Alternative myosin hinge regions are utilized in a tissue-specific fashion that correlates with muscle contraction speed

Venetia L. Collier, William A. Kronert, Patrick T. O'Donnell, Kevin A. Edwards, and Sanford I. Bernstein¹

Biology Department and Molecular Biology Institute, San Diego State University, San Diego, California 92182 USA

By comparing the structure of wild-type and mutant muscle myosin heavy chain (MHC) genes of Drosophila melanogaster, we have identified the defect in the homozygous-viable, flightless mutant Mhc¹⁰. The mutation is within the 3' splice acceptor of an alternative exon (exon 15a) that encodes the central region of the MHC hinge. The splice acceptor defect prevents the accumulation of mRNAs containing exon 15a, whereas transcripts with a divergent copy of this exon (exon 15b) are unaffected by the mutation. In situ hybridization and Northern blot analysis of wild-type organisms reveals that exon 15b is used in larval MHCs, whereas exons 15a and/or 15b are used in adult tissues. Because Mhc¹⁰ mutants fail to accumulate transcripts encoding MHC protein with hinge region a, analysis of their muscle-specific reduction in thick filament number serves as a sensitive assay system for determining the pattern of accumulation of MHCs with alternative hinge regions. Electron microscopic comparisons of various muscles from wild-type and Mhc¹⁰ adults reveals that those that contract rapidly or develop high levels of tension utilize only hinge region a, those that contract at moderate rates accumulate MHCs of both types, and those that are slowly contracting have MHCs with hinge region b. The presence of alternative hinge-coding exons and their highly tissue-specific usage suggests that this portion of the MHC molecule is important to the isoform-specific properties of MHC that lead to the different physiological and ultrastructural characteristics of various Drosophila muscle types. The absence of other alternative exons in the rod-coding region, aside from those shown previously to encode alternative carboxyl termini, demonstrates that the bulk of the myosin rod is not involved in the generation of isoform-specific properties of the MHC molecule.

[Key Words: Myosin heavy chain; Drosophila; muscle mutant; alternative RNA splicing]

Received January 12, 1990; revised version accepted March 6, 1990.

Muscle myosin heavy chain (MHC) is a contractile protein whose rod-like domain forms the backbone of thick filaments and whose globular head domain acts as a bridge to actin-containing thin filaments (for review, see Harrington and Rodgers 1984). ATP-dependent deformation of the myosin cross-bridge results in muscle contraction. Functionally different muscles usually accumulate different isoforms of MHC (for review, see Emerson and Bernstein 1987). The amino acid variations among MHC isoforms may impart alternative myosin ATPase activities, actin or myosin light-chain-binding affinities, or thick filament assembly properties. Drosophila melanogaster is unusual in that a single muscle MHC gene, rather than a multigene family, encodes all forms of muscle MHC via alternative RNA splicing (Bernstein et al. 1986; Rozek and Davidson 1986; Wassenberg et al. 1987; George et al. 1989; Hess et al. 1989). Analysis of the location of these alternative exons and their tissue-specific pattern of expression should help to elucidate the regions of the protein that impart isoformspecific properties to the MHC molecule. Alternative splicing at the 3' end of the MHC transcript has been well documented, and it results in the production of proteins with alternative carboxyl termini (Bernstein et al. 1986; Rozek and Davidson 1986). MHC proteins containing one carboxyl terminus accumulate in larval muscles and in some adult muscles, whereas those containing the other carboxyl terminus are mostly found in adult thoracic musculature, as well as in some muscles of the adult head (Bernstein et al. 1986; Kazzaz and Rozek 1989).

The Drosophila MHC locus is haploinsufficient for flight muscle function (Bernstein et al. 1983), and mutations in the MHC gene that induce a flightless pheno-

¹Corresponding author.

Collier et al.

type have been isolated (Mogami and Hotta 1981; Mogami et al. 1986). A large proportion of these mutations are dominant flightless and homozygous lethal; the recessive lethality of these mutations likely arises as a result of disrupting MHC synthesis in all muscle types (Mogami et al. 1986; O'Donnell and Bernstein 1988). In contrast to these recessive-lethal alleles, several homozygous-viable MHC mutations cause muscle-specific reductions in MHC accumulation (Chun and Falkenthal 1988; O'Donnell et al. 1989). These mutations likely affect alternative exons that are used in muscles whose functions are not essential to viability.

Here, we report how the analysis of one of these mutants, Mhc^{10} , led us to examine alternative splicing in the rod-coding region of the *Drosophila* MHC gene. MHC gene and cDNA sequence analysis indicates that the hinge domain of the rod is encoded by two mutually exclusive alternative exons. We found that the homozygous-viable Mhc^{10} mutation is a defect in the splice acceptor of one of these exons. Ultrastructural analyses of mutant and wild-type organisms, along with in situ hybridization studies, indicate that some muscles contain MHCs with one hinge region, whereas other muscles have MHCs of both types. These observations suggest the MHC hinge plays a key role in the musclespecific function of the myosin molecule.

Results

Analysis of the molecular defect in the Mhc¹⁰ allele

The homozygous-viable, dominant flightless mutant Mhc¹⁰ fails to accumulate MHC RNA, MHC protein, or thick filaments in the indirect flight muscles (IFM) and jump muscles of the adult and has reduced levels of MHC in leg muscles (O'Donnell et al. 1989). We were interested in determining the molecular lesion responsible for these defects and reasoned that it was probably associated with an alternative exon that was specifically included in transcripts of the affected muscles. We molecularly cloned the mutant MHC gene and sequenced the region containing the alternatively spliced 3' penultimate exon, which had been shown previously to be utilized in the affected muscle tissues (Bernstein et al. 1986; Kazzaz and Rozek 1989). However, no defects in the coding region or splice junctions of this or its flanking exons were found.

Having failed to detect a mutation in the penultimate exon of the Mhc^{10} allele, we examined upstream regions of the MHC gene to search for additional sites of alternative splicing. By comparison of the deduced amino acid sequence to previously sequenced MHCs, we were able to derive the exon/intron structure of the Drosophila MHC rod-coding region (Fig. 1A). The vast majority of the rod is encoded by constitutive exons, and the rod displays the expected 28-amino-acid residue repeat that is important to the assembly properties of the myosin molecule (McLachlan and Karn 1982). Examination of the gene sequence, along with restriction mapping and sequencing of several cDNA clones, revealed only one additional site of alternative splicing, which is within the hinge-coding region. Two 79-nucleotide alternative exons (exons 15a and 15b, according to the notation of George et al. 1989) are used in a mutually exclusive manner and encode amino acid sequences that differ by 72% (Fig. 1B). Sequence analysis of these exons in Mhc^{10} DNA revealed a single-base-pair change that mutates the consensus 3' splice site of exon 15a from CAG to CAA (Fig. 1C).

Stage- and tissue-specific use of the alternative hingecoding regions

We examined the stage-specific usage of the alternative hinge-coding exons by preparing antisense RNA probes and hybridizing them to electrophoretic gel blots of RNA isolated from larvae and pupae (Fig. 2). At the larval stage, two size classes of transcripts (6.1 and 6.6 kb) accumulate in the wild type; these arise from the use of alternative polyadenylation sites and exclusion of the 3' penultimate exon (Bernstein et al. 1986; Rozek and Davidson 1986). Both larval transcript classes hybridize to exon 15b, but neither hybridizes to exon 15a. The pattern for Mhc10 larvae is identical to that of wild type. At the pupal stage, both larval size classes of mRNA accumulate in wild-type organisms, as do two additional transcripts (6.6 and 7.1 kb) that result from inclusion of the 3' penultimate exon and polyadenylation at either of two sites (Bernstein et al. 1986; Rozek and Davidson 1986). Note that the smaller-sized pupal-specific mRNA containing the 3' penultimate exon comigrates with the larger sized mRNA lacking that exon. As in larvae, exon 15b hybridizes to the 6.1- and 6.6-kb mRNAs in both the wild-type and mutant pupae. Exon 15a, however, which hybridizes to the 6.6- and 7.1-kb mRNAs in the wildtype, does not detectably hybridize to transcripts in Mhc^{10} pupae (Fig. 2C). Prolonged exposure of Northern blots probed with exon 15a (not shown) reveals that this exon is included in 6.6- and 7.1-kb mRNAs of Mhc10 at \sim 5% the level of wild type.

The 95% reduction in exon 15a-containing transcripts in the mutant likely arises because the 3' splice junction of exon 15a has been mutated from CAG to CAA. Note, however, that the mutation results in the production of an AG dinucleotide (CAAG). This cryptic splice junction may be used with low efficiency as a result of the lack of a consensus pyrimidine before the AG junctional signal (unspliced transcripts would likely be degraded). It is also possible that the cryptic splice junction is used efficiently (Smith et al. 1989) but that the spliced transcripts are unstable due to the presence of a stop codon in exon 15a resulting from a translational frameshift (see Fig. 1C). Transcript instability caused by a truncated open reading frame has been demonstrated previously in other systems (Baserga and Benz 1988). Any truncated MHC proteins produced from the mutant mRNA would likely be degraded, as sarcomeric MHCs require their rod regions for stability (Dibb et al. 1985; O'Donnell and Bernstein 1988).

А	'	EXON 12			
	VF	FRAG	V L G	Q M E E	FR
14170 14180 14190 14200 14210 14220 14220	GGTGTTCT	TCCGCGCCGG1	GTCCTGGGTC	AGATGGAGGA	GTTCCGT
DERLGKIMSWMONWARCYISDECE	114240		14260	142/0	14280
GATGAGCGTCTGGGCAAGATCATGTCCTGGATGCAGGCATGGGCCCCGTGGTTACCTGTCCCGTAAGGCGCTTCAA		V L V R Aggàacagege	CTCCCCCCA	ACCTTCTCCA	CCGCAAT
14290 14300 14310 14320 14330 14340 14350	14360	14370	14380	14390	14400
L R K Y L Q L R T W P W Y K L W O K V K P L L N V	SR	IEDE	IA		
CTGCGCAAGTACCTGCAGCTCCGTACCTGGCCCTGGTACAAACTGTGGCAGAAGGTCAAGCCCCTCCTCAACGT	CAGCCGCA	TCGAGGATGAC	ATTGCOGTGA	GTATTCCCCA	GAACGAA
14410 14420 14430 14440 14450 14460 S-1/Rod	14480	14490	14500	14510	14520
TGAGAACGAGACCAGCCTGGGTCTGGGGACTTAGTGGTCTTTGGTGATTCGATGCGATGAACTTGGACAGTACAG	GTTGCTGG	TGGCAATCATT	TGGTGACCGC	AGCACGTTGA	TGTAGCC
14530 14540 14550 14560 14570 14580 14590	14600	14610	14620	14630	14640
ATAGCCCTGGTGTGTCCATTCCATTTGGTCACTGTCTGGGGTAATTGGATGAGCTGGGGCAGCTGCAAGCAGCA	ATCAAATC	CCCAACTCAA1	CTTATTTAG	CCAATGGTCA	TTATCTT
14650 14660 14670 14680 14690 14700 14710	14720	14730	14740	14750	14760
14770 14780 14790 14800 14810 14820 14820		ACCCACTACCA	GCTGCCCGAA	ACATATCGAT	ACGATCT
GATTCCAATCGCTGGGCTAGCCAAAAATAGACGCTGTCATAATCTTTTTCCCACTCACACCCAATCGCCCCAAAAATAGACGCTGTCATAATCTTTTTCCCCCCCC	1989U	000000000000000000000000000000000000000	14000	14570 TCC33TCC33	14050
14890 14900 14910 14920 14930 14940 14950	14960	14970	14980	14990	15000
TCACCTCTCACATTTTCATACACCCAAACCATCCCCCCATACGTACATTGTATGTA	TCGAAAGT'	TATGACTTGT	GTTTGGGATC	TATTCTTGGC	CCTGCCA
15010 15020 15030 15040 15050 15060 15070	15080	15090	15100	15110	15120
TATAGCTGGGGATTTTTTAAAGAGCCACAATTAGTTTGCCAATTAGCTTGCTAATTAAATAATAAAAAAATGACCG	CGAGGATG	CCCACGTATAG	TTCCGATCCC	CGATGAGTAA	TTTACCA
15130 15140 15150 15160 15170 15180 15190	15200	15210	15220	15230	15240
15250 15260 15270 15280 15280 15280 15280	CGCACAAG	TGCACTTGAT	GGCCCCATAC	CGACGCAATA	TAGCATT
ATAGCAATGCCAGCAATTCGTGAATTATCATGGCTACCTCCTAACCAAAGCCAATTCGCGAATTCGTGAATTACCATGGCTAACCCCAACCCAATTGCGCGAATTCGCGCAACCCCAA	15320		15340	15350	15360
15370 15380 15390 15400 15410 15420 15430	15440	15450	15460	15470	15480
CTGCTCGCATCGGGGCGTTTTTGCTAATTTTATAGCCAAAGTGGCCCAGTGCATTTGCAGCAGCCATTCCAGTTC	GTATTTGA	CCAGTCGCCG1	CGCGGTAACG	CGTACCACAA	AAAATCG
15490 15500 15510 15520 15530 15540 15550	15560	15570	15580	15590	15600
AACCCCTAACGAAGAACCAATCGATCTGAGCAAAAGAGCCACATTAGCCACAGAGCATCTTGACTGGCATCAAGA	ATAAC <u>GAA</u>	<u>etc</u> aaattgaa	ATTGTATTGT.	AGTTGACAAT	TGTGAAA
15610 15620 15630 15640 15650 15660 15670	Eco	RI	15700	15710	15720
AGCCCCTTTCCAAGAGAGTTTCACATTTAATTAGGAGTTATTGAAAGTGCTTAATCATCAACATGGCCGATGAA	AAGAAAGC	CAAGAAAACGA	AAAAGTCCAC	CGAATCGACC	ACACCCA
15730 15740 15750 15760 15770 15780 15790	15800	15810	15820	15830	15840
15850 15860 15870 15870 15870 15870 15870 15870	ATCGACCG	CCGTCGAACCA	CCACAAAATC	CCCAGCCAGC	TGACGAG
15656 15666 15676 15666 15890 <u>15900 15910</u>	15920	19930	13940	19990	123900
	K A I	F F T. H		v 	K F
CTCAGCTCCGGCAACAACCCTTCTAACCCATCAAATGCCTCTAATGACTTACACGTCTGGAGGAGAGGCCAAC	GAAGGCTG	AGGAACTGCAT	GCCGCTGAAG	TGAAGGTGCG	CAAGGAG
15970 15980 15990 16990 16990	16040				
<u> </u>	10040	16050	16060	16070	16080
LEALNAKLLAEKTALLDSLSGEKGA	L Q I	16050 DYQE	16060 RNA	16070 KLTA	16080 Q K
L E A L N A K L L A E K T A L L D S L S G E K G A CTCGAGGCCCTCAACGCCAAGCTTTTGGCTGAGAAGACCGCTCTGCTGGAGACCCCTGTCCGGCGAGAAGGGGGGCC	L Q I CCTGCAGG	16050 D Y Q E ACTACCAGGAG	16060 RNA CGCAACGCCA	16070 K L T A AGTTGACCGC	16080 Q K CCAGAAG
L E A L N A K L L A E K T A L L D S L S G E K G A CTCGAGGCCCTCAACGCCAAGCTTTTGGCTGAGAAGACCGCTCTGCTGGACTCCCTGTCCGGCGAGAAGGGTGCC 16090 16100 16110 16120 16130 16140 16150	L Q I CCTGCAGGI 16160	16050 D Y Q E ACTACCAGGAG 16170	16060 R N A CGCAACGCCA 16180	16070 K L T A AGTTGACCGC 16190	16080 Q K CCAGAAG 16200
LEALNAK LIAEKTALIGUE 16010 16010 16020 16030 LEALNAKLIAEKTALIDSLSGEKGA CTCGAGGCCCTCAACGCCAAGCTTTGGCTCAGAAGACCGCTCTGCTGCAGCTGCACGCGAAGGGGGGGC 16090 16100 16110 16120 16130 16140 16150 NDLENQLCR	L Q I CCTGCAGGI 16160	16050 DYQE ACTACCAGGAG 16170	16060 R N A CGCAACGCCA 16180	16070 K L T A AGTTGACCGC 16190	16080 Q K CCAGAAG 16200
LEAD 1330 1330 16010 16010 16020 16030 L E A L N A K L L A E K T A L L D S L S G E K G A CTCGAGGCCTCAACGCCAAGCTTTIGGCTGAAAGACGCCTCGGCTCGG	L Q I CCTGCAGGA	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA	16060 R N A CCGCAACGCCA 16180	16070 K L T A AGTTGACCGC 16190 CGAAATTGAT	16080 Q K CCAGAAG 16200 CAGAGGC
LEAD 15350 15350 16010 16010 16020 16030 L E A L N A K L L A E K T A L L D S L S G E K G A CTCGAGGCCCTCAACGCCAAGCTTTTGGCTGAGAAGACCGCTCTGGGGCTCCCGGCGAGAAGGGTGCC 16090 16100 16110 16120 16130 16140 16150 N D L E N Q L R AACGACTCGAGAACCAGCTGCGGGGTAAGTATCCCATTAATCCCATACACTAAATCCCATTATGCGATGGCATA 16210 16220 16230 16240 16250 16260 16270 AATCGAGCATGGAAGTATCCCCCTATTCCCCCCCCCATACTACGCATACACCACTACACCACTACACCACTACACCACTACACCAC	L Q I CCTGCAGGA 16160	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290	16060 R N A CCGCAACGCCA 16180 	16070 K L T A AGTTGACCGC 16190 CGAAATTGAT 16310	16080 Q K CCAGAAG 16200 CAGAGGC 16320
LEAD 15350 15350 16000 16010 16020 16030 L E A L N A K L L A E K T A L L D S L S G E K G A CTCGAGGCCTCAACGCCAAGCTTTTGGCTGAGAAGACCGCTCTGCTGGACTCCCGGCGAGAAGGTGCC 16090 16100 16110 16120 16130 16140 16150 N D L E N Q L R AACGACCTCGAGAACCACCTGCGGGTAAGTATCCCATTAATCCCATACACTAAATCCCATTTATGCGATGGCAT 16210 16220 16230 16240 16250 16260 16270 AATCGAGCATGAAAGTTGGTCGGCTATTTCTGGGCTCTGGCGCACTTGGCTAAATTTAGGAGGTGACCCACTGCT 	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTC/ 16400	16050 DYQE ACTACCAGGAG <u>16170</u> GTGCCACGAAA 16290 AGCCACTTAAC	16060 R N A CCGCAACGCCA 16180 CCGGAAAACG 16300 ACGCCCACAC 16420	16070 K L T A AGTTGACCGC 16190 CGAAATTGAT 16310 AATTTTGTTC 16430	16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA
LEAR <thlear< th=""> LEAR <th< td=""><td>L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTC/ 16400</td><td>16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410</td><td>16060 R N A CCGCAACGCCA 16180 GCGGAAAACG 16300 ACGCCCACAC 16420</td><td>16070 K L T A AGTTGACCGC 16190 CGAAATTGAT 16310 AATTTTGTTC 16430</td><td>16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16440</td></th<></thlear<>	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTC/ 16400	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410	16060 R N A CCGCAACGCCA 16180 GCGGAAAACG 16300 ACGCCCACAC 16420	16070 K L T A AGTTGACCGC 16190 CGAAATTGAT 16310 AATTTTGTTC 16430	16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16440
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	16040 L Q I CCTGCAGGJ 16160 16280 AAATCCAAC 16280 AACTTCTCJ 16400 D Q E	16050 DYQE ACTACCAGGAQ 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI	16060 R N A CCGCAACGCCA 16180 GCGGAAAACG 16300 ACGCCCACAC 16420 K K D	16070 K L T A AGTTGACCGC 16190 CGAAATTGAT 16310 AATTTTGTTC 16430 I E D	16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16440 L E L
LEADUL 1999U 1999U 16000 16010 16020 16030 L E A L N A K L L A E K T A L L D S L S G E K G A CTCGAGGCCCTCAACGCCATGCTGTGGGGGGGGGGGGGG	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTCJ 16400 D Q E ATCAGGAGJ	16050 DYQE ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI ATCTCTGGCCT	16060 R N A CCCAACGCCA 16180 GCCGGAAAACG 16300 ACCGCCACAC 16420 K K D GAAGAAGGAC	16070 K L T A AGTTGACCGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC	16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16440 L E L TGGAATT
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTCJ 16400 D Q E ATCAGGAGJ 16520	16050 DYQE ACTACCAGGAG 516170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI ATCTCTGGCCT 16530	16060 R N A CCGCAACGCCA 16180 GCGGGAAAACG 16300 ACGCCCACAC 16420 K K D GAAGAAGGAC 16540	16070 K L T A AGTTGACCGC 16190 CGAAATTGAT 16310 AATTTTGTTC 16430 I E D ATCGAGGATC 16550	16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16440 L E L TGGAATT 16560
LEAR <thlear< th=""> LEAR LEAR <thl< td=""><td>L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTCJ 16400 D Q E ATCAGGAGJ 16520 4 Q D</td><td>16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCC1 16530 E L I N</td><td>16060 R N A CGCAACGCCA 16180 </td><td>16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K</td><td>16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q</td></thl<></thlear<>	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTCJ 16400 D Q E ATCAGGAGJ 16520 4 Q D	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCC1 16530 E L I N	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K	16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	L Q I CCTGCAGG 16160 AAATCCAAG 16280 AACTTCTC/ 16400 D Q E ATCAGGAG 16520 4 Q D ACCAGGAT	16050 D Y Q E ACTACCAGGAG 5TGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCCT 16530 E L I N GAGCTCATCAA	16060 R N A :CGCAACGCCA 16180 :GCGGAAAACG 16300 :ACGCCCACAC 16420 K K D :GAAGAAGGAC 16540 K L N :CAACTTGAAC	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTCC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA	Q K Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	L Q I CCTGCAGG 16160 AAATCCAAG 16280 AACTTCTC/ 16400 Q E ATCAGGAG 16520 4 Q D ACCAGGATC 16640	16050 DYQE ACTACCAGGAG 16170 STGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI ISGI ELIN GAGCTCATCAG	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670	16080 Q K CCAGAAG 16200 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA 16680
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	L Q I L Q I CCTGCAGGJ 16160 16280 AACTCTC/ 16400 O Q E ATCAGGAGJ 16520 H Q D ACCAGGAT(16640 K A K	16050 DYQE ACTACCAGGAG 16170 STGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI 16530 ELIN SAGCTCATCAA 16650 LEQT	16060 R N A CCGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTCGACGATT	16080 Q K CCAGAAG 16200 AATTCCA 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA 16680 S L E CCCTCCA
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTCJ 16400 O Q E ATCAGGAGJ 16520 H Q D ACCAGGATG 16640 K A K AGGCCAAGG	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCCT 16530 E L I N GAGCTCATCAA 16650 L E Q T CTCGAGCAGAC	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTGGAGGATT 16790	16080 Q K CCAGAAG 16200 AATTCCA 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA 16680 S L E CGCTGGA 1680
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	L Q I L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTCJ 16400 O Q E ATCAGGAGJ 16520 H Q D ACCAGGATC 16640 K A KAGCCAAGG 16760 E A V	16050 DYQE ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI ATCTCTGGCCT 16530 ELIN SAGCTCATCAA 16650 LEQT TCTCGAGCAGAC 16770 ADLE	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTGGAGGATT 16790 K E L	16080 Q K CCAGAAG 16200 CAGAGGC 16320 AATTCCA 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA 16680 S L E CGCTGGA 16800 E Q T
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTCJ 16400 D Q E ATCAGGAGJ 16520 H Q D ACCAGGATC 16640 K A K AGGCCAAGG 16760 E A V AGGCTGTTC	16050 DYQE ACTACCAGGAG 15170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI 16530 ELIN GAGCTCATCAA 16650 LEQT TCGGACCAGAC 16770 ADLE GCCGATCTGGA	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTGGAGATT 16790 K E L AAGGAGCTCG	16080 Q K CCAGAAG 16200 AATTCCA 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA AGATGCA 16680 S L E CGCTGGA 16800 E Q T AGCAGAC
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AAATCCCAAG 16280 AACTTCTCJ 16400 C Q E ATCAGGAGJ 16520 H Q D ACCAGGATG 16520 H Q D ACCAGGATG 16640 K A K AGGCCAAGG 16760 E A V AGGCTGTTC 16880	16050 D Y Q E ACTACCAGGAGA 16170 5TGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCC1 16530 E L I N GAGCTCATCAA 16650 L E Q T CTCGAGCAGAC 16770 A D L E SCCGATCTGGA 16890	16060 R N A CCGCAACGCCA 16180 CGCGGAAAACG 16300 CACGCCCACAC 16420 K K D GAAGAAGGAC 16540 K L N CCAGTTGAAC 16660 L D E CCTCGATGAAC 16780 R N K GCCCAACAACA 16900	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGATC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAAA 16670 L E D CTGGAGGATT 16790 K E L AAGGAGCTCG 16910	16080 Q K CCAGAAG 16200 CAGAGGC 163200 AATTCCA 16440 L E L TGGAATT 16550 K M Q AGATGCA K M Q AGATGCA 16800 E Q T AGCAGAG 16920
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTCTC7 16400 D Q E ATCAGGAGJ 16520 H Q D ACCAGGAT 16520 H Q D ACCAGGAT 16640 K A K AGGCCAAGG 16760 E A V AGGCTGTTC 16880 H Q R	16050 D Y Q E ACTACCAGGAG STGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCC1 16530 E L I N GAGCTCATCAA 16650 L E Q T CTCGAGCAGAC 16770 A D L E SGCCGATCTGGA 16890 Q I K E	16060 R N A :CGCAACGCCA 16180 :GGGGAAAACG 16300 :ACGCCCACAC 16420 . K K D :GAAGAAGACA 16540 :K L N :CAAGTTGAAC 16560 :L D E :CCTCGATGAAC 16780 . R N K :GCGCAACAACA 16900 . L Q A	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGATC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTGGAGGATT 16790 K E L AAGGAGCTCG 16910 R I E	16080 Q K CCAGAAG CCAGAAG 163200 AATTCCA 16440 L E L TGGAATT 16550 K M Q AGATGCA 16680 S L E CGCTGGA 16800 E Q T AGCAGAC 16920 E L E
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTCTCJ 16400 O Q E ATCAGGAGJ 16520 4 Q D ACCAGGAGT 16540 K A K AGGCCAAGG 16760 E A V AGGCTGTTC 16880 4 Q R ACCAGGGCC	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCC1 16530 E L I N GAGCTCATCAA 16650 L E Q T CTCGACCAGAC 16770 A D L E GCCGATCTGG A D L E GCCGATCTGG Q I K E CAGATCAAGGA	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTCC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTGGAGGATC 16790 K E L AAGGAGCTCG 16910 R I E CGCATCGAGG	16080 Q K CCAGAAG CCAGAAG 16320 AATTCCA 16320 AATTCCA 16320 L E L TGGAATT 16550 K M Q AGATGCA 16680 S L E CCCTGGA 16800 E Q T AGCAGAC 16920 E L E L G AGCTCGA
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	L Q I CCTGCAGG 16160 AAATCCAAG 16280 AAATCCAAG 16280 AACTCTC7 16400 Q E ATCAGGAGJ 16520 4 Q D ACCAGGAGT 16540 K A K AGGCCAAGG 16760 E A V AGGCCGAGG 16880 4 Q R ACCAGCGCC 17000	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCCT 16530 E L I N GAGCTCATCAA 16650 L E Q T TCCGACCAGAC 16770 A D L E GCCGATCTGGA 16890 Q I K E CAGATCAAGGA 17010	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTCC 16310 ATCTAGACA I E D ATCGAGGATC 16570 K E K AAGGAGAAGA 16670 L E D CTGGAGGATT 16790 K E L AAGGAGCTCG 16910 R I E CGCATCGAGG 17030	16080 Q K CCAGAAG 16200 AATTCCA 1640 L E L 16560 K M Q AGATGCA 16680 S L E CGCTGGA 16800 E Q T AGCAGAC 16800 E Q T AGCAGAC 16900 E L E L S L
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	L Q I CCTGCAGG 16160 AAATCCAAG 16280 AACTCTC7 16400 O Q E ATCAGGAG 16520 H Q D ACCAGGAG 16520 H Q D ACCAGGAG 16760 E A V AGGCCAGGC 16880 H Q R ACCAGCGCC 17000 E L	16050 DYQE ACTACCAGGAG 16170 STGCCACGAAA 16290 AGCCACTTAAC 16410 ISGCCACTTAAC 16410 ISGCCACTTAAC ELIN SAGCTCATCACAA 16650 LEQT CTCGAGCAGAC 16770 ADLE SCCGATCTGGA 16890 QIKKE CAGATCAAGGA 17010 GERL	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTGGAGGATT 16790 K E L AAGGAGCTCG 16910 R I E CGCATCGAGG 17030 G G A	16080 Q K CCAGAAG CCAGAAG CCAGAAG CCAGAGG 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA 16680 S L E CGCTGGA 16800 E Q T AGCAGAC 16920 E L E L E L E CGCTGGAC
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	L Q I L Q I CCTGCAGG 16160 AAATCCAAG 16280 AACTCTC7 16400 Q E ATCAGGAG 16520 H Q D ACCAGGATG 16540 K A K AGGCCAAGG 16700 E A V AGGCTGTTC 16880 H Q R ACCAGGGCC 17000 E L L AGGAATTGC	16050 DYQE ACTACCAGGAG 16170 STGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI AGCCACTTAAC 16530 ELIN SAGCTCATCAA 16650 LEQTI CTCGAGCAGAC 16770 ADLE SCCGATCAGGA 16890 QIKE CAGATCAAGGA 17010 GERL SGCGACCGTCT	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTGGAGGATT 16790 K E L AAGGAGCTCG 16910 R I E CGCATCGAGG 17030 G G A 17150	16080 Q K CCAGAAG CCAGAAG 16200 AATTCCA 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA 16680 S L E CGCTGGA 16680 S L E CGCTGGA 16680 E Q T AGCAGAC 16920 E L E L E L E L S CGCTGGA 17040 T S A CCCTCGC
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	L Q I L Q I CCTGCAGG 16160 AAATCCAAG 16280 AACTTCTC/ 16400 Q E ATCAGGAG 16520 H Q D ACCAGGATC 16540 K A K AGGCCAGGAT 16700 E A V AGGCTGTTC 16880 H Q R ACCAGGGCC 17000 E L AGGAATTGC 17120 D H F	16050 DYQE ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 ISGI AGCCACTTAAC 16530 ELIN SAGCTCATCAA 16650 LEQI CTCGAGCAGCAC 16770 ADLE GCCGATCTGGA 16690 QIKE CAGATCAAGGA 17010 GERL 3GCGACCGTCT 17130 STLA	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAGAAGA 16670 L E D CTGGAGGATT 16790 K E L AAGGAGCTCG. 16910 R I E CGCATCGAGG 17030 G G A 17150 K K H	16080 Q K CCAGAAG 16200 AATTCCA 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA 16680 S L E CGCTGGA 16800 E Q T AGCAGAC 16820 E L E L E L E L S CGCTGGA 17040 T S A CCTCTGC 17160 N D A
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AACTTCTCJ 16400 D Q E ATCAGGAGJ 16520 H Q D ACCAGGAT 16520 H Q D ACCAGGAT 16640 K A K AGGCCAGGT 16880 H Q R ACCAGCGCG 17000 E E L AGGAATTGC 17120 2 H E AGGCACGAG	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCCT 16530 E L I N GAGCTCATCAA 16650 L E Q T CTCGAGCAGAC 16770 A D L E GCCGATCTGGA 16890 Q I K E CAGATCAAGGA 17010 G E R L GGCGAGCGTCT 17130 S T L A TCCACCCTGGC	16060 R N A CCGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTTC 16430 I E D ATCGAGGATC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAGAAGA 16670 L E D CTGGAGGATT 16790 K E L AAGGAGCTCG 16910 R I E CGCATCGAGG 17030 G G A GGCGGTGCCA 17150 K K H H	16080 Q K CCAGAAG 16200 AATTCCA 16320 AATTCCA 16440 L E L TGGAATT 16560 K M Q AGATGCA 16680 S L E CGCTGGA 16800 E Q T AGCAGAC 16920 E L E AGCTCGA 17040 T S A CCTCTGC 17160 N D A AGCATGC
LSTO LSTO LSTO LSOU <thlsou< th=""> LSOU LSOU <thl< td=""><td>L Q I L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AAATCCAAG 16520 4 Q D ACCAGGAGJ 16520 4 Q D ACCAGGAG 16520 4 Q D ACCAGGAG 16640 4 A K AGGCCAGGAG 16880 4 Q R ACGCCAGCGCC 17000 5 E L AGGAATTGC 17120 2 H E AGGCACGAG 17240</td><td>16050 D Y Q E ACTACCAGGAG 16170 STGCCACGAGAG 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCC1 16530 E L I N GAGCTCATCAA 16650 A D L E GCCGATCTGGA 16770 A D L E GCCGATCTAGGA 16890 Q I K E CAGATCAAGGA 17010 G E R I GGCGAGCGTCT 17130 S T L A ICCACCCTGGC 17250</td><td>16060 R N A CGCAACGCCA 16180 </td><td>16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGATC 16310 AATTTGATC 16530 K E K AAGGAGAACA 16670 L E D CTGGAGGAATC 16790 K E L AAGGAGCTCGA 16910 R I E CCCATCGAGG 17030 G G A GGCGGTGCCA 17150 K K H AAGGAGCACA 17270</td><td>16080 Q K CCAGAAG 16200 CAGAGGC 163200 AATTCCA 16440 L E L TGGAATT 16550 K M Q ACATGCA 16800 S L E CGCTGGA 16800 E Q T AGCAGAGC 17040 T S A CCTCTGC 17160 N D A ACGATGC 17280</td></thl<></thlsou<>	L Q I L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AAATCCAAG 16520 4 Q D ACCAGGAGJ 16520 4 Q D ACCAGGAG 16520 4 Q D ACCAGGAG 16640 4 A K AGGCCAGGAG 16880 4 Q R ACGCCAGCGCC 17000 5 E L AGGAATTGC 17120 2 H E AGGCACGAG 17240	16050 D Y Q E ACTACCAGGAG 16170 STGCCACGAGAG 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCC1 16530 E L I N GAGCTCATCAA 16650 A D L E GCCGATCTGGA 16770 A D L E GCCGATCTAGGA 16890 Q I K E CAGATCAAGGA 17010 G E R I GGCGAGCGTCT 17130 S T L A ICCACCCTGGC 17250	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGATC 16310 AATTTGATC 16530 K E K AAGGAGAACA 16670 L E D CTGGAGGAATC 16790 K E L AAGGAGCTCGA 16910 R I E CCCATCGAGG 17030 G G A GGCGGTGCCA 17150 K K H AAGGAGCACA 17270	16080 Q K CCAGAAG 16200 CAGAGGC 163200 AATTCCA 16440 L E L TGGAATT 16550 K M Q ACATGCA 16800 S L E CGCTGGA 16800 E Q T AGCAGAGC 17040 T S A CCTCTGC 17160 N D A ACGATGC 17280
LATAG LSTRU <th< td=""><td>L Q I L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AAATCCAAG AACTCTC7 16400 D Q E ATCAGGAGJ 16520 H Q D ACCAGGAGT 16520 H Q D ACCAGGAGT 16640 K A K AGGCCAGGGC 16760 E A V AGGCCAGCGCC 17000 E E L AGGAATTGC 17120 2 H E AGGAACGAGJ 17240</td><td>16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAGAG 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCCT 16530 E L I N GAGCTCATCAG 16550 E L I N GAGCTCATCAG 16770 A D L E Q T CTCGAGCAGCAGAC 16770 A D L E Q T CTCGAGCAGCAGAC 16770 A D L E Q T CTCGAGCAGCAGCC 16770 A D L E Q T CTCGAGCAGCAGCC 17130 S T L A CCCCCCGGC 17250</td><td>16060 R N A CGCAACGCCA 16180 </td><td>16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGGAGAAGA 16670 L E D CTGGAGGAATA 16790 K E L AAGGAGCTGA 16910 R I E CGCATCGAGG 17030 G G A GGCGGTGCCAA 17150 K K H AAGAAGCACA 17270</td><td>16080 Q K CCAGAAG CCAGAAG 163200 AATTCCA 16440 L E L TGGAATT 165500 K M Q AGATGCA 16800 E Q T AGCAGAG 16800 E Q T AGCAGAG 16800 E L E CGCTGGA 16800 E L E AGCTCGA 17040 T S A ACCATGC 17160 N D A ACGATGC 17280</td></th<>	L Q I L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AAATCCAAG AACTCTC7 16400 D Q E ATCAGGAGJ 16520 H Q D ACCAGGAGT 16520 H Q D ACCAGGAGT 16640 K A K AGGCCAGGGC 16760 E A V AGGCCAGCGCC 17000 E E L AGGAATTGC 17120 2 H E AGGAACGAGJ 17240	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGAGAG 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCCT 16530 E L I N GAGCTCATCAG 16550 E L I N GAGCTCATCAG 16770 A D L E Q T CTCGAGCAGCAGAC 16770 A D L E Q T CTCGAGCAGCAGAC 16770 A D L E Q T CTCGAGCAGCAGCC 16770 A D L E Q T CTCGAGCAGCAGCC 17130 S T L A CCCCCCGGC 17250	16060 R N A CGCAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGGAGAAGA 16670 L E D CTGGAGGAATA 16790 K E L AAGGAGCTGA 16910 R I E CGCATCGAGG 17030 G G A GGCGGTGCCAA 17150 K K H AAGAAGCACA 17270	16080 Q K CCAGAAG CCAGAAG 163200 AATTCCA 16440 L E L TGGAATT 165500 K M Q AGATGCA 16800 E Q T AGCAGAG 16800 E Q T AGCAGAG 16800 E L E CGCTGGA 16800 E L E AGCTCGA 17040 T S A ACCATGC 17160 N D A ACGATGC 17280
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	L Q I L Q I CCTGCAGGJ 16160 AAATCCAAG 16280 AAATCCAAG 16280 AACTCTC7 16400 C Q E ATCAGGAGJ 16520 4 Q D ACCAGGAGC 16760 E A V AGGCCAAGG 16760 E A V AGGCCAGCGCC 17000 E E L AGGATGTTCG 17120 2 H E AGGACGAGJ 17240 ACTTCTGGG	16050 D Y Q E ACTACCAGGAG 16170 GTGCCACGGAA 16290 AGCCACTTAAC 16410 I S G I ATCTCTGGCC1 16530 E L I N GAGCTCATCAA 16650 E L I N GAGCTCATCAA 16650 C L E Q T CTCGACCAGAC 16770 A D L E GCCGATCTGGA 16890 Q I K E CAGATCAAGGA 17010 G E R L GGCGACGTCT 17130 S T L A ICCACCCTGGC 17250 CTAGCTTTTTC	16060 R N A CCGAACGCCA 16180 	16070 K L T A AGTGACGC 16190 CGAAATTGAT 16310 AATTTGTC 16430 I E D ATCGAGGATC 16550 K E K AAGGAGAAGA 16670 L E D CTGGAGGATC 16790 K E L AAGGAGATC 16910 R I E CGCATCGAGG 17030 G G A 17150 K K H 17270 GCTATCGAGA	16080 Q K CCAGAAG CCAGAAG 16200 AATTCCA 16320 AATTCCA 16320 L E L TGGAATT 16550 K M Q AGATGCA 16800 E Q T AGCAGAC 16800 E Q T AGCAGAC 16800 E Q T AGCAGAC 16800 E L E AGCTCGA 16800 E L E CCTCTGC 17160 N D A ACGATGC 17280 TAGAGAGG

Figure 1. Part I (See p. 890 for legend.)

EXON 15a EHDR А ATCTTGAAGGATCTACAGTTTACAGTCTTCTTTCGAAGAGCTTTGGTTGCCGAACCCAACATAATAGATATTTGCTCTTTTTTACCCATTTGCTAATCCAGGGCTGAACACGATCGCCA Bal II 17440 17450 17460 17470 17480 17510 Pvu I T C H N E L N Q T R T A C D Q L G R D K GACTTGCCACAACGAGCTGAATCAGACTCGTACCGCCTGCGAT<u>CAGCTG</u>GGTCGCGATAAGCTAATATGTCGTGATAACTGGCG<u>CCCGAG</u>CAGGACGTCCAGCGATTCATACATATACAG 17<u>540</u> 17550 17560 Pvu II 17580 Ava I AAATTGAAGGTTATTTGGAAAAAAGGTTCAACGCAAGGAAAGCAATTCTCTCGCTTCTCGCCTCTGTATTTGTCTCTGTCTCTGTCTTTATCCCCCCACAAAACATGAACTAACCCGTGAA 17820 17830 17860 17870 17880 EXON 15b A E K E K N E Y Y G Q L N D L R A G V D H I T N E K **20 17930 17940 17950 17960 17970** Sal I ATTGAATTTGATCTCTACATATATTATTGCTCTCATCGCGATGCTGTAAATTTATATGCTGAAAGTTCTAGTTTAGCCACGTCTACTACATCTGTGTATAGTATCAACGTAGTTGCAACG CAACCACCACGAAACTCCAAAAGAAAATTTGATTATGTAAAAACTAAAACGACCCACCACCACTAAAGTAATATCC<u>AAGCTT</u>GATCTACTAAACCCCCAAATGCCTTGTACAGCCACTTGA Hind III 18210 EXON 16 AAQEKIA CACGAAACCAAAATGCCACACTTACAATCAATTATGAATGTTCCCCAATCCCCATCGTGAAAACTAATGAACATCCATTGCTTGAATCACTAATAGCTGCCCAGGAGAAGATCGCCAAGG 18250 18260 18270 18280 18290 18300 18310 18320 Q L Q H T L N E V Q S K L D E T N R T L N D F D A S K K <u>K L</u> S I E N S D L L R Q AGCTGCAGCACCCCTCAACGAGGTGCAGTCGAAACTGGATGAGACCAACAGGACTCTGAACGACTTCGATGCCAGCAAGAAGAAGCTGTCCATTGAGAACTCCGATCTGCTCCGCCAGC 18370 18380 18390 18400 18410 18420 18430 18440 Hinge/LMM 18460 18470 1841 E E A E S Q V S Q L S K I K I S L T T Q L E D T K R L A D E E S R E R A T L L TGGAGGAGGCCGAGTCCCAGGTGTCTCAGCTGTCCAAGATCAAGATCTCTCTGACCACCCAGTTGGAGGACACCAAGCGTCTGGCCGACGAGGAGTCGCGCGAGCGTGCCACCCTTTTG 18490 18500 18510 18520 18530 18540 18550 18560 18570 18580 18590 GKFRNLEHDLDNLREQVEEEAEGKADLQRQLSKANAEA 18590 18600 GCAAGTTCCGCAACTTGGAGCACGACCTGGACAATCTGCGCCAGCAGGTTGAGGAGGAGGAGGCCGAGCCGATCTGCAGCGCCAGCTGAGCCAAGGCCAACGCTGAGGCCCAGGTGT 18610 18620 18630 18640 18650 18660 18670 18680 18690 18700 18710 SKYESDGVARSEELEEAKRKLQARLAEAEETIESLN RSKY о к GGCGCAGCAAGTACGAGTCCGATGGCGTTGCCCGCTCTGAGGAGGTGGAGGAAGCCAAGAGGAAGCTGCAGGCCCGTTTGGCCGAGGCCGAGGAGACCATCGAGTCCCTCAACCAGAAGT 18740 18750 18760 18770 18780 18790 18800 18810 18820 18830 K T K Q R L S T E V E D L Q L E V D R A N A I A N A A E K K Q K A KIGLE D GCATTGGCCTGGAGAAGACCAAGCGTCTGTCCACCGAGGTGGAGGATCTGCAGCTGGAGGTCGACCGTGCCAACGCCATTGCCAACGCTGCCGAGAAGAAGCAGAAGGCCTTCGACA 18850 18860 18870 18880 18890 18900 18910 18920 18930 18940 18950 1896 IGEWKLKVDDLAAELDASQKECRNYSTELFRLKGAYEE AGATCATCGGCGAGTGGAAGCTCAAGGTCGACGATCTGGCTGCTGAGCTGGATGCCTCCCAGAAGGAGTGCCGCAACTACTCCACCGAGCTGTTCCGTCTTAAGGGCGCCTACGAGGAGG 18970 18980 18990 19000 19010 19020 19030 19040 19050 19060 19070 1908 GQEQLEAVRRENKNLADEVKDLLDQIGEGGRNIHEIEKAR GCCAGGAGCAGTTGGAGGCTGTGGGTGGGAGAACAAGAACCTGGCCGATGAGGTCAAGGATCTGCTCGACCAGATCGGTGAGGGTGGCCGCAACATCCATGAGATCGAGAAGGCCCGCA 19090 19100 19110 19120 19130 19140 19150 19160 19170 19180 19190 K R L E A E K D E L Q A A L E E A E A A L E Q E E N K V L R A Q L E L S Q V R O AGCGCCTGGAAGCCGAGAAGGACGAGCTCCAGGCTGCCCTCGAGGAGGCTGAGGCCGCTCTTGAGCAGGAGGAGAACAAGGTGCTCCGCGCTCAGCTTGAGCTGTCCCAGGTGCGCCAGG 19210 19220 19230 19240 19250 19260 19270 19280 19290 19300 EIDRRIQEKEEEFENTRKNHQRALDSMQASLEAEA KGK AGATCGACCGCCGCATCCAGGAGAAGGAGGAGGAGGAGTTCGAGAACACCCGCAAGAACCACCAGCGTGCCCTCGACTCCATGCAGGCTTCCCTCGAAGCCGAGGCCAAGGGCAAGGCTGAGG 19330 19340 19350 19360 19370 19380 19390 L R M K K K L E A D I N E L E I A L D H A N K EXON 17 A N A E A Q K N I K R Y Q Q Q L K D I Q T A L E E E Q R A R D D A R <u>19570</u> <u>1958</u>0 19590 19600 19610 19620 19630 19640 19650 19660 19670 Q L G I S E R R A N A L Q N E L E E S R T L L E Q A D R G R R Q A E Q E A D 19690 19700 19710 19720 19730 19740 19750 19760 19770 19780 19790 19800 A H E Q L N E V S A Q N A S I S A A K R K L E S E L Q T L H S D L D E L L N E A GCCCACGAGCAGCTGAACGAAGTGTCCGCCCAGAACGCCTCCATCTCCGCTGCCAAGAGGAAGCTGGAGTCCGAGCTGCAGACCCTGCACCTCGACGTGGACGAACTCCTGAACGAAGCC 19810 19820 19830 19840 19850 19860 19870 19880 19890 19900 19910 19920 KNSEEKAKKAMVDAARLADELRAEQDHAQTQEKLRKALEQ 19930 19940 19950 19960 19970 19980 19990 20000 20010 20020 20030 20040 Q I K E L Q V R L D E A E A N A L K G G K K A I Q K L E Q R V R E L E N E L D G

Figure 1. Part II (See p. 890 for legend.)

Alternative myosin hinge regions

											_
GAGCAGAGGAG	GCACGCCGATG	A Q K N CCCAGAAGAA	L R K	S E R I	R V K E	L S F	Q S E E	C D R K	N H E	R M Q	D L V ATCTGGTC
20170	20180	20190	20200	20210	20220	20230	20240	20250	20260	20270	2
DKL	QKI	КТҮК	RQI	EEAB	EEIA	A L N	LAKF	' R K A	QQE	LEE	AEE
20290	ACAGAAGATCA 20300	AGACATACAA	SAGGCAGATCO	GAGGAGGCTG/	AGGAAATCGC	CGCCCTCAAC	TTGGCCAAATT	CCGCAAGGC	ICAGCAGGAGC	TTGAGGAGG	CCGAGGAG
RADI	AEO	ATSK	20320 F R A	20330 K C P 1	20340 A C S V	20350	20360	20370	20380	20390	20400
CGCGCCGATCT	GGCCGAGCAGG	CCATCAGCAA	TTCCGCGCC	AGGGACGTG	CCGGTTCTGT	CGGTCGTGGT	SCCAGCCCAG	GTAAGTTA	TTGAACAATGG	CATCAAATG	CCTTCATC
20410	20420	20430	20440	20450	20460	20470	20480	20490	20500	20510	20520
ATCACTACCCI	TTAGCCCTTAA	GACCCCACAA	GACCTTACCO	CACTCAGAGA	AAAAAGTAAA	TATGAAAGCCO	CATTIGAACTI	CTCTCAATCO	GAACACTTCTG	AGCCTTCAT	TTTGCTGC
20530	20540	20550	20560	20570	20580	20590	20600	20610	20620	206 <u>30</u>	20640
										_	EXON 18
ATCAGCGCCAT	CTCCATTCACG	TTATTTCTGG	TCAGTTATTO	AATAATGTC	PATATTCTTC	CTTTCATAT	CCTCCCCTAT	GCTCTGCTT	CACTTACCAA	AACAGATCT	AAGCCTCA
20650	20660	20670	20680	20690	20700	20710	20720	20730	20740	20750	20760
GGCATCACCAT	CGAATAAAAA	TCTATATAGCO	SATCCAATTAT	GGTTCATTT	ACGATAATGA	GAGAACAAAAJ	ACCATGCAAAG	AGAATTATA	GCTTATAAGAA	TTATTATAA	ААТАААТА
20770	20780	20790	20800	20810	20820	20830	20840	20850	20860	20870	20880
20890	AATATAGATGC	ATCACATGAC	ATTCGCTAGCA	GACAAAAATO	GTATAAATTAC	GAGGCCACCC	CAATCAAAAAT	CAATGGATC	ATTCATCGAA	TACCAACCA	AGAATGCG
TTCATTTTGAC	AGAAAACCAGA	CAAAAAGCCAJ	20920 AAAAATATTC	20930	20940	20950	20960	20970 TCAATCGAA	20980	20990 'ACCACTATT	21000 8888287
21010	21020	21030	21040	21050	21060	21070	21080	21090	21100	21110	21120
AGAGAAAAAAC	AATACTATTCT	TCCACTATAT1	TATTATAAACI	GTATTTGTAC	CATGTGTACTO	GCAGCCGACAG	CGTCAAAGGG	CCGCAAGAG	GCGCTGCTGG	AGCAGTAGA	AACTTTCT
21130	21140	21150	21160	21170	21180	21190	21200	21210	21220	21230	21240
TGAAACATTTG	TAAGTGGACGT	GTTCGTGCAC	STCTTGACCCG	STICTIGTIGI	TGTCGGAACTO	CATCCCTCCC	ACTCCCGCACA	CGCACACCA	STCAGTCAATC	AACCGCAGC	TCATCTAG
TTGTCTCGCTC	ACTCGTTCGTT	ZIZ/U GTCGTTGCTCI	21280	21290	21300	21310	21320	21330	21340	21350	21360
21370	21380	21390	21400	21410	21420	21430	21440	21450	21460	21470	21480
CGAAGAGCATT	TACGTACCTCA	GTTACCGTTCT	TTCACCGCGC	CACTCGCTGG	GCACTACAAAA	ACGTGTTCAG	GAGTGCTAGT	CGTACTCGT	TATCGTTCTAT	CCCGTGCAG	TCGATCCA
21490	21500	21510	21520	21530	21540	21550	21560	21570	21580	21590	21600
									EXON 19		1
ACATGGCGACT	Thereconnece	~~~~						P	RAT	SVR	PQF
21610	21620	21630	21640	21650	21660	21670	21680	2169d	21700	21710	21720
DGLA	FPP	RFDL	APE	NEF *	+				21100	21/10	21/20
GACGGATTGGC	CTTCCCACCAA	GATTCGACCTI	GCTCCTGAAA	ACGAATTCT	AATGCCATTI	CATTTTTTA	TTTATTTTAA	ATGATATTT	CATAATGATTA	TGTTTATGA	TTTAATT
21730	21740	21750	21760	Eco RI	21780	21790	21800	21810	21820	21830	21840
TAATTTCTTAA	TTTAAAAACAA	AATAATAAAAC	TATAACAAAA	TAAATATCGA	AAACGACGGG	GAGGCAAACC	CCAACAACGC	CAGATGCAC	TAGGCAAAAA	AATAAAATC.	ATATAACA
ACAATCGATCA	Z186U CCATCGATTAG	21870 Сатасатаати	21880 Сатаастсат	21890	21900	21910	21920	21930	21940 P	oly A	21960 TTCTATT
21970	21980	21990	22000	22010	22020	22030	22040	22050	22060	22070	22080
ATATATTTTA	TACAACTTTCG	AAATGCATTC	ACCAACTAAC	TATGAAAGAA	CCATCGACTO	TCAACCGAAC	AACATGAGAA	TAATCGCGA	ACGTGAGCCC	CGTCAAAAT	GAAGGCCC
22090	22100	22110	22120	22130	22140	22150	22160	22170	22180	22190	22200
22210	22220	GTCGCCCGCAG	TTCGAATATT	TTGCTCTATA	ATGTACTCCCA	ATATTGTTTC	TCATGCAGAC	CAGACAAGAG	TTACGGACGA	CACGAGCTA	CACACTGO
GCTTATGTTTA	TGTTTAACTCT	ZZZJU SACTCTATCTZ	ACAATGATAG	22230	2226U	22210 	22260 2888678786	22290 3333336CCA1	22300 	22310 CTABATCAS	22 <u>320</u> 1777 8 - 8 - 8 - 8
22330	22340	22350	22360	22370	22380	22390	22400	22410	22420	22430 P	
ATGTGTATTAT	TTTCCGTGATC	AAATCATTCGA	CTGGGACATA	TTCATTCATA	TTTACTGGCC	CATACGTGTT	AGAAATTGTC	ACAGGTCTTT	TGTAAACTAA	TAAGATGTT	STGGATAT
22450	22460	22470	22480	22490	22500	22510	22520	22530	22540	22550	22560
B											
	* *	***	* **	* * * * * * *	*** * ** *	* * *	****** *	* * ** *	* * *****	* *****	*** **
Rat emb. MH	C LKKKDFEYS	QLQSKVEDEQI	LSLQLQKKIN	ELQARIEELE	EEEIEAER AT H	RAKTEKORSD	ARELEELSEF	LEEAGGVTS	TQIELNKKREA	EFLKLRRDL	EEATLQHE
C. elegans	LKKKESELH	SVSSRLEDEH	LVSKLQRQIK	DGQSRISELE	EEELENEROS	RSKADRAKSDI	LORELEELGEN	LDEOGGATA	AQVEVNKKREA	ELAKLRRDL	EEANMNHE
Dros. MHC	IORKDKELS.	SITAKLEDEQ	VVLKHQRQIK	ELQARIEELE	EEEVEAEROA	RAKAEKQRADI	LARELEELGEF	LEEAGGATS	AQIELNKKREA	ELSKLRRDL	EEANIQHE
*****	* ** * *	* * * * *	*	**	4S-2/H1 * **	tnge	* * *	* *	* * * *	EXON	14
ATVATLRKKHA	DSAAELAEQID	NLQRVKQKLEI	EKSEFKLEID	DLSSSVESVS	SKSKANLEKI	CRTLEDOLSE	ARGKNEETORS	LSELTTOKS	RLOTEAGELSE	10	
NQLGGLRKKHT	DAVAELTDQLD	QLNKAKAKVE I	DKAQAVRDAE	DLAAQLDQE1	SGKLNNEKL	AKOFELQLTE	LOSKADEOSRO	LODFTSLKG	RLHSENGDLVF	Q	
STLANLRKKHN	DAVAEMAEQVD	OLNKLKAKAEH	DRQICHNELN	QTRTACDOLO	GRDKAAQEKI	AKQLQHTLNE	QSKLDETNRI	LNDFDASKK	KLSIENSDLLF	RQ	
		AEK	EKNEYYGQLN	DLRAGVDHIT	NEK	EXON 16		Hinge/	LMM	-	
		16	NON 124/12	0	1						
C											
~											
			1						_		
Wild time	₩₩₩ ₩₩₩₩₩		► A E H	D R Q	T C H N	ELN	QTR	T A C E	Q L G	K D K	
	TIACCATT	GUINATUCAG	GGCIGAACAC	GATUGUUAGA	UTTGCCACAA	GAGCTGAAT	CAGACTCGTA	CUBUCTGUGA	LICAGCTGGGT	UGUGATAAGO	JIAAT
			◆ L N T	IAR	LAT	T S *				Ļ	
Mhc ¹⁰	TTTACCCATT	GCTAATCCAA	GGCTGAACAC	GATCGCCAGA	CTTGCCACAA	CGAGCTGAAT	CAGACTCGTA	CCGCCTGCGA	TCAGCTGGGT	CGCGATAAG	STAAT
			Figure	1. Part II	I (See follo	wing page i	for legend.)				
			0		, ,	0 0 - /	0/				

GENES & DEVELOPMENT 889

Collier et al.

We used the antisense RNA probes for the hingecoding regions to determine the tissue specificity of alternative exon usage via in situ hybridization to wildtype organisms. As expected, exon 15b hybridized to transcripts in all larval muscles, whereas exon 15a failed to hybridize to larval muscle (not shown). In agreement with these observations, 23 independent cDNA clones isolated from a late embryo cDNA library (Brown and Kafatos 1988) all contained exon 15b. In adults, a more complex pattern emerges (Fig. 3). Exon 15a, but not exon 15b, is utilized in IFM and jump muscles; however, both exons are used in various head and leg muscles. The failure of Mhc10 mutants to accumulate mRNAs containing exon 15a, along with the tissue-specific usage of this exon in wild-type flies, correlates well with our previous demonstration that Mhc10 pupae do not have detectable levels of MHC mRNA and protein in IFM and jump muscles and show reduced accumulation in leg muscles (O'Donnell et al. 1989).

Ultrastructural analysis of the Mhc¹⁰ mutant is a highly sensitive assay for usage of alternative MHC hinge regions

From our in situ hybridization studies, it is clear that there is a tissue-specific pattern of accumulation of MHCs with alternative hinge regions. In situ hybridization is inadequate for determining the ratio in which alternative hinge-encoding MHC transcripts are coexpressed. However, the Mhc10 strain provides an opportunity to examine the ultrastructure of muscles lacking MHC with hinge region a, and one can thereby infer the relative proportions of MHCs with each of the alternative hinge regions. Only exon 15a is used in IFM and jump muscles (Fig. 3), and these muscles in Mhc^{10} flies do not accumulate MHC transcripts, MHC protein, or thick filaments (Fig. 4B; O'Donnell et al. 1989). In contrast, muscles of the upper leg (coxa) are only partially affected by the mutation. Intracoxal levator muscles (seen in cross section in Fig. 4H) show an \sim 50% reduction in thick/thin filament ratio compared to wild type (Fig. 4E). Intracoxal depressor muscles are more severely affected, as they lack rectangular myofibrils (Fig. 4F) and have only occasional thick filaments that are abnormal in shape (Fig. 4G). Finally, the dorsal body wall muscles of the abdomen appear completely normal in the mutant (Fig. 4K). On the basis of our in situ hybridization and electron microscopic evidence, we conclude that the severity of reduction in thick filament number in the mutant muscles correlates with the use of exon 15a in the wild-type muscle.

Homozygous-viable MHC mutations that display different tissue-specific defects from Mhc¹⁰ are not within exon 15

We characterized previously three other homozygous-viable MHC mutations (Mhc⁷, Mhc⁹, and Mhc¹¹) which, like Mhc¹⁰, prevent myosin accumulation in the IFM (O'Donnell et al. 1989; see also Chun and Falkenthal 1988). Unlike Mhc¹⁰, the Mhc⁷, Mhc⁹, and Mhc¹¹ mutations do not affect MHC accumulation in leg muscles and have less severe effects on thick filament accumulation in the jump muscle. The different tissue specificities of the Mhc⁷, Mhc⁹, and Mhc¹¹ mutations suggested that they should not affect exon 15a. We cloned and sequenced exons 15a and 15b and their surrounding introns from these mutants and failed to find any differences from the wild-type sequence. These results indicate that the genetic lesions in the Mhc⁷, Mhc⁹, and Mhc^{11} alleles likely arise from mutations in alternative exons other than 15a and 15b.

Discussion

Our results indicate that there is remarkable complexity to the accumulation of transcripts encoding MHCs with alternative hinge regions in *Drosophila*. Embryonic and larval muscles as well as abdominal body wall muscles of the adult use only the hinge encoded by exon 15b. Adult leg muscles and the proboscis musculature apparently accumulate some MHCs containing each hinge

Figure 1. (A) DNA sequence analysis of the rod-coding region of the Drosophila muscle MHC gene. The nucleotide numbering system corresponds to that used previously (Wassenberg et al. 1987), and exon numbers are assigned on the basis of sequencing of the entire MHC gene (George et al. 1989; W. Kronert and S. Bernstein, unpubl.). Sequencing of the rod-coding region of wild-type genomic DNA and analysis of several cDNA clones revealed only a single set of alternative exons (15a and 15b). Exon/intron boundaries were determined by their homology to consensus splice sequences (Mount 1982), by comparison of the encoded amino acid sequence to MHCs of other organisms (Karn et al. 1983; Strehler et al. 1986), and by partial sequencing of a larval and a pupal cDNA clone. The borders of the hinge regions are denoted, and the amino acids that correspond to the proteolytic cleavage sites in vertebrate MHC protein (Strehler et al. 1986) are shown in boldface type. The penultimate exon of the gene (exon 18) is included in some adult transcripts, where it encodes a single carboxy-terminal amino acid. This exon is omitted from other MHC mRNAs, in which case exon 19 encodes the carboxyl terminus of the protein (see Fig. 2A; Bernstein et al. 1986; Rozek and Davidson 1986). (B) Comparison of the amino-acid-coding potential of Drosophila, rat embryonic skeletal muscle (Strehler et al. 1986), and C. elegans unc-54 MHC (Karn et al. 1983) hinge regions. The region encoded by the alternatively spliced Drosophila exons is embedded in the central portion of the hinge. Asterisks (*) indicate when both the rat and nematode sequences contain the same amino acid as at least one of the Drosophila MHCs. (C) Sequence comparison of exon 15a from wild type and the flightless mutant Mhc^{10} . The single-nucleotide change (G \rightarrow A) that eliminates the consensus 3' splice junction (PyAG) is underlined. This also creates a nonconsensus 3' splice junction (AAG) that may be used for splicing (see text). As shown, this would result in the production of a truncated protein due to shifting of the translational reading frame. Splice junctions are indicated by arrows.

Alternative myosin hinge regions



Figure 2. Hybridization of an RNA gel blot with probes specific to alternative exons 15a and 15b. (A) Alternative splicing and polyadenylation patterns that give rise to the various size classes of MHC mRNA. Only the 3' end of the gene is shown; alternative inclusion of the penultimate exon (exon 18) and polyadenylation at pA1 or pA2 result in four transcript classes. Alternative splicing of upstream exons does not affect the size of the mRNAs (Wassenberg et al. 1987; George et al. 1989; Hess et al. 1989). (B) Hybridization of the exon 15b probe reveals transcripts of 6.1 and 6.6 kb in wild-type larval (lane 1) and pupal (lane 3) RNA, as well as in Mhc^{10} larval (lane 2) and pupal (lane 4) RNA. (C) Hybridization of the exon 15a probe to the same blot shown in B (following removal of the exon 15b probe) reveals MHC transcripts of 6.6 and 7.1 kb in the wild-type pupal RNA (lane 3) but no hybridization to larval RNA (lanes 1 and 2) or to RNA from Mhc^{10} pupae (lane 4). On longer exposures of this blot, a very minor amount of the 6.6- and 7.1-kb transcripts in the Mhc^{10} pupal lane could be observed (not shown).

type. Ultrastructural analysis of two different types of leg muscles in the Mhc^{10} mutant reveals that the ratio of the two MHCs can vary dramatically between such muscles. At the other end of the spectrum, the highly specialized IFM and jump muscles use solely exon 15a. In agreement with our results, George et al. (1989) showed by S1 nuclease mapping that only exon 15a is included in IFM transcripts, whereas exon 15b is predominant in larvae and adult abdomen. Although it is possible that the tissue-specific inclusion of exon 15a or 15b is serendipitous, we believe it is more likely that the alternative hinge regions encoded by these exons directly affect the functional properties of the MHC molecule.

The MHC rod has an α -helical structure that permits the interaction of MHC molecules to form dimers and the interaction of MHC dimers to form thick filaments (McLachlan and Karn 1982). The two major helixbreaking domains of the rod are the hinge and the carboxy-terminal tailpiece. Interestingly, our results and those published recently by George et al. (1989) indicate that these are the only regions of the Drosophila MHC rod that are encoded by alternatively spliced exons. On the basis of our Northern blots and in situ hybridizations of both wild-type and Mhc10 organisms, as well as previous studies (Bernstein et al. 1986; George et al. 1989; Kazzaz and Rozek 1989), it appears that the use of exon 15a generally correlates with the inclusion of the 3' penultimate exon of the MHC gene (which encodes an alternative tailpiece). Thus, the hinge and tailpiece may act in concert with each other and, perhaps, with regions of the globular head encoded by alternative exons (Wassenberg et al. 1987; George et al. 1989; Hess et al. 1989), to produce functional differences between MHC isoforms. Our results rule out the possibility that regions of the rod other than the hinge impart isoform-specific properties to the MHC molecule. This conclusion may apply to MHC isoforms encoded by multigene families as well.

Heterogeneity in the MHC rod or tailpiece might be expected to affect assembly properties of myosin molecules (Kuczmarski and Spudich 1980; McLachlan and Karn 1982; Kiehart et al. 1984), the morphology of thick filaments, and the number of thin filaments that assemble around each thick filament (Hayashi et al. 1983) and, perhaps, force generation (Ueno and Harrington 1986). Thin filament orbital number correlates roughly with use of exon 15b, it being lowest in IFM (which never utilizes exon 15b) and greatest in larval and abdominal body wall muscles (which use only exon 15b). Hayashi et al. (1983) found that thick filaments assemble in vitro with specific numbers of thin filaments, depending on the type of myosin included in the thick filaments. Our results suggest that the MHC hinge and/ or carboxyl terminus may be implicated in this process. Another correlation we have observed is that muscles required to contract quickly (IFM) and/or generate high levels of tension (jump muscle) accumulate solely transcripts encoding hinge region a, whereas muscles that contract slowly (larval muscles and abdominal muscles of the adult) only accumulate hinge region b-coding transcripts. Muscles of an intermediate contraction speed (leg and proboscis muscles) contain some transcripts of each type. Alternative myosin hinges may



Figure 3. In situ hybridization of radiolabeled RNA probes to cryosections of wild-type *Drosophila* pupae. Anterior is to *left* in A-E. (A) A control (sense) probe transcribed from the MHC gene (exons 4-6; see O'Donnell et al. 1989) and hybridized to parasagittal section does not result in silver grain accumulation over muscle tissues. (IFM) Indirect flight muscle; (P) proboscis muscles; (L) leg tissues; (A) antennal muscles. (B) Exon 15a antisense probe hybridizes to IFM, some proboscis muscles, and leg muscles. (C) Dorsal view of pupa hybridized with exon 15a antisense probe shows grains over IFM and jump muscle [(TDT) tergal depressor of the trochanter muscle]. (D) Exon 15b antisense probe fails to hybridize to IFM but clearly labels proboscis muscles. (E) Exon 15b probe hybridizes to proboscis (P), antennal (A), and leg muscles (L) but not to IFM. (A, B, and D) Bar, 440 μ_i (C and E) bar, 250 μ .

therefore play a role in the generation of different levels of contractile force or speed. Indeed, Harrington and colleagues (Harrington and Rodgers 1984; Ueno and Harrington 1986) argue that force generation is a result of conformational changes within the hinge. However, Kishino and Yanagida (1988) and Spudich and colleagues (Hynes et al. 1987; Toyoshima et al. 1987) observe myosin force production or movement in vitro in the

Figure 4. Ultrastructures of muscles from wild-type and Mhc^{10} adults reveal the pattern of exon 15a usage. Reduced numbers of thick filaments result from failure to correctly splice exon 15a in the mutant. Wild-type IFM (A) displays a highly ordered arrangement of six thin filaments surrounding each thick filament. In the mutant (B), no normal thick filaments are observed. In C, the two muscle groups of the coxal segment of the leg, the intracoxal levators (IL), and the intracoxal depressors (ID) are indicated (Crossley 1978). The wild-type intracoxal depressor is shown in oblique section (D) and has a regular myofilament lattice. This muscle in Mhc^{10} legs (F) contains extremely disorganized myofibers whose myofilaments are in complete disarray. At higher magnification, thick filaments are seen only occasionally (G), and these are abnormal in shape. Z disc material (Z) is poorly organized. The mutant intracoxal levator muscle (H) shows an ~50% decrease in the thick/thin filament ratio, with some disruption of the thick filament linear array seen in wild type (E). In I, the dorsal body wall muscles (DM) of the adult abdomen are shown. At higher magnification, wild-type (I) and Mhc^{10} (K) dorsal body wall muscles have identical myofilament arrays. Glycogen granules (G), mitochondria (M), and sarcoplasmic reticulum (S) are indicated. Bars for electron micrographs, 0.2 μ ; bars for light micrographs (C and I), 16 μ .

Alternative myosin hinge regions

absence of the hinge-containing rod region. These experiments, although clearly demonstrating that the S1 globular head of MHC is capable of generating movement and force, do not rule out the possibility that the hinge modulates these properties directly or by changing the structure of the thick filament and/or the geometry of the interaction between thick and thin filaments. The presence of alternative hinge regions in *Drosophila*



Figure 4. (See facing page for legend.)

Collier et al.

MHC and the general correlation with contraction speed and/or tension development suggest that the hinge may indeed have an effect on muscle contraction.

Use of the Mhc^{10} null mutant in conjunction with MHC gene/cDNA fusions introduced by P-element-mediated germ line transformation should permit the analysis of IFM and jump muscles forced to express hinge b. Subsequent mechanical measurements of contraction speed and force would indicate the effect of the alternative hinges on these parameters. Likewise, ultrastructural examination would show whether the morphology of the thick filament or the thin filament orbital number is affected by the hinge region of the MHC molecule. This in vitro mutagenesis/gene transformation approach should therefore be a useful and unique way of analyzing MHC domain function in vivo.

Materials and methods

Isolation and sequencing of genomic and cDNA clones

Wild-type genomic MHC DNA (Bernstein et al. 1983) was subcloned into plasmid vectors prior to sequence determination. BamHI-digested DNA from the homozygous Mhc^{10} strain was cloned into λ vector EMBL3 (Stratagene) and packaged into phage heads by use of the Gigapack extract (Stratagene). The recombinant library was screened by use of a nick-translated probe containing the 3' end of the MHC gene. A positive phage containing a 16-kb BamHI fragment that encompasses the 3' half of the MHC gene was isolated. Sequences were determined on plasmid subclones of genomic MHC DNA by the dideoxynucleotide chain-termination method (Hattori and Sakaki 1986) with synthetic oligonucleotide primers. Late embryonic cDNA clones were obtained by screening a cDNA plasmid library (Brown and Kafatos 1988), using a nick-translated MHC gene probe.

RNA isolation, probe preparation, and hybridization

RNA was isolated from second and third-instar larvae or from late-stage pupae (following eye pigment deposition) by homogenization of 50 organisms in 1 ml of freshly prepared 6 M urea and 3 M LiCl and incubation overnight on ice. RNA was collected by centrifugation and then resuspended in 300 μ l of 10 mM Tris-HCl (pH 7.5), 10 mM EDTA, and 1% SDS. The sample was phenol/chloroform-extracted twice and extracted once with chloroform. RNA was precipitated with ethanol, pelleted by centrifugation, and resuspended in 25 μ l of sterile distilled water. Ten micrograms of each RNA sample was electrophoresed and blotted as described previously (Zyskind and Bernstein 1989). Prior to blotting, gels were stained with ethidium bromide and observed under ultraviolet light to ensure that equivalent amounts of RNA had been loaded.

Northern blot hybridization probes were ${}^{32}P$ -labeled antisense RNAs, prepared as described by Melton et al. (1984). To construct the exon 15a- and 15b-containing vectors for exonspecific probe preparation, the Bg/II-HindIII fragment encompassing both exons (see Fig. 1A) was cloned into the BamHI and HindIII sites of the pKS vector (Stratagene). By digesting with PvuII to linearize the plasmid and employing T7 RNA polymerase for transcription, an exon 15b-specific probe could be produced. An exon 15a clone was made by inserting the Bg/II-Aval fragment (see Fig. 1A) into the BamHI site and SmaI site of pKS (the AvaI site was first treated with S1 nuclease to create blunt ends). The exon 15a-specific probe was made by use of T7 RNA polymerase, following linearization in the polylinker with *Eco*RI. Northern blot hybridization and washing were essentially as described by DeLeon et al. (1983). Probes labeled with ³⁵S were prepared for in situ hybridization, which was performed as described previously (O'Donnell et al. 1989).

Light and electron microscopy

Fixation, embedding, sectioning, and staining of tissues were done as described previously (O'Donnell and Bernstein 1988; O'Donnell et al. 1989).

Acknowledgments

We thank Roger Sabbadini, Dianne Hodges, and Phillip Singer for their critical comments concerning the manuscript, Nicholas Brown for providing the late embryo cDNA library, Eric Fyrberg, Charles Emerson, and Anne Glenn for providing pupal MHC cDNA clones, and Elizabeth George, Margaret Ober, and Charles Emerson for sharing their manuscript with us prior to publication. Our research was supported by grants from the National Institutes of Health (GM-32443) and the Muscular Dystrophy Association. S.I.B. is an Established Investigator of the American Heart Association.

Note added in proof

Sequence data described in this paper have been submitted to the EMBL/GenBank Data Libraries under accession number X53155.

References

- Baserga, S.J. and E.J. Benz Jr. 1988. Nonsense mutations in the human β -globin gene affect mRNA metabolism. *Proc. Natl.* Acad. Sci. 85: 2056–2060.
- Bernstein, S.I., K. Mogami, J.J. Donady, and C.P. Emerson Jr. 1983. Drosophila muscle myosin heavy chain encoded by a single gene in a cluster of muscle mutations. Nature 302: 393-397.
- Bernstein, S.I., C.J. Hansen, K.D. Becker, D.R. Wassenberg II, E.S. Roche, J.J. Donady, and C.P. Emerson Jr. 1986. Alternative RNA splicing generates transcripts encoding a thoraxspecific isoform of *Drosophila melanogaster* myosin heavy chain. *Mol. Cell. Biol.* 6: 2511-2519.
- Brown, N.H. and F.C. Kafatos. 1988. Functional cDNA libraries from *Drosophila* embryos. J. Mol. Biol. 203: 425-437.
- Chun, M. and S. Falkenthal. 1988. Ifm(2)2 is a myosin heavy chain allele that disrupts myofibrillar assembly in the indirect flight muscle of Drosophila melanogaster. J. Cell Biol. 107: 2613-2621.
- Crossley, A.C. 1978. The morphology and development of the Drosophila muscular system. In The genetics and biology of Drosophila (ed. M. Ashburner and T.R.F. Wright), vol. 2b, pp. 499-560. Academic Press, London.
- DeLeon, D.V., K.H. Cox, L.M. Angerer, and R.C. Angerer. 1983. Most early-variant histone mRNA is contained in the pronucleus of sea urchin eggs. *Dev. Biol.* 100: 197-206.
- Dibb, N.J., D.M. Brown, J.Karn, D.G. Moerman, S.L. Bolten, and R.H. Waterston. 1985. Sequence analysis of mutations that affect the synthesis, assembly and enzymatic activity of the unc-54 myosin heavy chain of *Caenorhabditis elegans*. J. Mol. Biol. 183: 543-551.
- Emerson, C.P., Jr. and S.I. Bernstein. 1987. Molecular genetics of myosin. Annu. Rev. Biochem. 56: 695-726.
- George, E.L., M.B. Ober, and C.P. Emerson Jr. 1989. Functional

Alternative myosin hinge regions

domains of the *Drosophila* muscle myosin heavy chain gene are encoded by alternatively spliced exons. *Mol. Cell. Biol.* **9**: 2957–2974.

- Harrington, W.F. and M.E. Rodgers. 1984. Myosin. Annu. Rev. Biochem. 53: 35-73.
- Hattori, M. and Y. Sakaki. 1986. Dideoxy sequencing method using denatured plasmid templates. *Anal. Biochem.* 152: 232-238.
- Hayashi, T., P.M. Wozniak, M.L. Cayer, and D.S. Smith. 1983. Actin-myosin interaction: the role of myosin in determining the actin pattern in self-assembled "hybrid" contractile units. *Tissue Cell* **15**: 955–963.
- Hess, N., W.A. Kronert, and S.I. Bernstein. 1989. Transcriptional and post-transcriptional regulation of *Drosophila* myosin heavy chain gene expression. In *Cellular and molecular biology of muscle development.* (ed. L.H. Kedes and F.E. Stockdale), pp. 621-631. A.R. Liss, New York.
- Hynes, T.R., S.M. Block, B.T. White, and J.A. Spudich. 1987. Movement of myosin fragments in vitro: domains involved in force production. *Cell* **48**: 953-963.
- Karn, J., S. Brenner, and L. Barnett. 1983. Protein structural domains in the *Caenorhabditis elegans* unc-54 myosin heavy chain gene are not separated by introns. *Proc. Natl. Acad. Sci.* 80: 4253-4257.
- Kazzaz, J.A. and C.E. Rozek. 1989. Tissue-specific expression of the alternately processed *Drosophila* myosin heavy-chain messenger RNAs. *Dev. Biol.* 133: 550-561.
- Kiehart, D.P., D.A. Kaiser, and T.D. Pollard. 1984. Direct localization of monoclonal antibody-binding sites on Acanthamoeba myosin-II and inhibition of filament formation by antibodies that bind to specific sites on the myosin-II tail. *I.* Cell. Biol. 99: 1015-1023.
- Kishino, A. and T. Yanagida. 1988. Force measurements by micromanipulation of a single actin filament by glass needles. *Nature* **334**: 74–76.
- Kuczmarski, E.R. and J.A. Spudich. 1980. Regulation of myosin self-assembly: Phosphorylation of *Dictyostelium* heavy chain subunits inhibits formation of thick filaments. *Proc. Natl. Acad. Sci.* 77: 7292-7296.
- McLachlan, A.D. and J. Karn. 1982. Periodic charge distributions in the myosin rod amino acid sequence match crossbridge spacings in muscle. *Nature* **299**: 226-231.
- Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. 12: 7035-7056.
- Mogami, K. and Y. Hotta. 1981. Isolation of *Drosophila* flightless mutants which affect myofibrillar proteins of indirect flight muscle. *Mol. Gen. Genet.* **183**: 409-417.
- Mogami, K., P.T. O'Donnell, S.I. Bernstein, T.R.F. Wright, and C.P. Emerson, Jr. 1986. Mutations of the Drosophila myosin heavy-chain gene: effects on transcription, myosin accumulation, and muscle function. Proc. Natl. Acad. Sci. 83: 1393-1397.
- Mount, S.M. 1982. A catalogue of splice junction sequences. Nucleic Acids Res. 10: 459-472.
- O'Donnell, P.T. and S.I. Bernstein. 1988. Molecular and ultrastructural defects in a *Drosophila* myosin heavy chain mutant: Differential effects on muscle function produced by similar thick filament abnormalities. *J. Cell. Biol.* 107: 2601-2612.
- O'Donnell, P.T., V.L. Collier, K. Mogami, and S.I. Bernstein. 1989. Ultrastructural and molecular analyses of homozygous-viable *Drosophila melanogaster* muscle mutants indicate there is a complex pattern of myosin heavy-chain isoform distribution. *Genes Dev.* **3:** 1233-1246.

- Rozek, C.E. and N. Davidson. 1986. Differential processing of RNA transcribed from the single-copy *Drosophila* myosin heavy chain gene produces four mRNAs that encode two polypeptides. *Proc. Natl. Acad. Sci.* 83: 2128-2132.
- Smith, C.W.J., E.B. Porro, J.G. Patton, and B. Nadal-Ginard. 1989. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. Nature 342: 243-247.
- Strehler, E.E., M.-A. Strehler-Page, J.-C. Perriard, M. Periasamy, and B. Nadal-Ginard. 1986. Complete nucleotide and encoded amino acid sequence of a mammalian myosin heavy chain gene. J. Mol. Biol. 190: 291-317.
- Toyoshima, Y.Y., S.J. Kron, E.M. McNally, K.R. Niebling, C. Toyoshima, and J.A. Spudich. 1987. Myosin subfragment-1 is sufficient to move actin filaments in vitro. *Nature* **328**: 536-539.
- Ueno, H. and W.F. Harrington. 1986. Local melting in the subfragment-2 region of myosin in activated muscle and its correlation with contractile force. J. Mol. Biol. 190: 69-82.
- Wassenberg II, D.R., W.A. Kronert, P.T. O'Donnell, and S.I. Bernstein. 1987. Analysis of the 5' end of the Drosophila muscle myosin heavy chain gene: alternatively spliced transcripts initiate at a single site and intron locations are conserved compared to myosin genes of other organisms. J. Biol. Chem. 262: 10741-10747.
- Zyskind, J.W. and S.I. Bernstein. 1989. *Recombinant DNA laboratory manual*. Academic Press, San Diego.



Alternative myosin hinge regions are utilized in a tissue-specific fashion that correlates with muscle contraction speed.

V L Collier, W A Kronert, P T O'Donnell, et al.

Genes Dev. 1990, 4: Access the most recent version at doi:10.1101/gad.4.6.885

References	This article cites 31 articles, 12 of which can be accessed free at: http://genesdev.cshlp.org/content/4/6/885.full.html#ref-list-1
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.



Streamline your research with Horizon Discovery's ASO tool