The human unc-33-like phosphoprotein (hUlip/CRMP-4) is a member of a family of developmentally regulated genes that are highly expressed in the nervous system. Mutations in the C. elegans unc-33 gene lead to worms with abnormal movements. The hUlip gene encodes a 570-amino acid protein with 98% homology to its murine (Ulip) (Byk, T., Dobransky, T., Cifuentes-Diaz, C., and Sobel, A. (1996) J. Neurosci. 16, 688–701) and rat (CRMP-4) (Wang, L. H., and Strittmatter, S. M. (1996) J. Neurosci. 16, 6197–6207) counterparts. hUlip lacks sequences corresponding to the first six exons found in unc-33. unc-33 exons correspond to homologous hUlip exons as follows: VII to 1 and 2, VIII to 3–9, IX to 10–12, and X to 13 and 14. Using the hUlip clone 1 phage, fluorescence in situ hybridization analysis indicates that the hybridization signal localizes to human chromosome 5q32. Deletion analysis of 5′-flanking sequences delineated the sequences sufficient to express a reporter gene in both neuroblastoma cells and myoblasts. A consensus MyoD/myogenin binding site is located in a region of the downstream promoter that is nearly identical to its mouse homologue. Mutagenesis shows that this conserved MyoD/myogenin site is necessary for full promoter activity in both myoblasts and neuroblastoma cells.

To study the genetics of animal behavior, Caenorhabditis elegans were mutagenized, and mutants were identified that had uncoordinated movements and partial defects in their egg-laying behavior (5, 6). The unc-33 mutations were unique in that they were associated with alterations in axonal outgrowth and guidance of several classes of neurons yet did not involve any defects in neural cell lineages, numbers, or positioning (7, 8). The gene responsible for this defect encoded a novel intracellular protein, unc-33, that was highly expressed within neuronal processes during early embryogenesis (1). The first putative mammalian homologue of unc-33 was identified functionally as a 62-kDa protein, collapsin response-mediated protein (CRMP-62) that mediated collapsin signaling (9). During neural development, collapsins participate in axonal pathfinding by stimulating growth cone collapse and preventing neurite extension. The identification of other unc-33-related genes followed; TOAD-64 (turned on after cell division) was identified as a 64-kDa protein that was highly expressed in the brain, particularly in the cerebral granular neurons, after cessation of cell proliferation (10); Ulip (unc-33-like phosphoprotein) (2) was identified as a highly expressed abnormal protein in the brain detected by anti-stathmin antibodies; hUlip was identified as an anonymous gene induced during the neuronal differentiation of human neuroblastoma cells (4); 5,6-dihydroyridine amidohydrolase (an enzyme involved in uracil and thymine catabolism) and several genes termed 5,6-dihydroyridine amidohydrolase-related proteins (DRP-1–3) (11) were found to be homologous by sequence analysis.

A systematic study indicated that there are at least four mammalian homologs of unc-33 (Uliph/Ulip/CRMP-4/DRP-3; Ulip2/CRMP-2/CRMP-62/TOAD-64/DRP-2; CRMP-1; and CRMP-3) that are expressed in distinct yet partially overlapping patterns of expression during the development of the nervous system (3). CRMP-1 and Ulip (CRMP-4) have the most restricted pattern of expression, being turned on in E15 and off in P15 rat brain tissue. Ulip/hUlip/CRMP-4/DRP-3 is highly expressed in fetal but not adult neural tissues by in situ analysis and evaluation of total RNA, yet it is not detected at comparable levels in fetal or adult muscle tissue (3, 4). However, when poly(A)⁺ Northern blots of normal tissues are examined, Ulip/hUlip/DRP-3/CRMP-4 is detected in newborn rat muscle and in the adult human heart and skeletal muscles (2, 11).

The hUlip protein encodes a 62-kDa protein with no signal sequence, transmembrane domain, or classical protein interaction domains. However, hUlip contains a number of consensus phosphorylation sites that may serve as substrates for signaling molecules such as Cdk, protein kinase C, and proline-directed kinases (4). Increases in phosphorylation and dephosphorylation are required for maximal expression in neuroblastoma and myoblasts.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† The abbreviations used are: CRMP, collapsin response-mediated protein; TOAD, turned after cell division; Ulip, unc-33-like phosphoprotein; hUlip, human Ulip; mUlip, murine Ulip; AO, anaesthetic oligodendroglialoma; kb, kilobase(s); bp, base pair(s); SSC, sodium chloride; citrate; DMEM, Dulbecco's modified Eagle's medium; PCR, polymerase chain reaction; EST, expressed sequence tag; RA, retinoic acid; DRP, dihydropyrimidinase amidohydrolase-related protein.
phorylation of hUlip and Ulip are detected upon induction of neuritogenesis by retinoids in neuromastoma cells or by nerve growth factor in PC12 cells (2). Collapsin-induced neurite retraction is blocked in cells injected with antibodies to CRMP-62 (9), and experiments indicate that CRMP-62 is involved in G protein-mediated transduction of the collapsin signal (9) as well as other non-G protein-dependent effects on neurite extension and growth cone morphology. Functional studies in other tissues have not been evaluated.

To evaluate the developmental regulatory mechanism of hUlip, we have isolated the hUlip genomic sequence and studied its structure, chromosomal location, 5'-flanking sequences, and expression. We demonstrate that two promoters exist in the 5' region of the hUlip gene and that an evolutionarily conserved MyoD/myogenin site is necessary for optimal expression of hUlip in vitro.

**EXPERIMENTAL PROCEDURES**

Isolation and Characterization of hUlip Genomic Clones—Two million phage clones from a human leukocyte genomic library in EMBL3 were screened using a 12^{32}P-labeled 1.8-kb Xhol–XapI fragment of the hUlip cDNA (4), which encompasses the entire coding region. Five positive clones were identified, and phage DNA was isolated with a Lambda DNA isolation kit (Qiagen). For the characterization of selected hUlip cDNA (4), which encompasses the entire coding region. Five positive clones was executed using a neural network algorithm (14, 15). With TESS (Transcription Elements Search Software; University of AF246692). A computer search for putative cis-elements was performed to identify genomic DNA fragments that hybridized with a 3^{23}P-labeled hUlip cDNA fragment. These fragments were subcloned into pBluescript SK– (Stratagene) and analyzed by restriction endonuclease digestion and sequencing. Synthetic oligonucleotide sequencing primers used to define exon/intron boundaries are listed in Table I.

Identification of a Genomic Fragment Containing hUlip Exon 2—To identify the putative hUlip exon 2 in the human genome, the PCR-based genome walking strategy was utilized (12). Sets of sense and antisense primers that are specific for exon 2 sequence and adaptor primers were used to amplify human genomic gene fragments with adaptor-ligated genomic DNA as templates (GenomedWalker Kit, CLONTECH). The amplified products were subcloned into the PCR 2.1 vector (Invitrogen), and their sequences were verified.

Sequence Analysis of hUlip 5'-Flanking Region—A 2.5-kb BamHI fragment of phage clone 4 that contained 2.3 kb 5' to hUlip exon 1 was subcloned into pBluescript SK– (Stratagene) and sequenced by restriction endonuclease digestion. Sequence was determined in both orientations (GenBankTM accession number AF246692). A computer program for putative cis-elements was performed with TESS (Transcription Elements Search Software; University of Pennsylvania) (13). Additionally, a search for putative transcriptional start sites was conducted using a neural network algorithm (14, 15).

Fluorescence in Situ Hybridization—Phage clone hUlip-1 DNA was labeled with biotin using a T3/T7 DNA labeling kit (Bio Rad) and coprecipitated with 10 μg of biotin-coupled avidin (Vector) and amplified with one layer of anti-avidin antibody (Vector). The signal was detected by two layers of fluorescein isothiocyanate-conjugated goat anti-rabbit IgG (Vector) with 4%,6-diamidino-2-phenylindole dihydrochloride (Roche Molecular Biochemicals) in an antifade solution. A Zeiss Axioshot microscope equipped with a cooled charge-coupled device camera (Photometrics) was used for image acquisition. Image analysis was carried out with IP Laboratories Spectrum software (Signal Analyses). Fractional length measurements were made using the Fract Length 1 Probe Macro in the NIH Image software (Life Sciences Division, Los Alamos National Laboratory). 5'-untranslated region (Fig. 1A). The hUlip phage clone 4, whose insert size was 10.5 kb, contained a portion of the hUlip 5'-untranslated region, a region encoding the initiation methionine, an additional 12 amino acids, and a splice donor site. Eighty-nine base pairs of hUlip coding region sequence were not identified in the phage clones. Several library screens using the cDNA fragment corresponding to the putative exon 2 as well as the 5' intronic region of exon 3 and the 3' intronic region of exon 1 failed to isolate a genomic fragment containing exon 2 in the EMBL3 library. Therefore, we em-
employed the PCR-based walking strategy to identify exon 2 sequences in the human genome, and two successive rounds of PCR with the sets of sense or antisense primers each amplified 1.5-kb products. Sequence analysis reveals that these products overlap one another and contain the missing 89 bp of hUlip cDNA sequences (data not shown). Putative RNA splice acceptor and donor sites detected in these PCR products are compatible with the exon 1 donor and exon 3 acceptor splice sites.

From these results, we conclude that these PCR products contain hUlip exon 2. Table I lists the oligonucleotide sequences used to define the intron/exon boundaries of the hUlip gene, and the sequences at the intron/exon junctions are given in Table II. The 1.71-kb open reading frame of the hUlip cDNA corresponds to sequences conserved in all mammalian -half of exon VII as well as exons VIII–X, and this transcript corresponds to sequences conserved in all mammalian unc-33 family genes including the hUlip cDNA (1). The 5’ regions of the unc-33 transcripts are distinct, while the 3’ regions are identical, with all predicted transcripts being translated in the same open reading frame. The unc-33 exon VII and VIII and exon IX and X splice sites are identical to the hUlip exon 2 and 3 and exon 12 and 13 splice sites, respectively. hUlip exons 1–3 include the protein subdomain region A (Fig. 1A and Table III), which contains a region highly conserved with unc-33. This region is functionally important for the collapsin-induced Ca\(^{2+}\) influx response (9). An overall comparison of the exonic structure of hUlip with unc-33 reveals four highly homologous regions (Table III). Amino acids encoded by hUlip exons 2 and 3 (region A), exons 6 and 7 (region B), and exons 10 and 11 (region C and the D region within C) showed significant sequence similarities of 41–56% with unc-33 (Fig. 1A). These regions correspond to domains that were previously identified (Fig. 1A) (2, 4). An additional conserved domain encoded by hUlip exon 14 (region E, Fig. 1A and Table III) shows a 41.4% homology with unc-33.

**Chromosomal Location**—hUlip phage clone 1 was used to determine its chromosomal location. Specific hybridization was detected on the long arm of chromosome 5 in all 20 metaphases examined. Fractional length measurements (n = 30) yielded an average FLpter value of 0.823 ± 0.021, corresponding to a regional assignment of 5q32 (Fig. 2). In support of our finding, DRP-3 has been mapped to chromosome 5 by radiation hybrid mapping by the Whitehead Institute Center for Genome Research.

**Promoter Region of hUlip**—A 2.5-kb BamHI fragment containing exons AO and 1 was sequenced to study the putative promoter region of the hUlip gene (Fig. 3, A and B). The outlined letters represent exonic sequence from hUlip cDNA. The shaded boxes designate extended exon 1 and exon AO sequences derived from an anaplastic oligodendroglioma EST clone (accession number A1570709). Rapid amplification of cDNA ends was used to map transcriptional start sites in RNA from RA-treated neuroblastoma cells. Several putative start sites were found, although all of these were 3’ (within exon 1) to the 5’-end of a hUlip cDNA from a fetal brain cDNA library (data not shown). The 5’-end of the Ulip cDNA corresponds to 136 (relative to A\(^{+}\)TGG) in the hUlip gene (Fig. 3B). Additionally, this position corresponds to a putative start site predicted by a neural network algorithm (14, 15). A search for a -30
TATA box correlating with the 5’ terminus of known exon AO sequence did not reveal a TATA sequence. Although a consensus TATA sequence is found at −1166 bp, it seems unlikely that it is required for expression of exon AO sequences, since the size of the hUlip mRNA has been estimated to be between 5.5 and 5.8 kb by Northern analysis, which corresponds to the sum total of known exon sequences (4). The 2.8-kb unc-33 mRNA, which is most homologous to the hUlip mRNA, does not contain a TATA box in its 5’ region (1). These findings along with the known characteristics of TATA-less promoters suggest

| Exon | Synthetic oligonucleotide sequencing primers used to define exon/intron boundaries |
|------|----------------------------------------------------------------------------------|
| 1 (sense) | ATTCACTCCACCTGATCTCGGGGCGCTGTG |
| 3 (sense) | CTTTGTCCTTGGAAGAAGTC |
| 4 (sense) | GCACCTTCCTGCTTGA |
| 5 (sense) | TGGTGCTGCCGCTGTGAGT |
| 6 (sense) | CTTCGAGTGAGTGGGAGGAGGAGC |
| 7 (sense) | GGCAGTTATCCCTTATTTCCCAA |
| 8 (sense) | ATCCCAATTGCAGCC |
| 9 (sense) | TGTTGGAAATGGGGATAA |
| 10 (sense) | ATCCGAGCTTCCTCACACC |
| 11 (sense) | ACCGAGTCCTCCTGCCAAG |
| 12 (sense) | CGCCTCCTCGCCGCC |
| 13 (sense) | GTGTTGACGTAGAGA |
| 14 (sense) | CTCTGAGCCTGTCAGAA |

* Sense primers to define 3’ end of exons 5 and 6 were designed from the nucleotide sequence of introns 4 and 5, respectively.

| Exon | Intron/exon junction sequences |
|------|-------------------------------|
| Exon size | Sequence at exon/intron junction | Amino acid(s) interrupted | Intron size bp |
| AO | GGAAG/GTAGG ... ... | 206 | Thr13 | 417 |
| 1 | TCACG/GTCCAC ... ... | 206 | Ser14 | 26,838 |
| 2 | ATAAA/GTATG ... ... | 185 | Lys43 | 6244 |
| 3 | GATCA/GTAAAG ... ... | 165 | Ile105 | 2574 |
| 4 | CAGAG/GTAGG ... ... | 165 | Gly180 | 2049 |
| 5 | CCCAG/GTACAG ... ... | 165 | Leu230 | 906 |
| 6 | AAGAG/GTAAG ... ... | 165 | Ser323 | 1438 |
| 7 | TAAAG/GTGA ... ... | 165 | Gly371 | 1236 |
| 8 | GCCAG/GTTGGG ... ... | 165 | Val370 | 1943 |
| 9 | CTGTG/GTAAAG ... ... | 165 | Ala371 | 1435 |
hUlip Genomic Structure and Promoter Analysis

Table III

| Exonic structure of hUlip and unc-33 |
|-------------------------------------|
| **Exon** | **Amino acid residues** | **Exon** | **Amino acid residues** | **Homology** |
| 1        | 1–13                   | VII      | 313–325                | 30.7        |
| 2        | 14–43                  | VII      | 326–361                | 53.3\*      |
| 3        | 43–105                 | VIII     | 361–387                | 37.7        |
| 4        | 105–160                | VIII     | 387–471                | 23.6        |
| 5        | 160–180                | VIII     | 471–486                | 23.8        |
| 6        | 181–207                | VIII     | 487–513                | 40.7        |
| 7        | 208–220                | VIII     | 514–536                | 52.2        |
| 8        | 231–271                | VIII     | 537–577                | 27.5        |
| 9        | 271–323                | VIII     | 577–627                | 26.4        |
| 10       | 323–370                | VIII, IX | 627–673                | 41.5        |
| 11       | 371–427                | IX       | 674–730                | 56.1        |
| 12       | 428–487                | IX       | 731–790                | 26.7        |
| 13       | 488–542                | X        | 791–824                | 11.1        |
| 14       | 542–570                | X        | 824–854                | 41.4        |
| Overall  |                        |          |                        | 36.2        |

* Percentage homologies above the overall value are shown with boldface numbers.

![Fig. 2](http://www.jbc.org/)

**Fig. 2. Chromosomal localization of hUlip maps to 5q32 by fluorescence in situ hybridization.** Partial metaphase (A) illustrates a fluorescence signal on both homologues of chromosome 5 (arrows). The enlarged image (B) demonstrates regional assignment to 5q32.

that a primary start site exists at −136 with many heterogeneous starts occurring downstream in neuroblastoma cells. In anaplastic oligodendroglioma, presumably an upstream promoter is used, giving rise to a larger primary transcript beginning with exon AO, which is spliced to exon 1.

Given that Ulip is expressed in nervous system tissue during development and is selectively expressed in adult muscle, the presence of a consensus MyoD/myogenin binding site at −288 bp in the putative hUlip promoter may be important. To determine if the MyoD/myogenin site is conserved evolutionarily, the corresponding 5′ region of the murine Ulip gene was sequenced (GenBank™ accession number AF246699). The sequence MGD 1 (5′-CACTGCACCCCTCCCTCCTC-3′) located −253 bp upstream of the putative initiation codon was used as a sequencing primer. This primer was used in a dyeoxy sequencing reaction with the mUlip genomic subclone MG 5 to obtain a sequence encompassing the putative transcriptional start site (of the previously published mUlip cDNA; accession number X87817) and a consensus E-box located further 5′. Alignment of the sequence with the corresponding human Ulip genomic sequence by the GCG Lite version of the program Gap (Genetics Computer Group) demonstrates a region of near identity between +1 and −120, including identical E-box sequences represented by the boxed nucleotides in Fig. 3C. The MyoD/myogenin site as well as the c-Myc, Oct-1, and Ets-1 binding sites are conserved both in sequence and spatial relationship to one another.

**hUlip Expression and Promoter Analysis—**Some members of unc-33 family are expressed in adult heart and skeletal muscle (2, 11). Consistent with this finding, we detect differential expression of hUlip mRNA in poly(A)+ Northern blots from heart and skeletal muscle and low levels of expression in other adult tissues including brain, kidney, lung, liver, placenta, and pancreas (Fig. 4A). As a prerequisite to promoter analysis by transient transfection, we evaluated the expression of hUlip message in total RNA isolated from C2C12 myoblasts and KCNR neuroblastoma cells (Fig. 4B). Although it appears that hUlip may be more highly expressed in myoblasts compared with the neuroblastoma cells, 5-fold more myoblast total RNA was evaluated. PhosphorImager analysis, normalized to total RNA, indicated that in total RNA the neuroblastoma cells express almost twice the level of Ulip mRNA that is detected in myoblasts (Fig. 4B). Additionally, myogenic differentiation of myoblasts into myotubes produced a down-regulation of hUlip expression to almost undetectable levels by Northern analysis. Consistent with our previous results, retinoic acid-induced neurogenic differentiation of KCNR neuroblastoma cells showed a 2.5-fold increase in hUlip expression after 4 days of treatment (4).

Since the hUlip gene is active in both neuroblastoma and myogenic cells, the neuroblastoma cell line NGP and C2C12 myoblasts were chosen as model systems for hUlip promoter analysis. The 2.5-kb hUlip promoter was cloned upstream of the luciferase reporter gene. Nested deletions and point mutations were generated from the promoter using the constructs −2300UlipLuc (Fig. 5). Luciferase activity was normalized to β-galactosidase activity, and the −2300UlipLuc activity was set to 100% in both cell lines such that relative comparison can be made between cell lines. Computer data base analysis shows a perfect MyoD/myogenin binding (CCAGCTGGC) site in close proximity (−150) to the transcriptional start site. This motif could have implications for muscle and neuroblastoma expression, since factors other than myogenic transcription factors such as NeuroD and Myc and other members of the bHLH family are known to bind to E-boxes (20). Insertional mutagenesis of the hUlip E-box created clones E-GCUlipLuc and E-TAUlipLuc that disrupted the core binding site. Activity was decreased approximately 2-fold from wild type levels in both C2C12 myoblasts and NGP cells with both mutations. This demonstrates that in both neuroblastoma and myoblasts, the E-box element is necessary for full activity of the promoter.

Deletion of far 5′-flanking sequences to −940 in the hUlip promoter showed little or no difference from the wild type. Deletion from −940 through −540 showed a 2-fold reduction of activity in both myoblast and neuroblastoma cells. This deletion removes the region that includes the possible promoter/start site for exon AO and may be responsible for the decrease in activity if it is contributing to the transcriptional output in neuroblastoma and myoblasts. This interval also contains several sites for ubiquitous factors such as AP-2, Sp1, NF1, AP-1, and C/EBP. Tissue-specific binding sites for MFB1 and MEF-2 also appear in this region. Between −540 and −290, there is an approximately 10–20% decrease in both myoblast and neuroblastoma cells, corresponding to the possible deletion of more putative Sp1 binding sites. The deletion to −150 disrupts the MyoD/myogenin binding site and shows no significant decrease...
in neuroblastoma expression but gives a 30% decrease in myo-
blast expression. This indicates that elements contributing to
myoblast-specific expression reside in the region between
2290 and 2150 such as the MyoD/myogenin site. Further deletion to
220 gives 20–25% loss of activity in both neuroblastoma and
myoblast cells to levels no more than 2-fold above background,
consistent with the activity expected of the basal promoter.

DISCUSSION

Previously, we identified hUlip as a cDNA up-regulated dur-
ing retinoic acid-induced neuritogenesis of neuroblastoma (4).
In this report, we characterize the structure, chromosomal location, and ex-
pression and delineate the promoter sequences sufficient for
expression of the hUlip gene in culture.

To determine the genomic structure of the hUlip gene, we
screened a human leukocyte library. A genomic sequence, en-
compassing all sequences included in the hUlip cDNA (GenBank™ accession number AF246692). Outlined lettering
designates exon 1 sequences. Shaded sequences denote additional exonic sequence defined by the anaplastic oligodendroglioma EST. Boldface
letters indicate putative TATA boxes. Coordinates are shown relative to the initiator codon (double underlined). C, sequence alignment of murine
and human Ulip exon 1 and 5'-flanking sequences. Utilizing the TESS-String-Based Search (Transfac version 3.2), we analyzed the hUlip genomic
5' sequences for known transcription factor binding sites, which are in boldface type. The sequence coordinates are given to the
left, with −1 designating the first nucleotide immediately 5' of the initiator codon. The mouse sequence is a composite of the 5' portion of a Ulip cDNA reported
by Byk et al. (2) (GenBank™ accession number X57817, presented in lowercase letters with a filled triangle located below its 5' terminus) and the
mUlip gene (shown in uppercase letters) (GenBank™ accession number AF246693). hUlip cDNA sequences are indicated by outline lettering,
and the amino acid sequence is shown below the respective codons. Boxed sequences designate regions of the Ulip promoter that are highly conserved.

hUlip Genomic Structure and Promoter Analysis

Fig. 3. Structure and sequence of hUlip 5' region. A, schematic of the 2.5-kb BamHI fragment containing the putative promoter region of
hUlip. B, sequence analysis of the 2.5-kb BamHI genomic hUlip fragment cDNA (GenBank™ accession number AF246692). Outlined lettering
designates exon 1 sequences. Shaded sequences denote additional exonic sequence defined by the anaplastic oligodendroglioma EST. Boldface
letters indicate putative TATA boxes. Coordinates are shown relative to the initiator codon (double underlined). C, sequence alignment of murine
and human Ulip exon 1 and 5'-flanking sequences. Utilizing the TESS-String-Based Search (Transfac version 3.2), we analyzed the hUlip genomic
5' sequences for known transcription factor binding sites, which are in boldface type. The sequence coordinates are given to the
left, with −1 designating the first nucleotide immediately 5' of the initiator codon. The mouse sequence is a composite of the 5' portion of a Ulip cDNA reported
by Byk et al. (2) (GenBank™ accession number X57817, presented in lowercase letters with a filled triangle located below its 5' terminus) and the
mUlip gene (shown in uppercase letters) (GenBank™ accession number AF246693). hUlip cDNA sequences are indicated by outline lettering,
and the amino acid sequence is shown below the respective codons. Boxed sequences designate regions of the Ulip promoter that are highly conserved.
ately large intron also exists between CRIMP-1 exons 1 and 2, indicating that there are genomic structural similarities among these genes.

Comparison of the hUlip locus with *C. elegans unc-33* reveals that hUlip exon 1–14 coding sequences are homologous to worm exons VII–X. The 5′-half of *unc-33* exon I is highly conserved with hUlip. This region does not appear to function in the collapsin response, because antibodies to peptides made from this region do not block collapsin-mediated Ca²⁺ flux, and this region was also not included in the original cDNA functionally identified that mediated collapsin-induced Ca²⁺ flux (9). This region is most highly conserved among *unc-33*, hUlip, human dihydropyrimidinase, and the β-hydantoinase enzyme from *Bacillus stearothermophilus* and *Pseudomonas putida*. Amino acids encoded by exons 6 and 7, exons 10 and 11, and exon 14 also have a significant homology of 41–56% with *unc-33*. A high degree of sequence conservation often implies functional importance; thus, it is possible that these regions may encode another functional activity of these proteins independent from their activity in the mediation of collapsin signals. Hamajima et al. (11) have raised the possibility that these proteins may also function as amidohydrolases, although no such activity has been reported.

In the *C. elegans unc-33* mutants, mutational insertions occur within exon VII just 5′ to the start of the 2.8-kb *unc-33* mRNA that is homologous with hUlip and other members of the *unc-33*-like genes identified in mammals. These mutations result in deletion of the highly charged N-terminal region of the products encoded by the 3.8- and 3.3-kb *unc-33* mRNAs, although the protein-encoding sequences in the 2.8-kb *unc-33* mRNA are intact. It is not known whether alterations in one or all of the proteins encoded by these *unc-33* mRNAs are needed for the uncoordinated phenotype. While it is possible that the loss of the products of the 3.8- and 3.3-kb *unc-33* mRNAs are key, it is also possible that the altered regulation of the 2.8-kb

---

**FIG. 4.** hUlip expression in adult muscle tissue. *A*, Northern analysis of 1 µg of poly(A)+ from adult human tissues (CLONTECH) hybridized with 32P-labeled hUlip cDNA. *B*, Northern analysis of 50 µg of total RNA from murine myoblast cell line C2C12 (MB) and C2C12 myotubes (MT), 10 µg of total RNA from human neuroblastoma cell line KCNR (NB), and KCNR induced to differentiate with 5 µM retinoic acid (NB + RA) hybridized with 32P-labeled hUlip cDNA. The numbers below each lane are the relative densitometric units (RDU) as determined by PhosphorImager analysis using a Storm 840 PhosphorImager and ImageQuant software (Molecular Dynamics, Inc., Sunnyvale, CA) normalized to the quantity of total RNA loaded in each lane.

**FIG. 5.** hUlip promoter analysis. *A*, schematic of the transcriptionally important sequences of the hUlip promoter showing putative factor binding sites. The angled arrow represents the putative start site of the hUlip cDNA. *B*, deletion analysis of the hUlip promoter in both myoblast and neuroblastoma cells. All constructs are drawn to scale. Luciferase activities are normalized to cotransfected β-galactosidase and then set relative to the full-length promoter (−2300UlipLuc set to 100).
unc33 mRNA caused by mutations in its putative promoter region contribute to alterations in the expression of this gene that result in the uncoordinated phenotype of the worm.

The finding of expression of hUlip in adult muscle suggests that unc-33 related molecules may also play a role in muscle development or function. If this is true, then the original unc-33 defect in *C. elegans* may not be restricted to solely premature termination of axonal processes, which raises the possibility that the alterations in muscle development or function may contribute to the uncoordinated movements characteristic of these mutant worms. Interestingly, our expression analysis of hUlip mRNA during *in vitro* myogenesis shows hUlip message decreasing with differentiation into myotubes. This expression pattern may be explained by the lack of inner- or other environmental factors in *in vitro* that are otherwise present during normal development. Retinoic acid-induced differentiation of KCNR cells induces the slow accumulation of hUlip mRNA with the first detectable increase at 48 h and increasing to 2.5-fold after 96 h of treatment. The slow accumulation of hUlip mRNA and our previous studies showing no significant increase in hUlip transcription at 2 days (4) (22) indicate that RA does not significantly regulate hUlip transcription. In preliminary studies, transient transfection analysis of the hUlip 2.5-kb promoter also indicates that RA does not significantly alter hUlip transcription *in vitro* (data not shown).

RNase protection assays and 5'-rapid amplification of cDNA ends with neuroblastoma RNA suggest that the start site of transcription is heterogeneous and occurs near the 5' terminus of exon 1 in neuroblastoma cell line KCNR (data not shown). The predicted molecular weight of a message beginning in this vicinity is consistent with the size observed by Northern analysis. Sequence inspection shows that exon 1 5'-flanking sequences are GC-rich and do not contain an associated TATA box. These data support the hypothesis that hUlip transcription in neuroblastoma can initiate from an intragenic TATA-less promoter coincident with exon 1.

Given its high expression in nervous system tissue during development, the presence of a consensus MyoD/myogenin binding site at −288 bp in the putative promoter region of hUlip suggested that it also participated in muscle-specific expression. However, both *Drosophila* and vertebrate neurogenesis use a core differentiation network of transcription factors that strongly resembles the myogenic network of transcription factors. It is possible that this site also interacts with MASH, neurogenin, or NeuroD (20). The importance of the MyoD site for hUlip transcription is also supported by our finding that the 5' region of the murine Ulip gene has a region of near identity between −120 and +1 including identical E-box sequences. Aside from the murine and human Ulip MyoD/myogenin site, c-Myc, Oct-1, and Ets-1 binding sites are conserved showing no significant degradation in DNA polymerase activity compared with −290, yet it gives a 30% decrease in myoblast expression. This indicates that elements contributing only to myoblast-specific expression reside in the region between −290 and −150 such as the MyoD/myogenin site. Alternatively, the E-box could be necessary only to mediate the effects of upstream elements but has no transactivation activity on its own in neuroblastoma cells. Deletion to −20 results in a decrease to 2-fold over background and is consistent with the deletion of basal promoter elements such as binding sites for Sp1 and C/EBP.

In summary, we have presented the isolation and structural analysis of the human Ulip gene and initial characterization of its mechanism of expression. Although Ulip functional and expression studies have previously demonstrated its association with the neuronal phenotype (2), we have shown that in adult tissues it is most highly expressed in cardiac and skeletal muscle tissues. Additionally, we have ascertained that the transcriptional mechanism of hUlip expression relies on the presence of an E-box motif in both myoblast and neuroblastoma cells. Studies are in progress to further analyze hUlip function by gene knockout and the pattern of its expression by β-galactosidase knock-in. This will allow a better understanding of the role of hUlip in neuronal and muscular development and function.

Acknowledgment—We thank Patricia Harley for the construction and preparation of the hUlip reporter plasmids.

REFERENCES

1. Li, W., Herman, R. K., and Shaw, J. E. (1992) *Genetics* **132**, 675–689
2. Byk, T., Dobransky, T., Cifuentes-Diaz, C., and Sobel, A. (1996) *J. Neurosci.* **16**, 688–701
3. Wang, L. H., and Strittmatter, S. M. (1996) *J. Neurosci.* **16**, 6197–6207
4. Gaetano, C., Matsuo, T., and Thiele, C. J. (1997) *J. Biol. Chem.* **272**, 12195–12201
5. Brenner, S. (1973) *Br. Med. Bull.* **29**, 269–271
6. Brenner, S. (1973) *Genetics* **77**, 71–94
7. Hedgecock, E. M., Culotti, J. G., Thomson, J. N., and Perkins, L. A. (1985) *Dev. Biol.* **111**, 158–170
8. Hedgecock, E. M., Culotti, J. G., Hall, D. H., and Stern, B. D. (1987) *Development* **106**, 365–382
9. Goshima, Y., Nakamura, F., Strittmatter, P., and Strittmatter, S. M. (1995) *Nature* **376**, 509–514
10. Minturn, J. E., Fryer, H. J., Geschwind, D. H., and Hoefkild, S. (1995) *J. Neurosci.* **15**, 6757–6768
11. Hamajima, N., Matsuda, K., Sakata, S., Tamaki, N., Sasaki, M., and Nonaka, M. (1996) *Gene (Amst.)* **180**, 157–163
12. Siebert, P. D., Chenchik, A., Kellogg, D. E., Lukyanov, K. A., and Lukyanov, S. A. (1995) *Nucleic Acids Res.* **23**, 1087–1088
13. Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A. E., Kel, O. V., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Kolpakov, F. A., Podkolodny, N. L., and Kelchanov, N. A. (1998) *Nucleic Acids Res.* **26**, 362–367
14. Reese, M. G., and Eeckman, F. H. (1995) *The Seventh International Genome Sequencing and Analysis Conference, Hilton Head Island, SC, October 13–14, 1995*, pp. 152, American Mathematical Society, Providence, RI
15. Reese, M. G., Harris, N. L., and Eeckman, F. H. (1996) *Bioinformatics: Proceedings of the 1996 Pacific Symposium, Hawaii, January 3–6, 1996* (L. Hunter and T. E. Klein, eds) pp. 120–130, World Publishing, Co., Singapore
16. Pinkel, D., Straume, T., and Gray, J. W. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 2934–2938
17. Marchuk, D., Drumm, M., Saulino, A., and Collins, F. S. (1991) Nucleic Acids Res. 19, 1154
18. Wigler, M., Pellicer, A., Silverstein, S., Axel, R., Urlaub, G., and Chasin, L. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 1373–1376
19. Breathnach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383
20. Yun, K., and Wold, B. (1996) Curr. Opin. Cell. Biol. 8, 877–889
21. Torres, R., and Polymeropoulos, M. H. (1998) DNA Res. 5, 393–395
22. Thiele, C. J., Reynolds, C. P., and Israel, M. A. (1985) Nature 313, 404–406
23. Lyons, G. E., Micales, B. K., Schwarz, J., Martin, J. F., and Olson, E. N. (1995) J. Neurosci. 15, 5727–5738
24. Luscher, B., Mitchell, P. J., Williams, T., and Tjian, R. (1989) Genes Dev. 3, 1507–1517
Fig. 3. Structure and sequence of hUlip 5′ region. A, schematic of the 2.5-kb BamHI fragment containing the putative promoter region of hUlip. B, sequence analysis of the 2.5-kb BamHI genomic hUlip fragment cDNA (GenBank™ accession number AF246692). Outlined lettering designates exon 1 sequences. Shaded sequences denote additional exonic sequence defined by the anaplastic oligodendroglioma EST. Boldface letters indicate putative TATA boxes.
FIG. 3—continued