Ubiquitin-specific Proteases of *Saccharomyces cerevisiae*

CLONING OF UBP2 AND UBP3, AND FUNCTIONAL ANALYSIS OF THE UBP GENE FAMILY*

(Received for publication, May 29, 1992)

Rohan T. Baker‡, John W. Tobias‡, and Alexander Varshavsky¶

From the Division of Biology, California Institute of Technology, Pasadena, California 91125

In eukaryotes, both natural and engineered ubiquitin (Ub) fusions to itself or other proteins are cleaved by processing proteases after the last (Gly\textsuperscript{95}) residue of ubiquitin. YUH1 and UBP1, the genes for two ubiquitin-specific proteases of the yeast *Saccharomyces cerevisiae*, have been cloned previously and shown to encode nonhomologous proteins. Using an *Escherichia coli*-based genetic screen, we have isolated two other yeast genes for ubiquitin-specific proteases, named *UBP2* and *UBP3*. Ubp2 (1,264 residues), Ubp3 (912 residues), and the previously cloned Ubp1 (809 residues) are largely dissimilar except for two short regions containing Cys and His which encompass their putative active sites. Neither of these proteases has sequence similarities to Yuh1. Both Ubp2 and the previously identified Ubp1 cleave *in vitro* at the C terminus of the ubiquitin moiety in natural and engineered fusions irrespective of their size, poly-Ub being the exception. However, both Ubp1 and Ubp2 are also capable of cleaving poly-Ub when coexpressed with it in *E. coli*, suggesting that such cleavage is largely cotranslational. Although inactive in *E. coli* extracts, Ubp3 was active with all of the tested ubiquitin fusions except poly-Ub when coexpressed with them in *E. coli*. Null yuh1 ubp1 ubp2 ubp3 quadruple mutants are viable and retain the ability to deubiquitinate ubiquitin fusions, indicating the presence of at least one more ubiquitin-specific processing protease in *S. cerevisiae*.

Ubiquitin, a highly conserved 76-residue protein, is present in eukaryotic cells either free or covalently joined to a great variety of proteins. The posttranslational coupling of ubiquitin to other proteins is catalyzed by a family of ubiquitin-specific ATP-independent proteases capable of cleaving ubiquitin from its relatively short C-terminal extensions but is virtually inactive with larger fusions such as ubiquitin-β-galactosidase (Ub-βgal). Wilkinson et al. (1989) have also cloned a cDNA encoding a mammalian homolog of the yeast Yuh1 protease. Tobias and Varshavsky (1991) reported the cloning and functional analysis of another yeast gene, named *UBP1*, which encodes a ubiquitin-specific processing protease whose amino acid sequence is dissimilar to those of the Yuh1 protease and other known proteins.

We used an *E. coli*-based genetic screen to isolate two other genes that encode ubiquitin-specific proteases of *S. cerevisiae*, thereby bringing to four the number of such proteases known in yeast. We also report a functional analysis of the *UBP* gene family that indicates the presence of at least one more ubiquitin-specific protease in *S. cerevisiae*.

*This work was supported by National Institutes of Health Grants AG08891 and GM31330 (to A. V.). 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.

† Supported by fellowships from the Life Sciences Research Foundation and the Fulbright Foundation. Present address: Molecular Genetics Group, John Curtin School of Medical Research, Australian National University, P.O. Box 334, Canberra, ACT 2601, Australia.

‡ Present address: Dept. of Microbiology, Harvard Medical School, Boston, MA 02115.

§ To whom correspondence should be addressed: Division of Biology, 147-75, Caltech, Pasadena, CA 91125. Tel.: 818-505-3785; Fax: 818-440-9821.

1 The abbreviations used are: Ub, ubiquitin; kb, kilobase(s) or kilobase pair(s); hp, base pair(s); ORF, open reading frame; DHFR, mouse dihydrofolate reductase; Ub-X-DHFR, ubiquitin-X-DHFR, where X is an amino acid residue at the Ub-DHFR junction; βgal, *E. coli* β-galactosidase; Ub-X-βgal, ubiquitin-X-βgal; PAGE, polyacrylamide gel electrophoresis; XGal, 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside; IPTG, isopropyl-1-thio-β-D-galactopyranoside.
Ubiquitin-specific proteases cleave ubiquitin from its C-terminal extensions irrespective of the identity of the extension's residue abutting the cleavage site (proline being the single exception) (Bachmair et al., 1986; Gonda et al., 1989). This property, together with high activity, selectivity, and robustness of at least some of these enzymes, should make them preferred reagents for site-specific proteolytic processing of engineered protein fusions. It also allows the in vivo or in vitro generation of polypeptides bearing predetermined N-terminal residues, a method with applications in basic research and biotechnology (Varshavsky, 1992; Lu et al., 1990; Butt et al., 1989; Sabin et al., 1989; Ecker et al., 1989; Yoo et al., 1989; Mak et al., 1989).

**EXPERIMENTAL PROCEDURES**

**Strains, Media, and Genetic Techniques—**E. coli strains MC1061, JM101, and DH5αF′ (Ausubel et al., 1989), used as hosts for plasmids and M13 phage derivatives, were grown in Luria Broth or in M9 minimal medium in the presence of xanthine antibiotics and anti- biotics as required. S. cerevisiae strains (Table I) were grown in rich (YPD) or synthetic (SD) media prepared according to Sherman et al. (1986). Transformations of S. cerevisiae and E. coli were carried out, respectively, by the lithium acetate method (Ito et al., 1983; Baker, 1989), and by the calcium phosphate method (1) in 0.1 M calcium phosphate buffer (Ausubel et al., 1989). The yeast genomic DNA library RB237, used for the isolation of UBP2 and UBP3, has been produced by cloning a partial Sau3AI digest of S. cerevisiae genomic DNA into the BamHI site of the YRp35 vector (Rose et al., 1987). Yeast mating, sporulation, and tetrad analyses were carried out as described (Sherman et al., 1986; Guthrie and Fink, 1991).

**DNA Manipulation and Sequencing—**Standard procedures involving recombining DNA were carried out as described by Ausubel et al. (1989). For nucleotide sequencing, DNA fragments produced by restriction endonucleases were subcloned into M13mp18 and M13mp19 phages and processed for hybridization as described by Finley et al. (1983), followed by electrophoresis in a formaldehyde-containing agarose gel as described by Finley et al. (1989). Fractionated RNA was blotted onto Genescreen in 25 mM sodium phosphate (pH 6.5) and hybridized (Ausubel et al., 1989; 1990). DNA probes were labeled using [α-32P]dCTP and the method of Feinberg and Vogelstein (1986). A Southern blot of yeast chromosomal DNA was hybridized with a 1-kb BamHI genomic DNA fragment containing the yeast LEU2 gene. First, pYUHI was digested with SalI and PstI, and the large (~3.7-kb) fragment was ligated to the ~2.2-kb SalI/PstI fragment containing the S. cerevisiae LEU2 gene to yield pYUHI::LEU2. The LEU2-containing fragment was isolated from pRB1, which was constructed by subcloning the ~2.2-kb SalI/XhoI LEU2-containing fragment of YEpl3 (Brock et al., 1979) into the SalI site of pUC19, in the orientation that yielded LEU2 upon subsequent digestion with SalI and PstI, which were named YRB21 and YRB22, respectively. The resulting plasmid, pKYYH2, was sequenced as described by Finley et al. (1989). Further sequencing analysis of several Leu' transformants (expected to be heterozygous at the YUHI locus) confirmed the predicted structure of the yuh1 allele by showing the presence of restriction sites diagnostic of the transplacement (data not shown) (Robinson et al., 1981). One Leu' transformant, named YRB101, was sequenced and subjected to nucleotide sequencing analysis. In all of the 16 (complete) tetrad examined, the LEU2 and URA3 genes (markers for the yuh1 and ubp1 null alleles, respectively) each segregated 2:2. Also, ~25% of the haploid segregants were double (yuh1 ubp1) mutants, as expected for unlinked genes.

To construct the plasmid pKYYH1, which expressed YUHI in E. coli, a YUHI-containing Sau3AI/EcoRI fragment of pYUHI was blunted with Klenow polymerase I and dNTPs, followed by its ligation to the Smal-digested pKK223-3, in the orientation that placed the start codon of YUHI adjacent to the Pm promoter and 60 bp downstream of the Shine-Dalgarno sequence (see below), yielding pKHHUb2. The start codon of the ubiquitin coding region in these plasmids is 36 bp downstream of the Shine-Dalgarno sequence in pKK223-3 (Brosius and Lupsky, 1987).

**Cloning, Deletion, and Heterologous Expression of YUHI—**To clone YUHI from a M13-mer oligodeoxynucleotides isolated from the published sequence of YUHI (Miller et al., 1989), were synthesized. The first oligodeoxynucleotide, Y1, 5' (CGGAAAAAAACCTACATGACC) 3', spanned an HindIII site upstream of YUHI, while the sequence of the second oligodeoxynucleotide, Y2, 5' (CGGATCCTTTTGCACATTGG) 3', was complementary to the sense strand of YUHI over the final 23 bp of the reported sequence of the YUHI locus (which terminated at a BamHI site); two additional C residues at the 5' end of Y2 were added to facilitate the subsequent digestion with BamHI. Both oligodeoxynucleotides (25 pmol each) were annealed to 35 ng of S. cerevisiae genomic DNA from the strain YC1 (1.5 μg/ml KCl; pH 8.3); and 0.2 mM each of dGTP, dATP, dCTP, and dTTP. Taq DNA polymerase (25 units, Bethesda Research Laboratories) was then added, and the sample was subjected to 30 cycles of the polymerase chain reaction, each round consisting of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 60 s. Agarose gel electrophoresis analysis of the polymerase chain reaction products showed a single fragment of the expected size (~1.7 kb; Miller et al., 1989). The fragment was digested with BamHI and HindIII and subcloned into BamHI/HindIII-digested pUC19, yielding pYUHI.

A yuh1 null allele was made in vitro by replacing 68% of the YUHI coding region and 37 bp of its 3'-flanking region with the yeast LEU2 gene. First, pYUHI was digested with Sall and PstI, and the large (~3.7-kb) fragment was ligated to the ~2.2-kb Sall/PstI fragment containing the S. cerevisiae LEU2 gene to yield pYUHI::LEU2. The LEU2-containing fragment was isolated from pRB1, which was constructed by subcloning the ~2.2-kb SalI/XhoI LEU2-containing fragment of YEpl3 (Broach et al., 1979) into the SalI site of pUC19, in the orientation that yielded LEU2 upon subsequent digestion with Sall and PstI, which were named YRB237 and YRB238, respectively. The resulting plasmid, pKYYH1, was sequenced and subjected to nucleotide sequencing analysis. In all of the 16 (complete) tetrad examined, the LEU2 and URA3 genes (markers for the yuh1 and ubp1 null alleles, respectively) each segregated 2:2. Also, ~25% of the haploid segregants were double (yuh1 ubp1) mutants, as expected for unlinked genes.

To construct the plasmid pKYYH1, which expressed YUHI in E. coli, a YUHI-containing Sau3AI/EcoRI fragment of pYUHI was blunted with Klenow polymerase I and dNTPs, followed by its ligation to the Smal-digested pKK223-3, in the orientation that placed the start codon of YUHI adjacent to the Pm promoter and 60 bp downstream of the Shine-Dalgarno sequence (see above), yielding pKHHUb2. The start codon of the ubiquitin coding region in these plasmids is 36 bp downstream of the Shine-Dalgarno sequence in pKK223-3 (Brosius and Lupsky, 1987).
either white or paler blue than the control (YCp50-transformed) colonies (see "Results" for the design of this XGal screen). Restriction analysis of plasmid isolates (named pRBW1-pRBW6) prepared from these six colonies showed that each of the six samples contained pACYC184 (pUC19-based) and only three of the samples contained an insert-bearing YCp50. Further analysis of the inserts by restriction digestion of the insert-bearing plasmids showed that pRBW1 and pRBW6 contained overlapping yeast DNA inserts, whereas pRBW2 contained a different yeast DNA insert. Extracts of E. coli transformed with pRBW1 or pRBW2 were then assayed as described below for their ability to deubiquitinate Ub-Met-pgal. The plasmid pRBW2 could reproducibly confer the deubiquitinating activity in vitro (in extracts from E. coli carrying the plasmid; data not shown). The apparent ability of these plasmids to confer deubiquitinating activity on Ub-Arg-pgal in vivo was therefore verified by pulse-chase analysis of Ub-Arg-pgal, using a monoclonal antibody to pgal (Bachmair et al., 1991) in E. coli transformed with either pRBW1 or pRBW2. This test (see "Results") confirmed that both plasmids conferred the deubiquitinating activity on E. coli in vivo.

The relevant regions of the inserts in pRBW1 and pRBW2 were identified by subcloning restriction fragments of each insert into pUC19 and assaying the resulting subclones using the above XGal screen. Subclone pRB5, containing a -1.8-kb HindIII fragment of pRBW2, and subclone pRB7, containing a ~5.3-kb SphI/EcoRI fragment of pRBW1, were found to confer deubiquitinating activity on E. coli (data not shown). Restriction fragments of the inserts in pRB5 and pRB7 were subcloned into M13mp18 and M13mp9 and sequenced as described above. To describe the region of pRB7 above. To determine the sequencing of an open reading frame (ORF) a portion of which was absent from pRB5, a -2-kb HindIII fragment of pRB6, another subclone of pRBW2, and a 0.95-kb SphI/HindIII fragment of pRB9, another (adjoining) pRBW2 subclone, were sequenced as well. The 3,792-bp pRB5 from pRB5/6/9 was named UBP2; the 2,736-bp ORF from pRB7 was named UBP3.

Assays for Ubiquitin-specific Protease Activity—For the in vitro assays, purified, 35S-labeled Ub-Met-DHFR was prepared as described by Tobias and Varshavsky (1991). Purified, 35S-labeled Ub-Met-Arg was a gift from Dr. D. Gonda (Yale University).

Extracts of E. coli containing the plasmids pRBW1 and pRBW2 were prepared as described by Tobias and Varshavsky (1991), using a monoclonal antibody to pgal (Bachmair et al., 1991), in E. coli transformed with either pRBW1 or pRBW2. This test (see "Results") confirmed that both plasmids conferred the deubiquitinating activity on E. coli in vivo.

The relevant regions of the inserts in pRBW1 and pRBW2 were identified by subcloning restriction fragments of each insert into pUC19 and assaying the resulting subclones using the above XGal screen. Subclone pRB5, containing a -1.8-kb HindIII fragment of pRBW2, and subclone pRB7, containing a ~5.3-kb SphI/EcoRI fragment of pRBW1, were found to confer deubiquitinating activity on E. coli (data not shown). Restriction fragments of the inserts in pRB5 and pRB7 were subcloned into M13mp18 and M13mp9 and sequenced as described above. To determine the sequencing of an open reading frame (ORF) a portion of which was absent from pRB5, a -2-kb HindIII fragment of pRB6, another subclone of pRBW2, and a 0.95-kb SphI/HindIII fragment of pRB9, another (adjoining) pRBW2 subclone, were sequenced as well. The 3,792-bp pRB5 from pRB5/6/9 was named UBP2; the 2,736-bp ORF from pRB7 was named UBP3.

Assays for Ubiquitin-specific Protease Activity—For the in vitro assays, purified, 35S-labeled Ub-Met-DHFR was prepared as described by Tobias and Varshavsky (1991). Purified, 35S-labeled Ub-Met-Arg was a gift from Dr. D. Gonda (Yale University).

Extracts of E. coli containing the plasmids pRBW1 and pRBW2 were prepared as described by Tobias and Varshavsky (1991), using a monoclonal antibody to pgal (Bachmair et al., 1991), in E. coli transformed with either pRBW1 or pRBW2. This test (see "Results") confirmed that both plasmids conferred the deubiquitinating activity on E. coli in vivo.

The relevant regions of the inserts in pRBW1 and pRBW2 were identified by subcloning restriction fragments of each insert into pUC19 and assaying the resulting subclones using the above XGal screen. Subclone pRB5, containing a -1.8-kb HindIII fragment of pRBW2, and subclone pRB7, containing a ~5.3-kb SphI/EcoRI fragment of pRBW1, were found to confer deubiquitinating activity on E. coli (data not shown). Restriction fragments of the inserts in pRB5 and pRB7 were subcloned into M13mp18 and M13mp9 and sequenced as described above. To determine the sequencing of an open reading frame (ORF) a portion of which was absent from pRB5, a -2-kb HindIII fragment of pRB6, another subclone of pRBW2, and a 0.95-kb SphI/HindIII fragment of pRB9, another (adjoining) pRBW2 subclone, were sequenced as well. The 3,792-bp pRB5 from pRB5/6/9 was named UBP2; the 2,736-bp ORF from pRB7 was named UBP3.

Assays for Ubiquitin-specific Protease Activity—For the in vitro assays, purified, 35S-labeled Ub-Met-DHFR was prepared as described by Tobias and Varshavsky (1991). Purified, 35S-labeled Ub-Met-Arg was a gift from Dr. D. Gonda (Yale University).

Extracts of E. coli containing the plasmids pRBW1 and pRBW2 were prepared as described by Tobias and Varshavsky (1991), using a monoclonal antibody to pgal (Bachmair et al., 1991), in E. coli transformed with either pRBW1 or pRBW2. This test (see "Results") confirmed that both plasmids conferred the deubiquitinating activity on E. coli in vivo.
Retention of Deubiquitinating Activity in a yuh1 ubp1 Double Mutant—YUH1, one of the two previously isolated genes for ubiquitin-specific processing proteases of *S. cerevisiae*, encodes a 236-residue protease that deubiquitinates relatively short C-terminal extensions of ubiquitin but is virtually inactive with larger ubiquitin fusions such as Ub-β-gal (Miller et al., 1989; Tobias and Varshavsky, 1991). By contrast, the more recently cloned yeast *UBP1* gene encodes an 807-residue ubiquitin-specific protease that deubiquitinates natural and engineered ubiquitin fusions irrespective of their size (Tobias and Varshavsky, 1991). The null *ubp1* mutant is viable and retains the ability to deubiquitinate Ub-β-gal, suggesting that the family of ubiquitin-specific proteases in yeast is not limited to Ubp1 and Yuh1 (Tobias and Varshavsky, 1991). To test this directly, we constructed a null *yuh1 ubp1* mutant (YRB111; see Table I and “Experimental Procedures”). Extracts prepared from *yuh1* and *ubp1* single mutants, from wild-type *S. cerevisiae*, and from a congenic *yuh1 ubp1* double mutant were indistinguishable in their ability to deubiquitin-ate Ub-Met-DHFR (Fig. 1). Since ubiquitin is an essential protein that is encoded in yeast exclusively by gene fusions (Finley et al., 1989; Finley and Chau, 1991), these results indicate that the family of ubiquitin-specific processing prote-ases in *S. cerevisiae* is not limited to Ubp1 and Yuh1.

**Isolation of UBP2 and UBP3**—To isolate the genes encoding ubiquitin-specific proteases other than Ubp1 and Yuh1, we took advantage of the absence of ubiquitin and ubiquitin-specific enzymes from bacteria such as *E. coli* (Tobias and Varshavsky, 1991) and also of the recent finding that the N-end rule, a relation between the in vivo half-life of a protein and the identity of its N-terminal residue, operates in *E. coli* as well (Tobias et al., 1991). In eukaryotes, ubiquitins fusions to test proteins such as β-galactosidase are deubiquitinated by ubiquitin-specific processing proteases irrespective of the identity of a residue at the Ub-β-gal junction (proline being the single exception), making it possible to expose in vivo different residues at the N termini of otherwise identical test proteins (Bachmair et al., 1986; Varshavsky, 1992). This technique, required for detection and analysis of the N-end rule in eukaryotes, has been made applicable in bacteria through the isolation of the yeast *UBP1* gene, inasmuch as *E. coli* transformed with *UBP1* acquires the ability to deubiquitinate ubiquitin fusions (Tobias and Varshavsky, 1991). The finding that an X-β-gal test protein such as Arg-β-gal is short-lived in *E. coli* (Tobias et al., 1991), whereas Ub-Arg-β-gal is long-lived (Bachmair et al., 1986), made possible an *E. coli*-based in vivo screen for ubiquitin-specific proteases. *E. coli* expressing the (long-lived) Ub-Arg-β-gal fusion protein form blue colonies on plates containing XGal, a chromogenic substrate of β-gal. However, if a deubiquitinating activity is present in the cells as well, Ub-Arg-β-gal is converted into the short-lived Arg-β-gal (Tobias et al., 1991) whose low steady-state level results in *E. coli* colonies that are white on XGal plates (see “Experimental Procedures”).

To be cloneable by this strategy using a conventional yeast genomic DNA library, a yeast gene must have a promoter that functions in *E. coli* (a minority of yeast promoters can do so), must lack introns in its coding region (most yeast genes lack introns), and must encode a ubiquitin-specific processing protease that functions as a monomer or a homooligomer. One advantage of this in vivo screen over the previously used in vitro screen that yielded *UBP1* (Tobias and Varshavsky, 1991) is that the present screen requires a relevant protease to be active in vivo but not necessarily in *E. coli* extracts.

**E. coli** carrying a plasmid expressing Ub-Arg-β-gal were transformed with a *S. cerevisiae* genomic DNA library carried in the plasmid YCP50, plated on XGal plates containing antibiotics that selected for the presence of both plasmids, and incubated overnight at 37 °C (see “Experimental Procedures”). Of ~800 colonies thus screened, six were white or pale blue, whereas the other colonies were as dark blue as control colonies of *E. coli* transformed with pACUB-R-β-gal.

### RESULTS

**Mutant- YUH1**,
and the YCp50 vector. Three of the six candidate colonies were found to be false positives, two contained plasmids (termed pRBW1 and pRBW6) with overlapping inserts of yeast DNA, and the remaining colony contained a plasmid (termed pRBW2) with a distinct yeast DNA insert (see “Experimental Procedures”). Plasmids pRBW1 and pRBW2 were isolated and retransformed into E. coli expressing either Ub-Arg-βgal or Ub-Met-βgal. Transformants expressing Ub-Arg-βgal formed white colonies on XGal plates, confirming the original results, whereas transformants expressing Ub-Met-βgal formed blue colonies on these plates, indicating that the metabolic destabilization of Ub-Arg-βgal by inserts in pRBW1 and pRBW2 was N-end rule-specific. (Arg and Met are, respectively, destabilizing and stabilizing residues in the E. coli N-end rule (Tobias et al., 1991).)

Surprisingly, extracts of E. coli carrying pRBW1 or pRBW2 were inactive in an in vitro deubiquitinating assay with Ub-Met-DHFR (data not shown; see “Experimental Procedures” and Fig. 1), suggesting that ubiquitin-specific proteases encoded by pRBW1 and pRBW2 were either inactivated in cell extracts or, alternatively, could deubiquitinate ubiquitin fusions cotranslationally but not posttranslationally. The ubiquitin-specific protease activities conferred by pRBW1 and pRBW2 on E. coli were therefore assayed in vitro by pulse-chase analyses with Ub-Met-βgal (see “Experimental Procedures”). The results (not shown, but see below) confirmed that pRBW1 and pRBW2 (but not the YCp50 vector alone) did confer the deubiquitinating activity on E. coli. Subsequent overexpression of the ubiquitin-specific protease encoded by pRBW2 made possible its detection in E. coli extracts (see below).

The UBP2 Gene—The ORF encoding the deubiquitinating activity of pRBW2 was identified by subcloning experiments and nucleotide sequencing as described under “Experimental Procedures” and was named the UBP2 gene (Figs. 2A and 3).

The position of the start (ATG) codon in the UBP2 was inferred so as to yield the longest (3,715 bp) ORF, which encodes an acidic (calculated pI of 4.95), 1,264-residue (145-kDa) protein (Fig. 3). No ATGs occur in any of three forward reading upstream frames until position −203 (Fig. 3). The codon adaptation index of UBP2 (calculated according to Sharp and Li, 1987) is 0.153, which is close to that of UBPl and UBP2 (see above), 0.180, which is close to those of UBPl and UBP2 (see above), and is characteristic of weakly expressed yeast genes. Computer-aided comparisons of the UBP2 nucleotide sequence and of the predicted amino acid sequence of Ubp2 to sequences in databases showed no statistically significant similarities to sequenced genes or gene products, except for the yeast UBPl gene (see below).

Northern hybridization analysis with a UBP2-specific DNA probe and RNA isolated from growing yeast cells identified a −4.0-kb transcript. This transcript was neither induced nor significantly down-regulated after a 90-min heat stress (Fig. 4B). A partially sequenced ORF (ORFb in Fig. 2A) that begins 253 bp downstream of UBP2 is unusual in containing a high proportion of codons of the consensus sequence GAN that encode either Asp or Glu; 78 of the 327 codons in the sequenced portion of ORFb are the GAN codons (Fig. 3). About 50 “islands” of such repeating triplets are present in the S. cerevisiae genome; they are known as the “CAT” repeats (ATG or “frame-shifted” GAT repeats appear as CAT repeats on the complementary strand). Three sets of CAT repeats have been sequenced at unmapped genomic loci (Wildeman et al., 1986); their occurrence within transcribed ORFs (see below) has not been reported previously. ORFb encodes a highly acidic protein containing several potential phosphorylation sites for casein-type protein kinases (Ser or Thr flanked by acidic residues) (Fig. 3). These sites are similar to the sites observed by Wickner (1988) in Makl6, a S. cerevisiae protein involved in cell cycle control. Another adjacent, partially sequenced ORF (ORFc) is also oriented in the same direction as UBP2 and terminates 357 bp upstream of UBP2. ORFc encodes a protein of at least 198 residues (Figs. 2A and 3). Northern analysis with RNA from growing yeast cells showed that both ORFb and ORFc are transcriptionally active, producing, respectively, −1.5-kb (Fig. 4C) and −0.85-kb (Fig. 4A) mRNAs. Neither of these ORFs encodes a protein with significant similarities to sequences in data bases.

The UBP3 Gene—The ORF encoding the deubiquitinating protease of pRBW1 was identified by subcloning experiments and nucleotide sequencing as described under “Experimental Procedures” and was named the UBP3 gene (Figs. 2B and 5). The position of the start (ATG) codon was inferred so as to yield the longest (2,736 bp) ORF, which encodes a slightly basic (calculated pI of 7.92), 912-residue (102-kDa) protein (Fig. 3). About 50 “islands” of such repeating triplets are present in the S. cerevisiae genome; they are known as the “CAT” repeats (ATG or “frame-shifted” GAT repeats appear as CAT repeats on the complementary strand). Three sets of CAT repeats have been sequenced at unmapped genomic loci (Wildeman et al., 1986); their occurrence within transcribed ORFs (see below) has not been reported previously. ORFb encodes a highly acidic protein containing several potential phosphorylation sites for casein-type protein kinases (Ser or Thr flanked by acidic residues) (Fig. 3). These sites are similar to the sites observed by Wickner (1988) in Makl6, a S. cerevisiae protein involved in cell cycle control. Another adjacent, partially sequenced ORF (ORFc) is also oriented in the same direction as UBP2 and terminates 357 bp upstream of UBP2. ORFc encodes a protein of at least 198 residues (Figs. 2A and 3). Northern analysis with RNA from growing yeast cells showed that both ORFb and ORFc are transcriptionally active, producing, respectively, −1.5-kb (Fig. 4C) and −0.85-kb (Fig. 4A) mRNAs. Neither of these ORFs encodes a protein with significant similarities to sequences in data bases.
Drosophila melanogaster; small font size.

repeats in the analyzed regions of this satellite DNA ranges another such sequence is present 20 bp upstream of this region (Fig. 5). AATAC is the core repeat in a satellite DNA of UBP3-specific DNA probe and RNA from growing yeast cells. Between 50 and 75 bp downstream of the AATAC repeat is a GC repeat, and 75 to 100 bp downstream is another such sequence.

Fig. 3. Nucleotide sequence of the UBP2 gene and deduced amino acid sequence of the Ubp2 protein. The nucleotide residues are numbered on the left, beginning at the A of the presumed start codon. Amino acid residues are numbered also on the right, with a smaller font size. The asterisk indicates the TAG stop codon. The presumed start codon of ORFb and the TAA stop codon of ORFc (Fig. 2A) are double underlined.

UBP3-specific DNA probe and RNA from growing yeast cells identified a ~3.0-kb transcript. This transcript was slightly induced after a 90-min heat stress (Fig. 4E).

Between 50 and 75 bp downstream of the UBP3 stop codon, there are five tandem repeats of the sequence AATAC, and another such sequence is present 20 bp upstream of this region (Fig. 5). AATAC is the core repeat in a satellite DNA of Drosophila melanogaster; the number of contiguous AATAC repeats in the analyzed regions of this satellite DNA ranges from 18 to 41 one (Lohe and Brutlag, 1987). AATAC is also present in the repeats are located within 75 bp of the DNA region that is adjacent to a cluster of same organism (Spradling 1987), whereas 12 contiguous Alu repeats in the human Y chromosome.

The nucleotide residues may also be present in the repeats in the region that is adjacent to a cluster of the same organism (Spradling 1987), whereas 12 contiguous Alu repeats have not been described previously in S. cerevisiae. Since the AATAC repeats are located within 75 bp of the UBP3 stop codon, they may also be present in the UBP3 mRNA.
prior to RNA isolation total yeast RNA were hybridized with "P-labeled DNA probes specific for ORF2 (panel A; see Fig. 2A), UBP2 (panel B; see Fig. 2A), ORFb (panel C; see Fig. 2A), ORF2 (panel D; see Fig. 2B), and UBP3 (panel E; see Fig. 2B). RNA was prepared from S. cerevisiae exponentially growing at 30 °C (lanes 30) or shifted to 39 °C for 90 min prior to RNA isolation (lanes 39) (see "Experimental Procedures"). 28 S and 18 S indicate the positions of the corresponding rRNAs.

Computer-assisted analysis of the 3'-flanking region of UBP3 detected a putative tRNA(GU) gene and a truncated Ty δ element, with the tRNA gene located 154 bp downstream of the UBP3 stop codon and oriented in the opposite direction (Figs. 2B and 5). This 78-bp, putative tRNA gene contains the anticodon AAG recognizing the GUU Val codon, and its nucleotide sequence is identical to those of the two previously reported tRNA(GU) genes of S. cerevisiae (Baker et al., 1982; Labbe-Bois, 1990). The three tRNA(GU) genes are clearly distinct, as they have nonhomologous 3'- and 5'-flanking regions (data not shown). The UBP3-linked tRNA(GU) gene was named TGV3 (a third tRNA gene for a Val trna). Like many other yeast tRNA genes (Hauber et al., 1988; Warmington et al., 1986), TGV3 is associated with a Ty δ element, which is located 63 bp upstream from the 5' end of TGV3 (Fig. 5). This δ element, flanked by 5-bp direct repeats (Fig. 5), is incomplete: it consists of the last 67 bp of the full-length, -330-bp δ element (Hauber et al., 1988; Warmington et al., 1986). Detailed analysis of the TGV3-containing region will be published elsewhere.②

A sequence upstream of UBP3 contains a 1,046-bp ORF (ORF2), whose stop codon is 231 bp upstream of the UBP3 start codon (Figs. 2B and 5). Northern analysis using an ORF2-specific DNA probe identified three mRNA species of ~2.9, ~1.5, and ~1.0 kb (Fig. 4D). Comparison of ORF2 with nucleotide sequences in the GenBank/EMBL database data base showed that ORF2 is identical to a previously reported ORF (also called ORF2) located 53 bp downstream of the PET122 gene (Ohmen et al., 1990). PET122 encodes a translational regulator of the mitochondrial cytochrome oxidase III mRNA (Ohmen et al., 1990). PET122 is transcribed as a part of bicistronic mRNAs of ~2.9 and ~1.5 kb which also contain ORF2. A distinct ~1.0-kb transcript that contains only ORF2 and starts within the 3' end of the PET122 gene is produced as well (Ohmen et al., 1990). These three transcripts are clearly the same as those observed by us using the ORF2-derived probe (Fig. 4D). The function, if any, of the bicistronic organization of the PET122-encoding mRNAs is unknown. The PET122/ORF2 transcripts do not extend into UBP3, as we could not detect any UBP3-linked, ORF2-containing transcripts by Northern hybridization, using a probe derived from the 5' end of UBP3 (data not shown). There is also a partially sequenced, divergently transcribed ORF (ORF3) that starts 215 bp upstream of PET122 (Fig. 2B) (Ohmen et al., 1990). The functions of ORF2 and ORF3 are unknown; neither of these genes has significant similarities to sequences in databases.②

Chromosomal Mapping of UBP1-UBP3—DNA probes specific for UBP2 and UBP3 that were used in the Northern analyses (Fig. 4) and a UBP1-specific probe (a ~2.2-kb EcoRI fragment of UBP1; Tobias and Varshavsky, 1991) were hybridized to Southern blots of S. cerevisiae chromosomes fractionated by pulse-field electrophoresis (see "Experimental Procedures"). This analysis indicated an unlinked arrangement of UBP genes, with UBP1 located on chromosome IV, UBP2 on chromosome VII, and UBP3 on chromosome V (Fig. 6). Although more detailed mapping has not been carried out, it is known that UBP1 is not close to the centromere-proximal TRP1 gene on chromosome IV because the segregation of Trp" and Ura" phenotypes was indistinguishable from random upon dissection of 22 tetrads from an UBP1/ubp1- Δ1::URA3 TRP1/Δtrp1 diploid (data not shown).

Sequence Comparisons of Ubiquitin-specific Proteases—Sequence alignment of the 809-residue Ubp1, 1,264-residue Ubp2, and 912-residue Ubp3 demonstrated the lack of overall sequence similarity among these proteins, as well as the presence of two short regions of statistically significant similarity which are spaced a few hundred residues apart in each of the Ubp proteases (Fig. 7A). The two regions of similarity contain a Cys residue (Fig. 7B) and two His residues (Fig. 7C), respectively, and may form portions of the active sites of these thiol proteases (Fig. 7B) (Hempel et al., 1991; Kamphuis et al., 1985). As has been seen with Ubp1 (Tobias and Varshavsky, 1991), neither Ubp2 nor Ubp3 has significant sequence similarities to the fourth ubiquitin-specific protease of yeast, Yuh1 (Miller et al., 1989), or its mammalian homologs (Wilkinson et al., 1989). The region in Yuh1 and its homologs that contains a putative active-site Cys residue (Wilkinson et al., 1989) is not similar to the conserved "Cys" region of Ubp1-Ubp3: apart from the Cys residue, only 1 other residue position is occupied by an identical residue (Asn) in all six proteins (Fig. 7B). No such similarities are seen in an analogous alignment of the 2 conserved His residues in Yuh1-like proteases with either of the conserved His residues in Ubp1-Ubp3 (data not shown).

In Vitro Properties of Ubiquitin-specific Proteases—The previously characterized Ubp1 protease can efficiently deubiquitinate in vitro a variety of linear ubiquitin fusion proteins, including the natural ubiquitin precursors Ubi1-Ubi3 and engineered fusions such as Ub-X-Δgal and Ub-X-DHFR (Tobias and Varshavsky, 1991). Similar assays, in which an extract of E. coli carrying an overexpression vector-based plasmid expressing either Ubp2 (pRB105), Ubp3 (pRB143), or Yuh1 (pKKYUH1) is incubated with ubiquitin-containing test proteins, were used to analyze in vitro the substrate specificity of these proteases (see "Experimental Procedures"). Extracts of E. coli carrying the UBP1-expressing plasmid pJT70 (Tobias and Varshavsky, 1991) or vector alone were also used in these assays. The cleavage products were fractionated by SDS-PAGE and visualized by immunoblotting, using anti-ubiquitin antibodies (Fig. 8C) or, with 35S-labeled test proteins, directly by fluorography (Fig. 8, A and B).

In these in vitro assays, the Ubp2 protease efficiently deubiquitinated Ub-Met-Δgal (Fig. 8A) and Ub-Met-DHFR (Fig. 8B), as well as Ubi2 and Ubi3, the natural precursors of ubiquitin, in which it is fused to specific ribosomal proteins (Fig. 8C). Both Ubp1 and Ubp2 released the Cys residue from Ub-Ub-Cys (di-Ub bearing a 1-residue C-terminal extension) but were unable to cleave at the Ub-Ub junction in Ub-Ub-Cys (Fig. 8C). Ubp1 and Ubp2 were also unable to cleave at the Ub-Ub junctions in the yeast poly-Ub, a natural ubiquitin precursor containing five head-to-tail ubiquitin repeats (Fig.

② R. T. Baker, manuscript in preparation.
Ubiquitin-specific Proteases

Fig. 5. Nucleotide sequence of the UBP3 gene and deduced amino acid sequence of the Ubp3 protein. The nucleotide residues are numbered on the right, beginning at the A of the presumed start codon. Amino acid residues are numbered also on the right, with a smaller font size. The asterisk indicates the TAA stop codon. The TAG stop codon of ORF2 upstream of UBP3 (Fig. 2B) is double underlined. Also double underlined are six repeats of AATTCC that flank the d element. A tRNA^GU (GUU) gene (termed TGO3) between UBP3 and Ty d is single underlined. See also Fig. 2B and main text.

Although the expression of Ubp3 in E. coli from the plasmid pRB143 resulted in a substantial overproduction of a protein with the expected molecular mass (data not shown), extracts of Ubp3-expressing E. coli lacked deubiquitinating activity (Fig. 8). Since Ubp3 is certainly active in E. coli in vivo (see below), it is either inactivated in E. coli extracts or is able to cleave ubiquitin fusions exclusively during or shortly after
their ribosome-mediated synthesis.

In agreement with the findings by Miller et al. (1989), extracts of E. coli expressing Yuh1 efficiently deubiquitinated short ubiquitin fusions such as Ubi2 and Ubi3 (Fig. 8C). However, Yuh1 was much less active against the larger fusion Ub-Met-DHFR (a 229-residue C-terminal extension of ubiquitin), deubiquitinating at most ~50% of the fusion even after a prolonged incubation (Fig. 8D and data not shown), and was virtually inactive against Ub-Met-βgal (a 1045-residue C-terminal extension of ubiquitin) (Fig. 8A).

In Vivo Properties of Ubiquitin-specific Proteases—As expected from their activities in E. coli extracts, Ubp1, Ubp2, and Yuh1 were active in vivo against the natural ubiquitin fusions Ubi2 and Ubi3 (Fig. 8D). Ubp3, which was inactive in E. coli extracts, efficiently deubiquitinated Ubi2 and Ubi3 when coexpressed with them in E. coli (Fig. 8D).

Although Ubp1 and Ubp2 were unable to cleave at the Ub-Ub junction in poly-Ub in vitro, both of them were active against the yeast poly-UB when coexpressed with it in E. coli (Fig. 8D). In contrast, the Ubp3 protease, although active in vivo against ubiquitin fusions such as Ubi2 and Ubi3 (Fig. 8D), was inactive, under the same conditions, against poly-Ub (Fig. 8C). These distinctions among ubiquitin-specific processing proteases indicate subtle differences in their requirements for the conformation of protein domains in the vicinities of Ub-X peptide bonds (see “Discussion”).

The in vivo deubiquitination of ubiquitin fusions such as Ub-Met-βgal by Ubp2 and Ubp3 was also followed by pulse-chase analysis, in part to confirm the findings of the original XGal screen (see above). As expected, both proteases deubiquitinated Ub-Met-βgal in vivo, except that the cleavage by Ubp3 was incomplete, and a significant proportion of pulse-labeled Ub-Met-βgal remained intact 15 min after the pulse (Fig. 8E). These results are consistent with the pattern of deubiquitination by Ubp3 that is more strictly cotranslational than that by Ubp2 and Ubp1. In a similar pulse-chase assay, Yuh1 was unable to deubiquitinate Ub-Met-βgal in vivo (Fig. 8E), indicating that an apparently greater susceptibility of the Ub-Met peptide bond in a nascent (as distinguished from mature) Ub-Met-βgal is insufficient to allow its in vivo deubiquitination by Yuh1. By contrast, this difference is sufficient to allow a cotranslational (but apparently not post-translational) deubiquitination of Ub-Met-βgal by Ubp3 (see above).

S. cerevisiae Contains at Least Five Distinct Ubiquitin-specific Processing Proteases—The observation that a S. cerevisiae strain lacking both Ubp1 and Yuh1 was able to deubiquitinate ubiquitin fusions (Fig. 1) has led to the isolation of the UBP2 and UBP3 genes. To determine whether Yuh1 and Ubp1–Ubp3 comprise the full set of ubiquitin-specific proteases in S. cerevisiae, we constructed a haploid yuh1 ubp1 ubp2 ubp3 quadruple mutant, YRB141 (Table I), which lacks all of these proteases (see “Experimental Procedures”). Since deubiquitination of at least one of the four yeast ubiquitin precursors Ubi1–Ubi4 is required for producing mature ubiquitin, an essential protein (Finley and Chau, 1991), the viability...
of the quadruple mutant YRB141 indicates that at least one other ubiquitin-specific protease must exist in yeast. To verify this conclusion, extracts from YRB141 (yuh1 ubp1 ubp2 ubp3) and from other ubp mutants were tested for their ability to deubiquitinate Ub-Met-βgal. All of the single ubp mutants examined were able to deubiquitinate Ub-Met-βgal in vitro with indistinguishable efficiencies (Fig. 9). However, the YRB141 extract, as well as an extract from a ubp1Δ ubp2Δ double mutant, were slightly but detectably less active in this deubiquitination assay, as evidenced by increased levels of the uncleaved Ub-Met-βgal (Fig. 9). The approximately equal decrease in the activity of extracts from the quadruple-deletion strain and the ubp1Δ ubp2Δ strain (Fig. 9) is consistent with the evidence that Yuh1 and Ubp3 cannot cleave Ub-Met-βgal in vitro (Fig. 8). These results also indicate that S. cerevisiae contains at least one more ubiquitin-specific processing protease that is active against large ubiquitin fusions such as Ub-Met-βgal.

While viable, the yuh1 ubp1 ubp2 ubp3 quadruple mutant is hypersensitive to a variety of stresses, including a chronic heat stress and exposure to amino acid analogs such as canavanine (data not shown). Since one common feature of these stresses is an increased demand for ubiquitin (Finley et al., 1989), the above phenotypes suggest a reduced capacity for deubiquitination in the quadruple deletion mutant.

**DISCUSSION**

Using an E. coli-based genetic screen, we have cloned UBP2 and UBP3, the genes encoding two new ubiquitin-specific processing proteases of S. cerevisiae. The Ubp2 (1,264 resi-
Ubiquitin-specific Proteases

Fig. 9. *S. cerevisiae* that lacks Yuh1, Ubp1, Ubp2, and Ubp3 retains deubiquitinating activity. Purified, 35S-labeled Ub-Met-βgal was added to a lysate buffer alone (control) or to a *S. cerevisiae* extract from YRR110 (wild-type), YRB109 (ubp2Δ), YRB173 (ubp3Δ), YRB175 (ubp3Δa), YRB147 (ubp2A ubp3A), or YRB141 (ubp2Δ ubp2Δ ubp2A yuh1Δ), as indicated above the lanes. After an incubation for 1 h at 36°C, the samples were fractionated by SDS-PAGE, followed by fluorography. See “Experimental Procedures” and Table I. Bands of Ub-Met-βgal and Met-βgal are indicated on the right.

dues), Ubp3 (912 residues), and the previously cloned Ubp1 (809 residues) proteases (Tobias and Varshavsky, 1991) are largely dissimilar in sequence, except for two short regions containing Cys and His that encompass their putative active sites. Neither of these proteases has sequence similarities to Yuh1, the other previously identified ubiquitin-specific protease of *S. cerevisiae* (Miller et al., 1989).

We have also constructed a null yuh1 ubp1 ubp2 ubp3 quadruple mutant, which is viable and retains the ability to deubiquitinate ubiquitin fusions, indicating the presence of at least one more ubiquitin-specific processing protease in *S. cerevisiae*. Thus, some of the members of the at least five-member UBP gene family have overlapping functions. That these functions do not overlap completely is indicated not only by the stress-sensitive phenotype of the yuh1 ubp1 ubp2 ubp3 quadruple mutant (such phenotype might be due solely to insufficient expression of the remaining ubiquitin-specific proteases) but also by the distinct substrate specificities of Ubp1–Ubp3 and Yuh1. For example, whereas Ubp1 and Ubp2 are able to deubiquitinate in vitro all of the tested ubiquitin fusions, Yuh1 under the same conditions is unable to remove ubiquitin from either poly-Ub or relatively large ubiquitin fusions such as Ub-βgal (see “Results”). Further distinctions between ubiquitin-specific proteases are suggested by the finding that Ubp3, in contrast to Yuh1, efficiently deubiquitinating in vitro large ubiquitin fusions such as Ub-βgal; however, like Yuh1 but unlike Ubp1 and Ubp2, the Ubp3 protease fails to cleave poly-Ub.

Although the in vivo (in *E. coli*) substrate specificity of Yuh1 is not significantly different from those observed in vitro (in *E. coli* extracts), the Ubp3 protease could be characterized thus far only in vivo, as it is inactive in *E. coli* extracts. On the other hand, Ubp1 and Ubp2, although active both in vitro and in vivo, are able to cleave poly-Ub when coexpressed with it in *E. coli* but not in an *E. coli* extract. The dependence of activity on coexpression with a substrate strongly suggests that the deubiquitination of poly-Ub by Ubp1 or Ubp2 is cotranslational or, more generally, occurs before the folding of the newly formed, multidomain poly-Ub protein is complete. Since the N-terminal region of ubiquitin, and in particular its N-terminal Met residue, is an integral part of the compact ubiquitin globule (Vijay-Kumar et al., 1987; Wilkinson, 1988; Bamesei et al., 1991), the cleavable Gly-Met bond between the adjacent ubiquitin moieties in the fully folded poly-Ub protein may be sterically shielded to a greater extent than are analogous Gly-X bonds in ubiquitin fusions that lack the Ub-Ub motif. Similar considerations may also account for the apparently cotranslational cleavage of Ub-βgal by Ubp3 in vivo (unlike Ubp1 and Ubp2, the Ubp3 protease is unable to cleave Ub-βgal in *E. coli* extracts) and for the inability of Yuh1 to remove ubiquitin from its larger C-terminal extensions such as βgal under in vivo or in vitro (see “Results”).

When expressed in *E. coli*, neither Yuh1 nor Ubp1–Ubp3 proteases are active against the added poly-Ub in an *E. coli* extract (Fig. 8C). By contrast, an extract from *S. cerevisiae* is able to cleave the added poly-Ub (Fig. 8C; see also Tobias and Varshavsky, 1991). One explanation of these results is that the cleavage of poly-Ub in yeast extracts is mediated by an as yet unidentified Ub-specific protease (at least one such protease remains to be isolated; see above). Another testable possibility is that, unlike bacteria (which lack the ubiquitin system), *S. cerevisiae* contains an accessory factor that is able to perturb the conformation of mature (completely folded) poly-Ub in a way that makes it susceptible to cleavage by ubiquitin-specific proteases such as Ubp1 and Ubp2.

A priori, ubiquitin-specific proteases that deubiquitinate linear ubiquitin fusions may also have an isopeptidase activity, i.e. the ability to deubiquitinate posttranslationally formed, branched ubiquitin-protein conjugates. A multi-Ub chain is one such conjugate, in which several ubiquitin moieties are attached sequentially to the initial acceptor protein, forming a chain of Ub-Ub conjugates in which the C-terminal Gly of one ubiquitin is linked to Lys of the adjacent ubiquitin (Chau et al., 1989). Formation of a multi-Ub chain on a targeted protein is the required predegradation step for substrates of the N-end rule pathway and other ubiquitin-dependent proteolytic pathways (Chau et al., 1989; Johnson et al., 1992; Varshavsky, 1992). Thus, overexpression of a ubiquitin-specific protease that is active against multi-Ub would be expected to perturb the in vivo dynamics of multi-Ub chains and therefore, probably, the degradation of normally multiubiquitinated substrates as well. Such effects have indeed been observed: overexpression of the Ubp2 protease in *S. cerevisiae* was found to result in a strong metabolic stabilization of normally short-lived N-end rule substrates and a decrease of their steady-state multiubiquitination; as would be expected from these findings, cells that overexpress Ubp2 are hypersensitive to stresses whose containment is known to require enhanced intracellular proteolysis.

Acknowledgments—We are indebted to Tom Shrader and Jürgen Dohmen for assistance with the development of the screen protocol. We thank Barbara Doran for secretarial assistance.

REFERENCES


Broach, J. R., Strathern, J. N., and Hicks, J. B. (1979) Gene (Amst.) 8, 121–133


* R. T. Baker and A. Varshavsky, manuscript in preparation.
Ubiquitin-specific Proteases


