Characterization and in Vivo Functional Analysis of Splice Variants of Cypher*

Chengqun Huang‡, Qiang Zhou‡, Peihua Liang‡, Melinda S. Hollander‡, Farah Sheikh‡‡, Xiaodong Li‡, Marion Greaser‡, G. Diane Shelton¶, Sylvia Evans‡, and Ju Chen‡§**

From the Institute of Molecular Medicine and Department of Medicine, University of California at San Diego, School of Medicine, La Jolla, California 92039-0641, §Muscle Biology Laboratory, University of Wisconsin, Madison, Wisconsin 53706, and the Department of Pathology, University of California at San Diego, School of Medicine, La Jolla, California 92039-0613

Previously, we reported two splice variants of Cypher, a striated muscle-specific PDZ-LIM domain protein, Cypher1 and Cypher2. We have now characterized four additional splice isoforms, two of which are novel. The six isoforms can be divided into skeletal or cardiac specific classes, based on the inclusion of skeletal or cardiac specific domains. Short and long isoforms share an N-terminal PDZ domain, but the three C-terminal LIM domains are unique to long isoforms. By RNA and protein analysis, we have demonstrated that Cypher isoforms are developmentally regulated in both skeletal and cardiac muscle. We have previously shown that knockout of Cypher is neonatal lethal. To investigate the function of splice variants in vivo, we have performed a rescue experiment of the Cypher null mutant by replacing the endogenous Cypher gene with cDNAs encoding either a short or long skeletal muscle isoform. In contrast to Cypher null mice, a percentage of mice that express only a short or long skeletal muscle-specific isoform can survive to at least 1 year of age. Although surviving mice exhibit muscle pathology, these results suggest that either isoform is sufficient to rescue the lethality associated with the absence of Cypher.

The human genome project has estimated that there are only 32,000 genes in the human genome, far fewer than anticipated (1, 2). This relative paucity of genes may be compensated for by alternative splicing to yield more protein products per gene as substantiated by bioinformatics screening, which suggests that 35–60% of human genes have at least one alternative splice form (3). These observations suggest that to understand the functions of an alternatively spliced gene, it will be important to characterize and understand the diverse biological roles of its specific protein isoforms. To date, only several hundred alternatively spliced genes have been identified by molecular biological approaches. Very few of them have been characterized at a functional level (3, 4).

During our previous studies of a striated muscle-specific PDZ-LIM domain protein, Cypher, we reported two splice variants that were expressed to significant levels, designated Cypher1 and Cypher2 (5). The longer isoform, Cypher1, contains an amino-terminal PDZ and three C-terminal LIM domains. The shorter, Cypher2, shares only the PDZ domain (5, 6). Both PDZ and LIM domains are believed to be important in protein-protein interactions (7–9). Cypher1 and Cypher2 interact with α-actinin 2 through their PDZ domain, whereas Cypher1 interacts with distinct protein kinase C isoforms through its LIM domains (5).

The relatively high levels of expression of each isoform and their distinct biochemical functions suggest that each may have distinct as well as overlapping biological roles. To investigate this further, we have undertaken a detailed characterization of Cypher isoforms in skeletal and cardiac muscle. In these studies, we have characterized four additional Cypher splice isoforms, two of them novel, to bring the total of Cypher splice isoforms in mouse to six. The six isoforms can be divided into two classes based on the inclusion of a domain that is specific to either skeletal or cardiac muscle. We have also found that these isoforms are exquisitely regulated both at RNA and protein levels during successive stages of embryonic and postnatal development.

Our previous studies demonstrated that Cypher plays an essential role in striated muscle structure and function (6). Mice which are homozygous null for Cypher die within 1–5 days after birth from multiple striated muscle failure with symptoms that include decreased milk intake, limb muscle weakness, and cyanosis. Transmission electron microscopy studies revealed that skeletal and cardiac muscle from mutant mice is severely disorganized with discontinuous/punctuate Z-discs (6). To understand the requirement for short or long Cypher isoforms in striated muscle function, we have performed a rescue experiment for Cypher null mice by genetically engineering mice that express only a short or long skeletal muscle-specific isoform rather than the six isoforms normally expressed by the endogenous gene. Results of this analysis suggest that either isoform can partially rescue the lethality associated with the absence of Cypher.

EXPERIMENTAL PROCEDURES

RNA Analyses—Total RNA was isolated from mouse cardiac ventricle and gastrocnemius muscle using RNAzol B (Tel-Test). 10 μg of RNA was electrophoresed on a 1% agarose gel, blotted, and hybridized with α-32P-labeled probe in QuickHyb solution as described by suggested protocols. Reverse transcription-PCR (RT-PCR) was performed using

1 The abbreviations used are: RT-PCR, reverse transcription-PCR; aa, amino acids; nt, nucleotide(s).
Cypher Splice Isoforms

Fig. 1. Schematic representation of cypher genomic structure and corresponding Cypher splice variants. The structure of the cypher gene is schematically represented, with the 17 coding exons shown as boxes and introns as lines between them. Cypher protein isoforms are shown below, color-coded to illustrate the genomic origin of each protein domain. Numbers above each protein represent the amino acid found at the beginning of each domain. Major structural features are noted, including the PDZ domain, cardiac (C) and skeletal (S) specific regions, and LIM domains. Note that the skeletal domain in Cypher1s is 5 amino acids shorter than that in the other two skeletal isoforms. Sizes of each exon are as follows. Exon 1, 31 aa; exon 2, 50 aa + 2 nt; exon 3, 25 aa + 1 nt; exon 4, 119 aa + 2 nt; exon 5, 7 aa + 2 nt; exon 6, 68 aa; exon 7, 5 aa; exon 8, 56 aa + 2 nt; exon 9, 12 aa + 1 nt; exon 10, 31 aa + 1 nt + TAA; exon 11, 62 aa; exon 12, 48 aa + 2 nt; exon 13, 148 aa + 1 nt; exon 14, 60 aa + 1 nt; exon 15, 40 aa + 1 nt; exon 16, 38 aa + 2 nt; exon 17, 29 aa + TAA.

RESULTS

Cloning Cypher Isoforms—We had shown previously that Cypher has two distinct splice variants, Cypher1 and Cypher2. Western blot analysis of proteins from skeletal and cardiac muscle utilizing antibody specific for Cypher1 revealed that Cypher1 in cardiac muscle was ~6 kDa larger than Cypher1 in skeletal muscle (data not shown).

Our original cypher cDNA clone had been obtained from a heart library. To investigate the observed size difference between cardiac and skeletal Cypher1 proteins, RT-PCR was performed with RNAs from skeletal and cardiac muscle utilizing a set of primers that would amplify the coding region of Cypher1 (see “Experimental Procedures”). The resulting PCR products were subcloned, and 10 clones from each striated mRNA source were sequenced. 8 of 10 clones from heart were identical to Cypher1. The other two clones were identical to Cypher1 with the exception that DNA sequences encoding amino acids 296–357 of Cypher1 were missing (see Fig. 1, Cypher1c and -3c). On the other hand, 10 out of 10 clones from skeletal muscle differed from Cypher1 by the absence of DNA sequences encoding amino acids 108–226 of Cypher1c. In nine cases, this region of Cypher1 was replaced by DNA sequences identical to Cypher1. The other two clones were identical to Cypher1 with the exception that DNA sequences encoding amino acids 296–357 of Cypher1 were missing (see Fig. 1, Cypher1c and -3c). On the other hand, 10 out of 10 clones from skeletal muscle differed from Cypher1 by the absence of DNA sequences encoding amino acids 108–226 of Cypher1c. In nine cases, this region of Cypher1 was replaced by DNA sequences identical to Cypher1. The other two clones were identical to Cypher1 with the exception that DNA sequences encoding amino acids 296–357 of Cypher1 were missing (see Fig. 1, Cypher1c and -3c).
of Cypher1 and DNA sequences encoding 108–187 of Cypher2 were cardiac and skeletal muscle-specific domains, respectively.

We then performed Northern blot analyses utilizing cDNA probes encoding cardiac and skeletal muscle-specific domains. In concert with the RT-PCR results, the probe encoding amino acids 108–226 of Cypher1, gave a positive signal only in cardiac tissue (Fig. 2A). Conversely, the probe encoding amino acids 108–187 of Cypher2 gave a positive signal only in skeletal muscle (Fig. 2B).

Following these results, we have designated Cypher isoforms as either c (cardiac) or s (skeletal muscle) (Fig. 1). We have renamed the original Cypher1 and Cypher2 as Cypher1c and Cypher2c, respectively. The isoform that is the cardiac counterpart of Cypher2s is Cypher2c. The isoform that is the skeletal muscle counterpart of Cypher1c is Cypher1s. The cardiac isoform lacking amino acids 296–357 of Cypher1c is Cypher3c. Its skeletal counterpart is Cypher3s (Fig. 1).

9 of 10 clones from skeletal muscle were Cypher3s, whereas 8 of 10 clones from cardiac tissue were Cypher1c, suggesting preferential expression of each of these in the respective striated muscle tissue. To confirm this result, Northern blot analysis was performed with a probe specific for Cypher1 isoforms. As shown in Fig. 2C, the probe gave a strong signal in the heart, consistent with previous RT-PCR results. After prolonged exposure, a faint band for Cypher1 was detected in skeletal muscle (data not shown).

We next performed Western blot analyses with antibodies raised against the cardiac specific domain or the skeletal muscle-specific domain (Fig. 1). The cardiac specific antibody detected protein in heart extracts but not skeletal muscle (Fig. 2D). The reverse was true for the skeletal muscle-specific antibody (Fig. 2E). These results were consistent with our RNA data.

A blastp search of the NCBI data base (12) demonstrated that Cypher1c is equivalent to Oracle1 (13) and an ortholog of human KIAA0613; Cypher3c is equivalent to Oracle2; and Cypher2s and Cypher3s sequences correspond to ZASP and ZASP variant 3, respectively, with the exception that the mouse isoforms include a 5-amino acid exon at the C terminus of the skeletal muscle-specific domain that is missing in their human counterparts (see below and Fig. 1 for exon structure). Cypher1s and Cypher2c are novel isoforms.

Genomic DNA Organization of cypher—To understand how distinct Cypher splice isoforms are generated from the cypher gene, we obtained cypher genomic DNA by screening a BAC genomic DNA library (Resgen, Huntsville, AL), utilizing cypher cDNA as a probe. The exon-intron junctions of cypher were determined by PCR analyses of the BAC clones with primers localized at distinct regions of cypher followed by sequencing analyses. The result is shown in Fig. 1. Cypher contains a total of 17 coding exons. Interestingly, the cardiac specific region is encoded by a single exon, exon 4, whereas the skeletal muscle-specific region of cypher is encoded by three exons, exons 5–7. However, in the skeletal isoforms, exon 7 is only included in Cypher2s and -3s and is differentially spliced out in Cypher1s. Exon 10 is the last exon of both cypher2 isoforms. Exon 11 is differentially spliced to generate Cypher1 (included) or Cypher3 (excluded) isoforms.

Cypher Isoforms Are Developmentally Regulated—Splice isoforms of other striated muscle genes are developmentally regulated (14–16). To investigate whether Cypher isoforms are also developmentally regulated, we performed both Northern and Western blot analyses. Northern blot analysis of mRNAs from heart and skeletal muscle demonstrated that expression of cypher2 mRNAs is dramatically up-regulated postnatally (Fig. 3, A and B). To assess mRNA expression of the longer isoforms cypher1 and -3 in heart and skeletal muscle, RT-PCR analysis was carried out, utilizing a set of primers within exons 8 and 13, bracketing exon 11, which is specific to cypher1 isoforms (Fig. 1). In cardiac muscle, although expression of cypher3c is up-regulated developmentally, cypher1c is the predominant long isoform expressed throughout development into adulthood (Fig. 3C). However, in skeletal muscle, expression of cypher1s is gradually replaced by expression of cypher3s postnatally. Western blot analysis utilizing Cypher antibodies specific to cardiac and skeletal muscle isoforms was consistent with RNA analyses (Fig. 3, E and F).

Cypher2s or -3s Alone Can Partially Rescue the Cypher Null Mutant—To investigate the function of short versus long Cypher isoforms, we performed a rescue experiment by creating two mouse lines in which cDNAs encoding either Cypher2s (Cypher2sKI) or Cypher3s (Cypher3sKI) were knocked in to replace the endogenous cypher gene. Previously, to generate our Cypher knockout, we had knocked lacZ cDNA into the endogenous cypher locus. This knock-in disrupted expression of all Cypher isoforms. Additionally, expression of lacZ recapitulated expression of the endogenous cypher gene. Therefore, cypher2s or cypher3s cDNAs were inserted where the lacZ cDNA had been inserted, disrupting the endogenous cypher gene (Fig. 4) (6).

Mice that were heterozygous for Cypher2sKI or Cypher3sKI had no apparent phenotype and were crossed to generate mice that were homozygous for Cypher2sKI or Cypher3sKI. Western blot analysis demonstrated that Cypher2s or Cypher3s were the only Cypher isoforms expressed in mice that were homozygous for Cypher2sKI or Cypher3sKI, respectively (Fig. 4, D and E).

Genotyping of 294 mice from 35 independent Cypher2sKI
heterozygous crosses at weaning (4 weeks) indicated that
~49% of the expected number were homozygous for
Cypher2sKI (Table I). Analysis of 230 mice from 32 independ-
ent Cypher3sKI heterozygous crosses at weaning indicated
that ~19% of the expected number were homozygous for
Cypher3sKI. Genotyping of 58 newborns from seven independ-
ent Cypher2sKI heterozygous crosses and 66 newborns from
nine independent Cypher3sKI heterozygous crosses revealed
that homozygous Cypher2sKI or homozygous Cypher3sKI oc-
curred at the expected frequency of 25%. These data indicated
that ~50% of Cypher2sKI and 80% of Cypher3sKI homozygous
mice die between birth and 4 weeks of age. Since no dead
animals were found after 1 week, labile homozygotes were
dying within the first week following birth. Cypher null mice
also die within the first week of birth, and homozygous
knock-in mice that die during this time period exhibited muscle
phenotypes comparable with those previously seen with
Cypher null mice (data not shown).

**FIG. 3.** Detection of cardiac and skeletal muscle-specific isoforms of Cypher during mouse development by RNA and protein analysis. Total RNA or protein was extracted and ana-
lyzed from embryonic day 16.5 (E16.5), 1-day-old (D1), and 14-day-old (W2) neo-
mates and 2-month-old adult (M2) mouse hearts (A, C, and B) or gastrocnemius
skeletal muscle (B, D, and F). A and B, RNA (10 μg) samples were electrophore-
sed, blotted, and probed with the PDZ domain of cypher. RNA bands of 5 and 2
kb were detected in heart (A) and skeletal muscle (B), representative of Cypher1/3
and Cypher2, respectively. 28 and 18 S RNA bands in each sample were visual-
ized by ethidium bromide and photographed before blotting as shown (bottom
panel of A and B). C and D, RNA samples were subjected to RT-PCR analysis using
primers that distinguish Cypher1c from Cypher3c (C) and Cypher1s from
Cypher3s (D) (see “Experimental Procedures”). E and F, Western blot analysis
was performed with Cypher antibodies specific for cardiac isoforms (E) or skeletal
isoforms (F). Protein bands of 78, 72, and 32 kDa and 72, 70, and 29 kDa, corre-
sponding to Cypher1, -3, and -2, respectively, were detected in heart (E) and
skeletal muscle (F).

**FIG. 4.** Targeted generation of Cypher2/3s knock-in mice. A, targeting strategy. A restriction map of the relevant genomic region of Cypher (top panel), targeting constructs for Cypher2/3s (middle panel), and the mutated locus after recombination (lower panel) are shown. The
arrows indicate the orientation of the Cypher2/3s and neomycin resistance gene.
B, BamHI; C, ClaI; E, EcoRI; H, HindIII; P, PstI; S, SstI; Sa, SalI; X, XbaI. neo, neomycin resistance gene. B, detection of wild type and targeted
alleles by Southern blot analysis. DNAs from electroporated Cypher2s and -3s ES cells were digested with BamHI and analyzed by Southern blot
analysis by a previously described probe (6). The 5- and 7-kb bands for Cypher2sKI (B) as well as the 5- and 8-kb bands for Cypher3sKI (C)
represent wild type and targeted alleles, respectively. Detection of Cypher protein in Cypher2sKI (D) and Cypher3sKI (E) mice was analyzed by
protein blot analysis. Protein was extracted from 1-month-old gastrocnemius skeletal muscle from wild type (left) and Cypher2sKI and
Cypher3sKI homozygous mice (right) and analyzed with skeletal muscle-specific Cypher antibodies as well as chemiluminescence.
Surviving homozygous knock-in mice were growth-retarded and their weight was ~60% that of littermate controls. Some surviving homozygous knock-in mice have been allowed to age and are now 1 year old. Histopathological and histochemical analyses were performed on fresh frozen muscle specimens from mutant and wild-type littermates from 1 month to 10 months of age. No significant age-dependent differences were observed. Representative results for skeletal muscle are shown in Fig. 5.

Skeletal muscle from homozygous knock-in mutants displayed cytoarchitectural abnormalities that were more marked in Cypher3sKI relative to that obtained from Cypher2sKI. Compared with wild type muscle, mutant muscle fibers exhibited a marked variation in myofiber size with numerous atrophic fibers (Fig. 5, A–C). Several mutant fibers had a lobulated appearance with subsarcolemmal invaginations highlighted by the modified Gomori trichrome stain (Fig. 5, D–F). Several fibers also showed multifocal loss of the normal sarcoplasmic pattern with a clear, hyaline appearance. Reaction for myofibrillar ATPase localization at pH 4.3 demonstrated that atrophic fibers were of both fiber types (Fig. 5, G–I) with no difference in the fiber type composition compared with wild type muscle. Lobulations were also present in both fiber types. Mitochondrial cytochrome c oxidase staining demonstrated that enzyme activity was prominent in the sarcolemmal invaginations and reduced or absent in the multifocal areas having the clear, hyaline appearance with the modified Gomori trichrome stain. This pattern of reactivity confirmed the presence of mitochondrial accumulations in subsarcolemmal areas and loss in clear areas.

In cardiac muscle, fibrosis was detected in tissue from both Cypher3sKI and Cypher2sKI homozygous mutants, again with greater severity for Cypher3sKI (data not shown).

**DISCUSSION**

We have discovered two novel isoforms of Cypher, Cypher1s, and Cypher2c. The human orthologue of Cypher has been designated as ZASP, and four splice variants have been described (17). Three of the human isoforms, KIAA0613, ZASP, and ZASP variant 3, are the counterparts of Cypher1c, Cypher2s, and Cypher3s, respectively. The fourth human splice isoform, ZASP variant 2, has not yet been described in mice. Cypher has also been referred to as Oracle (13), and two splice forms have been described. Oracle 1 and Oracle 2 correspond to Cypher1c and Cypher3c, respectively. It is possible that there are more novel isoforms of Cypher to be discovered.

Our analysis of the six isoforms of Cypher expressed in striated muscle has revealed for the first time the existence of two classes of splice variants that are skeletal and cardiac muscle-specific. Furthermore, we have also demonstrated that splicing of Cypher mRNA is developmentally regulated both in skeletal and cardiac muscle. In both cardiac and skeletal muscle, Cypher2, the short isoform, is expressed to barely detectable levels during embryogenesis, with expression increasing postnatally, suggesting an important role for this isoform in mature muscle. In cardiac muscle, the predominant long isoform throughout development is Cypher1c. However, in skeletal muscle, the predominant long isoform switches from Cypher1s during embryogenesis and in neonatal muscle to Cypher3s with progressive aging.

The discovery of cardiac and skeletal muscle-specific domains in Cypher suggests that mutations in the human cypher gene could result in human disease with symptoms restricted to either heart or skeletal muscle or present in both, depending on which exons are affected by mutation. Mutations affecting exons specific to short Cypher2 isoforms or long Cypher1/3 isoforms or all isoforms might be expected to have varying severity.

To investigate the ability of either short or long Cypher

**TABLE I**

| Type of Crosses | No. of Crosses | Age at genotyping | Genotype<sup>a</sup> | +/+ | +/KI | KI/KI |
|----------------|---------------|------------------|---------------------|-----|-----|------|
|                |               |                  | Observed no. | Expected no. | Observed no. | Expected no. | Observed no. | Expected no. |
| Cypher2sKI     | 35            | 4 weeks          | 88           | 74            | 170           | 147            | 36           | 74            |
| Cypher3sKI     | 32            | 4 weeks          | 76           | 57.5          | 143           | 115            | 11           | 57.5          |
| Cypher2sKI     | 7             | Newborn          | 12           | 14.5          | 30            | 29             | 16           | 14.5          |
| Cypher3sKI     | 8             | Newborn          | 19           | 16.5          | 29            | 33             | 18           | 16.5          |

<sup>a</sup> +/+; wild type; +/KI, heterozygous; KI/KI, homozygous.
isoforms to rescue Cypher function in the null mutant, we substituted cDNAs encoding cypher2s or cypher3s for the endogenous cypher gene. These two skeletal isoforms have in common their N-terminal 255 amino acids containing the PDZ domain yet differ in their C-terminal sequences, with that of Cypher3s containing three LIM domains. No Cypher null mice can survive past 5 days postnatally. In contrast, a percentage of Cypher2sKI and Cypher3sKI homozygous mice can survive to adulthood, indicating that, in some contexts, the first 255 amino acids can provide sufficient Cypher function to ensure survival and that in this respect the remainder of the protein sequence, including the three LIM domains, is dispensable. However, only 50% of Cypher2sKI and 20% of Cypher3sKI mice survive, and survivors all display growth retardation and pathological striated muscle phenotypes, suggesting that each isoform has distinct functional roles that are required for normal muscle structure and function. Pathological changes within the mutant muscle were of a myopathic, noninflammatory nature and could fit into a dystrophic phenotype.

RNA and protein analyses demonstrated that both Cypher2s and Cypher3s are up-regulated during the course of development and are the two predominant isoforms in adult skeletal muscle (Fig. 3, B, D, and F). However, at postnatal day 1, Cypher1s and Cypher2s are the predominant isoforms in skeletal muscle, with Cypher3s being expressed at very low levels, suggesting that Cypher1s and/or Cypher2s are required for muscle function at this period of time. This may explain why ~80% of Cypher3sKI and only 50% of Cypher2sKI homozygous mice die within the first week after birth.

The fibrotic phenotype we observe in cardiac muscle could reflect the absence of cardiac muscle-specific isoforms or the requirement for both long and short isoforms. The knock-in approach utilized here does not allow us to discriminate between these two possibilities. To address this issue, we are currently knocking out cardiac or skeletal muscle-specific cypher exons.

Acknowledgments—We thank Dr. Jim Lin for tropomyosin antibody and Norma Prades for technical assistance.

REFERENCES

1. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., et al. (2001) Nature 409, 880–921
2. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wendl, M. C., Wilson, R. K., and Venter, J. C. (2001) Science 291, 1596–1604
3. Dubowitz, V. (1995) in Muscle Biopsy: A Practical Approach (Dubowitz, V., ed) pp. 19–40, Bailliere Tindale, London
4. Chen, J., Kubalak, S. W., Minamisawa, S., Price, R. L., Becker, K. D., Hickey, R., Ross, J., Jr., and Chien, K. R. (1998) J. Biol. Chem. 273, 1252–1256
5. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Nucleic Acids Res. 25, 3389–3402
6. Passier, R., Richardson, J. A., and Olson, E. N. (2000) Mech. Dev. 92, 277–284
7. Gomes, A. V., Guzman, G., Zhao, J., and Potter, J. D. (2002) J. Biol. Chem. 277, 35341–35349
8. Cazorla, O., Freiburg, A., Helmes, M., Centner, T., McNabb, M., Wu, Y., Trombitas, K., Labert, S., and Granzierl, H. (2000) Circ. Res. 86, 59–67
9. Trevisan, S., Bortolotto, G., Scannapieco, P., Salamon, M., Mouly, V., Vallee, G., and Lanfranchi, G. (1999) J. Cell Biol. 146, 465–475