The human \textit{REV1} gene codes for a DNA template-dependent dCMP transferase

Wensheng Lin, Hua Xin, Yanbin Zhang, Xiaohua Wu, Fenghua Yuan and Zhigang Wang*

Graduate Center for Toxicology, University of Kentucky, Lexington, KY 40536, USA

Received July 12, 1999; Revised and Accepted September 29, 1999 DDBJ/EMBL/GenBank accession nos AF151538, AF153594

\section*{ABSTRACT}

DNA is frequently damaged by various physical and chemical agents. DNA damage can lead to mutations during replication. In the yeast \textit{Saccharomyces cerevisiae}, the damage-induced mutagenesis pathway requires the Rev1 protein. We have isolated a human cDNA homologous to the yeast \textit{REV1} gene. The human \textit{REV1} cDNA consists of 4255 bp and codes for a protein of 1251 amino acid residues with a calculated molecular weight of 138 248 Da. The human \textit{REV1} gene is localized between 2q11.1 and 2q11.2. We show that the human \textit{REV1} protein is a dCMP transferase that specifically inserts a dCMP residue opposite a DNA template G. In addition, the human \textit{REV1} transferase is able to efficiently and specifically insert a dCMP opposite a DNA template apurinic/apyrimidinic (AP) site or a uracil residue. These results suggest that the \textit{REV1} transferase may play a critical role during mutagenic translesion DNA synthesis bypassing a template AP site in human cells. Consistent with its role as a fundamental mutagenic protein, the \textit{REV1} gene is ubiquitously expressed in various human tissues.

\section*{INTRODUCTION}

DNA damage can lead to mutations during replication. In the yeast \textit{Saccharomyces cerevisiae}, it appears that the majority of induced mutations are generated through the damage-induced mutagenesis pathway (1,2). The required yeast genes in this pathway include: \textit{RAD6}, \textit{RAD18}, \textit{REV1}, \textit{REV3}, \textit{REV6}, \textit{REV7} and \textit{NGM2} (1–7), most of which have been isolated by gene cloning. As expected, inactivating these mutagenesis genes dramatically decreases the mutation frequency following DNA damage (3,8).

Rad6 is a ubiquitin-conjugating enzyme (9) and forms a complex with Rad18 (10–12). It has been proposed that this complex may play an important role in the initial steps of the damage-induced mutagenesis pathway (10). \textit{Rev3} protein is a DNA polymerase (DNA polymerase \(\zeta\)) capable of translesion DNA synthesis (13). In contrast to the replicative DNA polymerases, deletion of the yeast \textit{REV3} gene does not lead to lethality (1). Hence, this polymerase is specifically required for damage-induced mutagenesis in yeast. Rev1 belongs to the UmuC family of proteins (14). It possesses a deoxyxycytidyl (dCMP) transferase activity in a template-dependent reaction, which can efficiently insert a dCMP opposite a template apurinic/apyrimidinic (AP) site (15). Yeast Rad30, an \textit{Escherichia coli} DinB homolog, is another member of the UmuC family (14,16,17). However, unlike Rev1, Rad30 is not a component of the damage-induced mutagenesis pathway, but appears to be involved in a novel error-free lesion bypass mechanism (16,17). Most recently, Rad30 was shown to be a non-essential DNA polymerase (pol \(\eta\)) capable of error-free translesion DNA synthesis opposite a TT dimer \textit{in vitro} (18). Apparently, the UmuC family of proteins are involved in different mechanisms in the damage tolerance response to unrepaired DNA lesions during replication.

It is only very recently that the damage-induced mutagenesis pathway in humans has been investigated. Two human homologs of the yeast \textit{RAD6} gene have been identified: HHR6A and HHR6B (19,20). Additionally, hREV3 has been isolated as the human homolog of the yeast mutagenic DNA polymerase \(\zeta\) (21,22). Thus, it is most likely that a damage-induced mutagenesis pathway similar to that in yeast is operational in humans. Given the genetic complexity of the yeast mutagenesis pathway, it is certain that more human mutagenesis genes remain to be identified. Since mutations are the building blocks of human cancers, understanding the damage-induced mutagenesis pathway in humans is a key to the understanding of carcinogenesis. Isolating the human mutagenesis genes and elucidating the activities of these gene products are essential steps in these studies.

In this report, we describe the isolation of a full-length cDNA representing the homolog of the yeast Rev1 mutagenesis protein. We determine the chromosomal location of the human \textit{REV1} gene and show its ubiquitous expression in various human tissues. Furthermore, we demonstrate that the human \textit{REV1} protein is a dCMP transferase capable of inserting a dCMP opposite a template AP site.

\section*{MATERIALS AND METHODS}

\textbf{Materials}

Human bone marrow and leukocyte cDNA libraries in \lambda gt11 were purchased from Clontech Laboratories. A human T cell cDNA library in the ZAP Express vector and a human placenta genomic DNA library were purchased from Stratagene.

*To whom correspondence should be addressed. Tel: +1 606 323 5784; Fax: +1 606 323 1059; Email: zwang@pop.uky.edu
**Escherichia coli** uracil-DNA glycosylase was obtained from New England BioLabs. The vector pUC19M1 was constructed from Stratagene. The E.coli cloning region of plasmid pUC19 immediately after the adding by Deepak Rajpal (University of Kentucky, Lexington, KY) by New England BioLabs. The vector pUC19M1 was constructed.

**Isolation of human REV1 cDNA**

Based on a human EST sequence (GenBank accession number AA029147), a 59mer oligonucleotide (probe I) was derived from CL1265-7C (1) by deleting the uracil-DNA glycosylase was obtained from *Escherichia coli* strain XL1-Blue MRF² and XLOLR were purchased from Stratagene. The yeast strain CL1265rev1Δ (MATα rev1Δ arg4-17 leu2-3,112 his3-D1 trp ura3-52) was derived from CL1265-7C (1) by deleting the REV1 gene.

**Construction of a full-length human REV1 cDNA**

The insert cDNA from two *lgt11* clones with overlapping partial human REV1 cDNA sequences was excised from the phage DNA with Salt restriction endonuclease and subcloned into the plasmid vector pUC19M1. The resulting plasmids, pWL269 and pWL270, contain the 5’ and 3’ halves of the human REV1 cDNA, respectively. The full-length human REV1 cDNA was constructed by ligating a 1.7 kb XbaI–SpIh fragment from pWL270 and a 2.7 kb SpIh–SacI fragment from pWL269 into the SacI–XbaI sites of the plasmid vector pUC19. The resulting recombinant plasmid, pWL296, contained the full-length human REV1 cDNA.

**Isolation of human REV1 genomic clones**

A human REV1 clone from the human T cell cDNA library was excised in vivo from the ZAP Express vector in *E.coli* XL1-Blue MRF² cells infected with a M13 helper phage. The resulting packaged phagemid particles were used to infect *E.coli* XLOLR cells to convert the phagemid into a double-stranded plasmid containing the human REV1 cDNA. The recovered plasmid was then digested with EcoRI restriction endonuclease, releasing the 0.8 kb human REV1 cDNA insert (corresponding to the human REV1 cDNA position –178 to +646) from the plasmid vector. After isolating it from an agarose gel, this cDNA fragment was used as the template to prepare 32P-labeled DNA probes by randomly primed DNA synthesis. Approximately 2 million clones from a human placenta genomic DNA library were screened with the human REV1 probes. Two clones were isolated. The DNA inserts of ~20 kb each were subcloned into the NorI site of PCR-Script vector and partially sequenced.

**Northern blot analysis of the human REV1 mRNA**

A 59mer oligonucleotide probe, corresponding to the human REV1 cDNA position –126 to –184, was synthesized and labeled with 32P at its 5’ end by T4 polynucleotide kinase. A human mRNA blot (Invitrogen) was hybridized with the probe in a buffer containing 50% formamide, 0.25 M NaCl, 0.25 M sodium phosphate, pH 7.2, 1 mM EDTA and 5% SDS at 42°C for 18 h. The blot was washed with 15 mM NaCl, 1 mM sodium phosphate, pH 7.4, and 0.1 mM EDTA at 60°C for 1 h. The hybridized human REV1 mRNA was visualized by autoradiography at –80°C with an intensifying screen.

**Detection of the human REV1 expression**

Expression of the REV1 gene in various human tissues was detected by RT–PCR. Poly(A) mRNA samples were isolated from various human tissues and used for first strand cDNA synthesis by reverse transcriptase. These cDNA samples were then normalized against glyceraldehyde-3-phosphate dehydrogenase cDNA. Such human multiple tissue cDNA panels were purchased from Clontech Laboratories and used for PCR. Two PCR primers, CCCAGGAGGAGGATAAGGCTG and GTCTGGTAGTTATGGACAAACTCAGTC, were used to amplify a 360 bp region of the human REV1 cDNA. PCR reactions (20 µl) contained 0.4 ng cDNA, 5 pmol each of the primers, 2 mM MgCl₂, 20 mM Tris–HCl, pH 8.0, 50 mM KCl, 200 µM each of dATP, dCTP, dGTP and dTTP, and 1 U Taq DNA polymerase. After heating at 94°C for 20 s, 35 cycles of amplification were performed according to the following conditions: 20 s denaturation at 94°C, 30 s annealing at 60°C and 45 s extension at 68°C. Reaction products were separated by electrophoresis on a 1% agarose gel containing 0.5 µg/ml ethidium bromide.

**Purification of the human REV1 protein**

The 5’ 2.3 kb open reading frame of the human REV1 cDNA was amplified from pWL296 by PCR, generating a DNA fragment flanked by XhoI (added by the 5’ PCR primer) and HindIII sites. This DNA fragment was cloned into the XhoI–HindIII sites of the yeast expression vector pEGLh6 (24). At its HindIII site, a 2 kb HindIII DNA fragment containing the missing 3’ end of the human REV1 was then transferred from pWL296. The resulting plasmid pEGLh6-hREV1 codes for the full-length human REV1 protein tagged with six histidine residues at its N-terminus. A yeast rev1 deletion mutant strain was transformed with pEGLh6-hREV1 for regulated human REV1 expression under the control of the GAL1/10 promoter. Yeast cells containing pEGLh6-hREV1 were grown in minimum medium containing 2% sucrose to late logarithmic phase. Expression of the human
**RESULTS**

**Cloning of the human REV1 cDNA**

Using the yeast Rev1 protein sequence, we searched the non-redundant GenBank CDS database for its homologs. A *Caenorhabditis elegans* hypothetical protein (ZK675.2) was identified. Alignment of 710 amino acid residues showed 27% identity and 44% similarity to the yeast Rev1. Thus, we conclude that this *C. elegans* protein is a homolog of the yeast Rev1. Using the *C. elegans* REV1 protein sequence, we subsequently searched the GenBank EST (expressed sequence tag) database and identified a human EST clone (GenBank accession number AA029147). Based on this clone, we searched the EST database again and identified two related EST clones (GenBank accession numbers AA393888 and T08134). Combining the three EST sequences, a partial 3′ cDNA sequence of the putative human REV1 gene was obtained.

To isolate the full-length human REV1 cDNA, we synthesized a 59mer oligonucleotide probe and screened three human cDNA libraries. As a result, an additional 5′ sequence of the putative human REV1 gene was obtained. Subsequently, two additional 59mer oligonucleotide probes were synthesized and used to screen the human cDNA libraries. Forty cDNA clones were isolated, the largest of which contained an insert of 4 kb. The 5′ sequence of the putative human REV1 gene was generated after sequencing. Finally, a cDNA clone containing both the 5′ and the 3′ sequences was reconstructed from partial cDNA clones (GenBank accession number AF151538). This cDNA (4255 bp) codes for a protein of 1251 amino acid residues with a calculated molecular weight of 138 248 Da and pI of 8.76. Upon searching the GenBank with our protein sequence, we identified the yeast Rev1 as its homolog (Peu = 5 × 10^-46). Hence, we conclude that our cDNA clone codes for a human homolog of the yeast mutagenesis protein Rev1. Accordingly, this cDNA and its gene is referred to as the human REV1.

The sequence context of the putative ATG start codon (CCACCATGA) in our human REV1 clone matches well with the Kozak consensus sequence (CCACCATGG), which is commonly found surrounding the mammalian ATG initiator codon (26). However, the 5′ untranslated region of this human REV1 cDNA does not contain an in-frame termination codon. Furthermore, two cDNA clones contained an intron-like sequence 5′ upstream of the position −10, in which the sequence context at the junction closely resembles the consensus sequence of the 3′ splicing site. These observations raised the question whether we have isolated the full-length human REV1. To answer this question, we first determined the size of the human REV1 mRNA by a northern blot analysis. As shown in Figure 1, the human REV1 mRNA was estimated to be 4.5 kb. This is in good agreement with the size of our cDNA clone (4.3 kb). Secondly, we screened a human genomic library using the human REV1 cDNA as the probe. Two overlapping genomic clones were isolated. Sequencing these clones confirmed the presence of the 5′ splicing site and revealed multiple termination codons upstream of the cDNA sequence (GenBank accession number AF153594). These results show that we have isolated the full-length cDNA of the human REV1. Additionally, the results indicate that the first exon of the human REV1 gene is non-coding.
Comparison of various REV1 proteins revealed a protein sequence conservation was found among these proteins also identified from the genomic sequencing projects. Again, accession numbers AC002342 and AL035548, respectively.) were Schizosaccharomyces pombe the efficiency of the human aborted human REV1 protein synthesis. Thus, the translational acids. Translation from this first ATG codon would lead to an potentially direct the synthesis of a polypeptide of 12 amino located 32 nt upstream of the initiator codon, which could had cloned the human identities of 21–35% and similarities of 43–59% (Fig. 2). After we conservation of REV1 protein sequences from yeast to humans	Sequence alignment between the yeast and the human REV1	identities of 21–35% and similarities of 43–59% (Fig. 2). Four conserved regions were identified with amino acid sequence identities of 21–35% and similarities of 43–59% (Fig. 2). After we had cloned the human REV1 cDNA, the Arabidopsis thaliana and the Schizosaccharomyces pombe REV1 homologs (GenBank accession numbers AC002342 and AL035548, respectively) were also identified from the genomic sequencing projects. Again, protein sequence conservation was found among these proteins (Fig. 3). Comparison of various REV1 proteins revealed a BRCT (BRCA1 C-terminus) domain at their N-terminal regions and five sequence motifs (Fig. 3).

Chromosomal localization of the human REV1 gene
Using the human REV1 cDNA as the probe, we isolated two human REV1 genomic clones from a library. One clone contained a sequence tagged site (STS), EST164698 (GenBank accession number G25709), upstream from the 5’ end of the human REV1 gene. The distance between this STS and the 5’ sequence of the human REV1 cDNA was estimated to be 20 kb by PCR using either the genomic clone or the total genomic DNA isolated from human placenta (data not shown). The location of this STS was assigned to 512.6 cR from the top of chromosome 2 linkage group by radiation hybrid mapping of marker SGC33758 by the Whitehead Institute/MIT Center for Genome Research. On GeneMap’98, it was further mapped to physical position: 355.80 cR99 (P > 3.00) between reference intervals D2S113–D2S176 (115.3–120.8 cM). These markers are localized between 2q11.1 and 2q11.2 on the cytogenetic ideogram. Therefore, we conclude that the human REV1 gene is located between 2q11.1 and 2q11.2.

Expression of the REV1 gene in human tissues
In yeast, the Rev1-involved mutagenesis pathway is a major mechanism for generating mutations after DNA damage. However, it is not known whether this pathway functions in various human tissues. Thus, we examined the expression of the REV1 gene as an indication of the importance of this putative mutagenesis pathway in various human tissues. As shown in Figure 4, the human REV1 expression was detected by RT–PCR in all of the 16 tissues examined. Hence, we conclude that the REV1 gene is ubiquitously expressed in humans.

The human REV1 protein is a dCMP transferase
The yeast Rev1 protein possesses a deoxycytidyl transferase activity, which transfers a dCMP residue to the 3’ end of a DNA primer opposite a template G or an AP site (15). To determine whether the human REV1 protein is a dCMP transferase, we first partially purified the protein and then assayed for dCMP transferase activity. To facilitate detection and purification of the human REV1, we tagged the protein with six histidine residues at its N-terminus, and expressed it in yeast rev1 deletion mutant cells. The tagged human REV1 was purified by affinity chromatography on a nickel–Sepharose column. As a control, we used rev1 deletion mutant extracts for identical purification. Using a primed 40mer DNA template (Fig. 5A), we assayed the transferase activity of the partially purified human REV1. As shown in Figure 5B (lane 10), a transferase activity was detected that extended the 32P-labeled primer by 2 nt opposite the two template G residues. In contrast, the control sample without the human REV1 did not contain any detectable transferase activity (Fig. 5B, lane 9), indicating that the transferase activity is specific to the human REV1 protein. To identify the nucleotides transferred opposite the template G residues, we performed the transferase assays with dATP, dCTP, dGTP or dTTP individually, rather than all four dNTPs together. Only dCTP supported the transferase activity (Fig. 5B, lane 4). Again, the transferase activity was not detected with the control sample without the human REV1 protein (Fig. 5B, lanes 1, 3, 5 and 7). The transferase activity was not observed opposite a template A, C or T (data not
Hence, we conclude that the human REV1 protein is a dCMP transferase, which transfers dCMP opposite a template G. Supporting this conclusion, the transferase activity co-purified with the human REV1 as revealed by western blots during nickel–Sepharose column chromatography (data not shown). In the control purification from rev1 deletion mutant extracts, none of the fractions contained the transferase activity (data not shown).

To examine whether the transferase activity of the human REV1 functions opposite a template AP site, we prepared a site-specific uracil-containing template (Fig. 6A). Treatment...
with uracil-DNA glycosylase completely converted the uracil-containing templates into AP site-containing templates, as revealed by the AP site cleavage with the *E.coli* endonuclease III (Fig. 6B, lane 2). Transferase activity of the human REV1 was detected opposite the template AP site (Fig. 6C, lane 9). A template U also supported the human REV1 transferase activity (Fig. 6C, lane 10). To identify the nucleotides transferred opposite the template AP site or uracil, we performed the transferase assays with only one deoxynucleoside triphosphate, dATP, dCTP, dGTP or dTTP. As shown in Figure 6C (lanes 3 and 4), only dCTP supported the human REV1 transferase activity opposite the template AP site or uracil. These results show that the human REV1 protein is a template-dependent dCMP transferase that is active opposite a template G, U or AP site.

The yeast REV1 gene had been deleted from the host cells used for the human REV1 expression and purification. Thus, the yeast Rev1 could not have contaminated our human REV1 protein preparations. Nevertheless, to provide further support to our conclusion that the human REV1 is a dCMP transferase, we purified the protein to apparent homogeneity (Fig. 7A and B). Again, a transferase activity opposite an AP site was observed with the pure human REV1 preparation (Fig. 7C). Additionally, we performed the transferase assay opposite a template G using the pure human REV1 protein. The dCMP transferase activity was detected (data not shown). These results show that the observed dCMP transferase activity is associated with the human REV1 protein.

**DISCUSSION**

DNA damage-induced mutagenesis is an important cellular response to unrepaired DNA lesions during replication. The biological outcome of this pathway is enhanced cell survival and increased mutations following DNA damage. The yeast *S.cerevisiae* has served as the most informative model organism in studies of the damage-induced mutagenesis pathway in eukaryotes. Yeast genetic analyses have implicated at least seven genes in this mutagenesis pathway, including *REV1* (27).

We have isolated a full-length cDNA of the yeast *REV1* counterpart in humans. The REV1 protein is conserved from yeast to humans. Some regions share over 30% identity and more than 50% similarity between the yeast and the human proteins. The REV1 protein is additionally found in *S.pombe, C.elegans* and *A.thaliana*, with significant sequence homologies among them. Thus, the function of REV1 protein in the damage-induced mutagenesis pathway may have been preserved in various eukaryotic organisms during evolution. REV1 proteins of various sources all contain an N-terminal BRCT domain. It was originally identified in the breast cancer suppressor protein BRCA1 and subsequently found in some other proteins involved in cell cycle checkpoints, DNA repair and in recombination, such as Rad9, p53-binding protein, XRCC1 and DNA ligases III and IV (28–30). This structural domain is important in protein–protein interactions (31). Thus, it is likely that REV1 may interact with other proteins during damage-induced mutagenesis, although none of the REV1 interactions have been identified. Additionally, REV1 proteins contain several conserved sequence motifs (I–V), which...
Most recently, by employing the yeast two-hybrid system, Wixler et al. (32) identified a partial human cDNA whose polypeptide interacts with the cytoplasmic domain of the ε3A integrin subunit. This cDNA clone (α3 integrin interacting protein 80, AIBP80) corresponds to the 2.6 kb of the 3′ end of our human REV1 cDNA, with a few sequence discrepancies. This sequence was localized by the Sanger Centre between 2q11.1 and 2q11.2, a region identical to our human REV1 chromosomal location. Since integrins are transmembrane receptors that provide a link between extracellular matrix proteins such as laminins and the intracellular cytoskeleton (32), it is not obvious at present as to how such an integrin–REV1 interaction is involved in damage-induced mutagenesis. Thus, the physiological relevance of this protein interaction needs to be further examined.

We found that the human REV1 protein is a dCMP transferase capable of inserting a dCMP opposite a template AP site. This activity provides evidence supporting a role of the REV1 protein in damage-induced mutagenesis in humans. In vitro, the human REV1 dCMP transferase functions efficiently opposite a template AP site. Thus, the human REV1 transferase may play a critical role during mutagenic translesion DNA synthesis opposite a template AP site in vivo. Supporting this notion, Johnson et al. (33) recently demonstrated that AP site-induced mutagenesis in yeast requires the Rev1 protein. Our results also suggest that the damage-induced mutagenesis pathway will incorporate a C residue opposite an AP site during human DNA replication, regardless of the original base identity previously residing at the AP site. Since REV1 is also needed for UV-induced mutagenesis in yeast (27) (D. Rajpal, X. Wu and Z. Wang, unpublished results), additional function of the protein must be involved during mutagenesis opposite other DNA lesions.

Yeast Rev1 is also a dCMP transferase (15). Thus, detection of the dCMP transferase activity of the human REV1 indicates that it is both a structural and a functional homolog of the yeast protein. Additionally, humans contain highly conserved homologs of the yeast mutagenesis proteins Rad6 (19,20) and Rev3 (21,22). Taken together, these observations clearly indicate the existence of a damage-induced mutagenesis pathway in humans in response to DNA lesions. Further supporting this conclusion, Gibbs et al. (21) showed that UV-induced mutagenesis in human cells requires the human REV3. We suspect that this damage-induced mutagenesis pathway is likely a fundamental and major mechanism for generating mutations in humans after DNA damage. Consistent with this hypothesis, we observed ubiquitous expression for both REV1 and REV3 (22) in various human tissues. The isolation of the human REV1 cDNA and identification of its dCMP transferase activity should facilitate molecular and biochemical studies of the damage-induced mutagenesis pathway in humans.

ACKNOWLEDGEMENTS

We thank Richard Cunningham and Christopher Lawrence for providing us with purified E.coli UmuC-related proteins (14). Identifying these conserved motifs should be useful for future structure–function studies of REV1.

Examination of the 5′ untranslated region of the human REV1 cDNA revealed the presence of an out-of-frame ATG at nucleotide position −35 which initiates an ORF of 12 codons and terminates at position +2. The stop codon of this mini ORF overlaps with the human REV1 initiator ATG codon. The sequence context of this upstream ATG is close to the consensus Kozak sequence (26). Thus, it is likely that the translational efficiency of the human REV1 message may be reduced by the presence of this upstream mini ORF. Structural features of the human REV3 gene also suggest a low-level expression (21,22). These features imply that under normal growth conditions human cells may contain limited amounts of the mutagenesis proteins.
Kentucky and a New Investigator Award in Toxicology from the Burroughs Wellcome Fund.

REFERENCES