THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 267, No. 10, Issue of April 5, pp. 6495-6501, 1992
Printed in U.S.A.


# A Rabbit Reticulocyte Ubiquitin Carrier Protein That Supports Ubiquitin-dependent Proteolysis ($E2_{14k}$) Is Homologous to the Yeast DNA Repair Gene RAD6*

(Received for publication, April 30, 1991)

**Simon S. Wing‡, France Dumas, and Denis Banville**

*From the Molecular Biology Sector, Biotechnology Research Institute, National Research Council of Canada, Montreal, Quebec H4P 2R2, Canada*

**The two isoforms of the 14-kDa ubiquitin carrier protein ($E2_{14k}$) are unique among rabbit *E2*s in efficiently supporting ubiquitin-protein ligase (*E3*)-mediated ubiquitination of proteins destined for degradation. To begin determining the structural basis for this property, we have isolated a cDNA encoding the predominant reticulocyte isoform of the *E2* from a rabbit skeletal muscle library. The sequence predicts a protein of 152 amino acids with a molecular weight of 17,293. Expression of the cDNA in *Escherichia coli* and purification of the recombinant protein revealed an *E2* with high affinity for *E3* and ubiquitin activating enzyme (*E1*). The latter high affinity interaction appears to be between the ubiquitin charged form of *E1* and the uncharged form of *E2* and does not result in a stable complex between these two enzymes. The predicted sequence shows regions of strong homology with other sequenced *E2*s, suggesting that these regions may be involved in binding to *E1* and/or in ubiquitin transfer from *E1*, functions common to all *E2*s. Surprisingly, the $E2_{14k}$ sequence is markedly more similar to *Saccharomyces cerevisiae* RAD6 (69% identity) than to its proposed homologs UBC4/UBC5 (38% identity). The sequence is identical to that recently reported for a human 17-kDa *E2* which can complement rad6 mutants thereby identifying rabbit $E2_{14k}$ as a RAD6 homologue. The biochemical properties of this previously uncharacterized human 17-kDa *E2* are now defined and its misassignment as a homologue of rabbit $E2_{17k}$ is corrected. Our findings resolve current confusion regarding relationships among *E2*s and define yeast RAD6, rabbit $E2_{14k}$, and the human 17-kDa *E2* as a subclass of *E2*s which biochemically support *E3*-mediated conjugation and ubiquitin-dependent proteolysis and physiologically play a role in DNA repair.**

The polypeptide ubiquitin plays a critical role in a variety of cell functions such as intracellular proteolysis, DNA repair, regulation of the cell cycle, protein synthesis, and response to heat shock (reviewed in Ref. 1). Most of these functions are ascribed to posttranslational linking of ubiquitin to the ε-amino groups of lysine residues of proteins (2). This isopeptide form of ubiquitination occurs by a multistep mechanism (3). Initially, the carboxyl terminus of ubiquitin is activated in a reaction coupled to the hydrolysis of ATP (4). This reaction is catalyzed by ubiquitin-activating enzyme (*E1*)[1] and results in the formation of a high energy thiol ester linkage between the carboxyl terminus of ubiquitin and a cysteine residue of the enzyme and in the production of AMP (5). Subsequently, the ubiquitin moiety is transferred, in a transthiolation reaction, to a cysteine residue of one of a family of ubiquitin carrier proteins (*E2*) (also known as ubiquitin-conjugating enzymes) (reviewed in Ref. 6, 7-9), which ligates the ubiquitin directly to proteins or requires the participation of ubiquitin-protein ligase (*E3*).

The best understood function of ubiquitin is its role in intracellular proteolysis, where certain proteins, when ligated with ubiquitin, are recognized and degraded by a 1500-kDa proteolytic complex (10-12). In reticulocyte extracts, this proteolytic pathway generally requires the presence of *E3* in the conjugation step (3). Although there are multiple *E2* isoforms, they appear to have distinct functions as only the two $E2_{14k}$ isoforms support efficiently *E3*-mediated conjugation and therefore ubiquitin-dependent proteolysis (7-9). $E2_{25k}$ appears unique in its ability to catalyze the synthesis of multiubiquitin chains (13) and $E2_{20k}$, $E2_{32k}$, (14, 15) and $E2_{230k}$ (16) catalyze polyubiquitination of histones. Although originally described as being ineffective in supporting *E3*-mediated conjugation (7), recent studies have suggested that $E2_{20k}$ and $E2_{32k}$ can do so, but at markedly lower rates than that performed by $E2_{14k}$ (8, 17). Similarly in *S. cerevisiae*, of the multiple *E2* genes identified, only mutants of the UBC4, UBC5 genes which encode almost identical 16-kDa *E2*s and of the UBC1 gene which encodes a 24-kDa *E2* show defects in proteolysis (18, 19). Mutant forms of the RAD6 gene, which encodes a 20-kDa *E2*, are defective in DNA repair, sporulation, and DNA damage-induced mutagenesis (20) and show enhanced retrotransposition (21), whereas mutations in the CDC34 gene, which encodes a 34-kDa *E2* are defective in transition from the $G_1$ to S phase of the cell cycle (22).

The distinct biochemical properties of the rabbit *E2*s and the distinct phenotypes of various yeast *E2* mutants indicate that *E2*s play a critical role in specifying the function of ubiquitin. As the $E2_{14k}$ isoforms appear unique among rabbit *E2*s in efficiently supporting *E3*-mediated conjugation and ubiquitin-dependent proteolysis (7-9), it is of great interest

* 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™/EMBL Data Bank with accession number(s) M62387.*

‡ Recipient of a Clinician-Scientist Award from the Medical Research Council of Canada. To whom reprint requests and correspondence should be addressed.

[1] The abbreviations used are: *E1*, ubiquitin-activating enzyme; *E2*, ubiquitin carrier protein or ubiquitin-conjugating enzyme; *E3*, ubiquitin-protein ligase; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; DTT, dithiothreitol.

to determine the structural basis underlying these specific biochemical properties. As an essential first step, we have isolated and sequenced a cDNA encoding the predominant reticulocyte $E2_{14k}$. In so doing, we have been able to explain the recently described ability of the yeast RAD6 gene product to support $E3$-mediated conjugation (17, 23) and have resolved confusion regarding the relationships between the mammalian and yeast $E2$s. We have also expressed the $E2_{14k}$ in *E. coli* in active form and have used the recombinant protein to characterize further the high affinity interaction between this protein and $E1$.

## MATERIALS AND METHODS

*Purification of Rabbit Reticulocyte E2 Enzymes*—Reticulocytosis was induced in female New Zealand White rabbits by phenylhydrazine injection (24). The preparation of fraction II and its separation into 0–30% and 30–85% ammonium sulfate fractions were as described (8). $E1$ and $E2$ were isolated from the 30–85% ammonium sulfate fractions by covalent affinity chromatography on an Affi-Gel ubiquitin column (8) and were then resolved by sequential chromatography on fast protein liquid chromatography Mono Q HR 5/5 and Superose 12 HR10/30 (Pharmacia LKB Biotechnology Inc.) columns (8). All five major reticulocyte $E2$s appeared as homogeneous bands on Coomassie Blue-stained SDS-polyacrylamide gels (Fig. 4). Of the two 14-kDa enzymes, only the predominant one, distinguished by its elution at 0.21 M NaCl (in the presence of 50 mM Tris, pH 7.5) (designated as $E2_{14Kb}$ in Ref. 8, $E2_{14K}$ in Ref. 9 and in this paper) on the Mono Q HR 5/5 column was purified.

*Fragmentation and Microsequencing of $E2_{14k}$*—Quantitation of the purified protein was by amino acid analysis, using acid hydrolysis followed by ninhydrin detection and analysis by a Beckman 7300 high performance analyzer. An attempt at sequencing the intact purified $E2_{14k}$ by applying 100 pmol directly to the filter disk of the sequencer was unsuccessful, indicating that the amino terminus was probably blocked. Therefore 540 pmol of $E2_{14k}$ was fragmented with cyanogen bromide (25). The fragments were resolved by SDS-PAGE on a 15% acrylamide gel and electroblotted onto a polyvinylidene difluoride membrane (Immobilon-P, Millipore) as described previously (26) with the modification that the gels were cast 3 days in advance, stored at 4 °C, prerun for 2 h at 3 mA with upper chamber buffer containing 50 μM glutathione, and run with upper chamber buffer containing 1 mM sodium thioglycolate. Such modifications appear to improve the reproducibility and yield of peptides resolved and sequenced by this method (27). Two bands on the membrane were detected by staining with Coomassie Blue, excised, and sequenced by the Edman degradation method using an Applied Biosystems 470A gas phase sequencer equipped with an on-line Applied Biosystems 120A phenylthiohydantoin analyzer. Preliminary analysis of the cyanogen bromide fragmentation products by SDS-PAGE had predicted the presence of one other smaller fragment. This smaller fragment was probably lost due to the known propensities of small protein fragments to leak from gels, to bind poorly to membranes, and to stain poorly. Since good sequence data were obtained for the two larger peptides, information about the third could be obtained by sequencing an unresolved mixture of the fragmentation products and eliminating the two known sequences from the multiple amino acids identified at each cycle of the degradation reaction. Thus, the initial 100 pmol of intact protein applied directly to the sequencer filter disc was cleaved *in situ* by applying 30 μl of cyanogen bromide (25 mg/ml in 70% formic acid) to the filter. The filter was then suspended in a 7-ml glass scintillation vial containing 100 μl of the same solution and incubated under argon in the dark at room temperature for 24 h. Subsequently the filter was dried under argon and analyzed by the sequencer.

*Cloning of a Rabbit $E2_{14k}$ cDNA*—The peptide sequence data obtained were found to be identical to that predicted from two regions of a gene encoding a human 17-kDa $E2$ (28), indicating that the human gene was probably a homologue of the desired rabbit gene. Thus, the initial approach to cloning this cDNA was to amplify a DNA fragment in a polymerase chain reaction using as primers, two oligonucleotides encoding sequences 353 base pairs apart in the human sequence (forward primer: 5′-TGGAATGCAGTTATATTTGGACCAGAAGGGAC-3′, reverse primer: 5′-ATGAATCATTCCAGCTTTGTTCAAC-3′), and as a template, first strand cDNA reverse-transcribed from total rabbit muscle RNA. A fragment of approximately 350 base pairs was obtained and purified from an agarose gel. A single-stranded DNA probe was produced in a polymerase chain reaction using the fragment as a template, [$\alpha^{32}$-P]ATP, unlabeled dCTP, dGTP, dTTP, and only the reverse primer. This probe was used to screen a rabbit neonatal skeletal muscle cDNA library in the plasmid pCDX (29) by colony hybridization (30). From 50,000 colonies, we identified one positive colony and prepared plasmid DNA from it. The insert was then subcloned into the plasmid pTZ18R (Pharmacia) at the *Bam*HI site to create plasmid p14kE2 and sequenced by the dideoxy chain termination method using the T7 polymerase sequencing kit (Pharmacia). Both strands were sequenced completely using both the deletion cloning method and new oligonucleotide primers based on previously obtained sequence data (Fig. 2*A*). Sequences were analyzed using the Microgenie (Beckman Instruments) and PC Gene (Intelligenetics) software packages.

*Expression of the $E2_{14k}$ cDNA in E. coli*—A DNA fragment containing the complete coding sequence of $E2_{14k}$ was prepared by the polymerase chain reaction using oligonucleotides corresponding to the sequences at the 5′ and 3′ ends of the protein and the cloned rabbit cDNA as template. (Forward primer: 5′-ATGTCGCCCCGGCCCGGAGGAGGCTCATG-3′; reverse primer: 5′-ATGAATCATTCCAGCTTTGTTCAAC-3′. These primers are actually derived from the human sequence (28) but encoded amino acids identical to the rabbit sequence.) This fragment was subcloned into the *Eco*RI site of the expression plasmid pKK223-3 (Pharmacia) by filling in the cohesive ends of the cut plasmid using the Klenow fragment of DNA polymerase I followed by blunt end ligation of the polymerase chain reaction-generated fragment and transformation of the plasmid into *E. coli* DH5α. Transformants were grown in Terrific Broth (30), and cell lysates were assayed for the ability to form a thiol ester linkage with $^{125}$I-ubiquitin. Plasmid DNA was prepared from a positive transformant and sequenced to confirm that the inserted DNA fragment had been correctly amplified in the polymerase chain reaction.

*Purification of Recombinant $E2_{14k}$*—To allow a detailed characterization, the recombinant protein was purified. Bacteria were grown overnight to stationary phase ($A_{600}$: 10.5) in 250 ml of Terrific Broth (30). The cells were collected by centrifugation and washed by resuspension in 375 ml of 50 mM Tris, pH 7.5, 1 mM dithiothreitol (DTT). Following centrifugation, the bacteria were resuspended in 70 ml of the buffer and lysed by sonication. The lysate was centrifuged at 10,000 × *g* for 10 min. The supernatant was then centrifuged at 100,000 × *g* for 1 h. The supernatant was concentrated to 35 ml by ultrafiltration through an Amicon YM-3 membrane. $E2$ was isolated by covalent affinity chromatography on a Affi-Gel-ubiquitin column in the same manner as was done for the redissolved 30–85% ammonium sulfate pellet from reticulocyte fraction II (8). However, purified rabbit reticulocyte $E1$ was added to the bacterial preparation to a final concentration of 10 nM to activate the immobilized ubiquitin and catalyze its linkage to $E2$. The column was eluted with AMP/pyrophosphate and DTT sequentially as described previously. The bulk of the $E2$ was eluted by DTT, and this eluate was concentrated to 7 ml prior to application on a Mono Q HR5/5 column and elution as previously described (8). The fractions containing $E2$ activity were pooled and concentrated by ultrafiltration in a Centricon-3 unit (Amicon).

*Biochemical Assays*—All assays were conducted at 37 °C. The $E1$ and $E2$ enzymes were quantitated by measuring the initial burst of radioactive pyrophosphate formed following incubation in the presence of [$\gamma^{32}$-P]ATP and ubiquitin (5, 8) or, for $E1$, by measuring the $E1$-AMP complex formed following incubation in the presence of [$^3$H]ATP and ubiquitin (5). The bovine ubiquitin (Sigma) was labeled with Na$^{125}$I to a specific activity of 13,000 cpm/pmol as described (31). Affinity of $E2$ for $E3$ was determined by measuring the initial rates of $E3$-mediated conjugation at various $E2$ concentrations. Rates of $E3$-mediated conjugation were measured by determining the initial rates of ligation of $^{125}$I-ubiquitin to endogenous reticulocyte proteins. Assays contained 1.75 mg/ml of the 0–30% ammonium sulfate fraction of fraction II as a source of both $E3$ and substrates, 50 nM $E1$, 5 μM $^{125}$I-ubiquitin, 2 mM ATP, 2 mM $MgCl_2$, 1 mM DTT, 50 mM Tris, pH 7.5, and various concentrations of $E2$. Rates were found to be linear over the first 4 min of the reaction. The products were resolved from free $^{125}$I-ubiquitin by SDS-PAGE on 10% acrylamide gels. Following autoradiography, the lanes were cut out and counted to determine the amount of ubiquitin conjugated to proteins. Affinity of $E1$ for $E2$ was determined by measuring rates of $E2$-mediated ligation of $^{125}$I-ubiquitin to histone H2A (Boehringer Mannheim) at varying $E1$ concentrations (14, 15). Assays were carried out for 25 min (rates

were linear during this time) in the presence of 120 nM *E*2, 20 μM H2A, 5 μM $^{125}$I-ubiquitin, 40 units/ml inorganic pyrophosphatase, 2 mM ATP, 10 mM $MgCl_2$, 0.2 mM DTT, 50 mM Tris, pH 7.5, and various concentrations of *E*1. The conjugation of $^{125}$I-ubiquitin was determined by SDS-PAGE resolution of the reaction products on 12.5% acrylamide gels, autoradiography, and excision and counting of the radiolabeled H2A bands. Apparent $K_m$ values were determined by double-reciprocal plots.

To test whether *E*1 and *E*2 form a complex in the presence of ATP, 30 pmol of *E*1 and 60 pmol of *E*2 were mixed in a total volume of 100 μl of 50 mM Tris, pH 7.4, 10 mM $MgCl_2$, 2 mM ATP, 0.5 mM DTT, 1 mg/ml bovine serum albumin, and 20 units/ml of inorganic pyrophosphatase for 3 min at 37 °C. The mixture was then chilled on ice and applied to a 7.5–30% (v/v) glycerol gradient (4.5 ml) that contained 50 mM Tris, pH 7.0 (measured at 20 °C), 10 mM $MgCl_2$, 2 mM ATP, 1 mM DTT, and 1 mg/ml bovine serum albumin. Following centrifugation for 16 h at 4 °C at 38,000 rpm in a Beckman SW55.1 rotor, 100-μl fractions were collected. Aliquots of 20 μl were assayed for *E*1 and *E*2 activity by ability to form thiol ester linkages with $^{125}$I-ubiquitin in the absence and presence of exogenous *E*1, respectively (see legend to Fig. 4 for details). To test whether the ubiquitin charged forms of *E*1 and *E*2 could form a complex, *E*1 and *E*2 were mixed together as above but also in the presence of 5 μM $^{125}$I-ubiquitin. Following incubation for 90 s at 37 °C, the mixture was applied to the same glycerol gradient, but lacking DTT (to avoid hydrolysis of the thiol esters), ATP, and $MgCl_2$. Following centrifugation 40-μl aliquots of the fractions were mixed with Laemmli sample buffer without mercaptoethanol and electrophoresed on 12.5% SDS-polyacrylamide gels at 4 °C. Following drying, the gels were autoradiographed to determine the positions of the ubiquitinated *E*1 and *E*2 in the glycerol gradient.

## RESULTS

*Fragmentation and Microsequencing of $E2_{14k}$*—The preparation of $E2_{14k}$ by this method (8) results in highly purified material which appeared as a single band of 15,300 kDa in size on a silver-stained denaturing polyacrylamide gel (Fig. 1) and was the smallest of the five predominant reticulocyte *E*2s (Fig. 4). Fragmentation by the cyanogen bromide method produced three peptides of approximately 11,000, 5,000, and 2,400 in molecular weight (Fig. 1). However, as discussed under "Materials and Methods," when the fragments were electroblotted onto the polyvinylidene difluoride membrane, only the two larger bands were detected by Coomassie Blue staining and were sequenced (Fig. 1). To gain sequence information of the smaller fragment, the protein was fragmented *in situ* on the filter disc of the sequencer with cyanogen bromide. From the multiple amino acids identified at each cycle of the Edman degradation reaction, the ones identified by sequencing of the two large fragments were removed, leaving the potential sequence of the lost fragment (Fig. 1). Although an unequivocal sequence could not be established by this approach, it provided data for comparison with the predicted sequence from the cloned cDNA (see below). The two definitive sequences obtained were found to be 67 and 81% identical to residues 37–51 and 78–93, respectively, of *S. cerevisiae* RAD6 (20) and completely identical to the same blocks of residues in a human 17-kDa *E*2 (28).

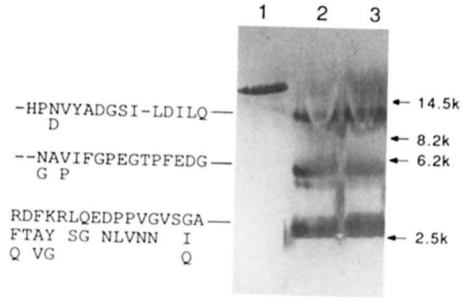


FIG. 1. **Fragmentation of rabbit and recombinant 14-kDa *E*2 by cyanogen bromide and peptide sequencing.** Approximately 20 pmol of purified rabbit and recombinant 14-kDa *E*2 were fragmented with cyanogen bromide as described previously (25). Aliquots of the purified rabbit *E*2 (*lane 1*), fragmented rabbit *E*2 (*lane 2*), and fragmented recombinant *E*2 (*lane 3*) were electrophoresed on a 20% polyacrylamide high density Phast Gel (Pharmacia) followed by silver staining. In a separate fragmentation reaction using 540 pmol, sequences were obtained from the two larger peptides and are as indicated. The possible sequences assigned to the smallest peptide is deduced from eliminating the sequences of the two larger peptides from the sequence data generated from analysis of the unresolved cyanogen bromide reaction products (see "Materials and Methods").

*Sequence of $E2_{14k}$*—Using the protein sequence data, a clone was isolated from a rabbit neonatal skeletal muscle cDNA library in the plasmid pCDX (see "Materials and Methods"). The insert and some adjoining plasmid DNA (total length approximately 1100 base pairs) was subcloned into the vector pTZ18R to produce plasmid p14kE2 (Fig. 2*A*). The actual rabbit sequence in the subclone, excluding the distal poly(A) region, is 885 nucleotides in length (Fig. 2*B*). The largest open reading frame is from bases −3 to 456. The two definitive peptide sequences determined (Fig. 1) can be found in the predicted sequence (Fig. 2*B*), indicating that this is the correct reading frame. Furthermore, the sequence predicted following the methionine at base 31 can be found in the mixture of amino acids identified by the cleavage of the protein *in situ* on the filter disk of the sequencer (Figs. 1 and 2*B*). Four methionines are present in this open reading frame at bases 1, 28, 100, and 226. As three fragments were detectable by gel electrophoresis after cyanogen bromide cleavage, translation could theoretically begin at either bases 1 or 28. (The nine-amino acid fragment encoded between bases 4 and 30 would be too small to be retained by the gel.) Initiation of translation at base 1 would predict 9 arginines in the protein, whereas initiation at base 28 would predict the presence of 6 such

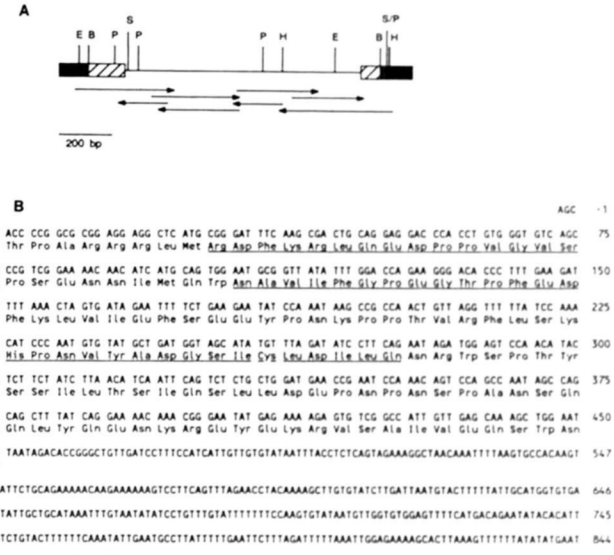


FIG. 2. **Sequence of rabbit 14-kDa *E*2 cDNA.** *A*, linear representation of the plasmid p14kE2. The *thick line* represents regions of the plasmid pTZ18R. The *hatched area* indicates plasmid pCDX sequences. The *thin line* represents the 14-kDa *E*2 cDNA. Several restriction sites are indicated. *Arrows* indicate the region and direction of sequencing. *B*, *Bam*HI; *E*, *Eco*RI; *H*, *Hin*dIII; *P*, *Pst*I; *S*, *Sal*I. *B*, nucleotide and predicted amino acid sequence of the 14-kDa *E*2 cDNA. The sequences corresponding to the cyanogen bromide generated peptides are *underlined*. The *dashed segment* in the third fragment at the cysteine residue indicates a blank cycle in the output from the sequenator. Numbering of the nucleotide sequence begins with the first residue of the coding region.

residues. Amino acid analysis of two different samples of the purified protein indicated the presence of 8.8 and 8.3 arginine residues. Thus the methionine at base 1 is the likely start of translation and this was confirmed by showing that protein generated by expression of this cDNA was indistinguishable from the purified rabbit protein (see below). This coding sequence would predict a protein of 152 amino acids with a pI of 5.7 and a molecular weight of 17,293. This is somewhat greater than the molecular weight of 15,300 estimated by gel electrophoresis (Fig. 4), but is not inconsistent considering the known inaccuracies of molecular weight estimation of small proteins and peptides by electrophoresis. Although the actual molecular weight is 17,293, we will continue to refer to this *E2* as $E2_{14k}$ until a consensus on *E2* nomenclature is reached. The 14-kDa designation is well established in this area of research and avoids confusion with another rabbit *E2* referred to as 17 kDa.

The predicted sequence contains only 1 cysteine at residue 88 and is therefore the site of the thiol ester linkage with ubiquitin. The sequence surrounding this region is strongly conserved with other *E2*s (Fig. 3). Surprisingly, 14-kDa *E2* appears homologous to the DNA repair gene RAD6 with 69% identical residues and 80% similarity if conservative substitutions are considered. By dot matrix analysis, the regions of similarity are diffuse (Fig. 3) with the notable exception of a strongly acidic carboxyl terminus present only in the RAD6 gene. The extent of similarity between 14-kDa *E2* and its presumed yeast homologues UBC4 and UBC5 and with UBC1, another yeast gene implicated in proteolysis, is markedly less impressive with the proportions of identical amino acids being only 38, 39, and 36%, respectively. Very similar sequences are found in three short regions in these proteins (Fig. 3). Of these, only two regions are common amongst all of them (surrounding residues 65 and 88 of 14-kDa *E2*). If the region containing the critical cysteine (residue 88) required for thiol ester formation, found in all sequenced *E2*s, is excluded, then the proportions of identical amino acids are only 29, 30, and 32% for UBC1, UBC4, UBC5, respectively. These values are actually slightly less than that obtained (36%) when the sequence of CDC34, an *E2* of different function (22), is similarly analyzed. The homology between the rabbit $E2_{14k}$ and the human 17-kDa *E2*, suspected after peptide sequencing, was confirmed with the finding that their protein sequences deduced from their cDNAs are identical. The nucleotide sequences in the coding regions were very similar with 95% of the bases being identical.

*Expression of $E2_{14k}$ in E. coli*—The ability to express the cloned gene is essential for future studies involving mutational analysis to determine the structural basis of the various biochemical properties of the $E2_{14k}$. The coding sequence was inserted into the pKK223 expression vector and transformed into *E. coli*. Functional *E2* was detected in the *E. coli* lysates at approximately 40 nmol/liter of bacterial culture. Approximately 70% of the enzymatic activity loaded on the Affi-Gel ubiquitin column was recovered in the DTT eluate. The protein eluted from the subsequent fast protein liquid chromatography Mono Q column at 0.21 M NaCl, confirming that this was the major isoform of $E2_{14k}$ (8, 9). The final yield was 15 nmol/liter of initial culture. The protein produced was pure as determined by Coomassie Blue staining of a polyacrylamide gel. It migrated in a manner identical to that of the purified rabbit protein (Fig. 4*A*).

As all known *E2* sequences show approximately 30–40% amino acid identity, detailed structural and functional analy-

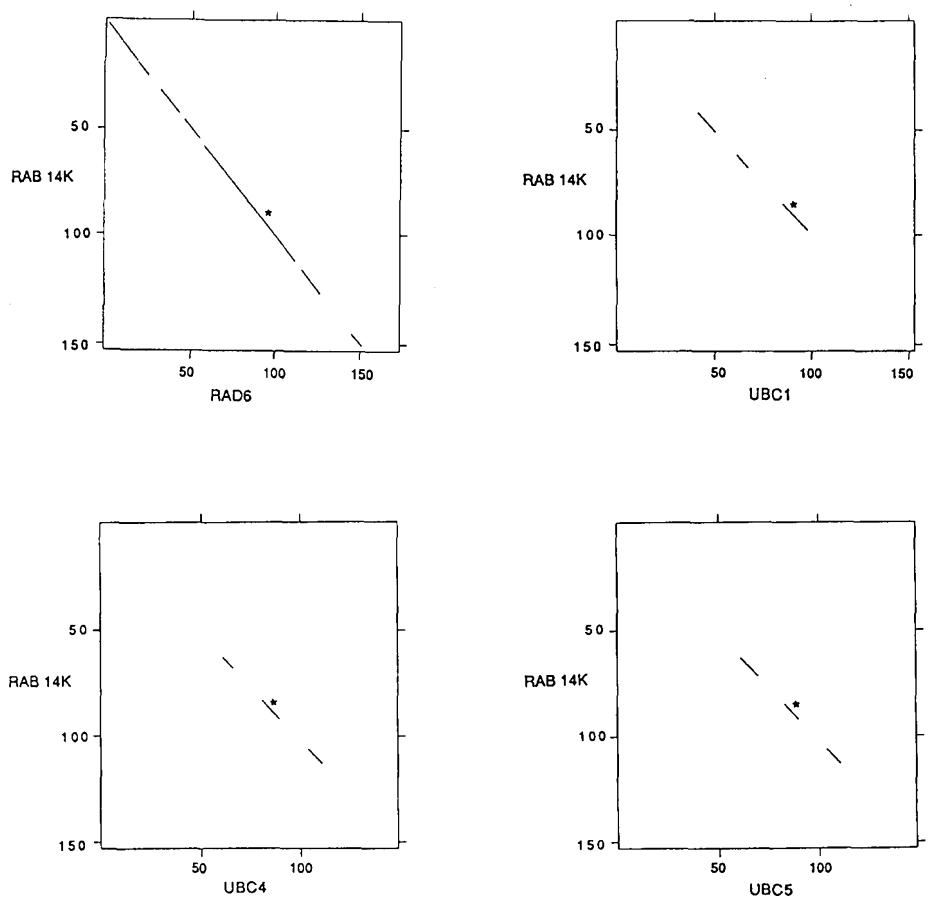


FIG. 3. **The protein sequence of rabbit 14-kDa *E2* is markedly more similar to RAD6 than to UBC1, UBC4, or UBC5.** Comparisons were made by the dot matrix method. Each dot identifies regions of similar sequence (defined as five out of six amino acids identical) in the indicated proteins. The *asterisk* denotes the region containing the cysteine required for thiol ester formation with ubiquitin. Numbering on the axes refer to amino acid residue numbers in the indicated sequence. For ease of comparison, only the first 152 amino acids of UBC1 are shown. No areas of similarity were seen in more distal regions.

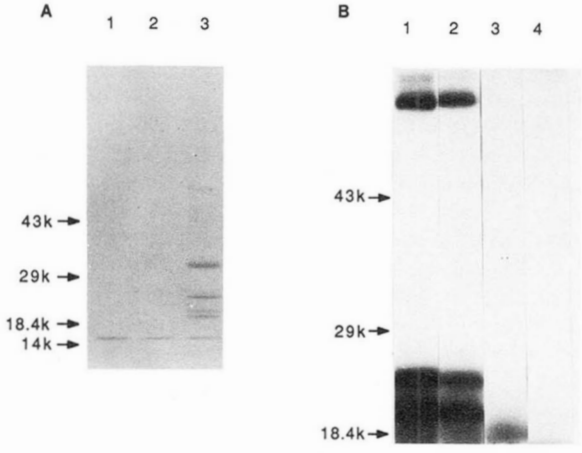


FIG. 4. **The recombinant *E2* migrates and forms a thiol ester linkage in a manner identical to that of rabbit 14-kDa *E2*.** *A*, purified recombinant *E2* (*lane 1*), rabbit 14-kDa *E2* (*lane 2*), and a mixture of the five major rabbit reticulocyte *E2*s (14, 17, 20, 24, and 32 kDa) (*lane 3*) were electrophoresed on a 10–15% gradient SDS-polyacrylamide gel (Phast Gel, Pharmacia) and stained with Coomassie Blue. *B*, thiol ester assays were performed by incubation of rabbit 14-kDa *E2* (*lane 1*) and recombinant protein (*lanes 2–4*) in the presence of 60 nM rabbit *E1*, 5 μM $^{125}$I-ubiquitin, 20 units/ml inorganic pyrophosphatase, 2 mM ATP, 10 mM $MgCl_2$, 50 mM Tris, pH 7.5, and 0.25 mM DTT for 90 s at 37 °C. To preserve the labile ubiquitin thiol ester, reactions were quenched with Laemmli sample buffer without mercaptoethanol and kept on ice until electrophoresis on a 10% SDS-polyacrylamide gel at 4 °C. The ubiquitin thiol esters were detected by autoradiography of the dried gel. In *lane 3*, ubiquitin activating enzyme (*E1*) was omitted. In *lane 4*, to hydrolyze the thiol ester, the sample was boiled for 5 min in the presence of 5% mercaptoethanol prior to loading on a separate gel run at room temperature.

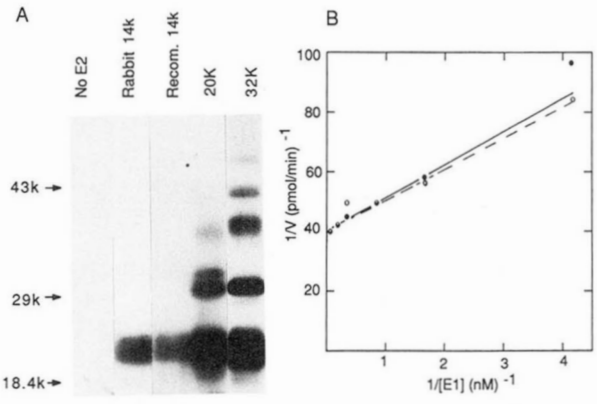


FIG. 5. **Recombinant *E2* performs monoubiquitination of histone H2A and has high affinity for *E1*.** Assays of ubiquitination of histone H2A were carried out as described under "Materials and Methods." To demonstrate the preferential monoubiquitination by recombinant 14-kDa *E2*, parallel incubations were carried out using 120 nM concentrations of each of the indicated *E2*s in the presence of 9 nM *E1*. *B*, double-reciprocal plot of the rate of histone conjugation at various *E1* concentrations using rabbit (●—●) or recombinant (○—○) purified *E2*s. The apparent $K_m$ values determined were 106 pM (rabbit) and 95 pM (recombinant).

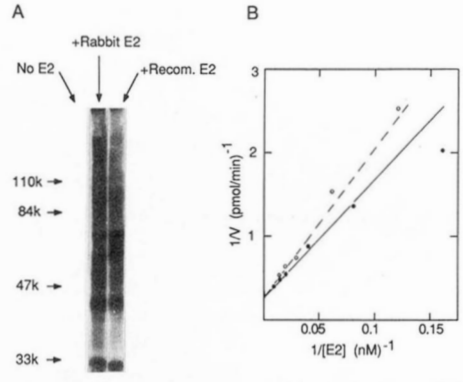


FIG. 6. **Recombinant *E2* supports *E3*-mediated ligation of ubiquitin to reticulocyte proteins and has high affinity for *E3*.** *A*, assays were performed as described under "Materials and Methods" in the absence or presence of the recombinant or rabbit proteins. *B*, double-reciprocal plot of the rate of *E3*-mediated conjugation at various concentrations of rabbit (●—●) or recombinant (○- - -○) *E2*. The apparent $K_m$ values of *E2* for *E3* determined were 53 nM (rabbit) and 62 nM (recombinant).

sis of the purified recombinant protein was carried out to rigorously identify this *E2*. The purified expressed protein was fragmented with cyanogen bromide and this yielded three fragments which migrated identically to the fragments generated from the rabbit $E2_{14k}$ protein (Fig. 1). Functionally, the recombinant protein formed a thiol ester linkage with ubiquitin in an *E1*-dependent mechanism (Fig. 4*B*).

The $E2_{14k}$ can be distinguished from other *E2*s by its preference for monoubiquitination of histone H2A (14, 15) and by its ability to efficiently support *E3*-mediated ligation (8). Accordingly, the recombinant and rabbit *E2*s carried out identical preferential monoubiquitination of histone H2A (Fig. 5*A*). This apparent preference for monoubiquitination was not simply due to short assay times or short exposure during autoradiography as parallel reactions with identical concentrations of $E2_{20k}$ and $E2_{32k}$ readily generated polyubiquitinated histone (Fig. 5 *A*). Furthermore, the $K_m$ values of *E1* for our recombinant and rabbit proteins, determined from these assays, were 95 and 106 pM, respectively (Fig. 5*B*), in excellent agreement with the value of 100 pM previously reported for the $E2_{14k}$ (14). The recombinant protein supported *E3*-mediated ubiquitination to endogenous rabbit reticulocyte proteins identically to that of rabbit $E2_{14k}$ (Fig. 6*A*). The two proteins showed identical $V_{max}$ for this reaction and the apparent $K_m$ values for binding to *E3* were similar (rabbit 53 nM, recombinant 62 nM) given the experimental variability in quantitating the reticulocyte and recombinant *E2*s in the functional burst assay (Fig. 6*B*).

As a result of the very low apparent $K_m$ of *E1* for *E2*, we evaluated whether these two enzymes form a tight complex. Following mixing in the presence of ATP and $Mg^{2+}$, *E1*, and *E2* resolve as distinct peaks 1 ml apart on a 4.5-ml 7.5–30% glycerol gradient suggesting that the ATP-bound form of *E1* does not bind tightly to *E2*. When $^{125}$I-ubiquitin is included in the mixture to generate ubiquitinated forms of *E1* and *E2*, these forms also appear, following centrifugation, as distinct peaks. Assaying the fractions in thiol ester assays with exogenous *E1* and $^{125}$I-ubiquitin confirmed that there was no uncharged *E2* in any *E1* containing fractions or any other fractions outside the ubiquitinated *E2* peak (data not shown).

## DISCUSSION

Residing within the structure of $E2_{14k}$, the smallest reticulocyte *E2* described, are three functions: a site for thiol esterification to ubiquitin, an *E1* binding site, and an *E3* binding site (7, 14, 32). The thiol esterification site is easily localized as there is only one cysteine, at residue 88, in the determined sequence. This cysteine residue is situated in a cluster of amino acids that is highly conserved among all *E2*s of known sequence and suggests that this region likely serves a function common to all *E2*s. This is likely to be related to its interaction with *E1*, either with the transfer of ubiquitin from *E1* or with the binding of *E1*.

This interaction of $E2_{14k}$ with $E1$ is of very high affinity (Fig. 5, Ref. 14), suggesting the possibility of complex formation. The formation of ubiquitin charged $E1$ follows a sequence (33) in which ATP binding is followed by ubiquitin binding. A tightly bound ubiquitin-adenylate is formed with the release of pyrophosphate. The $E1$-ubiquitin adenylate complex is converted into an $E1$ with ubiquitin bound in a thiol ester linkage to an $E1$ cysteine residue and AMP is released. A second ubiquitin is then activated to form a ternary complex of $E1$ containing two ubiquitin moieties per molecule of $E1$, one bound in a thiol ester linkage and the other bound noncovalently as ubiquitin-adenylate. Transfer of ubiquitin to $E2$ occurs from the thiol ester bound moiety. The native forms of these enzymes do not appear to interact as they resolve on gel filtration chromatography during the purification (7, 8). We have demonstrated that $E1$ bound with ATP, the first cofactor to bind to $E1$ during its thiol esterification (33), does not complex with $E2$. We also observed that the mixture of the ubiquitin charged forms of $E1$ and $E2$ do not form a complex. This data provide experimental support for a previous hypothesis that $E1$ has higher affinity for $E2$ than for ubiquitin charged $E2$ (14). Thus the high affinity observed is likely related to the binding of ubiquitin charged $E1$ with the uncharged form of $E2$. As ubiquitin charging of $E1$ is rapid (33), the binding of $E2$ is probably to the ternary form of $E1$. This ternary $E1$-$E2$ complex would be difficult to demonstrate as it is certainly transient as shown by the ability of $E1$ to ubiquitinate rapidly an excess of $E2$ within 60 to 90 s (Fig. 4, Refs. 7–9). This rapid ubiquitination of $E2$ prevents a stable complex and permits $E1$ to ubiquitinate the various $E2$ isoenzymes, which in sum, in reticulocytes, appear to be in severalfold excess with respect to $E1$ (8).

The third functional property of $E2_{14k}$ is its ability to bind tightly to $E3$ (Fig. 6, Ref. 17). Although it is not known whether the low apparent $K_m$ observed represents binding of ubiquitinated $E2$ or uncharged $E2$ to $E3$, the two native enzymes can form a complex (32). RAD6, which has recently been reported to support $E3$-mediated conjugation (17, 23), differs strikingly from $E2_{14k}$ in its possession of an acidic carboxyl-terminal extension. Thus, this acidic tail of RAD6 is unessential in this interaction with $E3$, and this has been confirmed by studies using tailless forms of the enzyme (23). Sequences that are important in this function are likely located outside of the highly conserved regions. Our demonstrated ability to express the active protein in large quantities will permit the construction and expression of mutant forms of $E2_{14k}$ to critically determine the locations of these binding sites.

Our findings resolve current confusion regarding the structural and functional relationships of RAD6 to several mammalian $E2$s. RAD6 $E2$ has been suggested to be a homologue of the rabbit $E2_{20k}$ (6) based on similarity in size and ability to polyubiquitinate histones. UBC4 and UBC5 have been proposed to be homologs of rabbit $E2_{14k}$ based on similarities in size and their common roles in intracellular proteolysis (6, 19). Thus the recently described ability of the RAD6 $E2$ to support $E3$-mediated conjugation and ubiquitin-dependent proteolysis (17, 23), the key biochemical characteristics of rabbit $E2_{14k}$, was surprising in view of the differences between RAD6 and UBC4/UBC5, the putative homologs of $E2_{14k}$, in both structure and in phenotypes of mutant strains (18, 20). rad6 mutants are defective in DNA repair, DNA damage-induced mutagenesis, and sporulation (20), whereas UBC4 and UBC5 mutants show defects in degradation of short-lived and abnormal proteins (18). To add further confusion, a human 17-kDa $E2$ has been shown to be 70% identical in sequence to RAD6 (28, 34, 35) and able to complement rad6 mutants (34), indicating that these two genes are indeed homologous. Although biochemically uncharacterized, this human 17-kDa $E2$ has been proposed (28, 35) to be a homologue of rabbit $E2_{17k}$. However, rabbit $E2_{17k}$ is strikingly dissimilar to RAD6 protein in being unable to support $E3$-mediated ubiquitination or conjugation to histones (8, 14, 15).

Thus, in observing the marked similarity of the $E2_{14k}$ sequence to that of RAD6, the ability of RAD6 $E2$ to support $E3$-mediated conjugation and ubiquitin-dependent proteolysis is explained. The observation of sequence identity between rabbit $E2_{14k}$ and the human 17-kDa $E2$ which complements the DNA repair defects of rad6 mutants indicates that rabbit $E2_{14k}$ is a functional homolog of RAD6 and may play a role in DNA repair in mammalian cells. This identity between rabbit $E2_{14k}$ and the human 17-kDa $E2$ now permits the assignment of the former's biochemical properties to the latter and corrects the latter's misassignment as a homolog of rabbit $E2_{17k}$ (28, 35).

The sequence and functional homologies between RAD6 and rabbit 14-kDa $E2$ now render doubtful the previous presumption of homology between $E2_{14k}$ and UBC4 and UBC5 (6, 18). In support of this we have, using oligonucleotides derived from the yeast UBC4 sequence as primers, amplified from reverse-transcribed rat muscle RNA, a 0.4-kilobase pair DNA. The deduced sequence is distinct from rat $E2_{14k}$, but 83% identical to the corresponding region in UBC4,[2] and therefore this DNA may encode part of a true homologue of UBC4. However, the notion that $E2_{14k}$ and UBC4/UBC5 are unlikely to be true functional homologues should not necessarily be viewed as arguing against a role of $E2_{14k}$ in intracellular proteolysis. Rather, the relative lack of sequence similarity of $E2_{14k}$ to UBC1, UBC4, or UBC5 may imply that degradation of different substrates and/or binding to different $E3$s involves structurally distinct $E2$s.

Our inferences regarding the structure-function relationships of $E2_{14k}$ and our resolution of this confusion among $E2$s were dependent on clear evidence that our cloned cDNA does encode rabbit $E2_{14k}$. Since all $E2$s share approximately 30–40% identity in sequence, extensive structural and functional studies of the recombinant protein were performed and are essential for unequivocal identification of this cDNA. The findings can be summarized as follows and offer overwhelming evidence that this is the correct primary sequence for rabbit $E2_{14k}$. (*a*) Well established purification protocols for $E2_{14k}$ were used and produced protein which was highly purified and bore physical charge and functional characteristics similar to those reported previously. (*b*) The protein sequence data (involving 32% of the protein sequence) derived from the highly purified protein is completely consistent with the predicted sequence. (*c*) Structurally, the expressed and rabbit purified proteins migrated identically on gel electrophoresis, generated an identical peptide pattern following cyanogen bromide cleavage, and eluted at the same salt concentration on ion exchange chromatography. (*d*) Functionally, the expressed and rabbit purified proteins were indistinguishable with respect to their formation of ubiquitin thiol ester linkages, in their affinities for $E1$ and $E3$ and in their abilities to monoubiquitinate histone H2A and to support more generalized ubiquitination of proteins by $E3$. (*e*) The elution properties on ion exchange chromatography (8, 9), the $K_m$ of $E1$ for $E2$ (14), the preference for monoubiquitination (14, 15), and the efficient support of $E3$-mediated ligation (8) are similar to those reported by other groups and therefore essen-

[2] S. S. Wing and D. Banville, unpublished data.

tially excludes the possibility that the incorrect protein was purified.

Our findings now permit the assignment of *S. cerevisiae* RAD6, human 17-kDa *E2*, and rabbit $E2_{14k}$ to a subclass of *E2*s which structurally have highly conserved sequences, which biochemically mediate *E3*-dependent ligation and ubiquitin-dependent proteolysis and which physiologically are involved in DNA repair. The *Schizosaccharomyces pombe* rhp6$^+$ gene (36) and one other human RAD6 homolog (34), although they have not been tested biochemically for *E2* activity, are structurally and functionally homologous to RAD6 and are therefore likely to be members of this subclass. The exact roles of *E3*-dependent ligation and proteolysis in the mechanisms underlying DNA repair are issues of significant interest that remain to be addressed.

*Acknowledgments*—We are grateful to Dr. A. Haas for helpful advice regarding the purification and characterization of the *E2*s. Dr. D. H. MacLennan generously provided the rabbit neonatal skeletal muscle cDNA library. We thank Dr. Y. Konishi for protein sequencing. During the preparation of the manuscript, helpful advice was provided by Drs. D. Thomas, K. Murthy, and A. Haas.

## REFERENCES

1. Rechsteiner, M. (ed) (1988) *Ubiquitin*, Plenum Press, New York
2. Hershko, A., Ciechanover, A., Heller, H., Haas, A. L., and Rose, I. A. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 1783–1786
3. Hershko, A., Heller, H., Elias, S., and Ciechanover, A. (1983) *J. Biol. Chem.* **258,** 8206–8214
4. Ciechanover, A., Heller, H., Katz-Etzion, R., and Hershko, A. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 761–765
5. Haas, A. L., Warms, J. V. B., Hershko, A., and Rose, I. A. (1982) *J. Biol. Chem.* **257,** 2543–2548
6. Jentsch, S., Seufert, W., Sommer, T., and Reins, H.-A. (1990) *Trends Biochem. Sci.* **15,** 195–198
7. Pickart, C. M., and Rose, I. A. (1985) *J. Biol. Chem.* **260,** 1573–1581
8. Haas, A. L., and Bright, P. M. (1988) *J. Biol. Chem.* **263,** 13258–13267
9. Pickart, C. M., and Vella, A. T. (1988) *J. Biol. Chem.* **263,** 12028–12035
10. Eytan, E., Ganoth, D., Armon, T., and Hershko, A. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 7751–7755
11. Hough, R., Pratt, G., and Rechsteiner, M. (1987) *J. Biol. Chem.* **262,** 8303–8313
12. Driscoll, J., and Goldberg, A. L. (1990) *J. Biol. Chem.* **265,** 4789–4792
13. Chen, Z., and Pickart, C. M. (1990) *J. Biol. Chem.* **265,** 21835–21842
14. Haas, A. L., Bright, P. M., and Jackson, V. E. (1988) *J. Biol. Chem.* **263,** 13268–13275
15. Pickart, C. M., and Vella, A. T. (1988) *J. Biol. Chem.* **263,** 15076–15082
16. Klemperer, N. S., Berleth, E. S., and Pickart, C. M. (1989) *Biochemistry* **28,** 6035–6041
17. Haas, A. L., Reback, P. B., and Chau, V. (1991) *J. Biol. Chem.* **266,** 5104–5112
18. Seufert, W., and Jentsch, S. (1990) *EMBO J.* **9,** 543–550
19. Seufert, W., McGrath, J. P., and Jentsch, S. (1990) *EMBO J.* **9,** 4535–4541
20. Jentsch, S., McGrath, J. P., and Varshavsky, A. (1987) *Nature* **329,** 131–134
21. Picologlou, S., Brown, N., and Liebman, S. W. (1990) *Mol. Cell. Biol.* **10,** 1017–1022
22. Goebl, M. G., Yochem, J., Jentsch, S., McGrath, J. P., Varshavsky, A., and Byers, B. (1988) *Science* **241,** 1331–1335
23. Sung, P., Berleth, E., Pickart, C., Prakash, S., and Prakash, L. (1991) *EMBO J.* **10,** 2187–2193
24. Rabinowitz, M., and Fisher, J. M. (1964) *Biochim. Biophys. Acta* **91,** 313–322
25. Charbonneau, H. (1989) in *A Practical Guide to Protein and Peptide Purification for Microsequencing* (Matsudaira, P. T., ed) pp. 15–30, Academic Press, New York
26. Matsudaira, P. (1987) *J. Biol. Chem.* **262,** 10035–10038
27. Moos, M., Jr., Nguyen, N. Y., and Liu, T.-Y. (1988) *J. Biol. Chem.* **263,** 6005–6008
28. Schneider, R., Eckerskorn, C., Lottspeich, F., and Schweiger, M. (1990) *EMBO J.* **9,** 1431–1435
29. Brandl, C. J., Green, N. M., Korczak, B., and MacLennan, D. H. (1986) *Cell* **44,** 597–607
30. Sambrook, J., Fritsch, E. M., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor Press, Cold Spring Harbor, NY
31. Ciechanover, A., Heller, H., Elias, S., Haas, A. L., and Hershko, A. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77,** 1365–1368
32. Reiss, Y., Heller, H., and Hershko, A. (1989) *J. Biol. Chem.* **264,** 10378–10383
33. Haas, A. L., and Rose, I. A. (1982) *J. Biol. Chem.* **257,** 10329–10337
34. Koken, M. H., Reynolds, P., Jaspers-Dekker, I., Prakash, L., Prakash, S., Bootsma, D., and Hoeijmakers, J. H. J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 8865–8869
35. Wolffendin, C., Chen, Z., Staskus, K., Retzel, E. F., and Plagemann, P. G. W. (1991) *Biochim. Biophys. Acta* **1090,** 81–85
36. Reynolds, P., Koken, M. H. M., Hoeijmakers, J. H. J., Prakash, S., and Prakash, L. (1990) *EMBO J.* **9,** 1423–1430