J. Biochem. 124, 35-39 (1998)

Cloning of Human Polyubiquitin cDNAs and a Ubiquitin-Binding Assay Involving Its *In Vitro* Translation Product[1]

Nam-Soon Kim,* Tomoko Yamaguchi,* Shingo Sekine,* Mihoro Saeki,† Shawichi Iwamuro,† and Seishi Kato*,†,[2]

*Sagami Chemical Research Center, 4-4-1 Nishi-Ohnuma, Sagamihara, Kanagawa 229-0012; and †Kato Cytoprotein Network Project, ERATO, JST, Sagamihara, Kanagawa 229-0012

Received for publication, December 3, 1997

During large-scale *in vitro* translation analysis of a human full-length cDNA bank, we found a clone producing a remarkably smaller translation product than that expected from the open reading frame. The cDNA encodes a polyubiquitin, UbC, composed of nine tandem repeats of the ubiquitin unit. The bank contained twelve UbC cDNAs including four full-length ones. Sequencing analysis of these clones showed that UbC cDNAs can be classified into two types, UbC1 and UbC2, in each of which there are six polymorphic nucleotide variations. The present UbC cDNA was *in vitro* translated in a rabbit reticulocyte or wheat germ extract to produce a free ubiquitin labeled with [^{35}S]methionine. The labeled ubiquitin could be used as a substrate for thiol ester formation with ubiquitin-activating enzyme E1 or ubiquitin-conjugating enzyme E2.

Key words: cDNA, human, *in vitro* translation, polymorphism, ubiquitin.

We have been performing large-scale *in vitro* translation of human full-length cDNAs registered in the Homo·Protein cDNA Bank (*1*). During this process, we have found many clones showing unexpected mobilities on an SDS-PAGE gel. This paper describes one of these cDNAs, of which the *in vitro* translation product has a smaller molecular mass than that expected from the size of the open reading frame of the cDNA. Full-sequencing of this cDNA revealed that it encodes a polyubiquitin composed of nine tandem repeats of a ubiquitin unit.

Ubiquitin (Ub) is used in the cell as a labeled molecule that is covalently bound to a target protein to be degraded by the proteasome system (*2*). Ub is a 76-amino acid protein encoded by genes consisting of multiple Ub repeats or by genes fused to ribosomal protein genes. To date, four human Ub genes have been identified by Northern analysis (*3*) and genomic and cDNA cloning: UbA52 and UbA80 each coding for a single Ub fused to ribosomal proteins, UbB coding for three tandem repeats of Ub, and UbC coding for nine tandem repeats of Ub. The human genomic sequences have been reported for UbA52 (*4*), UbB (*5*), and UbC (*3, 6*). Human cDNAs for UbA52 (*4*) and UbA80 (*7*), and truncated cDNAs for UbC (*8*) have been isolated, but there has been no report on human full-length cDNAs for UbB and UbC.

These genes comprising multiple Ub repeats express a Ub precursor that is processed by C-terminal hydrolase to produce a free Ub. The free Ub enters the Ub pathway in which, first, Ub is activated by enzyme E1 to yield an E1-Ub thiol ester (*2, 9*). The activated Ub is then transferred from E1 to a specific thiol of E2, an E2-Ub thiol ester being formed. Finally, the Ub of E2-Ub is donated to a target protein either alone or together with an E3 protein. The ubiquitinated protein is degraded by 26S proteasomes in an energy-dependent manner.

In order to investigate the Ub pathway, a labeled Ub is useful as a substrate. ^{125}I-labeled Ub was used to detect the Ub thiol ester formation by E1 or E2s in a rabbit reticulocyte lysate (*10*). But iodination of tyrosine 59 of Ub was reported to abolish the acceptor activity of Ub in diubiquitin synthesis (*11*). Recently, a recombinant Ub fused to a phosphorylation tag was produced in *Escherichia coli* cells and then phosphorylated *in vitro* with cAMP-dependent protein kinase to prepare a ^{32}P-labeled Ub (*12, 13*). This ^{32}P-labeled Ub was shown to have the ability to form a thiol ester with E1 or E2s. But conjugation of epitope-tagged Ub was reported to inhibit the proteolysis of a ubiquitinated target protein by proteasomes (*14*).

In this study, we show that human UbC has two polymorphic variants. The reduction in size of the *in vitro* translation product of the Ub cDNA resulted from rapid processing in the rabbit reticulocyte lysate or wheat germ extract. Furthermore, the ^{35}S-labeled free Ub produced on *in vitro* translation was applied to a Ub-binding assay for thiol ester formation by E1 or E2s.

MATERIALS AND METHODS

Materials—Restriction enzymes and modification enzymes were purchased from Takara Shuzo. The radioiso-

[1] This work was partly supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture of Japan. The nucleotide sequences reported in this paper have been submitted to the DDBJ/GenBank™/EMBL data banks under accession number AB009010.

[2] To whom correspondence should be addressed. Phone: +81-427-42-5091, Fax: +81-427-42-5091, E-mail: seishi@sagami.or.jp

Abbreviation: Ub, ubiquitin.

tope-labeled compounds, [^{35}S]methionine and [^{14}C]methylated protein molecular weight markers, were purchased from Amersham.

cDNA Libraries—The human fibrosarcoma cell line HT-1080 and histiocytic lymphoma cell line U937 cDNA libraries were described in the previous paper (*1*). Human gastric adenocarcinoma and liver cDNA libraries were prepared using a vector primer, pKA1, carrying a T7 RNA polymerase promoter according to the DNA-RNA chimeric oligonucleotide capping method (*1*).

Sequencing—The sequencing reaction was performed by the dideoxy method using a dye primer cycle sequencing kit (Applied Biosystems). The reaction mixture was analyzed with a model 377 automated DNA sequencer (Applied Biosystems). The entire sequence of the cDNA was determined by sequencing of overlapping deleted clones prepared using a deletion kit (Takara Shuzo). The pKA1 vector has another sequencing primer site followed by *Pst*I and *Sph*I in the upstream region from the cDNA cloning site (*1*). The pKA1-polyubiquitin cDNA vector was digested with *Eco*RI and *Sph*I, and then the cDNA was deleted from the *Eco*RI site using exonuclease III. Deleted clones for antisense strand sequencing were prepared using the vector in which the cDNA insert was subcloned reversely into the pKA1.

In Vitro Translation—*In vitro* translation was carried out in a reaction mixture containing [^{35}S]methionine using a T_NT coupled reticulocyte lysate system (Promega) or a T_NT coupled wheat germ extract system (Promega) according to the manufacturer's instructions. The translated product was analyzed by 10-20% SDS-PAGE.

Preparation of ^{35}S-Labeled Ub—The *in vitro* translation of polyubiquitin cDNA was performed at room temperature for 90 min in a 100 μl reaction mixture comprising 10 μg plasmid, 8 μl [^{35}S]methionine (370 MBq/ml), and 50 μl rabbit reticulocyte lysate. The reaction product was applied to a Sephadex G-50 column with a bed volume of 18 ml. Elution was carried out with a column buffer comprising 20 mM Tris-HCl (pH 8.0), 10 mM NaCl, and 1 mM EDTA. The fractions giving a Ub band on an SDS-PAGE gel were collected and concentrated with Microcon-3 (Amicon).

Preparation of Mouse E1 and Human Ubc E2—The $(His)_6$-tagged mouse recombinant E1 was a gift from Dr. Yamao (National Institute of Genetics). The preparation of human recombinant UbcH-ben was described in the previous paper (*12*).

Ub-Binding Assay—^{35}S-labeled Ub was mixed with $(His)_6$-tagged mouse recombinant E1 and/or human recombinant UbcH-ben in a reaction mixture comprising 50 mM Tris (pH 7.5), 10 mM $MgCl_2$, 1 mM ATP, and 0.1 mM DTT for 20 min at room temperature. The reaction was stopped by the addition of the SDS gel sample buffer from which DTT had been omitted. A replicate sample was boiled for 1 min in the presence of 0.1 M DTT. All samples were analyzed by SDS-PAGE and ^{35}S-labeled Ub-containing bands were visualized by autoradiography.

RESULTS

Identification of Polyubiquitin cDNAs—The Homo·Protein cDNA bank contained twelve cDNAs including four full-length ones encoding putative polyubiquitin UbC. One full-length clone, CL1, comprises a total 2,280 bp, including a 5′-untranslated region of 68 bp, an open reading frame of 2,058 bp, a 3′-untranslated region of 66 bp, and a poly(A) tail of 88 bp (Fig. 1). The open reading frame encodes a protein of 685 amino acid residues, which comprise nine tandem repeats of Ub. The amino acid sequence of each repeat is identical except that the ninth repeat has an additional amino acid residue at its C-terminal. The nucleotide sequence of each repeat, however, slightly differs from those of the others.

The 5′-regions of the remaining clones were sequenced and the obtained nucleotide sequences of about 600 bp were compared with each other (Fig. 2). Clone CL2 has an identical sequence to clone CL1 in this region, but the two other full-length clones, CL3 and CL4, have three variations. Full-sequencing of clone CL3 showed that it has six nucleotide variations, including the above three, compared with CL1. The remaining truncated cDNAs belong to either type, because the sequenced region contained at least one variation site. It should be noted that each cDNA was isolated from various cDNA libraries prepared from different cells. These results suggest that UbC has two polymorphic variants, named UbC1 for clone CL1 and UbC2 for clone CL3.

In Vitro Translation of Polyubiquitin cDNA—Clone CL1 was *in vitro* translated in a rabbit reticulocyte lysate or wheat germ extract. In both systems, the size of the translation product was <10 kDa, which is smaller than the size expected, 77 kDa, from the polyubiquitin encoded by the cDNA. With short exposure on autoradiography, only the <10 kDa band was observed, but with long exposure many ladder and smear bands of high molecular weight appeared. Figure 3A shows the time course of *in vitro* translation in the rabbit reticulocyte lysate. In this case, the Ub band appeared after 20 min and the amount of the translated Ub gradually increased. Even at an early stage we could not observe the polyubiquitin band, suggesting that the processing of polyubiquitin rapidly occurred just after translation. This figure also shows that many ladder bands of less than 40 kDa were produced in one hour and then smear bands of high molecular weight increased. On the other hand, the wheat germ system produced ladder bands corresponding to multimers of Ub, especially a relatively strong band corresponding to a Ub dimer of 18 kDa (Fig. 3B).

Conjugation between E1 and In Vitro Translated Ub—*In vitro* translation experiments suggested that *in vitro* translated and processed Ub molecules may bind to various proteins in a lysate *via* a Ub-binding pathway to yield ladder bands and a high molecular weight smear band. This led us to use the *in vitro* translated Ub as a substrate for an *in vitro* Ub-binding assay. To examine the thiol ester formation ability of the *in vitro* translated Ub, the ^{35}S-labeled Ub partially purified by gel filtration was reacted with a mouse Ub-activating enzyme, E1, produced in *E. coli* cells. When the reaction mixture was separated on an SDS-PAGE gel, a band corresponding to an E1-Ub conjugate was observed under non-reducing conditions, but not under reducing conditions (Fig. 4).

The Transfer of Ub from E1 to E2—The Homo·Protein cDNA bank contains several Ub-conjugating enzyme E2s. To examine the transfer of Ub from E1 to E2, UbcH-ben, that is a human E2 enzyme recently identified by us (*12*), was added to the above system. By adding the UbcH-ben

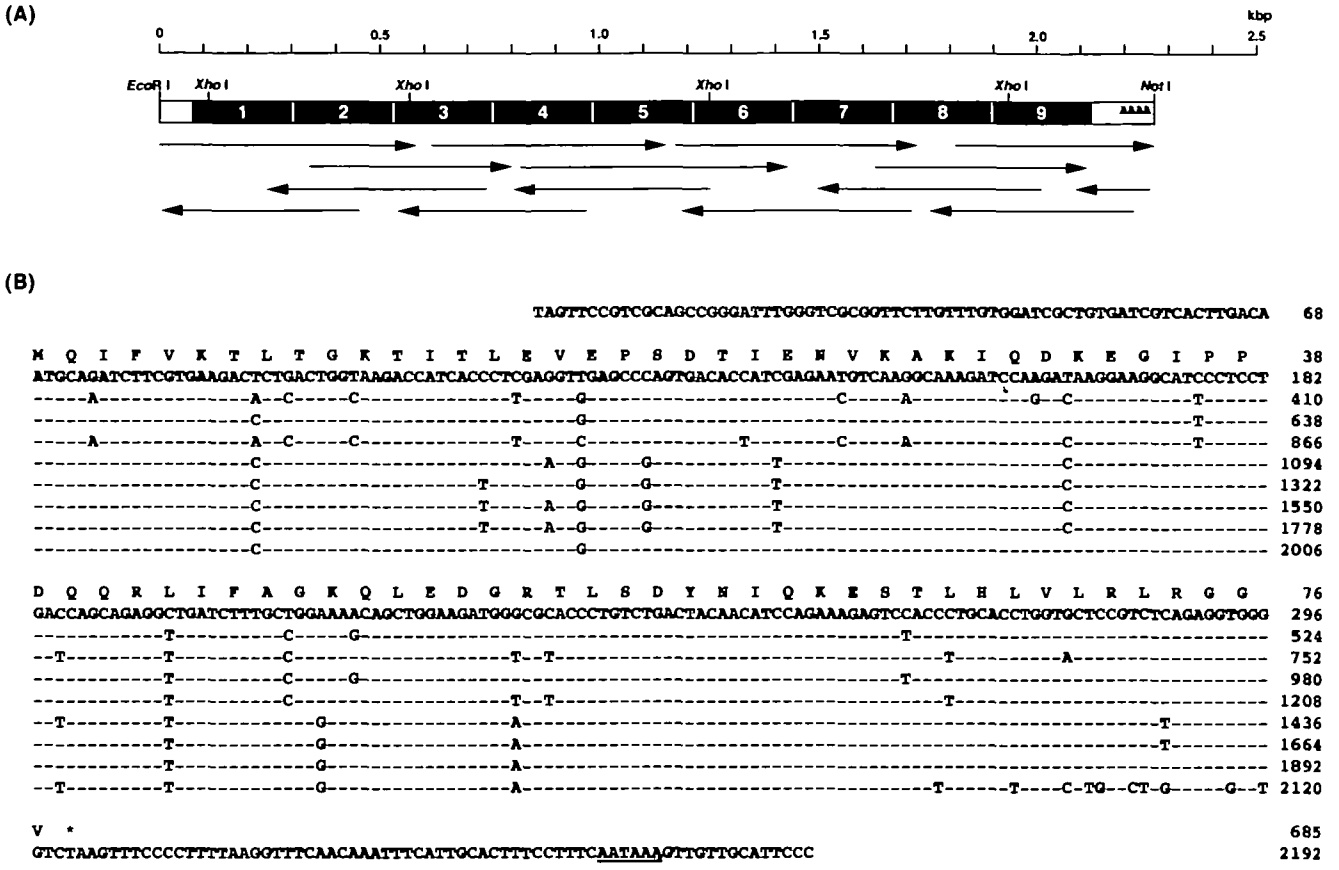

Fig. 1. **Structure of the polyubiquitin cDNA.** (A) Restriction map of clone CL1 and the sequencing strategy. Each shaded box represents an open reading frame. The number in each open reading frame is that of the unit of Ub. (B) The nucleotide sequence of clone CL1 and the deduced amino acid sequence. The putative poly (A) addition signal is double-underlined. Since each of the nine tandem repeats encodes an identical amino acid sequence of Ub, only different nucleotides are shown.

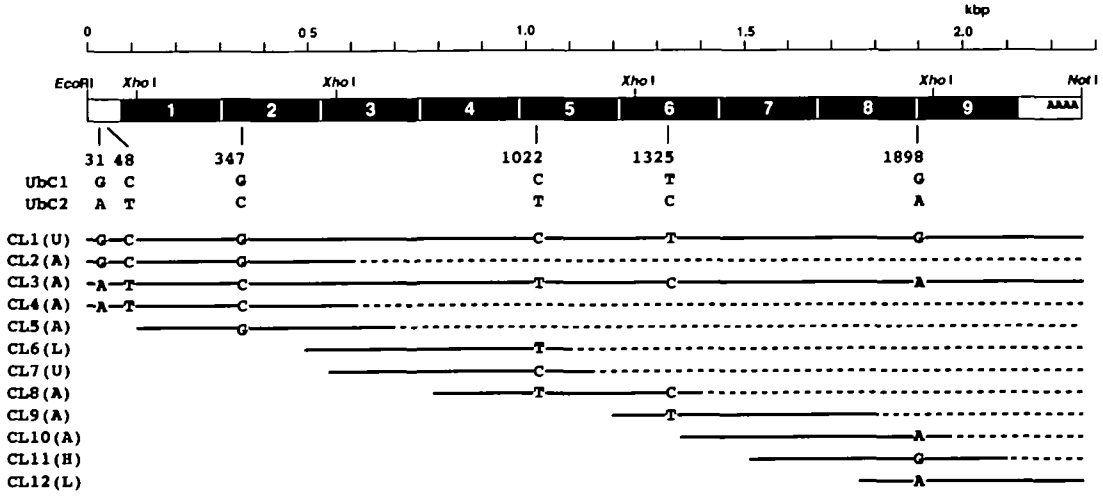

Fig. 2. **Comparison of two polymorphic variants.** Clones CL1 to CL12 were partially or fully sequenced. The solid lines represent the sequenced parts. Only polymorphic nucleotides of the sequence-determined region are shown. The letters in parentheses denote the tissues or cell lines from which the cDNA libraries were prepared: U, U937; A, gastric adenocarcinoma; L, liver; H, HT-1080.

produced in *E. coli* cells to the reaction mixture containing the partially purified ^{35}S-labeled Ub and the mouse recombinant E1, we could observe a band of 26 kDa, which corresponded to the size of an E2-Ub conjugate (Fig. 5A). In this reaction, the disappearance of the E1-Ub band suggests that Ub was transferred rapidly from E1 to E2. Unexpect-

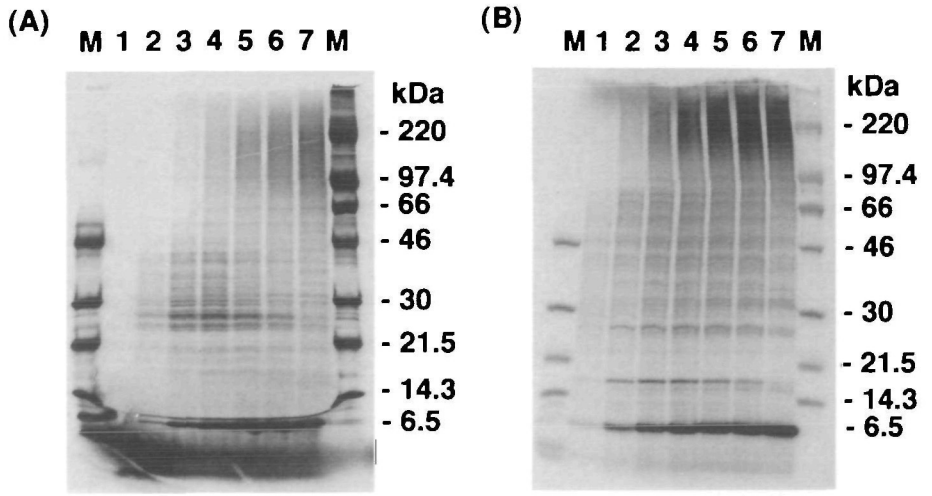

Fig. 3. **SDS-PAGE analysis of *in vitro* translated polyubiquitin cDNA.** (A) *In vitro* translation in a rabbit reticulocyte lysate. The *in vitro* translation was carried out at room temperature in a 50 μl reaction solution containing 4 μg polyubiquitin cDNA plasmid, 4 μl [^{35}S]methionine (370 MBq/ml), and 25 μl T_NT coupled reticulocyte lysate. An aliquot of 3 μl was taken from the reaction mixture after 10 min (lane 1), 20 min (lane 2), 30 min (lane 3), 40 min (lane 4), 50 min (lane 5), 60 min (lane 6), and 90 min (lane 7), respectively, and then analyzed by 10–20% SDS-PAGE. (B) *In vitro* translation in a wheat germ extract. 25 μl wheat germ extract was used. The other procedures were the same as for the rabbit reticulocyte system.

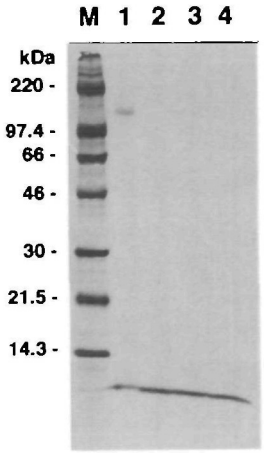

Fig. 4. **Conjugation between E1 and *in vitro* translated Ub.** Partially purified *in vitro* translated Ub (lanes 2 and 4) was reacted with 8 μg mouse recombinant E1 (lanes 1 and 3). The reaction product was analyzed by SDS-PAGE under non-reducing conditions (lanes 1 and 2) or reducing conditions (lanes 3 and 4).

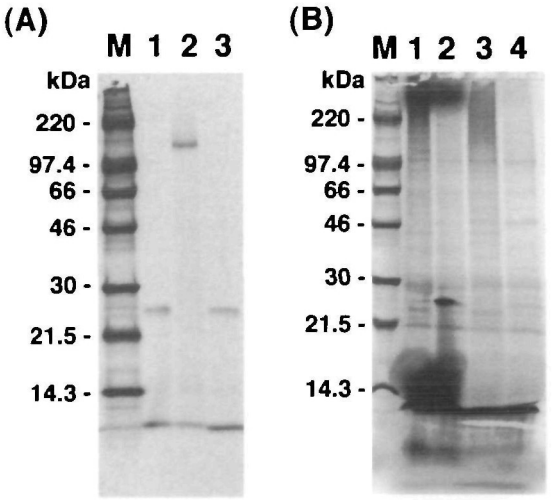

Fig. 5. **Transfer of Ub from E1 to E2.** (A) Partially purified *in vitro* translated Ub was reacted with both 8 μg mouse recombinant E1 and 3.4 μg human recombinant E2 (lane 1), with only E1 (lane 2), or with only E2 (lane 3). SDS-PAGE was carried out under non-reducing conditions. (B) The *in vitro* translation reaction mixture of polyubiquitin cDNA (lanes 1 and 3) was reacted with human recombinant E2 (lanes 2 and 4). The reaction product was analyzed by SDS-PAGE under non-reducing conditions (lanes 1 and 2) or reducing conditions (lanes 3 and 4).

edly, the E2-Ub band was observed even without the addition of E1. This reaction may be catalyzed by rabbit E1 contaminating the partially purified Ub sample. In fact, when the recombinant UbcH-ben was added to the *in vitro* translation reaction mixture of polyubiquitin cDNA without the addition of E1, the E2-Ub band was only observed under non-reducing conditions (Fig. 5B).

DISCUSSION

Analysis of the *in vitro* translation products of full-length cDNAs provides us with much information about the proteins encoded by the cDNAs. We routinely performed the *in vitro* translation experiment to confirm that a clone has a full-length cDNA. Most clones gave a band of the size predicted from the open reading frame of the cDNA on an SDS-PAGE gel, but some products showed abnormal mobilities. Analysis of these unusual clones is expected to provide information about the post-translational modification of the protein. In this study, we showed an example, the polyubiquitin UbC cDNA producing a remarkably smaller product than that expected.

The Homo•protein cDNA bank contained twelve UbC cDNAs, including four full-length cDNAs, isolated from four different cDNA libraries. Sequencing analysis showed that these cDNAs can be classified into two polymorphic types, UbC1 and UbC2, in each of which there are six nucleotide variations. The entire sequence reported for truncated polyubiquitin cDNA (*8*) is identical to the region from 1,317 to 2,140 of UbC1. The sequence of UbC1 cDNA is almost identical to the genomic sequence of the *UbC* gene isolated from HeLa S3 cells (*6*), but there are four differences in the nucleotide sequence: A to G at position 1082, C to T at 1088, C to G at 1244, and T to N at 1258. Since the six polymorphic sequences are identical to UbC1, this genomic clone may encode UbC1. Comparison between the genomic sequence and our cDNA ones showed that the first

intron of 811 bp is located between positions 65 and 66 of the cDNA sequence and that there is no intron downstream. The transcription initiation site determined by means of primer extension was reported to be A at position 2 of the UbC1 cDNA (*6*). Two of our full-length cDNA clones, CL1 and CL3, start from T at position 1, CL2 and CL4 starting from A at position 2. The two sites may act as an alternative transcription initiation site. Nenoi *et al.* (*6*) reported the existence of short transcripts lacking the eighth repeat. Our cDNAs including truncated ones have both eighth and ninth repeats.

The observed polymorphism is expected to exist in the human population for the following reasons. (a) The possibility of sequencing error or cloning artifacts is low, because multiple clones for each variant were isolated from different cDNA libraries and the polymorphic sites were checked by double or triple sequencing using deleted clones of the entire regions of both strands. (b) Both variants were cloned from multiple cells including not only cell lines but also tissues isolated on operation. (c) One of the variants, UbC1, is identical to the genomic sequence in the nucleotide sequence of the polymorphic region. (d) EST databases contain partial polyubiquitin cDNA sequences with polymorphic changes corresponding to each variant. It should be noted that four changes in the coding region occur at the same positions as the nucleotide changes among the nine tandem repeats. This may suggest that these polymorphic changes are related to the process of generation of tandem repeats through gene duplication.

The reduction in size of the translation product could be explained by processing by C-terminal hydrolase present in the lysate. This phenomenon was suggested by Agell *et al.* (*15*) in a search for polyubiquitin processing activity. They found that an artificial mRNA encoding chicken diubiquitin produced the free Ub in a rabbit reticulocyte lysate or wheat germ extract system. Our results not only supported their data but also suggested that the following processes, such as the activation of Ub and its transfer to a Ub-conjugating system, can be experimentally observed in the *in vitro* system. The experiments involving recombinant E1 and E2 showed that the free Ub produced in the lysate can be activated by E1 and then transferred to E2. The smear band of high molecular weight may imply the existence of polyubiquitinated products produced *via* their own E1 and E2s present in the lysate.

The processing of polyubiquitin by C-terminal hydrolase was reported to occur rapidly in yeast cells with no intermediates (*16*). We have shown that the processing of *in vitro* translated polyubiquitin also occurred rapidly in a rabbit reticulocyte lysate or wheat germ extract system. These results suggested that polyubiquitin translated in mammalian or plant cells may be rapidly processed and that most free Ub molecules may exist as a monomer in the cells. When the wheat germ extract was used, ladder bands corresponding to Ub multimers, especially the dimer, were observed. However, since the amounts of these ladder bands were so small compared with that of the free Ub monomer, they seem not to have any physiological meaning. Rather, this insufficient processing may result from a decrease in the C-terminal hydrolase activity in this lot or from inappropriate experimental conditions for processing.

The labeled Ub will be useful for studying various processes concerning the Ub pathway. The ^{125}I-labeled Ub and ^{32}P-labeled Ub were successfully applied to E2-Ub conjugation studies. One drawback of these modified Ubs was that they did not act in some processes, probably because of stereochemical hindrance by attached molecules. Since the ^{35}S-labeled Ub in this study has the same structure and properties as the native Ub, it might be an ideal substrate for Ub-conjugating reactions. Another merit of the ^{35}S-labeled Ub is its easy preparation. As shown in the above experiments, it is not necessary to purify the labeled Ub if we only want to determine the Ub-binding ability of the E2 enzyme. Thus, the present system for preparing ^{35}S-labeled Ub is expected to become a useful tool for studying a Ub-conjugating system.

REFERENCES

1. Kato, S., Sekine, S., Oh, S.-W., Kim, N.-S., Umezawa, Y., Abe, N., Yokoyama-Kobayashi, M., and Aoki, T. (1994) Construction of a human full-length cDNA bank. *Gene* **150**, 243-250
2. Hershko, A. and Ciechanover, A. (1992) The ubiquitin system for protein degradation. *Annu. Rev. Biochem.* **61**, 761-807
3. Wiborg, O., Pedersen, M.S., Wind, A., Berglund, L.E., Marcker, K.A., and Vuust, J. (1985) The human ubiquitin multigene family: some genes contain multiple directly repeated ubiquitin coding sequences. *EMBO J.* **4**, 755-759
4. Baker, R.T. and Board, P.G. (1991) The human ubiquitin-52 amino acid fusion protein gene shares several structural features with mammalian ribosomal protein genes. *Nucleic Acids Res.* **19**, 1035-1040
5. Baker, R.T. and Board, P.G. (1987) The human ubiquitin gene family: structure of a gene and pseudogenes from the Ub B superfamily. *Nucleic Acids Res.* **15**, 443-463
6. Nenoi, M., Mita, K., Ichimura, S., Cartwright, I.L., Takahashi, E., Yamauchi, M., and Tsuji, H. (1996) Heterogeneous structure of the polyubiquitin gene *UbC* of HeLa S3 cells. *Gene* **175**, 179-185
7. Lund, P.K., Moats-Staats, B.M., Simmons, J.G., Hoyt, E., D'Ercole, A.J., Martin, F., and Van Wyk, J.J. (1985) Nucleotide sequence analysis of a cDNA encoding human ubiquitin reveals that ubiquitin is synthesized as a precursor. *J. Biol. Chem.* **260**, 7609-7613
8. Einspanier, R., Sharma, H.S., and Scheit, K.H. (1987) Cloning and sequence analysis of a cDNA encoding poly-ubiquitin in human ovarian granulosa cells. *Biochem. Biophys. Res. Commun.* **147**, 581-587
9. Jentsch, S. (1992) The ubiquitin-conjugation system. *Annu. Rev. Genet.* **26**, 179-207
10. Haas, A.L. and Bright, P.M. (1988) The resolution and characterization of putative ubiquitin carrier protein isozymes from rabbit reticulocytes. *J. Biol. Chem.* **263**, 13258-13267
11. Pickart, C.M., Haldeman, M.T., Kasperek, E.M., and Chen, Z. (1992) Iodination of tyrosine 59 of ubiquitin selectively blocks ubiquitin's acceptor activity in diubiquitin synthesis catalyzed by $E2_{25K}$. *J. Biol. Chem.* **267**, 14418-14423
12. Yamaguchi, T., Kim, N.-S., Sekine, S., Seino, H., Osaka, F., Yamao, F., and Kato, S. (1996) Cloning and expression of cDNA encoding a human ubiquitin-conjugating enzyme similar to the *Drosophila bendless* gene product. *J. Biochem.* **120**, 494-497
13. Osaka, F., Seino, H., Seno, T., and Yamao, F. (1997) A ubiquitin-conjugating enzyme in fission yeast that is essential for the onset of anaphase in mitosis. *Mol. Cell. Biol.* **17**, 3388-3397
14. Ellison, M.J. and Hochstrasser, M. (1991) Epitope-tagged ubiquitin. A new probe for analyzing ubiquitin function. *J. Biol. Chem.* **266**, 21150-21157
15. Agell, N., Bond, U., and Schlesinger, M.J. (1988) *In vitro* proteolytic processing of a diubiquitin and a truncated diubiquitin formed from *in vitro*-generated mRNAs. *Proc. Natl. Acad. Sci. USA* **85**, 3693-3697
16. Finley, D., Özkaynak, E., and Varshavsky, A. (1987) The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation, and other stresses. *Cell* **48**, 1035-1046