

Multiple mRNA isoforms of the transcription activator protein CREB: generation by alternative splicing and specific expression in primary spermatocytes

Siegfried Ruppert¹, Timothy J. Cole, Michael Boshart, Erika Schmid and Günther Schütz

German Cancer Research Center, Institute of Cell and Tumor Biology, D-6900 Heidelberg, Im Neuenheimer Feld 280, FRG

¹Present address: Howard Hughes Medical Institute, University of California at Berkeley, Department of Molecular and Cell Biology, 401 Barker Hall, Berkeley, CA 94720, USA

Communicated by G. Schütz

We have characterized cDNA clones representing mouse CREB (cyclic AMP responsive element binding protein) mRNA isoforms. These include CREB Δ and CREB α , of which the rat and human homologues have been previously identified. Both encode proteins with CRE-binding activity and identical transactivation potential. The additional CREB mRNA isoforms potentially encode CREB related proteins. From the structural organization of the mouse CREB gene we conclude that the multiple transcripts are generated by alternative splicing. Furthermore we show that specific CREB mRNA isoforms are expressed at a high level in the adult testis. Expression of these isoforms is induced after commencement of spermatogenesis. *In situ* hybridization suggests that this expression occurs predominantly in the primary spermatocytes. Comparison of the CREB gene with the recently isolated CREM (cAMP responsive element modulator) cDNAs illustrates that the two genes have arisen by gene duplication and have diverged to encode transcriptional activators and repressors of the cAMP signal transduction pathway.

Key words: alternative splicing/CREB gene/CREM gene/gene duplication/spermatocytes

Introduction

In response to extracellular signals, such as hormonal stimulation, cyclic AMP (cAMP) levels change and lead to a variety of responses including the alteration of transcription of many eukaryotic genes. This transcriptional activation is mediated through cAMP-responsive elements (CRE; Comb *et al.*, 1986; Montminy *et al.*, 1986; Delegeane *et al.*, 1987; Silver *et al.*, 1987; Tsukada *et al.*, 1987; Bokar *et al.*, 1988; Quinn *et al.*, 1988; Fisch *et al.*, 1989; Boshart *et al.*, 1990; Weih *et al.*, 1990). The CRE is recognized by a family of DNA binding proteins, the CREB/ATF family, among which the CRE binding protein (CREB) is best characterized (Montminy and Bilezikjian, 1987; Deutsch *et al.*, 1988; Yamamoto *et al.*, 1988). Recently, cDNAs encoding CREB were isolated from human placenta (Hoeffler *et al.*, 1988) and from rat brain (Gonzalez *et al.*, 1989), identifying CREB as a bZIP protein (basic region/leucine zipper protein; reviewed by Johnson and McKnight, 1989; Kerppola and

Curran, 1991). Both the basic region and the leucine zipper dimerization domain are essential for DNA binding and transcriptional activation (Dwarki *et al.*, 1990). While the human and rat cDNA clones were very homologous in their overall structure, the rat CREB cDNA clone contained an insertion of 42 bp, termed α . Recently, Yamamoto *et al.* (1990) reported the isolation of a rat CREB cDNA clone, Δ CREB, from PC12 cells that appeared to be homologous to the human CREB cDNA clone isolated by Hoeffler *et al.* (1988). Although the CREB gene structure was not analysed, the possibility was discussed that these different cDNA clones might have arisen by alternative splicing from a single gene transcript. This hypothesis was supported by the isolation of both CREB cDNA isoforms in a human T cell cDNA library (Berkowitz and Gilman, 1990). While rat CREB α and CREB Δ showed a different transactivation potential in F9 teratocarcinoma cells (Yamamoto *et al.*, 1990) this difference was not observed in a different test system (Berkowitz and Gilman, 1990).

We are analysing the hormone dependent and liver-specific expression of the rat tyrosine aminotransferase (TAT) gene (reviewed by Nitsch *et al.*, 1991). A cAMP-responsive element was identified as an essential component of a liver-specific enhancer of the TAT gene (Boshart *et al.*, 1990) which was also shown to be the target for the tissue-specific extinguisher-1 (*TSE-1*; Boshart *et al.*, 1990). Following stimulation by cAMP we observed increased protein binding over the TAT–CRE using *in vivo* footprinting (Weih *et al.*, 1990) and *in vitro* binding studies suggested that CREB binds to this sequence (M. Nichols and F. Weih, unpublished). It was evident from Northern blot analysis (S. Ruppert, unpublished) that multiple CREB mRNAs exist, in addition to the previously identified mRNAs for CREB Δ (Hoeffler *et al.*, 1988) and CREB α (Gonzalez *et al.*, 1989). Here we describe the isolation and characterization of mouse CREB cDNA and genomic clones. This allows us to analyse the significance of this CREB mRNA diversity, to identify which CREB isoforms interact with the TAT–CRE and to elucidate the particular role of CREB isoforms in the antagonism between cAMP signalling and *TSE-1* (Boshart *et al.*, 1990, 1991). Furthermore, knowledge of the structure of the CREB gene will enable us to manipulate CREB gene expression by homologous recombination.

Several CREB and CREB related cDNA clones, including the homologues of the previously reported human and rat cDNA clones (Hoeffler *et al.*, 1988; Gonzalez *et al.*, 1989; Berkowitz and Gilman, 1990) were isolated. In addition we identified other CREB mRNA isoforms which would encode different CREB related proteins. A detailed analysis of cDNA and genomic clones proved that the CREB gene encodes multiple mRNAs which are generated by alternative splicing. Specific CREB mRNA isoforms are specifically expressed at higher levels in the primary spermatocytes of the testis.

Results

Isolation and analysis of mouse CREB cDNA clones reveals many isoforms

We used Northern blotting to analyse RNAs from various tissues and cells for CREB expression. Under high stringent conditions multiple mRNAs were identified (data not shown) suggesting that they may have arisen from the same gene by alternate processing. To elucidate the nature of these different RNA species, mouse CREB cDNAs were isolated from various cDNA libraries (F9 teratocarcinoma, B16 melanoma and liver). A total of 13 different cDNAs, ranging in size from 1 to >5.5 kb, were isolated and characterized by restriction enzyme analysis, crosshybridization to individual subprobes and by sequencing of their entire coding regions. A schematic representation of these cDNA clones is shown in Figure 1.

The most abundant cDNA isoform CREB Δ , represented by pmcCREB Δ , pmcCREB21, pmcCREB36, pmcCREB37 and pmcCREB28, was identified as the mouse homologue of the human and rat CREB cDNA clones (Hoeffler *et al.*, 1988; Gonzalez *et al.*, 1989; Berkowitz and Gilman, 1990). Analyses of the other cDNA clones revealed that they contained insertions or deletions with respect to the CREB Δ isoform changing the coding information. Clones pmcCREB27 and pmcCREB41, contain a 42 bp insertion leading to the incorporation of 14 additional amino acid residues, termed the α -domain. The rat and human homologues of these clones have been described in PC12 cells and in T cells (Gonzalez *et al.*, 1989; Berkowitz and Gilman, 1990). Two cDNA clones pmcCREB30 and pmcCREB γ , contain another insertion, the γ -domain. This 113 bp insertion occurs at the identical position as the α -domain. However, due to the presence of internal translation stop codons it would encode a protein of only 90 amino acid residues lacking the basic region and the leucine zipper dimerization domain but containing the glutamine-rich (Q) domain, postulated to be involved in CREB mediated transactivation (Gonzalez *et al.*, 1991). The variant CREB cDNA clone, pmcCREB Ψ , is characterized by a 175 bp insertion close to its 5' end which would result in an even shorter product—a 61 amino acid protein, containing only one half of the Q-domain. The isoform pmcCREB Ω reveals an in-frame deletion of 183 bp and a variable sequence at its 3' end. This altered sequence would lead to a modification of the basic region which is devoid of the leucine zipper dimerization domain. Another clone, pmcCREB28, contains a short, 11 bp sequence (5'-GCCTTCTTCAG-3') at its 5' end which occurs exactly at the exon 1–intron A boundary. Oligonucleotide hybridization has shown that this 11 bp sequence is not present in the cloned CREB contig (data not shown). It might be hypothesized that it is part of a far upstream exon which is transcribed from a different promoter. Clones pmcCREB28 and pmcCREB47 contain long extended regions of 3' noncoding sequences indicating the use of alternate polyadenylation sites.

Analysis of the structural mouse CREB gene provides evidence for alternative splicing

In order to understand the basis for the diversity of the CREB mRNA isoforms, we compared the isolated cDNAs with the structural organization of the gene. A mouse genomic λ DNA library was screened using the cDNA inserts of pmcCREB γ

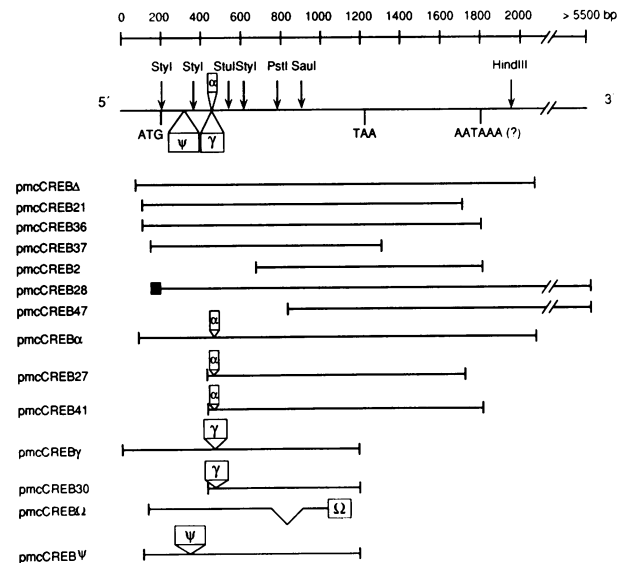


Fig. 1. Schematic representation of the isolated mouse CREB cDNA clones. cDNA inserts are aligned based upon restriction mapping and nucleotide sequencing. Boxes denoted α , γ , Ψ and Ω indicate insertions with respect to pmcCREB Δ . The filled box in pmcCREB28 indicates an additional variation in the 5' untranslated region of CREB, unique to this cDNA clone. For completeness pmcCREB α which has been generated *in vitro* from pmcCREB Δ is included in this scheme.

and pmcCREB28 as probes and eleven independent λ clones were isolated. All phages formed one overlapping contig and indicated the presence of one single gene. The cloned DNA spans more than 80 kb of which ~ 70 kb comprise the gene (Cole *et al.*, 1992).

The DNA fragments of the genomic subclones containing exons were subcloned and partially sequenced using appropriate designed oligonucleotides derived from the mouse CREB cDNA sequence. The sequences determined from the mouse CREB gene, including sequences of the exon–intron boundaries, are shown in Figure 2. The genomic sequence did not show any deviation from the cDNA sequence and all exon–intron boundaries fit perfectly to the splice consensus sequence (Padgett *et al.*, 1986). In total 11 exons and 10 introns were identified.

A comparison of the different cDNA species to the sequential order of the genomic exon sequences indicates that the different CREB mRNAs were generated by alternative splicing of a single gene transcript. This is shown schematically in Figure 3. The most abundant transcript, the CREB Δ isoform (see also Figure 5), is generated by a splicing reaction in which exon 3 (termed Ψ), exon 5 (termed α) and exon 6 (termed γ) are eliminated. For the production of CREB α , which differs from CREB Δ by a 42 bp insertion, exon 5 (α) is incorporated. These two mouse CREB cDNA isoforms are homologous to those which have been described recently (Hoeffler *et al.*, 1988; Gonzalez *et al.*, 1989; Berkowitz and Gilman, 1990). The CREB isoform CREB γ incorporates exon 6 (γ) at the same position where the α -domain (exon 5) is incorporated in CREB α . The isoform CREB $\alpha\gamma$, identified by PCR analysis (see Figure 5), represents a combination of these two isoforms which contains both the α and the γ exons. Another isoform, CREB Ψ , is the only RNA species identified so far containing exon 3 (Ψ). The CREB Ω isoform deletes exon 3 (Ψ), 4 (α), 6 (γ) and 9, and in addition, it contains a modification of

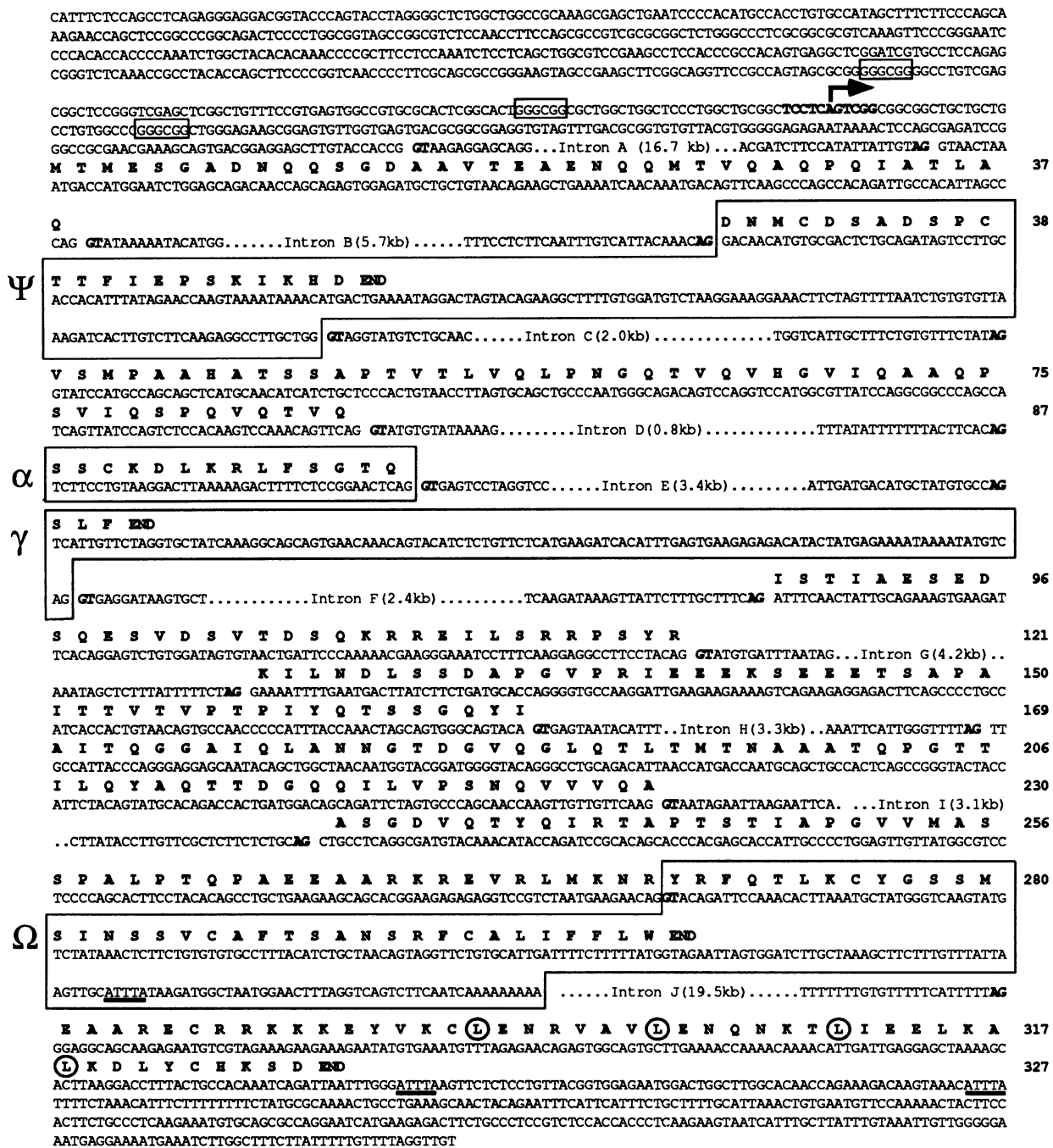


Fig. 2. Nucleotide and deduced amino acid sequence of the mouse CREB cDNA isoforms, including 5' flanking and intron sequences. The sequences were derived from mouse CREB cDNAs and from genomic subclones. Boxes denoted Ψ, α, γ and Ω indicate inserts specific for CREB cDNA clones pmcCREBΨ, pmcCREBα, pmcCREBγ and pmcCREBΩ, respectively. Intron boundaries (GT and AG) are indicated by bold letters. Similar to the human and rat CREB cDNA sequence (Hoeffler *et al.*, 1988; Gonzalez *et al.*, 1989) we assigned the first ATG of the mouse CREB cDNA as the translation initiation codon. Immediately upstream is an in-frame translation termination codon. A second ATG within the sequence ACCATGG, two amino acid residues C-terminal fits much better the Kozak consensus sequence PuCCATGG (Kozak, 1986) and therefore it is most likely that this second ATG will be used more efficiently *in vivo* for translation initiation. The numbering of the amino acid sequence given on the right refers to the most abundant isoform CREBΔ. Amino acid residue specific to other CREB isoforms are given in the respective boxes and are not numbered. Leucines putatively involved in forming the leucine zipper are indicated by circles. Within the initiator sequence Inr 5'-TCCTCAGTCCG-3' (Smale and Baltimore, 1989), indicated in bold letters the arrow points to the putative transcription initiation site. In the promoter region possible binding sites for the transcription factor SPI, based on the consensus sequence 5'-GGGCGG-3' (Mitchell and Tjian, 1989) are boxed. Within the 3' untranslated region the ATTTA motifs, the selective mRNA degradation signals (Shaw and Kamen, 1986; Raymond *et al.*, 1989) are underlined.

exon 10. To generate CREBΩ the donor splice site of exon 10 is not used resulting in the incorporation of additional sequences (see Figure 2). In all other CREB mRNA variants this splice site is used very efficiently and the 151 bp exon 10 is included. Thus, the mature CREBΩ species incor-

porates immediately adjacent intron sequences into the mRNA as deduced by sequence identity of the cDNA clone pmcCREBΩ and the corresponding genomic region (data not shown).

Analysis of the cloned CREB contig with rare cutting

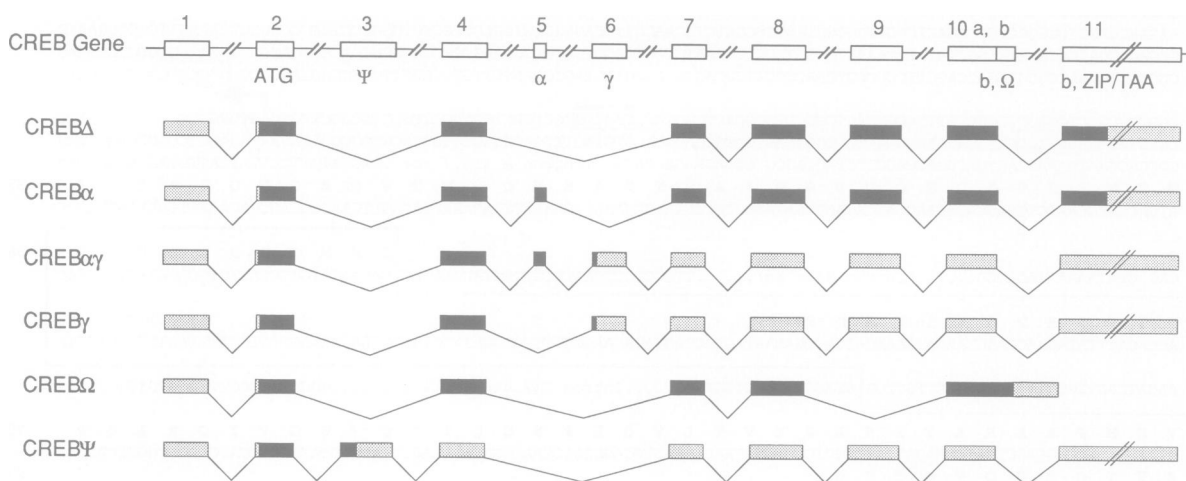


Fig. 3. Organization of the mouse CREB gene and the alternatively spliced transcripts. Exons, numbered 1–11, are shown as open boxes, 5' and 3' flanking sequences as lines and introns as disconnected lines. ATG, translation initiation; TAA, translation termination; Ψ, α, (γ) and (Ω) denote alternatively spliced exons; b, basic region; ZIP, leucine zipper dimerization domain. For completeness pmcCREB $\alpha\gamma$, which has been identified by PCR analysis (see Figure 5) is included in this scheme. Coding regions are indicated by black boxes; noncoding regions by stippled boxes.

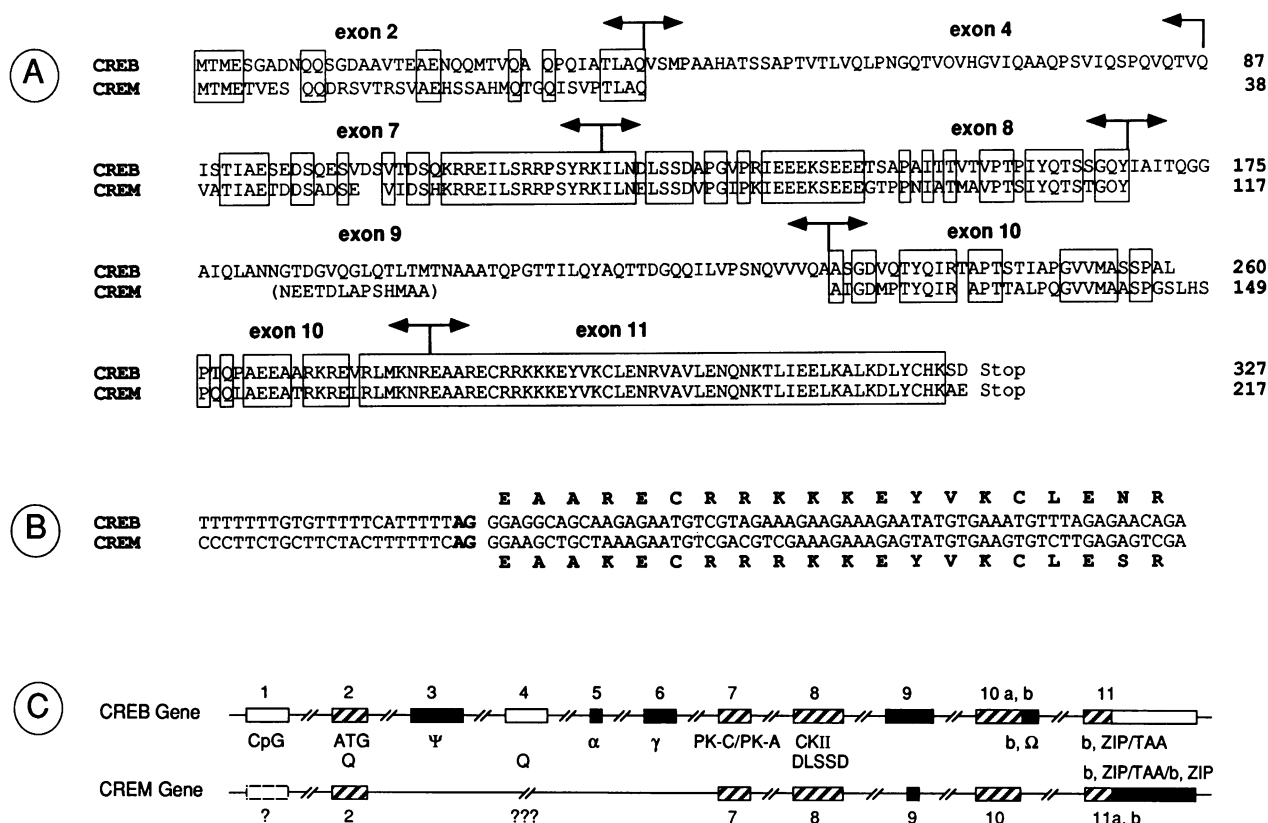


Fig. 4. Comparison of the mouse CREB gene with the CREM cDNA sequences indicates homology in the gene structure. (A) Comparison of the 327 amino acid sequence of mouse CREB Δ and the 217 amino acid sequence of CREM γ (Foulkes *et al.*, 1991) in the one letter amino acid code. For optimal alignment gaps have been introduced. Numbers of amino acids are given on the right. Boxed amino acids indicate sequence identity. The positions of arrows indicate exon–intron boundaries identified in the mouse CREB gene (Figure 2). The 12 amino acid sequence present in mouse CREM α and CREM β and absent in CREM γ , which exhibits no obvious homology to mouse CREB exon 9, is given in parentheses. (B) Nucleotide sequence comparison of the mouse CREB exon 11 and the sequence of the leucine zipper domain II of the CREM α cDNA clone indicates that the exon–intron boundary occurs at the identical position. For both regions part of the intron sequence and part of the exon sequence with the amino acid sequence are shown. The consensus AG of the splice site is shown in bold letters. (C) Schematic representation of the genomic organization of the mouse CREB and CREM genes. Introns, shown as disconnected lines, are not drawn to scale. Exons are indicated by numbers with additional highlighted features: CpG, CpG rich exon; ATG, translation initiation; Q, glutamine-rich transactivation domain; Ψ, α, γ and Ω, alternatively spliced exons of the CREB gene; DLSSD, additional transactivation domain; PK-C, protein kinase C consensus sequence; PK-A, protein kinase A consensus sequence; CKII, casein kinase II consensus sequence; b, Ω, basic region; ZIP, leucine zipper dimerization domain; TAA, translation termination (Hoeffler *et al.*, 1988; Gonzalez *et al.*, 1989, 1991; and this work). Striped boxes denote exons which are homologous between CREB and CREM. Black boxes indicate exons which are alternatively spliced in either CREB or CREM transcripts. White boxes denote exons which have been identified for the CREB gene but not for the CREM gene.

restriction enzymes identified a CpG island (Bird, 1987) at the 5' end of the cloned contig (Cole *et al.*, 1992). Subsequent DNA sequence analysis indicated that the putative mouse CREB gene promoter also resides in this CpG rich RNA region (Figure 2). There is no suppression of methylatable CpG dinucleotides, i.e. the percentage of CpG dinucleotides within the sequenced region is identical to GpC dinucleotides, which is thought to be indicative for promoters of housekeeping genes (Bird, 1987). The CREB gene promoter is devoid of obvious TATA or CCAAT box sequences, but possesses multiple putative binding site for the transcription factor SP1 (see Figure 2; Mitchell and Tjian, 1989). Within the putative promoter region we noticed the nucleotide sequence 5'-TCCTCAGTCGG-3', closely resembling the Inr (initiator) sequence 5'-PyCCTCAPy-TCTG-3', which has been identified by Smale and Baltimore (1989) to represent a discrete element in RNA polymerase II transcribed promoters lacking a TATA box. This element can act either alone or in concert with upstream elements, greatly increasing specific transcription initiation at the conserved A nucleotide within the Inr (Smale and Baltimore, 1989). Two SP1 binding sites are located at positions -36 and +36 with respect to the A in the Inr (Figure 2). This would be a favourable distance (Smale and Baltimore, 1989) to allow strong transcription from the CREB gene promoter. One of the major transcriptional start sites of the CREB gene was indeed mapped to the Inr sequence (Cole *et al.*, 1992).

The mouse CREB and CREM genes are highly homologous

Recently, Foulkes *et al.* (1991) reported the isolation of a gene, termed CREM (cAMP-responsive element modulator), encoding a protein highly homologous to the nuclear factor CREB. Within one cDNA clone, CREM α , they identified two almost identical DNA binding/leucine zipper domains of which the second motif was out of frame. By PCR analysis they detected a second CREM isoform, CREM β , in which this second DNA binding/leucine zipper motif was in frame. By DNA sequence analysis of the genomic region and by hybridization of a specific probe for the CREB leucine zipper domain, the presence of such a second leucine zipper motif was not evident in the cloned CREB contig (data not shown). Although the genomic CREM contig was not characterized, Foulkes *et al.* (1991) speculated that these CREM isoforms were generated by alternative splicing.

Knowledge of the genomic structure of the CREB gene strongly supports this hypothesis. The CREB exon 2, containing the ATG translation initiation codon, is identical in size and is similar in its amino acid composition to that of CREM (Figure 4A). Furthermore, by aligning the amino acid sequences of CREB and CREM we observed that homologous regions can be unambiguously assigned to individual exons identified for the mouse CREB gene (see Figures 2-4). It was obvious that in addition to the homology in exon 2 the corresponding sequences of exons 7, 8, 10 and 11 are nearly identical (Figure 4A). The homology of CREM to CREB breaks off exactly at the positions of identified CREB exon boundaries. The size of the CREM protein isoform CREM γ (217 amino acid residues) is also in agreement with the overall coding capacity of the CREB exons 2, 7, 8, 10 and 11. The 12 unique amino acid residues of the CREM α , missing in CREB γ (Foulkes

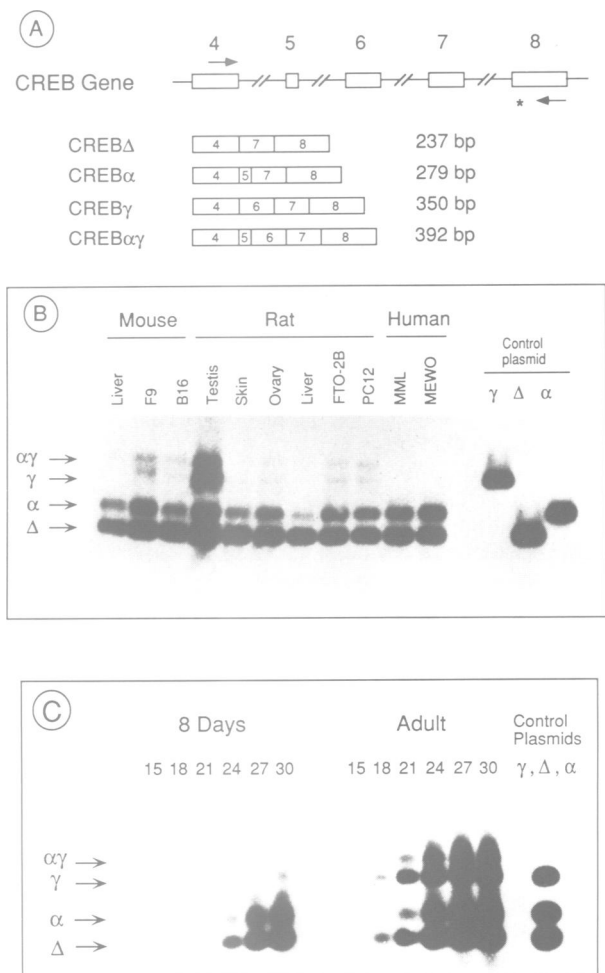


Fig. 5. PCR analysis of CREB isoform transcripts reveals high expression of CREB γ mRNA in testis and identifies a new mRNA isoform, CREB $\alpha\gamma$. (A) Schematic representation of part of the mouse CREB gene. Boxes numbered 4-8 indicate exon sequences. Introns are shown as disconnected lines. The position and direction of the two PCR primers located in exon 4 and 8 are indicated by arrows. The asterisks denotes the position of the oligonucleotide used as a probe for hybridization in B. The expected exon compositions and sizes of the CREB isoforms are indicated. (B) Southern blot hybridization analysis of PCR products. The mouse, rat and human tissues and cell lines used as a source of RNA preparation, for reverse transcription and subsequent PCR are indicated above the lanes. Plasmids used as controls were pmcCREB γ (γ), pmcCREB Δ (Δ) and pmcCREB α (α). Arrows on left point to Δ , α , γ and $\alpha\gamma$ specific amplification products. The PCR products were hybridized with an oligonucleotide indicated by the asterisks in A, detecting all CREB isoforms. For unambiguous assignment of individual CREB isoform structures within the amplified DNA fragments, the PCR products were analysed in addition by Southern blot hybridization using CREB isoform-specific oligonucleotides as probes (see Materials and methods; data not shown). (C) Southern blot hybridization analysis of PCR products from testis RNA. Polyadenylated RNA for reverse transcription and subsequent PCR was prepared from testis of 8-day-old and adult mice. PCR was performed as in (B) at various cycles of amplification as indicated. Arrows on the left point to Δ , α , γ and $\alpha\gamma$ specific amplification products. The PCR products were hybridized with an oligonucleotide indicated by the asterisks in A, detecting all CREB isoforms. Plasmids used as controls were as described in (B).

et al., 1991) and occurring at the equivalent position of CREB exon 9 (encoding 61 amino acid residues) reveals no obvious homology to the mouse CREB gene and therefore most likely represents an alternatively spliced exon of the

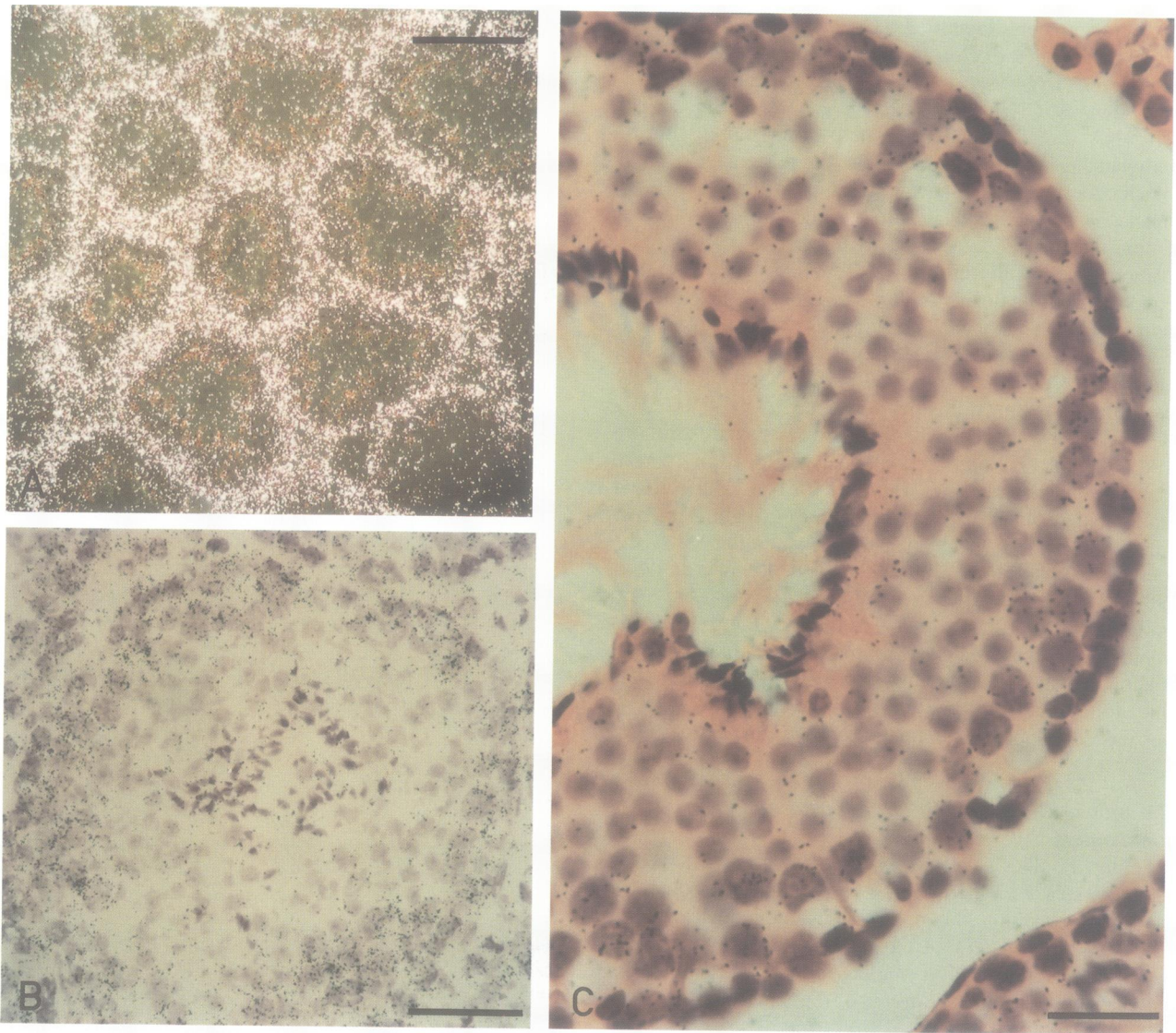


Fig. 6. CREB is specifically expressed in primary spermatocytes. (A) Dark-field photograph of a cryostat section from mouse testis hybridized with a ^{35}S -labelled riboprobe as described in Materials and methods. The bar represents 200 μm . (B) Same as A but bright-field photography showing strong hybridization signals in the outer cell region of a single tubulus. The bar represents 65 μm . (C) Bright-field photography of paraffin embedded mouse testis. Strongest hybridization is observed in cells identified as spermatocytes I. Secondary spermatocytes and mature spermatides reveal expression levels close to background. The bar represents 23 μm .

CREM gene. Also note that exon 9 of the mouse CREB gene is alternatively spliced in the CREB Ω isoform (Figure 3).

As inferred from the CREB α cDNA isoform (Foulkes *et al.*, 1991), the overall similarity in the gene structure between CREB and CREM is best documented by the exon–intron boundary of CREB exon 11. This exon–intron boundary separating the basic region from the leucine zipper domain is positioned in the CREB gene where the sequences of the two isoforms of CREM begin to diverge (Figure 4B; Foulkes *et al.*, 1991). The striking homologies described here, lead us to postulate a similar genomic organization for both the CREB and the CREM gene (see Figure 4C and Discussion).

Specific expression of CREB isoforms in the testis

Isolation of multiple CREB isoforms from a single cDNA library (for example cDNAs corresponding to CREB Δ , CREB α , and CREB γ) were isolated from a F9 teratocarcinoma cDNA library and CREB Δ , CREB γ , and CREB Ω isoforms were isolated from a B16 melanoma cDNA library)

suggests that multiple isoforms are expressed within the same cell. To characterize further the expression of CREB isoforms, we used reverse transcription/PCR and *in situ* hybridization to analyse expression in various tissues and cell lines. To detect expression of CREB isoforms in RNAs prepared from mouse, rat and human samples primers were chosen from conserved sequences located in exon 4 and exon 8 (see Figure 5A and Materials and methods). The results of this PCR analysis, shown in Figure 5B, clearly demonstrate that CREB α and CREB Δ are expressed in all tissues and cell lines tested. CREB Δ showed a 3 to 5-fold higher level of expression compared to CREB α , reflecting the relative abundance of the two isoforms isolated by cDNA screening (Figure 1) and correlates with the RNase protection experiments of Berkowitz and Gilman (1990). With the exception of human melanoma cell lines tested, the γ -isoform is also ubiquitously expressed, however, at a lower level compared to CREB Δ and CREB α . In addition to the Δ -, α - and γ -isoforms, hybridizations to defined oligonucleotide probes (data not shown) identified a species,

called CREB $\alpha\gamma$, which was not isolated by our cDNA screening. All tissues and cell lines expressing the α - and γ -isoforms also expressed this CREB $\alpha\gamma$ isoform but at a low level.

Surprisingly, we found a high level of expression of the γ isoforms in testis. Since it was observed that specific isoforms of the CREM gene are generated by alternative splicing during testis differentiation (Foulkes *et al.*, 1992), the expression of CREB isoforms in testis was analysed further before and after the onset of spermatogenesis. Testis RNA from 8-day-old and adult mice was analysed by reverse transcription/PCR as described above at different PCR cycle numbers. Results are shown in Figure 5C. In testis RNA from 8-day-old mice the level of expression of the γ - and $\alpha\gamma$ - isoforms was found to be similar to that of other tissues tested. In adult mice, after the onset of spermatogenesis, levels of the γ - and $\alpha\gamma$ - isoforms increased dramatically to a similar level of the Δ - and α - isoforms. To be certain that our PCR reaction only amplified DNA fragments arising from CREB mRNAs, another CREB-specific 5' PCR primer from CREB exon 2 and a CREM-specific primer from the homologous part of the CREM gene was prepared. PCRs were repeated on testis RNA samples and confirmed the amplification and detection of only CREB-specific PCR products (data not shown).

Northern blot analysis detected a high level of CREB expression in testis and brain (data not shown). We were particularly interested in identifying the cell types expressing the CREB gene and therefore analysed expression of CREB mRNA by *in situ* hybridization. Whereas CREB gene expression in brain does not show any cell type preference (data not shown), CREB mRNA is abundant in a defined region of the testis (Figure 6A). The primary spermatocytes (predominantly those of the pachytene stage) express high CREB mRNA levels (Figure 6B and 6C) whereas the secondary spermatocytes and mature spermatids either do not express CREB or express it at an undetectable level. Thus, based on our PCR analysis (Figure 5) which showed a high expression of the γ -isoform, we conclude that the primary spermatocytes express higher levels of the CREB γ and CREB $\alpha\gamma$ mRNA isoforms in addition to CREB Δ and CREB α .

CREB α and CREB Δ display identical transactivation potentials

The differentially spliced CREB isoforms with intact DNA binding and dimerization domains, CREB α and CREB Δ , were tested for their CRE-mediated transactivation potentials in the transient cotransfection assay used by Gonzalez and Montminy (1989). We compared cDNA isoforms differing solely by the inclusion of individual exon sequences within an otherwise identical sequence content. Full-length mouse cDNAs for CREB α and CREB Δ , cloned into the eukaryotic expression plasmid pHD (Müller *et al.*, 1988), were cotransfected into F9 teratocarcinoma cells together with the somatostatin promoter-CAT plasmid ($\Delta(-71)$ SOM-CAT; Montminy *et al.*, 1986) as a reporter and an expression plasmid encoding the catalytic subunit of protein kinase A (PKA) (Uhler and McKnight, 1987). With increasing amounts of the CREB expression plasmids we observed a steady increase of transactivation with both CREB isoforms, up to 70-fold (Figure 7). A much lower, but significant transactivation was observed when the catalytic subunit of

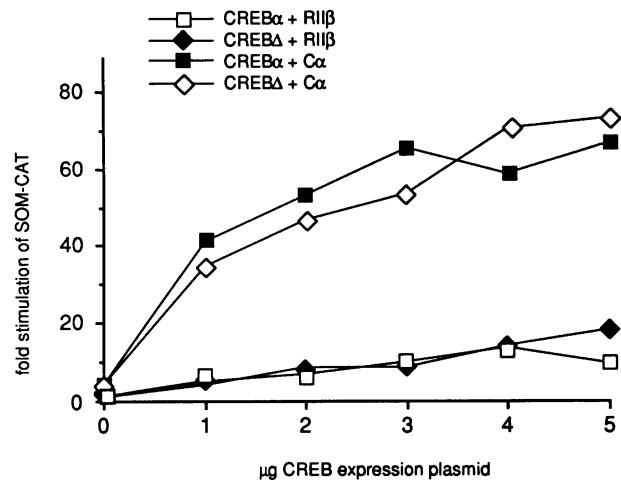


Fig. 7. The isoforms CREB α and CREB Δ have similar transactivation potentials. Increasing amounts of CREB α and CREB Δ expression plasmids were cotransfected with a somatostatin promoter construct ($\Delta(-71)$ SOM-CAT; Montminy *et al.*, 1986) as a reporter, together with expression plasmids encoding either the catalytic (C α) or regulatory (RII β) subunit of protein kinase A into undifferentiated F9 teratocarcinoma cells. CAT activities are corrected for β -galactosidase activity of an internal reference and calculated as fold stimulation, setting the control without CREB expression plasmid to 1.

protein kinase A was omitted (data not shown) or when the regulatory subunit RII β (Büchler *et al.*, 1990) was cotransfected (Figure 7). We conclude that the mouse CREB α and CREB Δ proteins are equally able to transactivate a CRE-containing reporter construct in this cotransfection system.

As the CREB γ isoform lacks the DNA binding domain and the leucine zipper (see Figure 2) but contains the glutamine-rich transactivation domain (Gonzalez *et al.*, 1991; Courey and Tjian, 1988) we considered the possibility that this form could function by modulating CREB activity, e.g. in a dominant negative fashion, like squelching (Ptashne, 1988). Therefore an expression plasmid encoding CREB γ , pHDmCREB γ , was cotransfected with the $\Delta(-71)$ SOM-CAT reporter plasmid into rat PC12 cells which contain endogenous CREB (Yamamoto *et al.*, 1990). No effect of CREB γ on CRE-mediated transcription could be detected (data not shown).

Discussion

Functional significance of mouse CREB mRNA isoforms generated by alternative splicing

Six different mouse CREB mRNAs, CREB Δ , CREB α , CREB γ , CREB $\alpha\gamma$, CREB Ψ and CREB Ω , which differ in their coding information were identified. These isoforms are generated by alternative splicing of a single primary gene transcript and lead to the inclusion or exclusion of different exon sequences, termed Ψ , α , γ and Ω . CREB Δ and CREB α are the most abundant CREB isoforms. CREB Δ , which was expressed in all cells and tissues tested, encodes a protein of 327 amino acids (Figure 2). Incorporation of the α -domain, yielding CREB α , does not interfere with the amino acid reading frame, however, inclusion of the exon sequences designated Ψ and γ would lead to termination of translation upstream of the assigned stop codon (Figure 2).

The generation of multiple mRNA transcripts of a single

gene by alternative splicing has been documented for a variety of transcription factors, including the CCAAT box binding proteins CTF/NF-1 (Santoro *et al.*, 1988), retinoic acid receptors α , β and γ (Leroy *et al.*, 1991; Zelent *et al.*, 1991; Krust *et al.*, 1989), the *Oct-2* gene (Wirth *et al.*, 1991) and most likely the CREM gene (Foulkes *et al.*, 1991, 1992). Differential activities of alternate CTF/NF-1 isoforms have been described (Santoro *et al.*, 1988), but in many cases the cell and tissue-specific expression and the functional significance of various isoforms of a particular gene is not known.

We analysed the two CREB isoforms, CREB α and CREB Δ , which include the basic region and the leucine zipper dimerization domain, in a transactivation assay. Both mouse CREB isoforms showed a very similar potential to transactivate a CRE-containing reporter construct in undifferentiated F9 cells (Figure 7). This result is in agreement with transfection experiments performed by Berkowitz and Gilman (1990). They transfected human CREB α and CREB Δ isoforms, combined with the DNA binding domain of the yeast transcriptional activator protein GAL4 into BALB/c3T3 cells and also did not observe a functional difference between these CREB isoforms. However, in one report using rat CREB isoforms, the α -domain was implicated in enhancing CREB mediated transactivation from a somatostatin promoter-CAT construct (Yamamoto *et al.*, 1990). As mouse and rat CREB proteins are very similar, we do not consider the possibility that the different outcome of our experiments and those described by Yamamoto *et al.* (1990) could be attributed to the minor amino acid residue exchanges. Thus, at least by these experiments the role of the α -domain in transactivation seems to be questionable.

The incorporation of exon 3 (Ψ) or exon 6 (γ) into the mature mRNA would generate shorter CREB proteins, lacking DNA binding and leucine zipper domains. These proteins would include the glutamine-rich transactivation domain (CREB γ) and in addition the α -domain (CREB $\alpha\gamma$). The abundance of the CREB γ mRNA isoform in a specific cell type, e.g. the primary spermatocytes of the testis (see below), might be taken as an argument for a specific role. Some transcription factors such as SP1, Oct-1 and Oct-2 stimulate transcription through similar glutamine-rich domains (Courey *et al.*, 1989; Tanaka and Herr, 1990). This suggests that CREB γ might downregulate CREB mediated activation by acting as a transdominant inhibitor (Ptashne, 1988). In a series of cotransfection experiments, however, we could not obtain evidence to support this hypothesis. Thus, whereas CREB Δ and CREB α encode transcription factors with DNA binding activity and transactivation potential (Figure 7), the functional significance of additional CREB isoforms in the cAMP signalling process remains elusive.

Specific CREB mRNA isoforms are more highly expressed in primary spermatocytes

Elevated expression of the CREB gene was detected in primary spermatocytes by *in situ* hybridization (Figure 6) and expression of at least four CREB mRNA species, CREB Δ , CREB α , CREB γ , and CREB $\alpha\gamma$, has been demonstrated by PCR in testis (Figure 5). In all other cells and tissues tested, mRNAs for the CREB γ and CREB $\alpha\gamma$ isoforms are minor splicing variants. During spermatogenesis a significant switch in splicing of CREB transcripts leads

to higher levels of these two CREB mRNA isoforms and suggests a specific function for these isoforms in spermatogenesis. As discussed above, the functional implication of splicing products like CREB γ and CREB $\alpha\gamma$ in the cAMP signalling pathway remains elusive. Nevertheless, we can speculate that the switch to generation of higher levels of CREB γ mRNAs in testis may reflect an initial event of downregulation of CREB-mediated gene expression by lowering mRNA species encoding active CREB forms. Consistent with this is the observation that the regulatory subunit RI α of PKA, which when bound to the catalytic subunit prevents activation of CREB by phosphorylation, shows a particularly high level of expression in the primary spermatocytes (Oyen *et al.*, 1987). Both the high level of CREB γ isoforms and RI α expression would lead to a reduction in CREB-dependent gene transcription. The CREM gene has recently been shown to be expressed in testis where cell-specific alternative splicing generates a novel CREM isoform, called CREM τ (Foulkes *et al.*, 1992). Unlike the previously characterized CREM isoforms, α , β , and γ which function as antagonists of the cAMP transcription activation pathway, CREM τ functions as a transcriptional activator. During the onset of spermatogenesis there is an abrupt switch in CREM gene expression from the generation of the CREM antagonists to the generation of the CREM τ activator. Therefore this coincides with the generation of CREB γ isoforms in testis. Taken together with the similarities in gene structure of CREB and CREM, this may reflect a common regulatory switch in the mechanism of alternate splicing for these two genes leading to production of a CREM activator and down regulation of CREB activation functions.

The CREB and CREM gene have a similar organization in genomic structure

Determination of the structural organization of the mouse CREB gene has allowed alignment of similar regions of CREB mRNAs to the defined exons of the CREB gene. The structural organization of the mouse CREM gene has yet to be determined but based on these comparisons we suggest a similar gene structure for the mouse CREM gene (Figure 4C). The DNA binding and leucine zipper dimerization domains of CREB and CREM are nearly identical and both bind to a CRE with the same affinity (Foulkes *et al.*, 1991). The similarity in structural organization is also in keeping with the speculation made by Foulkes *et al.* (1991) that expression of the CREM gene is subject to alternate splicing.

Recent characterization of another isoform of the CREM gene, called CREM τ (Foulkes *et al.*, 1992), strengthens the similarity found between CREM and CREB. CREM τ contains two additional segments, compared with the other CREM isoforms and these two insertions correspond to exons 4 and 9 of the CREB gene. In structure CREM τ most closely resembles CREB Δ . Unlike the other CREM isoforms, but similar to CREB, CREM τ is an activator of the cAMP induced signal transduction pathway. The inserted sequence in CREM τ similar to exon 4 of CREB is rich in glutamine residues. This domain has recently been implicated in the transactivation function of CREB (Gonzales *et al.*, 1991). The presence of the second inserted sequence in CREM τ , similar to exon 9 of CREB, could be important in correct spacing of domains and thus important in

Acknowledgements

We thank J.P.Hoeffler and J.F.Habener for providing the human CREB G1 cDNA clone, M.Montminy for providing the somatostatin promoter-CAT construct $\Delta(-71)\text{SOM-CAT}$, G.S.McKnight for providing the expression plasmid encoding the catalytic subunit of PKA ($C_{\alpha}\text{EV}$); U.Walter and M.Meinecke for providing the expression plasmid encoding the regulatory subunit RII_{β} of PKA (REV) and W.Fleischer for oligonucleotide synthesis and photographic work. We are thankful to I.Grummt, C.de Vack, U.Walter and Th.Wirth for helpful suggestions and comments on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft through SFB 229 and the Leibniz Programm and the Fonds der Chemischen Industrie.

References

- Berkowitz, L.A. and Gilman, M.Z. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5258–5262.
- Bird, A.P. (1987) *Trends Genet.*, **3**, 342–347.
- Bokar, J.A., Roesler, W.J., Vandenbark, G.R., Kaetzel, D.M., Hanson, R.W. and Nilson, J.H. (1988) *J. Biol. Chem.*, **263**, 19740–19747.
- Boshart, M., Weih, F., Schmidt, A., Fournier, R.E.K. and Schütz, G. (1990) *Cell*, **61**, 905–916.
- Boshart, M., Weih, F., Nichols, M. and Schütz, G. (1991) *Cell*, **66**, 849–859.
- Büchler, W., Meinecke, M., Chakraborty, T., Jahnsen, T., Walter, U. and Lohmann, S.M. (1990) *Eur. J. Biochem.*, **188**, 253–259.
- Cole, T., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Schütz, G. and Ruppert, S. (1992) *Genomics*, in press.
- Comb, M., Birnberg, N.C., Seasholtz, A., Herbert, E. and Goodman, H.M. (1986) *Nature*, **323**, 353–356.
- Courey, A.J. and Tjian, R. (1988) *Cell*, **55**, 887–898.
- Courey, A.J., Holtzman, D.A., Jackson, S.P. and Tjian, R. (1989) *Cell*, **59**, 827–836.
- Delegeane, A.M., Ferland, L.H. and Mellon, P.L. (1987) *Mol. Cell. Biol.*, **7**, 3994–4002.
- Deutsch, P.J., Hoeffler, J.P., Jameson, J.L. and Habener, J.F. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7922–7926.
- Duboule, D. and Dollé, P. (1989) *EMBO J.*, **8**, 1497–1505.
- Dwarki, V.J., Montminy, M. and Verma, I.M. (1990) *EMBO J.*, **9**, 225–232.
- Fisch, T.M., Prywes, R., Simon, M.C. and Roeder, R.G. (1989) *Genes Dev.*, **3**, 198–211.
- Foulkes, N.S., Borrelli, E. and Sassone-Corsi, P. (1991) *Cell*, **64**, 739–749.
- Foulkes, N.S., Mellström, B., Benusiglio, E. and Sassone-Corsi, P. (1992) *Nature*, **355**, 80–84.
- Gonzalez, G.A., Yamamoto, K.K., Fischer, W.H., Karr, D., Menzel, P., Biggs, W. III, Vale, W.W. and Montminy, M.R. (1989) *Nature*, **337**, 749–752.
- Gonzalez, G.G. and Montminy, M.R. (1989) *Cell*, **59**, 675–680.
- Gonzalez, G.A., Menzel, P., Leonard, J., Fischer, W.H. and Montminy, M.R. (1991) *Mol. Cell. Biol.*, **11**, 1306–1312.
- Graham, F.L. and van der Eb, A.J. (1973) *Virology*, **52**, 456–467.
- Hoeffler, J.P., Meyer, T.E., Yun, Y., Jameson, J.L. and Habener, J.F. (1988) *Science*, **242**, 1430–1433.
- Johnson, P.F. and McKnight, S.L. (1989) *Annu. Rev. Biochem.*, **58**, 799–839.
- Kerppola, T.K. and Curran, T. (1991) *Cur. Opin. Struct. Biol.*, **1**, 71–79.
- Kozak, M. (1986) *Cell*, **44**, 283–292.
- Krust, A., Kastner, Ph., Petkovich, M., Zelent, A. and Chambon, P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5310–5314.
- Lee, C.Q., Yungdae, Y., Hoeffler, J.P. and Habener, J.F. (1990) *EMBO J.*, **9**, 4455–4465.
- Leroy, P., Krust, A., Zelent, A., Mendelsohn, C., Garnier, J.-M., Kastner, P., Dierich, A. and Chambon, P. (1991) *EMBO J.*, **10**, 59–69.
- Luckow, B. and Schütz, G. (1991) In Mcpherson, M. (ed.), *Directed Mutagenesis: A Practical Approach*, IRL Press, Oxford, pp. 101–133.
- Mitchell, P.J. and Tjian, R. (1989) *Science*, **245**, 371–378.
- Montminy, M.R., Sevarino, K.A., Wagner, J.A., Mandel, G. and Goodman, R.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 6682–6686.
- Montminy, M.R. and Bilezikjian, L.M. (1987) *Nature*, **328**, 175–178.
- Müller, G., Ruppert, S., Schmid, E. and Schütz, G. (1988) *EMBO J.*, **7**, 2723–2730.
- Nitsch, D. *et al.* (1991) In Cohen, P. and Foulkes, J.G. (eds), *The hormonal control regulation of gene transcription*. Elsevier Science Publishers B.V., The Netherlands, pp. 223–234.
- Oyen, O., Froyosa, A., Sandberg, M., Eskild, W., Joseph, D., Hansson, V. and Jahnsen, T. (1987) *Biol. Reprod.*, **37**, 947–956.
- Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) *Annu. Rev. Biochem.*, **55**, 1119–1150.
- Ptashne, M. (1988) *Nature*, **335**, 683–689.
- Quinn, P.G., Wong, T.W., Magnuson, M.A., Shabb, J.B. and Granner, D.K. (1988) *Mol. Cell. Biol.*, **8**, 3467–3475.
- Raymond, V., Atwater, J.A. and Verma, I.M. (1989) *Oncogene Res.*, **5**, 1–12.
- Ruppert, S., Müller, G., Kwon, B. and Schütz, G. (1988) *EMBO J.*, **7**, 2715–2722.
- Ruppert, S., Boshart, M., Bosch, F.X., Schmid, W., Fournier, R.E.K. and Schütz, G. (1990) *Cell*, **61**, 895–904.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Santoro, C., Mermod, N., Andrews, P.C. and Tjian, R. (1988) *Nature*, **334**, 218–224.
- Schöler, H.R., Ruppert, S., Suzuki, N., Chowdhury, K. and Gruss, P. (1990) *Nature*, **344**, 435–439.
- Shaw, G. and Kaman, R. (1986) *Cell*, **46**, 659–667.
- Short, J.M., Fernandez, J.M., Sorge, J.A. and Huse, W.D. (1988) *Nucleic Acids Res.*, **16**, 7583–7600.
- Silver, B.J., Bokar, J.A., Virgin, J.B., Vallen, E.A., Milsted, A. and Nilson, J.H. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 2198–2202.
- Smale, S.T. and Baltimore, D. (1989) *Cell*, **57**, 103–113.
- Tanaka, M. and Herr, W. (1990) *Cell*, **60**, 375–386.
- Taylor, J.W., Ott, J. and Eckstein, F. (1985) *Nucleic Acids Res.*, **13**, 8765–8785.
- Tsukada, T., Fink, J.S., Mandel, G. and Goodman, R.H. (1987) *J. Biol. Chem.*, **262**, 8743–8747.
- Uhler, M.D. and McKnight, G.S. (1987) *J. Biol. Chem.*, **262**, 15202–15207.
- Weih, F., Stewart, F., Boshart, M., Nitsch, D. and Schütz, G. (1990) *Genes Dev.*, **4**, 1437–1449.
- Wirth, T., Priess, A., Annweiler, A., Zwilling, S. and Oeler, B. (1991) *Nucleic Acids Res.*, **19**, 43–51.
- Yamamoto, K.K., Gonzalez, G.A., Biggs, W.H. III and Montminy, M.R. (1988) *Nature*, **334**, 494–498.
- Yamamoto, K.K., Gonzalez, G.A., Menzel, P., Rivier, J. and Montminy, M.R. (1990) *Cell*, **60**, 611–617.
- Zelent, A., Mendelsohn, C., Kastner, P., Krust, A., Garnier, J.-M., Ruffenach, F., Leroy, P. and Chambon, P. (1991) *EMBO J.*, **10**, 71–81.

Received on September 11, 1991; revised on January 15, 1991