Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 8865–8869, October 1991
Biochemistry

Structural and functional conservation of two human homologs of the yeast DNA repair gene *RAD6*

(ubiquitin conjugation/E2 enzyme/DNA damage/UV mutagenesis/sporulation)

MARCEL H. M. KOKEN*, PAUL REYNOLDS†, IRIS JASPERS-DEKKER*, LOUISE PRAKASH†, SATYA PRAKASH‡§, DIRK BOOTSMA*, AND JAN H. J. HOEIJMAKERS*

*Department of Cell Biology and Genetics, Erasmus University, P.O. Box 1738, 3000DR Rotterdam, The Netherlands; †Department of Biophysics, University of Rochester School of Medicine, Rochester, NY 14642; and ‡Department of Biology, University of Rochester, River Campus Station, Rochester, NY 14627

Communicated by James V. Neel, July 1, 1991

ABSTRACT **The *RAD6* gene of *Saccharomyces cerevisiae* encodes a ubiquitin-conjugating enzyme (E2) that is required for DNA repair, damage-induced mutagenesis, and sporulation. We have cloned the two human *RAD6* homologs, designated *HHR6A* and *HHR6B*. The two 152-amino acid human proteins share 95% sequence identity with each other and ≈70% and ≈85% overall identity with the homologs from yeasts (*S. cerevisiae* and *Schizosaccharomyces pombe*) and *Drosophila melanogaster*, respectively. Neither of the human RAD6 homologs possesses the acidic C-terminal sequence present in the *S. cerevisiae* RAD6 protein. Genetic complementation experiments reveal that *HHR6A* as well as *HHR6B* can carry out the DNA repair and mutagenesis functions of *RAD6* in *S. cerevisiae rad6Δ* mutants.**

The *Saccharomyces cerevisiae RAD6* gene plays a key role in DNA repair and DNA damage-induced mutagenesis. *rad6* mutants are extremely sensitive to a plethora of DNA-damaging agents, including UV irradiation, x-rays, and alkylating agents; they are defective in postreplication repair of UV-damaged DNA, in mutagenesis induced by DNA damaging agents, and in sporulation (for a review, see ref. 1). Transposition of Ty elements is enhanced in *rad6* mutants (2). *RAD6* encodes a protein of 172 amino acids (3) with a globular domain consisting of approximately the first 149 amino acids and an extended, predominantly acidic C terminus (4). RAD6 is a ubiquitin-conjugating enzyme (5, 6). Ubiquitin, a highly conserved, 76-amino acid polypeptide is covalently attached to many cellular proteins and targets them for selective degradation, (re)folding, or stabilization (for recent reviews, see refs. 7–9). Ubiquitination is carried out by a family of proteins in a multistep reaction involving a ubiquitin-activating enzyme (E1) that binds and subsequently transfers a ubiquitin moiety to one of a set of ubiquitin-conjugating enzymes (E2). The E2 enzyme ligates ubiquitin directly to a target protein, with or without the help of a ubiquitin protein ligase (E3). RAD6 polyubiquitinates histones H2A and H2B *in vitro* without the involvement of E3, and the acidic domain of RAD6 is required for multiple ubiquitination of histones (6). It is possible that the protein is implicated in modifying chromatin structure as part of the processes that are disturbed in a *rad6* mutant, including repair and mutagenesis.

RAD6 is highly conserved among eukaryotes. Previously, we cloned the *rhp6*$^+$ gene of the fission yeast *Schizosaccharomyces pombe* and showed that it is a structural and functional homology of *RAD6* (10). We have also isolated a *RAD6* homolog, *Dhr6*, from *Drosophila melanogaster* (11). In this paper, we present the cloning, sequence analysis, and functional studies with the *RAD6* homologs from human.¶ In contrast to yeast and *Drosophila*, where *RAD6* is a single copy gene, interestingly, in human, the *RAD6* homologous gene is duplicated, and the proteins encoded by the two genes *HHR6A* (human homolog of RAD6) and *HHR6B* share 95% identical amino acid residues. We also show that the *HHR6A* and *HHR6B* genes complement the DNA repair and UV mutagenesis defects of the *S. cerevisiae rad6* mutant.

MATERIALS AND METHODS

Restriction Enzyme Digests and Southern Blot Library Hybridizations. Restriction enzyme digestions were performed according to the manufacturer's descriptions. Blots were prepared on Zeta-Probe (Bio-Rad) using the alkaline-blotting procedure as recommended by the manufacturer. DNA probes were labeled by the random-priming method as described (12). Unless stated otherwise, hybridization of *Sc. pombe* and *Drosophila* probes to human DNA occurred overnight at 55°C and hybridization of human probes to human DNA was at 65°C in a hybridization mixture containing 10× Denhardt's solution (2% Ficoll/2% bovine serum albumin/2% polyvinylpyrrolidone)/10% dextran sulfate/0.1% SDS/3× standard saline citrate (SSC)/50 mg of sonicated salmon sperm DNA per liter. Washings for hybridizations involving different species were performed for 5 min in 3× SSC twice and for 5 min in 1× SSC once at 55°C. For hybridizations within a species, washings were done twice for 20 min each in 3× SSC, twice for 20 min each in 1× SSC, and twice for 20 min each in 0.3× SSC at 65°C.

Northern Blotting and Nucleotide Sequence Analysis. Isolation of total RNA by the LiCl/urea method, preparation of poly(A)$^+$ RNA by two consecutive passages over oligo(dT) cellulose columns, and Northern blotting protocols were all according to Sambrook *et al.* (13). Sequence analysis on double-stranded DNA was done by the T7 polymerase modification (Pharmacia) of the dideoxynucleotide chain-termination method (14) using sequence-derived oligonucleotides and exonuclease III prepared deletion clones for sequencing both strands. For separation of the fragments we used Hydrolink (AT Biochem, Malvern, PA) sequencing gels.

Yeast Strains, Media, and Genetic Analyses. The *S. cerevisiae* haploid strains used in this study were the *rad6Δ* strains EMY1 (*MAT***a** *leu2-3 leu2-112 trp1 ura3-52 rad6::LEU2*$^+$) and EMY8 (*MATα ade5 his7 leu2-3 lys1 met14 pet15 ura3 trp1::URA3*$^+$ *rad6::LEU2*$^+$). The *rad6Δ/rad6Δ* diploid EMY28 was constructed by mating EMY1 and EMY8. UV irradiation conditions and media for determining survival and mutagenesis after UV exposure and sporulation

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

Abbreviations: ORF, open reading frame; UTR, untranslated region.
§To whom reprint requests should be addressed.
¶The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M74524 and M74525).

media were as described (4). Standard genetic techniques for *S. cerevisiae* (15) were used.

Plasmids. The following yeast plasmids were used in this study: pR67 contains the *RAD6* gene within a 2-kilobase (kb) *Hin*dIII/*Bam*HI DNA fragment inserted into the yeast *CEN4* plasmid YCp50 as described (4). pR611 is derived from pR67 by deleting the 0.6-kb *RAD6 Eco*RI fragment containing the entire *RAD6* open reading frame (ORF). Plasmids pR67 and pR611 are maintained in yeast as low copy plasmids. For expression of the human *RAD6* homologs *HHR6A* and *HHR6B* in *S. cerevisiae*, the human genes were cloned downstream of the highly expressed *S. cerevisiae ADC1* promoter in the yeast expression vector described previously (16), yielding plasmids pRR510 and pRR518, respectively.

RESULTS

Cloning of Human cDNAs Cross-Hybridizing to *RAD6* Derivatives. Southern and Western blot analyses indicated that the *RAD6* gene and protein are conserved in eukaryotes. For cloning the human *RAD6* homolog, a human λ cDNA library prepared from human testis RNA was screened. One set of filters was hybridized with the *Drosophila Dhr6* probe and a duplicate filter set was hybridized with the *Sc. pombe rhp6*$^+$ gene probe. Of the many clones cross-hybridizing, those reacting to some extent with both probes (≈30 in 10^6 plaques) were picked and examined by restriction enzyme analysis. Unexpectedly, restriction maps indicated the presence of two classes, corresponding to the *HHR6A* and *HHR6B* genes as described below.

Northern Blot Analysis. Representative cDNA inserts of each class were hybridized to Northern blots containing total or poly(A)$^+$ RNA from various cell lines. As shown in Fig. 1 (lane 2) the *HHR6A* probe detects transcripts of 1.7 and 0.8 kb in HeLa RNA; the *HHR6B* probe (lane 1) mainly hybridizes to a mRNA species of 4.4 kb. Similar results were obtained with RNA from the myelocytic cell line K562, a primary human fibroblast line and mouse and rat tissues (data not shown). Clones of each class with insert sizes expected for full-length cDNAs (two for the two transcripts of *HHR6A*, one for the 4.4-kb RNA of *HHR6B*) were selected for further analysis.

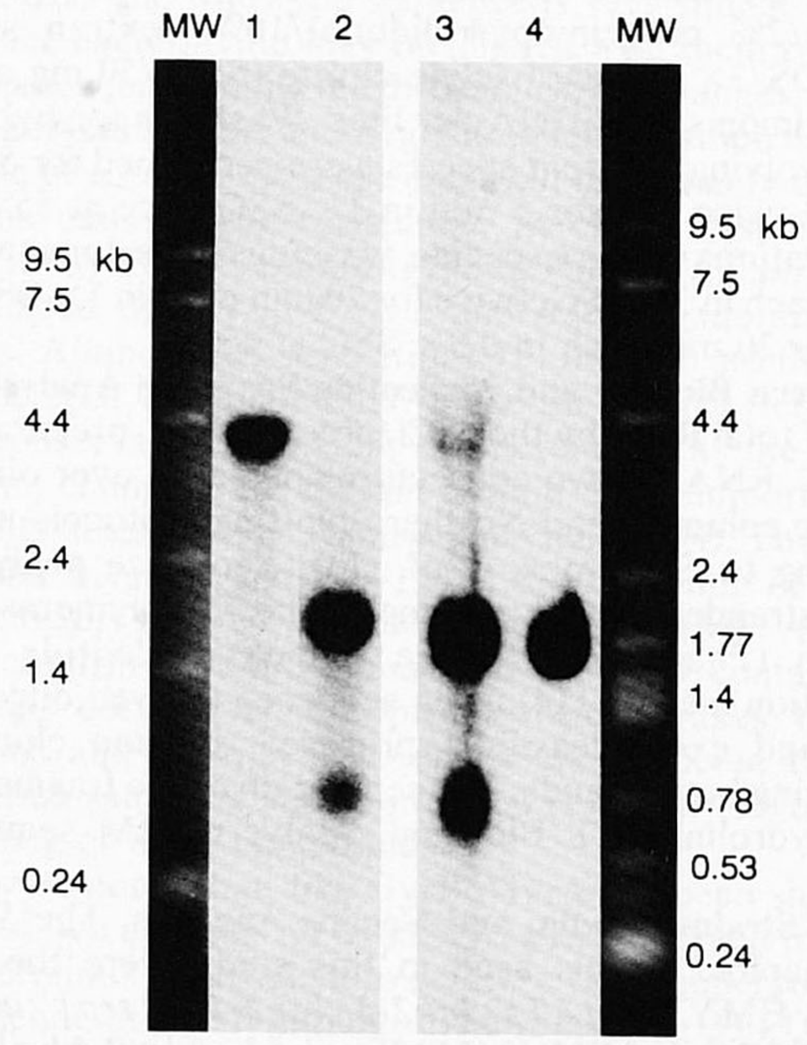
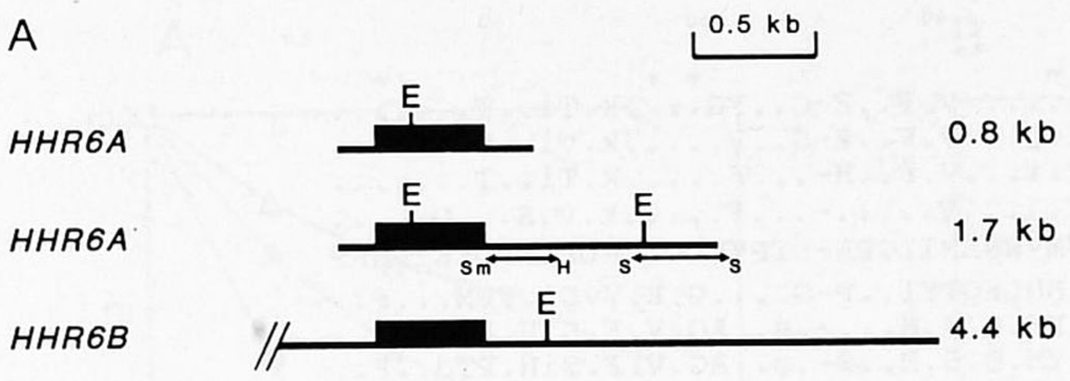


FIG. 1. Northern blot analysis of human *HHR6A* and *HHR6B* transcripts. Poly(A)$^+$ RNA was size fractionated on a 1% agarose gel containing formaldehyde. Lanes: MW, RNA molecular size markers; 1, hybridization with a *Pvu* II/*Hin*dIII fragment containing the coding region of *HHR6B*, recognizing a 4.4-kb mRNA; 2, hybridization with a probe (1.2-kb *Eco*RI fragment) of human *HHR6A* cDNA, which hybridizes to both a 0.8- and a 1.7-kb mRNA; 3 and 4, hybridization with the 293-bp *Sma* I/*Hin*dIII and the 392-bp *Sac* I *HHR6A* cDNA probe, respectively (see Fig. 2 *A* and *B* for location on the *HHR6A* cDNA maps).

Nucleotide and Predicted Amino Acid Sequences of *HHR6A* and *HHR6B*. Restriction maps for the three cDNA inserts representing the *HHR6A* and *HHR6B* genes are shown in Fig. 2*A*. The nucleotide and deduced amino acid sequences of the regions of interest in *HHR6A* and *HHR6B* are shown in Fig. 2 *B* and *C*, respectively.

Sequence analysis of the *HHR6A* cDNAs indicates that the minor 0.8-kb mRNA is identical to the major 1.7-kb mRNA species for the 5′ untranslated region (UTR), ORF, and the first part [≈100 base pairs (bp)] of the 3′ UTR. The 3′ UTR of the larger transcript extends for an additional ≈1 kb. This mRNA species harbors an AATAAA polyadenylylation signal (directly followed by a suboptimal one: AATAAC) 13 bp before the presumed polyadenylylation site (18, 19). The short 0.8-kb transcript also has two potential but suboptimal polyadenylylation signals close to the 3′ terminus. Hence, the difference between the two *HHR6A* transcripts can be explained as a result of alternative polyadenylylation site selection. This is confirmed by the Northern blot hybridization shown in Fig. 1 (lanes 3 and 4) using 3′ UTR probes derived from the region common to both cDNAs (293-bp *Sma* I/*Hin*dIII probe; lane 3) and from the area unique to the 1.7-kb species (392-bp *Sac* I probe; lane 4). The *HHR6A* sequence contains a single long ORF that encodes a protein of 152 amino acids with a calculated M_r of 17,243. The *HHR6B* ORF specifies a protein of 152 amino acids with a calculated M_r of 17,312. The expected sizes of both proteins are consistent with the results of the Western blot analysis (data not shown). The *HHR6B*-encoded protein shares a high degree of identity (95%) with the HHR6A amino acid sequence with only eight amino acid substitutions, two of which are conservative changes (Fig. 3, top two lines). At the nucleotide level, the coding sequence is much less conserved (80%) and the 5′ and 3′ UTR sequences are very different. The HHR6B protein sequence is identical to the predicted gene product [termed E2 (M_r 17,000)] of a partial cDNA clone recently described by Schneider and coworkers (17). As shown by the alignment in Fig. 3 (*Upper*) and the quantitative data summarized in Fig. 4, both the HHR6A and HHR6B polypeptides share extensive amino acid sequence similarity with RAD6 homologs of other species. However, both human proteins resemble those of *Drosophila* and *Sc. pombe* in lacking the acidic C terminus characteristic of *S. cerevisiae* RAD6. In addition, there is significant similarity to the other ubiquitin-conjugating enzymes (Fig. 3 *Lower*).

Functional Complementation of the *rad6* Mutation of *S. cerevisiae* by the *HHR6A* and *HHR6B* Genes. The high degree of amino acid identity between the *HHR6*- and *RAD6*-encoded proteins suggests that the human genes function similarly to *RAD6*. To examine this possibility, we determined whether the *HHR6A* and *HHR6B* genes complement the DNA repair, UV mutagenesis, and sporulation defects of the *rad6*Δ mutation of *S. cerevisiae*. Plasmids pRR510 and pRR518, containing the human *HHR6A* and *HHR6B* genes fused to the yeast *ADC1* promoter, respectively, were introduced into the *S. cerevisiae rad6*Δ strain EMY8 by transformation and the level of HHR6 protein was examined by Western analysis using affinity-purified anti-rad6-149 antibodies. The level of HHR6A and HHR6B proteins in *rad6*Δ cells was somewhat higher than the level of RAD6 protein present in the wild-type *S. cerevisiae* cells (data not shown). Both the *HHR6A* and *HHR6B* genes substantially increase the UV resistance of the *rad6*Δ strain. At 10 J/m^2, the survival of the *rad6*Δ strain is enhanced >500-fold and



B

```
GCGGTTCCTCTGGGTGCTTCCGCCTCCCCTTCTCCTGCTTCTCCAGCCTCTTCGGCCTCC  - 61
TCGCCCGCCGCGGGAACCCGAGACCCCAGTGTATGCCCCACCCCTGACCCCGCTCGCGAC
1   M  S  T  P  A  R  R  R  L  M  R  D  F  K  R  L  Q  E  D  P
ATGTCCACCCCGGCTCGGCGGCGCCTCATGCGGGACTTCAAGAGGTTGCAGGAGGATCCT   60
     P  A  G  V  S  G  A  P  S  E  N  N  I  M  V  W  N  A  V  I
CCAGCCGGAGTCAGCGGGGCTCCGTCCGAGAACAACATAATGGTGTGGAACGCGGTCATT
41   F  G  P  E  G  T  P  F  G  D  G  T  F  K  L  T  I  E  F  T
TTCGGGCCTGAAGGGACCCCGTTTGGGGATGGAACATTTAAACTTACAATAGAATTCACT  180
     E  E  Y  P  N  K  P  P  T  V  R  F  V  S  K  M  F  H  P  N
GAAGAATATCCAAATAAACCACCTACAGTTAGATTTGTCTCTAAGATGTTCCATCCAAAT
81   V  Y  A  D  G  S  I  C  L  D  I  L  Q  N  R  W  S  P  T  Y
GTCTATGCAGATGGTAGTATATGTCTGGACATACTTCAGAACCGTTGGAGTCCAACCTAT  300
     D  V  S  S  I  L  T  S  I  Q  S  L  L  D  E  P  N  P  N  S
GATGTGTCTTCCATTCTAACATCCATACAGTCTCTGTTGGATGAACCCAATCCCAATAGT
121  P  A  N  S  Q  A  A  Q  L  Y  Q  E  N  K  R  E  Y  E  K  R
CCAGCAAACAGCCAGGCTGCTCAGCTGTACCAGGAGAACAAACGGGAATATGAAAAGCGT  420
     V  S  A  I  V  E  Q  S  W  R  D  C  *
GTTTCTGCAATAGTAGAACAAAGCTGGCGTGATTGTTGACCCCGGGTACAGTTTAAAGAA
                                    SmaI
GCTGGCCATAAGAAAAATATATATTGATGTGTTTGTCACCTCCCTACTCCTGTCATTACA  540
TTTACTTTATTAAAAGCAAAATAACTGTTGTGCTGTTTCCATCTTCCTTGCCAAGTTTTC
CTACCCCTTCTACCCTCTCCTTAAACATCAGAAAACACCCTCTATGAAATCAAATGTACT  660
GTACCTGGGTTACTTGCAAAAATTACTAATGCTTCAGTTTTTCTGTTGTATTTCATTTCC
AGTTTTCAGGCAGTTATTTTTATTATTGTACTTTAAGCTTTTAAGATGAATTGTTATACA  780
                              HindIII
AGAGGTGCTTATGCTTAGCTTGATGACCAGGATGTTATTTTTAACAAAATGATTGCTGAA
GTGTTTCATCCTGGCTGGTCCTTCACTTGTGTTGGATTTAGAAGTGAATGTGTTTGGAAT  900
ATGGCCTACAGAGAATAGAAACAAATCCATGTAAACAATTTTGAAGGAGGCATGGGAGCT
AAAAATCCTGTGATACTAAGATCTCAGTCATATGAATTACAACGTAGTATTTACTGGCAA 1020
GAAGGAGAAAGTTGAAGGACTCAGCTAAAGGAGTACAGCAATTGTAGTAACTGACACATC
CTCTCTTTGCAAGCTGCTGACTGGGCACACTCATGCCAAGTTTCAGAATTATTGGTCTTC 1140
TGGGTTTTTGCTTTTTAAAAGAGGTGTGGGAGCAGAGGAATGGAAACAATCGTGAGTTTT
TGAGCTAGGGAAAGTTGGAGCTCCTTTAATCTTTTTAAAGGATCAGTGCTGCCCTAAGTG 1260
                          SacI
AATAAACTCAATTGTCCATCTTTATTTTAGAGTTTTAATGAATTCAAGGAAGGGAGCATA
GCATATCTGTGGCAAACTATTTTCCACTCAAATCCTGAGTTATTGCTGCATGCTTTAATT 1380
TCTTCCCTTTCAGCATCTGAGAACCTTAAAGCCAATGTCTGCGATCTTTTTTTGGATATT
TATACTTTTAGATATATAGTACCTTTAAGTAGCAGTATGGGACAAGGCTTGTAAATGTTT 1500
TGTCTAATGTTCTATTGTCACCTTTTATGCATTTATCACTTCCAAATCTAACTTTGCACA
AGTAACCCATGTAAAAAAAAATGTACATTTTTCAAAAGTTGTAAATAAAAATAACCTTAA 1620
AAA... SacI#
```

C

```
     ...CGTCCCTCTTTCTCCACCCCGCCGAGCCTAAACTAGTGACGGGGAGGGAGAC
GGGATAGTGTTTCTGTTTCCGTGGTCTTTGAATCCACAACCTCTAGTCTGAACACAGAGA -310
ACACAAATTTGTCGATTACTTCAGGAAGCCCAGGGGGAACCGCGCCCCGCTGAAGGGCTC
GGGGAAGGCGCGCTCCCGCGTCCCGCCCCCTCCTCCCCCAGCACTCACACTGTGGTAGCG -190
GCGGAGGGCGGAGGGATCCGGAGGCGGCGGAGGGAGACGTCATTGCAGGGTTGTTTGTCA
GTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGATCCACAGTGATTCGGCCGCG - 70
CCGCCGGGGGGTGGGGGGGCTGCGCGGGACTTTTTTTTTTTTTCAGACTGACCGCGGGGCAG
1            M  S  T  P  A  R  R  R  L  M  R  D  F  K  R  L  Q
CTGCGGAGCATGTCGACCCCGGCCCGGAGGAGGCTCATGCGGGATTTCAAGCGGTTACAA  51
     E  D  P  P  V  G  V  S  G  A  P  S  E  N  N  I  M  Q  W  N
GAGGACCCACCTGTGGGTGTCAGTGGCGCACCATCTGAAAACAACATCATGCAGTGGAAT
38   A  V  I  F  G  P  E  G  T  P  F  E  D  G  T  F  K  L  V  I
GCAGTTATATTTGGACCAGAAGGGACACCTTTTGAAGATGGTACTTTTAAACTAGTAATA 171
     E  F  S  E  E  Y  P  N  K  P  P  T  V  R  F  L  S  K  M  F
GAATTTTCTGAAGAATATCCAAATAAACCACCAACTGTTAGGTTTTTATCCAAAATGTTT
78   H  P  N  V  Y  A  D  G  S  I  C  L  D  I  L  Q  N  R  W  S
CATCCAAATGTGTATGCTGATGGTAGCATATGTTTAGATATCCTTCAGAATCGATGGAGT 291
     P  T  Y  D  V  S  S  I  L  T  S  I  Q  S  L  L  D  E  P  N
CCAACATATGATGTATCTTCTATCTTAACATCAATTCAGTCTCTGCTGGATGAACCGAAT
118  P  N  S  P  A  N  S  Q  A  A  Q  L  Y  Q  E  N  K  R  E  Y
CCTAACAGTCCAGCCAATAGCCAGGCAGCACAGCTTTATCAGGAAAACAAACGAGAATAT 411
     E  K  R  V  S  A  I  V  E  Q  S  W  N  D  S  *  *
GAGAAAAGAGTTTCGGCCATTGTTGAACAAAGCTGGAATGATTCATAATAGACAACTGGT
CTGTTAATCTTTTTCATCATTGTTGTGTATAATTTACCTCTCATTAGAAAGGCTAACAAA 531
TTTTAAGTGCCACAGGTTTTAAGGATTCTGCAGAAAAAAAAGAAAAAAGTCCTTCAGTTT
AGAACCTACAAAAGCTTGTGTATCTTGATTAATGTACTTTTTATTGCATGGTGTGAACTA 651
AGTTATTGCTGCATAAATTTGTAATATATCCTGTTTGTATTTTTTTCCAAGTGTATAATG
TTGGTGTGGAGTTTTCATGACAGAATATACACATTTTGTAAATCTGTACTTTTTTCAAAT 771
ATTGAATGCCTTATTTTTGAATTCTTTAGATTTTTAAATTGGAGAAAAGCACTTAAAGTT
TTTTATATATGAATATTACATGTAAAGCTGTTAAAATACATAACTTCAGTGCAAGAGACT 891
TTGTCACTTATTTCCTTATGTGTGTAGGAGGGGTTAATAAGTCTCTAGCTCTCCATCTAT
TGATAGTTTCATTTACAATTTCAAAAGAACATTCTTATATTTTATCAAGGAAGTCTTCAA 1011
ATTTGATTCTAAATAGCGATTATAATCTCCAACTTTATTTTGAATGTACCTCTATTAGTT
TCAATTGAGTAATTCTAGACATAACTGGTTTGACTCTGTCCAACTCTGTATTTAGGCCAT 1131
TTGTTACAGTTTCTTCATGCATTACTTACTGTTAAAACTGTACCTTTTGCGATTTCACAG
TTGGCACTTCTGCCATGAGCAGAGAACTGATGCGACTTGTTTTGCTGCTTGGTAGCACTT 1251
TAAAAAATTTTTTGATTAATGAAGAAAGTAAAACCATAAACATTGCCAAAAATTCATGCC
CCAGTATTAGCAATGAATTAGTTGAATTGGTTTGAGAAAGGCACATATTGGAGGGAAATC 1371
TTGGTGTAACTTAAATATTTGAAAATTACCTTTAATGCAATGCATATCTGTTTATTCTGG
GAAATGTTTTAATGCCAGGGCCTGCTGAGTTGCTTCTTCTTGTGGAGATTTTTTTTTTAA 1491
TCTCCTGAGTTGTATAAAAGTTGTACTGCATCTTAGTTTACTGGATAAATTTAAAACACA
GTATTGTAGAAAGCTAATACAAAACTATCCTATGCCTTCAAATAGTATAGAAAATGGAAA 1611
ATATACAAGTAAATTCTGTTGAACCACCTGTGTAGTCTCTCTAGTAGTTAAAACAGCTAT
TTTAGTAACCCCAAGAGTTTCTTCTCATATCCATAAAATTGGATTACAGTATGGCAATAT 1731
CTACAGCTTCTATTCATTATGCAACGTTTAACATTGATTGGATAAATGCTTATAGGCTGG
GCCAGGTGGCTCCCACCTGTAGTCCCAGCACTTTGGGAGGCAAAGGCGGGAGGATCACTT 1851
GAGGCCAGGACTTTGAGACCAGCCAGGGCAACATAATAAGACTTTTCTCTACTTTAAATT
TTTTAAAAAATTAGTTGGGCATAGTGGCACTTGCCTGTAGTCTCAGCTACTTAGGAGGCT 1971
GAGGTGGCAGAATAGCTTATGCCTAGGAGGTTGAGGCTGCAGTGAGCTGTGATTATGCCA
TTGCACTCCAGCCTGGGTGACAGAGCAATACCCTGTCTTTAAAAATAAAAGCTATTACTT 2091
GGCAAAAAATTCCGCACTATATAATTCGCACACATTAATTAGGGTTTATGTACCATACAA
AAAAAAATTAATTAGGGCT...
```

FIG. 2. Physical map and nucleotide/amino acid sequence of *HHR6A* and *HHR6B*. (*A*) Physical map of the two types of *HHR6A* cDNA and a partial map of the *HHR6B* cDNA. E, *Eco*RI; H, *Hin*dIII; S, *Sac* I; Sm, *Sma* I. Position of the 293-bp *Sma* I/*Hin*dIII and the 392-bp *Sac* I *HHR6A* cDNA probes used in Fig. 1 is indicated. (*B*) Nucleotide sequence of the human *HHR6A* cDNA. Start codon ATG at position +1 (and −28) as well as the stop codon TGA at position +456 (and −16) are doubly underlined. The restriction enzyme sites used for probe preparation (see *A*) are indicated (the *Sac* I# site is artificial). Presumed polyadenylylation signals are doubly underlined. Arrowhead points to the position where the cDNA for the 0.8-kb mRNA terminated. Amino acids are given in the single-letter code. (*C*) Nucleotide sequence of the human *HHR6B* cDNA (not the entire 5′ and 3′ UTR sequence is shown). Start codon and stop codon are doubly underlined. A trinucleotide tandem repeat $(CGG)_{11}$ is indicated by interrupted underlining. Dotted underlining points to a region with very strong secondary structure. The 3′ UTR contains an *Alu* repeat (singly underlined). ATTTA boxes are in boldface. The segment from nucleotides −39 to +875 is identical with the sequence of a partial cDNA clone published by Schneider *et al.* (17) except for the presence of an extra G residue at −2 in our sequence, which changes the −3 position important for translation initiation from a G into a more optimal A. Amino acids are indicated in the single-letter code.

>1000-fold by *HHR6A* and *HHR6B*, respectively (Fig. 5*A*). The *HHR6A* and *HHR6B* genes also restore UV mutagenesis in the *rad6*Δ strain to wild-type levels (Fig. 5*B*). In contrast, the two human homologs confer only a low level of sporulation ability (≈5%) to the *rad6*Δ/*rad6*Δ strain.

DISCUSSION

In this paper, we have identified two closely related homologs of the *S. cerevisiae RAD6* gene in human, one of them being identical to the E2 (M_r 17,000) protein recently described by Schneider *et al.* (17), who isolated an incomplete cDNA on the basis of a partial amino acid sequence. Our extensive analysis of a large number of independent genomic DNA clones points to the existence of only a single *RAD6* gene in *S. cerevisiae*, *Sc. pombe*, and *D. melanogaster*. The very high degree of amino acid sequence conservation throughout eukaryotic evolution points to extremely strong sequence constraints imposed on the RAD6 protein. As shown in Fig. 4, the human and yeast *RAD6* homologs share ≈70% sequence identity and the *Drosophila* homolog is the one most closely related to the human HHR6 proteins (85–87% identity). The Dhr6 and HHR6 proteins share almost the same degree of sequence homology (68–69% identity) to RAD6, whereas the rhp6$^+$ gene product is only somewhat more homologous to the *S. cerevisiae* protein (77% identity). Based on the degree of divergence between the various RAD6 homologs, we calculate that the duplication found in humans (and also in mouse and kangaroo; unpublished results) must have occurred ≈200 × 10^6 years ago, in the Jurassic era.

Fig. 3 (top five lines) shows that among the RAD6 homologs, the N-terminal part and the central region, in particular, have been highly conserved. The middle portion contains the invariant Cys-88 residue that is involved in thiol ester linkage with ubiquitin and that is crucial for all *RAD6* functions, as its substitution by valine or alanine produces a *rad6* null phenotype (24). The C terminus, on the other hand, has diverged much more. The *S. cerevisiae* RAD6 protein is unique in harboring an acidic tail sequence. Mutational analysis has shown the acidic domain to be essential for sporulation in *S. cerevisiae* (4). A possible explanation for the absence of an acidic C-terminal extension in other RAD6

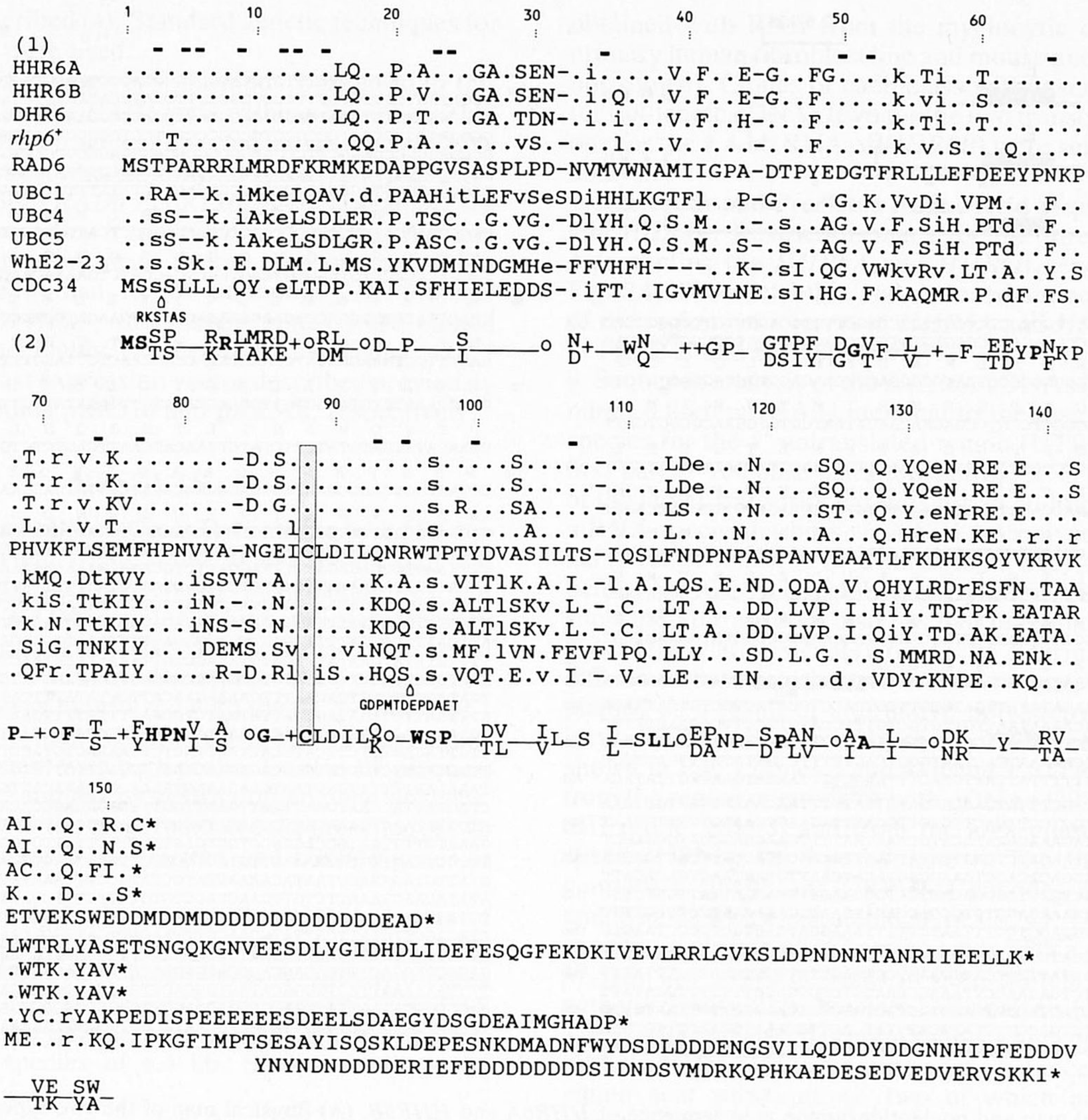


FIG. 3. Comparison of amino acid sequences of HHR6A and HHR6B proteins with various RAD6 homologs and with other ubiquitin-conjugating enzymes. (*Upper*) Comparison of various RAD6 homologs: *S. cerevisiae* RAD6 (3), *Sc. pombe* rhp6$^+$ (10), *D. melanogaster* Dhr6 (11), and human HHR6A and HHR6B (human E2; M_r 17,000) (this paper; ref. 17). (*Lower*) Comparison of the other published ubiquitin-conjugating (E2) proteins; *S. cerevisiae* UBC1 (20); UBC4 and UBC5 (21), involved in protein degradation; *S. cerevisiae* CDC34, involved in cell cycle regulation (22); and wheat E2 (M_r 23,000) (23). Dots indicate identity, whereas lowercase letters indicate strongly conserved residues compared with the yeast RAD6 protein. Conserved amino acids: R and K; E and D; I, V, and L; T and S. (1), Horizontal bars, amino acid residues exclusively conserved in all members of the *RAD6* family; (2), consensus sequence present in all 10 E2 enzymes. Boldface letters, amino acid residues occurring at this position in all 10 ubiquitin-conjugating enzymes; lightface letters, the most likely possibility at this position (occurring in 80% or more of the cases); circles, hydrophilic residue at this position in all the proteins; crosses, hydrophobic residues in all 10 positions. Cys-88 residue, used for ubiquitin attachment, is boxed in all E2 family members.

homologs is that in the other species this domain may have evolved into a protein of its own or it may have become incorporated into a different protein.

The comparison of RAD6 with the other ubiquitin-conjugating enzymes presented in Fig. 3 (bottom six lines) reveals marked similarity, especially in the central part around the Cys-88 residue (see overall consensus sequence 2 in Fig. 3). This segment is likely involved in binding of ubiquitin and/or interaction with the ubiquitin-activating enzyme E1 that donates a ubiquitin moiety from an internal cysteine residue to the cysteine in E2 enzymes. The amino acid sequence around Cys-88 in E2 enzymes bears resemblance to the sequence context of Cys-908 and -866 of the recently cloned ubiquitin-activating enzymes (E1) of wheat and human, respectively (25, 26), and may define a ubiquitin binding domain in E1 enzymes as well.

	RAD6	rhp6$^+$	DHR6	HHR6A	HHR6B	
S. cerevisiae RAD6	-	77	68	68	69	
S. pombe rhp6$^+$	84	-	70	71	71	
D. melanogaster DHR6	74	77	-	87	85	% Identity
Man HHR6A	74	76	90	-	95	
Man HHR6B	74	76	89	96	-	
			% Similarity			

FIG. 4. Identical and similar amino acid residues shared among *RAD6* homologs. Percentage identity is given above the diagonal, and percentage similarity is given below the diagonal. See Fig. 3 legend for classification of conserved residues.

The strict conservation of the N terminus among RAD6 homologs does not extend to the other E2 enzymes. This part may therefore be implicated in important RAD6-specific functions such as interaction with protein components of the DNA repair and mutagenesis machinery. Finally, it is remarkable that all E2 proteins begin with the sequence MS(S/T). Proteins starting with serine are frequently subject to N-terminal acetylation (27). It is not known whether RAD6 or any other E2 enzyme is acetylated at the N terminus.

The high degree of amino acid sequence conservation of RAD6 is also reflected at the functional level. Both human

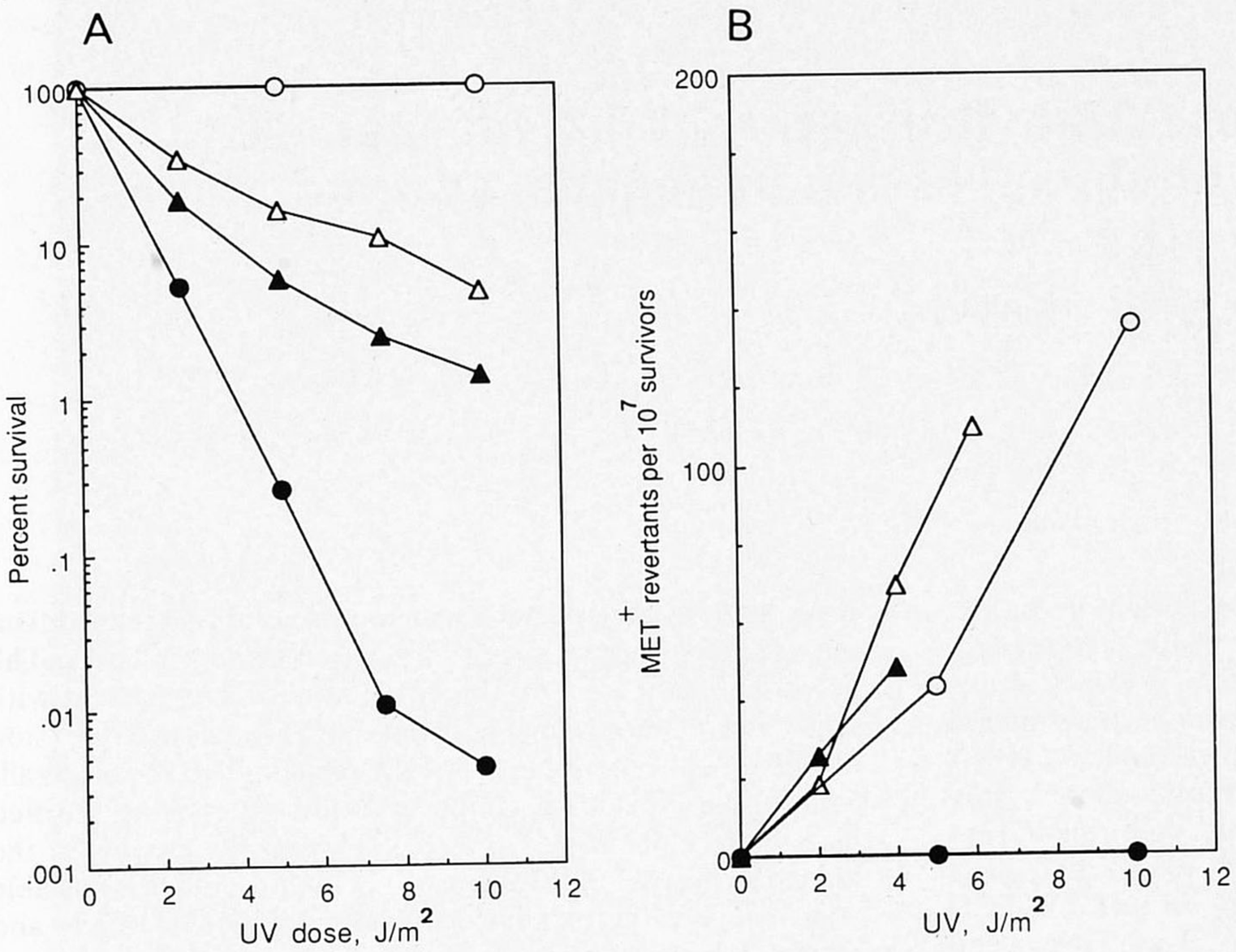


FIG. 5. Complementation of UV sensitivity and UV immutability of the *S. cerevisiae rad6*Δ mutation by human *HHR6A* and *HHR6B* genes. Survival after UV irradiation (*A*) and UV-induced reversion of *met14* (*B*) in the *S. cerevisiae rad6*Δ strain EMY8 carrying the *HHR6A* or *HHR6B* gene on the *ADC1* plasmid. Cells were grown in synthetic complete medium lacking tryptophan for selection of the plasmid and were harvested in midexponential phase. After plating on appropriate medium, cells were irradiated with UV light at a dose rate of 0.1 $J \cdot m^{-2} \cdot sec^{-1}$ and incubated in the dark to avoid photoreactivation. ○, EMY8 + pR67 (*CEN RAD6*); ●, EMY8 + pR611 (*rad6*Δ); ▲, EMY8 + pRR510 (*ADC HHR6A*); △, EMY8 + pRR518 (*ADC HHR6B*).

homologs restore normal levels of UV mutagenesis and effect a substantial increase in UV resistance in *S. cerevisiae rad6* mutants. On the other hand, human homologs confer only a very low level of sporulation ability to *rad6/rad6* mutants. This result is expected in view of the absence of the acidic tail sequence in the human proteins and previous observations that this domain is essential for sporulation but not for DNA repair or UV mutagenesis (4).

The availability of *HHR6* genes should make it possible to examine their role in various cellular processes in mammals such as mutagenesis, postreplication repair, and recombination. Because of the involvement of *RAD6* in sporulation, it will be of special interest to examine whether the *HHR6* genes are implicated in meiosis and gametogenesis. At the final stages of spermatogenesis, histones are replaced by protamines. One can envisage that the capability of RAD6 to polyubiquitinate histones is utilized at this stage to mark histones for degradation by the ATP-dependent ubiquitin-specific protease complex. For these studies, it will be necessary to obtain *HHR6* mutants. One way toward identifying such mutants will be to screen mutant cell lines from human DNA repair disorders or from the existing collection of *in vitro* generated repair-deficient rodent cell lines. Alternatively, *HHR6* mutants could be generated by gene disruption utilizing recently developed methods of gene replacement (28). It is possible to perform this in totipotent mouse embryonic stem cells and in that way to create an *HHR6* defective mouse model. An obvious complication, however, is the presence of two genes, whose function is likely to overlap considerably, necessitating the simultaneous inactivation of both genes.

We thank Mirko Kuit and Tom de Vries Lentsch for photography and Sjozef van Baal for computer assistance. This work was supported by the Dutch Cancer Society (Project IKR 88-2 and 90-20), the European Community Contract B16-141-NL, and U.S. Public Health Service Grants GM19261 and CA41261 from the National Institutes of Health and DE-FG02-88ER60621 from the Department of Energy.

1. Prakash, S., Sung, P. & Prakash, L. (1990) *The Eukaryotic Nucleus*, eds. Straus, P. R. & Wilson, S. H. (Telford Press, Caldwell, NJ), Vol. I, pp. 275–292.
2. Picologlou, S., Brown, N. & Lieberman, S. (1990) *Mol. Cell. Biol.* **10,** 1017–1022.
3. Reynolds, P., Weber, S. & Prakash, L. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 168–172.
4. Morrison, A., Miller, E. J. & Prakash, L. (1988) *Mol. Cell. Biol.* **8,** 1179–1185.
5. Jentsch, S., McGrath, J. P. & Varshavsky, A. (1987) *Nature (London)* **329,** 131–134.
6. Sung, P., Prakash, S. & Prakash, L. (1988) *Genes Dev.* **2,** 1476–1485.
7. Hershko, A. (1988) *J. Biol. Chem.* **263,** 15237–15240.
8. Rechsteiner, M. (1988) *Ubiquitin* (Plenum, New York).
9. Jentsch, S., Seufert, W., Sommer, T. & Reins, H.-A. (1990) *Trends Biochem. Sci.* **15,** 195–198.
10. Reynolds, P., Koken, M. H. M., Hoeijmakers, J. H. J., Prakash, S. & Prakash, L. (1990) *EMBO J.* **9,** 1423–1430.
11. Koken, M. H. M., Reynolds, P., Bootsma, D., Hoeijmakers, J. H. J., Prakash, S. & Prakash, L. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3832–3836.
12. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132,** 6–13.
13. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
14. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **82,** 168–172.
15. Sherman, F., Fink, G. R. & Hicks, J. B. (1986) *Methods in Yeast Genetics: Laboratory Course Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
16. Sung, P., Prakash, L., Weber, S. & Prakash, S. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6045–6049.
17. Schneider, R., Eckerskorn, C., Lottspeich, F. & Schweiger, M. (1990) *EMBO J.* **9,** 1431–1435.
18. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41,** 349–359.
19. Wickens, M. (1990) *Trends Biochem. Sci.* **15,** 277–281.
20. Seufert, W., McGrath, J. P. & Jentsch, S. (1990) *EMBO J.* **9,** 4535–4541.
21. Seufert, W. & Jentsch, S. (1990) *EMBO J.* **9,** 543–550.
22. Goebl, M. G., Yochem, J., Jentsch, S., McGrath, J. P., Varshavsky, A. & Byers, B. (1988) *Science* **241,** 1331–1335.
23. Sullivan, M. L. & Vierstra, R. D. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 9861–9865.
24. Sung, P., Prakash, S. & Prakash, L. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2695–2699.
25. Hatfield, P. M., Callis, J. & Vierstra, R. D. (1990) *J. Biol. Chem.* **265,** 15813–15817.
26. Handley, P. M., Mueckler, M., Siegel, N. R., Ciechanover, A. & Schwartz, A. L. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 258–262.
27. Persson, B., Flinta, C., von Heijne, G. & Jörnvall, H. (1985) *Eur. J. Biochem.* **152,** 523–527.
28. Capecchi, M. R. (1989) *Science* **244,** 1288–1292.