Vol. 138, No. 1, 1986 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 16, 1986 Pages 392-398

THE COMPLETE NUCLEOTIDE SEQUENCE OF CANINE BRAIN B CREATINE KINASE mRNA: HOMOLOGY IN THE CODING AND 3' NONCODING REGIONS AMONG SPECIES

Joseph J. Billadello*, Daniel P. Kelly*, Dragos G. Roman+, and Arnold W. Strauss+‡

Departments of *Internal Medicine, +Pediatrics,
and ‡Biological Chemistry,
Washington University School of Medicine,
St. Louis, MO 63110

Received May 15, 1986

SUMMARY: To define the structure of canine B creatine kinase, clones were isolated from a library prepared from dog brain mRNA and constructed in the vector λgt11. The entire coding portion, the complete 3' nontranslated region, and 43 bp of the 5' noncoding region are reported. Comparison of the predicted amino acid sequence of canine B creatine kinase with the sequence of canine M creatine kinase shows 81% identity. When compared to cDNAS encoding B creatine kinase isolated form other species unusual and striking nucleotide sequence identity in the 3'noncoding region is present. Moreover, two B creatine kinase clones (BCK2 and BCK38) demonstrate microheterogeneity within the 3' nontranslated region indicating variable processing of B creatine kinase pre-mRNA or the existence of multiple genes encoding canine B creatine kinase.

Creatine kinase (EC 2.7.3.2), a pivotal enzyme of cellular energy metabolism, occurs as three cytoplasmic isoenzymes as well as a mitochondrial form (1). Each cytoplasmic form is a dimer composed of M and/or B subunits. Mitochondrial creatine kinase is composed of two identical subunits immunologically distinct from the cytoplasmic isoenzymes (2). Immunologic and amino acid sequence data suggest that at least three separate genes encoding the subunits exist.

The characterization of isoenzymes of CK has proven to be of clinical as well as biochemical interest. Measurement of the activity of MB in plasma serves as a specific marker of the course and extent of myocardial infarction in patients (3). Expression of isoenzymes of creatine kinase is developmentally and hormonally regulated in various tissues. In skeletal muscle undergoing maturation M content increases as the B isoenzyme declines to undetectable levels in adult muscle (4).

0006-291X/86 $1.50

Recently we reported the complete nucleotide sequence of dog heart M creatine kinase mRNA and noted a very high degree of sequence identity between M subunits from different species (5). Preliminary data revealed extensive amino acid sequence identity between the canine M and B subunits, a finding which was not previously suspected (6).

To determine the extent of homology between the M and B subunit and in order to understand the regulatory processes determining the tissue specificity and developmental expression of isoenzymes of creatine kinase and their post-translational processing to isoforms we isolated cDNA clones encoding B creatine kinase from a library constructed from dog brain mRNA. We report the complete sequence of canine B cDNA and present a detailed description of extensive homology among vertebrate species, including both the coding and 3' nontranslated regions of the mRNA.

METHODS

Total RNA was prepared from adult canine brains by the proteinase K technique (7). Poly(A)+RNA was selected by chromatography on oligo (dT)-cellulose (8). Double-standard cDNA was prepared as described (9). The cDNA was ligated to λgt11 vector DNA, which had been digested with EcoRI and treated with phosphatase, using T_4 DNA ligase. The recombinant DNA was packaged into phage and plated on E. coli strain Y1088.

The library was searched for B cDNA clones with a synthetic oligonucleotide probe corresponding to amino acids 368-380 of dog B which had been determined by automated Edman degradation (10). Further screening was performed using restriction fragments derived from cDNA clones encoding canine myocardial M creatine kinase characterized in our laboratory (5). This strategy was based on our preliminary observations of amino acid identity between the M and B subunit of canine creatine kinase and the assumption of similar nucleotide sequence identity. Plaque hybridization was done at 67°C using 6XSSC 5x Denhardt's solution and 100 µg/ml denatured, sonicated salmon sperm DNA.

Restriction fragments were subcloned into the M13 phage vectors mp18 and mp19. DNA sequencing reactions were performed using the dideoxynucleotide chain termination procedure (11). Either the universal M13 sequencing primer or synthetic oligonucleotides derived from B creatine kinase cDNA were used for sequence determinations.

B creatine kinase was purified from dog brain to a specific activity of 650 IU/mg protein (12). After cleavage with cyanogen bromide, trypsin, or Staphylococcal V8 protease, peptides were separated by reverse phase HPLC and subjected to automated Edman degradation (10). Phenylthiohydantion amino acids were identified by HPLC (13).

RESULTS AND DISCUSSION

The dog brain cDNA library contained at least 500,000 independent recombinant phages. Screening with the synthetic oligonucleotide probe and

with restriction fragments derived from M creatine kinase cDNA yielded several positive signals. Three putative B clones BCK2, BCK3, and BCK38 were purified to homogeneity, subjected to restriction enzyme analysis and the complete nucleotide sequences determined. The three clones are overlapping and together comprise the entire coding sequence of canine B, the 3'nontranslated region, the poly(A) tail and 43 nucleotides of the 5' nontranslated region (Figure 1).

Comparison of the complete amino acid sequence of canine B with the sequence of canine M reported previously (5) reveals 81% amino acid identity between the M and B subunits (Table I). While this work was in progress cDNA clones and derived amino acid sequences for B subunits were

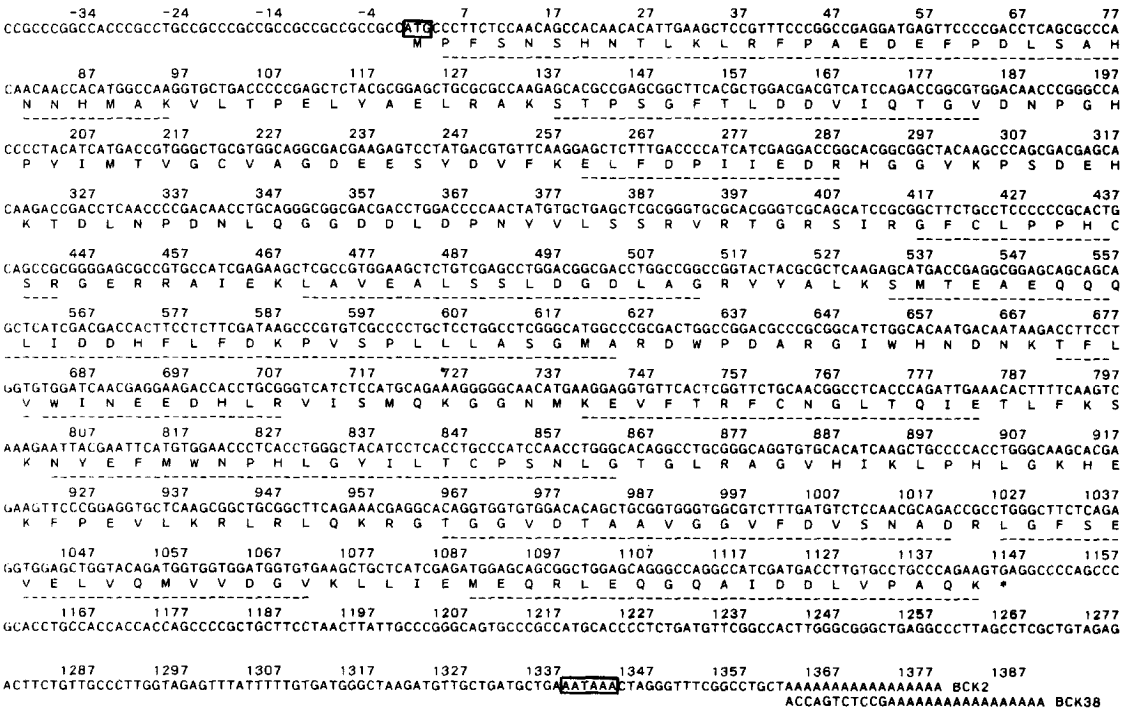

FIGURE 1: Complete nucleotide sequence of canine brain B creatine kinase with the derived amino acid sequence. The single letter code for the amino acids is shown. Numbering indicates nucleotides and begins at the initiation codon, with negative numbers to show the 5' noncoding region. The start and poly(A) addition sites are outlined with boxes. The (*) indicates the translation stop site. Predicted amino acid sequences confirmed by automated degradation of peptides derived from B creatine kinase are underlined. Clone BCK38 containes 11 additional nucleotides between the poly(A) addition site and the poly(A) tails when compared to clone BCK2 but is otherwise identical.

TABLE 1

HOMOLOGY OF CREATINE KINASES

	NUCLEIC ACID LEVEL (CODING REGION)	PROTEIN LEVEL
DOG B vs DOG M	78%	81%
DOG B vs RABBIT B	93%	97%
DOG B vs RAT B	90%	96%
DOG B vs CHICK B	81%	90%

reported for rabbit (14), rat (15) and chick (16). Comparison of the derived amino acid sequences of the B subunits shows sequence identity ranging from 90 to 97% (Table 1). More than half of the 48 amino acid substitutions between the B subunits occur within the first 90 residues. However, most of the substitutions are conservative and only 12 result in a change in charge. It should be noted that all protein sequences of M and B subunits reported to date are of equal length.

We have previously shown a high degree of nucleotide sequence identity in the coding region of M subunits among different species (5). However, the M subunit nucleotide sequences in the 3' nontranslated region are highly variable (5). Comparison of the nucleotide sequences between the M and B subunits of canine creatine kinase shows striking sequence identity in the coding region (Table 1). In contrast the 3' noncoding regions are of different lengths and are not homologous. Comparison of the nucleotide sequences of the B subunit mRNAs of different species shows even more striking (81-93%) nucleotide conservation which is observed throughout the coding region. Although the 3' noncoding regions differ slightly in length there is extraordinary nucleotide sequence identity in this region (Figure 2).

Comparison of the complete nucleotide sequences of clones BCK2 and BCK38 demonstrate that the coding regions are identical. The only difference between the two clones is in the 3' nontranslated region. The presence of 11 additional bp of sequence in clone BCK38 between the polyadenylation signal and the poly(A) tails indicate that there may be microheterogeneity of the site of cleavage-polyadenylation of B pre-mRNA.

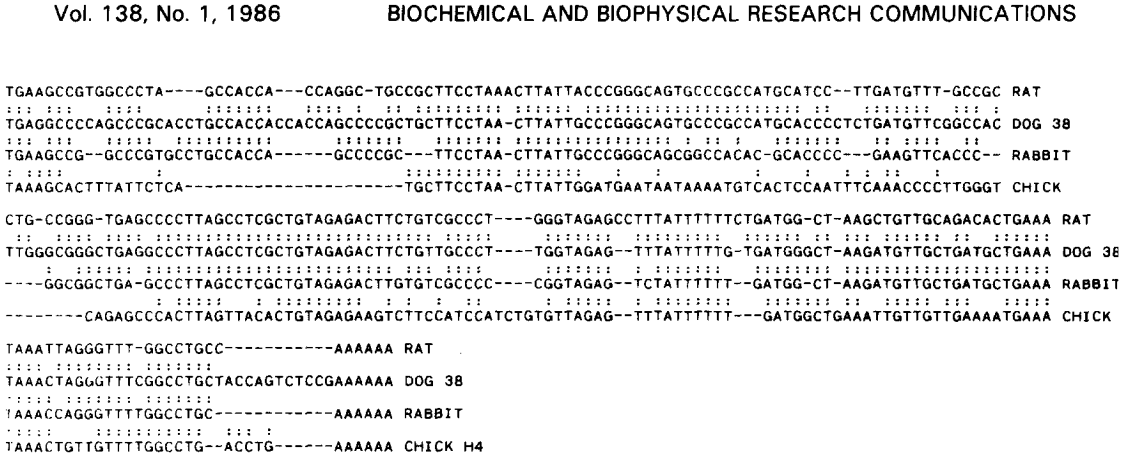

FIGURE 2: Comparison of the 3' nontranslated region of the cDNAs coding rat, dog (clone BCK38), rabbit and chick (clone H4) B creatine kinase. The alignment was made to obtain maximal homology by the introduction of deletions (-) into the sequence. The (:) indicates identical nucleotides. All sequences start with the translation stop signal. Only 6 residues of the poly(A) tail are shown for all clones.

The 3' nontranslated regions of mRNAs from related genes are, in general, less conserved than the coding sequences (17). However, a few examples of sequence identity in the 3' nontranslated region between mRNAs encoding highly conserved homologous proteins in distantly related vertebrates have been found and include actin (18), tubulin (19) and myosin (20). B creatine kinase provides a further example of conservation of both the coding and the 3' nontranslated regions of mRNAs between diverse species. The striking conservation of 3'nontranslated sequences of B cDNAs among diverse species suggests the existence of evolutionary selective pressure that maintains homology in a noncoding region whose function is presently unknown. However, sequences present in the 3' noncoding region may have a role in regulation of the tissue specific expression of these mRNAs (17).

A second unusual feature of the 3' noncoding sequences of canine B cDNAs is the presence of microheterogeneity. This finding is very unusual and suggests variabilty in the site of cleavage-polyadenylation of transcripts. Microheterogeneity in the 3' nontranslated region of hepatitis B viral surface antigen (21), chicken lens beta-B1-crystallin polypeptide (22) and chick B CK (16) was noted when several cDNAs containing the 3' nontranslated region were sequenced. Polyadenylation in higher eukaryotes

may be directly coupled to 3' processing (23). RNA processing involves an endonucleolytic cut in the pre-mRNA followed by polyadenylation usually at a single site downstream of the signal sequence AATAAA. Characterization of genomic clones encoding B creatine kinase will determine whether micro-heterogeneity in the 3' nontranslated region of B cDNA is due to variable processing of a single gene transcript or due to the presence of multiple genes encoding B creatine kinase.

ACKNOWLEDGEMENTS: We thank Dr. Burton E. Sobel for critical review of the manuscript, Catherine Ovitt for help with construction of the canine brain cDNA library, Barbara Donnelly for secretarial assistance, and Ann Grace for help with purification of canine B protein. This work was supported in part by Grant HL-17646, Specialized Center for Research in Atherosclerosis in Ischemic Heart Disease, from the National Institutes of Health.

REFERENCES

1. Watts, D.C. (1973) In Boyer, P. D. (ed.) The Enzymes. Academic Press, New York, pp. 383-455.
2. Grace, A. M., Perryman, M. B., and Roberts, R. (1983) J. Biol. Chem. 258, 15346-15354.
3. Roberts, R., Gowda, K. S., Ludbrook, P. A., and Sobel, B. E. (1975) Am. J. Cardiol. 36, 433-437.
4. Rosenberg, U. B., Hunz, G., Frischauf, A., Lehrach, H., Mahr, R., Eppenberger, H. M., and Perriard, J.-C. (1982) Proc. Natl. Acad. Sci., USA 79, 6589-6592.
5. Roman, D., Billadello, J., Gordon, J., Grace, A., Sobel, B., and Strauss, A. (1985) Proc. Natl. Acad. Sci. USA 82, 8394-8398.
6. Perryman, M. B., Strauss, A. W., Buettner, T. L., and Roberts, R. (1983) Biochem. Biophys. Acta 747, 284-290.
7. Dillman, W. H., Barrieux, A., Neeley, W. E., and Contreras, P. (1983) J. Biol. Chem. 258, 7738-7745.
8. Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1417.
9. Watson, C. J. and Jackson, J. F. (1985) In: Glover, D. M. (ed.) DNA Cloning, pp 79-88, IRL Press, Oxford.
10. Billadello, J. J., Roman, D. G., Grace, A. M., Sobel, B. E., and Strauss, A. W. (1985) J. Biol. Chem. 260, 14988-14992.
11. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
12. Grace, A. and Roberts, R. (1982) Clin. Chim., Acta 123, 59-71.
13. Gordon, J. I., Smith, D. P., Andy, R., Alpers, D. H., Schonfeld, G., and Strauss, A. W. (1982) J. Biol. Chem. 257, 971-978.
14. Pickering, L., Pang, H., Biemann, K., Munro, H., and Schimmel, P. (1985) Proc. Natl. Acad. Sci. USA 82, 2310-2314.
15. Benfield, P. A., Henderosn, L., and Pearson, M. L. (1985) Gene 39, 263-267.
16. Hossle, J. P., Rosenberg, W. B., Schafer, B., Eppenberger, H. M., and Perriard, J.-C. (1986) Nucleic Acids Res. 14, 1449-1463.
17. Yaffe, D., Nudel, U., Mayer, Y., and Neuman, SA. (1985) Nucleic Acids Res. 13, 3723-3737.

18. Ordahl, C. P. and Cooper, T. A. (1983) Nature 303, 348-341.
19. Cowan, N. J., Dobner, P. R., Fucks, E. V., and Cleveland, D. W. (1983) Mol. Cell. Biol. 3, 1738-1745.
20. Saidapet, C., Khandekar, P., Mendola, C., and Siddiqui, M. A. Q. (1984) Arch. Biochem. Biophys. 233, 565-572.
21. Simonsen, C. C. and Levinson, A. D. (1983) Mol. Cell Biol. 3, 2250-2258.
22. Hejtmancik, J. F., Thompson, M. A., Winstow, G., and Piatigorsky, J. (1986) J. Biol. Chem. 261, 982-987.
23. Birnstiel, M. L., Busslinger, M., and Strub, K. (1985) Cell 41, 349-359.