Metabolism of the Polyubiquitin Degradation Signal: Structure, Mechanism, and Role of Isopeptidase T[†]

Keith D. Wilkinson,*[,‡] Valery L. Tashayev,[‡] Lydia B. O'Connor,[‡] Christopher N. Larsen,[‡] Eileen Kasperek,[§] and Cecile M. Pickart[§]

Department of Biochemistry, 4017 Rollins Research Building, Emory University School of Medicine, Atlanta, Georgia 30322, and Department of Biochemistry, 140 Farber Hall, State University of New York at Buffalo, Buffalo, New York 14214

Received June 20, 1995; Revised Manuscript Received September 5, 1995[⊗]

ABSTRACT: A necessary step in ubiquitin-dependent proteolysis is the addition of a polyubiquitin chain to the target protein. This ubiquitinated protein is degraded by a multisubunit complex known as the 26S proteasome. The polyubiquitin chain is probably not released until a late stage in the proteolysis by the proteasome. It is subsequently disassembled to yield functional ubiquitin monomers. Here we present evidence that a 93 kDa protein, isopeptidase T, has the properties expected for the enzyme which disassembles these branched polyubiquitin chains. Protein and cDNA sequencing revealed that isopeptidase T is a member of the ubiquitin specific protease family (UBP). Isopeptidase T disassembles branched polyubiquitin chains (linked by the G76−K48 isopeptide bond) by a sequential *exo* mechanism, starting at the proximal end of the chain (the proximal ubiquitin contains a free carboxyl-terminus). Isopeptidase T prefers to disassemble chains in which there is an intact and unblocked RGG sequence at the C-terminus of the proximal subunit. Rates of disassembly are reduced when G76 of the proximal ubiquitin is modified, for example, by ligation to substrate protein, by esterification, by replacement of the proximal glycine with alanine (G76A), or by truncation. Linear proubiquitin is only a poor substrate. Observed rates and specificity are consistent with isopeptidase T playing a major role in disassembly of polyubiquitin chains. The high discrimination against chains that are blocked or modified at the proximal end indicates that the enzyme acts after release of the chains from conjugated proteins or degradation intermediates. Thus, the proteolytic degradation signal is not disassembled by isopeptidase T before the ubiquitinated protein is degraded. These (and earlier) results suggest that UBP isozymes may exhibit significant substrate specificity, consistent with a role in the regulated catabolism of the polymeric ubiquitin, including the polyubiquitin protein degradation signal.

Ubiquitin-dependent proteolysis is a complex pathway of protein metabolism which has been implicated in a host of cellular functions, including maintenance of chromatin structure, ribosome biogenesis, cell cycle regulation, receptor function, oncoprotein degradation, the stress response, and degradation of abnormal proteins (Ciechanover & Schwartz, 1994; Jentsch, 1992a; Hershko & Ciechanover, 1992). Well over 50 enzymes are known to be involved in this pathway, and most are highly conserved between species. The molecular mechanism of ubiquitin action requires its covalent attachment to proteins, including other molecules of ubiquitin. This post-translational ubiquitination[1] is dependent on the formation of an isopeptide bond between the C-terminal glycine (G76) of ubiquitin and side chain amino groups of other ubiquitins or other proteins (Hershko & Ciechanover, 1992). The isopeptide bond is formed by the sequential action of three enzymes: E1, the ubiquitin-activating enzyme; E2, one of a family of ubiquitin-conjugating enzymes; and E3, one of a family of ubiquitin protein ligases. Polyubiquitinated proteins are degraded by the 26S protease (Hershko & Ciechanover, 1992). The specificity of this multienzyme complex toward ubiquitinated substrates derives in part from the presence of a 50 kDa subunit (S5a) which has high affinity for branched polyubiquitin chains linked by a G76−K48 isopeptide bond (Deveraux *et al.,* 1994).

Monoubiquitinated protein is converted to a polyubiquitinated protein by the condensation of additional ubiquitin

[†] This work was supported by grants from the National Institutes of Health, GM30308 to K.D.W. and DK02026 (RCDA) and DK46984 to C.M.P. C.N.L. received support from GM08367 (BCMB Training Grant).

* Address correspondence to Keith D. Wilkinson, Ph.D., Professor, Department of Biochemistry, 4017 Rollins Research Building, Emory University School of Medicine, Atlanta, GA 30322. Phone: (404) 727-5980. Fax: (404) 727-3452. E-mail: genekdw@bimcore.emory.edu.

[‡] Emory University School of Medicine.

[§] State University of New York at Buffalo.

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1995.

[1] The nomenclature used here is as follows. Proubiquitin is the polymeric ubiquitin product of the UBI4 gene in yeast and its homologue in other eukaryotes, while ubiquitin-CEP's are the products of the UBI1, UBI2, and UBI3 genes in yeast and their homologues in other eukaryotes (Ozkaynak *et al.,* 1987). Processing is the hydrolysis of the peptide bond at glycine 76 of ubiquitin to yield monomeric ubiquitin from these "linear" ubiquitin gene products. Ubiquitination refers to the formation of isopeptide bonds between the C-terminus of ubiquitin and ϵ-amino side chains of proteins. Multiubiquitination refers to the addition of single ubiquitins to several lysines on a protein. Polyubiquitination refers to the addition of several ubiquitins to a single lysine on a protein. Polyubiquitin chains are linked by isopeptide bonds between the C-terminus of ubiquitin and a lysine (usually K48) on another ubiquitin. They are also referred to as "branched" chains, since all of the ubiquitin amino-termini are free but only one C-terminus is free. The subunit containing this free C-terminus is the proximal subunit (*i.e.,* the site of attachment of the chain to target proteins), and the other end of the chain is referred to as the distal end. Disassembly of polyubiquitin chains occurs by hydrolysis of the isopeptide bond at glycine 76 of ubiquitin.

0006-2960/95/0434-14535$09.00/0

molecules. In many cases, these linkages are between K48 of the ubiquitin at the terminus of the growing chain and the C-terminal G76 of the incoming ubiquitin (Gregori *et al.*, 1990). Target proteins which do not bear polyubiquitin chains can also be degraded (Haas *et al.*, 1990). In at least one case, there is a degradative rate enhancement of 10-fold or more when a single ubiquitin is replaced by a K48-linked polyubiquitin chain (Gregori *et al.*, 1990; Chau *et al.*, 1989). Under certain circumstances, polyubiquitin chains can also be linked by isopeptide bonds between G76 and either K63 (Arnason & Ellison, 1994; Spence *et al.*, 1995), K29 (Arnason & Ellison, 1994), or K11 (A. L. Haas, private communication). The significance of these linkages is not yet known.

There are multiple E2 and E3 isozymes, and there are several enzymatic routes to polymeric ubiquitin conjugates (Jentsch, 1992b; Ciechanover & Schwartz, 1994). For instance, some E2's can catalyze protein monoubiquitination in the absence of E3, while other E2's can catalyze the formation of polyubiquitin chains of the types described above. There are multiple E2's in most organisms, and mutations in individual isoforms may confer very specific phenotypes (Jentsch, 1992a), for example, radiation sensitivity and defects in DNA repair (RAD6), cell cycle arrest (CDC34), defects in peroxisome biogenesis (PAS2), and reduced stress response (UBC4,5). Finally, several E3 ligases have been reported (Hershko & Ciechanover, 1992), including one which participates in the ubiquitination of p53 (Scheffner *et al.*, 1993). Thus, while the details and specificity of the ubiquitination process are as yet unclear, it is likely that there are several semi-independent polyubiquitination pathways, and perhaps even different types of polyubiquitin signals.

In addition to polyubiquitinated conjugates which are intermediates in protein degradation, the cell also contains a variety of monoubiquitinated conjugates of poorly defined function. For example, monoubiquitinated histone 2A is the most abundant single conjugate in higher eukaryotes. It is not an intermediate in degradation but is reversibly disassembled during mitosis by an isopeptidase (Matsui *et al.*, 1982). Additional types of ubiquitin derivatives also exist. All of the genes encoding ubiquitin and ubiquitin-related proteins direct synthesis of fusion proteins. One type of ubiquitin gene (Ozkaynak *et al.*, 1987) encodes proubiquitin, a head-to-tail fusion of five or more ubiquitin monomers. A second class of ubiquitin genes directs synthesis of fusions of ubiquitin with ribosomal proteins (Ub-CEPs). Both types of fusions must be processed to the mature form by proteolytic cleavage at G76.

The structural and functional diversity of ubiquitin polymers, fusions, and conjugates suggests that the catabolism of these species should be regulated. In particular, there must be efficient processing of primary ubiquitin gene products linked by peptide bonds, without premature disassembly of the isopeptide-linked polyubiquitin degradation signal. The existence of a large family of putative isopeptidases (Papa & Hochstrasser, 1993) suggests that different isozymes may have different specificity, but there are as yet little data to clarify this hypothesis. To examine the mechanisms underlying the specificity of polymeric ubiquitin catabolism, we have undertaken cDNA cloning and characterization of isopeptidase T. Isopeptidase T is a ubiquitin-binding protein first purified from reticulocytes by adsorption to, and elution from, a ubiquitin affinity column (Pickart & Rose, 1985a). A partially purified preparation has been reported to disassemble free polyubiquitin chains with the K48 linkage (Chen & Pickart, 1990), to stimulate degradation of ubiquitinated proteins by the 26S proteasome complex (Hadari *et al.*, 1992), and to process the proubiquitin gene product (Falquet *et al.*, 1995). On the basis of these results, it was postulated that this enzyme acts at the terminal step of ubiquitin-dependent proteolysis, disassembling the polyubiquitin chain following the action of the proteasome on the target protein (Hadari *et al.*, 1992). The present results confirm this hypothesis and strongly suggest that the physiological role of isopeptidase T is to disassemble polyubiquitin chains after the chains are released from the target protein. The results also provide insight into the molecular basis of the enzyme's specificity for free branched polyubiquitin chains *vs* polyubiquitinated protein conjugates.

EXPERIMENTAL PROCEDURES

Purification of Isopeptidase T. Isopeptidase T was purified from bovine erythrocytes. Fraction II was prepared from fresh bovine erythrocytes as described for reticulocytes (Hershko *et al.*, 1983), except that the volume of the DE-52 column was reduced by a factor of 4 to compensate for the lower content of acidic proteins in erythrocytes relative to reticulocytes. Fraction II, 9 mL containing 138 mg of protein, was made 0.1 mM in EDTA[2] and applied to a 3-mL ubiquitin−Sepharose column (5 mg of ubiquitin/mL of resin) previously equilibrated (at room temperature) with TDE buffer (50 mM Tris-HCl, 10% base, 0.1 mM EDTA, and 0.2 mM DTT). The column was first washed with 4 volumes of TDE buffer and then eluted with 4 volumes of elution buffer (50 mM Tris-HCl, 90% base, 0.1 mM EDTA, 0.2 mM DTT, and 0.1 mg/mL ovalbumin). The eluate was collected into 0.1 volume of 1 M Tris-HCl (10% base) on ice, concentrated, and dialyzed against TDE buffer. The concentrated and dialyzed eluate (0.6 mL) was applied to a 1-mL FPLC monoQ column previously equilibrated with buffer A (50 mM Tris-HCl, 5% base, pH 7.5/5 °C, 0.1 mM EDTA, and 0.2 mM DTT). The column was eluted at 5 °C with a linear gradient of NaCl (0 to 0.5 M in 40 mL of buffer A) at a flow rate of 0.5 mL/min. The absorbance of the eluate was continuously monitored at 280 nm. Isopeptidase T eluted at 21 mL (0.27 M NaCl). The peak fractions were pooled, concentrated, and exchanged into TDE buffer. For experiments with ubiquitin ethyl ester, the isopeptidase was further purified on a Superose-6 gel filtration column (Pharmacia) in order to remove any residual UCH. The concentration of isopeptidase T was estimated by SDS−PAGE and Coomassie blue staining (comparison to a series of BSA standards). The yield of isopeptidase was approximately 1.5 mg/g of fraction II protein. The final preparations were estimated to be significantly greater than 98% pure by SDS−PAGE, nondenaturing PAGE, and isoelectric focusing and capillary liquid chromatography.

[2] Abbreviations: desGGUb, residues 1−74 of ubiquitin; DTT, DL-dithiothreitol; EDTA, ethylenediaminetetraacetic acid; PAGE, polyacrylamide gel electrophoresis; PEG, poly(ethylene glycol); PMSF, phenylmethylsulfonyl fluoride; SDS, sodium dodecyl sulfate; Tris, 2-amino-2-(hydroxymethyl)-1,3-propanediol; UBP, ubiquitin specific proteases (Tobias & Varshavsky, 1991; Baker *et al.*, 1992; Papa & Hochstrasser, 1993); UCH, ubiquitin carboxyl-terminal hydrolase (Wilkinson *et al.*, 1989; Wilkinson, 1994); UV, ultraviolet.

Characterization of Isopeptidase T. The molecular mass of bovine isopeptidase T was determined by electrospray mass spectrometry (Dr. Clive A. Slaughter, Howard Hughes Medical Institute, Dallas, TX) to be 93 267 ± 13 Da. Nondenaturing PAGE in gradient gels (Pharmacia Phast gels, 8−25% acrylamide) gave an estimated molecular mass of 92 400 ± 7900 Da, and SDS−PAGE gave an estimated molecular mass of 100 000 ± 5000 Da. Peptides were prepared by digesting isopeptidase T with trypsin (2 × 5% w/w for 30 min each) in 0.05 M sodium bicarbonate, 10 mM $CaCl_2$. Peptides were separated by HPLC as described previously (Cox *et al.*, 1986) and sequenced by the Emory University Microchemical Facility.

Cloning and Sequencing a cDNA-Encoding Human Isopeptidase T. It was noted that four of the tryptic peptide sequences determined above were encoded by the 5′ end of a cDNA from a human fetal brain library (Adams *et al.*, 1994). PCR amplification of this sequence (manufacturer's instructions, Perkin-Elmer Corp., Norwalk, CT) using primers complimentary to the 5′ end (GenBank:T08022) and the 3′ end (GenBank:T08021) of this cDNA yielded a 1.4 kB clone (FexRi.1) from several cDNA libraries (Figure 1). Futher studies were conducted with a cDNA library from HSC93 lymphoblast cells, generously donated by Dr. Manuel Buchwald (Strathdee *et al.*, 1992). Using oligonucleotides complimentary to the middle of the coding region (RC1 and RC2) as reverse PCR primers and a vector specific primer complimentary to the RSV region as a forward primer for the PCR amplification, two more partial clones were obtained. This result was confirmed by using another cDNA library derived from HL60 cells donated by Dr. David Uhlinger, Emory University. Independent amplification of HL60 cDNA with the same gene specific primer (RC2) and the appropriate vector specific primer (complimentary to the T7 RNA polymerase binding site) resulted in the clone with the same internal sequence as the one obtained previously from HSC93 cDNA library. A second partial clone, RsvRc.1, was generated using a primer complimentary to the 5′ end of FexRi.1 and another to the SV40 promoter region of the vector. Finally, several clones 2.8 kB in length were amplified using primers complimentary to the 5′ end of the gene and to the 3′ UTR. All PCR products were rescued in the TA Cloning System (Invitrogen). DNA sequences were determined using the Sanger dideoxy techniques (Sequenase Kit, United States Biochemical) or on an automated sequencer (Applied Biosystems 373A instrument). Two clones from independent amplification reactions were completely sequenced on both strands, and ambiguities were resolved by sequencing additional clones from independent amplifications. This cDNA sequence has been deposited in the GenBank database under the accession number U35116.

Assays. Enzyme activity was detected by SDS−PAGE or HPLC of reaction mixtures; both assays gave similar results. Assays were conducted in a volume of 12−70 μL at 37 °C and contained 50 mM Tris-HCl (24% base, pH 7.3), 0.1 mM EDTA, and 0.5 mM DTT. Isopeptidase was diluted in a buffer containing ovalbumin (1 mg/mL) and DTT (0.5 mM) to counteract absorption and prevent thiol oxidation. The presence of carrier ovalbumin in enzyme and Ub_2 (below) resulted in ~1 mg/mL ovalbumin in the assays. For SDS−PAGE assay, aliquots (usually 10 μL) were removed at timed intervals and quenched by addition to an equal volume of SDS−PAGE sample buffer. After being heated to 100 °C, aliquots (usually 15 μL) were electrophoresed on gels containing 13.5% acrylamide. The gels were dried and autoradiographed; the bands corresponding to Ub_2 and Ub were excised and counted. HPLC-based assays were done as described previously (Wilkinson *et al.*, 1986). Under the conditions used, the elution times of these chains were as follows: monoubiquitin, 6.2 min; diubiquitin, 8.22 min; and triubiquitin, 9.9 min. Substrate concentrations were varied from 0.1 to 30 μM as necessary. Protein concentrations were determined using a dye binding assay (Bearden, 1978).

Synthesis of Polymeric Ubiquitin Chains and Derivatives. Unless otherwise noted, all ubiquitin derivatives were linked by an isopeptide bond between the C-terminus of one subunit with the ε-amino group of K48 on the next subunit. Recombinant bovine $E2\text{-}_{25K}$ was purified as described (Pickart *et al.*, 1992). E1 was purified from rabbit reticulocytes or bovine erythrocytes (Pickart & Vella, 1988). It was found that the gradient anion exchange step (FPLC monoQ) during E1 purification was critical in order to avoid isopeptidase T contamination. Recombinant murine $E2\text{-}_{20K}$ was purified as described (Wefes *et al.*, 1995). Unless otherwise indicated, chain synthesis reactions were carried out at 37 °C, and the mixtures contained 0.4 μM E1, 10 μM $E2\text{-}_{25K}$, 0.4 mM DTT, 50 mM Tris-HCl (50% base), 2.5 mM $MgCl_2$, 1 mM ATP, 5 mM phosphocreatine, and 0.3 u/mL each of CPK and PPase.

Ub_2-OEt was synthesized in a volume of 1 mL containing Ub (0.5 mg/mL) and Ub-OEt [2 mg/mL (Wilkinson *et al.*, 1986)]. After 60 min, E1 and E2 were removed by passage through Q-Sepharose, and Ub_2-OEt was separated from residual Ub-OEt by FPLC cation exchange chromatography at pH 4.5 (Pickart *et al.*, 1992). To facilitate recovery in chromatography here and elsewhere, fractions were collected in tubes containing BSA or ovalbumin (0.1 mg/tube).

The end-labeled tetramer ({^{125}I-K48R-Ub}-Ub_3-OEt) was generated by using $E2\text{-}_{25K}$ to condense the free Lys-48 side chain of Ub_2-OEt with the free C-terminus of end-labeled {^{125}I-K48R-Ub}-Ub. Purified recombinant K48R-Ub was a generous gift of Robert Cohen (University of Iowa). To make {^{125}I-K48R-Ub}-Ub, ^{125}I-K48R-Ub (~20 cpm/pmol, 1.3 mg/mL) was incubated with unlabeled Ub-OEt (1.3 mg/mL) as described above, except that the volume was 0.5 mL and incubation was for 1 h. These two species can only assemble in one way, since K48R-Ub lacks a free Lys-48 and Ub-OEt lacks a free C-terminus. Ion exchange chromatography (see above) yielded 0.5 mg of {^{125}I-K48R-Ub}-Ub-OEt. To generate the capped, end-labeled tetra-ubiquitin {^{125}I-K48R-Ub}-Ub-OEt (0.4 mg) was incubated with 20 mM NaOH (37 °C, 30 min), to generate a free C-terminus. Following neutralization with HCl, Ub_2-OEt (0.4 mg) was added, together with enzymes and other components as above (volume, 0.5 mL; incubation for 100 min). The two dimers can only be assembled in one way, since the labeled dimer lacks a free Lys-48 and the unlabeled dimer lacks a free C-terminus. Since the label is associated with K48R-Ub, the label must be in the distal Ub unit. Ion exchange chromatography (see above) yielded electrophoretically homogeneous end-labeled {^{125}I-K48R-Ub}-Ub_3-OEt (0.1 mg, 20 cpm/pmol).

Ub-Ub, Ub-{G76A-Ub}, and {G76A-Ub}-Ub were synthesized in single turnover reactions. The $E2\text{-}_{25K}$ thiol ester adduct of Ub or G76A-Ub was generated in a pulse (330

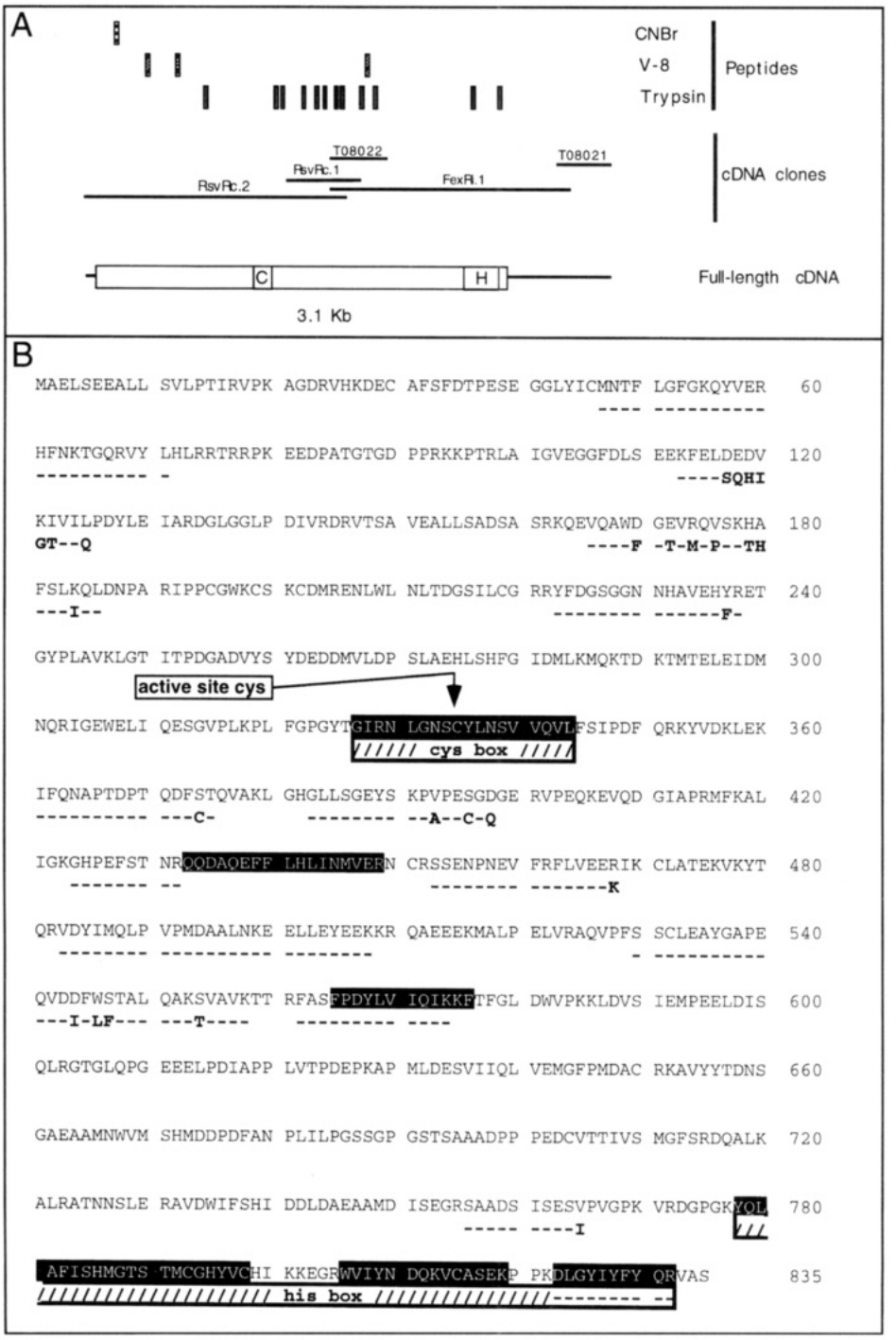

FIGURE 1: Sequence of human isopeptidase T. (A) Arrangement of sequenced peptides (from the bovine protein), cDNA clones generated, and the structure of the putative full-length mRNA. The coding sequence is indicated by the box, and the position of the Cys domain (C) and the His domain (H) are indicated. Sequences T08022 and T08021 were previously reported to be present in the sequence-expressed tag HIBAA66 (Adams *et al.*, 1994). (B) The deduced amino acid sequence of human isopeptidase T is shown in the top sequence. The sequences similar in all UBP isozymes are highlighted in black. The Cys box and the His box sequences are indicated. The lower sequences are those of the bovine peptides sequenced. Residues identical in the bovine and human proteins are indicated by dashes, and differences are noted. The bovine and human proteins are 88% identical in the 16 peptides (221 residues) compared.

μL, pH 7.3, 37 °C), containing 32 μM of the relevant Ub, ~3 μM of the relevant ^{125}I-Ub, 40 μM E2-$_{25K}$, and 0.5 μM E1 (other conditions as described above). Ub specific radioactivity in the pulse was ~400 cpm/pmol. After incubation for 15 min, the chase was initiated by adding 0.1 volume of 0.2 M EDTA (pH 7.0), followed immediately by 1 mg of unlabeled Ub or G76A-Ub (in 90−100 μL). Incubation was continued for 50 min more at 37 °C. In the pulse, ≥90% of the input Ub was linked to the E2; all of the E2-bound Ub appeared in Ub_2 during the chase. The respective dimers were purified as above. As expected, the final specific radioactivity of each Ub_2 was ~400 cpm/pmol. Overall recovery of Ub_2 was 20−40%.

The $\{G76A\text{-}Ub\}_2$ dimer was synthesized *in situ* in bovine reticulocyte fraction II. The incubation contained (200 μL, pH 7.3, 37 °C) 3 mg/mL fraction II protein; ATP, Tris-HCl buffer, $MgCl_2$, pyrophosphatase, and the ATP-regenerating components (see above) were present at 2-fold higher

concentrations than in the pulses for the single turnover reactions. G76A-Ub was present at 5 mg/mL. After 3 h of incubation, the reaction mixture was passed through a 1 mL Q-Sepharose column to remove fraction II protein. The $\{G76A\text{-}Ub\}_2$ dimer was resolved from Ub by gradient cation exchange chromatography as described above.

Ub_3 was generated using $E2\text{-}_{25K}$ and purified by cation exchange chromatography as described above. $^{125}I\text{-}Ub_3$ (~5 μM) was conjugated to yeast cytochrome *c* (2 mg/mL, Sigma) according to published procedures (Pickart & Rose, 1985b), using E1 (0.1 μM) and recombinant murine $E2\text{-}_{20K}$ (1.5 μM) 110 μL, incubation for 30 min). Ub_3–cytochrome-*c* was separated from residual Ub_3 by cation exchange chromatography on S-Sepharose at pH 7.6. A 1 M NaCl eluate, containing free cytochrome-*c* and Ub_3–cytochrome-*c*, was dialyzed prior to use for assay of isopeptidase T.

Proubiquitin oligomers of different length were prepared by expressing the yeast pentaubiquitin gene in *Escherichia coli* (Jonnalagadda *et al.*, 1987). Linear oligomers (n = 2, 3, 4, and 5) were purified by ion exchange chromatography and gel filtration as described above for the polyubiquitin chains. Pretreatment of these oligomers with ubiquitin carboxyl-terminal hydrolase (UCH-L3) is expected to remove the C-terminal aspargine coded for by the pentaubiquitin gene. This treatment had no effect on the kinetics of processing by isopeptidase T.

RESULTS

Structure of isopeptidase T. We have purified isopeptidase T to apparent homogeneity from bovine erythrocytes and sequenced 12 unique tryptic peptides. Additional sequence information was forwarded to us by Martin Rechsteiner and his collaborators (private communication). Initial data base searches using these short peptide sequences revealed no statistically significant similarity with other known proteins (but see below). However, it was noted that the 5′ end of a cDNA sequence from a human fetal brain library (Adams *et al.*, 1994) encoded a peptide sequence with a high degree of identity to four of the peptides sequenced. PCR primers based on the sequence of the reported cDNA readily amplified a 1.4 kB fragment from several human cDNA libraries. This fragment contained the C-terminal half of the protein and approximately 600 base pairs of 3′-untranslated sequence. Additional clones were identified by PCR from both human lymphoblast and HL60 cDNA libraries, and a putative full-length sequence was determined by sequencing several independent PCR products. Figure 1A shows the relationship between the peptides, the individual cDNA clones isolated, and the full-length cDNA. The deduced sequence predicts a protein of 835 amino acids, with a molecular mass of 93 334 Da. Figure 1B shows the deduced human sequence and the sequence of 16 bovine peptides (containing 221 residues) determined in the course of this work. In this figure, identities are indicated by a dash, and the observed differences in the bovine peptides and the human deduced sequence are indicated. This protein is highly conserved between these two species, with only 25 differences in 221 residues sequenced (12%).

The DNA sequence 5′ to the putative initiator methionine, as determined from clones originating from several different libraries, is quite short (45 nucleotides, no methionine, and no stop codons). As the purified bovine protein migrates on SDS–PAGE with an apparent molecular mass of 100 000 ± 5000 Da, it was a formal possibility that the true initiator codon was 5′ to the sequence identified. Numerous attempts to identify clones with longer 5′ sequences (using PCR or hybridization techniques) failed to indicate evidence of a longer human 5′-untranslated sequence. The N-terminus of the bovine protein is blocked and could not be directly sequenced. We thus determined the apparent molecular mass of purified bovine isopeptidase T using nondenaturing PAGE in 8–25% acrylamide gradient gels. In this technique, proteins migrate until they reach a "pore limit", defined by the size and shape of the protein but independent of the charge on the molecule. Using this technique, the apparent molecular mass of the native protein is 92 400 ± 7900 Da (average and standard deviation of four independent runs). In confirmation of this estimate, electrospray mass spectrometry on the purified protein gave a molecular mass of 93 267 ± 12.79 Da. If a more 5′ ATG codon were used, a protein product of at least 94 900 Da would have been expected. Thus, all of these data support the use of the indicated methionine to initiate translation of a protein of 93 334 Da.

Data base searches with the complete deduced sequence indicated a strong homology with several sequences of the ubiquitin carboxyl-terminal hydrolases, family 2. These proteins are also known as ubiquitin specific proteases (UBP). Regions of homology present in most, or all, of these sequences and also in isopeptidase T are boxed in Figure 1. It should be noted that initial data base searches with the short peptides described above failed to convincingly identify isopeptidase T as a member of this family. This is because of the short nature of the peptides and the sequence variabity of the "consensus" UBP sequences to which these peptides corresponded. Figure 2 shows the relationship between the putative catalytic domains of isopeptidase T and 20 other recognizable members of the UBP family. These sequences were detected in the protein and nucleic acid data bases using the BLAST network service at the National Center for Biotechnology Information (Altschul *et al.*, 1990) with the consensus sequence of the active-site peptide (GLXNLGNT-CYMNSVLQCL). Full-length coding sequences with scores of greater than 48 (probability <0.95) were further examined. All 20 sequences detected contained the previously described His box [Figure 1, residues 771–831 and also Papa and Hochstrasser (1993)]. The 21 sequences (beginning 27 residues N-terminal to the active site Cys and extending through the His box sequence) were submitted to the BLOCK MAKER server (URL http://www.blocks.fhc.org/blockmkr/make_blocks.html), and the six conserved sequences identified from this algorithm are reported in Figure 2. These conserved blocks include the Cys box, a QQDAQEF motif, a region with the consensus sequence LPQILVIHLKRF, and three blocks from the His box region. These similarities have previously been noted in a smaller subset of UBP isozymes (Papa & Hochstrasser, 1993). These sequence comparisons define a large number of UBP family members and suggest common structural or functional importance for these regions of these proteins. By these criteria, isopeptidase T is related to this family of ubiquitin-processing proteases. The similarities and differences among these family members are discussed in more detail below.

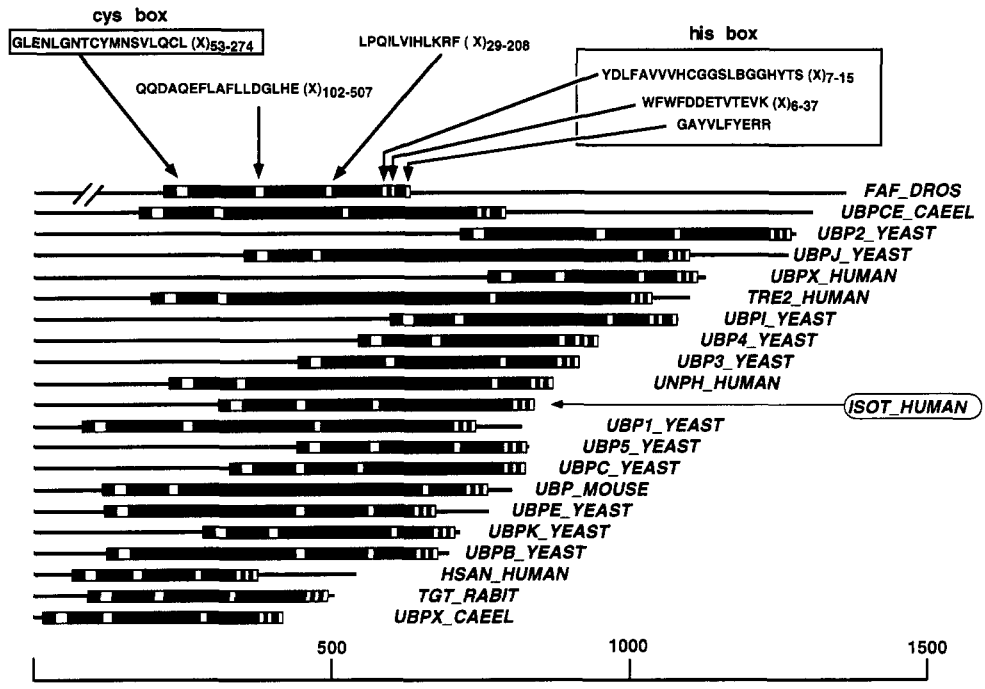

FIGURE 2: Human isopeptidase T is a member of the UBP gene family. The structure of 21 UBP family members is shown here schematically. The putative catalytic domains from each span the sequence between the Cys box (containing the active-site cysteine) and the highly conserved His box. Blocks of sequence conserved among all family members were identified using the BlocksMaker server (URL http://www.blocks.fhc.org/ blockmkr/make_blocks.html), and their positions are indicated by unfilled boxes. The consensus sequence of these blocks was determined using the ProfileMake subroutine of the GCG package (Genetics Computer Group, Madison, WI). These consensus sequences and the spacing between them are indicated above the aligned sequences. Sequence names and accession numbers are as follows: TRE2_HUMAN, SwissProt:p35125; UBPJ_YEAST, SwissProt:p39538; UBP1_YEAST, SwissProt:p25037; UNPH_HUMAN, GenPept:u20657; UBP_MOUSE, SwissProt:p35123; UBPCE_CAEEL, GenPept_z47811; UBPE_YEAST, SwissProt:p39967; UBPB_YEAST, SwissProt:p38187; UBP2_YEAST, SwissProt:q01476; ISOT_HUMAN, submitted; UBPC_YEAST, SwissProt:p38237; UBPI_YEAST, SwissProt:p40453; UBP3_YEAST, SwissProt:q01477; UBPK_YEAST, SwissProt:p36026; FAF_DROS, PIR:a49132; UBP4_YEAST, SwissProt:p32571; TGT_RABIT, SwissProt:p40826; UBPX_CAEEL, SwissProt:p34547; UBP5_YEAST, SwissProt:p39944; UBPX_HUMAN, SwissProt:p40818; HSAN_HUMAN, EMBL:d38378.

Substrate Specificity: Disassembly of K48-Linked Polyubiquitin Chains. Previous studies on this enzyme used lysozyme-bearing ubiquitin chains of heterogeneous length and linkage (Hadari *et al.*, 1992). We used homopolymeric K48-linked polyubiquitin chains (Chen & Pickart, 1990) and other polymeric ubiquitin substrates to probe the specificity of isopeptidase T. As reported previously (Chen & Pickart, 1990), branched polyubiquitin chains with the K48 linkage are good substrates for this enzyme, producing free ubiquitin with $k_{cat} \sim 40$ min^{-1} (see below). The results indicated that K_m was <1 μM (data not shown), and unless otherwise specified, all rate constants given below are k_{cat} values. The enzyme also exhibits considerable activity ($k_{cat} = 8$ min^{-1}) on the generic substrate, ubiquitin ethyl ester (Wilkinson *et al.*, 1986). In contrast, the enzyme is several orders of magnitude less efficient at processing large ubiquitin fusion proteins, including UbCEP52 (Monia *et al.*, 1989) or Ub-β-Gal (Bachmair *et al.*, 1986) (data not shown). This confirms previous work (Chen & Pickart, 1990; Hadari *et al.*, 1992; Falquet *et al.*, 1995) which suggested that the enzyme exhibits considerable specificity for the nature of the amide linkage and/or the precise structure of the leaving group.

Further characterization of this specificity was accomplished using branched polyubiquitin chains of defined lengths, some of which bore specific modifications (Figure 3). Branched diubiquitin bearing a G76A substitution in the distal ubiquitin (*i.e.*, at the scissile bond) was disassembled with a k_{cat} of 5 min^{-1}, only 8-fold lower than wild type. In contrast, branched diubiquitin bearing a G76A substitution in the proximal ubiquitin was disassembled with a 100-fold diminished k_{cat}. These effects were additive; there was a rate decrease of ~500-fold for dimer bearing G76A substitutions at both positions. These results suggested that isopeptidase T might discriminate among branched polyubiquitin substrates in part on the basis of the status of the C-terminus of the proximal ubiquitin in the chain.

The rates of disassembly of other branched di-, tri-, and tetraubiquitin derivatives strongly support this hypothesis. Thus, any change in the C-terminus greatly diminished the rate of disassembly (Figure 3). This included the removal of one or two glycines from the proximal C-terminus (>4000-fold), the G76A substitution (100-fold), and esterification (40−80-fold). These results suggest that the normal function of isopeptidase T is to act on unanchored chains, rather than on chains conjugated to substrate proteins in which the proximal C-terminus is necessarily blocked. Indeed, conjugation of triubiquitin to cytochrome-*c* diminished the rate of disassembly by at least 500-fold (Figure 3). Since these are k_{cat} values, this specificity suggests that interactions between the enzyme and the C-terminus of the proximal ubiquitin are important for proper positioning of the substrate or the catalytic groups on the enzyme, rather than for binding *per se*.

Substrate Specificity: Processing of Linear Proubiquitin. As this paper was being prepared, it was reported that

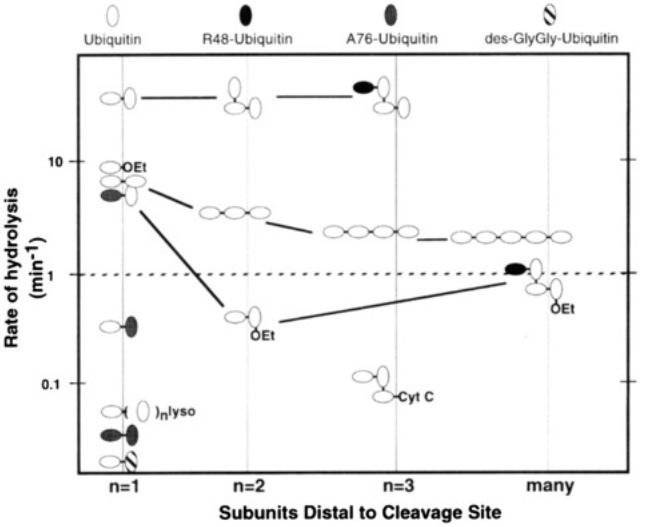

FIGURE 3: Rates of disassembly of various polymeric ubiquitin substrates. Substrates are grouped by the number of subunits distal to the scissile bond. The vertical lines indicate the position of cleavage in the indicated substrates. Ubiquitin is shown by a white oval, R48-ubiquitin by a black oval, A76-ubiquitin by a gray oval, and des-Gly-Gly-ubiquitin by a striped oval. The isopeptide-linked polymers are branched and have adjacent subunits perpendicular to each other (rates given are k_{cat} values). The linear α-amide-linked fusion proteins are arranged head-to-tail (rates given are k_{obs} values at 30 μM substrate). Note that the scale is logarithmic and spans over 3 orders of magnitude. In the cases of linear ubiquitin polymers and the branched $(Ub)_3$-cytochrome *c*, the substrate has been arbitrarily drawn to minimize the size of the leaving group. This would be consistent with the conclusions drawn from the branched polyubiquitin polymers.

isopeptidase T was able to process the linear diubiquitin fusion protein, *i.e.*, an α-linked amide bond (Falquet *et al.*, 1995). We have confirmed this activity with our preparations of protein. The rates of processing of linear diubiquitin (at a substrate concentration of 30 μM) was approximately 6 min^{-1}, as compared to an estimated 3 min^{-1} based on the data previously presented [Figure 2 (Falquet *et al.*, 1995)]. Other data presented in that paper indicate that the rate of branched chain processing by isopeptidase T was about 20−40 min^{-1}, in good agreement with our estimate of 40 min^{-1} for k_{cat}. Further studies in our laboratory have demonstrated that the rate of processing of α-linked ubiquitin polymers was length-dependent at 30 μM substrate, with k_{obs} decreasing from 6 min^{-1} for linear dimer to 1 min^{-1} for the linear pentamer (Figure 3). In contrast, the rate of processing of the branched K48-linked chains was 10−100-fold higher and independent of chain length. It is assumed, though not proven, that linear pentaubiquitin is cleaved from the proximal end as is shown below for branched polyubiquitin chains.

Mechanism of Polyubiquitin Processing. The requirement for an intact C-terminus on the proximal ubiquitin in the branched polyubiquitin chain suggested that isopeptidase T might act as an *exo*-isoamidase that releases ubiquitin sequentially from the proximal end of the chain. To test this possibility, we used a tetramer labeled in the distal ubiquitin with ^{125}I. If isopeptidase T acted sequentially from the proximal chain end, the initial product would be labeled trimer; this would be followed by labeled dimer and finally labeled monomer. In contrast, if disassembly occurred sequentially from the distal chain end, the initial (and only detectable) product would be labeled monomer. Significant *endo*-isopeptidase activity would produce labeled dimer as an early product. The results (Figure 4) conform precisely to the predictions of the first model, *exo* cleavage from the proximal end. Transient accumulation of 20−30% of the substrate in the form of cleavage intermediates during disassembly (Figure 4) indicates both that the enzyme is not processive and that the rates of cleavage of branched di-, tri-, and tetraubiquitin are similar.

Extent of the Polyubiquitin Binding Site. These results and those of previous workers suggest that isopeptidase T has either multiple active sites or multiple binding pockets for the subunits of the branched polyubiquitin chain (Hadari *et al.*, 1992). To examine this, we measured the effect of added ubiquitin monomer on the disassembly of polyubiquitin labeled in its distal subunit (Figure 5). In the absence of ubiquitin (solid squares), the sequential nature of the cleavage is apparent (see also Figure 4). In the presence of a 350-fold molar excess of monomeric ubiquitin (open squares), loss of tetramer is slightly slowed (panel A), while disassembly of trimer is significantly inhibited (panel B), and disassembly of dimer is blocked (panels C and D). Thus, the effectiveness of ubiquitin as an inhibitor is inversely related to the chain length of substrate. This data would suggest that there are indeed multiple subunit binding sites on isopeptidase T. The substantially weaker inhibition of tetramer disassembly by ubiquitin (as compared to its effect on disassembly of smaller chains) may indicate the presence of four sites, since ubiquitin monomer would be expected to compete poorly when all binding sites are occupied by subunits of a chain. Finally, it should be noted that this inhibition is not expected to be physiological, since the levels of ubiquitin used here are at least 5-fold higher than the measured levels of free ubiquitin in cells (Haas & Bright, 1985).

DISCUSSION

It has recently become clear that the ubiquitin pathway is important for the regulation of the cell cycle, for the repair of damaged protein, and for a variety of other cellular functions (Ciechanover & Schwartz, 1994; Jentsch, 1992a; Hershko & Ciechanover, 1992). Further, it is apparent that the mechanisms of, and balance between, synthesis and disassembly of the polyubiquitin degradation signal will be an important determinant in the specificity and rate of ubiquitin-dependent proteolysis. A reasonable model for the disassembly of the polyubiquitin degradation signal invokes the following sequential events: release of branched polyubiquitin from the terminal degradation intermediate (polyubiquitin linked to a small peptide), catalyzed by a proteasome-associated isopeptidase (Eytan *et al.*, 1993; Papa & Hochstrasser, 1993); dissociation of the free polyubiquitin chain from the S5a subunit of the 26S proteasome (Deveraux *et al.*, 1994); and disassembly of the free chain by an isopeptidase. This study characterizes the structure, specificity, and mechanism of isopeptidase T, an enzyme which appears to be responsible for the disassembly of free polyubiquitin chains.

The UBP Gene Family. The sequence of isopeptidase T (Figure 1) clearly identifies it as a member of the UBP gene family (Figure 2). The UBP gene family (ubiquitin C-terminal hydrolases, family 2) is a large family of related proteins, with many different members already recognized

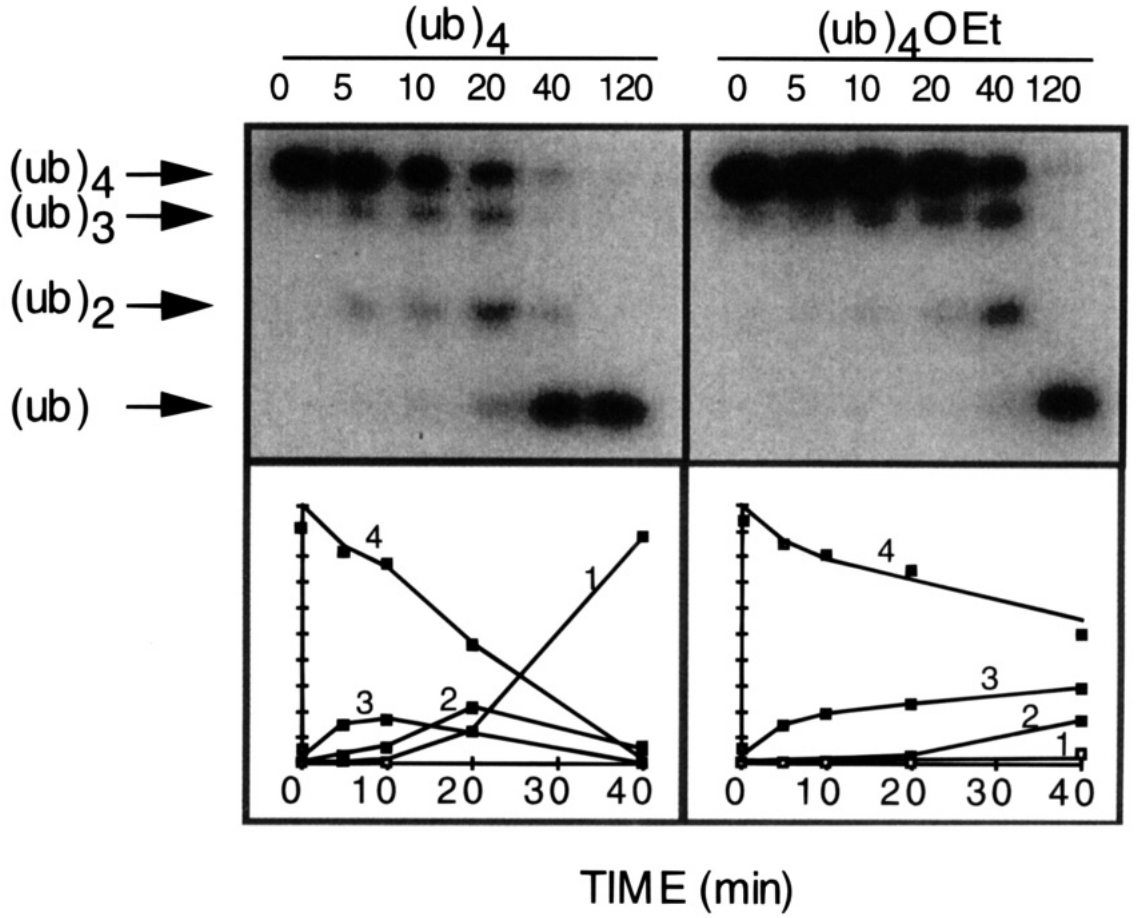

FIGURE 4: Disassembly of end-labeled tetraubiquitin by isopeptidase T. Isopeptidase was present at 0.6 nM, and the substrates were present at 0.8 μM. Incubations were terminated by addition of SDS sample buffer followed by SDS−PAGE. The radioautograms are shown in the upper panels; densitometric quantitation generated the data plotted in the lower panels. Substrates were tetraubiquitin derivatives linked by R48 isopeptide bonds: $(ub)_4$OEt, ^{125}I-R48Ub-$(Ub)_2$-UbOEt; and $(ub)_4$, ^{125}I-R48Ub-$(Ub)_2$-Ub.

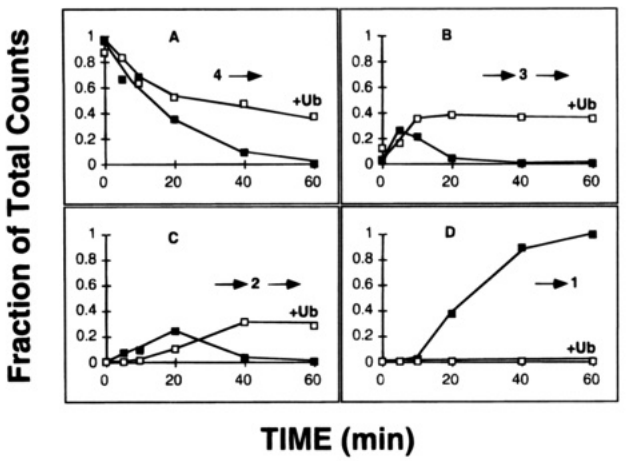

FIGURE 5: Inhibition of polyubiquitin disassembly by free ubiquitin. ^{125}I-R48Ub-$(Ub)_2$-Ub was subjected to the action of isopeptidase T, and the relative amounts of labeled ubiquitin species are shown: (A) $(Ub)_4$, (B) $(Ub)_3$, (C) $(Ub)_2$, and (D) (Ub). Filled symbols represent the disassembly observed in the absence of ubiquitin, while open symbols represent that observed presence of added unlabeled ubiquitin. Conditions are as in Figure 4, except that, where present, ubiquitin was 280 μM (350-fold molar excess over substrate).

(Baker *et al.*, 1992; Papa & Hochstrasser, 1993). Twenty-one of these are shown schematically in Figure 2, while several other partial sequences are detected in the data bases. Eleven of these complete sequences are from yeast, and five are from human sources. Thus, most organisms are likely to have multiple UBP isozymes. The sequences of all of these enzymes are similar in only six regions, encompassing approximately 100 residues.

The physical arrangement of these blocks of similarity suggest, that all UBP enzymes (including isopeptidase T) have a minimum catalytic core of approximately 300 residues, interrupted by insertions of various size. These catalytic domains span the sequence between the active site Cys and the His box (Papa & Hochstrasser, 1993) and range in size from 314 to 840 residues. Isopeptidase T is within this range with a catalytic domain of 527 residues. In addition, many of the isozymes have N-terminal extensions and, less frequently, C-terminal extensions. It seems likely that the insertions and extensions in different isozymes serve to control specificity, protein−protein interactions, or localization. Isopeptidase T has a 310-residue N-terminal extension with no apparent homology to non-UBP sequences. Isopeptidase T is most closely related to a yeast sequence UBPC, being 31% identical and 53% similar. The similarity extends throughout the length of the molecule, including both the N-terminal and catalytic domains. This suggests that UBPC may be the yeast homologue of isopeptidase T.

Substrate Specificity of Isopeptidase T. The specificity studies reported above can be summarized as follows. Branched polyubiquitin chains with free C-termini are the best substrates with k_{cat} values of 40 min^{-1} and K_m values below 10^{-6} M. Modification of the C-terminus of polyubiquitin chains results in a rate reduction ranging from 10

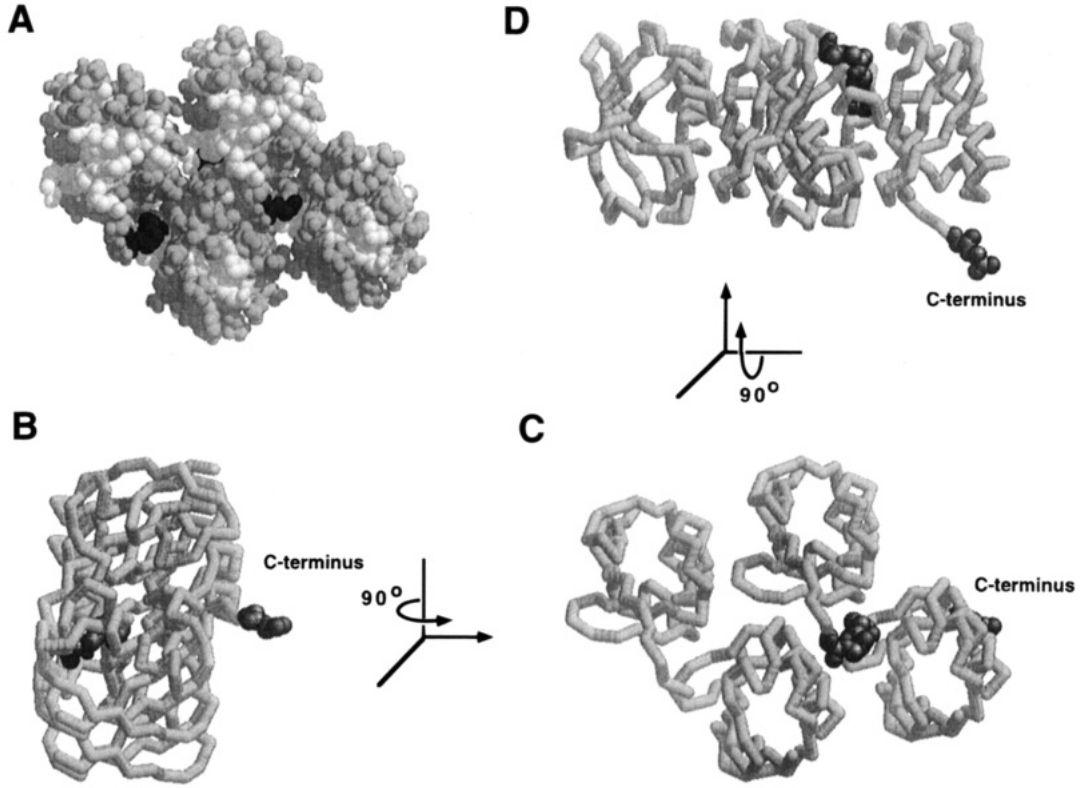

FIGURE 6: Location of the C-terminus in the polyubiquitin chain. The tetraubiquitin structure (Cook *et al.,* 1994) is shown in a space-filling representation (A). The residues forming the isopeptide bond are colored black, while hydrophobic residues are colored white. In panels B−D, the backbone is represented by a tube and the scissile bond and the free C-terminus of the chain are shown in a space-filling representation. The C-terminus is disordered in the crystals and so was model-built in an extended conformation. Panel C is derived from panel B by a rotation of 90° about the *y*-axis, and panel D is obtained by a rotation of panel C 90° about the *x*-axis. Panels A and C show the molecule in the same orientation. Steric constraints prevent the C-terminus from approaching closer than within 10−15 Å of the scissile bond.

to 10^4. Linear proubiquitin, linked by an α-amide bond, is a poor substrate with a k_{cat} of 6 min^{-1} and K_m values of over 10^{-5} M for the dimer. Finally, linear ubiquitin fusion proteins are poor substrates indeed, as shown by the inability of isopeptidase T to process Ub-β-Gal and Ub-CEP fusion proteins. Thus, any model explaining the specificity of isopeptidase T action must consider the dependence on a free C-terminus and the nature of the amide bond hydrolyzed.

The physiological importance of linear proubiquitin processing *vs* branched polyubiquitin disassembly by isopeptidase T may be estimated on the basis of the relative kinetic parameters and the abundance of the putative substrates. The specificity constant (k_{cat}/K_m) for branched polyubiquitin chains as substrate is approximately 5×10^5 M^{-1} s^{-1} [Falquet *et al.* (1995) and this data]. The comparable number for linear proubiquitin is approximately 4×10^3, some 100-fold lower. This means that, at low concentrations of substrates, polyubiquitin chains would be greatly preferred substrates. The steady-state concentrations of substrates is also relevant. There is no evidence for any significant steady-state level of proubiquitin. In fact, pulse chase studies after heat shock show very little, if any, proubiquitin accumulates. By contrast, polyubiquitin chains of the length studied here are a major steady-state species in yeast and in mammalian cells (Hodgins *et al.,* 1992; Arnason & Ellison, 1994; Spence *et al.,* 1995; Haldeman *et al.,* 1995). It can be conservatively estimated that polyubiquitin chains are at least 1 order of magnitude more abundant than proubiquitin fusion proteins. Thus, polyubiquitin chain disassembly by isopeptidase T would be expected to be quantitatively much more important than proubiquitin processing. Consistent with this interpretation, deletion of the yeast gene most closely related to isopeptidase T (UBPC) results in the accumulation of free polyubiquitin chains but no evidence for proubiquitin accumulation (M. Hochstrasser, unpublished observations). Presumably, one or more of the other UBP enzymes have a complementary specificity, *i.e.*, efficiently processing ubiquitin gene products and not chains.

Structure of the Bound Substrate. The marked dependence of the rate of disassembly on the presence of an intact C-terminus suggests that the C-terminus is involved in forming a catalytically functional complex. One unlikely possibility is that the free C-terminus could be involved in acid base catalysis at the scissile bond. In the published crystal structure of K48-linked tetraubiquitin (Cook *et al.,* 1994), the proximal C-terminus is disordered and could not be observed, but if one model builds the last four residues in an extended conformation, it is apparent that the C-terminus cannot approach more closely than 10 or 20 Å to the scissile bond; moreover, the C-terminus is on the opposite face of the molecule (Figure 6). These considerations exclude the possibility that the importance of the C-terminus derives from a mechanism in which the enzyme positions the C-terminus near the scissile bond in order to generate a catalytically active complex. Rather, it is likely that the C-terminus is bound to a specific site on the enzyme, with this binding serving to position the substrate properly for catalysis.

Because of the symmetry of the polyubiquitin structure, all of the subunits in the bottom row of Figure 6C have

similar conformations and orientations, as do all subunits in the top row. To recognize the C-terminal end of the chain, isopeptidase T would have to bind the free C-terminus itself, which projects from a face of ubiquitin that is blocked by intersubunit contacts in the body of the chain. This face includes residues S20, E24, D32, D39, E51, and D52. It is interesting to note that most of the residues which form intersubunit contacts lie near to, or are identical with, residues that vary in ubiquitin from different organisms (Wostmann *et al.*, 1992). Since all known substitutions at these positions involve conservation of charge, it is possible that there is a significant ionic interaction between the enzyme and the proximal end of the polyubiquitin chain.

Finally, we can suggest that the purpose of the enzyme binding to the free C-terminus is to alter the orientation (or conformation) of the proximal subunit of a branched polyubiquitin chain. The absence of an intact C-terminus prevents this reorientation and results in less efficient catalysis. The effect of the G76A substitution at the scissile bond may be due to the fact that the steric bulk of the side chain methyl group prevents the substrate from adopting the altered orientation. Glycine, which lacks a side chain, has much more conformational flexibility and may be better able to adopt the proper orientation for peptide bond hydrolysis.

Architecture of Polyubiquitin Chain Disassembly. Figure 4 demonstrates that isopeptidase T catalyzes the disassembly of free polyubiquitin chains by the sequential removal of one subunit at a time from the proximal end of the chain. A proximal *exo* mechanism is fully consistent with the substrate properties of a variety of branched di- and triubiquitins (above) which indicate a high degree of discrimination on the basis of the status of the C-terminus of the proximal ubiquitin. Similarly, esterification of the proximal C-terminus in branched polyubiquitin diminishes the rate of disassembly by at least 40−80-fold (Figure 3). Hence, the reduced ability of isopeptidase T to release a chain from the conjugate of triubiquitin with cytochrome-*c* (rate of <0.1 min^{-1}) is probably a simple reflection of the requirement for a free C-terminus.

These results are also consistent with the data of Hadari *et al.* (1992), who found that the disassembly of crude lysozyme conjugates by isopeptidase T was more rapid if the conjugates were exposed to the 26S protease. This is presumably because proteasome-associated isopeptidase released free chains (Eytan *et al.*, 1993; Papa & Hochstrasser, 1993). They observed "trimming" of crude lysozyme conjugates by isopeptidase T alone; the calculated rate of this reaction, measured as appearance of ubiquitin monomer, is ~ 0.05 min^{-1} (Hadari *et al.*, 1992). The rates we observed for C-*exo* disassembly of short polyubiquitin chains are at least 3 orders of magnitude faster.

It has been reported that expression of G76A-ubiquitin in yeast cells leads to the accumulation of free polyubiquitin chains (Hodgins *et al.*, 1992). On the basis of the results shown in Figure 3, this is consistent with isopeptidase T playing a major role in the disassembly of polyubiquitin chains. However, the present results suggest that mutant chain accumulation depends more upon the presence of the mutant residue at the free C-terminus of the proximal ubiquitin than upon the presence of the mutant residue at the scissile bond as was previously suggested (Hodgins *et al.*, 1992).

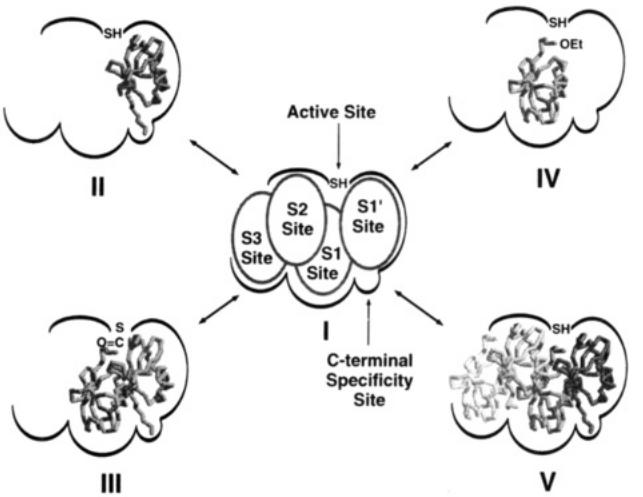

FIGURE 7: Model of isopeptidase T mechanism. The subsite nomenclature was adapted from that used for specifying recognition sites in peptidases (Schecter & Berger, 1967). The sites are drawn to correspond to the shape of the tetraubiquitin chain (Cook *et al.*, 1994). The S1′ site is the position of free ubiquitin binding (II) and the proximal subunit of the polyubiquitin chain (V). Ubiquitin ethyl ester apparently binds in the S1 site as shown in IV with its scissile bond positioned near the active-site thiol. Structure III depicts the simultaneous binding of ubiquitin to the S1′ site and ubiquitin aldehyde to the S1 site. This complex may be stabilized by formation of a thiol−hemiacetal between the active-site thiol and the aldehyde functionality.

Finally, it has been reported (Falquet *et al.*, 1995) and confirmed here that linear proubiquitin chains are also poor substrates for isopeptidase T. The α-amino group of M1 and the ϵ-amino group of K48 are ~21 Å apart. Presumably, proubiquitin can adopt a conformation somewhat similar to that required for catalysis, allowing disassembly with a slow k_{cat} and weak K_m.

A Model for Isopeptidase T Action. Figure 7 presents a model for the isopeptidase active site which incorporates all of the currently available information. The model postulates a specificity pocket that binds the free C-terminus of the chain within the S1′ site (Schecter & Berger, 1967), the site which accommodates the proximal subunit. The existence of a specificity pocket follows from the finding that any modification of the chain C-terminus impairs enzyme activity. The strong discrimination against the presence of Ala at the C-terminus suggests that this specificity pocket is sterically or conformationally hindered. Free ubiquitin binds to the enzyme, as shown directly (Hadari *et al.*, 1992) and by its ability to block the disassembly of chains of $n \leq 3$ (Figure 5). Ubiquitin binding is not competed with by des-GG-ubiquitin or by ubiquitin C-terminal alcohol (Hadari *et al.*, 1992), suggesting that a free C-terminus is an important determinant of monomeric ubiquitin binding, *i.e.*, that ubiquitin monomer binds preferentially in the S1′ site (Figure 7, II). The enzyme presumably has a nucleophilic active site thiol, since it is a member of the UBP family of thiol proteases and is sensitive to both thiol-alkylating reagents and the generic inhibitor ubiquitin aldehyde (Hadari *et al.*, 1992). The catalytic thiol should be positioned at some distance from the chain C-terminus, on the basis of considerations of chain structure (above). Thus, it is likely that ubiquitin aldehyde binds differently than unmodified ubiquitin and that both can be bound simultaneously [Figure 7, III, and Hadari *et al.* (1992)]. These results suggest that there are at least two sites (S1 and S1′) on the enzyme which

interact with ubiquitin, perhaps in a fashion similar to the binding of the two C-terminal subunits of polyubiquitin. The enzyme also catalyzes a very slow reaction of ubiquitin ethyl ester and glycine ethyl ester to give ubiquitin−glycine ethyl ester (data not shown). No such reaction occurred with free ubiquitin and glycine ethyl ester, suggesting that ubiquitin ethyl ester binds in the S1 site (where it is positioned for attack by the catalytic thiol) rather than in the S1′ site which free ubiquitin will occupy by virtue of its C-terminus (Figure 7, IV). Finally, polyubiquitin interacts extensively with the enzyme up to a chain length of at least four subunits (Figure 7, V). This binding places the proximal subunit in the S1′ site (*via* interaction between the C-terminus and the specificity pocket); disassembly occurs between the S1 and S1′ sites. This model also explains the nonprocessive nature of the enzyme; after disassembly of the first subunit, the enzyme and substrate must rotate by 180° with respect to each other, and translocate by one binding site, in order to place the newly formed C-terminus in the S1′ specificity pocket. This extensive rearrangement may not be possible without first dissociating the shortened polyubiquitin chain.

The precise requirements for binding to the S1′ specificity pocket are not well-defined. Clearly, the enzyme binds the C-terminus of the proximal subunit. However, the S1′ site can also accommodate proubiquitin linked by an α-amide bond. It is not known if the processing of proubiquitin requires an intact C-terminus, but it seems unlikely that proubiquitin could be bound by the C-terminus and still position the scissile bond in the same location as adopted by the best substrates, free polyubiquitin chains. The structure of proubiquitin has not been determined, but it is possible that the conformation of α-linked ubiquitin bound to the S1′ site positions the scissile bond properly without interacting with the C-terminus.

Physiological Implications. Our results help to clarify the mechanisms by which the cell discriminates among different ubiquitin conjugates. Isopeptidase T is unlikely to function in disassembly of substrate-conjugated polyubiquitin, since the ubiquitin(s) in such conjugates have blocked C-termini. The polyubiquitin signal is elaborated specifically on proteins destined for degradation and ideally should be disassembled only after degradation of the target protein. This can be accomplished by the sequential reactions of the proteasome (degradation of the target), a proteasome-associated isopeptidase (cleavage of the residual peptide from the polyubiquitin chain), and isopeptidase T (chain disassembly *via* a C-terminal *exo* mechanism). It is obviously desirable to avoid trimming and/or disassembling the polyubiquitin degradation signal before the proteasome acts; this is accomplished by having isopeptidase T recognize only chains with unblocked C-termini. However, it is less clear why the polyubiquitin chain needs to be disassembled after degradation, since polyubiquitin can be activated and ligated to proteins (Chen & Pickart, 1990; Van Nocker & Vierstra, 1993). Presumably, one of the functions of isopeptidase T is to prevent the accumulation of a large fraction of cellular ubiquitin in the form of chains. This could be deleterious because it would lead to a deficit of the monomeric ubiquitin required for certain conjugative reactions (*e.g.*, histone ubiquitination) and could promote proteasome inhibition through binding of unanchored chains to S5a.

ACKNOWLEDGMENT

We thank Jane Chastain and Steven T. Warren for help with the DNA sequencing, William J. Cook for furnishing the tetraubiquitin coordinates, and Daniel Finley for helpful discussions. We would especially like to thank Mark Hochstrasser for drawing our attention to the EST sequences, for sharing his data on the yeast UBPC mutations, and for many helpful suggestions.

REFERENCES

Adams, M. D., Soares, M. B., Kerlavage, A. R., Fields, C., & Venter, J. C. (1994) *Nat. Genet. 4*, 373−380.
Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
Arnason, T., & Ellison, M. J. (1994) *Mol. Cell. Biol. 14*, 7876−7883.
Bachmair, A., Finley, D., & Varshavsky, A. (1986) *Science 234*, 179−186.
Baker, R. T., Tobias, J. W., & Varshavsky, A. (1992) *J. Biol. Chem. 267*, 23364−23375.
Bearden, J. C. (1978) *Biochim. Biophys. Acta 533*, 525−529.
Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., & Varshavsky, A. (1989) *Science 243*, 1576−1583.
Chen, Z., & Pickart, C. M. (1990) *J. Biol. Chem. 265*, 21835−21842.
Ciechanover, A., & Schwartz, A. L. (1994) *FASEB J. 8*, 182−191.
Cook, W. J., Jeffrey, L. C., Kasperek, E., & Pickart, C. M. (1994) *J. Mol. Biol. 236*, 601−609.
Cox, M. J., Shapira, R., & Wilkinson, K. D. (1986) *Anal. Biochem. 154*, 345−352.
Deveraux, Q., Ustrell, V., Pickart, C., & Rechsteiner, M. (1994) *J. Biol. Chem. 269*, 7059−7061.
Eytan, E., Armon, T., Heller, H., Beck, S., & Hershko, A. (1993) *J. Biol. Chem. 268*, 4668−4674.
Falquet, L., Paquet, N., Frutiger, S., Hughes, G. J., Hoang-Van, K., & Jaton, J. C. (1995) *FEBS Lett. 359*, 73−77.
Gregori, L., Poosch, M. S., Cousins, G., & Chau, V. (1990) *J. Biol. Chem. 265*, 8354−8357.
Haas, A. L., & Bright, P. M. (1985) *J. Biol. Chem. 260*, 12464−12473.
Haas, A., Reback, P. M., Pratt, G., & Rechsteiner, M. (1990) *J. Biol. Chem. 265*, 21664−21669.
Hadari, T., Warms, J. V., Rose, I. A., & Hershko, A. (1992) *J. Biol. Chem. 267*, 719−727.
Haldeman, M. T., Finley, D., & Pickart, C. M. (1995) *J. Biol. Chem. 270*, 9507−9516.
Hershko, A., & Ciechanover, A. (1992) *Annu. Rev. Biochem. 61*, 761−807.
Hershko, A., Heller, H., Elias, S., & Ciechanover, A. (1983) *J. Biol. Chem. 258*, 8206−8214.
Hodgins, R. R., Ellison, K. S., & Ellison, M. J. (1992) *J. Biol. Chem. 267*, 8807−8812.
Jentsch, S. (1992a) *Trends Cell. Biol. 2*, 98−103.
Jentsch, S. (1992b) *Annu. Rev. Genet. 26*, 179−207.
Jonnalagadda, S., Butt, T. R., Marsh, J., Sternberg, E. J., Mirabelli, C. K., Ecker, D. J., & Crooke, S. T. (1987) *J. Biol. Chem. 262*, 17750−17756.
Matsui, S., Sandberg, A. A., Negoro, S., Seon, B. K., & Goldstein, G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1535−1539.
Monia, B. P., Ecker, D. J., Jonnalagadda, S., Marsh, J., Gotlib, L., Butt, T. R., & Crooke, S. T. (1989) *J. Biol. Chem. 264*, 4093−4103.
Ozkaynak, E., Finley, D., Solomon, M. J., & Varshavsky, A. (1987) *EMBO J. 6*, 1429−1439.
Papa, F. R., & Hochstrasser, M. (1993) *Nature 366*, 313−319.
Pickart, C. M., & Rose, I. A. (1985a) *J. Biol. Chem. 260*, 7903−7910.
Pickart, C. M., & Rose, I. A. (1985b) *J. Biol. Chem. 260*, 1573−1581.
Pickart, C. M., & Vella, A. T. (1988) *J. Biol. Chem. 263*, 15076−15082.

Pickart, C. M., Haldeman, M. T., Kasperek, E. M., & Chen, Z. (1992) *J. Biol. Chem. 267*, 14418−14423.
Schecter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 2*, 157−162.
Scheffner, M., Huibregtse, J. M., Vierstra, R. D., & Howley, P. M. (1993) *Cell 75*, 495−505.
Spence, J., Sadis, S., Haas, A. L., & Finley, D. (1995) *Mol. Cell. Biol. 15*, 1265−1273.
Strathdee, C. A., Gavish, H., Shannon, W. R., & Buchwald, M. (1992) *Nature 356*, 763−767.
Tobias, J. W., & Varshavsky, A. (1991) *J. Biol. Chem. 266*, 12021−12028.
Van Nocker, S., & Vierstra, R. D. (1993) *J. Biol. Chem. 268*, 24766−24773.
Wefes, I., Mastrandrea, L., Haldeman, M., Koury, S. T., Tamburlin, J., Pickart, C. M., & Finley, D. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 4982−4986.
Wilkinson, K. D. (1994) in *Heat Shock Proteins in the Nervous System* (Mayer, R. J., & Brown, I. R., Eds.) pp 191−234, Academic Press, London.
Wilkinson, K. D., Cox, M. J., Mayer, A. N., & Frey, T. (1986) *Biochemistry 25*, 6644−6649.
Wilkinson, K. D., Lee, K. M., Deshpande, S., Duerksen-Hughes, P. J., Boss, J. M., & Pohl, J. (1989) *Science 246*, 670−673.

BI9513923