Gene Duplication Events Producing Muscle (M) and Brain (B) Isoforms of Cytoplasmic Creatine Kinase: cDNA and Deduced Amino Acid Sequences from Two Lower Chordates

Nicole A. Graber and W. Ross Ellington

Institute of Molecular Biophysics and Department of Biological Science, Florida State University

Creatine kinase (CK) is coded for by at least four loci in higher vertebrates—two cytoplasmic isoforms, muscle (M) and brain (B), and two mitochondrial isoforms, sarcomeric and ubiquitous. M is expressed primarily in skeletal muscle, while B is expressed in a variety of cells, including cardiac and smooth muscle fibers, neurons, transport epithelia, and photoreceptors. M and B subunits form very stable homodimers (MM [M-CK], BB [B-CK]) and heterodimers (MB). M-CK is capable of binding to the M line of the myofibril, thereby creating an energy transfer microcompartment; BB and MB CKs are not. M- and B-like CKs are present in all vertebrates yet examined, including fish. Cytoplasmic, dimeric CKs are widely distributed in the invertebrates. The only available amino acid sequence for an invertebrate dimeric CK, that of the protostome polychaete Chaetopterus variopedatus, is just as similar to the vertebrate M isoform as to the B isoform. Echinoderms lack dimeric, cytoplasmic CKs, which appear to be replaced by a dimeric arginine kinase which evolved secondarily from CK. Thus, it is likely that the gene duplication event producing the M and B isoforms occurred after the divergence of the chordates from echinoderms.

To narrow down the timing of this duplication event, we obtained the cDNA and deduced amino acid sequences of dimeric CKs from the tunicate Ciona intestinalis (subphylum Urochordata) and the lancelet Branchiostoma floridae (subphylum Cephalochordata). Our results show that these CKs are strikingly similar to both invertebrate and vertebrate CKs. However, phylogenetic analyses by neighbor-joining and parsimony show that these two enzymes appeared to have diverged before the point of divergence of the M and B isoforms. Thus, the gene duplication event for formation of the muscle and brain isoforms of CK most likely occurred during the radiation of the fish, a time noted for gene duplication events at a variety of other loci.

Introduction

Creatine kinase (CK) catalyzes the reversible transfer of phosphate from creatine phosphate (CP) to ADP yielding ATP. CK is typically found in cells which display high and variable rates of energy turnover such as muscle fibers, neurons, transport epithelia, photoreceptors, and primitive-type spermatozoa (Wallimann et al. 1992; Wallmann and Hemmer 1994). CK is a member of a highly conserved enzyme family called the guanine (phosphagen) kinases, also consisting of arginine kinase (AK), glycyocynamine kinase (GK), lombricine kinase (LK), taurocyamine kinase (TK), and hypotaurocyamine kinase (HTK). Vertebrates express CK only, and this enzyme is widely distributed in both lower and higher invertebrate groups and in the lower chordates (Watts 1971, 1975).

Four different isoforms of CK are known to exist in the cells of mammals and birds. Two of these, muscle type (M) and brain type (B), are localized in the cytoplasm and form homodimeric (MM, BB) and heterodimeric (MB) isoenzymes (Eppenberger, Dawson, and Kaplan 1967). MM-CK (M-CK) has the unique property of binding to subcellular structures such as the myofibrillar M line (Turner, Wallimann, and Eppenberger 1973), while this binding capacity is lacking in MB and BB (B-CK) isoenzymes (Wallimann, Moser, and Eppenberger 1983; Schäfer and Perriard 1988). The binding epitope for M-CK is present in the N-terminal region of this protein (Stolz and Wallimann 1998), and binding has recently been shown to be mediated by a "lysine charge clamp" unique to the M subunit (Hornemann, Stolz, and Wallimann 2000).

Two additional isoenzymes of CK in higher vertebrates have been identified, both of which are primarily octameric and are found in the mitochondrial intermembrane compartment (reviewed by Wyss et al. 1992). One mitochondrial CK isoenzyme is found in nonmuscle tissue (ubiquitous MiCK), while the other is restricted to muscle (sarcomeric MiCK). In chickens, these two isoenzymes are referred to as Mi,a and Mi,b-CK for acidic and basic forms, respectively.

Cladistic analyses of amino acid sequences of vertebrate CKs have shown that the M and B isoforms each display distinct clusters which share a common immediate ancestor, while the Mi,a and Mi,b isoforms each form distinct clusters which share a common ancestor with the M and B isoforms (Mühlebach et al. 1994). This could be evidence that two separate gene duplication events occurred in the CK cluster: one, from an original ancestral gene, giving rise to ancestral mitochondrial and cytosolic CK genes, and the next giving rise to two mitochondrial (ubiquitous and sarcomeric) and two cytosolic (M and B) CK isoforms (Qin et al. 1998). Extensive characterization of the creatine kinase gene has provided new evidence which further supports the idea of these two distinct gene duplication events (Mühlebach et al. 1996; Qin et al. 1998). The genes for the two cytosolic CK isoforms are identical in size and location of coding region exons, whereas they differ from the mitochondrial CK gene exons in length and location in all but one exon (Qin et al. 1998).
The timing of the gene duplication event creating the M and B isoforms is not clear. Trout CK is B-like (Garber, Winkfein, and Dixon 1990), while Torpedo (Giraudat et al. 1984) and carp (Sun, Hui, and Wu 1998) CKs cluster with the M isoform in phylogenetic analyses (Pineda and Ellington 1999). Thus, both isoforms are present in fish, so it is entirely appropriate to look at more primitive groups for the point of divergence. As indicated previously, CK is widely distributed in the invertebrates and in nonvertebrate chordates (Watts 1971). However, cDNA and deduced amino acid sequences exist for only two invertebrate CKs, and none exist for lower chordate CKs. One such sequence, the dimeric, cytoplasmic CK from the protostome polychaete Claeotopterus variopedatus (Pineda and Ellington 1999), is equally similar to vertebrate B and M isoforms and in a phylogenetic analysis was shown to have diverged prior to the appearance of these latter two forms. The deuterostome echinoderms lack dimeric CKs but, instead, have either a dimeric arginine kinase (which evolved secondarily from CK [Suzuki et al. 1999]) or an unusual monomeric CK consisting of three contiguous CK domains (Ratto, Shapiro, and Christen 1989; Wothe, Charbonneau, and Shapiro 1990).

The available evidence would suggest that the gene duplication event forming the M and B isoforms occurred after the divergence of the chordates from the echinoderm clade. To narrow down the timing of this event, we determined the cDNA and deduced amino acid sequences of dimeric CKs from a urochordate, Ciona intestinalis, and a cephalochordate, Branchiostoma flordia. Studies of 18S rDNA and mitochondrial DNA sequences support an extremely close relationship between the cephalochordates and the vertebrates (Turbeville, Schultz, and Raff 1994; Wada and Satoh 1994; Spruyt et al. 1998). Our results show that these two CKs, although very similar to both the M and the B isoforms, diverged before the appearance of the predominant cytoplasmic isoforms in the vertebrates. Thus, it seems likely that the gene duplication event producing M and B occurred during the early radiation of the fish.

Materials and Methods

Animals and Chemicals

Specimens of the tunicate C. intestinalis were purchased from the Marine Biological Laboratory, Woods Hole, Mass., while the specimens of the lancelet (=Amphioxus) Branchiostoma flordia were obtained from the Gulf Marine Specimen Laboratory, Panama, Fla. Organisms were either frozen in liquid nitrogen and stored at −80°C or used immediately. Biochemicals were purchased from Boehringer Mannheim (Indianapolis, Ind.) and Sigma Chemical Co. (St. Louis, Mo.). All other chemicals were of reagent grade quality.

Properties of C. intestinalis and B. flordia CKS

Preparation of Tissue Extracts

Fresh or previously frozen tissue was homogenized in varying volumes of extraction buffer (50 mM imidazole/HCl [pH 7], 1 mM EDTA) containing 14 mM or 1.4 mM 2-mercaptoethanol or 0.5 mM DTT-dithiothreitol (depending on protocol). Homogenization was accomplished using either a Brinkman Polytron or an Ultra Turrax. The resulting homogenates were clarified by high-speed centrifugation, and the resulting supernatants were used immediately for activity and relative molecular mass (Mr) determinations.

Enzyme Assays

CK activity was determined by the protocol of Strong and Ellington (1993), substituting creatine phosphate for arginine phosphate. Assays were conducted in a Gilford 252-I spectrophotometer.

Native Mr Determinations

Native Mr values for both C. intestinalis and B. flordia CKs were determined by size exclusion chromatography using a Superdex 200HR FPLC column (Amersham Pharmacia, Arlington Heights, Ill.). Running buffer was 50 mM sodium phosphate (pH 7), 150 mM NaCl, 0.5 mM EDTA, 1.4 mM 2-mercaptoethanol. Flow rate was 0.4 ml/min. Chromatography was conducted using a Beckman System Gold apparatus. All runs were done in triplicate.

For C. intestinalis CK, enzyme was purified by ion exchange chromatography (Graber 2000) prior to Mr determinations. Ciona intestinalis CK samples (100 μl) were injected into the Superdex 200HR column and elution monitored at 280 nm using a diode array detector. The size exclusion column was calibrated using protein standards (thyroglobulin 669 kDa; ferritin, 440 kDa; aldolase, 158 kDa; MM-CK, 81 kDa; ovalbumin, 45 kDa). Native Mr was obtained from a plot of log Mr of standards versus elution time.

In the case of B. flordia CK, frozen tissue was homogenized in five volumes of extraction buffer and centrifuged, and the supernatant was passed through a 0.22-μm centrifugal filter. A small volume of the filtrate (50 μl) was injected into the Superdex 200HR column, and 0.1-ml fractions were collected. The CK elution profile was determined by assaying each fraction for enzyme activity. Native Mr was determined by a plot of the log Mr of standards (aldolase, 158 kDa; MM-CK, 81 kDa; bovine serum albumin, 66 kDa; ovalbumin, 45 kDa) versus elution time.

Molecular Biology Procedures

RNA Isolation

Total RNA was isolated from fresh tissue from each species using the TRIZOL Reagent (Life Technologies) extraction protocol. Body wall muscle from one large Ciona was homogenized in 1 ml of TRIZOL Reagent per 50–100 mg of tissue using a Brinkmann Instruments (Westbury, N.Y.) Polytron. The cellular debris was pelleted and the supernatant containing the RNA was subjected to chloroform extraction and further isopropanol purification. A final centrifugation step pelleted the RNA, and this sample was aliquoted among several microfuge tubes and stored at −80°C in 75% ethanol pre-
pared in diethyl pyrocarbonate (DEPC)-treated water. One working tube contained RNA stored in DEPC water. Five whole Branchiostoma were homogenized, and RNA was isolated once again by the TRIZOL Reagent protocol.

**RT/PCR**

Reverse transcription reactions were carried out using total RNA as the template and a lock-docking-DT primer (5’-GGCCACGGTG CACTAGTACT 17(A,C,G)(A,C,G))-3’) to produce cDNA products (Borson, Salo, and Drewes 1992). This procedure was carried out using Ready-to-Go You-Prime First-Strand Beads (Amersham-Phar-macia Biotech).

**Rapid Amplification of cDNA Ends Amplification of the cDNA for Ciona CK**

Rapid amplification of cDNA ends (RACE) protocols were based on those of Frohman, Dush, and Martin (1988) using kits from Life Technologies (Grand Isle, N.Y.). For 3’ RACE PCR amplification of the cDNA of Ciona, the lock-docking primer was used to generate the first-strand cDNA. Subsequently, a “universal” forward primer (5’-GT(A,C,G,T)TGGA(A,G)T(A,C,G,T)AA(T,C) GA(A,G)GA(A,G)GA(T,C)-CA-3’) was used in PCR amplifications of the cDNA. This primer corresponds to the amino acid sequence IWV/INEEDH, which is highly conserved in all phosphagen kinases (Suzuki and Furuokohri 1994; Suzuki et al. 1997, 1998, 1999). PCR amplification using a touchdown procedure was performed under the following conditions: 1 cycle of 95°C for 5 min; 2 cycles of 94°C for 1 min, 56°C for 1 min, and 72°C for 1.5 min; and two cycles of 94°C for 1 min, 54°C for 1 min, and 72°C for 1.5 min; the next cycles had the same denaturation and elongation temperatures and times, 94°C for 1 min and 72°C for 1.5 min, respectively. The annealing temperature dropped by two degrees every two cycles until it reached 48°C; then it dropped to 45°C for two cycles, and then it went back up to 50°C for the remaining 16 cycles. A final extension of one cycle of 72°C for 7 min ended the program. This procedure was carried out in a Hybaid Omn-E (Middlesex, U.K.) thermal cycler using Taq DNA polymerase (Life Technologies).

This ~550-bp product was then gel-purified using the Concert Rapid Gel Extraction System (Life Technologies), and the product was used in a subsequent PCR amplification utilizing the same touchdown protocol as before. This product was once again gel-purified and sequenced using an automated PE-ABI model 373A sequencer. The primer used for sequencing was the “universal” NEEDH primer.

Another product, of 650 bp, was identified when twice the amount of cDNA template was used in the PCR reaction. This product was digested with SmaI for blunt-end subcloning into the puC19 TA cloning vector (Life Technologies) and sequenced.

A gene specific primer (CK1: 5’-TGCAACCGGT-CTTCAACTG-3’) located about 60 nt downstream of the NEEDH region was synthesized. The gel-purified product which was used for direct sequencing was further amplified with the 3’ RACE protocol using CK1 and a different touchdown PCR protocol under the following conditions: one cycle of 94°C for 5 min and two cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 1.5 min; the next cycles once again had the same denaturation and elongation temperatures and times of 94°C for 1 min and 72°C for 1.5 min, but every two cycles the annealing temperature would drop by 2°C until 38°C was reached. The next two cycles were conducted at 35°C for 1 min, and the following 16 cycles were at 43°C for 1 min. There was a final extension of 72°C for 7 min.

A 650-bp product was generated and subcloned into a puC19 TA cloning vector. The cloning reaction was transformed into Invitrogen One Shot competent cells and plated on kanamycin plates overnight at 37°C. Positive clones were identified as white colonies and were screened on a lysis gel to check for inserts. Plasmid DNA was then isolated for the selected colonies using Promega (Madison, Wis.) Wizard Mini Prep. The plasmids were then digested with BamHI and EcoRI. Three independent clones with the correct inserts were then sequenced using the M13 forward and reverse primers.

The 5’ RACE amplification of Ciona cDNA yielded three clones approximately 900 bp in length. Two sets of 5’ RACE gene-specific primers were generated from the partial sequence obtained from the 3’ RACE analysis: CK2-1 (5’-ACCAGCTGCAACCTGTG-3’) and CK2-2 (5’-CTGTCGGAGGTCTGCTG-3’), and CK7-1 (5’-CTCACACACATCGATGACC-3’), and CK7-2 (5’-CTGCAGAGCTACCTTTG-3’). PCR analysis was conducted by mixing and matching these primer pairs. Inserts of the correct size were finally generated using the CK7 set of gene-specific primers for an initial 5’ RACE and then using CK2-2 in a subsequent reaplification.

The following PCR protocol was followed using the abridged anchor primer and CK7-2: 1 cycle of 94°C for 2 min; 30 cycles of 94°C for 15 s, 57°C for 30 s, and 72°C for 1 min; and a final extension of 72°C for 7 min. This PCR product was diluted 1:1,000 and underwent a second round of PCR amplification using an abridged universal anchor primer and CK2-2 under the following conditions: 1 cycle of 94°C for 2 min; 5 cycles of 94°C for 30 s, 60°C for 1 min, and 72°C for 1 min; 5 cycles of 94°C for 30 s, 55°C for 1 min, and 72°C for 1 min; 15 cycles of 94°C for 30 s, 50°C for 1 min, and 72°C for 1 min; and a final extension of 72°C for 7 min.

This product was gel-purified using the Concert Rapid Gel Extraction System and run on a gel to extract the corresponding fragment. The TOPO-XL kit protocol (Invitrogen, Carlsbad, Calif.) was used for extraction and cloning of this product. Three independent clones were identified and sequenced.

**RACE Amplification of the cDNA for Branchiostoma CK**

We performed 3’ RACE on cDNA of Branchios- toma utilizing the same forward “universal” phospha-
gen kinase primer as for Ciona in conjunction with the lock-docking primer used to generate the first-strand cDNA. The PCR protocol used was the same touchdown used to generate the 3′ end of Ciona. A 1,000-bp fragment was generated, gel-purified, and subcloned into the pUC19 TA cloning vector. Three clones were identified with positive inserts, and plasmid purification was achieved. These clones were sequenced.

We used 5′ RACE to determine the sequence for the 5′ end of Branchiostoma CK cDNA. Gene-specific primers were constructed using the sequence obtained from 3′ analysis. The primer sequence used for the initial generation of cDNA by RT was CK6-1 (5′-CA-GATGCTGGTTCCACATGTA-3′). The nested primer used in subsequent PCR analysis was CK6-2 (5′-TACTCGTGGCCCTTCTTCTTC-3′). This yielded an 850-bp product with a PCR protocol as follows: 1 cycle of 94°C for 2 min; 5 cycles of 94°C for 0.5 min, 55°C for 1 min, and 72°C for 1 min; 5 cycles of 94°C for 0.5 min, 53°C for 1 min, and 72°C for 2 min; 20 cycles of 94°C for 0.5 min, 51°C for 1 min, and 72°C for 2 min; and a final extension at 72°C for 7 min. This product was gel-purified and subcloned into pUC19 TA cloning vector. Two positive clones were obtained from this and sequenced. One additional clone was generated using Taq extension for 30 min, purifying the product with the Wizard PCR Prep vacuum protocol, and then proceeding with the TOPO-XL cloning and transformation utilizing electroelution. Transformants were plated on kanamycin plates, and clones were selected for plasmid purification. Clone 400-48 was purified and sequenced.

Computer Analyses

These data were analyzed using the GCG (Genetics Computer Group, Madison, Wis.) software package, version 9.1. Multiple-sequence alignments and phylogenetic analyses were studied. The PILEUP program constructed these multiple-sequence alignments and phosphagen kinases from lower chordates, invertebrates, and vertebrates used for comparisons. Manual adjustments in sequence alignments were made where needed. Tree reconstructions were performed by the neighbor-joining (based on pairwise amino acid distances) and parsimony methods using the Unix version 4.0d55 of PAUP as implemented in GCG. Bootstrapping analyses were conducted (1,000 replicates) using the basal taxa from each major cluster as outgroups (Neanthes GK, Eisenia LK, sea cucumber AK, both domains of sea anemone AK) to obtain bootstrap values for each node after conversion of the tree to midpoint rooting. In the case of parsimony analysis, the topology of the tree is based on the 50% majority-rule consensus tree of five equally parsimonious trees.

Results

The 3′ halves of the cDNAs for the CKs of Ciona and Branchiostoma were amplified by PCR with the phosphagen consensus primer and the lock-docking primer. Two new primers for the CKs from each organism were designed for amplification of the 5′ half. Three independent clones from the 5′ RACE and three independent clones from the 3′ RACE amplifications were produced and subsequently sequenced. The complete nucleotide and deduced amino acid sequences for Ciona and Branchiostoma CKs were thus determined (figs. 1 and 2, respectively). The nucleotides are numbered at the left of each row, above the numbers for amino acids. The termination codon is denoted by an asterisk, and the putative polyadenylation signal (AATAAA) is underlined. This polyadenylation signal is 14 nt from the polyA tail in Ciona and 23 nt from the polyA tail in Branchiostoma.

The cDNA for C. intestinalis CK consists of 1,371 bp. The open reading frame of 1,131 nt is flanked by a 5′ untranslated region of 72 bp and a 3′ untranslated region of 168 bp (fig. 1). The mature protein consists of 377 amino acids with an estimated isoelectric point (pI) of 5.9 and a calculated molecular mass of 42.4 kDa. The native molecular mass was calculated from the CK elution profile from FPLC against standards and was found to be 83.4 kDa. Thus, this CK is clearly dimeric. The above M<sub>r</sub> values are consistent with known CKs from other organisms. CK activity in Ciona averages 28.2 μmol/min/g wet weight of tissue.

The cDNA for Branchiostoma CK consists of 1,717 bp. The open reading frame of 1,137 nt is flanked by a 5′ untranslated region of 68 bp and a 3′ untranslated region of 512 bp (fig. 2). The mature protein consists of 379 amino acids with an estimated pI of 6.1 and a calculated molecular mass of 42.5 kDa. The native molecular mass was calculated from the CK elution profile obtained from FPLC analysis against standards and found to be 88.9 kDa. The above data are consistent with dimeric quaternary structure. CK activity averages 350 μmol/min/g wet weight of tissue.

Figure 3 shows a multiple-sequence alignment of selected dimeric CKs, as well as the third domain of sea urchin contiguous trimeric CK. This figure is a section of a much larger alignment of 26 phosphagen kinase sequences, including 21 CKs, a GK, an LK, and 3 AKs created using the program PILEUP of the GCG software package (for the entire alignment, see Graber [2000]). The alignment of dimeric CKs (fig. 3) shows extreme conservation of residues in comparing the Ciona and Branchiostoma CKs with other CKs including that of the polychaete which last shared a common ancestor with these animals roughly 700 MYA (Doolittle et al. 1996).

This conservation is readily apparent in a comparison of percentages of amino acids of these two lower chordate CKs identical to those of other phosphagen kinases using the program OldDistances of the GCG software package (table 1). Ciona and Branchiostoma CKs display similar and high percentage identities to polychaete (Chaetopterus) dimeric CK, as they do to various M, B, and fish dimeric CKs. A somewhat lower percentage of identity was observed for a domain of the sea urchin contiguous trimeric CK (table 1).

Phylogenetic analyses were conducted on the above multiple-sequence file of the 26 representative phosphagen kinases using PAUP as implemented in the GCG
Evolution of Cytoplasmic Creatine Kinases

Fig. 1.—cDNA and deduced amino acid sequences for CK from the tunicate Ciona intestinalis. These sequences have been deposited in GenBank (accession number AF258619).

software package. Both domains of the “contiguous dimeric” sea anemone AK constitute the outgroup cluster. Neighbor-joining (NJ) analysis is shown in figure 4. It has been previously shown that sea cucumber AK evolved secondarily from a CK-like ancestor and that GK and LK originated in the CK clade (Suzuki et al. 1999). All of these forms and the CKs are present in a large supercluster distinct from the largely monomeric AKs found in the protostome or lower groups (Suzuki et al. 1999).

The large supercluster in figure 4 shows that mitochondrial CKs diverged early and that the formation of the sarcomeric and ubiquitous mitochondrial isoforms occurred after the protostome-deuterostome split. The numbers represent the percentages of 1,000 bootstrap replicates which support the branch. The clustering of trout CK with the B-CKs and the Torpedo and carp CKs with the M-CKs is consistent with previous analyses (MuÈhlebach et al. 1994; Pineda and Ellington 1999). The clustering of the M and B clades is well supported. Of great interest is the fact that the NJ tree shows that both Ciona and Branchiostoma CKs diverged before the duplication event that produced the M and B isoforms (fig. 4).

Discussion

From a kinetic standpoint, the various cytoplasmic and mitochondrial isoenzymes of CK display minimal differences in properties (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992). These CKs do differ in terms of cellular expression, as well as in terms of localization within the cell. Thus, major structural differences appear to be related to the nature and capacity for intracellular targeting (Wallimann et al. 1992).
In the case of the cytoplasmic isoenzymes, M-CK is expressed in skeletal muscle fibers, and 5%–10% of the CK protein is not soluble but is specifically bound to myofibrillar M-line (Turner, Wallimann, and Eppenberger 1973). Myofibrillar M-CK rapidly phosphorylates ADP produced by myosin ATPase, thereby creating an efficient ATP regenerating system (Wallimann, Schloesser, and Eppenberger 1984; Ventura-Clapier, Veksler, and Hoerter 1994). The other cytoplasmic isoenzymes (MB, BB) are incapable of such binding (Wallimann, Moser, and Eppenberger 1983; Schafer and Perriard 1988). Interestingly, a significant amount of B-CK is not soluble and is associated with membrane-bound ion transport ATPases in such cells as neurons and glia (Kaldis et al. 1996).

M- and B-CKs are present in all higher vertebrates thus far examined. Phylogenetic analyses show that trout CK is more closely related to the B isoform, while Torpedo and carp CKs are more closely related to the M isoform (see figs. 4 and 5). As indicated previously, CK is widely distributed in the invertebrates and lower chordates. Sponges, the most primitive of all metazoans, express a dimeric CK which shows sequence homology to CKs from higher forms (Ellington 2000). The cytoplasmic dimeric CK from the polychaete *C. variopedatus* displays roughly the same percentage of identical amino acids to the B and M isofoms (Pineda and Ellington 1999; Pineda 2000) and clusters outside of the M/B supercluster (figs. 4 and 5). Thus, the gene duplication event producing the cytoplasmic isoforms of CK in vertebrates occurred within the deuterostomes.

The topology of the NJ and parsimony trees (figs. 4 and 5) deserves some comment. Suzuki et al. (1999) have shown that CKs, LK, GK, and echinoderm AK are much more similar to each other than the remaining nondeuterostome AKs. These latter AKs are mostly monomeric and form a large cluster consisting of all protostome and lower invertebrate AKs (Suzuki et al. 1999). Our present phylogenetic analyses are consistent with observations showing that LK, GK, and echinoderm AK diverged rather early. Both trees show that the mitochondrial and cytoplasmic isoforms of CK evolved before the divergence of the major metazoan groups and that the contiguous trimeric form diverged later (figs. 4 and 5). However, it has recently been shown that sperm of the polychaete *C. variopedatus* also contain this unique CK (Ellington, Roux, and Pineda 1998). Thus, this form arose early; details of the evolution of this CK isomerase are under study.
Evolution of Cytoplasmic Creatine Kinases

**Fig. 3.**—Multiple-sequence alignments of *Ciona intestinalis* and *Branchiostoma floridae* CKs with other CKs. Abbreviations are as follows: *Chaetopterus* dimCK (accession number AF139588), human BCK (accession number 180570), chicken BCK (accession number X03509), trout CK (rainbow trout, accession number X53859), human MCK (accession number M14780), chicken MCK (accession number X00954), carp M1CK (carp MCK isoenzyme 1, accession number AF055288), carp M2CK (carp MCK isoenzyme 2, accession number AF055289), carp M3CK (carp MCK isoenzyme 3, accession number AF055289), *Torpedo* CK (accession number P00566), and sea urchin D3 (sea urchin domain 3, accession number 161473).

**Table 1**

Percentages of Identical Amino Acid Sequences for Selected Creatine Kinases (CKs)

<table>
<thead>
<tr>
<th>CK</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <em>Ciona</em> CK</td>
<td>66.31</td>
<td>66.67</td>
<td>63.81</td>
<td>70.29</td>
<td>68.70</td>
<td>67.94</td>
<td>66.79</td>
<td>69.50</td>
<td></td>
</tr>
<tr>
<td>2. <em>Branchiostoma</em> CK</td>
<td>74.19</td>
<td>62.73</td>
<td>63.81</td>
<td>69.13</td>
<td>69.39</td>
<td>69.35</td>
<td>62.73</td>
<td>80.58</td>
<td></td>
</tr>
<tr>
<td>3. <em>Chaetopterus</em> dimCK</td>
<td>63.71</td>
<td>70.29</td>
<td>68.28</td>
<td>70.29</td>
<td>68.28</td>
<td>68.28</td>
<td>70.29</td>
<td>68.28</td>
<td></td>
</tr>
<tr>
<td>4. Sea urchin D1 CK</td>
<td>69.13</td>
<td>69.39</td>
<td>69.35</td>
<td>62.73</td>
<td>70.29</td>
<td>68.28</td>
<td>68.28</td>
<td>68.28</td>
<td></td>
</tr>
<tr>
<td>5. Chicken B-CK</td>
<td>69.13</td>
<td>69.39</td>
<td>69.35</td>
<td>62.73</td>
<td>70.29</td>
<td>68.28</td>
<td>68.28</td>
<td>68.28</td>
<td></td>
</tr>
<tr>
<td>6. Chicken M-CK</td>
<td>69.13</td>
<td>69.39</td>
<td>69.35</td>
<td>62.73</td>
<td>70.29</td>
<td>68.28</td>
<td>68.28</td>
<td>68.28</td>
<td></td>
</tr>
<tr>
<td>7. <em>Torpedo</em> CK</td>
<td>80.58</td>
<td>78.42</td>
<td>78.48</td>
<td>78.42</td>
<td>78.48</td>
<td>78.48</td>
<td>78.48</td>
<td>78.48</td>
<td></td>
</tr>
<tr>
<td>8. Carp M3 CK</td>
<td>67.64</td>
<td>70.71</td>
<td>68.82</td>
<td>63.54</td>
<td>78.42</td>
<td>78.48</td>
<td>78.48</td>
<td>78.48</td>
<td></td>
</tr>
<tr>
<td>9. Trout CK</td>
<td>69.50</td>
<td>68.34</td>
<td>68.28</td>
<td>63.81</td>
<td>83.99</td>
<td>78.48</td>
<td>78.48</td>
<td>78.48</td>
<td></td>
</tr>
</tbody>
</table>

**Note.**—*Chaetopterus* dimCK = polychaete dimeric CK; sea urchin D1 CK = first CK domain of contiguous trimeric CK; carp M3 CK = third form of carp muscle CK.

The absence of a cytoplasmic, dimeric CK in the echinoderms is indeed puzzling, especially when viewed in the context of its apparent "replacement" by a dimeric AK (Suzuki et al. 1999) or a contiguous trimeric CK (Ratto, Shapiro, and Christen 1989; Wothe, Charbonneau, and Shapiro 1990). Thus, we turned to CKs from two lower chordates (tunicate and lancelet) to gain insight into the timing of the gene duplication event producing the M and B genes. Early electrophoretic and immunological studies of CK in fish and two lower chordates identified four distinct isofoms in fish (A, B, C, and D) which displayed a rather complicated distribution between species and a tissue-specific pattern of...
expression (Fisher and Whitt 1978a; Fisher et al. 1980). Subsequent work showed that the A and C likely correspond to the M and B loci, respectively, of higher vertebrates (Mühlebach et al. 1994). Interestingly, Fisher and Whitt (1978b) also showed that the tunicate Styleta plicata had a single CK locus which was identified as the A (M-like) locus in terms of immunological properties. Furthermore, it was shown that two species of cephalochordates appear to have both A (M-like) and C (B-like) loci (Fisher et al. 1980).

Cellulose acetate electrophoresis of crude tissue extracts of C. intestinalis (body wall muscle) and B. floridanae (whole animals) showed that the tunicate had a single CK band consistent with the earlier results of Fisher and Whitt (1978a), while the lancelet has a single major CK staining band with a very minor band of CK activity (Graber 2000). It is possible that this minor electrophoretic band in B. floridanae constitutes the second cytoplasmic CK locus that was observed by Fisher et al. (1980). A very intense 3'RACE RT-PCR product was consistently generated using our “universal” forward PCR primer and B. floridanae RNA. However, several other very minor products were observed but not characterized further (unpublished observations). Even if two CK loci are present in the lancelet, it is not altogether clear whether the second is cytoplasmic or mitochondrial; both loci were present in metazoans prior to the protostome-deuterostome divergence (Pineda and Ellington 1999).

Our analyses of the cDNA and deduced amino acid sequences for the dimeric CKs from the tunicate C. intestinalis and the lancelet B. floridanae show that these two forms diverged well before the origin of the M and B isoforms found in extant vertebrates. If we assume that only one cytoplasmic CK gene is present in tunicates such as C. intestinalis, then clearly the gene duplication event forming the major cytoplasmic isoforms of CK occurred after the divergence of this group from the cephalochordate-craniate lineage. As indicated previously, cephalochordates such B. floridanae may have two CK loci present. The presence of the second locus and its molecular characteristics (cytoplasmic or mitochondrial?) remain to be elucidated by modern approaches. If two cytoplasmic CK genes are indeed present, then the gene duplication event could have occurred within this group. This conclusion was reached by Fisher et al. (1980), who proposed a phylogenetic scheme for the cytoplasmic genes. However, our phylogenetic analyses (figs. 4 and 5) show that B. floridanae CK diverged prior...
to the appearance of the M and B loci and that this CK is slightly more similar to a protostome cytoplasmic CK than to lower chordate and vertebrate CKs (table 1).

If only one cytoplasmic CK locus is present in tunicates and cephalochordates, then our phylogenetic analyses indicate that the gene duplication event creating the M and B isoforms occurred after the divergence of the craniates from the lower chordates. If this is the case, then the gene duplication event forming the M and B genes of CK took place during the radiation of the fish, as both M- and B-like forms appear to be present in this group. There is extensive evidence for a large number of gene duplication events at other loci having taken place during the evolution of the major fish groups (Ohno 1970; Holland et al. 1994; Sharman and Holland 1996 [as cited in Stock et al. 1997]). The early electrophoretic and immunological work on fish CKs (Fisher and Whitt 1978a, 1978b; Fisher et al. 1980), coupled with the existing cytoplasmic CK sequences from fish, are insufficient to resolve exactly when this duplication took place. A promising group to look at would be agnathans. Lampreys appear to have a (M-like CK) and C (B-like CK) loci (Fisher et al. 1980). Surprisingly, Fisher et al. (1980) were unable to detect CK in the most primitive of all craniates, the hagfish (order Myxiniformes). This latter observation should be reinvestigated. Clearly, a more robust fish CK database, including representative agnath CK and additional elasmobranch and teleost CK sequences, is required to better resolve the timing of the duplication event.

Acknowledgment

This research was supported by a grant from the National Science Foundation (96-31907) to W.R.E.

LITERATURE CITED


STOLZ, M., and T. WALLIMANN. 1998. Myofibrillar interaction of cytosolic creatine kinase (CK) isoymes: allocation of


ANTONY DEAN, reviewing editor

Accepted March 16, 2001