Characterization of the testis-specific gene 'calmegin' promoter sequence and its activity defined by transgenic mouse experiments

Daisuke Watanabe, Masaru Okabe, Noriyuki Hamajima, Takashi Morita, Yukio Nishina, Yoshitake Nishimune*

Research Institute for Microbial Diseases, Osaka University, 3-1 Yamadaoka, Suita, Osaka 565, Japan

Received 1 May 1995; revised version received 15 June 1995

Abstract We have cloned the genomic DNA of calmegin [(1992) J. Biol. Chem. 269, 7744–7749] and analyzed its promoter region. It contained GC-rich sequences and potential binding sites for AP 2 and Sp 1, but lacked the TATA sequence. The 330 bp 5′ flanking sequence of calmegin genomic DNA fused with the CAT gene was used for the study of promoter activity in transgenic mice. The CAT gene activity was detected exclusively in testes, indicating that the 330 bp calmegin 5′ sequence was sufficient for the testis-specific expression. The existence of testicular nuclear factors specifically bound to the putative promoter sequence was also demonstrated.

Key words: Spermatogenesis; Testis; Transcription factor; Specific promoter; Ca^{2+} binding protein; Transgenic mouse

1. Introduction

Mammalian spermatogenesis is a complex process including the proliferation of spermatogonia, meiosis of spermatocytes and drastic morphological changes from spermatid to sperm. During the process many male germ cell-specific molecules were expressed utilizing the germ cell-specific transcriptional systems [2,3]. Recently some of these genes were isolated by using germ cell-specific antibodies or identified as related molecules of some somatic genes. However, the germ cell-specific genes reported so far are not so many. Moreover, the lack of male germ cell lines makes it difficult to study the germ cell-specific transcriptional regulation. These are the reasons why the structure and the regulatory mechanism of male germ cell-specific promoters are not well understood yet.

Calmegin is a Ca^{2+} binding protein specifically and abundantly expressed in meiotic germ cells from pachytene spermatocytes to spermatids. We have recently molecularly cloned the cDNA encoding calmegin using specific monoclonal antibody [1,4]. Although the precise function of calmegin is still unclear, a high homology to the calreticulin (the major Ca^{2+} binding protein of ER membrane) as well as to the calnexin (ER membrane phosphoprotein) [5–7] implied its physiological function involving Ca^{2+} binding. In fact, we have demonstrated a Ca^{2+} binding activity of this protein [1]. Furthermore, calmegin mRNA was expressed only in the male meiotic germ cells and not in other somatic cells [1].

From these lines of evidence it is conceivable that some *cis*-acting element in the calmegin promoter region and its *trans*-acting factors are present in testes for germ cell-specific expression.

In the present study, to examine these possibilities, we have isolated the genomic DNA of calmegin gene, and analyzed the promoter sequence and transcriptional starting sites. Furthermore, we produced transgenic mice containing a calmegin-CAT fusion gene and analyzed promotor activities in variuous tissues of the transgenic mice. Here we showed that the 330-bp calmegin 5′ flanking region was sufficient to direct testis-specific gene expression and that some testicular nuclear factors bound specifically to this promoter sequence.

2. Materials and methods

2.1. Molecular cloning and DNA sequencing of the calmegin genomic gene

Liver DNA obtained from 129 mice was digested with endonuclease *Sau*3AI and cloned into the *Bam*HI site of λ phage EMBL3. About 3×106 recombinant phage plaques were screened with a radio-labeled *Pst*I-*Eco*RI 700-bp 5′ region of the calmegin cDNA [1]. Recombinant phage DNA was isolated by the standard procedures. Three kbp DNA fragment containing a 178-bp calmegin 5′ untranslated region at 3′ end was isolated by digestion with *Bam*HI and subcloned into a pBluescript SK+ vector, and sequenced by the dideoxy chain termination method.

2.2. Plasmid construction and microinjection

The 330-bp *Sac*I/*Bam*HI fragment of the genomic calmegin gene containing the 178-bp 5′ untranslated first exon and 152-bp upstream sequence from the transcription initiation site was cloned into the *Bam*HI/*Sac*I restriction sites of pUC O CAT vector [8]. From the resulting fusion construct pMegSB-CAT, the DNA fragment was excised by *Pvu*II/*Sac*I digestion and injected into male pronucleus of fertilized mouse eggs [27]. Founder transgenic mice were identified by Southern blot hybridization using a ^{32}P-labeled CAT probe.

2.3. CAT enzyme assay

Various tissues from hemizygous transgenic mice were homogenized in 10 mM Tris-HCl, (pH 7.8), heated at 60°C for 10 min [28], and centrifuged at $10000 \times g$ for 5 min. CAT activity of extracted proteins (20 mg) in each supernatant was analyzed as described by Gorman et al. [29].

2.4. Primer extension analysis of the calmegin gene

A synthetic oligonucleotide complementary to the sequence from nucleotide position +93 to +122 of the calmegin cDNA (Fig. 1A) was endolabeled with γ[^{32}p] ATP by T4 polynucleotide kinase. Primer-RNA (30 μg) hybridization and extension reactions were performed as described by Bina-Stein et al. [30].

2.5. Gel mobility shift assay

Nuclear extracts were prepared from testis and liver essentially as described by Schreiber et al. [31]. Each nuclear extract (15 μg) was incubated at 4°C for 30 min with 1 μg non-specific DNA (poly dAdT) in a 30 μl volume containing 10% SDS, 12 mM HEPES (pH 7.9), 60 mM KCl, 0.5 mM EDTA, 0.6 mM DTT, and 0.1 μg of the 5′ endolabeled calmegin genomic DNA fragment from −152 to +18. The mix-

*Corresponding author. Fax: (81) (6) 879-8339.

SSDI 0014-5793(95)00729-6

tures were loaded on a 4% non-denaturing polyacrylamide gel and electrophoresed. The gel was then dried and autoradiographed.

3. Results

3.1. Molecular cloning of a 5′-calmegin flanking sequence and its primer extension analysis

From the λEMBL3 129 male mouse genomic liberary, we obtained a 5′ flanking region of the calmegin genomic DNA. The 330 bp calmegin genomic sequence is shown in Fig. 1A. The calmegin 5′ flanking sequence contained multiple GC-rich (about 73%) sequences, and putative binding sequences for transcription factors such as AP2 and SP I. It does not possess canonical TATA and CCAAT boxes in the typical locations upstream of the transcriptional starting site (Fig. 1A).

To identify the transcriptional initiation site of calmegin gene, we performed a primer extension experiment using a oligonucleotide complementary to a 5′ untranslated region of calmegin cDNA. One strong and three weak signals were detected only in the testis but not in the liver RNAs (Fig. 2). From the RNase mapping analysis, we also detected one major and some minor transcriptional starting sites in the same regions (data not shown), consistent with the data of the primer extension experiment. Although calmegin gene may have several minor transcription initiation sites, we concluded that the major transcription initiation site was one strongly identified in both primer extention and RNase mapping experiments (see '+1' in Fig. 1A).

3.2. Generation of transgenic mice carrying the 5′ flanking region of calmegin gene fused with the CAT gene

The fusion gene used in this experiment is shown in Fig. 1B. It contained both the 152-bp upstream and 178-bp downstream non-coding sequences from the transcriptional initiation site of the calmegin gene (Fig. 1B). This segment was fused to the *Sac*I/*Bam*HI sites of the vector pUC O CAT [8]. DNA with the construct was purified and injected into the single-cell fertilized mouse embryos, which were then transferred to the pseudopregnant foster mothers. The animals containing the transgene were identified by Southern blot hybridization of tail DNA using the CAT DNA as a probe. Two independent lines of transgenic mice TGcal01 and TGcal02 were established. Both lines were estimated to carry about 10 copies of the transgene per genome by Southern blotting analysis as compared with known standards (data not shown).

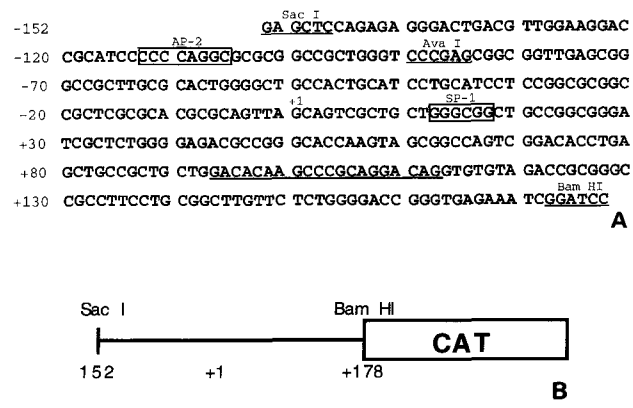

Fig. 1. A 5′ flanking DNA sequence of the genomic calmegin gene and a construction used for generating transgenic mice. (A) Nucleotides were numbered from the first nucleotide of the transcription initiation site (+1). Restriction enzyme sites were indicated by underlines. The putative binding sites for nuclear factors (AP 2 and SP 1) are indicated by boxes. A synthetic oligonucleotide complementary to the underlined sequence (from +93 to +122) was used for primer extention analysis. (B) The construct injected into fertilized mouse embryos. The construction of the calmegin-CAT fusion gene is described in section 2.

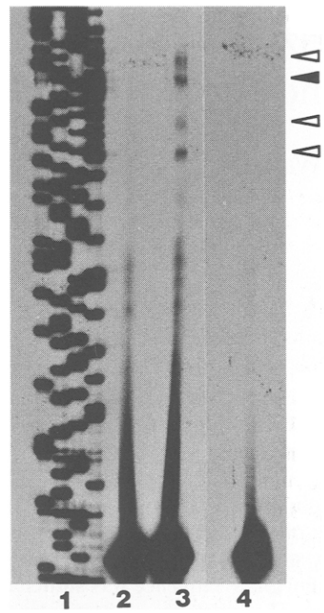

Fig. 2. Primer extension analysis of the calmegin gene. Total RNA (30 μg) was hybridized with 5′ ^{32}P-labeled synthetic oligonucleotides complementary to the DNA sequence of the calmegin cDNA from nucleotide position +93 to +122 . Primer extension products were synthesized and analyzed as described in section 2. Lane 1 = M13 DNA sequence standard to determine each size of primary extended products; lane 2 = 30 μg of total RNA of liver; lane 3 = 30 μg of total RNA from the testis; lane 4 = no RNA. A closed arrowhead indicates the major primary extended product and open arrowheads indicate the minor ones.

3.3. Tissue-specific expression of the recombinant gene

Tissues from the adult TGcal01 (hemizygous) mouse were homogenized and tested for CAT enzymatic activity by thin layer chromatography. As shown in Fig. 3, CAT activity was found only in testis, not in any other tissues examined (e.g. liver, lung, heart, spleen, kidney, muscle, brain, intestine and ovary). The same expression pattern was also observed in another transgenic mouse line TGcal02. These results indicated that the *cis*-acting element necessary for testis-specific expression of the calmegin gene was present within the 330 bp of the 5′ flanking sequence.

3.4. Identification of testicular nuclear factors specifically binding to calmegin promoter sequences

To examine factors binding specifically to the calmegin promoter region, a 170-bp calmegin genomic sequence from nucleotide position −152 to +18 was subjected to a gel shift assay using a testis and a liver nuclear extract. Three retarded bands were observed only when the testis nuclear extract was used (Fig. 4, lane 2). Futhermore, the three bands disappeared in the presence of 100-fold excess of the unlabeled fragments, which were used as a competitor (Fig. 4, lane 3). No retarded bands were observed with the nuclear extract prepared from the liver

(Fig. 4, lane 4). These results strongly suggested that the factors binding specifically to the calmegin promoter region existed in testes.

4. Discussion

Calmegin (Meg 1) is a testis-specific protein that was characterized by a monoclonal antibody TRA 369 [4]. Using the antibody, we have recently cloned and characterized the mouse calmegin cDNA [1]. It was found that calmegin has high homologies to the calnexin and calreticulins. Both of these proteins have a Ca^{2+} binding ability and were reported to work as a molecular chaperone of the membrane receptor proteins [9,10] and a transcriptional repressor of the glucocorticoid receptor proteins [11,12]. The Ca^{2+} binding ability of calmegin was also demonstrated by using the Ca^{2+} overlay experiment [1]. Although the precise function is still unclear, the result may indicate that calmegin has a function similar to that of either calreticulins or calnexin.

The transcription of the calmegin gene occurs exclusively in testicular germ cells and not in any other somatic tissues. Thus, expression of the calmegin gene is highly regulated. Moreover, the expression is limited to pachytene spermatocytes and spermatids. These results suggest that calmegin has some important roles in meiosis and spermatogenesis.

Using the calmegin cDNA, we have isolated the calmegin genomic DNA. The promoter region of the calmegin gene contained a number of GC-rich sequences and did not have a canonical TATA or CCAAT box in the typical location upstream of the transcriptional starting site (Fig. 1A). The GC-rich sequences were also found in some testis-specific genes such as mouse proenkephelin and T complex genes [13–15], and may have a role in testis-specific gene expression responsible for DNA methylation or chromatin conformation [16–18]. It appears that calmegin gene has multiple transcriptional initiation sites as demonstrated by primer extension (Fig. 2) and RNase mapping analyses (data not shown). The presence of multiple initiation sites was also reported in other testis-specific genes [14,19]. It was also reported that some genes with multiple promoters often exhibited developmental stage specific expression [20]. The presence of the GC-rich sequences and some promoters could allow an additional control and greater flexibility for calmegin gene expression. Flanking sequences of some genes at 5′ were responsible for testis-specific gene expression [21–24]. However, we could not detect any conserved sequences within the calmegin 5′ flanking sequence. The 330 bp 5′ flanking sequence of the calmegin gene characterized in the present study is very short for testis-specific transcription in comparison with those of the already reported genes. It contained only the 152 bp putative promoter sequence upstream of the transcription initiation site. Therefore, the *cis*-acting element regulating testis-specific transcription must exist within this sequence. Some testis-specific genes were controlled by a silencer-like element capable of inhibiting the transcription in other somatic tissues [25]. However, the calmegin promoter sequence seems to be too short to include such a negative element, indicating that the calmegin gene is positively regulated by unidentified *cis*-acting elements localizing in the 5′ flanking sequence. In this context, this promoter sequence could be useful for characterizing the mechanism of testis-specific gene expression.

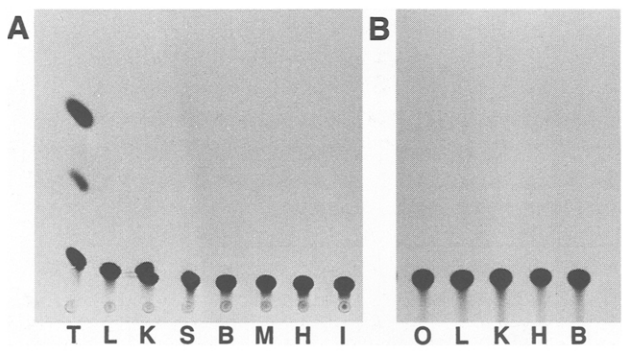

Fig. 3. CAT activities in various tissues of the transgenic mice. Twenty mg of protein obtained from the transgenic mouse TGcal01 was assayed as described in section 2. (A) Tissues from the male transgenic mouse; (B) tissues from the female transgenic mouse. T, testis; L, liver; K, kidney; S, spleen; B, brain; M, muscle; H, heart; I, intestine and O, ovary. CAT activity was found only in the testis of the transgenic mouse.

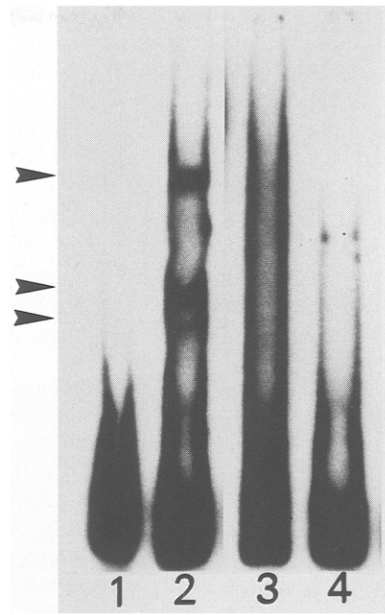

Fig. 4. Gel shift assay of calmegin promoter sequence. Gel shift assay was performed using the ^{32}P-labeled calmegin genome DNA fragment (−152 to +18) with a testis or liver nuclear extract (15 μg) as described in section 2. Lane 1 = the probe only; lane 2 = testicular nuclear extract; lane 3 = testicular nuclear extract with the 100-fold excess of unlabeled same oligonucleotide; lane 4 = liver nuclear extract. Retarded bands specific to the testicular nuclear extract were indicated by arrowheads.

In the present study, using the gel shift assay we have also demonstrated that the nuclear factors in the testis specifically bound to the calmegin promoter sequence (Fig. 4). Recently, the presense of many *cis*-acting elements and *trans*-acting factors has been demonstrated and some of the mechanisms of tissue-specific or development-specific gene expression were characterized. As for the testis, a few reports are available. Some testicular nuclear factors that bind to testis-specific gene promoters have been identified [24,26]. Although, the DNA fragment we used contained potential binding sites for SP 1 and AP2, any shift bands were not observed in our experimental condition with the liver nuclear extract. However, we were able to demonstrate specific binding of testicular extract to the DNA

fragment. Thus, the testicular nuclear factors identified in our experiment specifically bound to the calmegin promoter sequence. The molecular properties of the nuclear factors and their binding sequences are not fully characterized yet. However, we assume the factors would have some important roles in specific gene expressions during male germ cell differentiation in the testis. Characterzation of these testicular nuclear factors and their target sequences is now in progress.

Further systematic analyses of the calmegin promoter sequence and the nuclear factors identified in the present study will bring us a better understanding of the mechanism of testicular germ cell-specific gene express ion and the development.

Acknowledgments: We thank Dr. A. Tanaka for reading and correcting our manuscript.

References

[1] Watanabe, D., Yamada, Y., Nishina, Y., Tajima, Y., Koshimizu, U. and Nishimune, Y. (1992) J. Biol. Chem. 269, 7744–7749.
[2] Willison, and Ashworth, A. (1987) Trends Genet. 3, 351–355.
[3] Erickson, R.P. (1990) Trends Genet. 6, 264–269.
[4] Watanabe, D., Sawada, K., Koshimizu, U., Kagawa, T. and Nishimune, Y. (1992) Mol. Reprod. Dev. 33, 307–312.
[5] Fliegel, L., Burns, K., MacLennan, D.H., Reithmeier, R.A.F. and Michalac, M. (1989) J. Biol. Chem. 264, 21522–21528.
[6] Smith, M.J and Koch, G.L.E. (1989) EMBO J. 8, 3581–3586.
[7] Wada, I., Rindress, D., Cameron, P.H., Ou, W.J., Doherty, J., Louvard, D., Bell, A.W., Dignard, D., Thomas, D.Y. and Bergeron, J.J. (1991) J. Biol. Chem. 266, 19599–19610.
[8] Yamada, K., Noguchi, T., Miyazaki, J., Matsuda, T., Takenaka, M., Yamamura, K. and Tanaka, T. (1990) Biochem. Biophys. Res. Commun. 171, 243–249.
[9] Jackson, M.R., Cohen, D.M., Peterson, P.A. and Williams, D.B. (1994) Science 263, 384–387
[10] Rajagopalan, S., Xu, Y. and Brenner, M.B. (1994) Science 263, 387–390.
[11] Burns, K., Duggan, B., Atkison, E.A., Famulski, K.S., Nemer, M., Bleackley, R.C. and Michalak, M. (1994) Nature 367, 476–480.
[12] Dedhar, S., Rennie, P.S., Shago, M., Hagesteijn, C.L., Yang, H., Filmus, J., Hawlay, R.G., Bruchovsky, N., Cheng, H., Matusik, R.J. and Giguere, V. (1994) Nature 367, 480–483.
[13] Rappold, G.A., Stubbs, L., Labeit, S., Crkvenjakov, R.B. and Lehrach, H. (1987) EMBO J. 6, 1975–1980.
[14] Kilpatrick, D.L., Zinn, S.A., Fitzgerald, M., Higuchi, H., Sabol, S.L. and Meyerhardt, J. (1990) Mol. Cell. Biol. 10, 3717–3726.
[15] Mazarakis, N.D., Nelki, D., Lyon, M.F., Ruddy, S., Evans, E.P., Freemont, P. and Dudley, K. (1991) Development 111, 561–571.
[16] Bird, A.P. (1986) Nature 321, 209–213.
[17] Antequera, F., Boyes, J. and Bird, A. (1990) Cell 62, 503–514.
[18] Ariel, M., McCarrey, J. and Cedar, H. (1991) Proc. Natl. Acad. Sci. USA 88, 2317–2321.
[19] Teruya, J.H., Kutsunai, S.Y., Spear, D.H., Edwards, P.A. and Clarke, P.A. (1990) Mol. Cell. Biol. 10, 2315–2326
[20] Schibler, U and Sierra, F. (1987) Annu. Rev. Genet. 21, 237–257
[21] Robinson, M.O., McCarrey, J.R. and Simon, M.I. (1989) Proc. Natl. Acad. Sci. USA 86, 8437–8441.
[22] Zinn, S.A., Ebert, K.M., Mehta, N.D., Joshi, J. and Kilpatrick, D.L. (1991) J. Biol. Chem. 266, 23850–238505.
[23] Langford, K.G., Shai, S., Howard, T.E., Kovac, M.J., Overbeek, P.A. and Bernstein, K.E. (1991) J. Biol. Chem. 266, 15559–15562.
[24] Ewulonu, U.K., Buratynski, T.J. and Schimenti, J.C. (1993) Development 117, 89–95.
[25] Goto, M., Koji, T., Mizuno, K., Tamaru, M., Koikeda, S., Nakane, P.K., Mori, N., Masamune, Y. and Nakanishi, Y. (1990) Exp. Cell. Res. 186, 273–278.
[26] Tamaru, M., Nagao, Y., Taira, M., Tatibana, M., Masamune, Y. and Nakanishi, Y. (1990) Biochim. Biophys. Acta 1049, 331–338.
[27] Yamamura, K., Kikutani, H., Takahashi, N., Taga, N., Akita, S., Kawai, K., Fukuchi, K., Kumahara, K., Honjo, T. and Kishimoto, T. (1984) J. Biochem. 96, 357–363.
[28] Crabb, D.W. and Dixon, J.E. (1987) Anal. Biochem. 163, 88–92.
[29] Gormann, C.M., Moffat, L.F. and Howard, B.H. (1982) Mol. Cell. Biol. 2, 1044–1051.
[30] Bina-Stein, M., Thoren, M., Salzman, N. and Thompson, J.A. (1979) Proc. Natl. Acad. Sci. USA 76, 731–735.
[31] Schreiber, E., Matthias, P., Muller, M.M. and Schaffner, W. (1989) Nucleic Acids Res. 17, 6419.