Gene Expression, Vol. 4, pp. 19–28 © 1994
Printed in the USA.

1052-2166/94 $3.00 + .00

Structure and Expression of a Polyubiquitin Gene From the Crustacean *Artemia*

ENRIQUE FRANCO,[1,2] M. MARGARITA BEHRENS,[1] MARGARITA DÍAZ-GUERRA,[3] AND JAIME RENART[4]

Instituto de Investigaciones Biomédicas del CSIC and Departamento de Bioquímica de la Universidad Autónoma de Madrid, 28029-Madrid, Spain

We have characterized two polyubiquitin genes from the crustacean *Artemia franciscana*. One of them, Ubi1, has nine ubiquitin units and an intron of a minimum size of 3.5 kb that ends 7 bp before the initiator ATG. The 5′ end of the transcript from this gene has been identified by anchored PCR. The existence of the other gene (Ubi2) was inferred from several cDNA clones that differ from Ubi1 in the C-terminal extension and in the 3′ untranslated region as well as in the nucleotide sequence of the coding region. We find two transcripts of ubiquitin genes, of 2.7 and 3.3 kb. Hybridization of RNA blots with an oligonucleotide specific for Ubi2 gene demonstrates that this gene codes for the 3.3 kb transcript. Ubiquitin messenger RNAs are present in the dormant embryos and their steady-state levels are maximum at 8 h after resumption of development, declining thereafter. The Ubi2 gene transcripts are less abundant but its proportion in relation to the other transcript does not vary with development.

UBIQUITIN is a 76-residue protein found in all eukaryotic organisms studied to date, and it was recently found in at least one archaebacterium, *Thermoplasma acidophilum* (Wolf et al., 1993). It has been reported to be covalently bound to other proteins, as diverse as histones H2A and H2B (Busch and Goldknopf, 1981; Wu et al., 1981), the lymphocyte homing receptor (Siegelman et al., 1986), and the platelet-derived growth factor (Yarden et al., 1986) [see Rechsteiner (1987, 1988) for more comprehensive reviews]. Ubiquitin is involved in the in vivo degradation of cellular proteins, acting as a tag for an ATP-dependent proteolytic system (Hershko et al., 1980, 1984). This function is related to the stress response system (Finley et al., 1984) and to the progression in the cell cycle (Glotzer et al., 1991). Ubiquitin is also found in the neurofibrillary tangles typical of Alzheimer syndrome (Perry et al., 1987; Mori et al., 1987).

The conservation of ubiquitin throughout the philogenetic tree is remarkable. Its amino acid sequence is identical from arthropods to mammals, and is 96% similar (three changes out of the 76 residues) with that of yeast, *Dictyostelium discoideum*, protozoa, and plants.

Gene organization is also conserved. There are two types of ubiquitin genes. In the first, several ubiquitin coding units are ligated head-to-tail, without any interruption between them, forming a polyubiquitin gene. These genes do not normally have introns in the coding region and end with a small extension of several amino acids between the last ubiquitin unit and the stop codon. They also contain in their 5′ region a heat-shock promoter element (Bond and Schlesinger, 1986). The second type of genes code for a single ubiquitin unit followed by a C-terminal extension protein that is conserved in all species studied to date (Finley et al., 1989; Cabrera y Poch et al., 1990; Redman and Rechsteiner, 1989) and corresponds to a ribosomal protein. It has been suggested that the ubiquitin moiety could act as a chaperon for the correct assembly of the ribosomal subunits (Finley et al., 1989).

The brine shrimp *Artemia* is a crustacean for which gastrula can enter cryptobiosis when environmental conditions are adverse. After resumption of development, gastrulae develop into nauplii larvae without cell division and DNA synthesis. This pro-

Received September 29, 1993; revision accepted March 23, 1994.
[1]E.F. and M.M.B. are equal first authors.
[2]Present address: Departamento de Biología I, Universidad del Atlántico, Barranquilla, Colombia.
[3]Present address: Centro Nacional de Biotecnología, Cantoblanco, 28049-Madrid, Spain.
[4]Address correspondence to Jaime Renart, Instituto de Investigaciones Biomédicas del CSIC, Arturo Duperier 4, 28029-Madrid, Spain.

TABLE 1

DESCRIPTION OF THE OLIGONUCLEOTIDES USED

Oligo	Derived From	Positions	Sequence (5′ → 3′)
C-1	Ubi1	1351–1367 (as)	AATCTGCATAATTCAAC
G-16	Ubi1	1370–1391 (as)	CTTTCCTGTGAGCGTTTTCACG
gUbq	Ubi1	3510–3532 (as)	TGTCCGGAGCGAAAAAATCAAAAA
T-75	Ubi1	3425–3448 (as)	TAAGATCTTTATGAGTAAACTTCT
cUbq	pcUbq	705–728 (as)	TTCTGGTTGATACAATTGACGTTA
Ubi2-3′	pcUbq	856–874 (as)	AGATACGTTCGTTTTTTTT
C-155	pcUbq	649–668 (s)	CAAGACCTTTATGACAGAACTAAC
G-119	Anch-PCR	31–56 (s)	TTCTGTCTTCGGAGCGCATTGTTTGG
C-19	Kasturi and Bona (1991)		ACGAGCTCGGATCCGAATTCCCCCCCCCCCC
C-20	Kasturi and Bona (1991)		ACGAGCTCGGATCCGAATTCCC
GAGA	Strategene		$(GA)_{10}ACTAGTCTCGAGT_{18}$

(as): antisense direction; (s): sense direction.

cess is carried out at the expense of yolk-stored products. [See Browne et al. (1991) for recent reviews on *Artemia* biology.]

We present in this article our results concerning the structure of the polyubiquitin genes in *Artemia* as well as its expression during early development.

MATERIALS AND METHODS

Biological Material

Artemia cysts (dormant gastrulae) were from San Francisco Bay Brand. They were cultured at 30°C as described by Osuna and Sebastián (1980).

Screening of cDNAs Libraries

We used three different libraries made in our laboratory. Plasmid pcUbq was isolated from a library made in λgt11 from mRNA obtained from 20-h-old nauplii (Palmero et al., 1988).

pfu 5×10^4 (or 5×10^4 bacterial colonies in the case of the plasmid library) were grown and transferred in duplicate to nitrocellulose membranes. The membranes were hybridized overnight at 65°C with 10^6cpm/ml of probe [clone S9, kindly provided by Dr. M. Izquierdo (Izquierdo et al., 1984), that contains two ubiquitin units from *Drosophila melanogaster*], labeled by the random primer method (Feinberg and Vogelstein, 1983), in 50–100 μl/cm^2 of 4 × SSC, 5 mM EDTA, 20 mM Na phosphate buffer, pH 7.0, 0.1% SDS, 10 × Denhardt solution, and 200 μg/ml denatured and sonicated salmon sperm DNA. Filters were washed in the first two rounds at low stringency (2 × SSC, 0.1%, SDS, at 42°C) and the third round at high stringency (0.1 × SSC, 0.1% SDS, at 65°C).

Screening of Genomic Libraries

We used a genomic library made in λEMBL-3 (Frischauf et al., 1983) in our laboratory (Díaz-Guerra, 1987). Two genome equivalents (7×10^5 pfu) were grown in the appropriate host cell, plated, and transferred to nitrocellulose membranes as described (Sambrook et al., 1989).

Hybridization was carried out as described for the cDNA libraries using the pcUbq clone as probe, washing at high stringency conditions.

DNA Sequencing

We used the method of Chen and Seeburg (1985) for sequencing double-stranded DNA by the dideoxy technique. In some cases, the 373A DNA sequencer from Applied Biosystems was used, with fluorescent terminators as recommended by the suppliers. When needed, nested deletions with the Exo/Mung system of Stratagene were produced as suggested by the suppliers.

RNA Isolation and Analysis

RNA was isolated as described by Escalante and Sastre (1993) from *Artemia* embryos at the specified times after resumption of development. Electrophoresis was done in 1.5% agarose-2.2 M formaldehyde gels and transferred to nylon membranes (Thomas, 1980). Hybridization was as mentioned above.

Poly(A)$^+$ RNA was obtained from 25-h-old embryos with the Mini-QD™ kit from 5 Prime → 3 Prime, Inc.

Oligonucleotide Hybridizations

Blots were hybridized with oligonucleotides (labeled with ^{32}P with polynucleotide kinase) in 6 × SSC, 1% SDS, 1 × Denhardt solution, 0.05% sodium pyrophosphate, 100 μg/ml sonicated calf thy-

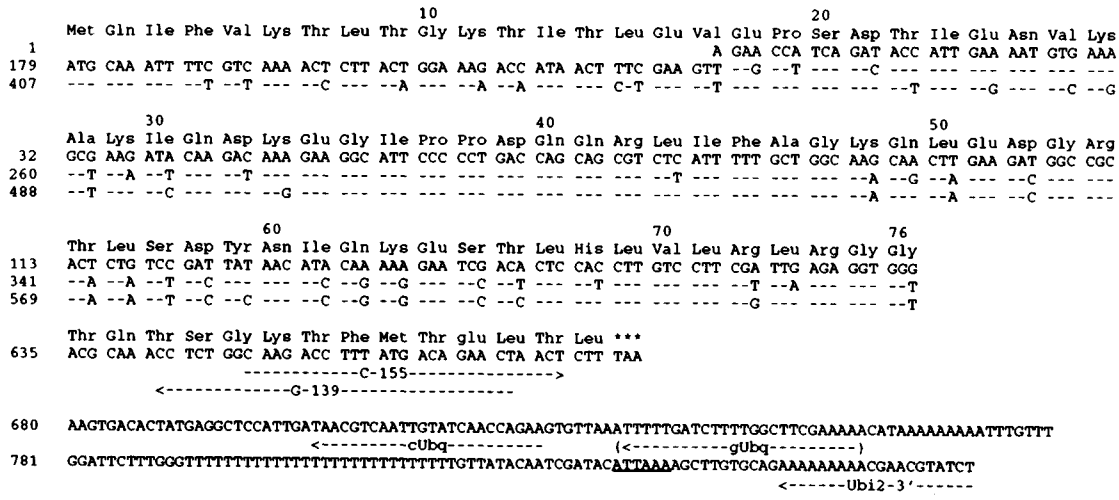


FIG. 1. Sequence and structure of the plasmid pcUbq. The sequence is aligned to show the ubiquitin units. Similar bases are denoted by a dash. The stop codon is marked with three asterisks. The oligonucleotides derived from pcUbq described in the Materials and Methods section are indicated below the sequence. The ATTAAA polyadenylation signal is underlined.

mus DNA, at 37°C for DNA blots and at 32°C for RNA blots. Washes for DNA blots were once with 6 × SSC, 0.1% SDS, 0.05% sodium pyrophosphate, 15 min at 37°C, and once with 2 × SSC, 0.1% SDS, 0.05% sodium pyrophosphate, 15 min at 42°C. RNA blots were washed with 6 × SSC, 0.1% SDS, 0.05% sodium pyrophosphate for 5 min at room temperature.

PCR Methods

We used a Gene ATAQ Controller from Pharmacia. Reverse transcription linked to PCR (RT-PCR) was done according to published protocols (Kawasaki, 1990). Anchored PCR was done according to Dorit and Ohara (1992), with minor modifications, using the primers described by Kasturi and Bona (1991).

The small DNA fragments used in the hybridizations described in Fig. 6 were labeled by the run-off method of Stürzl and Roth (1990).

Oligonucleotides used as primers are shown in Table 1 and also below the sequences of pcUbq (Fig. 1), Ubi1 (Fig. 2), and the 5′ end of Ubi1 (Fig. 3).

RESULTS

Isolation of a cDNA From Artemia *Coding for Ubiquitin*

A fragment encoding two ubiquitin units from *Drosophila melanogaster* was used to screen the library made in λgt11 described in the Materials and Methods section. We isolated a positive phage with a 0.9 kb insert that was subsequently cloned in the *Eco*RI site of pUC18 and referred to as pcUbq.

The sequence of the insert of pcUbq is shown in Fig. 1. The fragment is 872 bp in length, and contains almost three ubiquitin units (the first amino acid in pcUbq is Glu-18) plus a nonubiquitin coding extension corresponding to 14 amino acids and 234 bp of the 3′ nontranslated region.

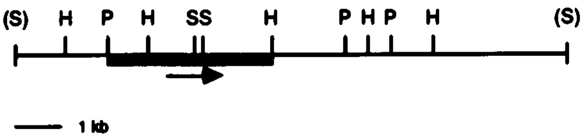
Identification of Genes Coding for Ubiquitin

pcUbq was used to screen an *Artemia* genomic library. Four positive phages were isolated. Upon digestion with *Sal*I, one of them, λgArtU5, gave two major bands, indicating that *Sal*I cuts within the coding region. Figure 2A shows the restriction map of this phage. Partial sequence data of the other three isolates indicated that they are pseudogenes, with significative amino acid changes and truncation of the coding sequences, so they were not considered further.

Figure 2B shows the sequence of the fragments (heavy line in Fig. 2A) that contain the coding sequence and 1.35 kb of 5′ upstream sequence from the initiator ATG. The gene present in λgArtU5 (that will be referred to as Ubi1 gene) is a polyubiquitin gene with nine ubiquitin units. The C-terminal extension is of 13 amino acids, very similar but not identical to that found in pcUbq.

The ubiquitin units of Ubi1 are identical in the amino acid sequence, but differ at the nucleotide level between 4% and 22%, depending on the individual units (the more homologous ones are units VII and VIII, with 4% differences, and the less ho-

A



B

```
   1                                          CTGCAGGTTCCAGCTTCTGCACAGCACTAATGACTGCCAACAATGGAATCTAGCAACTTCTTTGAAAAATTCAA
  75 ATAGAACTCGAGGTTCTCAGCCAAAAGTTCTCAAATAATAGTGGGGAAGAGTTTTTGTTCTTAATTAAATGATGGATACTAAGCTATCAAAACTGTGGAGACATTTT
 182 ACAAAAGAAAATAGCAAATTTATAAGACCAAGAACAGTTCATCCTCTTGTCCTATCTTATGAAAAAACAAGTTTTTCAACTGAAATTAAGGAGCAACATTAAAACTA
 289 AAAACGAACAGAAATTATTCTGCAAATGATGGAGCTGCCCCTTCCTCCACTCTTGCTATTGACGCAAAAATCGTAGAGCTTTTCTAAAAATATTATCTTAAAATTCA
 396 ATGGCCCTTGTGTTTGAGTTGTCTTAAAGATTCAAGACAGTCTTAAGTGTAATGAGGAAGGGTAGCGCCCCTCATACAATGATTTCTGTTTGATTAAATTTTTAATG
 503 TTGTTTCTTACTTTTGGTTGAAATACTTTTTTATTTAATTTCTTACTGTTTTTCAAATGATGCCATGAAATTACTTCCCCTCCCCATGTAGCGTTTCTCCTGGGAAA
 610 TTACCCAGAAAATTTTGAATTTACTAATGAACTTCATACATACGTGGAATCAGCATAAAAAGCCAATTGTTTTGATGAATCTATGGTTATCAACATCTGGCTTTTTG
 717 AGTTTCGGTTACTGTTGGACCGAGTTGTTCACTTACAATTCATTACCAGAACTGGTTGAGTCATTGTTATTACTTATTTTCTCCGAAAAAGTTGGTACTATAACCTA
 824 TAAATAATCTTAACAATCGCGTTTTGATGATGAAATATGTCAAATGGAATGACGTATACATCACGTCAAGGTGAAATACACCATAACTCAAATCTTGTGTCCCTTAT
 931 AGTGCTTTCACAATTATTCTTTAAGCCTTTCACACCATGAATGATCACACAATGTTTTTTTGGGCACTTGAAACATGGTATTTTCTATGTAATATTGTATAAGCTTG
1038 TCATTTACAATTTGTAAACAAAATTATGCTGAAAATGCTTTTTGGATAAAAAAAGACCTAACCATTCATTTTATCAATAAGTACATTTTATGTCTTTTGACCGTATT
1145 TGGAATGCTGTTTTTGAATGACCATAGCTGAGCTGATAAAATTAACCATTTAAAATACAAAGCGGAAAAATAGCAAAAATATAAGAATCTTTTTAATTCAAAGTATT
1252 ATTATATGCAGATAAGAAGCCTTGTTTTATAACCATCTTACTCTCTTGCTAGACAGTTTCTTATGATTAAGTCCTTGCTTAATTTTTTTTTTTTTCAGGTTGAATT
                                                                                                               <-------
                                                                                                               --------

                                             10                                      20
     Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys
1359 ATG CAG ATT TTC GTG AAA ACG CTC ACA GGA AAG ACA ATT ACG CTT GAA GTC GAG CCT TCG GAC ACA ATT GAA AAT GTC AAA
     ----C-1----  <----------G-16-------------
     -----------G-109------------------->
1587 --- --- --- --T --T --G --C T-G --- --G --- --- --- --T --- --- --A --A --A --- --T --T --- --G --- --- --G
1815 --- --A --- --- --C --- --C --- --G --- --A --C --A --T --C --- --T --- --- --C --T --T --- --- --- --T --G
2043 --- --A --- --- --C --- --C --T --T --- --- --C --A --T --C --- --T --- --- --C --T --T --- --- --- --A ---
2271 --- --A --- --- --C --- --C --T --T --- --- --C --A --T --C --- --T --- --- --A --T --T --- --- --- --A ---
2499 --- --A --- --T --C --G --C T-G --T --G --- --C --- --T --- --- --A --A --A --- --T --T --- --G --- --- --G
2727 --- --- --- --T --C --G --C T-G --T --G --- --T --- --T --- --- --A --A --- --T --T --C --- --- --- --T --G
2955 --- --- --- --T --T --G --C T-G --T --G --- --C --- --T --- --- --A --A --- --T --T --C --A --- --- --- --G
3183 --- --A --- --- --C --- --C --T --G --- --A --C --A --T --- --- --- --A --A --A --T --T --- --G --C --- --G

             30                              40                                  50
     Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg
1440 GCA AAG ATT CAA GAT AAA GAA GGC ATT CCT CCT GAT CAG CAG CGT CTG ATT TTC GCT GGC AAG CAG CTT GAA GAT GGC CGC
1668 --C --- --C --- --C --G --G --- --C --- --C --C --- --- --- T-G --- --T --- --- --A --- --A --- --- --- --A
1896 --C --A --- --- --- --- --- --- --- --C --- --C --- --- --- --C --- --T --- --- --- --A --- --- --- --T --T
2124 --T --A --- --- --- --- --- --- --- --C --- --C --- --- --- --C --- --T --- --- --- --- --- --- --C --- --T
2352 --T --A --- --- --- --- --- --T --- --C --- --C --- --- --- --T --- --T --- --T --- --- --- --- --C --- --T
2580 --T --- --C --- --C --G --- --- --C --T --C --C --- --A --C --C --- --- --- --T --A --- --A --- --- --- ---
2808 --C --- --C --- --C --- --G --- --C --- --C --- --A --- --- T-G --- --T --- --- --A --- --A --- --C --- --A
3036 --T --- --C --- --C --- --G --- --C --- --C --- --A --- --- T-G --- --T --- --T --A --- --A --- --C --- --A
3264 --- --- --C --- --C --- --- --T --C --- --C --C --- --- --- --C --- --T --- --- --A --- --A --- --C --- --T

                         60                                  70                      76
     Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
1521 ACT CTC TCC GAC TAT AAC ATA CAG AAG GAA TCG ACA CTT CAC CTG GTT CTT CGT TTG AGA GGG GGT
1749 --A T-A --T --- --- --- --C --- --- --- --C --C --C --T --T --C --- --- --- --- --T ---
1977 --- --- --T --T --- --T --C --A --A --G --- --- --C --- --T --C --- --A --- --- --T --G
2205 --- --- --T --- --- --- --- --A --- --G --- --- --G --- --T --C --- --A --- --G --T --G
2433 --- --- --T --- --- --- --- --A --- --- --A --C --- --- --T --G --- --- --- --- --C ---
2661 --A --A --T --- --- --- --C --A --- --- --C --T --- --T --T --- --- --- --- --- --T ---
2889 --A --A --T --- --- --T --T --- --- --- --C --C --C --T --T --C --- --- C-A --- --C ---
3117 --A --A --G --- --- --T --T --- --- --- --C --C --C --T --T --C --- --- --A --- --T ---
3345 --A --A --T --- --C --- --C --- --- --- --C --C --C --- --T --C --- --G --A --- --T ---

     Asn Gln Thr Ser Gly Lys Ile Phe Met Ser Lys Leu Leu ***
3411 AAT CAA ACC TCT GGT AAG ATC TTT ATG AGT AAA CTT CTT TAA
                  -------------T-75------------->

3453 AGTGGCACCGTGAGGTACCGTTGATAACGTCAATTGTATCAACCATAAGTGTTATATTTTTGATTTTTTGGCTCCGGACATAAAAAAAAATTTGTTTGGATTCTTTC
                             <-------(cUbq)----------            <----------gUbq---------
3560 AATTTTTGTTTGTTATATTTCAGATCGATAAATTAAAAGCTT
```

FIG. 2. Structure of the Ubi1 gene. (A) Restriction map of λgArtU5. The sequenced fragments are shown as a solid line. The arrow indicates the 5′ → 3′ direction of the gene. Restriction enzymes are: E, *Eco*RI; H, *Hind*III; P, *Pst* I; S, *Sal*I. (B) Sequence and structure of the Ubi1 gene. Conventions are the same as in Fig. 1. The intron acceptor site is double underlined, and the ATTAAA polyadenylation signal is underlined. Oligonucleotides described in the Materials and Methods section are also indicated.

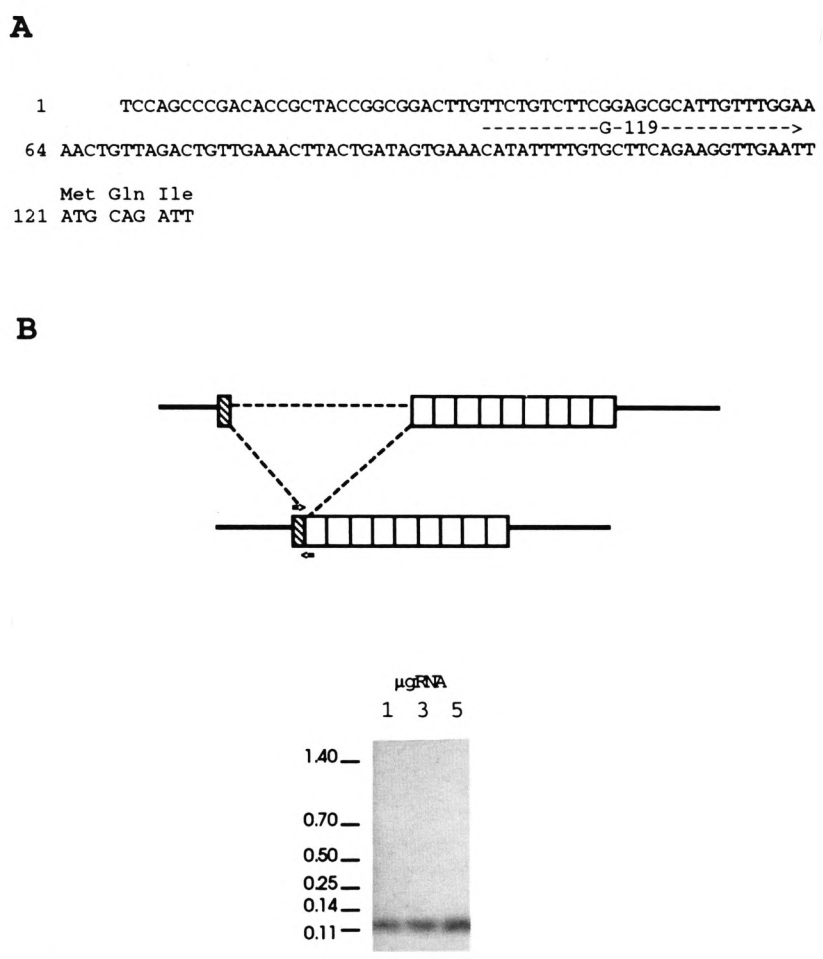


FIG. 3. Anchored PCR. (A) Sequence of the fragment obtained as described in the Materials and Methods section and in the text. The oligonucleotide derived from this sequence is indicated. (B) Diagram showing the setup of the experiment to demonstrate that the intron is processed. (C) Results of the experiment. RNA (1, 3, and 5 μg) from 8 h of development was used to perform a reverse transcription-linked PCR, as described in the text. The products of the PCR were electrophoresed in 3% agarose gel and transferred to a nitrocellulose membrane. The membrane was hybridized with oligonucleotide C-1. Molecular weight markers (in kb) pUC19 plasmid digested with *Hpa*II and *Taq*I.

mologous are units III and VIII, with 22% differences).

The same level of homology is found among the three units of pcUbq, and also between pcUbq and units VII, VIII, and IX of Ubi1. This fact strongly suggests that pcUbq is coded by a gene different from Ubi1. Two arguments favor this idea: a) the fact that cDNAs isolated with probes derived from Ubi1 or Ubi2 are identical in sequence to the probes used (data not shown), and b) the nonidentity of the C-terminal and 3′ nontranslated regions. We therefore call Ubi2 the gene coding for pcUbq.

We have sequenced 1358 bp of 5′ upstream nontranslated region, up to the *Pst*I site in search for promoter sequences. However, the most prominent signal in this region is an intron acceptor site at position 1334–1351, at −7 bp from the initiator ATG (double underlined in Fig. 2B). An intron in this position has been found in other polyubiquitin genes, as in chicken (Bond and Schlesinger, 1986), human (Baker and Bord, 1987), and *Drosophila* (Lee et al., 1988). In the *UBI4* gene from *S. cerevisiae* (Özkaynak et al., 1987), there is a consensus sequence for an acceptor site in the same region, and also consensus donor sites further upstream, although no mention is made by these authors about the real existence of the intron.

To assess the existence of the intron, we made an anchored PCR experiment. cDNA synthesis was primed by oligonucleotide G-16, at the beginning of the first ubiquitin unit (Fig. 2B). After adding oligo(dG) tails with terminal transferase, a first ampli-

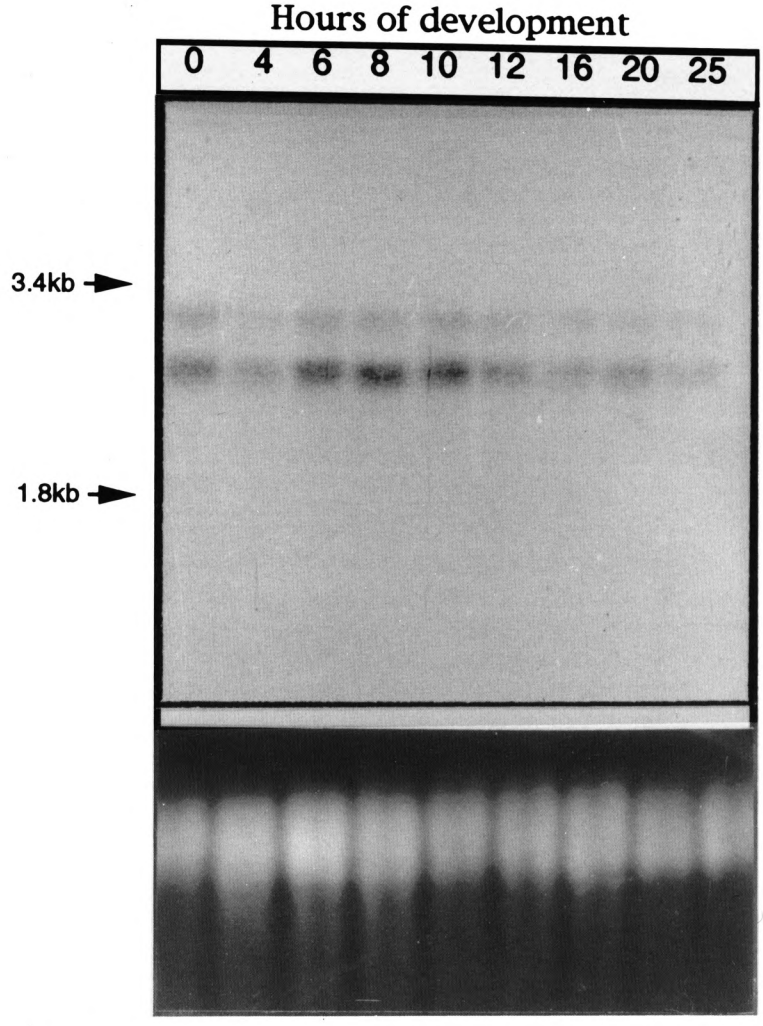


FIG. 4. Hybridization of RNAs from different times of development with an ubiquitin-coding sequence probe; 15 μg of total RNA was used per lane. Markers are ribosomal RNAs from the yeast *Saccharomyces cerevisiae*. The lower panel shows the ethidium bromide-stained gel. In this experiment, hatching of the nauplii was at 18 h of development.

fication was done using the same primer used for the cDNA synthesis (antisense) and primers complementary to oligo(dG) in the sense direction (primers C-19 and C-20 were used in a 1:9 ratio; these primers have restriction sites for *Sac*I, *Bam*HI, and *Eco*RI). The products of this first reaction were amplified again with a nested antisense primer (C-1) and the same sense primer (C-20), to increase specificity. The products of this reaction were blunted, digested with *Eco*RI, and cloned in pBlueskript (Stratagene) for sequencing. Figure 3A shows the sequence of the 130 bp fragment obtained using this technique. The fragment contains in the 3′ end the sequence of the primer used in the second amplification, which expands the initiator ATG. The sequence of this fragment is not found in the sequenced region shown in Fig. 2B, and moreover, it does not hybridize to the more upstream fragments of λgArtU5 (not shown).

To corroborate that this fragment is indeed the 5′ end of the Ubi1 gene, we performed a RT-PCR experiment using random primers for the reverse transcriptase reaction. The cDNA was then amplified with oligonucleotide G-119, derived from the anchored PCR fragment (sense, see Fig. 3) and oligonucleotide G-16, derived from the first ubiquitin unit in Ubi1, the one used for the reverse transcription reaction in the anchored PCR experiment described above. The PCR products were blotted and hybridized with oligonucleotide C-1, internal to the amplified sequence. As shown in Fig. 3B, the product of this reaction is a 120 bp fragment, that could only be derived from mature transcripts, because the probe

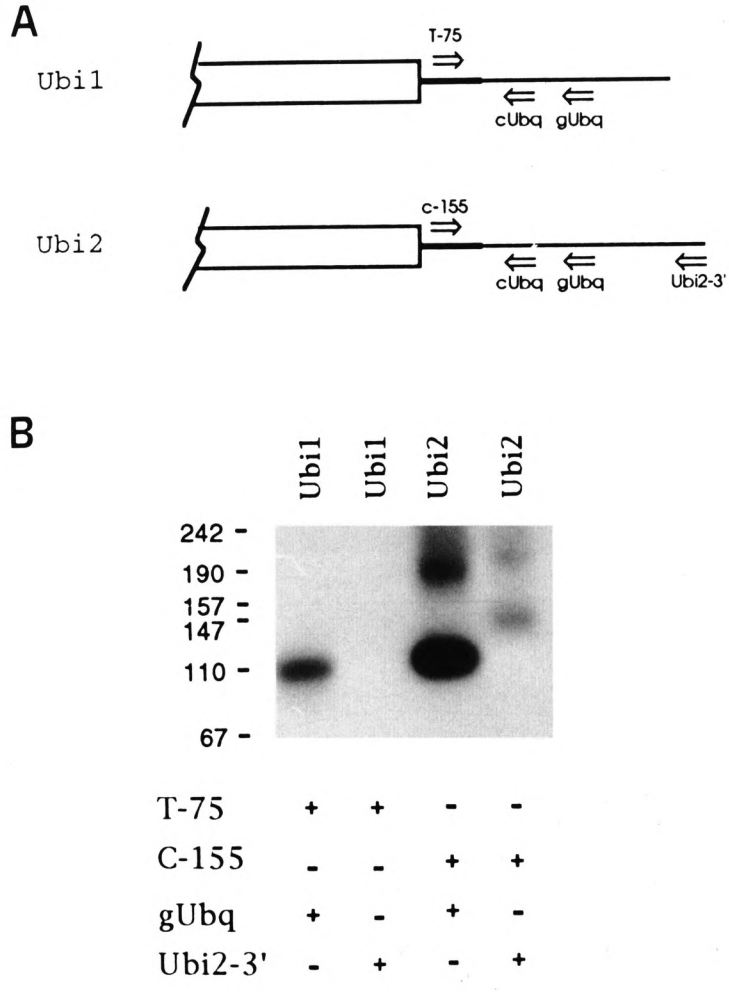


FIG. 5. Expression of Ubi1 and Ubi2 genes. (A) Diagram of the 3′ regions of Ubi1 and Ubi2 genes, showing the locations of the sense and antisense primers used. (B) cDNA made from RNA from 12 h of development was amplified with the primers indicated under each lane, separated by gel electrophoresis in 3% agarose, blotted, and hybridized with cUbq oligonucleotide. Molecular weight markers (in bp) are derived from pBlueskript plasmid digested with *Hpa*II.

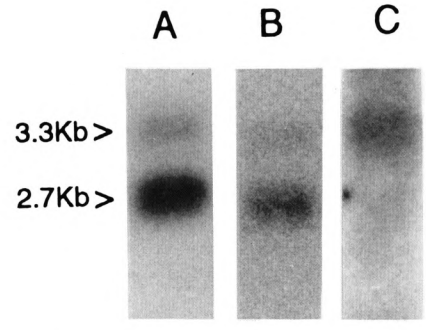


FIG. 6. Identification of Ubi2 mRNA as the 3.3 kb band. Poly(A) + RNA obtained from 25 h of development (15 μg) was electrophoresed and hybridized sequentially to (A), the 130 bp fragment of the 5′ end obtained by anchored PCR, (B) the 3′ end 137 bp fragment of pcUbq, and (C), the Ubi2-3′ oligonucleotide.

C-1 is internal to the amplified region. The same protocol was carried out using genomic DNA with negative results, confirming the presence of the intron. We have also carried out primer extension assays using oligonucleotide C-1, and again, the strongest band is of about 100 bp (not shown). Negative results were obtained making RNAse protection assays with RNA transcribed from the *Pst*I-*Hind*III fragment (0.9 kb) of λgArtU5, and also in hybridizations of RNA with this fragment or oligonucleotides derived from the proposed intron.

Expression of Ubiquitin Genes

Figure 4 shows the hybridization of RNAs isolated from different stages of development with a probe containing only ubiquitin coding sequences. There are two mRNAs, of 2.7 and 3.3 kb, respectively, that are more abundant the smaller the species. It is interesting to notice the fact that ubiquitin mRNAs are already present in dormant embryos and that the maximum level of expression is found at 8 h after resumption of development, about 8 h before hatching.

The experiment shown in Fig. 3 demonstrates that Ubi1 is transcribed and processed, and also rules out the possibility that the 3.3 kb mRNA is a nonprocessed messenger, as the intron of Ubi1 expands at least the 3.5 kb present in the upstream region of λgArtU5.

The experiment shown in Fig. 5 demonstrates that both Ubi1 and Ubi2 are expressed. cDNA was synthesized from RNA obtained at 12 h of development using an oligo(dT) primer and PCR was performed with sense primers derived from the C-terminal extensions (T-75 from Ubi1 and C-155 from Ubi2) and gUbq and Ubi2-3′ as antisense primers. As shown in the diagram of Fig. 5A (and also in the sequences in Figs. 1 and 2), gUbq is present in both genes (with four internal mismatches in Ubi2), whereas Ubi2-3′ is only present in Ubi2. Figure 5B shows the autoradiogram of the PCR products blotted and hybridized with cUbq oligonucleotide, present in both genes. As can be seen, the sense primer derived from Ubi2 gives positive bands with both gUbq (115 bp) and Ubi2-3′ (225 bp), whereas the Ubi1 sense primer gives only a positive band with gUbq (108 bp) but not with Ubi2-3′. (We think that the lower band present in the amplification with C-155 and Ubi2-3′ and the upper band in the reaction with C-155 and gUbq are artifacts due to the structure of the oligonucleotides and the low complexity of the 3′ end region.) These results confirm that there are transcripts from the two genes, and also that the Ubi1 transcripts do not contain in their 3′ end the sequences of Ubi2-3′.

The transcripts from Ubi1 and Ubi2 are nevertheless very similar. Figure 6 shows that the mRNAs of 3.3 and 2.7 kb hybridize with the 130 bp fragment from the 5′ end isolated by anchored PCR (Fig. 6A), and also with a 137 bp fragment of pcUbq that contains Ubi2-3′, from nucleotide 737–874 in Fig. 1. However, the Ubi2-3′ sequence is specific for this gene, and hybridizes only with the 3.3 kb transcript (Fig. 6C).

Although there are no canonical polyadenylation signals in the Ubi1 and Ubi2 sequences, there is a ATTAAA element [the second more abundant (Sheets et al., 1990) at positions 3591 in Ubi1 and 837 in Ubi2, underlined in Figs. 1 and 2].

DISCUSSION

We have found several cDNAs and a polyubiquitin gene in the crustacean *Artemia*. The polyubiquitin gene, that we call Ubi1, has nine ubiquitin units, all of them identical in terms of amino acid sequence, and also identical to the ubiquitin sequence of *D. melanogaster, X. laevis*, chicken, and humans. Less evolved eukaryotes (*T. cruzzi, D. discoideum*, yeast) and plants have two or three differences with respect to this sequence.

We have isolated cDNAs from two independent libraries. The identity between the sequence of the cDNA found and the one used for its isolation, in contrast with the variation within ubiquitin units in a given gene, points to *Artemia* having at least two polyubiquitin genes, Ubi1 and the one represented by pcUbq (Ubi2). Ubi1 has an intron in the 5′ untranslated region with a minimum size of 3.5 kb, in contrast to the size of this intron in other organisms, of much smaller size, 654 bp in chicken, 766 bp in *Drosophila*, or 715 bp in humans. It is noteworthy that other *Artemia* genes studied in our laboratory also have large introns when compared to their vertebrate homologs, as *ras* (M. Díaz-Guerra and J. Renart, in preparation), the Ca/Mg-ATPase (R. Escalante and L. Sastre, in preparation), and Na/K-ATPase (A. Garcia and L. Sastre, in preparation). This is not surprising given the large genome size of this crustacean, of 1.5×10^9 bp per haploid genome (Cruces et al., 1986).

We have detected the 5′ end of the transcript corresponding to Ubi1 by anchored PCR, and have also demonstrated that the intron is processed. Although we do not know whether Ubi2 also has an intron, it has the same 5′ end, as its corresponding mRNA (3.3 kb) hybridizes with the 5′ end derived from Ubi1. Similarly, both genes have very similar 3′ nontranslated regions.

The 3′ end of the Ubi1 transcript is not strictly

defined in our sequence; however, it cannot be too far away from the end of the sequenced region: the coding sequence has 2.1 kb, and the 5′ and 3′ ends around 100 bp each (i.e., 2.3 kb); the poly(A) tail would have about 400 nucleotides to give the 2.7 kb detected for the RNA band. The structure of the Ubi2 mRNA is more difficult to assess, although one could speculate with having more ubiquitin repeats, up to 12, assuming that all other regions are equivalent to the Ubi1 mRNA.

The mRNAs for ubiquitin are already present in the encysted *Artemia* embryo, and reach a maximum 8 h after resumption of development, around the time of emergence; the levels decrease thereafter. This pattern of expression is similar to that found in the *ras* gene (Díaz-Guerra et al., 1989). With respect to ubiquitin, this type of expression could be a prerequisite for increased levels of ubiquitin that could be needed for the early stages of development. A lysosomal degradation of yolk has been described during these stages (Perona et al., 1988), but the involvement of the ubiquitin degradation pathway has not been studied to date.

ACKNOWLEDGEMENTS

The photography work was done by Antonio Fernández. We thank Drs. Rafael Garesse and Leandro Sastre for helpful ideas and the critical reading of the manuscript. The technical assistance of Elvira Domínguez and Carmen Moratilla is also acknowledged. This work was financed by grant PB87-0208 from the CICYT, Spain.

REFERENCES

R. T. Baker and P. G. Bord (1987), Nucleic Acids Res 15, 443–463.

U. Bond and M. J. Schlesinger (1986), Mol Cell Biol 6, 4602–4610.

R. A. Browne, P. Sorgeloos, and C. N. A. Trotman (1991), *Artemia* Biology, CRC Press, Boca Raton, FL.

H. Busch and L. L. Goldknopf (1981), Mol Cell Biochem 40, 173–187.

H. L. Cabrera y Poch, C. Arribas, and M. Izquierdo (1990), Nucleic Acids Res 18, 3994.

E. J. Chen and P. H. Seeburg (1985), DNA 4, 165–170.

J. Cruces, M.-L. Wonenburger, M. Díaz-Guerra, J. Sebastián, and J. Renart (1986), Gene 44, 341–345.

M. Díaz-Guerra (1987), PhD thesis, Universidad Autónoma de Madrid.

M. Díaz-Guerra, M. Quintanilla, I. Palmero, L. Sastre, and J. Renart (1989), Biochem Biophys Res Commun 162, 802–808.

R. L. Dorit and O. Ohara (1992), in Current Protocols in Molecular Biology (F. M. Asubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, eds.), John Wiley & Sons, New York, pp. 15.6.1–15.6.8.

R. Escalante and L. Sastre (1993), J Biol Chem 268, 14090–14095.

A. P. Feinberg and B. A. Vogelstein (1983), Anal Biochem 132, 6–13.

D. Finley, B. Bartel, and A. Varshavsky (1989), Nature 338, 394–401.

D. Finley, A. Ciechanover, and A. Varshavsky (1984), Cell 37, 43–45.

A. Frischauf, H. Lehrach, A. Poustka, and N. Murray (1983), J Mol Biol 170, 827–842.

M. Glotzer, A. W. Murray, and M. W. Kirschner (1991), Nature 349, 132–138.

A. Hershko, A. Ciechanover, H. Heller, A. L. Haas, and I. Rose (1980), Proc Natl Acad Sci USA 77, 1783–1786.

A. Hershko, E. Leshinsky, D. Ganoth, and H. Heller (1984), Proc Natl Acad Sci USA 81, 1619–1623.

M. Izquierdo, C. Arribas, J. Galcerán, J. Burke, and V. M. Cabrera (1984), Biochim Biophys Acta 783, 114–121.

K. N. Kasturi and C. A. Bona (1991), Nucleic Acids Res 19, 6339–6340.

E. S. Kawasaki (1990), in PCR Protocols. A Guide to Methods and Applications (M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, eds.), Academic Press, San Diego, pp. 21–27.

H. Lee, J. A. Simon, and J. T. Lis (1988), Mol Cell Biol 8, 4727–4735.

H. Mori, J. Kondo, and Y. Ihara (1987), Science 235, 1641–1644.

C. Osuna and J. Sebastián (1980), Eur J Biochem 109, 383–389.

E. Özkaynak, D. Finley, M. Solomon, and A. Varshavsky (1987), EMBO J 6, 1429–1439.

I. Palmero, J. Renart, and L. Sastre (1988), Gene 68, 239–248.

R. Perona, J. C. Bés, and C. G. Vallejo (1988), Biol Cell 63, 361–366.

G. Perry, R. Friedman, G. Shaw, and V. Chau (1987), Proc Natl Acad Sci USA 84, 3033–3036.

M. Rechsteiner (1987), Ann Rev Cell Biol 3, 1–30.

M. Rechsteiner, ed. (1988), Ubiquitin, Plenum Press, New York.

K. L. Redman and M. Rechsteiner (1989), Nature 338, 438–440.

J. Sambrook, E. F. Fritsch, and T. Maniatis (1989), Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

M. D. Sheets, S. C. Ogg, and M. P. Wickens (1990), Nucleic Acids Res 18, 5799–5805.

M. Siegelman, M. W. Bond, W. M. Gallatin, T. St.John, H. T. Smith, W. A. Fried, and I. L. Weismann (1986), Science 231, 823–829.

M. Stürzl and W. K. Roth (1990), Anal Biochem 185, 164–169.

P. Thomas (1980), Proc Natl Acad Sci USA 77, 5201–5205.

S. Wolf, F. Lottspeich, and W. Baumeister (1993), FEBS Lett 326, 42–44.

R. S. Wu, K. W. Kohn, and W. M. Bonner (1981), J Biol Chem 256, 5916–5920.

T. Yarden, J. A. Escobedo, W.-J. Kuang, T. L. Yang-Fen, T. O. Daniel, M. P. Tremble, E. Y. Chen, M. E. Ando, R. M. Harkins, U. Franke, V. A. Fried, A. Ulrich, and L. T. Williams (1986), Nature 323, 226–232.