Transcript regulation and carboxyterminal extension of ubiquitin in *Dictyostelium discoideum*

M. Westphal, A. Müller-Taubenberger, A. Noegel and G. Gerisch

Max-Planck-Institut für Biochemie, 8033 Martinsried bei München, FRG

Received 10 September 1986

The sequence of a cDNA from *Dictyostelium discoideum* indicates linkage of a basic polypeptide to the C-terminal end of ubiquitin. Among the six RNA species recognized by a ubiquitin-specific fragment of the cDNA, four are increasing in quantity during development and two are decreasing. Six genomic *Eco*RI fragments of more than 3 kb are recognized by that probe, suggesting that the RNAs are transcribed from a family of separate genes. A fragment containing only the 3'-portion of the cDNA which includes part of the coding region for the basic polypeptide but not for ubiquitin, recognizes mainly two genomic fragments of which only one coincides in size with a fragment that is recognized by the ubiquitin-specific probe.

Ubiquitin *Precursor sequence* *(Dictyostelium)* *Gene regulation* *Multigene family*

1. INTRODUCTION

Ubiquitin is a protein with a highly conserved sequence of 76 amino acids which has been found in all eukaryotes examined: in yeast [1], *Drosophila* [2], *Xenopus* [3], chicken [4], bovine [5], and human [6]. It is involved in ATP-dependent non-lysosomal proteolysis [7], is one of the heat-shock induced proteins in chicken [4], and is coupled to the homing receptor on the surface of lymphocytes [8] as well as to histone H2A in nuclei [9]. The linkage of ubiquitin to histone may play a role in the control of chromatin structure since in *Drosophila* it appears to be preferentially associated with actively transcribed genes [10]. *Dictyostelium discoideum* is a microorganism proceeding from an amoeboid stage to multicellularity by cell aggregation, in which ubiquitin may serve different functions at subsequent stages of development. Here we report on the sequence of a ubiquitin cDNA clone and on developmental changes in the ubiquitin transcript pattern.

2. MATERIALS AND METHODS

2.1. *cDNA cloning and sequence analysis*

A cDNA library in plasmid p2732B was constructed as described previously using poly(A^+) RNA from *D. discoideum* AX2 cells harvested after 6 h of starvation [11]. For sequencing, cUB1 was cut with restriction enzymes as shown in fig.1, and subcloned into phages M13mp8 and mp9 [12] and sequenced using the dideoxynucleotide chain termination method [13].

2.2. *DNA and RNA isolation and hybridization*

For Southern blotting, DNA was isolated from purified nuclei as described [14], digested with *Eco*RI (Boehringer Mannheim), separated on 1% agarose gels in Tris-phosphate buffer [15], transferred to BA85 nitrocellulose (Schleicher and Schüll) and hybridized as indicated with nick-translated cDNA inserts for 16–18 h in 50% formamide, 2 × SSC, 4 × Denhardt's, 1% sarcosyl, 0.12 M sodium phosphate buffer, pH 6.8, and 0.1% SDS at 37°C.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

```
                                                 1                                            44
                                                 CC ATC ACT CTC GAA GTT GAA GGT TCA GAC AAC ATC GAA AAC GTC   cUB1
                                                    Ile Thr Leu Glu Val Glu Gly Ser Asp Asn Ile Glu Asn Val   DUB
Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr ... ... ... ... ... ... Pro ... ... Thr ... Glu ... ...   HUB
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... Ser ... ... Thr ... Asp ... ...   SUB

45                                                                                                 122
AAA GCC AAG ATT CAA GAC AAA GAA GGT ATC CCA CCA GAT CAA CAA CGT CTT ATC TTT GCC GGT AAA CAA TTA GAA GAT   cUB1
Lys 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   DUB
... Ala ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...   HUB
... Ser ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...   SUB

123                                                                                      ↓     200
GGT CGT ACT TTA TCC GAT TAT AAC ATC CAA AAG GAA TCA ACC CTC CAC TTA GTT TTA AGA TTA AGA GGT GGT ATT GAA   cUB1
Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Ile Glu   DUB
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... Ala Lys   HUB
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... Asn Stop  SUB

201                                                                                  ?             278
CCA TCC CTC GTC ATT CTT GCC CGT AAA TAC AAA TGT GAC AAA ATG ATT TGC AGA AAA TGT TAT GCT CGT TTA CAC CCA   cUB1
Pro Ser Leu Val Ile Leu Ala Arg Lys Tyr Lys Cys Asp Lys Met Ile Cys Arg Lys Cys Tyr Ala[Arg]Leu His Pro   DUB
Lys Arg Lys Lys Lys Ser Tyr Thr Thr Pro ... Lys Asn ... His Lys Arg Lys ... Val Lys Leu Ala Val Leu Lys   HUB

279                                                                                                356
CGT GCT GTT AAC TGT CGT AAG AAG GAA ATG TGG TCA CTC TAA CAA TTT AAG ACC AAA GAA GAA GTT ACT TAA ATA AAT   cUB1
Arg Ala Val Asn Cys Arg Lys Lys Glu Met Trp Ser Leu Stop                                                  DUB
Tyr Tyr Lys Val Asp Glu Asn Gly Lys Ile Ser Arg ... Arg Arg Glu Cys Pro Ser Asp Glu Cys Gly Ala Gly Val   HUB

357                                                                                                434
TCT TAA AGA AAT TAC TAA AAA AAG TAT ACT ATT CCT CAA AAT AAT AAA AAA TGT CTT TTA TTT TTT TAT TTA AAA AAA   cUB1
►                                                                                                          ◄
Phe Met Ala Ser His Phe Asp Arg His Tyr Cys Gly Lys Cys Cys Leu Thr Tyr Cys Phe Asn Lys Pro Glu Asp Lys   HUB

435                                 463
AAA AAA AAA AAA AAA AAA AAA AAA AAA AA                                                                    cUB1
►                                                                                                          ◄
Stop                                                                                                      HUB
```

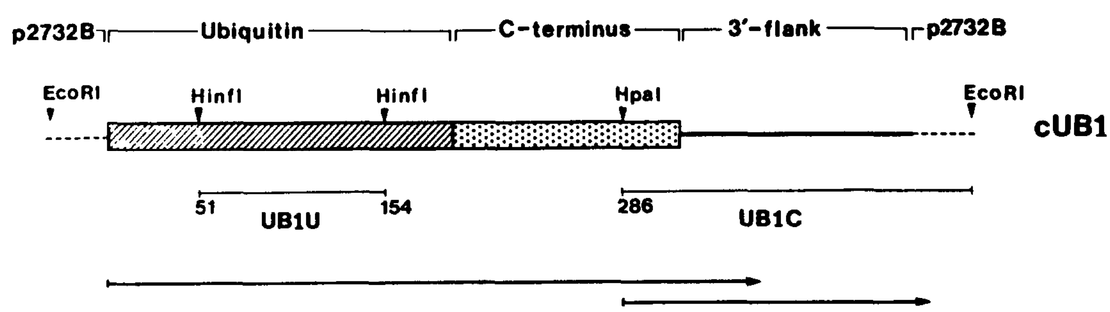

Fig.1. Sequence and structure of the *D. discoideum* clone cUB1. Top: cDNA sequence and comparison of the derived amino acid sequence of cUB1 with human (HUB) and yeast (SUB) ubiquitin sequences. The arrow after position 194 indicates the end of the ubiquitin sequence. Potential polyadenylation signals are underlined. Horizontal arrowheads separate the untranslated 3'-terminal nucleotide sequence of cUB1 from the HUB 3'-terminal amino acid sequence. Bottom: Organisation of the cUB1 cDNA, boundaries of two fragments, UB1U and UB1C, that have been used as hybridization probes, and strategy of sequencing as indicated by arrows.

For Northern blots, total cellular RNA (10 μg per lane) isolated from different developmental stages of *D. discoideum* strain AX2-214 cells was separated on 1.2% agarose gels in the presence of 6% formaldehyde [15], transferred to nitrocellulose and hybridized as described above. For development at 21°C, cells grown in a suspension of 1×10^{10} *E. coli* B/r cells per ml of 17 mM Soerensen phosphate buffer, pH 6.0, were washed and transferred onto Millipore filters Type HABG [16]. For the preparation of prespore or prestalk specific RNA, cells from the slug stage were separated on a Percoll-gradient [17]. The prespore cell fraction was identified by the use of a monoclonal antibody (Mud 1) that recognized a prespore-specific cell surface antigen [18]. Prespore cells were 95% enriched, prestalk-cell enriched fractions contained 18% prespore cells.

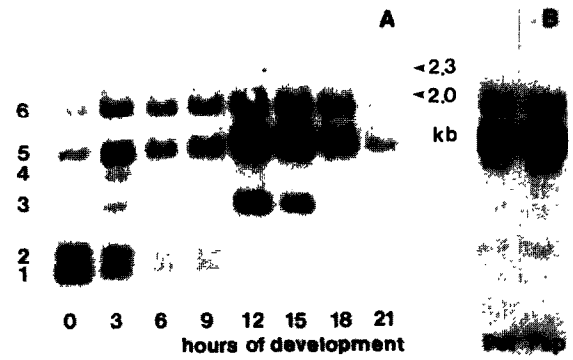

Fig.2. Northern blot analysis of ubiquitin transcripts. (A) Changes of transcripts during development. Cells were starved at 0 h and proceeded through aggregation (12 h), slug formation (15–18 h), and fruiting body formation (21 h). (B) Transcripts in prestalk (Pst) and prespore (Psp) cells separated from slugs. Blot A was probed with fragment UB1U, blot B with cUB1.

3. RESULTS

3.1. *Sequence of an ubiquitin cDNA clone*

Restriction sites and the sequence of a cDNA clone, cUB1, from *D. discoideum* are shown in fig.1. The 463 bp insert indicates one open reading frame of 317 bases, a non-translated region of 110 bases containing two potential polyadenylation signals in positions 350 and 399, respectively, and a poly(A) tail 24 bases downstream of the second polyadenylation signal. While comparison on the DNA level showed 73% homology to the human ubiquitin gene, the sequence of the first 64 amino acids of the derived protein sequence proved to be almost completely identical to amino acids 13–76 of human as well as yeast ubiquitin (fig.1). The ubiquitin sequence is immediately followed by a sequence of 41 amino acids which shows no significant homology with other known proteins.

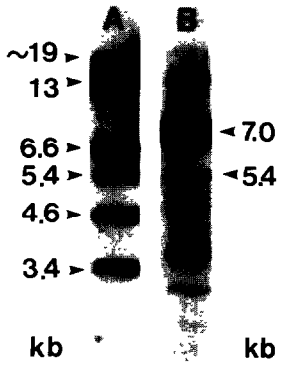

Fig.3. Southern blot analysis of *Eco*RI cut DNA. (A) Probed with the ubiquitin-specific fragment UB1U; (B) probed with the 3'-terminal fragment UB1C.

3.2. *Developmental regulation of multiple ubiquitin transcripts*

A ubiquitin-specific probe, UB1U, was prepared by digesting cUB1 with *Hin*fI and isolating a 104 bp fragment that contained part of the sequence encoding ubiquitin (fig.1). This probe was hybridized to Northern blots containing total RNA from different developmental stages. Six RNA species were recognized (bands 1 to 6 in fig.2A), the largest one having a size of approx. 2 kb. The two smallest RNA species were most prominent during the growth phase, their amount declined when development was initiated by starvation. The four larger RNA species showed variable changes during development. Their amount increased during the first 3 h and, except for band 4 RNA, again after 9 h of development. Stimulation of cells with pulses of 20 nm cyclic AMP every 6 min, a treatment which promotes development, accelerated the changes in relative abundance of ubiquitin RNA species (not shown).

The two major cell types of fruiting bodies were separated from the multicellular slug stage by den-

sity gradient centrifugation, and RNA was extracted from these prestalk and prespore cells. No cell-type specific differences were seen when a cUB1 probe was hybridized to Northern blots (fig.2B).

3.3. *Genomic DNA fragments recognized by the 5'- and 3'-portions of the cDNA*

The isolated cDNA did not contain *Eco*RI sites. Nevertheless, the ubiquitin-specific probe UB1U hybridized in genomic Southern blots to a minimal number of six different *Eco*RI fragments with sizes of more than 3 kb (fig.3A), which suggests the presence of multiple ubiquitin genes in the *D. discoideum* genome. When, after melting off the ubiquitin specific probe, the same blot was probed with the 3'-terminal fragment UB1C that did not include the ubiquitin-specific sequence (fig.1), primarily two fragments of 5.4 and 7.0 kb became visible (fig.3B). Only the smaller one coincided in size with a fragment recognized by the ubiquitin-specific probe UB1U.

4. DISCUSSION

Our results indicate that *D. discoideum* ubiquitin is encoded by a multigene family. Multiple ubiquitin genes are also present in the human genome [19] while *Saccharomyces cerevisiae* has a single ubiquitin gene [1]. At least some of the human genes contain multiple copies of ubiquitin encoding sequences in tandem, and the same is true for the yeast gene. Different strengths of hybridization of the ubiquitin-specific probe UB1U to the six *Eco*RI bands recognized in *D. discoideum* genomic DNA (fig.3A) suggest that the fragments contain varying numbers of ubiquitin encoding sequences which might be arranged in tandem. In Northern blots the same probe hybridized to at least six RNA species that differed from each other in size and developmental regulation, a situation which parallels that found for *Xenopus* ubiquitin transcripts [3].

The strong homology of the cDNA-derived sequence of *D. discoideum* ubiquitin with human as well as yeast ubiquitin (fig.1) is not matched by a similarly strong homology on the DNA level because of different codon usages. Of particular interest in our cDNA clone is the region coding for 41 amino acid residues that is linked to the 3'-end of the region coding for ubiquitin. A probe specific for the 3'-sequence did not hybridize to the majority of genomic *Eco*RI fragments coding for ubiquitin. The probe most strongly hybridized to a separate band (fig.3B). Since the sequenced cDNA contained no *Eco*RI site, the 3'-coding sequence appears to be linked in the genome either to ubiquitin-specific sequences through an intron that contains an *Eco*RI site, or to genes that code for proteins other than ubiquitin. The latter is less likely since we did not find any abundant RNA species that is recognized by the probe containing the 3'-coding sequence but not by the ubiquitin-specific probe.

The 3'-coding sequence in *Dictyostelium* is comparable to a ubiquitin-linked sequence in a human cDNA [20], which is represented in yeast ubiquitin cDNA only by a single codon for asparagine [1]. The polypeptide portion encoded by the *Dictyostelium* 3'-sequence is, like the human one, rich in basic amino acids and contains several cysteine residues. Although this general similarity is not complemented by a significant sequence homology, the *Dictyostelium* C-terminal polypeptide may serve the same function that has been suggested for the human one [20], to target the ubiquitin precursor into the nucleus where the ubiquitin may participate in gene regulation.

ACKNOWLEDGEMENTS

We thank our colleague Dr Peter Hirth and Andrea Ohmayer for providing us with the UB1 clone, and Dr Kees Weijer for prespore/prestalk separation.

REFERENCES

[1] Özkaynak, E., Finley, D. and Varshavsky, A. (1984) Nature 312, 663–666.
[2] Gavilanes, J.G., Gonzalez de Buitrago, G., Perez-Castells, R. and Rodriguez, R. (1982) J. Biol. Chem. 257, 10267–10270.
[3] Dworkin-Rastl, E., Shrutkowski, A. and Dworkin, M.B. (1984) Cell 39, 321–325.
[4] Bond, U. and Schlesinger, M.J. (1985) Mol. Cell. Biol. 5, 949–956.
[5] Schlesinger, D.H., Goldstein, G. and Niall, H.D. (1975) Biochemistry 14, 2214–2218.
[6] Schlesinger, D.H. and Goldstein, G. (1975) Nature 255, 423–424.

[7] Ciechanover, A., Finley, D. and Varshavsky, A. (1984) J. Cell. Biochem. 24, 27–53.
[8] Siegelman, M., Bond, M.W., Gallatin, W.M., St. John, T., Smith, H.T., Fried, V.A. and Weissman, I.L. (1986) Science 231, 823–829.
[9] Matsui, S.-I., Seon, B.K. and Sandberg, A.A. (1979) Proc. Natl. Acad. Sci. USA 76, 6386–6390.
[10] Levinger, L. and Varshavsky, A. (1982) Cell 28, 375–385.
[11] Gerisch, G., Hagmann, J., Hirth, P., Rossier, C., Weinhart, U. and Westphal, M. (1985) Cold Spring Harbor Symp. Quant. Biol. 50, 813–822.
[12] Messing, J. and Vieira, J. (1982) Gene 19, 269–276.
[13] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[14] Noegel, A., Welker, D.L., Metz, B.A. and Williams, K.L. (1985) J. Mol. Biol. 185, 447–450.
[15] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, NY.
[16] Newell, P.C., Telser, A. and Sussman, M. (1969) J. Bacteriol. 100, 763–768.
[17] Weijer, C.J., McDonald, S.A. and Durston, A.J. (1984) Differentiation 28, 13–23.
[18] Krefft, M., Voet, L., Gregg, J.H., Mairhofer, H. and Williams, K.L. (1984) EMBO J. 3, 201–206.
[19] Wiborg, O., Pedersen, M.S., Wind, A., Berglund, L.E., Marcker, K.A. and Vuust, J. (1985) EMBO J. 4, 755–759.
[20] Lund, P.K., Moats-Staats, B.M., Simmons, J.G., Hoyt, E., D'Ercole, A.J., Martin, F. and Van Wyk, J.J. (1985) J. Biol. Chem. 260, 7609–7613.