125 (1948).
 A. R. La Spada, E. M. Wilson, D. B. Lubahn, A. E.
- Harding, K. H. Fischbeck, Nature 352, 77 (1991). 14. Y.-H. Fu et al., Science 255, 1256 (1992).
- G. Jansen *et al.*, in preparation.
 We thank A. MacKenzie and L. Surh for their critical evaluation of this work and K. Noel for technical assistance. Supported by grants to R.G.K. from the Muscular Dystrophy Associations (MDA) of Canada and the United States, the Medical Research Council of Canada, and the Canadian National Centres of Excellence Genetic Disease Network; to B.W. from the Dutch Praeventiefonds and the U.S. MDA; and to P.deJ. from the U.S. Department of Energy. M.M. is supported by a Medical Re-search Council postdoctoral fellowship and L.S. is supported by an FCAR scholarship.

31 January 1992; accepted 14 February 1992

REFERENCES AND NOTES