THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 51, Issue of December 19, pp. 32280–32285, 1997

Printed in U.S.A.

A Ubiquitin-specific Protease That Efficiently Cleaves the Ubiquitin-Proline Bond*

(Received for publication, July 9, 1997, and in revised form, October 10, 1997)

Catherine A. Gilchrist‡, Douglas A. Gray§, and Rohan T. Baker‡¶

From the ‡Molecular Genetics Group, John Curtin School of Medical Research, Australian National University, GPO Box 334, Canberra, ACT 2601, Australia and §Cancer Research Group, Ottawa Regional Cancer Centre, Ottawa, Ontario, Canada K1H 8L6

Ubiquitin is a small eukaryotic protein that is synthesized naturally as one of several fusion proteins, which are processed by ubiquitin-specific proteases to release free ubiquitin. The expression of heterologous proteins as fusions to ubiquitin in either prokaryotic or eukaryotic hosts often dramatically enhances their yield, and allows the exposure of any amino acid following cleavage of ubiquitin. The single exception is when proline is the amino acid immediately following ubiquitin; the ubiquitin-proline bond is poorly cleaved by presently studied ubiquitin-specific proteases. We show that the mouse ubiquitin-specific protease Unp, and its human homolog Unph, can efficiently cleave the ubiquitin-proline bond in ubiquitin fusion proteins of different sizes. N-terminal sequencing of the cleavage products reveals that cleavage occurs precisely at the ubiquitin-proline junction. The biological significance of this cleavage activity is unclear, as ubiquitin-proline fusions do not occur naturally. However, it may indicate a different catalytic mechanism for these ubiquitin-specific proteases and/or that they can cleave ubiquitin-like proteins. Unp and Unph thus represent versatile ubiquitin-specific proteases for cleaving ubiquitin-fusion proteins in biotechnology and basic research, regardless of both the amino acid immediately following ubiquitin, and the size of the fusion partner.

Ubiquitin is a highly conserved eukaryotic protein that is invariably synthesized as a fusion protein either to itself or to one of two ribosomal proteins (1). One consequence of this fusion structure is that the action of an endopeptidase is required to cleave the fusion precursors to release free ubiquitin for its conjugation to other proteins as a marker for selective protein degradation (reviewed in Refs. 2 and 3). This cleavage is accomplished by members of a large family of enzymes variably termed ubiquitin-specific proteases (Ubps),[1] ubiquitin C-terminal hydrolases, or deubiquitinating enzymes (4–7). In the yeast *Saccharomyces cerevisiae*, a family of 16 Ubp enzymes can be identified in its completely sequenced genome, based on sequence similarity to conserved sequence domains observed in the first three yeast Ubp enzymes to be studied (5). Ubp activity has been demonstrated for 15 of these recombinant enzymes (4–6, 8, 9).[2] In addition, proteins from other species that contain these conserved sequence motifs have also been demonstrated to have Ubp activity (6, 10, 11), indicating that the presence of such motifs strongly correlates with Ubp activity. Yeast also contains an additional ubiquitin-cleaving enzyme termed Yuh1p (12), which is unrelated to the Ubp family by sequence, and which has several homologs in mammals (13, 14). The presence of such a large family of Ubp enzymes in yeast is not required merely to cleave ubiquitin precursors, and is indicative of other regulatory roles in the ubiquitin pathway.

The unusual ubiquitin fusion structure, and the ability of Ubps to cleave the ubiquitin moiety, have been exploited in several ways. The relative insensitivity of Ubp-mediated cleavage to the amino acid residue immediately following ubiquitin allows the synthesis of a protein with any residue at its N terminus. This technique was used to generate of a set of otherwise identical β-galactosidase (β-gal) proteins that differed only at their N-terminal residue, and subsequently led to the discovery of the N-end rule: the relationship between the half-life of a protein and the identity of its N-terminal residue (15–18). One limitation observed during these experiments was that the ubiquitin-proline bond was cleaved very inefficiently, either in yeast or in mammalian cells (rabbit reticulocyte extract), presumably due to the unusual structure of proline. The rate of cleavage was estimated to be 20 times slower than all other ubiquitin-β-gal fusions (17). Furthermore, the presence of the uncleaved ubiquitin moiety in a ubiquitin-proline-β-gal (Ub-P-β-gal) fusion in yeast, or an "uncleavable" Ub-β-gal fusion engineered by mutation of the Ubp cleavage site, actually targeted the fusion protein for degradation via ubiquitin-dependent proteolysis (19, 20). Thus uncleaved ubiquitin fusions are recognized as abnormal proteins and rapidly degraded.

The second feature of ubiquitin fusions is that they enhance (often dramatically) the yield of the protein to which ubiquitin is fused. This phenomenon was initially observed in yeast, where the natural ubiquitin-ribosomal protein fusion structure was found to enhance the yield of the ribosomal protein (21). Ubiquitin fusions have since been used to improve the yield of recombinant proteins in bacteria, yeast, and insect cells (reviewed in Ref. 22). That the yield enhancement occurs in bacteria, which lack ubiquitin and Ubp enzymes, indicates that this property is intrinsic to ubiquitin, and has been proposed to be due to a covalent chaperone effect of ubiquitin in allowing the fusion partner to fold and/or protect it from degradation

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) L00681 and U20657.

¶ Supported by a QEII Fellowship from the Australian Research Council and a grant from the Australian Capital Territory Cancer Society. To whom correspondence should be addressed. Tel.: 61 2 6249 3824; Fax: 61 2 6249 4712; E-mail: Rohan.Baker@anu.edu.au.

[1] The abbreviations used are: Ub, ubiquitin; Ubp, ubiquitin-specific protease, deubiquitinating enzyme; Unp, mouse ubiquitin-specific protease; Unph, human homolog of Unp, GST, glutathione *S*-transferase; β-gal, β-galactosidase; Ub-*X*-β-gal, ubiquitin-*X*-β-gal fusion, where *X* is one of the 20 amino acids; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis.

[2] R. T. Baker, unpublished data.

(21). The yield enhancement also occurs when a Ubp is co-expressed in bacteria with the fusion protein, which is rapidly cleaved co-translationally, indicating that the protective effect of ubiquitin is exerted in the small period of time that the fusion is intact (23).

We have recently been studying a mouse protein termed Unp, that is an oncoprotein and causes tumors when overexpressed in nude mice (24, 25). Unp, and also its human homolog Unph, both contain the conserved Ubp domains identified in the yeast Ubp family (5), and thus are potential Ubps (6, 26). During the characterization of the Ubp activity of Unp and Unph expressed in yeast and bacteria, we noted that it exhibited significant cleavage activity of the ubiquitin-proline bond. In this report, we characterize this novel cleavage activity.

EXPERIMENTAL PROCEDURES

Plasmids Expressing Ubps, Unp, and Unph—Escherichia coli was cultured and transformed with plasmids according to standard procedures (27). Clones containing Unp and Unph cDNAs have been described previously (24, 26). The 5′ portion of the Unp open reading frame from the initiation codon to the internal *Xba*I site was amplified by the polymerase chain reaction (PCR) using oligonucleotides UNPDAG (5′-d(CTCGAATTCGAGATGGCGGAAGG)) and UNPXBA (5′-d(GGCTTCTTCTTTACGCGGTTCAG)). The product was digested with *Eco*RI and *Xba*I, ligated into pBluescript, and the insert of one clone was sequenced (ThermoSequenase kit, Amersham Corp.) in full to reveal no PCR errors. This fragment and the remainder of the Unp clone 3′ of the *Xba*I site were ligated into the bacterial expression plasmid pKK261 (a derivative of pKK223-3) (23) to reconstruct the Unp open reading frame under control of the *tac* promoter (28), yielding plasmid pCG53. An analogous procedure was used to amplify and reconstruct Unph from a human Unph cDNA clone. Plasmids pCG43 and pCG42 were also initially constructed by analogous procedures, that expressed truncated Unp and Unph proteins, lacking residues 1–168 (UnpΔ1–168) and 1–55 (UnphΔ1–55), respectively. These truncated proteins were derived from deduced incomplete open reading frames due to sequencing errors in the original Unp and Unph sequencing reports (24, 26). Corrected sequences and full details will appear elsewhere[3] and in accession nos. L00681 (Unp) and U20657 (Unph). Plasmids expressing Ubp1p (pJT70), Ubp2p (pRB105, pRB173), Ubp3p (pRB143, pRB175), and Yuh1p (pKKYUH1) have been described previously (4, 5). A plasmid for the expression of Unp in yeast was constructed by ligation of the *Eco*RI-*Hin*dIII fragment from pCG53 downstream of the constitutive *ADH1* promoter (29) in the high copy vector YEplac181 (30).

Plasmids Expressing Ubiquitin-Proline Fusion Proteins—Plasmid pUb23-P expresses a Ub-P-β-gal fusion protein from a hybrid *CYC-GAL10* promoter on a yeast high copy plasmid (15). We have previously described a pKK223-3-based plasmid, pRB269, containing a single ubiquitin-coding region that contains a *Sac*II site at the 3′ end of ubiquitin to enable ligation of other gene fragments (5). A ubiquitin-proline-GSTP1 fusion (Ub-P-GSTP1) was constructed in pRB269 analogously to the ubiquitin-methionine-GSTP1 fusion described previously (5). The wild-type mature N terminus of human GSTP1 (Pro-Pro-Tyr) was mutated to Pro-Ala-Tyr, and a *Sac*II restriction site added, using the forward primer UBPGST (5′-d(GTGCCGCGGTGGTCCGGCGTATACCGTGG)), the reverse primer 3EX2 (5′-d(CTTAAGCTTCCCTCACTGTTTCCCG) (5), a GSTP1 cDNA clone (5), and the PCR. The PCR product was digested with *Sac*II and *Hin*dIII, ligated into pBluescript, sequenced in full to reveal no PCR errors, and ligated into pRB269 digested with *Sac*II and *Hin*dIII to yield pRB481. For coexpression of Ub-P-GSTP1 in *E. coli* with a Ubp-expressing plasmid, the *tac* promoter/Ub-P-GSTP1-rrnB terminator region of pRB481 was excised with *Eag*I and *Sca*I, and ligated into pACYC184 digested with *Eag*I and *Eco*RV to yield pRB486.

Oligonucleotide-directed Mutagenesis—The cysteine residue of Unp that lies within the conserved Ubp Cys domain (C311) was mutated to alanine using the oligonucleotide UNPCA (5′-d(CAGAGTTCATAAAGGCAGTGTTCCCC), an M13 subclone of the Unp 5′ *Eco*RI-*Xba*I fragment, and a *dut ung* mutagenesis kit (Bio-Rad). This oligonucleotide also destroys an *Rca*I (*Bsp*HI) site in this region. Mutagenized clones were screened by *Rca*I digestion, and then sequenced to confirm the correct mutation. The *Eco*RI-*Xba*I fragment containing the C311A mutation was used to replace the corresponding wild-type fragment in pCG53 to yield pCG54, which expresses UnpC311A.

Ubiquitin Cleavage Assays—Cleavage of Ub-P-β-gal fusions in yeast was assayed by pulse-labeling with [^{35}S]methionine, extraction, immunoprecipitation with a monoclonal antibody to β-gal (Promega), electrophoresis in a 6% SDS-polyacrylamide gel (31), and fluorography exactly as described elsewhere (32). For *in vitro* cleavage assays, Ub-P-β-gal was immunoprecipitated from yeast cells containing plasmid pUB23-P, the protein A-agarose pellet washed with TN buffer (50 mM Tris, pH 7.4, 150 mM NaCl) to remove SDS, and portions of the resuspended pellet slurry were mixed with extracts of *E. coli* expressing different Ubps. Where required, β-gal protein-containing bands were excised from dried gels, and the amount of ^{35}S in the bands was determined. Bacterial extracts were prepared from exponentially growing cultures that had been induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside for 3 h. Cells from 1.5-ml cultures were collected, resuspended in 0.2 ml of TN buffer containing 1 mM dithiothreitol and 0.25 mg lysozyme/ml, incubated on ice 10 min, sonicated for two 5-s bursts, and centrifuged for 5 min at 4 °C. The supernatant was collected as the soluble extract. Assays consisted of mixing an extract containing a Ubp (usually 15 μl) with an extract containing the substrate (5 μl) and incubating at 37 °C for 60 min. An antiserum against Ub-M-GSTP1 was raised in rabbits using purified Ub-M-GSTP1 (23). This antisera was used at 1:1000 dilution against immunoblots of cleavage reactions involving Ub-P-GSTP1 resolved in a 12.5% SDS-polyacrylamide gel. Where required, GSTP1-containing proteins were purified by affinity chromatography on GSH-agarose (Sigma) exactly as described previously (23). Protein sequencing was done at the Biomolecular Resource Facility, Australian National University.

RESULTS

Unp and Unph Are Ubps—Unp and Unph are mouse and human proteins, respectively, that contain conserved sequence motifs originally identified in the yeast ubiquitin-specific proteases Ubp1p, Ubp2p, and Ubp3p (5, 6, 24, 26). These domains contain a conserved cysteine residue, and two histidine residues, respectively, that presumably form part of the active site of these thiol proteases (5) (Fig. 1). Mutation of the corresponding cysteine and/or histidine residues in both the yeast Doa4p(Ubp4p), *Drosophila* fat facets Ubp, and yeast Ubp2p abolish their Ubp activity (6, 10).[2] To confirm that the presence of these conserved domains in Unp and Unph correlates with Ubp activity, we expressed these enzymes in *E. coli* and assayed crude extracts for Ubp activity against an artificial linear ubiquitin-GST fusion protein that has a methionine residue immediately following ubiquitin (UbMGSTP1) (23). While *E. coli* extracts alone exhibit no cleavage of the Ub-M-GSTP1 fusion protein, extracts containing either Unp or Unph cleaved a substantial proportion of this ubiquitin-fusion protein, demonstrating that they are bona fide Ubps (data not shown). Furthermore, mutation of the conserved cysteine residue in Unp to alanine (UnpC311A) abolished this cleavage activity (data not shown; see below), consistent with a role for this residue in the active site of the enzyme. A full characterization of the Ubp activity of Unp and Unph will be presented elsewhere.[4]

Unp Efficiently Cleaves the Ubiquitin-Proline Bond in a Ubiquitin-β-Galactosidase Fusion Both in in Vivo and in Vitro—Ubiquitin-*X*-β-galactosidase (Ub-*X*-β-gal) fusion proteins have been routinely used as model substrates of ubiquitin-dependent proteolysis in yeast (*e.g.* see Refs. 15, 19, 20, and 32). During initial experiments where Unp was coexpressed in yeast with Ub-P-β-gal to ascertain any effect of Unp on ubiquitin-dependent proteolysis, we noticed that substantially more cleavage of the ubiquitin-proline bond occurred in cells containing Unp than those without. This effect was further quantitated in a pulse-chase assay (Fig. 2*A*). In the presence of Unp,

[3] M. Di Fruscio, C. A. Gilchrist, R. T. Baker, and D. A. Gray, submitted for publication.

[4] C. A. Gilchrist, P. Blanchette, D. A. Gray, and R. T. Baker, manuscript in preparation.

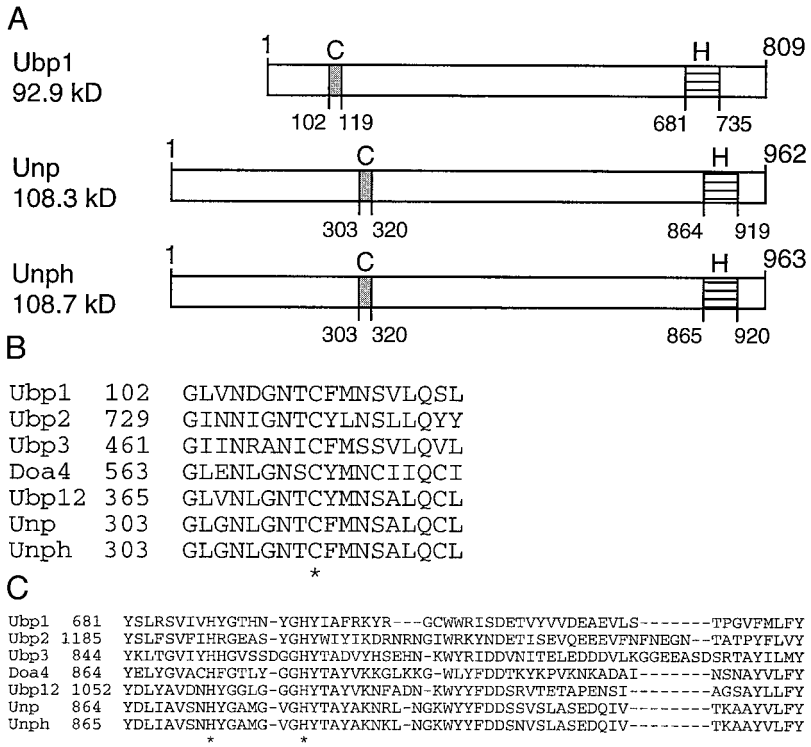

FIG. 1. **Structure and conserved sequence domains of Unp and Unph.** *A*, schematic structural representation of Unp, Unph, and Ubp1p, a yeast Ubp (4). Positions of the conserved sequence domains that contain a Cys residue (*C*), and two His residues (*H*), respectively, are shown. Amino acid residues are numbered. *B* and *C*, amino acid sequence of the conserved Cys (*B*) and His (*C*) domains (5, 6) are shown for Unp, Unph, and several yeast Ubps, in the standard single letter code. Ubp12p is a functional Ubp translated from open reading frame YJL197W on yeast chromosome X, and is the yeast Ubp with the highest sequence similarity to Unp/Unph (R. T. Baker and D. A. Gray, unpublished data) (8).

substantially more of the P-β-gal species is present at the zero time point than in control cells (compare *lanes 1* and *4*). The P-β-gal thus produced is a stable protein, in agreement with the N-end rule (15), and is not degraded over the 30-min chase. Unp does not have an isopeptidase activity in these experiments, in that the multiubiquitin chain assembled on Ub-P-β-gal (33) is not affected by the presence of Unp (Fig. 2*A*).

One explanation for the apparent enhanced cleavage at the ubiquitin-proline bond is that Unp may have a relaxed or degenerate cleavage specificity, and may be cleaving at an incorrect site other than the predicted glycine-proline junction. To test this, we used a Ub-*X*-β-gal fusion protein, Ub^{V76}-V-β-gal, where the Gly^{76} residue of ubiquitin had been mutated to valine, and the *X* residue was also valine. This mutation prevents cleavage by all known yeast Ubp enzymes, because when this fusion is expressed in yeast, no ubiquitin cleavage is observed, and the intact fusion is targeted for ubiquitin-dependent degradation via a pathway that recognizes noncleavable ubiquitin fusions (19, 20). Yuh1p, a different yeast ubiquitin-specific protease, is also unable to cleave the ubiquitin-proline bond in Ub-P-β-gal fusions (12). The presence of Unp did not result in any cleavage of this G76V mutant (Fig. 2*A*), indicating that Unp also requires the Gly^{76} residue for cleavage, and does not appear to have a relaxed cleavage site specificity.

To test whether Unp was capable of cleaving Ub-P-β-gal only when co-expressed with it *in vivo*, we performed an *in vitro* cleavage assay, whereby Ub-P-β-gal was immunoprecipitated from yeast cells grown with [^{35}S]methionine and then incubated with extracts of *E. coli* expressing different Ubps. As shown in Fig. 2*B*, an *E. coli* extract alone had no effect on Ub-P-β-gal, whereas extracts containing Unp resulted in 63% of the Ub-P-β-gal-containing species (including the multiubiquitinated species) being converted to P-β-gal (*lane 5*). In wild-type yeast, only 9% of the ^{35}S in β-gal-containing species was present as P-β-gal (*lane 1*). Given that Unp does not have an isopeptidase activity (see above), it appears to be able to cleave the ubiquitin-proline bond even when a large multiubiquitin chain is attached to the ubiquitin moiety. The yeast Ubps Ubp1p (*lane 3*), Ubp2p (*lane 4*), or Ubp3p (not shown) were not able to cleave the ubiquitin-proline bond in this assay, consistent with the poor cleavage of Ub-P-β-gal in yeast, and indicative of the unique cleavage specificity of Unp. (The Ubp1p and Ubp2p extracts used here were able to cleave a Ub-M-GSTP1 fusion protein *in vitro* (data not shown) (5). Ubp3p does not appear to have Ubp activity *in vitro* (5).) Notably, this assay revealed that Ubp1p has an apparent isopeptidase activity, in that it removed the bulk of the high molecular weight ubiquitin conjugates on Ub-P-β-gal (*lane 3*). This activity is being further investigated.

Unp and Unph Precisely Cleave at the Ubiquitin-Proline Junction in Vivo and in Vitro—We have previously used the ubiquitin fusion technique to express human glutathione *S*-transferase GSTP1 (23). While the mature N terminus of GSTP1 is proline, we were unable to use the ubiquitin fusion technique to expose proline at the N terminus of GSTP1 due to the poor efficiency of cleavage at this bond by known Ubps, and so retained the initiator methionine codon in the fusion (23). Given the ability of Unp to cleave this bond in Ub-P-β-gal, we constructed a Ub-P-GSTP1 fusion and tested its cleavage with Unp and Unph (see "Experimental Procedures"). For the purposes of this experiment, we mutated the wild type N terminus of GSTP1 (Pro-Pro-Tyr) (34, 35) to Pro-Ala-Tyr, so that the subsequent N-terminal sequencing to determine the cleavage site would not be ambiguous.

To assay *in vivo* cleavage of Ub-P-GSTP1, various Ubps were co-expressed in *E. coli* with Ub-P-GSTP1, cell extracts were resolved by SDS-PAGE, and an immunoblot was performed with a rabbit antiserum raised against purified Ub-M-GSTP1 (23). This antiserum detects uncleaved Ub-P-GSTP1, P-GSTP1, and (weakly) free ubiquitin. As shown in Fig. 3*A*, approximately 40% of the Ub-P-GSTP1 fusion was cleaved when co-expressed with full-length Unp (*lane 8*), while the truncated Unp (UnpΔ1–168) and Unph (UnphΔ1–55) enzymes also exhibited substantial cleavage (*lanes 5* and *6*). However, no detectable cleavage occurred with either Ubp1p (*lane 2*), Ubp2p (*lane 3*), Ubp3p (*lane 4*), or notably with the UnpC311A mutant (*lane 9*).

To assay *in vitro* cleavage of Ub-P-GSTP1, extracts of *E. coli*

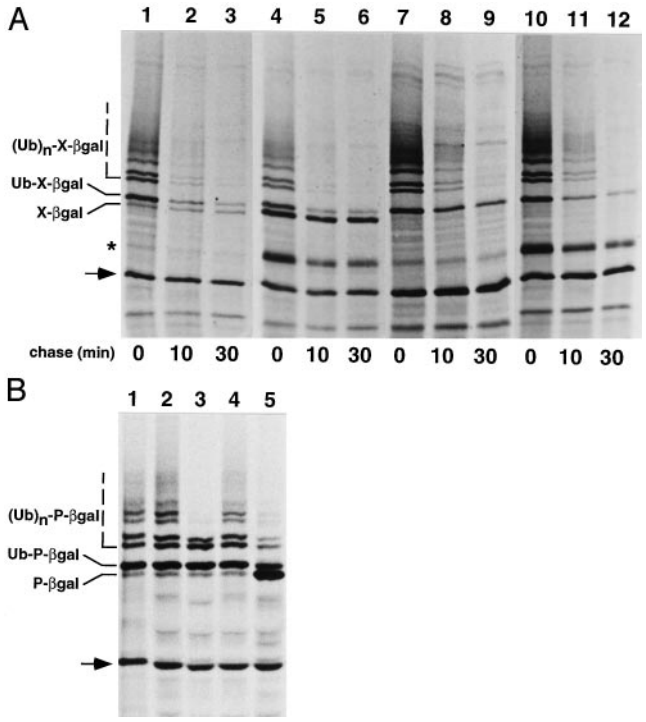

FIG. 2. **Cleavage of the ubiquitin-proline bond in Ub-P-β-gal by Unp.** *A*, *in vivo* assays. Yeast expressing either Ub-P-β-gal (*lanes 1–6*) or Ub^{V76}-V-β-gal (*lanes 7–12*) and either no additional Ubp (*lanes 1–3* and *7–9*) or Unp (*lanes 4–6* and *10–12*) were pulse-labeled with [^{35}S]methionine for 5 min, and chased with unlabeled methionine for either 0, 10, or 30 min, as indicated *below* the lanes. Extracts were immunoprecipitated with a monoclonal antibody to β-gal, resolved by SDS-PAGE, and fluorographed (see "Experimental Procedures"). Bands representing cleaved *X*-β-gal, uncleaved Ub-*X*-β-gal, and Ub-*X*-β-gal species bearing a multiubiquitin chain ($(Ub)_n$-*X*-β-gal) are indicated on the *left*. An *arrow* indicates a ~90-kDa degradation product of β-gal (32). An asterisk (*) indicates a ~95-kDa degradation product of β-gal usually observed with stable β-gal proteins (32), which appears to be enhanced by the presence of Unp. *B*, *in vitro* assay. Ub-P-β-gal bearing a multiubiquitin chain was immunoprecipitated from yeast cells in *lane 1*, *panel A*, the protein A-Sepharose pellet washed to remove SDS, divided into equal aliquots, and incubated with nothing (*lane 1*), *E. coli* extract (*lane 2*), or extracts of *E. coli* expressing either Ubp1 (*lane 3*), Ubp2 (*lane 4*), or Unp (*lane 5*). Other procedures and designations are as in *panel A*. Following fluorography, bands containing P-β-gal, Ub-P-β-gal, and $(Ub)_n$-P-β-gal were excised from the gel, and the amount of ^{35}S was determined by scintillation counting.

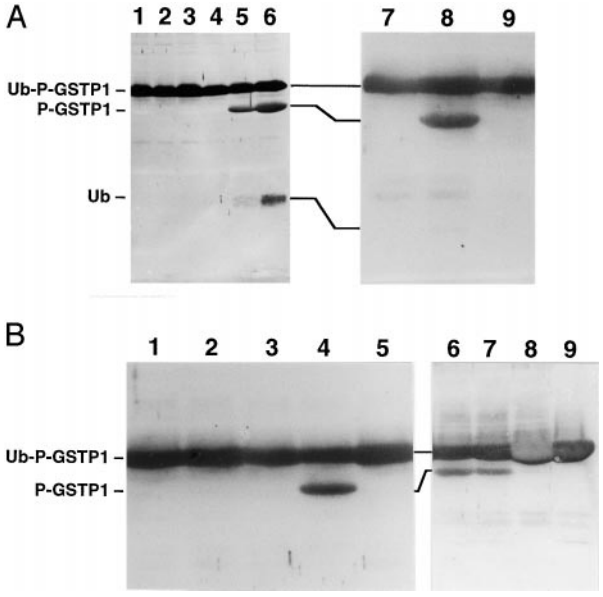

FIG. 3. **Cleavage of the ubiquitin-proline bond in Ub-P-GSTP1 by Unp and Unph.** *A*, *in vivo* assays. Extracts of *E. coli* cells containing a plasmid expressing Ub-P-GSTP1 and a second plasmid expressing either no Ubp (*lanes 1* and *7*), Ubp1p (*lane 2*), Ubp2p (*lane 3*), Ubp3p (*lane 4*), UnpΔ1–168 (*lane 5*), UnphΔ1–55 (*lane 6*), Unp (*lane 8*), or UnpC311A (*lane 9*) were resolved by SDS-PAGE in a 15% gel, electroblotted to a polyvinylidene difluoride membrane, and immunodetected with an antibody against Ub-M-GSTP1. Bands containing Ub-P-GSTP1, P-GSTP1, and Ub, are indicated on the *left*. *B*, *in vitro* assays. Extracts of *E. coli* cells expressing Ub-P-GSTP1 were mixed with extracts of *E. coli* cells expressing either no Ubp (*lane 1*), Ubp1p (*lane 2*), Ubp2p (*lane 3*), Unp (*lane 4*), UnpC311A (*lane 5*), UnpΔ1–168 (*lane 6*), UnphΔ1–55 (*lane 7*), Yuh1p (*lane 8*), or purified Yuh1p (*lane 9*), incubated at 37 °C for 1 h, and electrophoresed and immunoblotted as above. In *lane 8*, the abundant Yuh1p in the crude extract has displaced the Ub-P-GSTP1 band.

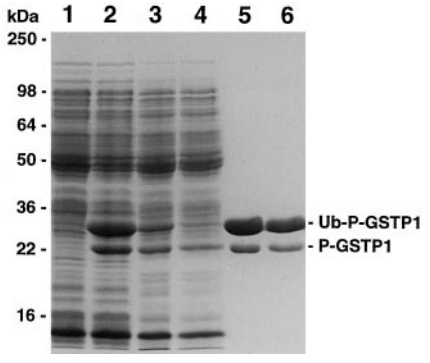

FIG. 4. **Purification of P-GSTP1 for N-terminal sequencing.** Samples from sequential steps in the purification of GSTP1-containing proteins were resolved by SDS-PAGE in an 11% gel, and stained with Coomassie Blue (see "Experimental Procedures"). *Lane 1*, crude extract of *E. coli* expressing Unp from pCG53; *lane 2*, crude extract of *E. coli* expressing Ub-P-GSTP1 from pRB486; *lane 3*, aliquot following mixture of extracts in *lanes 2* and *3* and incubation for 1 h at 37 °C; *lane 4*, proteins unbound to GSH-agarose; *lane 5*, first elution fraction with 5 mM GSH; *lane 6*, second elution fraction with 5 mM GSH. The positions of Ub-P-GSTP1 and P-GSTP1 are indicated on the *right*. The positions of molecular weight markers and their mass in kilodaltons (kDa) are indicated on the *left*. Bands containing P-GSTP1 were excised from the gel and sequenced. *Lanes 2*, *3*, and *4* contain chloramphenicol acetyl transferase, a 25-kDa protein expressed from the chloramphenicol acetyl transferase gene on pRB286 (pACYC184), that migrates near P-GSTP1 (23 kDa) but does not bind to GSH-agarose, and is not recognized by an anti-GSTP1-antibody (*e.g.* Fig. 3*B*, *lane 1*).

expressing Ub-P-GSTP1 were mixed with extracts of *E. coli* expressing different Ubps, and following incubation at 37 °C for 1 h, electrophoresis and immunodetection were performed as above. Similar results to the *in vivo* cleavage assays were observed (Fig. 3*B*), with full-length Unp cleaving approximately 40% of Ub-P-GSTP1 (*lane 4*), while the truncated Unp and Unph enzymes exhibited partial activity (*lanes 6* and *7*). No detectable cleavage was observed with either Yuh1p (*lanes 8* and *9*), Ubp1p (*lane 2*), Ubp2p (*lane 3*), or *E. coli* extract (*lane 1*). Each of these extracts (except for Yuh1p) exhibited cleavage activity against a Ub-M-GSTP1 fusion protein (data not shown). Notably, the UnpC311A mutant was again devoid of activity (*lane 5*).

To determine the fidelity of Unp and Unph-mediated cleavage of the ubiquitin-proline bond, the GSTP1-containing proteins arising from scaled-up cleavage reactions of UnpΔ1–168 cleavage of Ub-P-GSTP1 *in vivo* (Fig. 3*A*, *lane 5*), UnphΔ1–55 cleavage of Ub-P-GSTP1 *in vivo* (Fig. 3*A*, *lane 6*), and full-length Unp cleavage of Ub-P-GSTP1 *in vitro* (Fig. 3*B*, *lane 4*) were purified by glutathione affinity chromatography on GSH-agarose, eluted with 5 mM GSH, and resolved by SDS-PAGE. A representative purification is shown in Fig. 4 (full-length Unp cleavage of Ub-P-GSTP1 *in vitro*). The Coomassie-stained P-GSTP1-sized bands were excised from the gel (Fig. 4, *lanes 5* and *6*), eluted onto a polyvinylidene difluoride membrane, and

TABLE I
N-terminal sequencing of Unp and Unph cleavage products

The encoded sequence of Ub-P-GSTP1 is given, showing the C-terminal four residues of ubiquitin (Leu73-Gly76) and the expected site of Ubp cleavage (↓):

```
  P4   P3   P2   P1          P1′  P2′  P3′  P4′  P5′  P6′
Leu-Arg-Gly-Gly ↓       Pro-Ala-Tyr-Thr-Val-Val
 73   74   75   76
```

Ubp	Cleavage*	Observed sequence
UnpΔ1–168	*In vivo*	Pro-Ala-Tyr-Thr-Val-Val
UnphΔ1–55	*In vivo*	Pro-Ala-Tyr-Thr-Val-Val
Unp	*In vitro*	Pro-Ala-Tyr-Thr-Val-Val

* Cleavage occurred either *in vivo* (co-expression of Ub-P-GSTP1 and Ubp in *E. coli*) or *in vitro* (mixing of an extract of *E. coli* expressing Ub-P-GSTP1 and an extract of *E. coli* expressing Unp).

subjected to six cycles of automated N-terminal sequence analysis on a Applied Biosystems 494 Procise. The sequences obtained are listed in Table I. In all three cases, proline was the only residue detected at the N terminus, and the sequence obtained matched the expected sequence exactly. Neither glycine nor alanine was observed in the first cycle, which would be expected if cleavage had occurred either one residue N-terminal or C-terminal, respectively, of the glycine-proline bond. From this analysis, we conclude that at least 95% of the cleavage had occurred at the glycine-proline bond. Thus in contrast to other Ubps, Unp and its human homolog Unph exhibit very efficient and precise cleavage of the ubiquitin-proline bond in ubiquitin fusion proteins.

DISCUSSION

The ubiquitin-fusion technique has proven to be a versatile method of increasing the yield of recombinant protein expression, and as a way of synthesizing proteins with any desired N terminus, both for biotechnological and research applications (reviewed in Ref. 22). One limitation in this technique has been the extremely inefficient cleavage of the ubiquitin-proline bond, and thus proteins with an N-terminal proline residue have been excluded. For example, the status of N-terminal proline in the *E. coli* N-end rule has not been determined due to the inability of Ubp1p to deubiquitinate Ub-P-β-gal (36). Our demonstration that the mouse Ubp enzyme Unp, and its human homolog Unph, can efficiently and precisely cleave the ubiquitin-proline bond either *in vitro* or *in vivo*, allows circumvention of this problem, and improves the versatility of the ubiquitin fusion technique. Unp and Unph exhibit this activity irrespective of the size of the fusion partner (1045 residues in Ub-P-β-gal, 209 residues in Ub-P-GSTP1), and do not appear to be affected by the residue immediately following proline in the P-2′ position (His in Ub-P-β-gal, Ala in Ub-P-GSTP1). This activity is unique to Unp and Unph among known Ubps; cleavage of Ub-P-β-gal by Ubps present in yeast or in rabbit reticulocyte extract is very inefficient (15, 17), and recombinant yeast Ubps Ubp1p, Ubp2p, Ubp3p, and Yuh1p lack this activity (see "Results"). In addition, recombinant yeast Ubps Doa4p/Ubp4p, Ubp6p, Ubp12p, and Ubp15p cannot cleave the ubiquitin-proline bond (data not shown). Ubp12p is the most closely related yeast Ubp to Unp and Unph, and its inability to cleave this bond implies that yeast lack an efficient functional homolog of Unp/Unph with respect to ubiquitin-proline bond cleavage. While proline is only one of the 20 amino acids that may occur at the N terminus of a protein or peptide, several important proteins do have an N-terminal proline residue, including GSTP1, which is strongly induced in early stages of hepatocarcinogenesis in rats (reviewed in Hayes and Pulford (37)). Several biologically active peptides also have an N-terminal proline, such as angiogenesis inhibitor platelet factor 4 (38), basic fibroblast growth factor (39, 40), neuropeptide F (41), and neuropeptide Y derivatives (42). While organic synthesis is efficient for the production of short peptides *in vitro*, the ubiquitin fusion technique provides an efficient alternative for the synthesis of peptides of any length *in vivo* (43), and the ability of Unp and Unph to cleave the ubiquitin-proline bond enhances this technique.

Although the ability of Unp and Unph to cleave the ubiquitin-proline bond is a versatile research and biotechnology tool, the biological significance of this cleavage activity is unclear, as no natural ubiquitin fusions containing the ubiquitin-proline bond have been observed. It is likely that this cleavage activity reflects a unique property of Unp and Unph, in that they can access and cleave the C terminus of ubiquitin in fusions that other Ubps cannot. Presumably this reflects the structure of the natural substrate(s) of Unp/Unph, which may require a special catalytic mechanism of the Ubp that cleaves them. The unique structure of proline, with the side chain cyclized onto the backbone nitrogen atom, may sterically shield the bond to be cleaved from all Ubps other than Unp/Unph. It is also possible that Unp may cleave some or all of the ubiquitin-like proteins that are processed from linear precursors and/or post-translationally formed conjugates, such as the ubiquitin-like protein fused to ribosomal protein S30 (44) that is apparently conjugated to an immune suppressor factor (45), and the small ubiquitin-like modification (SUMO) of the Ran-GTPase-activating protein (46, 47). An ability to cleave ubiquitin-like proteins may signify a "relaxed" active site conformation, which may be able to tolerate the proline residue, whereas other Ubps cannot. In this respect, we note that both Unp and Unph have a glycine positioned six residues on the N-terminal side of the active site cysteine within the conserved Cys domain, whereas no other known Ubp from any species has a glycine in this position, having instead much bulkier residues (Fig. 1*B*) (8) (data not shown). Although we have no information on the structure that this domain adopts, the small glycine residue may allow toleration of the proline residue in the P-1′ position. One testable prediction of this model is that mutation of this Unp/Unph glycine to a non-glycine residue should greatly reduce its ability to cleave the ubiquitin-proline bond.

It is also noteworthy that in any ubiquitin-fusion where the P-1′ residue is not proline, be it a linear fusion or an isopeptide linkage, there is a proton on the nitrogen atom of the P-1′ residue, whereas when the P-1′ residue is proline, this proton is absent. It is possible that this proton is required for the catalytic mechanism of most Ubps, which cleave very inefficiently if this proton is absent, and that Unp and Unph have an alternate mechanism and/or can utilize a different proton, and thus can cleave fusions where proline is in the P-1′ position. In the model thiol protease papain, it has been proposed that the NH of the P-1′ leaving group is involved in hydrogen bonding with a carbonyl group of the papain peptide backbone, to aid in distortion of the scissile C–N bond to increase its electrophilicity and thus facilitate attack by the thiol group (48). The nitrogen of the P-1′ leaving group also has a role in the breakdown of the tetrahedral intermediate linking papain to its substrate, becoming protonated in the transition state (49). If the Ubps operate by an analogous mechanism, then the unusual structure of proline in the P-1′ position may not allow either of these roles to operate. In either case, it would appear that Unp and Unph have a different catalytic mechanism than other Ubps. The exploitation of this mechanism to cleave the ubiquitin-proline bond should prove fruitful for both biotechnology and research applications.

Acknowledgments—We thank Dr. Dennis Shaw for assistance and advice with protein sequencing, and Drs. Gareth Chelvanayagam and

Peter Milburn for advice on protein chemistry. We also thank a reviewer for helpful comments.

REFERENCES

1. Schlesinger, M. J., and Bond, U. (1987) *Oxf. Surv. Eukaryotic Genes* **4,** 77–91
2. Hershko, A., and Ciechanover, A. (1992) *Annu. Rev. Biochem.* **61,** 761–807
3. Hochstrasser, M. (1995) *Curr. Opin. Cell Biol.* **7,** 215–223
4. Tobias, J. W., and Varshavsky, A. (1991) *J. Biol. Chem.* **266,** 12021–12028
5. Baker, R. T., Tobias, J. W., and Varshavsky, A. (1992) *J. Biol. Chem.* **267,** 23364–23375
6. Papa, F., and Hochstrasser, M. (1993) *Nature* **366,** 313–319
7. Wilkinson, K. D., Tashayev, V. L., O'Connor, L. B., Larsen, C. N., Kasperek, E., and Pickart, C. M. (1995) *Biochemistry* **34,** 14535–14546
8. Hochstrasser, M. (1996) *Annu. Rev. Genet.* **30,** 405–439
9. Amerik, Alexander Yu., Swaminathan, S., Krantz, B. A., Wilkinson, K. D., and Hochstrasser, M. (1997) *EMBO J.* **16,** 4826–4838
10. Huang, Y., Baker, R. T., and Fischer-Vize, J. (1995) *Science* **270,** 1828–1831
11. Everett, R. D., Meredith, M., Orr, A., Cross, A., Kathoria, M., and Parkinson, J. (1997) *EMBO J.* **16,** 566–577
12. Miller, H. I., Henzel, W. J., Ridgway, J. B., Kuang. W.-J., Chisolm, V., Liu, C.-C. (1989) *BioTechnology* **7,** 698–704
13. Wilkinson, K. D., Lee, K., Deshpande, S., Duerksen-Hughes, P., Boss, J. M., and Pohl, J. (1989) *Science* **246,** 670–673
14. Larsen, C. N., Price, J. S., and Wilkinson, K. D. (1996) *Biochemistry* **35,** 6735–44
15. Bachmair, A., Finley, D., and Varshavsky, A. (1986) *Science* **234,** 179–186
16. Bachmair, A., and Varshavsky, A. (1989) *Cell* **56,** 1019–1032
17. Gonda, D. K., Bachmair, A., Wünning, I., Tobias, J. W., Lane, W. S., and Varshavsky, A. (1989) *J. Biol. Chem.* **264,** 16700–16712
18. Varshavsky, A. (1992) *Cell* **69,** 725–735
19. Johnson, E. S., Bartel, B., Seufert, W., and Varshavsky, A. (1992) *EMBO J.* **11,** 497–505
20. Johnson, E. S., Ma, P. C. M., Ota, I. M., and Varshavsky, A. (1995) *J. Biol. Chem.* **270,** 17442–17456
21. Finley, D., Bartel, B., and Varshavsky, A. (1989) *Nature* **338,** 394–401
22. Baker, R. T. (1996) *Curr. Opin. Biotechnol.* **7,** 541–546
23. Baker, R. T., Smith S. A., Marano R., McKee J., and Board P. G. (1994) *J. Biol. Chem.* **269,** 25381–25386
24. Gupta, K., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Gray, D. A. (1993) *Oncogene* **8,** 2307–2310
25. Gupta, K., Chevrette, M., and Gray, D. A. (1994) *Oncogene* **9,** 1729–1731
26. Gray, D. A., Inazawa, J., Gupta, K., Wong, A., Ueda, R., and Takahashi, T. (1995) *Oncogene* **10,** 2179–2183
27. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
28. Brosius, J., and Lupsky, J. R. (1987) *Methods Enzymol.* **153,** 54–68
29. Ammerer, G. (1983) *Methods Enzymol.* **101,** 192–201
30. Gietz, R. D., and Sugino, A. (1988) *Gene* (*Amst.*) **74,** 527–534
31. Laemmli, U. K. (1970) *Nature* **227,** 680–685
32. Baker, R. T., and Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 1090–1094
33. Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D., Gonda, D. K., and Varshavsky, A. (1989) *Science* **243,** 1576–1583
34. Ålin, P., Mannervik, B., and Jörnvall, H. (1985) *FEBS Lett.* **182,** 319–322
35. Suzuki, T., Coggan, M., Shaw, D. C., and Board, P. G. (1987) *Ann. Hum. Genet.* **51,** 95–106
36. Tobias, J. W., Shrader, T. E., Rocap, G., and Varshavsky, A. (1991) *Science* **254,** 1374–1377
37. Hayes, J. D., and Pulford, D. J. (1995) *Crit. Rev. Biochem. Mol. Biol.* **30,** 445–600
38. Maione, T. E., Gray, G. S., Petro, J., Hunt, A. J., Donner, A. L., Bauer, S. I., Carson, H. F., and Sharpe, R. J. (1990) *Science* **247,** 77–79
39. Abraham, J. A., Whong, J. L., Tumolo, A., Mergia, A., Friedman, J., Gospodarowicz, D., and Fiddes, J. C. (1986) *EMBO J.* **5,** 2523–2528
40. Esch, F., Baird, A., Ling, N., Ueno, N., Hill, F., Denoroy, L., Klepper, R., Gospodarowicz, D., Böhlen, P., Guillemin, R. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 6507–6511
41. Maule, A. G., Shaw, C., Halton, D. W., Thim, L., Johnston, C. F., Fairweather, I., and Buchanan, K. D. (1991) *Parasitology* **102,** 309–316
42. Taoudi-Benchekroun, M., St.-Pierre, S., Fournier, A., and Cadieux, A., (1993) *Br. J. Pharmacol.* **109,** 902–904
43. Yoo, Y., Rote, K., and Rechsteiner, M. (1989) *J. Biol. Chem.* **264,** 17078–17083
44. Baker R. T., Williamson, N. A., and Wettenhall, R. E. H. (1996) *J. Biol. Chem.* **271,** 13549–13555
45. Nakamura, M., Xavier, R. M., Tsunematsu, T., and Tanigawa, Y. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 3463–3467
46. Matunis, M. J., Coutavas, E., and Blobel, G. (1996) *J. Cell Biol.* **135,** 1457–1470
47. Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997) *Cell* **88,** 97–107
48. Lowe, G., and Yuthavong, Y. (1971) *Biochem. J.* **124,** 107–115
49. Lowe, G., and Yuthavong, Y. (1971) *Biochem. J.* **124,** 117–122