Volume 319, number 3, 244–248 FEBS 12258 March 1993

Intron and intronless transcription of the chicken polyubiquitin gene UbII

Jovita Mezquita, Blanca López-Ibor, Montserrat Pau and Cristóbal Mezquita

Molecular Genetics Research Group, Faculty of Medicine, University of Barcelona, Casanova 143, 08036 Barcelona, Spain

Received 26 January 1993; revised version received 8 February 1993

We have previously reported that the chicken polyubiquitin gene UbII is preferentially expressed during spermatogenesis and we show here that UbII is the predominant polyubiquitin gene expressed in early embryogenesis. Two main initiation sites were detected. Transcription from the initiation site used in early embryos results in the presence of an intron in the 5′-untranslated region of the transcripts as has been reported for other polyubiquitin messages. In mature testis, however, the use of a different initiation site, located within the intron, produces intronless transcripts. Distinct promoter sequences, present in each initiation site, may regulate the differential expression observed in this gene.

Ubiquitin; Initiation site; Spermatogenesis; Embryogenesis; tRNA; Ubiquitin-fusion gene

1. INTRODUCTION

Ubiquitin, a highly conserved, highly stable, 76-amino acid residue globular protein found in all eucaryotic cells, has a fundamental role in mediating intracellular selective protein degradation [1–3]. In yeast, ubiquitin conjugation is essential for cell cycle control, DNA repair, resistance to stress and sporulation [4,5]. Ubiquitin conjugation could be also essential for spermatogenesis [6,7]. Formation of ubiquitin conjugates may be crucial for proliferation of spermatogonia. In addition, ubiquitination may be needed in subsequent stages of spermatogenesis. We have postulated that the increase in the levels of histone ubiquitin conjugates observed during rooster spermiogenesis may participate in the mechanism of nucleohistone replacement by protamine [8,9] and in the final stages of condensation of chromatin after histones have been replaced by protamine [10]. We have shown previously that the UbII gene is preferentially expressed during chicken spermiogenesis [11]. In the present study we attempted to elucidate if the preferential expression of this gene in certain tissues could be dependent on special characteristics of its promoter region.

2. MATERIALS AND METHODS

2.1. *Separation of chicken testis cells by centrifugal elutriation*

Hubbard White Mountain roosters (8–50 weeks old) were used in the experiments. Chicken testis cells were prepared and separated by centrifugal elutriation as previously described [12].

2.2. *Isolation of RNA and hybridization analysis*

RNA was isolated by the guanidine isothiocyanate method [13] and purified through a CsCl gradient. For hybridization analysis two specific probes were prepared, one containing the *Xmn*I-*Sac*I 3′ region of the UbII gene and the other (BanII) containing the tail sequence and 3′ region of the Ub-t52 gene [11,14]. An additional *Bgl*II probe was obtained from the ubiquitin coding region. Due to the high similarity of the ubiquitin coding sequences, the *Bgl*II probe strongly hybridizes to the UbII and UbI transcripts and much less to the Ub-t52 transcript. To detect selectively the UbII and Ub-t52 transcripts the filters were stripped and hybridized sequentially with the *Xmn*I-*Sac*I (UbII) and BanII (Ub-t52) probes. The probes were labelled using the multiprime system from Amersham and [α-^{32}P]dCTP. Total RNA was electrophoresed in 1.2% agarose–2 M formaldehyde gels. RNA was transferred to nylon filters (Hybond N from Amersham). Blots were baked, prehybridized and hybridized as recommended by manufacturers (Amersham) using stringent conditions.

2.3. *cDNA cloning and DNA sequencing*

Positive clones for the ubiquitin gene were isolated from a chicken EMBL3 genomic library [11]. A 4 kb *Sac*I and 1 kb *Sac*I-*Hin*dIII subclones were constructed in the phagemid Bluescript SK+. They comprise the whole gene, extending 268 nucleotides at the 3′ and 4 kb at the 5′. Additional subclones were obtained after cutting with *Pst*I, *Sac*II, *Sma*I, *Bgl*II, *Apa*I and *Not*I.

Sequencing was performed from double and single stranded DNA and from both strands using the dideoxy method [15], T7 polymerase (Pharmacia or USB) and the primers T3, reverse primer, SK, KS, T7 or the universal primer from Stratagene.

2.4 *Sequence comparison*

Computation analysis was performed at the NCBI using the BLAST network service.

2.5. *Reverse transcriptase-PCR*

For the analysis of the 5′ end of the polyubiquitin UbII mRNAs, total RNA was purified through CsCl and 5 μg were used as a substrate for the synthesis of first-strand cDNA. Synthesis was performed by random hexanucleotides and murine reverse transcriptase as recommended by Pharmacia. Excess primer and dNTPs were removed using a Centricon 100 spin filter and the sample was further concentrated by speed vacuum centrifugation [16]. Oligo(dA) tailing was accomplished as recommended by Boehringer. The reaction was terminated by heating at 70°C and diluting to 300 μl. Aliquots of 10 μl were used for PCR amplification. PCR amplification was done using

Correspondence address: C. Mezquita, Facultad de Medicina, Fisiología, Casanova 143, 08036 Barcelona, Spain.

 Published by Elsevier Science Publishers B.V

Fig. 1. Nucleotide sequence of the 5′-flanking region of the chicken polyubiquitin UbII gene. The intron within the 5′ non-coding region is typed in lower case. The transcription start sites are indicated by arrows. The putative TATA box and cAMP responsive element (CRE) are boxed. Major transcription start sites are indicated by arrows. Minor heterogeneous start sites located upstream of the TATA box are indicated by dots. Sequences similar to the adenovirus E2 promoter, to the glucocorticoid responsive element, and several potential Sp1 binding sites are underlined. A sequence similar to the interferon regulatory factor promoter is overlined. The EMBL Data Library accession numbers for the sequence are X58195 and Z14958.

*Not*I d(T)$_{18}$ bifunctional primer from Pharmacia and a specific primer, synthesized by Isogen Bioscience, complementary to the first ubiquitin unit of the UbII gene (5′-AGGTGATGGTCTTGCCAGTG-3′). The nucleotide at the 3′ end is not complementary to any of the other ubiquitin units of the UbII and UbI genes, giving the primer the necessary specificity to prime only the 5′ region of the UbII gene. The extension product of the specific primer generates a *Bgl*II site located 14 nucleotides from the primer that was used for subsequent cloning of the amplified product. PCR amplification, 35 cycles, was carried out using 37 pmol of specific primer in a total volume of 100 μl. Cycles were performed at 95°C, 1 min; 55°C, 1 min and 72°C, 3 min. The cDNA's were cut with restriction enzymes *Bgl*II and *Not*I and inserted in *Bam*HI and *Not*I sites of Bluescript SK+.

3. RESULTS AND DISCUSSION

3.1. *Chicken polyubiquitin gene UbII contains two main initiation sites*

A recombinant clone containing the promoter region of the chicken polyubiquitin gene UbII has been isolated from a genomic library derived from chicken testis. A 4 kb fragment of the 5′-flanking region of the chicken polyubiquitin UbII gene has been sequenced, and 2.4 kb of this sequence are shown in Fig. 1. Using a reverse transcriptase-PCR method to determine the

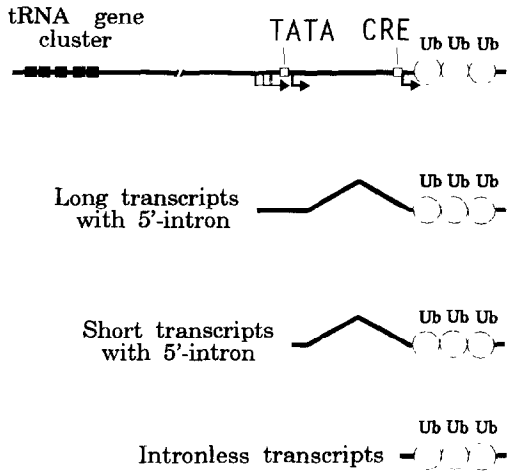

Fig. 2. Schematic representation of UbII initiation sites and transcripts. Specific initiation sites are used in different tissues giving rise to intron and intronless transcripts.

initiation site of transcription, two main initiation sites have been detected. In early chick embryos and adult kidney the initiation site of fourteen clones sequenced is located 514 bp or 516 bp upstream from the start of translation. A 448 bp intron within the 5′ untranslated leader region extends from the intron donor site, located 60 bp downstream of the transcription start site, to the acceptor site 6 bp upstream of the first codon ATG. In mature chicken testis, however, most of the transcripts (12 of 18 clones sequenced) initiate 80 bp upstream from the start of translation and contain no intron. Few transcripts (4 of 18 clones sequenced) start upstream of the TATA box (Fig. 1, marked with dots) and fewer (2 of 18 clones) at the same position found in chick embryo and kidney (Fig. 2).

The intronless transcription of the chicken polyubiquitin UbII gene in mature testis must permit the synthesis of ubiquitin, regardless of heat shock and other stresses that inhibit RNA splicing. Ubiquitin plays an essential role in the stress response system of eukaryotic cells [17–19]. Brief severe heat shock blocks the processing of intervening sequences of most eukaryotic genes and produces developmental abnormalities [20]. Heat shock proteins confer protection to the cell and their synthesis is not interrupted by heat shock because most of the genes encoding these proteins are free of introns and transcripts do not require splicing [21]. Several polyubiquitin genes do not follow this rule; an intron is present within their 5′ non-translated region [21,22] as we report here for the chicken polyubiquitin UbII transcripts of early embryos and adult kidney. In mature testis, however, the use of a different initiation site, produces intronless ubiquitin transcripts, that must allow the synthesis of ubiquitin in male germ cells in spite of heat shock and other stresses. This is particularly important due to the well known sensitivity of male germ cells to increased temperature. In most mammals, germ cell development is affected adversely if the temperature of the testis is raised even to body temperature [23]. Avian spermatogenesis takes place, however, at the high internal temperatures of birds (38.5–43.8°C) [24]. The molecular basis for thermosensitivity or thermotolerance are unknown at present. Selective intronless transcription of the polyubiquitin UbII gene during chicken spermatogenesis may be required for optimal differentiation of avian male germ cells at the elevated internal temperatures of birds.

The existence of different initiation sites of UbII in mature chicken testis could be explained by the heterogeneity of the cell population of this organ. Mature testis contains somatic cells and male germ cells at successive stages of spermatogenesis. During spermatogenesis the expression of certain genes is altered either quantitatively or qualitatively [25]. The use of testis specific promoters of the polyubiquitin gene UbII produces intronless transcripts and transcripts containing the TATA box and other putative promoter sequences incorporated in the 5′-non-translated region of the message. We have reported previously the striking presence of a TATA box and other promoter sequences in the chicken carbonic anhydrase CAII testis message [26]. The same observation has been reported for the human α-tubulin gene, H-44 [27]. The importance of the TATA box in correctly positioning the RNA polymerase for initiation of transcription has been well established. Why a different mechanism is used during spermatogenesis is not known at present.

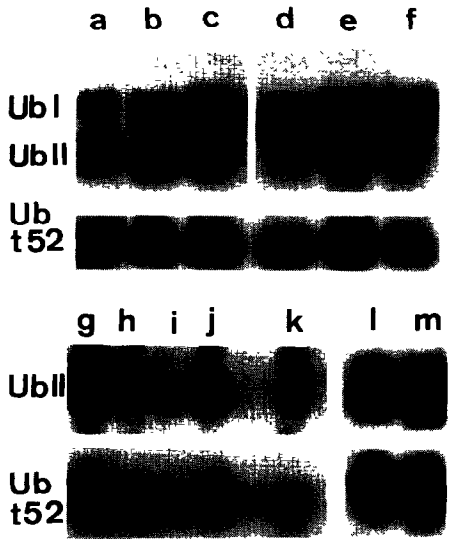

Fig. 3. Northern blotting analysis of the polyubiquitin UbII and Ubt52 expression during chicken embryogenesis, in different chicken tissues, and in male germ cells at successive stages of spermatogenesis. Chick embryo: (a) four days; (b) five days; (c) seven days; (d) muscle (12 days); (e) fibroblasts; (f) brain (15 days). Adult chicken tissues: (g) testis; (h) kidney; (i) heart; (j) brain; (k) liver. Stages of spermatogenesis: (l) meiotic and premeiotic cells. (m) round spermatids.

3.2. *Preferential expression of the polyubiquitin gene UbII and the gene coding for a ubiquitin hybrid protein Ub-t52 in early chick embryos and during spermatogenesis*

Ubiquitin gene Ub II is the predominant polyubiquitin gene expressed in early chick embryos (Fig. 3a). Later in embryogenesis similar levels of polyubiquitin UbI and UbII messages were detected (Fig. 3b,c). In late embryogenesis the expression of UbII sharply decreases (Fig. 3d–f). The expression of the polyubiquitin gene UbII during embryogenesis is parallel to the expression of a ubiquitin gene coding for a hybrid protein of ubiquitin and a ribosomal protein of 52 amino acids (Ub-t52) (Fig. 3a–f). The expression of UbII and Ub-t52 is also parallel in chicken adult tissues (Fig. 3g–k). Both mRNAs are also preferentially expressed during spermatogenesis. UbII and Ub-t52 mRNA's are abundant in a fraction containing meiotic and premeiotic cells and also in a fraction of non-replicating spermatids (Fig. 3l,m).

In yeast, two genes encoding hybrid proteins of ubiquitin and ribosomal proteins are expressed in exponentially growing cells, when ribosome synthesis is active, while their expression is repressed in stationary phase cells [18]. In addition to cell cycle control, ubiquitin may participate in other cell functions in early embryogenesis. Ubiquitin–protein conjugates are selectively present in early chick embryos in cells undergoing major cytomorfological reorganisation: cells of the lens, notochord and myotome [28].

Chicken spermatids do not divide as early embryo cells but grow during spermiogenesis with a concomitant development of the rough endoplasmic reticulum [29]. At the end of the process the endoplasmic reticulum cisternae break down into vesicles that dissociate from ribosomes and swell to give a vacuolated appearance to the cytoplasm of late spermatids [29]. In a similar way, during embryogenesis, ubiquitin conjugates are present in the notochord during the inductive events preceding vacuolisation [28].

3.3. *5′-Flanking region of the polyubiquitin UbII*

Preferential expression of polyubiquitin gene UbII during early embryogenesis and spermatogenesis may depend on special characteristics of the promoter elements present in the 5′-flanking regions of the start sites. The 5′-flanking region of the initiation site detected in early chick embryos contains a TATA box located 31 bp upstream from the initiation site and several potential Sp1 binding sites. The sequence GCGGGAAA, located 57 bp upstream of the TATA box, is identical to a sequence present within the Myc P2 promoter [29] and shows only a single G/C difference with promoters of the E2 gene of adenovirus. These sequences interact with the cellular transcription factor E2F [30]. Genes such as c-*myc*, N-*myc* and c-*myb* contain similar sites in their promoters [31]. These sites appear to be important for transcriptional activation of genes involved in cellular proliferation.

A sequence of 21 nucleotides, present 28 bp upstream of the TATA box, differs only at two positions from a sequence found in the promoter region of the virus inducible mouse interferon regulatory factor-1 (IRF-1) [32]. This nuclear factor induces the expression of genes as interferons IFN-β, IFN-α, MHC class I and other genes involved in immune responses [32]. Concomitant expression of the virus inducible IRF-1 and ubiquitin could participate in the processing of viral antigens through the postulated ubiquitin-dependent 26 S protease pathway [33]. The most frequent similarities found, comparing the 5′-flanking region of the UbII gene and the sequences of the EMBL/GenBank data libraries, correspond to DNA sequences of the genome of the herpes simplex virus and related viruses. It has been shown that lytic infection with herpes simplex results in specific transcriptional induction of a member of the polyubiquitin human gene family, mediated by the viral protein ICP4 [34].

The hexanucleotide TGTTCT of the glucocorticoid responsive element is present in a region of 27 nucleotides, located 461 bp upstream of the TATA box, 90% similar to the promoter sequence of the human liver-type arginase gene, a gene induced by glucocorticoids [35].

The 5′ flanking region of the testis initiation site contains a cAMP responsive element (CRE) and five potential Sp1 binding sites. The CRE is immediately followed by the dinucleotide CA. This feature has been observed in other genes expressed during spermatogenesis [36,37]. A sequence that differs only at the last position from the sequence GCGGGAAA of the Myc P2 promoter [30] is present 59 bp upstream of the CRE.

An interesting feature of the chicken polyubiquitin gene UbII is the presence of sequences coding for a tRNA gene cluster 1.9 kb upstream of the TATA box [38]. The presence of this cluster containing RNA polymerase III promoter sequences may exert a positive enhancer effect on the downstream adjacent RNA polymerase II promoters [39]. Moreover, the proximity of the tRNA gene cluster and the polyubiquitin UbII gene suggests the possibility of a coordinate expression of these genes. The tRNAs are key components of the protein synthesis machinery. Both in dividing (early embryogenesis) and non-dividing cells (spermatids) the expression of the polyubiquitin gene UbII is concomitant with very active protein synthesis. The expression of two ubiquitin genes in cells active in protein synthesis and the essential role of ubiquitin in selective proteolysis, suggest the existence of a link between protein synthesis and selective protein degradation. The putative coordinate expression of the tRNA genes and polyubiquitin gene UbII may serve a more specific function: the ubiquitin-dependent degradation of some proteins was found to require tRNA [40]. Recognition of target pro-

teins as substrates for ubiquitination depends on the identity of their N-terminal residues (N-end rule) [3]. Aminoacyl tRNA-protein transferases, which conjugate specific amino acids to the N-termini of acceptor proteins, may provide a mechanism for selective degradation or stabilization of proteins during cell differentiation [3]. Co-regulation of gene expression of tRNAs and ubiquitin could be important to assure the cell a proper ratio of tRNAs to ubiquitin for polyubiquitination of those target proteins which require a tRNA mediated addition of new N-terminal amino acids. The 26 S proteasome involved in selective degradation of ubiquitin conjugates contains tRNAs [41].

Acknowledgements: This work was supported by DGICYT Grant PM89-0051.

REFERENCES

[1] Hershko, A. (1991) Trends Biochem. Sci. 16, 265–268.
[2] Rechsteiner, M. (1991) Cell 66, 615–618.
[3] Varshavsky, A. (1992) Cell 69, 725–735.
[4] Finley, D., Ozkaynak, E. and Varshavsky, A. (1987) Cell 48, 1035–1046.
[5] Sung, P., Prakash, S. and Prakash, L. (1990) Proc. Natl. Acad. Sci. USA 87, 2695–2699.
[6] Mitchell, M.J., Woods, D.R., Tucker, P.K., Opp, J.S. and Bishop, C.E. (1991) Nature 354, 483–486.
[7] Kay, G.F., Ashworth, A., Penny, G.D., Dunlop, M., Swift, S., Brockdorff, N. and Rastan, S. (1991) Nature 354, 486–489.
[8] Agell, N., Chiva, M. and Mezquita, C. (1983) FEBS Lett. 155, 209–212.
[9] Agell, N. and Mezquita, C. (1988) Biochem. J. 250, 883–889.
[10] Esteban, D., Garcia-Valero, J., Mezquita, J. and Mezquita, C. (1989) J. Cell Biol. 109, 69a.
[11] Mezquita, J. and Mezquita, C. (1991) FEBS Lett. 279, 69–72.
[12] Roca, J. and Mezquita, C. (1989) EMBO J. 8, 1855–1860.
[13] Chirgwin, J.M., Przybyla, A.E., Macdonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[14] Mezquita, J., Pau, M. and Mezquita, C. (1988) Nucleic Acids Res. 16, 11838.
[15] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[16] Frohman, M.A. (1990) in: PCR Protocols. A Guide to Methods and Applications (Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. eds.) pp. 28–38, Academic Press, London.
[17] Bond, U. and Schlesinger, M.J. (1986) Mol. Cell. Biol. 6, 4602–4610.
[18] Özkaynak, E., Finley, D., Solomon, M.J. and Varshavsky, A. (1987) EMBO J. 6, 1429–1439.
[19] Fornace, A.J., Alamo, I., Hollander, M.C. and Lamoreaux, E. (1989) Nucleic Acids Res. 17, 1215–1230.
[20] Yost, H.J. and Lindquist, S. (1986) Cell 45, 185–193.
[21] Lee, H.S., Simon, J.A. and Lis, J.T. (1988) Mol. Cell Biol. 8, 4727–4735.
[22] Baker, R.T. and Board, P.G. (1987) Nucleic Acids Res. 15, 443–463.
[23] Chowdbury, A.K. and Steinberger, E. (1964) Am. J. Anat , 115, 509–524.
[24] Prinzinger, R., Prebmar, A. and Schleucher, E. (1991) Comp. Biochem. Physiol. 99A, 499–506.
[25] Willison, K. and Ashworth, A. (1987) Trends Genet. 3, 351–355.
[26] Mezquita, J., Pau, M., and Mezquita, C. (1991) J. Cell Biol. 115, 152a.
[27] Dobner, P.R., Kislauskis, E., Wentworth, B.M. and Villa-Komaroff, L. (1987) Nucleic Acids Res. 15, 199–218.
[28] Scotting, P., McDermott, H. and Mayer, R.J. (1991) FEBS Lett. 285, 75–79.
[29] Xia, L., Clermont, Y., Lalli, M. and Buckland, R.B.(1986) Am. J. Anat. 177, 301 312.
[30] Hiebert, S.W., Lipp, M. and Nevins, J.R. (1989) Proc. Natl. Acad. Sci. USA 86, 35–94.
[31] Mudryj, M., Hiebert, S.W. and Nevins, J.R. (1990) EMBO J. 9, 2179–2184.
[32] Miyamoto, M., Fujita, T., Kimura, Y., Maruyama, M., Harada, H. Sudo, Y., Miyata, T. and Taniguchi, T. (1988) Cell 54, 903–913
[33] Driscoll, J. and Finley, D. (1992) Cell 68, 823–825.
[34] Latchman, D.S., Estridge, J.K. and Kemp, L.M. (1987) Nucleic Acids Res. 15, 7283–7293.
[35] Takiguchi, M., Haraguchi, Y. and Mori, M. (1988) Nucleic Acids Res. 16, 8789–8802.
[36] Oliva, R. and Dixon, G.H. (1990) J. Mol. Evol. 30, 333–346.
[37] Tamura, T., Yasutaka, M., Mikoshiba, K. and Maramatsu, M. (1992) J. Biol. Chem. 267, 4327–4332.
[38] Mezquita, J. and Mezquita, C. (1992) Nucleic Acids Res. 20, 5477.
[39] Oliviero, S. and Monaci, P. (1988) Nucleic Acids Res. 16, 1285–1293.
[40] Ciechanover, A., Wolin, S.L., Steitz, J.A. and Lodish, H.F. (1985) Proc. Natl. Acad. Sci. USA 82, 1341–1345.
[41] Nothwang, H.G., Coux, O., Keith, G. Silva-Pereira, I. and Scherrer, K. (1992) Nucleic Acids Res. 20, 1959–1965.