Schiebel, W., & Schneck, U. (1974) *Hoppe Seyler's Z. Physiol. Chem. 355*, 1515-1525.
Scott, A. I. (1964) *Interpretation of the ultraviolet spectra of natural products*, p 38, Pergamon Press, Oxford.
Strominger, J. L., Park, J. T., & Thompson, R. E. (1959) *J. Biol. Chem. 234*, 3263-3268.
Thomm, M., & Stetter, K. O. (1985) *Eur. J. Biochem. 149*, 345-351.
Weber, C., Fischer, H., & Holler, E. (1988) *Eur. J. Biochem. 176*, 199-206.
Zabel, H.-P., Fischer, H., Holler, E., & Winter, J. (1985) *Syst. Appl. Microbiol. 6*, 111-118.

Molecular Organization of Developmentally Regulated *Dictyostelium discoideum* Ubiquitin cDNAs†

Tetsuo Ohmachi,‡ Roberto Giorda,§ David R. Shaw, and Herbert L. Ennis*
Roche Institute of Molecular Biology, Roche Research Center, Nutley, New Jersey 07110
Received December 9, 1988; Revised Manuscript Received March 9, 1989

ABSTRACT: *Dictyostelium discoideum* ubiquitin mRNAs are regulated in a complex fashion during spore germination and multicellular development. Species of mRNA of 1900, 1400, 1100, 840, 580, and 500 nucleotides (nt) are found which are expressed differentially during different stages of development. DNA blot analysis indicates that ubiquitin genes constitute a multigene family of at least six genes. cDNAs representing all the ubiquitin mRNA transcripts were isolated and sequenced. The *Dictyostelium* mRNAs are organized as tandem repeats of the 76 amino acid ubiquitin unit (228 nt). We isolated one cDNA containing seven of these tandem repeats, and two different five- and three-repeat cDNAs. In addition, 2 cDNAs containing a single ubiquitin repeat fused at its 3′ end to an unreleated 52 and 78 amino acid extension were identified. There is a remarkable similarity in the sequences of the non-ubiquitin extensions among yeast and mammalian counterparts. The extensions are very basic, containing approximately 30% lysine/arginine. Another common feature of these proteins is the presence of a common structural motif containing cysteine residues at conserved positions, suggesting a metal binding domain that matches a consensus sequence of *Xenopus* transcription factor TFIIIA and other nucleic acid biniding proteins. The characterization of ubiquitin cDNAs and genomic sequences in *D. discoideum* now makes the understanding of its developmental regulation feasible.

The simple eukaryote *Dictyostelium discoideum* has been used as a paradigm to study developmental phenomena. In this organism, the depletion of the food supply sets in motion a schedule of development which results in the orderly succession of developmental events, and which directs when and to what extent in the process, and where in the developing organism, a phenotypic trait will be expressed. Aggregation of individual cells in response to a chemotactic stimulus triggers new cell programs of gene expression that allow the construction of a multicellular structure. Once in the aggregate, the cells are now subject to signals depending on their position within the multicellular structure. These signals dictate types of programs of gene expression that will be followed, e.g., will the cell form a spore, stalk (Sussman & Brackenbury, 1976).

We have been interested in identifying proteins and mRNAs that are exclusively expressed during specific stages of spore germination or multicellular development (Dowbenko & Ennis, 1980; Kelly et al., 1983; Giorda & Ennis, 1987). During our studies, we isolated a cDNA, denoted pLK229, which showed an unusual pattern of regulation during *D. discoideum* development and which coded for a protein identical with human ubiquitin, except that proline in position 19 and threonine in position 22 of the human species were glycine and asparagine, respectively, in *Dictyostelium* (Giorda & Ennis, 1987). An interesting cDNA clone containing a basic polypeptide linked to the C-terminal end of ubiquitin has also been described (Westphal et al., 1986; Müller-Taubenberger et al., 1988a,b).

RNA blot analysis of RNA from vegetative cells, germinating spores, and multicellular development indicates that pLK229-specific mRNA is developmentally regulated (Giorda & Ennis, 1987; Westphal et al., 1986; Müller-Taubenberger et al., 1988a). mRNA species of about 1900, 1400, 1100, 840, 580, and 500 nucleotides (nt) which hybridized to pLK229 plasmid DNA were observed. The 1400-nt species was the only mRNA present in spores. At 1.5 h of spore germination, four mRNAs were observed of 1900, 1400, 1100, and 840 nt, whereas at 3 h germination the predominant species were the 1900- and 1400-nt mRNAs and only traces of the 1100- and 840-nt mRNAs were present. Little of the 1900- to 840-nt mRNAs was present in growing cells, but two new bands at 580 and 500 nt were observed. During multicellular development on filters, these two mRNAs disappeared, and the 1900- and 1400-nt species predominated.

Southern blots of *Dictyostelium* DNA digested with various restriction enzymes demonstrated that the pLK229 genes constituted a multigene family of at least six genes (Giorda & Ennis, 1987; Westphal et al., 1986). As had been described for other organisms, the ubiquitin genes and mRNAs are composed of identical tandem repeats of the 76 amino acid ubiquitin protein (polyubiquitin) (Dworkin-Rastl et al., 1984; Özkaynak et al., 1984; Vierstra et al., 1986; Arribas et al.,

† The nucleic acid sequences in this paper have been submitted to GenBank under Accession Number J02858.
* To whom correspondence should be addressed.
‡ Present address: Research Institute for Tuberculosis and Cancer, Tohoku University, 4-1 Seiryomachi, Sendai 980 Japan.
§ Present address: Pittsburgh Center Innstitute, Basic Research Facility, 3343 Forbes Ave., Pittsburgh, PA 15213.

0006-2960/89/0428-5226$01.50/0

1986; Wiborg et al., 1985). Since, as described above, a Northern blot indicated at least six transcripts, it seemed that we had the opportunity to isolate cDNAs representing all transcribed ubiquitin genes. Therefore, λgt10 cDNA libraries representing spore, 1.5-h spore germination, and growing cell mRNA were constructed and screened with pLK229 as probe. We were successful in isolating seven different cDNA clones, and the characterization of these is the subject of this study.

Materials and Methods

Organism. *D. discoideum* AX3 was grown axenically in HL-5 medium (Sussman & Sussman, 1967). Spore germination of strain B spores was carried out as previously described (Ennis & Sussman, 1975).

Preparation of cDNA Libraries. λgt10 libraries were prepared from poly(A)+ RNA isolated from spores, 1.5-h germinating spores, and growing cells, using an Amersham cDNA synthesis and cloning kit and the procedure described by the manufacturer. The libraries were screened by using nick-translated pLK229 cDNA as a probe. Positive clones were denoted DCUB for *Dictyostelium* cDNA *ub*iquitin. The numbers are merely individual isolates in the order they were obtained. DCUB19 and -14 were isolated from spore and 1.5-h libraries, DCUB14, -17, and -18 from a 1.5-h library, and DCUBV45 and -V20 from a growing cell library.

Sequencing. λgt10 DNA containing the insert was isolated as described (Maniatis et al., 1982). The insert was excised by digestion of the DNA with *Eco*RI and was isolated by size fractionation on a 1.0% agarose gel.

Insert derived from DCUBV45 was ligated into the *Eco*RI site of Bluescript SK- (Stratagene, San Diego, CA), while all the other inserts were ligated into the *Eco*RI site of M13mp18. Subclones were prepared by digestion with exonuclease III as previously described (Henikoff, 1984). Cloning and sequencing procedures followed were those of the manufacturer of the cloning and sequencing kits, with minor modifications, using the dideoxy chain termination procedure (Sanger et al., 1977).

The sequences obtained were stored and analyzed using DNASTAR programs (DNASTAR, Madison, WI).

Sources of Materials. Restriction endonucleases were obtained from Bethesda Research Laboratories, Inc. (Bethesda, MD), New England BioLabs, Inc. (Beverly, MA), and Boehringer Mannheim (Indianapolis, IN). Bethesda Research Laboratories was the sourruce of the nick-translation kit, *Escherichia coli* HB101-competent cells, M13 cloning and sequencing kits, the sequencing gel electrophoresis system, exonuclease III, T4 ligase, and Klenow fragment of DNA polymerase I, [α-^{32}P]dCTP, deoxyadenosine 5′-[α-^{35}S]thiotriphosphate, and the cDNA synthesis and λgt10 cloning kits were from Amersham Corp. (Arlington Heights, IL).

Results and Discussion

A large number of cDNAs that hybridized to pLK229 were isolated from spore, 1.5-h spore germination, and growing cell mRNA cDNA libraries. Seven different cDNA clones were resolved from this group, and each was sequenced. Figure 1 is a representation of the organization of the cDNAs based on sequence analysis and deduced proteins coded for by the observed open reading frames. The open boxes refer to the 76 amino acid (228-nt) ubiquitin repeat, and the extent of the 5′ and 3′ noncoding regions is presented.

Ubiquitin is a small protein containing 76 amino acids (Schlesinger et al., 1975). The amino acid sequences derived from each of the clones indicated that we isolated a single cDNA containing 7 tandem repeats of the 76 amino acid ubiquitin repeat, and 2 different 5- and 3-repeat cDNAs. Two

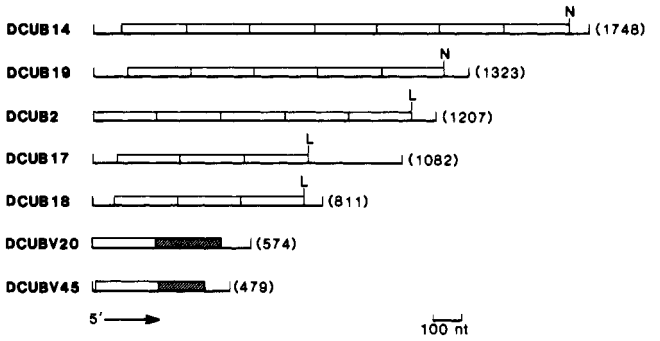

FIGURE 1: Organization of *D. discoideum* ubiquitin cDNAs. The figure depicts the sequences of all of the cDNAs isolated by us. The open rectangle depicts the ubiquitin 76 amino acid unit (228 nt). The hatched rectangle denotes the nonubiquitin extension (52 amino acids, 156 nt for DCUBV45; 78 amino acids, 234 nt for DCUBV20). The letters N for asparagine and L for leucine indicate the terminal non-ubiquitin amino acids in each polyubiquitin sequence. The thin lines represent the 5′ and 3′ noncoding regions. The 5′ region is at the left. The number in parentheses to the right of each diagram is the number of nucleotides in the respective cDNA. The clones are denoted by trivial names derived from the order in which they were isolated.

small cDNAs represent interesting fusions of 52 and 78 amino acids to the carboxyl terminus of one 76 amino acid ubiqutin unit. The 52 amino acid fusion has already been described (Westphal et al., 1986; Müller-Taubenberger et al., 1988a).

We previously reported, on the basis of Northern blot analysis, six species of ubiquitin sequence containing mRNAs: 1900, 1400, 1100, 840, 580, and 500 nt, respectively (Giorda & Ennis, 1987). Another group (Westphal et al., 1986; Müller-Taubenberger et al., 1988a,b) estimated the sizes to be 1900, 1400, 1200, 900, 700, and 600 nt. Thus, both groups are in reasonable agreement. However, Müller-Taubenberger (Müller-Taubenberger et al., 1988a,b) also reported a 1500-nt species which we have not observed. Perhaps this species was not resolved from the 1400-nt mRNA in our gels. The lengths of the cDNA clones we isolated are consistent with the sizes of the Northern analysis, although we do not know the exact size of the 5′ and 3′ noncoding regions. DCUB14 (1748 nt) probably is the 1900-nt mRNA species, DCUB19 (1323 nt) and DCUB2 (1207 nt) the 1500–1400-nt species, DCUB17 (1082 nt) the 1100-nt mRNA, DCUB18 (811 nt) the 840-nt mRNA, and DCUBV20 (574 nt) and DCUBV45 (479 nt) the 580-nt and 500- nt species, respectively. Consequently, we feel confident that we have been successful in identifying all the ubiquitin genes products transcribed during *D. discoideum* growth and development. It is possible, but we do not feel likely, that as with the case in yeast (Özkaynak et al., 1984) and *Arabidopsis* (Burke et al., 1988), some classes of mRNAs on the Northerns are composed of more than one species of mRNA. We previously reported the sequence of two genomic clones (Giorda & Ennis, 1987). Clone p229 is identical with DCUB19 (as well as the previously described cDNA pLK229). Clone λ229 is a genomic sequence we previously reported as having three ubiquitin repeats. We now find after more analysis that the sequence of this clone is identical with DCUB2 and contains five, not three, ubiquitin repeats.

In order to save space and to give the data in as informative a fashion as possible, we have presented the nucleotide sequence data in three sections. Figure 2 gives the complete nucleotide seuqence of the coding regions of all cDNAs. Each of the 228-nt (76 amino acid) repeats is compared to the DCUB14 first repeat, which is taken as an arbitrary reference, and only the differences between DCUB14 repeat 1 and the indicated repeat are presented. Figures 3 and 4 present the

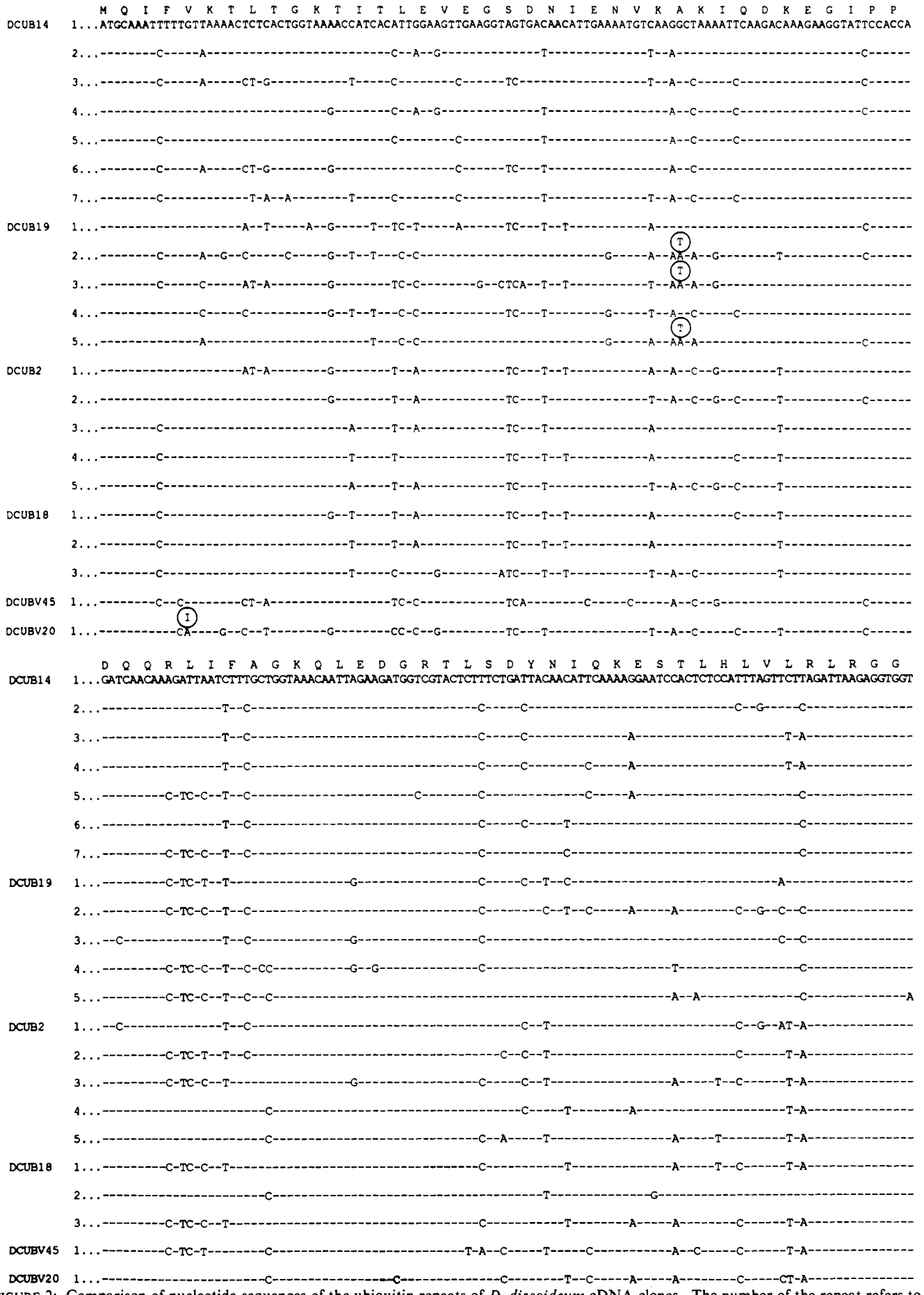

FIGURE 2: Comparison of nucleotide sequences of the ubiquitin repeats of *D. discoideum* cDNA clones. The number of the repeat refers to the order in which it appears in the cDNA. Only differences are given. An empty space indicates identity with DCUB14 repeat number 1, which was arbitrarily chosen as the reference sequence. Amino acids are denoted in the single-letter code.

```
DCUB14      ATTAATAATAATTTTTTAAATTCTTTATTTTTTTAAACATATAAAAATTTTTTTTTTCCTTTGTATATT
            ATTTTTTTTTCCATTTTTACATATTAATTAATTATG
DCUB19      AAATATTAAATTATATTTTATTTTTTTATATATTGTAACATTTCATAAAAAAAAAAAAAAAAAAAAAAT
            TAAAAATCAACTTATTTGAAAAAAAAAAAAAAAATAATAATTAATTATAATG
DCUB17      AAAATACAAATACAAATAACAAATACTTTACTATAGCTTTTTTTTTTCTTATTTATTTCTCCAAATAATTT
            TTTAATATG
DCUB18      AAATAACAAATACTTTACTATAGCTTTTTTTTTCTTATTTATTTCTCCAAATAATTTTTTAATATG
DCUBV45     AACAGATCAAGATG
```

FIGURE 3: 5′ noncoding ubiquitin cDNA sequences. The entire 5′ region is presented. The underlined ATG represents the first methionine codon in the deduced protein. Note that DCUB2 and DCUBV20 lack a 5′ noncoding region and the AT of the first ATG codon. DCUB17 and -18 are identical although the 5′ sequence of DCUB17 is larger than -18.

```
DCUB14      TAAATTTTGGTGCCAGAGTGAATAATAAATAAAAAAAAAAAAAAAA
DCUB19      TAAAATTACCCACACACAAAATTTATCTTATGTGACCTATTCCAAAAAAAAAAAAAAAAAAA
DCUB 2      TAAAAATAAATAAATAAAAAAAAAACTAAAACTTTAATCAGTTTTATTTCAACTTTTAAAAAAAAA
                                      ↓
DCUB18      TAAAAACTATAAATAAAACCACTTGTAAATATGTAAATCAAATTTCAAAAAAAAAAAAAAA
                                      ↓
DCUB17      TAAAAACTATAAATAAAACCACTTGTAAATATAAAAAAATAGATCCAATAGTAATAAATATATAAAAATTT
            TTTAATCTATTTTAGATTTTAAAAATAAAAAAAAAAAAGTTCAAGTTATTTTTTTTTTTATAAACATAACT
            TTATTTTATTATCATTATCCTTATTATCATATATATTAATTTTATTATTATTATTTTAATTATTATTTTT
            ATTACTATAGTTAATTTAATAAATATTAATAATAATAATAAAAATAATAATAATTTAAAAAAAATCTCTT
            TGTTATCAATTGATCTAAATTTCACTCGAATTTCAAAAT
DCUBV20     TAAACAATTTTCTTGAAAATATCATAATGCTTATCTAGTATGTGTTTTTCATGATTAGAGGACGTACTAT
            GCCATACCCAATTAAAACAAAAACTAAAAAAAAAAAAAAAAAAA
DCUBV45     TAAATAAATTCTTAAAGAAATTACTAAAAAAAGTATACTATTCCTCAAAATAATAAAAAATGTCTTTTAT
            TTTTTTAAAAAAAAAAAA
```

FIGURE 4: 3′ noncoding ubiquitin cDNA sequences. The entire 3′ region is presented. The stop codon, TAA, is underlined. Note the identity of the first 29 nucleotides of DCUB17 and -18. The sequences of the two cDNAs diverge after the arrows and are much longer in clone DCUB17.

5′ and 3′ noncoding sequences, respectively, of each clone.

We have sequenced 25 separate ubiquitin coding units. There is a remarkable uniformity in the amino acid sequence among them (Figure 2). The deduced amino acid sequences of all the ubiquitins were identical, except for a substitution of threonine for alanine in position 28 of repeats 2, 3, and 5 of clone DCUB19 [which is identical in sequence with the original cDNA pLK229 and a genomic clone, p229, isolated previously (Giorda & Ennis, 1987)], and an isoleucine (ATT) for valine (GTT) in position 5 of the DCUBV20 repeat.

By contrast, there were a reasonable number of nucleotide differences among the repeats (Figure 2). Except for DCUB17 and DCUB18 which have the identical coding region base sequences (see below for a discussion of this finding), no other ubiquitin codons possessed a completely identical sequence. Up to 33 base differences were observed (DCUB19 repeat 2) and as few as 13 (DCUB18 repeat 2) out of a total of 228 bases in the ubiquitin monomer.

In a recent publication (Sharp & Li, 1987), the molecular evolution of ubiquitin genes was probed, and some interesting conclusions were derived concerning the then few published sequences we had previously reported. Among the most provocative of them was the suggestion that the different serine codons found at position 20 in *Dictyostelium* ubiquitin, which differ in both position 1 and position 2 of the codons (e.g., TCT vs AGT), probably are derived one from the other by two simultaneous nucleotide substitutions. This double mutation should be a very rare event. The direction of change is inferred to be TCT → AGT because TCT is found at all other loci. In addition, there is a G → A transition in the first nucleotide of repeats 2, 3, and 5 at position 28, which results in the substitution of threonine for alanine. This is the only observed within-species ubiquitin amino acid change which arises in this manner. With the large amount of sequence data in the present paper, a more comprehensive study of the evolution of these genes can be accomplished.

Figures 3 and 4 give the 5′ and 3′ noncoding sequences of the cDNA clones. All the clones except DCUB2 and DCUBV20 possessed a 5′ noncoding region, and all had a 3′ region. It is noteworthy that clones DCUB17 and -18 had identical 5′ noncoding regions and 29 identical 3′ nucleotides (up to arrows in Figure 4). The sequences beyond this point were different. As was pointed out above, the nucleotide sequences of no other ubiquitin coding regions were identical (Figure 2). Consequently, this presents a dilemma to us. We are not certain whether these two cDNAs represent two different genes, two different transcripts of the same gene, or an artifact of the cDNA library preparation. We should be able to clarify this point when we isolate all the ubiquitin genes and more cDNAs representing these transcripts. Since DCUB2 and DCUBV20 did not have a 5′ noncoding region, we cannot unequivocally state that there are no more ubiquitin repeats on these isolates. However, DCUBV20 is similar in sequence with clones already described which have only one ubiquitin repeat (Lund et al., 1985; Özkaynak et al., 1984). Therefore, we feel that DCUBV20 possesses only one repeat.

We also isolated two cDNAs which represented a monoubiquitin sequence fused at its 3′ terminus to different peptides. One of these (our clone DCUBV45) is identical with that already described (Westphal et al., 1986; Müller-Taubenberger

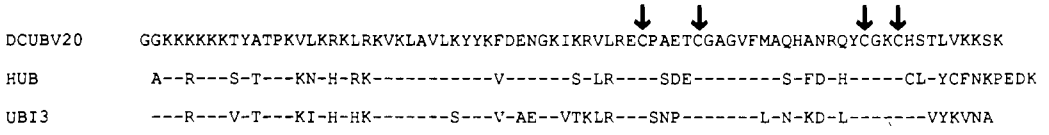

FIGURE 5: Comparison of the deduced amino acid sequence of the 78 amino acid extension of DCUBV20 with the 80 and 76 amino acid extensions of the corresponding human HUB (Lund et al., 1985) and yeast UBI3 (Özkaynak et al., 1984) proteins. The arrows indicate the conserved cysteines.

et al., 1988a) and has a 52 amino acid extension to the ubiquitin moiety. It is very similar in amino acid sequence to the corresponding yeast UBI1/UBI2 (Özkaynak et al., 1984) and mouse arf2 amino acid sequences (St. John et al., 1986).

The sequence of the other ubiquitin fusion protein (DCUBV20) is given in Figures 2 and 5. This protein possessed a 78 amino acid extension unrelated to the N-terminal monoubiquitin. It is very similar to the respective yeast UBI3 (Özkaynak et al., 1984) and a human protein (HUB in Figure 5) (Lund et al., 1985) whose amino acid sequences are compared in Figure 5. The DCUBV20 cDNA codes for a deduced protein containing 154 amino acids, with a molecular weight of 17 467. The sequence of the 76 N-terminal amino acids is identical with *Dictyostelium* ubiquitin (Figure 2), except, as already mentioned, for the substitution of an isoleucine for a valine in position 5. The 78 amino acid extension is very basic, containing 28% lysine/arginine (Figure 5). This is similar to UBI3 which contains 29% lysine/arginine in its 76 amino acid extension, and the human protein which contains 28% of these amino acids in its 80 amino acid extension.

The amino acid sequences of the protein extensions are conserved between the yeast, human, and *Dictyostelium* counterparts. For example, the 52 amino acid extension of DCUBV45 is identical with UBI1/2 except for 11 amino acids (79% identity) (Westphal et al., 1986; Müller-Taubenberger et al., 1988a; Özkaynak et al., 1984a). The 78 amino acid DCUBV20 extension shows 59% identity in amino acid sequence with yeast UBI3 (Özkaynak et al., 1984) and 61% with the human protein (HUB) (Lund et al., 1985) with no gaps introduced to align the sequences (Figure 5).

Another common feature of all the protein extensions is the presence of a common structural motif involving cysteine residues, suggesting a metal binding domain that matches the consensus sequence of the type Cys-X_{2-4}-Cys-X_{2-15}-Cys-X_{2-4}-Cys found in *Xenopus* transcription factor TFIIIA (Miller et al., 1985; Vincent, 1986). As shown in Figure 5, there is a remarkable similarity in amino acid sequence among all three proteins in the region containing the cysteines (marked by arrows). Sixteen of the 24 amino acids are identical in DCUBV20 and UBI3 and 17 between DCUBV20 and HUB. In addition, the position of a histidine residue is also conserved, and the histidine after the last cysteine in DCUBV20 is the same in UBI3 and is cysteine in HUB. The motif in UBI1/2 is identical with DCUBV45 (Westphal et al., 1986; Müller-Taubenberger et al., 1988a; Özkaynak et al., 1984). A search of other proteins for similar sequences has shown that this motif is present in many nucleic acid binding proteins (Vincent, 1986; Berg, 1986). Consequently, it is possible that, in conjunction with the basic amino acids which are in high amount, these proteins interact with nucleic acids and function in this capacity. In view of the similarity among these proteins, it is not unreasonable to infer that they have similar functions. Despite the similarity of the cysteine-containing regions among a large number of proteins, the *Dictyostelium* extension sequences show no other similarity to proteins in the data base. It has actually been shown that the fusion protein is cleaved to give ubiquitin plus the extensions (Redman & Rechsteiner, 1988).

Two unique features of the *Dictyostelium* cDNA clones are worthy of mention. A common feature of the ubiquitin repeats (Baker & Board, 1987) is the presence of a non-ubiquitin C-terminal amino acid preceding the stop codon. Asparagine (yeast), lysine (barley), tyrosine (chicken), and valine and cysteine (human) have been observed. Two of the *Dictyostelium* cDNA clones end in asparagine and three in leucine, the latter which has not been previously noted (Figure 1). Exceptions to this are *Xenopus* which contains no extra amino acid at the C-terminus (Dworkin-Rastl et al., 1984), *Arabidopsis* which contains two (Burke et al., 1988), and *Drosophila* which contains three (Lee et al., 1988).

One other common feature for all previously described polyubiquitin cDNAs is their incomplete 5′ region (Baker & Board, 1987). None extends beyond the ATG codon. This phenomenon has been ascribed to the presence of inverted repeats in the coding portion of the ubiquitin cDNAs which may form snap-back loop structures which are self-primed during cDNA synthesis (Baker & Board, 1987). Alternatively, the inverted repeats may form secondary structures that interfere with cDNA synthesis. Five of the *D. discoideum* cDNAs extended past the ATG codon. Perhaps this indicates that the structure of the coding region of *Dictyostelium* mRNAs does not contain inverted repeats, which seem to be prevalent in other organisms.

As has been found for humans (Wiborg et al., 1985), *Xenopus* (Dworkin-Rastl et al., 1984), *Drosophila* (Arribas et al., 1986), plants (Vierstra et al., 1986), and *Saccharomyces cerevisiae* (Özkaynak et al., 1984), the *D. discoideum* ubiquitin genes are organized as tandem repeats of the 76 amino acid protein, and they seem to be part of a multigene family (Giorda & Ennis, 1987; Westphal et al., 1986; Müller-Taubenberger et al., 1988a). What is very intriguing is the observation that ubiquitin mRNA species are differentially expressed during *D. discoideum* development (Giorda & Ennis, 1987; Westphal et al., 1986; Müller-Taubenberger et al., 1988a). The fact that ubiquitin mRNA is developmentally regulated (and by inference the synthesis of the protein) argues for a function for this protein in development. The recent observation that *Dictyostelium* ubiquitin is a heat-shock protein extends the possible requirement for the protein for other important cell functions (Müller-Taubenberger et al., 1988b). The characterization of ubiquitin cDNAs and genomic sequences in *D. discoideum* now makes the understanding of its developmental regulation feasible.

REFERENCES

Arribas, C., Sampedro, J., & Izquierdo, M. (1986) *Biochim. Biophys. Acta 868*, 119–127.

Baker, R. T., & Board, P. G. (1987) *Nucleic Acids Res. 15*, 443–463.

Berg, J. M. (1986) *Science 232*, 485–487.

Burke, T. J., Callis, J., & Vierstra, R. D. (1988) *Mol. Gen. Genet. 213*, 435–443.

Dowbenko, D. J., & Ennis, H. L. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1791–1795.

Dworkin-Rastl, E., Shrutkowski, A., & Dworkin, M. B. (1984) *Cell 39*, 321–325.

Ennis, H. L., & Sussman, M. (1975) *J. Bacteriol. 124*, 62–64.
Giorda, R., & Ennis, H. L. (1987) *Mol. Cell. Biol. 6*, 2097–2103.
Henikoff, S. (1984) *Gene 28*, 351–359.
Kelly, L. J., Kelly, R., & Ennis, H. L. (1983) *Mol. Cell. Biol. 3*, 1943–1948.
Lee, H., Simon, J. A., & Lis, J. T. (1988) *Mol. Cell. Biol. 8*, 4727–4735.
Lund, P. K., Moats-Staats, B. M., Simmons, J. G., Hoyt, E., D'Ercole, A. J., Martin, F., & Van Wyk, J. J. (1985) *J. Biol. Chem. 260*, 7609–7613.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Miller, J., McLachlan, A. D., & Klug, A. (1985) *EMBO J. 4*, 1609–1614.
Müller-Taubenberger, A., Westphal, A., Jaeger, E., Noegel, A., & Gerisch, G. (1988a) *FEBS Lett. 229*, 273–278.
Müller-Taubenberger, A., Hagmann, J., Noegel, A., & Gerisch, G. (1988b) *J. Cell. Sci. 90*, 51–58.
Özkaynak, E., Finley, D., & Varshavsky, A. (1984) *Nature (London) 312*, 663–666.
Redman, K. L., & Rechsteiner, M. (1988) *J. Biol. Chem. 263*, 4926–4931.
Sanger, F., Nicklen, N., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schlesinger, D. H., Goldstein, G., & Niall, H. D. (1975) *Biochemistry 14*, 2214–2218.
Sharp, P. M. & Li, W.-H. (1987) *Trends Ecol. Evol. 2*, 328–332.
St. John, T., Gallatin, W. M., Siegelman, M., Smith, H. T., Fried, V. A., & Weissman, I. L. (1986) *Science 231*, 845–850.
Sussman, M., & Brackenbury, R. (1976) *Annu. Rev. Plant Physiol. 27*, 229–265.
Sussman, R., & Sussman, M. (1967) *Biochem. Biophys. Res. Commun. 29*, 53–55.
Vierstra, R. D., Langan, S. M., & Schaller, G. E. (1986) *Biochemistry 25*, 3105–3108.
Vincent, A. (1986) *Nucleic Acids Res. 14*, 4385–4391.
Westphal, M., Müller-Taubenberger, A., Noegel, A., & Gerisch, G. (1986) *FEBS Lett. 209*, 92–96.
Wiborg, O., Pedersen, M. S., Wird, A., Berglund, L. E., Marker, K. A., & Vuust, J. (1985) *EMBO J. 4*, 755–759.

Binding of Actinomycin D to $[d(ATCGAT)]_2$: NMR Evidence of Multiple Complexes†

Ning Zhou,‡ Thomas L. James, and Richard H. Shafer*
Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143
Received November 3, 1988; Revised Manuscript Received March 3, 1989

ABSTRACT: Actinomycin D (actD) binds to the oligonucleotide $[d(ATCGAT)]_2$ with a hypochromatic and red-shifted visible absorbance band compared to free drug and a CD spectrum with double negative bands at 460 and 385 nm. These spectral features are similar to those of the actD–$[d(ATGCAT)]_2$ complex, while actD–$[d(AT)_5]_2$ gives spectra similar to those of free drug. Upon dilution or raising the temperature, the spectral characteristics accompanying complex formation disappear in the actD–$[(ATCGAT)]_2$ sample but remain in the actD–$[d(ATGCAT)]_2$ complex under the same experimental conditions. These results suggest that (a) sequence-specific binding of actD occurs with $[d(ATCGAT)]_2$ but not with $[d(AT)_5]_2$, (b) the binding is not as strong as with $[d(ATGCAT)]_2$, and (c) actD binds $[d(ATCGAT)]_2$ with the same mechanism as it binds $[d(ATGCAT)]_2$, i.e., by intercalation. From NMR spectra of the actD–$[d(ATCGAT)]_2$ complex, three types of signals can be detected below 20 °C, one major and two minor ones. At higher temperatures, exchange between the two minor ones becomes fast enough that only one type of minor signal was seen. Partial resonance assignments were made by using 2D nuclear Overhauser effect (NOE) and 2D homonuclear Hartmann–Hahn (HOHAHA) experiments. Proton chemical shift changes of the major complex are consistent with actD chromophore ring intercalation between hexamer base pairs. Data from NOE-detected dipolar interactions between actD and $[d(ATCGAT)]_2$ protons were interpreted in terms of a major complex with the actD chromophore ring system intercalated at the CG position and minor complexes with the drug intercalated off center at the GA positions. While the centrally intercalated complex still exists at a 1:2 $[d(ATCGAT)]_2$:actD ratio, there is also evidence for a complex consisting of two bound actD molecules per duplex.

Actinomycin D (actD) (Figure 1) is currently used as an antitumor drug in routine clinical treatment of a limited number of cancers. Its binding to DNA has been extensively studied [see review by Waring (1981)]. A widely accepted model of the actD–DNA complex is one with the drug chromophore intercalated between 5′-GpC-3′ and its peptide side chains lying in the minor groove. This model complex is stabilized by a strong hydrogen bond between G NH_2 and Thr C=O, a weak hydrogen bond between G N3 and Thr αH, and stacking forces between the actD chromophore ring and the G base ring, as well as by numerous hydrophobic interactions between the drug peptide chains and DNA minor groove

† This work was supported by USPHS Grant CA 27343 awarded by the National Cancer Institute, DHHS.
* Address correspondence to this author.
‡ Present address: Department of Biological Sciences, University of Calgary, AB T2N IN4, Canada.

0006-2960/89/0428-5231$01.50/0