The primary structure of chicken B-creatine kinase and evidence for heterogeneity of its mRNA

J.P.Hossle, U.B.Rosenberg*, B.Schäfer, H.M.Eppenberger and J.-C.Perriard**

Institute for Cell Biology, Swiss Federal Institute of Technology, CH-8093 Zürich, Switzerland

Received 30 October 1985; Revised and Accepted 2 January 1986

ABSTRACT

cDNA clones for chicken B-CK were isolated by immunoscreening from a gizzard cDNA library constructed in the expression vector λgt11. The entire coding portion in addition to the complete 3' untranslated region and 42 bp of the 5' noncoding part are represented in the clone H4. On RNA blots H4 insert DNA hybridized to a 1600 bp poly(A)+ RNA from gizzard, brain and heart but not to breast or skeletal muscle RNA. In vitro generated sense strand transcripts of H4 insert DNA were translated in vitro into a protein indistinguishable from isolated, authentic B-CK. The distinct nucleotide sequences of H4 insert DNA and M-CK cDNA were translated into 82% homologous amino acid sequences. Sequence heterogeneity among the B-CK cDNA clones within both the 3' noncoding and even in the coding region indicates the existence of multiple B-CK mRNA species.

INTRODUCTION

Creatine kinases (EC 2.7.3.2) play an important role in the energy metabolism of vertebrates (1) and of invertebrates (2). In invertebrates there are other guanidino phosphotransferases like arginine kinase, but their relationship to creatine kinase is not well understood.

In chicken as in most other higher vertebrates three different types of subunits of creatine kinase (CK) are known which are expressed in a tissue specific manner. B-CK occurs in adult brain, smooth muscle, heart and several other tissues and is also present in embryonic cells (3). M-CK, however, is accumulated in significant amounts in adult skeletal muscle of chicken, but contrary to mammalian heart in chicken heart no M-CK and only B-CK can be found. The third type of CK, the mitochondrial subunit Mi-CK, appears to be confined to mitochondria of cells that either express B-CK or M-CK. Between
the mitochondrial compartment and the cytosolic space a
creatine phosphate shuttle has been postulated (1,2).

B-CK and M-CK subunits form homodimeric isoenzymes MM-CK
and BB-CK as well as heterodimeric MB-CK, however both are
incapable of dimerizing with the mitochondrial subunit. The
molecular weight of the CK subunits were estimated to be around
40,000. M-CK has a more basic isoelectric point than B-CK and
both homodimeric isoenzymes give rise to double spots when
subjected to denaturing two dimensional gel electrophoresis
(4). Alkylation of CK results in loss of enzymatic activity (5)
and cysteine 283, embedded in a known 25 amino acid peptide
(6), is most easily modified.

Rabbit antibodies against chicken creatine kinases were
shown not to crossreact with heterologous native CK or CK from
other species. If however the antigens were partially or
completely denatured crossreaction was observed (7).

During development of muscle a switch from B-CK in
embryonic tissue to M-CK in differentiated muscle can be
observed. The isoenzyme transition can also be followed in
differentiating cultures of myogenic cells (8,9,10,11,12) but
does not go to completion for reasons not yet known.

We have recently reported the complete amino acid sequence
of M-CK (13) and present here the nucleotide and complete amino
acid sequence for chicken B-CK. B-CK clones were isolated by
screening a cDNA expression library made from gizzard RNA with
antibody against chicken BB-CK. Identification was done
independently from sequence analysis by translation of RNA
generated in vitro from a full length cDNA clone and subsequent
characterization of the B-CK product. The observation of
heterogeneity of other independent clones within the coding
portion, as well as in untranslated regions, indicates the
existence of multiple B-CK mRNAs.

MATERIAL AND METHODS
Construction of cDNA expression libraries
Total RNA was prepared from leg or gizzard muscle of 19-20 day
old chicken embryos by the guanidine-HCl technique (14) and
poly(A)$^+$RNA was subsequently purified from these preparations
by chromatography on oligo(dT)-cellulose (15). 5.5 ug of each poly(A)$^+$RNA were copied into cDNA with reverse transcriptase (J. Beard, Life Science, St. Petersburg, FL) (16) and double-stranded cDNA was prepared with the Klenow fragment of E. coli DNA polymerase I (BRL). After cleavage of the hairpin loops with nuclease S1, the cDNA was methylated with EcoRI methylase (New England Biolabs) and ligated to 0.6 ug of phosphorylated EcoRI linkers (Pharmacia-PL) with T4 ligase. The mixture was then digested with EcoRI and fractionated on a 1% agarose gel. cDNA larger than 500 bp was recovered by electrophoresis onto DEAE cellulose paper (17). One out of four aliquots was ligated to 2.4 ug of dephosphorylated $\lambda$gt11 vector DNA (18), packaged into the phage, and plated on E. coli Y1088 (19).

Immunoscreening of the chicken gizzard cDNA library

The $\lambda$gt11 expression library was screened with rabbit anti chicken BB-CK antibody by a modification of the procedure described (19). Instead of radioactive labeled antibody a peroxidase-conjugated goat antiserum against rabbit IgG (Calbiochem; working dilution 1:2'000) was used as secondary antibody. To remove components reactive with E. coli antigens from the antibody preparations, the undiluted first and secondary antibodies were preabsorbed with a lysate of BTA282 cells lysogenic for $\lambda$gt11 amp3 (gift of Dr. R. Joho, University of Zürich) covalently coupled to Sepharose 4B resin (Pharmacia Fine Chemicals).

In vitro transcription with SP6 RNA polymerase and cell free translation

Linear cDNA templates (2 ug) were transcribed with SP6 RNA polymerase (Boehringer Mannheim) as described (20). Capping was achieved by adding 500 uM m$^7$GpppG (Pharmacia Fine Chemicals) to the reaction while the GTP concentration was lowered to 24 uM. Usually 1-2 ug of homogenous mRNA were recovered.

Cell free translation was carried out in the rabbit reticulocyte lysate translation system (Promega Biotec) following the procedure supplied by the manufacturer.

An aliquot of the translation assay was removed and the peptides synthesized were separated on a 10% SDS-polyacrylamide gel and revealed by autoradiography (21). A second aliquot was
used for analysis by specific immunoprecipitation as described (21,22).

RNA blot hybridization
RNA blot analysis was performed using the glyoxal method (23).

DNA sequencing
Appropriate restriction fragments from the cDNA clones were subcloned into M13- phage vectors mp8 and mp9 for sequencing by the dideoxynucleotide chain termination method (24). Sequencing reactions were carried out with a M13 sequencing kit (Amersham) using deoxyadenosine 5'-[α[35S]thio) triphosphate (SJ.304; Amersham) as label.

RESULTS
Construction of a chicken leg muscle and a gizzard cDNA expression library
Two cDNA expression libraries prepared in vector λgt11 were constructed with cDNA derived from poly(A)+ RNA from leg muscle or gizzard of hatching chicken embryos (19-20 d of incubation) as outlined in the MATERIALS AND METHODS section. The leg library contained about 0.5x10⁶ independent phages of which 68% (0.34 x 10⁶) were recombinants. The gizzard library contained some 1x10⁶ phages with 60% (0.6x10⁶) recombinants.

Immunoscreening of the gizzard cDNA expression library with rabbit anti-chicken BB-CK antibody and characterization of the isolated clones
Preliminary experiments have shown that sensitivity of the detection of a protein was in the range of 100 pg BB-CK per 1 mm² spot. 3,1x10⁴ recombinants were screened and after the first round 10 signals could be identified. Plaque purification including 2-3 rounds of immunoscreening finally yielded 5 positive, homogeneous phages: 11a, 13a, 18b, 18c and 18e, which were further analyzed.

β-galactosidase fusion proteins from phages 13a and 18c were analysed on SDS-polyacrylamide gels, subsequently blotted onto nitrocellulose and were shown to react specifically with antibodies against BB-CK (data not shown).

In order to determine whether the complete sequence of the putative B-CK cDNA were represented in the recombinant phages,
DNA was prepared and digested with the restriction enzyme EcoRl, releasing the inserts. The digested DNA fragments were resolved by agarose gel electrophoresis and the insert sizes were determined and are listed in Table 1.

These inserts were subcloned into M13 mp8 and preliminary sequence analysis showed that the cloned cDNA inserts of 11a, 13a, 18b and 18e had a poly(A) tract as expected since cDNA synthesis was primed with oligo d(T). Clone 18c however did not display a poly(A) tract, but demonstrated a sequence overlap at its 3' end with the 5' most part of 18e and 18b. The clones 13a, 18b, 18c and 18e were partially sequenced and preliminary amino acid data all displayed a rather striking similarity with the already published M-CK amino acid sequences. The translated amino acid sequence derived from the DNA sequence of clone 11a however did not fit this pattern and therefore may represent a false positive clone. Further analysis of clone 18c showed no start codon in an appropriate position although the combined sequences of 18b or 18e and 18c would cover the entire putative B-CK cDNA.

**Isolation of a clone containing the almost full length cDNA for chicken B-CK**

In order to obtain a full length cDNA for B-CK the gizzard cDNA library was rescreened using the inserted cDNA from clone 18c as a hybridization probe. About $7 \times 10^3$ recombinant phages were

<table>
<thead>
<tr>
<th>Clone</th>
<th>Insert Size (bp)</th>
<th>Poly(A) Tract</th>
</tr>
</thead>
<tbody>
<tr>
<td>11a</td>
<td>approx. 650</td>
<td>+</td>
</tr>
<tr>
<td>13a</td>
<td>&quot; 600</td>
<td>+</td>
</tr>
<tr>
<td>18b</td>
<td>764</td>
<td>+</td>
</tr>
<tr>
<td>18c</td>
<td>680</td>
<td>-</td>
</tr>
<tr>
<td>18e</td>
<td>identical with 18b as shown by sequencing</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>detected by re-screening with 18c insert DNA:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>approx. 1300</td>
<td>+</td>
</tr>
<tr>
<td>H4</td>
<td>1400</td>
<td>+</td>
</tr>
<tr>
<td>H8</td>
<td>517</td>
<td>-</td>
</tr>
<tr>
<td>H10</td>
<td>approx. 1150</td>
<td>+</td>
</tr>
<tr>
<td>H52</td>
<td>&quot; 800</td>
<td>+</td>
</tr>
</tbody>
</table>
rescreened and 7 positive phages could be identified. The phages H1, H4, H8, H10 and H52 were purified to homogeneity, DNA was prepared, digested with EcoRI and the insert sizes were determined.

As shown in Table 1 insert sizes varied from 517 bp to 1400 bp and again all but H8 carried poly(A) tracts. It is likely that the longest insert from H4 represents a full length cDNA.

RNA blot hybridization
To further characterize these clones they were used in RNA hybridization experiments that allow estimation of the size of mature mRNA. At the same time the tissue specific expression was investigated since it is known that B-CK is accumulated in significant amounts in gizzard, brain and heart. No anti B-CK reacting material could be found in adult skeletal muscle (7) but the skeletal muscle specific form of M-CK is present. It has been shown earlier that M-CK cDNA probes hybridize to muscle RNA but do not crosshybridize to gizzard RNA (25) hence it is likely that B-CK cDNA probes will not crosshybridize with M-CK mRNA.

Samples of poly(A)^+ RNA were denatured and run on a agarose gel, blotted onto nitrocellulose and hybridized to plasmid DNA containing the insert of phage H4, 13a or 18b. The result for the hybridization with the H4 sequences is shown in Figure 1. RNA isolated from brain (lane 1), heart (lane 2) and gizzard (lane 4) show a prominent band with an estimated size of about 1600 nucleotides. RNA derived from breast muscle (lane 3) or leg muscle (lane 5) did not hybridize significantly to the putative B-CK cDNA probe. Essentially the same results were obtained for Northern blots developed with the probe derived from phages 13a or 18b (not shown). It is therefore likely that the recombinant DNA of H4, 18b or 13a represent B-CK cDNA sequences.

Cell free translation of in vitro generated homogeneous sense-strand transcripts derived from H4 insert DNA
The unambiguous identification of the recombinant DNA of phage H4 as B-CK cDNA was achieved by cell free translation of homogeneous sense-strand transcripts of the H4 insert DNA and subsequent immunological and electrophoretic analysis of the
resulting product. To this end the H4 insert DNA was subcloned into the plasmid pSP65 and this DNA was then used as template for the SP6-specific RNA polymerase, generating homogeneous sense-strand transcripts. These were then assayed in a cell free in vitro translation system derived from rabbit reticulocytes and the product was identified by immunoprecipitation and subsequent SDS gel electrophoresis as shown in Figure 2. The autoradiograph of the polyacrylamide gel (Figure 2, panel B) in lane 5 demonstrates that such RNA can be translated into a protein product migrating in an identical manner as purified B-CK protein (Figure 2A, lane 2) but it was clearly distinct from purified M-CK protein (Figure 2A, lane 1). The translation product could be immunoprecipitated with rabbit anti chicken BB-CK antibody (Figure 2B, lane 6) and again was indistinguishable of the immunoprecipitated product obtained by translation of poly(A)$^+$ RNA from gizzard tissue (Figure 2B,
Figure 2. Identification of clone H4 by cell-free translation of in vitro generated mRNA. pSP65 plasmid containing the cDNA insert from clone H4 was linearized with Hind III and templates in vitro transcribed with SP6 RNA polymerase. The RNA was then translated in a cell-free translation system and samples from cell-free protein synthesis and immunoprecipitated translation products were subsequently analysed on a 10% SDS-polyacrylamide gel. (A) Gel stained for protein. Lane 1: purified MM-CK (3 ug); 2: purified BB-CK 3 ug). (B) Autoradiograph of cell-free translation products. Lane 3: gizzard poly(A)+ RNA (0.3 ug); 4: immunoprecipitation with anti BB-CK antibody; 5: H4 derived RNA (20 ng) and 6: aliquot of H4 primed translation products immunoprecipitated with anti BB-CK antibody. Lane 7: leg muscle poly(A)+ RNA (0.1 ug) in vitro translated and immunoprecipitated with anti MM-CK antibody. The autoradiograph was exposed for 64 hrs.

lane 4) but distinct from the immunoprecipitated translated M-CK product of leg muscle poly(A)+ RNA (Figure 2B, lane 7). Hence these results indicate that the insert DNA of clone H4
Figure 3. Nucleotide and derived amino acid sequences of H4 insert DNA. Regions underlined indicate sequence heterologies between the isolated B-CK cDNA clones. The consensus translation start sequence and the poly(A) addition signal are overlined. The derived amino acid sequence is given in the single letter code above the DNA sequence. The putative active site region containing Cys 283 (circled) is shown in a box.

codes for B-CK sequences and contains all the coding information for the complete B-CK subunit.

Sequence analysis

The B-CK cDNA was now sequenced to fully characterize its nucleotide sequence and also to derive the until today unknown amino acid sequence of a chicken B-CK subunit. Various cDNAs and fragments thereof were subcloned into the phages M13 mp8 or M13 mp9 finally used as templates in dideoxy sequencing reactions. Both strands of almost all regions were sequenced at least once.

Figure 3 presents the nucleotide sequence and its derived amino acid sequence of the H4 insert DNA. The sequence contains...
1379 nucleotides, 42 nucleotides belong to the untranslated region at the 5' end and 194 untranslated nucleotides (including the termination codon TAA) are located at the 3' end. The coding portion of 1143 nucleotides codes for 381 amino acids including the methionine encoded by the initiation codon at position 43. The coding portion has the same length as all other CK sequences known so far (26). In the shown B-CK cDNA sequence one potential initiation site for protein synthesis could be characterized by the following criteria: It occurs at 381 codons upstream from the postulated stop codon and is embedded in the sequence at position 38: CAGCCATGC which is almost identical to the consensus sequence for the initiation of protein synthesis CC\textsuperscript{A}\textsubscript{T}GG (27). Furthermore the first amino acids are MPFSN and correspond exactly to the amino terminus of rabbit B-CK (28). The amino terminal peptide MPF is common to all other non-mitochondrial CK sequences known to date (26). The in vitro translation of the plasmid derived RNA demonstrates the synthesis of a protein not distinguishable from a fraction of the purified B-CK protein on SDS-polyacrylamide gels (see Figure 2).

At the 3' end two stop codons could be identified, TAA at position 1186 and TAG at position 1285. The first stop codon produces a CK protein of the expected size of 381 amino acids, with a C-terminal sequence of PAQK common to all known CK sequences (26), while the second stop codon would give rise to a completely different C-terminal peptide. Two canonical polyadenylation signals (AATAAA) occur in the sequence, one at position 1228, the other at position 1352. The second site appears to be used in all the independent cDNA clones that could be sequenced at the 3' end. Thus the 1379 nucleotide sequence most likely represents the major part of the mRNA sequence. If we assume that the average poly(A) segment contains about 200 nucleotides the resulting size would be some 1579 bases corresponding quite well with the observed size of about 1600 nucleotides as shown in the RNA blots of Figure 1. It is possible that a 20-30 nucleotide sequence is still missing at the 5' end of the mRNA sequence.
Figure 4. Heterogeneity of B-CK cDNA clones. The bars represent B-CK cDNA inserts from clone H4 and from clones 18b,e. The untranslated regions are shaded whereas the poly(A) tracts of the clones are shown as solid lines. The heterogeneous regions in 18b,e are indicated by solid blocks and the sequences for each of these regions are compared with the corresponding sequences from H4 between the bars. Derived amino acid sequences are given in the single letter code.

Heterogeneity of B-CK mRNA

Comparison of the presented H4 DNA sequence with partial sequences from the other independently derived B-CK cDNA clones revealed in most cases complete homology. However, in the identical clones 18b and 18e sequence differences within the coding as well as at the 3' end of the noncoding region were observed as shown in Figure 4.

The nucleotide sequence in H4 from positions 633-653 reads TGTTTCTCCTCTTCTGTTGGC. The corresponding nucleotides in the partial cDNA clones 18b,e read GGTCTGTGCTCATCAGTTGGT. Thus within a stretch of 21 nucleotides seven changes (underlined) occur. It is unlikely that this deviation has arisen due to a cloning artefact since most of the nucleotides are still conserved and the reading frame for the derived amino acid sequence is maintained. This heterogeneity however has striking consequences for the amino acid sequences. While in H4 DNA the peptide sequence is VSPLLALA it becomes in 18b,e VCAHQLV, therefore five out of seven residues are changed.

Furthermore, at position 957 a C is inserted in clones 18b,e instead of a T in the H4 DNA with no consequences for the amino acid sequence.
There is an additional heterogeneity at the 3' end since the pentanucleotide CCTGC at positions 1375-1379 in the H4 DNA is missing in the clones 18b, e and H52. The presence of this heterogeneity in the clone H52, otherwise identical to H4, indicates that this type of heterogeneity may be due to differential processing at the 3' end of the RNA.

DISCUSSION
The isolated B-CK cDNA sequences were identified by several criteria:
The fusion proteins of several clones reacted specifically with antibodies against chicken BB-CK.
RNA blot hybridization assays showed tissue specific expression but no significant crosshybridization with M-CK mRNA.
The cell free translation product of in vitro generated homogeneous B-CK mRNA was indistinguishable from isolated authentic B-CK protein and also reacted with anti BB-CK antibody.
The nucleotide sequences of M-and B-CK are distinct but the derived amino acid sequences show a high degree of conservation.

The nucleotide sequences of the coding portions are conserved to 72% between M-CK and B-CK. There are only two longer homologous stretches one of 17 nucleotides (residues 589-605 of the B-CK sequence) the other of 15 nucleotides (residues 1089-1103) and several shorter conserved segments.

The amino acid sequences derived from the nucleotide sequences show extensive homology. At 82% of the positions identical amino acids were found in both types of CK subunits. The changes are distributed over the entire sequence and are interspersed with blocks of homologous amino acids. Of the 70 changes, 36 are nonconservative and 16 result in the replacement of amino acids with positively charged or uncharged groups in the M-CK peptide by more negatively charged ones in B-CK, while only 9 replacements of residues give rise to a shift in the opposite sense. The changes are compatible with the observed biochemical differences of the purified isoenzymes. The B-CK subunit is more acidic than the M-CK subunit (4,29). The calculated overall amino acid composition
is in good agreement with the published data obtained by amino acid analysis (29). The amino acid composition allowed the calculation of the exact molecular weights. B-CK has a molecular weight of 42'878 while M-CK is 43'335 and is therefore only 457 (1.05%) units heavier than the B-CK subunit. The B-CK and M-CK subunits can be separated on SDS gels and B-CK migrates with an apparent molecular weight of about 3'000-5'000 larger than that of M-CK (11,21). The calculated Mr difference however cannot account for the observed difference in electrophoretic mobility since it is too small and in the wrong sense. It seems therefore that the differences in charge and secondary structure are more likely responsible for these phenomena than the differences in molecular weight.

B-CK typical sequence elements have been obtained by comparing of 6 muscle specific CK sequences with 3 B-CK sequences and will be discussed elsewhere (26).

The cloned fragments for B-CK showed heterogeneity at both the nucleotide level as well as within the derived amino acid sequences. The pentanucleotide missing in the 3' end of some clones may indicate termination of transcription of a B-CK gene at more than one site, heterogeneity in the 3' end processing or that the cDNA clones were derived from transcripts of more than one B-CK gene. One of the inhomogeneities discovered in the coding portion of clones 18b,e gives rise to an alternative peptide sequence. This indicates that amino acid sequences should be derived if possible from a full length cDNA clone, since combination of several independent clones or primer extensions may give rise to sequences not existing in nature. The amino acid replacements resulting from the nucleotide sequence heterogeneity would probably give rise to a B-CK subunit with a different isoelectric point. Indeed BB-CK from gizzard, heart, brain and myogenic cell cultures could be separated into two species with identical apparent molecular weights but different isoelectric points. In vitro translated poly(A)$^+$ RNA also gave rise to two spots comigrating on two-dimensional gels with the purified B-CK (4), therefore the two types of B-CK subunits might be due to the existence of two B-CK mRNA species. The detection of the two heterogeneous
B-CK coding sequences among the B-CK cDNA clones described above supports this hypothesis. The two mRNAs might arise by transcription of two different B-CK genes or by alternative processing of transcripts derived from a unique gene.

Genomic B-CK sequences were isolated and determination of their organization and elucidation of their regulation will help to answer the remaining questions.

ACKNOWLEDGEMENTS
We thank Drs. R. Ehrismann (Friedrich Miescher Institute, Basel) and P. Schimmel (MIT, Cambridge) for discussing their data on B-CK prior to publication. J.C.P. is very grateful to Dr. B. Nadal-Ginard (Harvard Medical School, Boston) for the laboratory facilities and encouragement and to M.-A. Strehler-Page and to Dr. E.E. Strehler (Harvard Medical School, Boston) for their cheerful support during a sabbatical leave. The work was supported by grant 3.491-0.83 from the Swiss National Science Foundation and a grant to J.C.P. from the Muscular Dystrophy Association of America.

* Current address: Lonza AG, CH-3930 Visp, Switzerland.

**To whom correspondence and reprint request should be addressed.

REFERENCES