Two tissue-specific isozymes of creatine kinase have closely matched amino acid sequences

LULU PICKERING*, HENRIANNA PANG†, KLAUS BIEMANN†, HAMISH MUNRO‡§, AND PAUL SCHIMMEL*

Departments of *Biology, †Chemistry, and ‡Nutrition and Food Science, Massachusetts Institute of Technology, Cambridge, MA 02139; and §USDA Human Nutrition Research Center on Aging at Tufts University, Boston, MA 02111.

ABSTRACT Creatine kinase activity is associated with different isozyme species. We have examined two of these: the cytoplasmic brain (B) isozyme that is expressed in many tissues and is reported to be induced by estrogen and the developmentally regulated cytoplasmic muscle (M) isozyme that is found predominantly in differentiated muscle tissue. Recently, we cloned and sequenced the cDNA for the M isozyme of rabbit creatine kinase. We now report the isolation of B-isozyme cDNAs and the deduced primary structure of the polypeptide. The translated cDNA nucleotide sequence was cross-checked by fast-atom bombardment/mass spectrometry of tryptic fragments from the protein. The sequence is exactly colinear with the rabbit M isozyme and the two isozymes have 80% nucleotide and amino acid sequence identity. There are blocks of 36 and 41 amino acids where the amino acid sequence is conserved exactly. The colinearity of the two sequences and the extent of their identity makes it unlikely that either isozyme has unique polypeptide domains that account for specialized functions. The rationale for the existence of these creatine kinase isozymes, with distinct biological features, evidently is at the level of regulation of individual isozyme expression.

Creatine kinases belong to a class of enzymes designated ATP:guanidino phosphotransferases, or guanidine kinases, which reversibly store energy as phosphagens (i.e., creatine phosphate) or regenerate ATP to maintain high ATP/ADP ratios. The major guanidine kinase found in invertebrates is arginine kinase, which many times occurs in association with creatine kinase (1). Guanidino kinases have been well conserved throughout evolution and share amino acid homologies in the region surrounding a reactive cysteine (1, 2). The guanidino kinase found in vertebrates is creatine kinase. The cytoplasmic forms, muscle isozyme (M isozyme) and brain isozyme (B isozyme), form dimers with MM being the major form in skeletal muscle and myocardium, MB existing in myocardium, and BB existing in many tissues, especially brain.

The B isozyme is found in embryonic tissue. The tissue-specific expression of creatine kinase isozymes and the isozyme switch that occurs during differentiation of myoblasts to myotubules has been well documented (3–5). Less is known about the regulation of isozyme expression, although the B isozyme is reported to be induced specifically by estrogens (6). The two isozymes are, therefore, distinguished not only by their tissue specificities but also by the sensitivity of B-isozyme expression to estrogen. This raises the possibility that the two isozymes, while each catalyzing the phosphorylation of ADP with creatine phosphate, have additional roles that are not held in common. Such roles should be reflected in domains of protein structure unique to one, but not the other, isozyme. The two isozymes are, indeed, immunologically distinguishable (2).

We recently cloned the cDNA for the rabbit M isozyme and established the primary structure of the protein (7). Upon the assumption that the two isozymes have some sections that are similar in sequence (e.g., in and around the

FIG. 1. RNA blot of creatine kinase mRNAs. RNA was isolated from rabbit muscle or brain by guanidinium thiocyanate precipitations (9), followed by two cycles of oligo(dT)-cellulose chromatography (10). Total muscle RNA (15 μg) or poly(A)+ brain mRNA (2.5 μg) was treated with glyoxal and dimethyl sulfoxide for RNA transfer blot analysis (8). Hybridization probes were produced by nick-translation to specific activities of >106 Cerenkov cpm/μg (11). Stable M-isozyme cDNA-B-isozyme mRNA hybrids (middle lane) were easily detected with a hybridization stringency of 30% formamide, 0.9 M NaCl/50 mM sodium phosphate, pH 7.4/5 mM EDTA, 5× concentrated Denhardt’s solution (100× concentration = 2% Ficoll/2% polyvinylpyrrolidone/2% bovine serum albumin), and 0.5% NaDodSO4 at 37°C for 12 hr, followed by washing the filters with 18 mM NaCl/1 mM sodium phosphate, pH 7.4/0.1 mM EDTA to 0.18 M NaCl/10 mM sodium phosphate, pH 7.4/1 mM EDTA and 0.1% NaDodSO4 at 37°C for 1 hr. Positions of end-labeled Ava II digests of pBR322 and 18S and 28S rRNAs run as standards in parallel lanes are indicated. Creatine kinase mRNA at 1600 nucleotides is indicated by the arrow. The higher molecular weight mRNA (1800 nucleotides) species in the middle lane may represent a mitochondrial creatine kinase mRNA species (12).

Abbreviations: FAB/MS, fast-atom bombardment/mass spectrometry; B isozyme, brain isozyme; M isozyme, muscle isozyme; bp, base pair(s).
active site), we used cDNA probes from the M isozyme to screen brain poly(A)⁺ RNA and its cognate cDNA library. This led to isolation of the B isozyme cDNA, which enabled us to address the issues raised above.

RESULTS AND DISCUSSION

RNA Blots of Brain Poly(A)⁺ RNA with a M-Isozyme cDNA Probe. Previously, our laboratory cloned and sequenced the cDNA for the M isozyme of rabbit creatine kinase (7). To use as a hybridization probe, we chose a M-isozyme cDNA region that encodes a reactive cysteine moiety that is known to be conserved between M and B isozymes of various creatine kinases (2). This region is contained in a 180-base-pair (bp) restriction fragment of M-isozyme cDNA clone pCKM19 (7). To establish the hybridization conditions that would allow formation of stable cDNA-RNA hybrids, we performed RNA blots (8) (under various stringency conditions) with the M-isozyme probes (plasmid pCKM19 or the 180-bp fragment of pCKM19) hybridizing to rabbit muscle or brain RNAs. A brain poly(A)⁺ RNA species was detected that has a size of about 1600 nucleotides (Fig. 1). This is also the size of the M isozyme mRNA that is detected by the same probe.

Isolation of Overlapping B-Isozyme cDNA Clones. Stringency conditions that gave stable hybrids were used to screen by colony hybridization a cDNA library (in pBR322) constructed from rabbit brain poly(A)⁺ RNA. Of 25,000 colonies, a total of 6 positive creatine kinase clones was detected, which suggests the abundance of B isozyme as ~0.024% of expressed brain mRNAs. Brain creatine kinase clone pCKB17 hybridizes to a 1600-nucleotide mRNA as does the muscle creatine kinase clone (Fig. 1). This clone contains an insert of about 950 bp. DNA sequencing showed that this insert encodes the reactive cysteine moiety that occurs in a group of 12 conserved amino acids that are present in both M and B isozymes. This 12-residue stretch is reported to have a single-position conservative substitution of isoleucine (in the

![Image of cDNA clones and sequencing subclones](image-url)

**Fig. 2.** Overlapping creatine kinase B-isozyme clones. Clone pCKB17 was isolated from a brain cDNA library by colony hybridization (14). The library was produced by converting 20 μg of poly(A)⁺ mRNA to double-stranded cDNA, followed by treatment with S1 nuclease, size fractionation on Sepharose CL-4B, and tailing with dC residues (15). This yielded 6.6 μg of cDNA, 400 ng of which was annealed to 1.6-μg Pst I-linearized dG-tailed pBR322 (Bethesda Research Laboratories) and transformed into Escherichia coli strain C600 (16) to produce a library of 20,000–25,000 tetracycline-resistant colonies. pCKB2-3E and pCKB4C-2 were produced by primer extending creatine kinase B-isozyme mRNA. Single-stranded B-isozyme-specific primers (bars) were prepared from M13 subclones and annealed to 22 μg of poly(A)⁺ mRNA (5.5 ng of B-isozyme mRNA assuming an abundance of 1:4000) in 90% formamide for 10 min at 50°C, followed by 3 hr at 50°C in 80% formamide/10 mM Pipes, pH 6.4/400 mM NaCl/1 mM EDTA (17). The annealed primer-mRNA hybrids were extended with reverse transcriptase, and the cDNA-mRNA hybrids were eluted from a 5% Tris-HCl/borate/EDTA/polyacrylamide gel, tailed with dC residues, annealed with Pst I-linearized dG-tailed pBR322 (Bethesda Research Laboratories), and transformed into E. coli C600 (18). Positive colonies were detected by colony hybridization (14). Subclones of B-isozyme cDNAs produced in the bacteriophage M13 for sequencing both sense and antisense cDNA strands are indicated by the arrows. The B-enzyme map is shown at the bottom. Restriction enzyme sites conserved between M and B isozymes are shown below the solid line, whereas those specific to B isozyme are shown above the line.
B isozyne) for valine (in the M isozyme), which is located 3 amino acids on the NH2-terminal side of the cysteine residue (2). The same substitution is found in pCKB17 and this finding gave strong evidence that we had isolated an authentic B-isozyne cDNA fragment. Further DNA sequencing showed that clone pCKB17 contains the COOH-terminal coding region of the protein, a 3'-untranslated region of 194 nucleotides, a polyadenylation consensus sequence (13), and an approximate 100-nucleotide poly(A) tail.

The NH2-terminal coding region of the isozyne is missing from pCKB17 as is the case for all of the positive clones detected in the library (data not shown). To obtain the complete coding region, brain mRNA was primer extended with internal B-isozyne primers derived from pCKB17, and clones pCKB4-2 and pCKB2-3E were subsequently generated from the primer extension products. The complete coding region of the B isozyne was obtained by overlapping sequences of clones pCKB4-2, pCKB2-3E, and pCKB17. Fig. 2 gives a diagrammatic representation of the 3 cDNA clones, their restriction maps, and the strategy for sequencing.

**Nucleotide and Polypeptide Sequence.** The complete nucleotide sequence of the B isozyne and its translated amino acid sequence are shown in Fig. 3. Where indicated, the sequence has been confirmed by peptide data obtained by fast-atom bombardment/mass spectrometry (FAB/MS) (24). The translated protein is 381 codons in length and is exactly colinear with the rabbit M isozyne. Note that clone pCKB4-2 extends only 5 nucleotides to the 5' side of the ATG initiator codon. Our designation of the NH2 terminus of the B isozyne is based on the exact colinearity with the M isozyne and by the presence of a eukaryotic translation initiation consensus sequence (25) at the ATG position (C-C-G-C-C-A-T-G-C). The molecular weight computed from the sequence is 42,530, which closely agrees with the value found experimentally (2). The reactive cysteine is codon 283 in both isozynes and codon 280 is isoleucine in the B isozyne and valine in the M isozyne.

The 3'-untranslated sequence of the B isozyne is 194 nucleotides in length compared to 271 nucleotides for the M isozyne. Unlike the coding regions of the two isozynes, the 3'-untranslated regions are not colinear and are not homologous. They should be useful as hybridization probes for studying tissue-specific isozyne expression.

Fig. 3. Nucleotide and amino acid sequences of creatine kinase B isozyne. M13 DNA templates were prepared (19) and the sequence of the B isozyne was determined by the method of Sanger (20) by using 35S-labeled dATP and 50-cm polyacrylamide gradient gels (21). DNA sequence data were analyzed by the ANALYSEQ programs of Staden (22, 23). Rabbit creatine kinase B isozyne purchased from Sigma (<50% pure) was purified further on DEAE-cellulose (Whatman DE-52) columns with very low yield. Purified protein was carboxymethylated, cleaved by trypsin (with some chymotryptic activity), and analyzed by FAB/MS (24) to determine the molecular weights of the fragments. Peptides confirmed by this analysis are indicated by the solid lines. These fragments were also subjected to one or two cycles of Edman degradation and the shortened peptides were analyzed again by FAB/MS (24) (solid arrows beneath the solid lines). Tryptic peptides 157-172 and 321-341 were detected only after the first Edman step. The tryptic peptide beginning with Val-Leu has a molecular weight of 1320, which fits the region 33-43 with tryosine in position 34, in disagreement with the DNA data as indicated by the dotted line. A peptide sequence that spans codons 280-292 and includes the reactive cysteine is confirmed by published data [solid arrows (2)]. A more thorough confirmation of the amino acid sequence has been hampered by the inability to obtain sufficient quantities of pure protein. The location of a polyadenylation consensus site, in the 3'-untranslated region, is underlined.
Comparison of Sequences of Isozymes. The nucleotide and deduced amino acid sequences of the M and B isozymes are compared in Fig. 4. Between the isozymes there are 215 nucleotide changes affecting 150 codons. Of these, 71 are silent codon changes due primarily to single nucleotide changes in the third positions of the respective codons. As a result, the isozymes are 80% identical at the protein level. Similarly, there is an 80% conservation of nucleotide sequence so that even some of the restriction enzyme sites are conserved (Fig. 2).

Due to the extensive amino acid homology, both isozymes have similar amino acid compositions, with about equal proportions of aromatic, hydrophobic, acid, amide, and hydroxyl residues. The observed difference in electrophoretic mobilities of MM, MB, and BB creatine kinase hybrids (2) is due mostly to 14 additional lysine residues in the M versus the B isozyme. Inspection of the amino acid sequences indicates that 10 tryptic fragments of a total of 38 for B isozyme and 46 for M isozyme have identical sequences, which further emphasizes the extent of homology and circularity in the differences that have been evolutionarily conserved for these rabbit isozymes.

The longest stretch (41 amino acids) of identical amino acids is located approximately at the center of the molecule and spans codons 183–222. Another long match (36 amino acids) occurs in the NH₂-terminal part of the polypeptide and runs from codons 47 to 82. Note also that the reactive cysteine codon (283) is in a region that has been shown to be conserved in guanido kinases (2, 2) and is located in a stretch of 24 amino acids in the COOH-terminal third of the sequence. Of these 24 amino acids, 21 are identical and the other 3 represent conservative changes (2 are isoleucine to valine and 1 is alanine to glycine). This region of homology is flanked on either side by areas containing many changes.

Although the two proteins show strong conservation of sequence, they are readily distinguished by polyclonal antibodies. This is the basis for diagnostic assays of the three possible dimeric species made of the two subunits (26). Antigenic determinants are located on protein surfaces where hydrophilic residues are concentrated (27). It is of interest that 9 of the 15 substitutions that give a lysine in the M isozyme are immediately adjacent to at least one additional amino acid substitution. These provide obvious targets for antigenic variability.

Recently, the cDNA sequence for rat muscle creatine kinase has been obtained (28). At the nucleotide level it is about 91% homologous to the rabbit M isozyme and 98% homologous to the amino acid level. There are only 6 amino acid changes between the M-isozyme sequences of rat and rabbit. These occur in regions that also show several changes between rabbit M- and B-isozyme sequences, but not within the highly conserved regions. Thus, these regions
may be variable regions for the creatine kinases, in general. Note also that the M isozymes between rabbit and rat are more highly conserved than are the M and B isozymes of rabbit.

Creatine Kinase mRNAs. From RNA blots, the size of both mRNAs is 1600 nucleotides. Previously, we reported the sequence of the M-isozyme cDNA (7), which included 54, 1140, and 271 nucleotides of 5'-untranslated, coding, and 3'-untranslated sequences, respectively, and probably about 100 nucleotides of a poly(A)^+ tail. This would make the mRNA the same size as the cDNA. For the B isozyme, the 3'-untranslated region is only 194 nucleotides, but it has the same size coding region of 1140 nucleotides and a poly(A)^+ tract of at least 100 nucleotides. This suggests, therefore, that the 5'-untranslated region is ~150 nucleotides long. The sequence of this region will be obtained by sequencing the 5' ends of genomic B-isozyme clones.

By RNA blot analysis with coding region probes, it is clear that both muscle and brain tissues express homologous creatine kinase genes. Due to the significant homology between M- and B-isozyme sequences, the best probes for studying tissue-specific expression of either isozyme will come from the nonhomologous 3'-untranslated regions and potentially from 5'-untranslated sequences. In RNA blots, coding region probes also hybridize to an RNA species larger than the 1600-nucleotide cytoplasmic creatine kinase mRNAs (Fig. 2). This mRNA of 1800 nucleotides may represent a mitochondrial creatine kinase mRNA in brain that is closely related to the cytoplasmic forms (12).

CONCLUDING REMARKS

We found no significant blocks of coding sequence where the two isozymes diverge greatly in sequence, even though the 3'-untranslated regions are completely divergent. From this we feel that, in spite of their distinguishable tissue specificities and the estrogen sensitivity of B-isozyme expression in the uterus, the two isozymes are unlikely to have special functions that are unique to one but not the other. The explanation for the unique biological features of each tissue-specific isozenzym evidently is due to a specialized framework for expression of each isozyme gene. What is now required is a fuller exploration of their biological roles and of the molecular basis for their tissue-specific expression.

Note Added in Proof. Inspection of the cDNA sequence suggests that stable hairpins in the mRNA may have arrested reverse transcription at defined points. These hairpins might account for the 5'-end points of cDNA clones pCKB17 and pCKB2-3E (Fig. 2).

We are indebted to Dr. Nancy Royal for the purification of commercial rabbit brain creatine kinase and assistance during the preliminary phases of the work. We thank Dr. Will Gilbert for assistance with the computer analysis. This work was supported by grants from the American Heart Association and the Muscular Dystrophy Association (P.S.) and by Grants GM05472 and RR00317 from the National Institutes of Health (K.B.). Lulu Pickering was a Postdoctoral Fellow of the Massachusetts Division of the American Cancer Society.