

## Isolation and characterization of *HCI*: a novel human DNA repair gene

D.O. Lopes<sup>1</sup>, F.C. Falconi<sup>1</sup>, A.M. Goes<sup>1</sup>, Y. Canitrot<sup>2</sup>, J.S. Hoffmann<sup>2</sup>,  
C. Cazaux<sup>2</sup>, G.R. Franco<sup>1</sup>, A.M. Macedo<sup>1</sup>, S.D.J. Pena<sup>1</sup> and C.R. Machado<sup>1</sup>

<sup>1</sup>Departamento de Bioquímica e Imunologia,  
Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais,  
Belo Horizonte, MG, Brasil

<sup>2</sup>Genetic Instability and Cancer,  
Institut de Pharmacologie et de Biologie Structurale, CNRS UMR 5089,  
Toulouse, France

Corresponding author: C.R. Machado  
E-mail: crmachad@icb.ufmg.br

Genet. Mol. Res. 8 (1): 247-260 (2009)  
Received December 11, 2008  
Accepted January 18, 2009  
Published March 3, 2009

**ABSTRACT.** Nucleotide excision repair (NER) acts on a broad spectrum of large lesions, while base excision repair removes individual modified bases. Although both processes have been well studied in human cells, novel genes involved in these DNA repair pathways have been described. Using a heterologous complementation approach, we identified a fetal human cDNA that complemented two *Escherichia coli* mutants that are defective in 3-methyl adenine glycosylase and in three endonucleases, all of which are enzymes with important roles in base excision repair. The central cDNA open reading frame complemented NER mutant strains and promoted an increase in survival rate of bacteria exposed to UV light. The corresponding protein was able to restore nucleotide-excision-repair activity when added to a cell extract from Chinese hamster ovary cells deficient in the ERCC1 protein, an enzyme known to promote incision at the 5' end of the lesion during NER. In contrast, that protein was not able to complement XPG Chinese hamster ovary cells deficient in the 3' incision step of NER. These data indicate a new human repair gene, which we named *HCI*; it is involved

in the recognition of two kinds of DNA lesions and it contributes to the 5' DNA incision step in NER.

**Key words:** DNA repair; Base excision repair; Nucleotide excision; Heterologous complementation; Genome maintenance; DNA damage

## INTRODUCTION

Evolution has selected several systems to protect the genetic material from damage by environmental agents or by products of metabolism, thus preventing the accumulation of mutations or cell death. These systems, which are DNA repair and lesion tolerance mechanisms, exist in all living organisms to maintain genetic stability. The vigorous efforts to isolate and characterize DNA repair genes have helped to understand some aspects of the cell mechanisms for DNA damage processing and their involvement with human syndromes and cancer development (Leibeling et al., 2006). The complexity of DNA repair mechanisms is far from being elucidated, and certainly a larger number of genes are involved in nuclear genome stability. The isolation and characterization of several related genes contribute not only to our knowledge of the cell and gene metabolism, but also help in the prevention of cancer, and this study has revealed an important degree of homology and analogy even in phylogenetically unrelated species (Friedberg, 2003; Kraemer et al., 2007).

The various pathways for repair include nucleotide excision repair (NER) mechanism, which repairs a broad spectrum of lesions, and a more specialized base excision repair (BER) mechanism, which is a predominant pathway for coping with a broad range of small base lesions (Baute and Depicker, 2008).

NER is a multistep process capable of removing a variety of DNA distorting lesions from prokaryotic and eukaryotic genomes. This is a complex but highly conserved process in which basically the following steps can be distinguished: i) recognition of a DNA lesion; ii) single strand incision at both sides of the lesion; iii) excision of the lesion-containing single-stranded DNA fragment; iv) DNA repair synthesis to replace the excised nucleotides; v) ligation of the remaining single strand nicks (Van Hoffen et al., 2003). In prokaryotic cells, the *uvrA* protein, in association with *uvrB*, is responsible for recognizing the lesion. The protein coded by *uvrC* is the endonuclease that produces the incision at 5' and 3' of the lesion; however, for the 3' incision, this protein needs to be in a complex with *uvrB*. After the *uvrD* helicase protein removes the lesion, DNA polymerase III synthesizes the new undamaged DNA (Verhoeven et al., 2000). In eukaryotic cells, the NER process requires more than 30 proteins. Recognition of the DNA lesion is carried out by XPC-HR23B, RPA-XPA or DDB1-DDB2. DNA unwinding is performed by the transcription factor TFIIH, and excision of the lesion, by XPG and XPF-ERCC1. Finally, resynthesis occurs with the help of Pol $\delta$  or Pol $\epsilon$ , and ligation, via DNA ligase I (Costa et al., 2003).

In BER, the first step is the removal of the damaged base, carried out by a group of enzymes termed glycosylases, each with its own unique specificity. The resultant abasic (AP) site is a substrate for AP endonuclease (APE1), which hydrolyzes the phosphodiester bond 5' of the AP site, generating a 5'-deoxyribose phosphate. A DNA polymerase then inserts a single nucleotide into the repair gap and the 5'-deoxyribose phosphate residue is removed prior to completion of repair by DNA ligase (Hitomi et al., 2007).

Massive DNA replication and rapid cell proliferation occur during embryonic development. A proficient DNA repair is thus important for normal organogenesis. Although most of the known DNA repair genes are expressed during organogenesis (Vinson and Hales, 2002), no specific embryonic DNA repair gene has been reported so far. In this study, we isolated a human DNA repair-related gene from a human fetal liver and spleen cDNA library by using complementation screening in bacteria defective in enzymes that participate in the initial steps of base excision repair. This gene, named *HCI* for human Heterologous Complementation 1, was able to complement bacterial strains deficient in different steps of BER and NER pathways. *In vitro* experiments confirmed its involvement in the NER pathway and demonstrated the proficiency of the HCI protein in promoting the incision step upstream of the lesion. Moreover, this protein is produced in the amniotic membrane cells, as shown by Western blot experiments, and probably has an important role in fetal DNA repair.

## MATERIAL AND METHODS

### Bacterial strains and media

The strains employed in this study were DNA repair mutant derivatives of the *Escherichia coli* K12 strain AB1157 (F- *thi1 his 4 Δ(gpt-proA) argE3 thr1 leu B6 ara 14 lac Y1 gal K2 xyl5 mtl1 tsc33 supE44 rpsL31*), which is the wild-type for DNA repair functions. The isogenic mutant strains were GC4803 (*alk::tag*), BW535 (*nfo::Kan nth :: Δ(xth-pncA)90*), AYM84 (*uvrC*), AB1886 (*uvrA*), AB1885 (*uvrB*), and AB2474 (*uvrA, lexA<sup>r</sup>*). The cells were routinely grown in 2X yeast tryptone (YT) in liquid medium or plates (containing 1.5% agar). Ampicillin (50 µg/mL), IPTG (Sigma) and X-gal (Sigma) were added to the medium for plasmid selection (Bachmann, 1972).

### Screening clones for methyl methane sulfate resistance

A normalized human fetal liver and spleen cDNA library kindly provided by Dr. Marcelo Bento Soares (University of Iowa) was used. The cDNA library was transformed into the *E. coli* GC4803 by the method of Hanahan (1983). Only one transformation was performed using 200 ng of the library, yielding a total of  $9 \times 10^5$  colonies. These were plated directly onto Ap-medium plates containing ampicillin (50 µg/mL) and 0.5 mM methyl methane sulfate (MMS) and incubated for 16 h at 37°C. The survival clones were re-tested in plates containing 0.5 and 5 mM MMS.

### Cloning procedures and DNA sequencing

The proofreading *Pfu* DNA polymerase (Stratagene) was used in polymerase chain reaction (PCR) amplifications for subcloning. Amplification products used in subcloning or transformation procedures were previously purified using the Wizard PCR-Prep kit (Promega). Cloning into pUC18 (Stratagene) vector was done using the SureClone Ligation Kit (Amersham Biosciences) or using the restriction enzymes *EcoRI* and *HindIII* sites (Pharmacia). Plasmid purifications were performed by alkaline lysis using the Wizard plus SV Miniprep kit (Promega). DNA sequencing reactions were prepared using Thermo

Sequenase Sequencing Kit (Amersham Biosciences) with fluorescent M13 primers and carried out in an A.L.F. DNA sequencer (Amersham Biosciences). The original insert was entirely sequenced on both strands and the GenBank (Altschul et al., 1990) database was searched for homologous genes. *HCI* cDNA was amplified using the following primers: E1 (5'-CGA GAA TTC AGA GCA AAT GCA TAG CAT TG-3') - P1 (5'-GAC GTC TGC TAG GTG CTC ATC GAA T-3') and E2 (5'-CGA GAA TTC AGA AAT GGA GAA ACA TGA GA-3') - P2 (5'-GAC GTC AAT ACT ATT GAC TGA GTG G-3') and cloned into pUC18. Thereafter, the fragments containing the first open reading frame (ORF) (E1P1) and the second one (E2P2) were recovered from pUC18 by digestion with *EcoRI/PstI* and cloned into the same restriction sites of the pMal-C2 (New England BioLabs) and pPRoEX-a, pPRoEX-b and pPRoEX-c (Gibco-BRL) for complementation studies and protein purification.

### Survival measurement after MMS and UV treatment

Mid-phase cells (OD = 0.6 at 580 nm) were pelleted and resuspended in the original volume of 10 mM MgSO<sub>4</sub>. Cell suspensions were treated with different concentrations of MMS for about 10 min and spread on 2X YT-Ap plates. After incubation for 16 h at 37°C, the colonies were counted to measure cell survival.

### UV irradiation

The bacterial cells were grown to OD = 0.6 at 580 nm, pelleted and resuspended in the original volume of 10 mM MgSO<sub>4</sub>. Cell suspensions were irradiated with UV light (254 nm), spread on 2X YT-Ap plates and incubated overnight at 37°C.

### Protein purification

The expression plasmid carrying the insert and the control plasmid (pMAL-c2) were used to transform *E. coli* strain BL21(DE3) competent cells. Transformed cells were grown at 30°C in LB medium, supplemented with ampicillin (50 µg/mL) to a cell density of A<sub>580nm</sub> = 0.4-0.6. The expression of recombinant protein was induced by 0.8 mM IPTG, and incubation was continued for 3-4 h at 30°C. The cells were harvested and the pellet was resuspended in 5-10 mL column buffer (20 mM Tris-HCl, pH 7.4, 200 mM NaCl and 1 mM EDTA). After sonication, the lysate was clarified by centrifugation at 14,000 rpm for 30 min at 4°C and then applied to an amylose column (5 mL). The bound proteins were eluted with the same buffer having now 10 mM maltose. The eluted fraction was quantified according to Bradford (1976).

### *In vitro* experiments

Chinese hamster ovary cell extracts were prepared according to a previously described protocol (Mandavilli et al., 2002). The 2959-bp pBS plasmid (pBluescript KS+; Stratagene, USA) and the related 3738-bp pHM14 plasmid (gift from R.D. Wood, ICRF, UK) were prepared by the alkaline lysis method from *E. coli* JM109. Both plasmids were carefully purified by one cesium chloride and two neutral sucrose gradient centrifugations

as described elsewhere (Biggerstaff et al., 1991). pBluescript KS+ plasmid was treated with UV light (300 J/m<sup>2</sup>) at 254 nm according to Hansson and Wood (1989). The repair synthesis assay was performed as described previously (Wood et al., 1988). The reaction mixture (50 µL) contained 300 ng each of damaged pBluescript KS+ and untreated pHM14 closed circular plasmids, 2 µCi α[<sup>32</sup>P]-dCTP (800 Ci/mmol), 200 µg cell extract protein, and 70 mM potassium glutamate in reaction buffer containing 45 mM HEPES-KOH, pH 7.8, 7.4 mM MgCl<sub>2</sub>, 0.9 mM DTT, 0.4 mM EDTA, 2 mM ATP, 20 µM each of dGTP, dATP and dTTP, 4 µM dCTP, 40 mM phosphocreatine, 2.5 µg phosphocreatine kinase, 3.4% glycerol, and 18 µg bovine serum albumin. Human HC1 protein purified in fusion with maltose binding protein (MBP) was then added (0.5, 1 and 2 µg). Incubation was carried out for 3 h at 30°C. Before electrophoresis on a 1% agarose gel containing 0.5 µg/mL ethidium bromide, plasmid DNA was purified from reaction mixtures. Gels were dried, processed on a PhosphorImager (Storm System, Molecular Dynamics, USA) and radioactivity was quantified (Image Quant 1.1 software). Photographs of the ethidium bromide-stained gels were scanned and processed to normalize for plasmid DNA recovery.

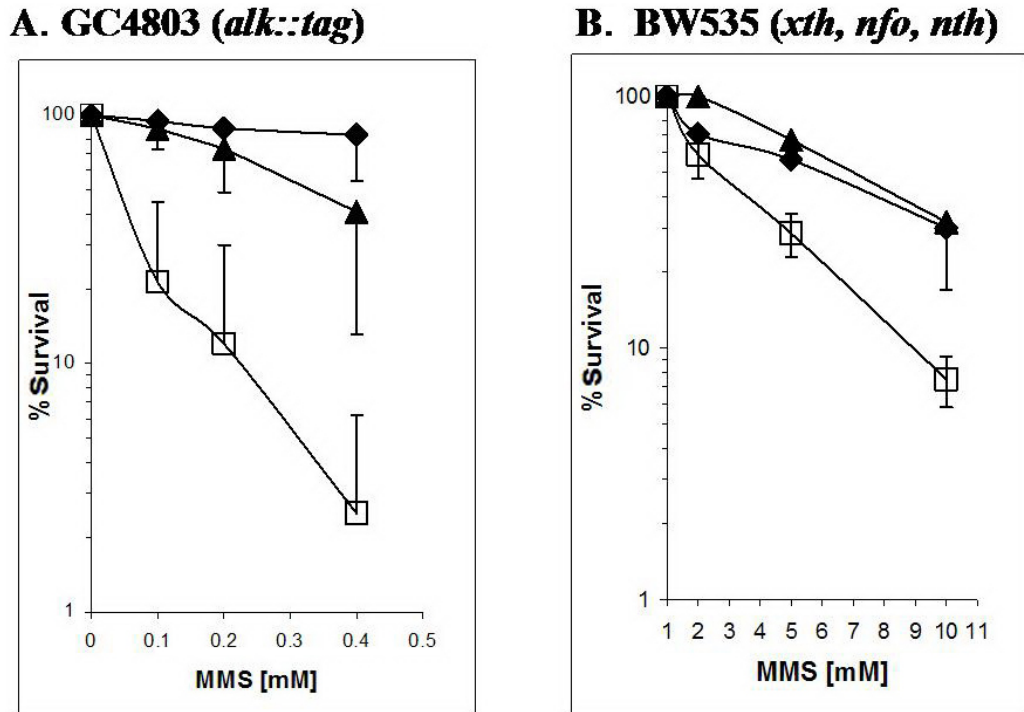
### Western blot

Equal amounts of embryonic cell extract (amniotic membrane), called Wish, were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to a nitrocellulose membrane. Polyclonal antibodies against HC1 protein were raised in rabbits using the purified protein (Makarova et al., 2003). Western blotting was carried out by the method of Towbin et al. (1992). For immunodetection, the primary antibody was tested in different dilutions and used under saturating conditions. Anti-rabbit IgG conjugated with alkaline phosphatase (Promega) was used as a secondary antibody with tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate as substrate of alkaline phosphatase.

## RESULTS

### Screening the fetal cDNA library for MMS resistance

We employed a fetal human cDNA library prepared as described by Bonaldo et al. (1996) inserted in the pT7T3 plasmid. The cDNA library was transformed into the *E. coli* strain GC4803, a double mutant defective for the *alk::tag* genes, which encode 3-methyl adenine glycosylase (Boiteux et al., 1984). Among 90,000 clones tested, only ten survived after the first treatment with the genotoxic agent MMS. These were again individually tested and in only one was MMS resistance confirmed. The cDNA was reintroduced into the GC4803 mutant to avoid a possible revertant phenotype due to bacterial modifications. First, the transformed bacteria were plated on medium without MMS and then tested for survival on MMS. The degree of complementation of the *E. coli* MMS-sensitive phenotype by the fetal gene was tested in several independent experiments. The plasmid abolished the MMS-sensitive phenotype of the GC4803 mutant (Figure 1A) at a level close to the wild type. The BW535 strain (Figure 1B), a triple mutant for three genes, *xth*, *nfo*, *nth*, which encode exonuclease III, endonuclease IV and endonuclease III, respectively (Cunningham et al., 1986), was also complemented by this clone. Both mutants are defective in important steps of BER.



**Figure 1.** Complementation of *Escherichia coli* strains exposed to methyl methane sulfate (MMS). **A.** Exponential-phase cells grown in media containing different concentrations of MMS for about 10 min. Survival of wild-type AB1157 (lozenges) and mutant strain GC4803 (*alk::tag*) transformed by vectors carrying target cDNA (triangles) or empty (squares). **B.** Survival of AB1157 (lozenges) and mutant strain BW535 (*xth, nfo, nth*) transformed with vectors carrying target cDNA (triangles) or empty (squares). The data are the average of 8 independent experiments.

### DNA sequence analysis

The complementing clone was isolated and submitted to automated sequencing; both strands were sequenced and all possible ambiguities were eliminated. The cDNA that had 1267 bp did not show significant homology to any known gene. This sequence, named HC1 (GenBank accession number: AY831680), is located on human chromosome 3 (region 3q13-11), a region that contains only a unigene (Hs.204621) with three expressed sequence tags (T7261.1, AI743005.1, BG186548.1). Analysis of the sequence revealed three possible ORFs, and we could not predict a protein continuously coded by this gene. As shown in Figure 2, the three possible ORFs are located at the beginning with 27 amino acids, in the middle with 60 amino acids, and at the end with 35 amino acids.

### ORF analyses by bacteria complementation studies

cDNA was subcloned to pUC18 and the region responsible for the BER activity was

```

1  GGCACGAGAAAAGAGCAAATGCATAGCATTGGCACTGGTGGTGTGAAAGCACAGCCAGGATGC
      M H S I G T G G V K A Q P R M
65  CTCTTCAGTCAAGTCACTCACTAAAGCTATTCGATGAGCACCTAGCATGGGCTAGGTACCTAG
      P L Q S Q S L T K A I R *
130 ACCAACACACAGTTTACCAACATCCCTGGCTAAGGTCACGTGATCCACAGAGATGCTTGTGTTG
194 CTGAGCATGAGTCAGATCTGGAGAAAAGTCCCTGAGAAAAGTCCCAAGGCTTCTATGTCTGAAAGTTC
258 CAAGCCAGGATGGAAGCAGCCAATTCACCTGCTATAGACTAAGTGTGTTGTCTTCTCCCTCCCTC
322 AATTTATATGCTGAAACCCTAACCCTCATGGTACTTGGAGGTAGGGGCTGGGTAGATTAAAGTC
386 ATGAGGCTGAGGGTGGGGCCCTCATGATGGGCTGAATGCCCTCATTAAAGAAATGGAGAAACAT
      M E K H
450 GAGATCTCTCTCCACCATGCGAGGATACAGCAAGAAGACGGCTATCTGCCAACCAGGAAGAC
      E I S L S T M R G Y S K K T A I C Q P G R
514 AGCCCTCACTAGACACAGGATCTGCCAGCACTTTTCATCTTGGGCTTCCAGCCTCCACAACGTG
      Q P S L D T G S A S T F I L G F P A S T T V
578 AAGAATAAATGCTGTGGCTTAAGCCACTCAGTCAATAGTATTTTGATATAGCAGCTTGGGCT
      R N K C L W L K P L S Q *
642 GAGATATCACCTGTATCGATTCCTTGGCAATGGATTGAGCAAAATGGTGGCTTTGTAGAATGATC
706 AGTTAACTGCAGGGATCAGTCGTGAGCAGGAGCTAAAAAGTCCCATGCTCGGGGAGCCTCACC
770 GAGTTCAGGAAAATGTAAAGAAGAATAACAACCTTAAGCTTCTACACGTCCTAACCTGCA
834 CCAAATGCTTTAGGTACGGTGTATTTCATTTAAATCCCAGAACATGCAACAGCAGGTTTCATCA
898 GATATATTTTATCCAGGGGTATATAAGGCTCAGAGAATAAAAAGGATATCCAAGGTACCAGG
962 TTGCAACCAATACAGGCAGGTTTCAACCAATTCCTCTGACTCAAAGCCCTGGCTCTGTTCTCC
1026 ACGGCTGTGCCAAGGAAGACTGTGCCGAAGATACTTTTGGTCTCAGATTGTGAAATTTCCCT
1090 TGACTCTGGGAAACAGTTTAAAGCAGCCGCAAAACATGGTTTCCACAACCTGCATTCTGAGCG
      M V S T T A F L K R
1154 CATACCCAGAGTTACTTCTGTAAACAGCACTTTCTCCAGATGACTGAGATGCGTCTGCTGSGTG
      I P R V T S C N S T F F Q M T E M R L L V
1218 ATAAGATCAGATTAACCTATTTCAACTTGATAGCTAGATTCTGACACACTG
      I R S D *

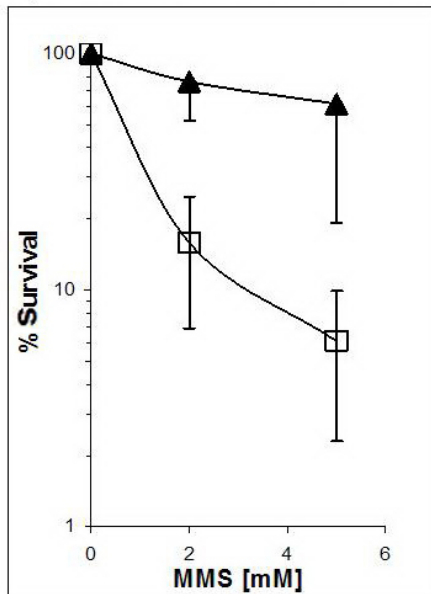
```

**Figure 2.** Analysis of *HCI* nucleotide sequence. The cDNA has 1267 bp and three possible open reading frames with 27, 61 and 35 amino acids, respectively. GenBank accession number: AY831680.

investigated by conducting MMS survival experiments with transformed bacteria. The first subclone, an *EcoRI-HindIII* fragment missing the ORF3 sequence, was still able to complement the MMS sensitivity of the BW535 strain (Figure 3A). The first and second ORFs were PCR amplified and cloned into pUC18. They were tested for MMS sensitivity also in the BW535 strain. As shown in Figure 3B, only the second ORF (E2P2) complemented this deficient repair strain after treatment with MMS. This ORF codes for a putative protein of 8.6 kDa with an isoelectric point of 9.8. This second ORF allowed the survival of an *E. coli* strain deficient in the *uvrA* gene (AB1886) after UV irradiation (Figure 4A). The *HCI* protein also acts in a bacterial strain (AYM84) deficient in the gene that codes for *uvrC* protein, an endonuclease protein responsible for making the 5' DNA incision in the NER pathway (Figure 4B). However, the *HCI* cDNA did not complement the bacteria AB1885, which are deficient in the *uvrB* gene (Figure 4C). Together with *uvrB*, *uvrC* is involved in the 3' DNA incision of the NER pathway (Verhoeven et al., 2000). The *uvrA* and *uvrC* proteins are important for lesion recognition and incision steps in NER, respectively. Therefore, this newly identified human gene could complement bacteria deficient in two important DNA repair pathways: BER and NER. To evaluate whether the *HCI* cDNA codes for a functional protein or a functional

RNA could be transcribed from that sequence, we cloned the second ORF into the plasmids pPRoEX-a, pPRoEX-b and pPRoEX-c, which allows expression of the second ORF in the three possible phases. Only the fragment cloned into the first phase, the pPRoEX-a vector, which codes for HC1 protein, showed a complementation phenotype after UV irradiation in *uvrA* deficient bacteria (Figure 4A). The activity of HC1 protein in lesion recognition and 5' DNA incision in the bacteria background is independent of SOS activation. Instead, the human cDNA complements the UV-sensitive phenotype of the AB2474 bacterial strain (Figure 4D). This strain is an NER (*uvrA*) deficient mutant that is unable to induce the SOS response because it has a *lexA* protein resistant to cleavage.

### A. BW535 (*xth*, *nfo*, *nth*)



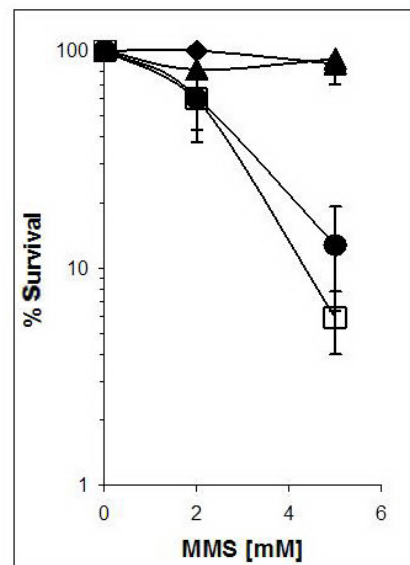
#### HC1



#### HC1 - EH



### B. BW535 (*xth*, *nfo*, *nth*)



#### HC1



#### E1P1



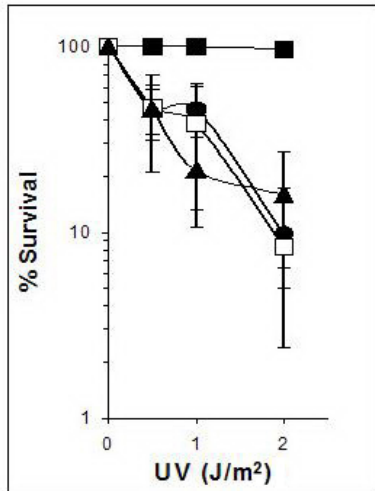
#### E2P2



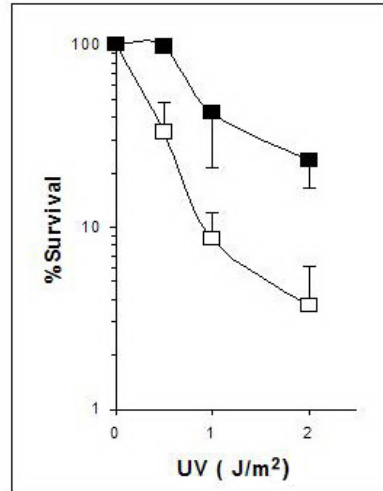
**Figure 3.** Selection of an open reading frame (ORF) responsible for complementation of mutant strains. **A.** Survival of mutant strain BW535 transformed by G1.4-EH clone (triangles) or empty vector (squares). Cells grown in media containing different concentrations of methyl methane sulfate (MMS) for 10 min. **B.** BW535 transformed with G1.4 (triangles), ORF1 (circles), ORF2 (lozenges), or empty vector (squares). The data are the average of 4 independent experiments.



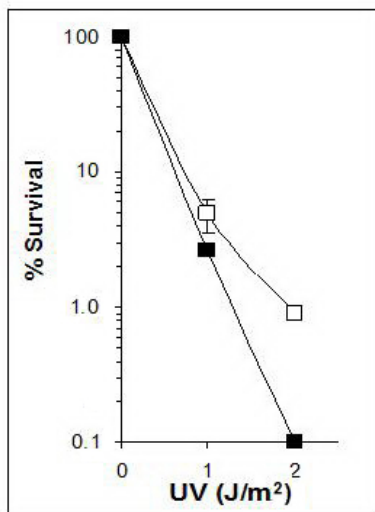
**A. AB1186 (*uvrA*)**



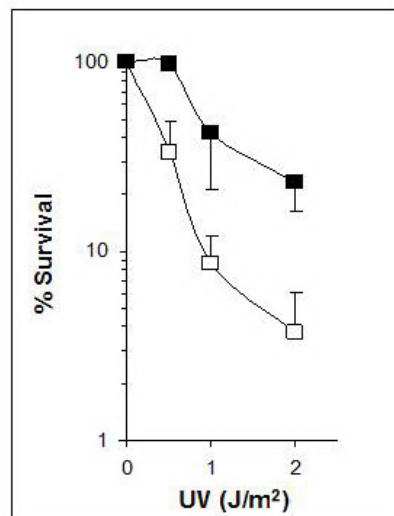
**B. AYM84 (*uvrC*)**



**C. AB1885 (*uvrB*)**



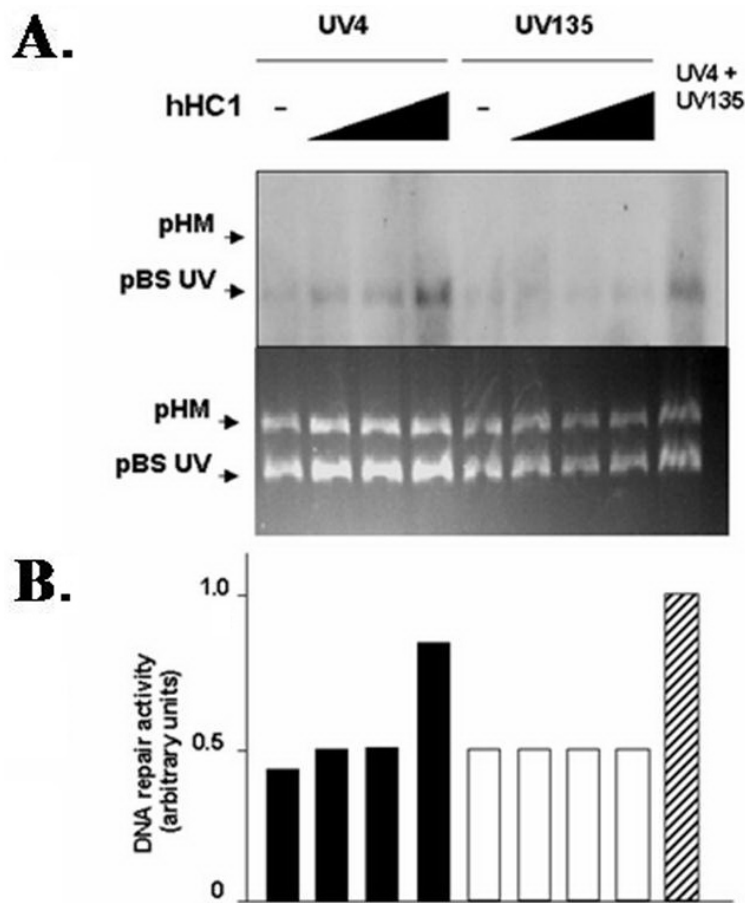
**D. AB2474 (*uvrA*, *lexA*<sup>r</sup>)**



**Figure 4.** Survival of nucleotide excision repair mutants exposed to UV radiation. **A.** AB1186 (*uvrA*) mutant strain was transformed with vectors carrying ORF2 in three possible phases: phase-1 (squares), phase-2 (circles), phase 3 (triangles), or empty vector (squares). Cells were exposed to different doses of irradiation. **B.** Mutants AYM84 (*uvrC*), **C.** AB1885 (*uvrB*) and **D.** AB2474 (*uvrA*, *lexA*<sup>r</sup>) were transformed with a vector carrying ORF2 phase-1 (*HCI*) (filled squares) or empty (open squares) and irradiated. The data are the average of 4 independent experiments.

### Functional complementation of HC1 protein in NER defective Chinese hamster ovary cells

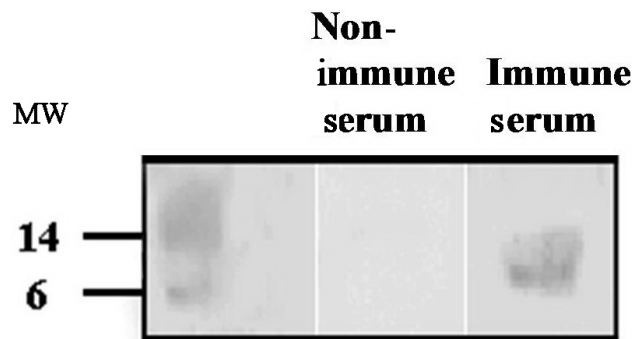
The HC1 protein was obtained in fusion with the MBP. The hybrid protein was purified in an amylose column and it was used in *in vitro* experiments. The HC1-MBP was able to complement the NER deficiency in an *in vitro* NER assay using extracts from hamster UV4 cells (Figure 5), known to be altered in the 5' incision during NER pathway (mutation in ERCC1) (Costa et al., 2003). No complementation by HC1-MBP was shown with extracts lacking the XPG protein (UV135 cell) (Figure 5).



**Figure 5.** DNA repair synthesis on UV-damaged plasmid. Chinese hamster ovary cell extracts (UV4, UV135 and mix) were incubated for 3 h at 30°C with damaged (pBS UV) and undamaged (pHM) plasmid in a repair synthesis reaction. Two  $\mu\text{Ci}$   $\alpha\text{[}^{32}\text{P]}\text{-dCTP}$  (800 Ci/mmol) was added to the normal pool of nucleotides. **A.** Photograph of ethidium bromide-stained gel (lower) and autoradiogram of the dried gel (upper). **B.** Specific incorporation expressed as the ratio of radiolabel incorporation into the damaged and undamaged plasmid.

### Expression of HC1 protein in cells

A protein of approximately 8 kDa (the size expected for HC1 protein) was detected in embryonic cells with antibodies against HC1 (Figure 6). Interestingly, this protein was not observed in normal adult cells such as fibroblasts (data not shown). In addition to these Western blot experiments, an SAGE tag (GTTTCCACAA) is colocalized with the *HC1* gene on chromosome 3. Even though this tag is not highly expressed, most libraries where it is shown correspond to libraries of tumor cells or cells with high proliferation rate (Table 1).



**Figure 6.** Western blot analyses. Wish cell extract (40 µg) was separated on an SDS-PAGE gel and transferred to a nitrocellulose membrane. It was incubated with non-immune and immune serum produced in rabbits. MW = molecular weight.

**Table 1.** SAGE library data obtained with the sequence of *HC1* gene.

Library name	Tags per million
TNF AlphaCultKC	75
Cultured human epidermal keratinocytes, stimulated with TNF-alpha	
SAGE Pancreas carcinoma	32
CL ASPC carcinoma pancreas epithelium ductal carcinoma cell line, non-normalized short SAGE	
SAGE OV1063 3	25
Ovary, carcinoma cell line OV1063	
SAGE Hiroshima GC W226T	22
Primary gastric cancer tissue stomach, early gastric cancer, non-normalized	
SAGE NC2	20
Normal colonic epithelium	
SAGE OVT 7	18
Primary tumor, ovary serous adenocarcinoma	
SAGE SW837	16
Colon, cancer cell line	
SAGE SciencePark MCF7 estradiol 10h Breast Carcinoma -	16
Pleural effusion mammary gland carcinoma cell line estrogen EST	
SAGE 95 348	16
Metastasis from breast tumor	
HPR2 SAGE	9
Human peripheral retina	
SAGE Brain astrocytoma grade II B H359	9
Astrocytoma grade II brain	

## DISCUSSION

We have isolated a novel human gene, named *HCI*, which is able to function in BER and NER DNA repair pathways. This gene was selected by its capacity to revert the MMS-sensitive phenotype of the *E. coli* strain GC4803, which lacks the 3-methyl adenine glycosylase (*alk::tag*), an enzyme that plays an important role in BER (Tudek et al., 1998). This fetal human cDNA also confers MMS-resistance to *E. coli* strain BW535, which is defective in three endonucleases (exonuclease III-*xth*, endonuclease IV-*nfo* and endonuclease III-*nth*) that also play important roles in BER (Demple et al., 1986). Interestingly, this human cDNA also complements the UV-sensitive phenotype of two NER-deficient bacterial strains affected in the recognition step of the lesion (*uvrA*) and in the 5' DNA incision by *uvrC* (Sancar, 1996). In bacteria, *HCI* seems to identify different types of lesions and cut the DNA at 5' from these lesions, since it complements the deficiency of proteins recognizing three different types of lesions (3-methyl adenine - *alk::tag*; AP-endonuclease site - *xth*, *nfo*, and UV lesions - *uvrA*) and also two endonucleases that cut the DNA 5' of the lesion (*xth*, *uvrC*). On the other hand, *HCI* was unable to complement the activity of the *uvrB* gene, an NER gene involved in 3' DNA incision, through complementation of the AB1885 strain.

The expression of some eukaryotic genes involved in DNA metabolism has been shown to induce the SOS system in bacteria (Perkins et al., 1999), leading to cell survival after the exposure to genotoxic agents. As the *HCI* gene is still able to complement NER deficient bacteria, which is unable to induce the SOS system, the DNA repair activity of this human gene may be independent of the induction of the bacterial SOS system.

Sequence analysis of the *HCI* revealed that this gene does not have homology with other known genes in the databank. It only showed homology to three human expressed sequence tags. The nucleotide sequence of *HCI* has three possible ORFs, and we could not predict a protein coded continuously by this cDNA. We found that only the longer ORF coding for 59 amino acids is responsible for the DNA repair activity detected in bacterium cells. We looked for homology in different protein and peptide databanks for this small protein coded by the second ORF and still could not find any significant homology. Despite the absence of such homology, the high isoelectric point of *HCI* protein indicates that it has a positive charge that would allow it to interact with DNA. Recently, a gene (REX1) of the alga *Chlamydomonas reinhardtii* that codes for a small protein of 8.9 kDa was identified and shown to participate in repair pathways caused by UV and MMS at a step preceding the excision (Cenkci et al., 2003). This protein has homologues in many organisms including humans. Although there is no homology between the *HCI* gene and REX1, the fact that such a small protein had been shown to work in a repair pathway supports our idea that *HCI* is in fact a small protein involved in the recognition and repair of DNA lesions.

Many other genes cloned by functional complementation of *E. coli* DNA repair mutants encode proteins directed to mitochondria or chloroplasts (Pang et al., 1992; Machado et al., 1996). The selection of genes in bacteria whose products function in these organelles may be explained by the likely prokaryotic origin of these organelles. Whether *HCI* protein is directed to mitochondria is an intriguing question, since no NER pathway has been identified so far in mitochondria (Mandavilli et al., 2002). Currently, we do not have any evidence that could be used to address this question.

The *in vitro* functional complementation of the HC1 protein is in agreement with its activity in bacterial cells. In the *in vitro* NER assay, we observed that HC1 only complements cell extracts deficient in the ERCC1 protein, an endonuclease responsible, together with XPF, for the 5' DNA incision of the damage. We did not find any evidence that the sole HC1 protein is able to cleave 5' of UV or cisplatin damage (not shown). As the XPF protein has an endonuclease catalytic activity but performs its activity only as heterodimers with ERCC1 (Marti and Fleck, 2004), it is possible that HC1 helps the XPF protein in 5' incisions. Based on our *in vivo* and *in vitro* data, we propose that HC1 protein may sense a DNA lesion before recruiting an endonuclease that cuts the DNA 5' from the damage. We are now investigating how this protein might be able to recognize a DNA lesion and with which partners does it interact.

The detection of HC1 proteins by specific antibodies only in embryonic cells leads to the speculation whether this protein would be specific to this developmental stage. During organogenesis there is an intense and fast cell division and any mutation or DNA blockade could be lethal. The importance of DNA repair at this stage of development is demonstrated by the phenotypes observed among null mutant mice lacking key repair genes (Vinson and Hales, 2002). A number of animals lacking DNA repair enzymes are not viable, with pre-implantation death being a common feature. Most of the embryonic losses occur during implantation, demonstrating the absolute requirement for repair ability when the embryonic cells start to proliferate rapidly. The presence of the HC1 protein may allow a more efficient repair at this stage. It is important to know if the role of this protein is to assist known DNA repair pathways or whether it is part of a new repair pathway present in embryonic cells.

The SAGE tag of this gene shows that it is poorly expressed but it is present only in proliferating cells, mainly different types of tumors. The gene was not found in any normal adult cells examined, which supports our idea of a specific embryonic gene and which also allows us to think that it could be turned on in some tumor cells, as with the telomerase gene (Cech, 2004). Thus, we have cloned a new human DNA repair gene that could be involved in lesion recognition and DNA incision and it appears to be specific for highly proliferative cells.

## ACKNOWLEDGMENTS

Research supported by Conselho Nacional de Pesquisa (Brazil-CNPq) and Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/COFECUB). We are grateful to Kátia Barroso Gonçalves (CNPq) and Neuza Antunes Rodrigues (CNPq) for technical support.

## REFERENCES

- Altschul SF, Gish W, Miller W, Myers EW, et al. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
- Bachmann BJ (1972). Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol. Rev.* 36: 525-557.
- Baute J and Depicker A (2008). Base excision repair and its role in maintaining genome stability. *Crit Rev. Biochem. Mol. Biol.* 43: 239-276.
- Biggerstaff M, Robins P, Coverley D and Wood RD (1991). Effect of exogenous DNA fragments on human cell extract-mediated DNA repair synthesis. *Mutat. Res.* 254: 217-224.
- Boiteux S, Huisman O and Laval J (1984). 3-Methyladenine residues in DNA induce the SOS function *sfhA* in *Escherichia coli*. *EMBO J.* 3: 2569-2573.
- Bonaldo MF, Lennon G and Soares MB (1996). Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res.* 6: 791-806.
- Bradford MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the

- principle of protein-dye binding. *Anal. Biochem.* 72: 248-254.
- Cech TR (2004). Beginning to understand the end of the chromosome. *Cell* 116: 273-279.
- Cenkci B, Petersen JL and Small GD (2003). REX1, a novel gene required for DNA repair. *J. Biol. Chem.* 278: 22574-22577.
- Christmann M, Tomicic MT, Roos WP and Kaina B (2003). Mechanisms of human DNA repair: an update. *Toxicology* 193: 3-34.
- Costa RM, Chigancas V, Galhardo RS, Carvalho H, et al. (2003). The eukaryotic nucleotide excision repair pathway. *Biochimie* 85: 1083-1099.
- Cunningham RP, Saporito SM, Spitzer SG and Weiss B (1986). Endonuclease IV (nfo) mutant of *Escherichia coli*. *J. Bacteriol.* 168: 1120-1127.
- Demple B, Johnson A and Fung D (1986). Exonuclease III and endonuclease IV remove 3' blocks from DNA synthesis primers in H<sub>2</sub>O<sub>2</sub>-damaged *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 83: 7731-7735.
- Friedberg EC (2003). DNA damage and repair. *Nature* 421: 436-440.
- Hanahan D (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166: 557-580.
- Hansson J and Wood RD (1989). Repair synthesis by human cell extracts in DNA damaged by cis- and trans-diamminedichloroplatinum(II). *Nucleic Acids Res.* 17: 8073-8091.
- Hitomi K, Iwai S and Tainer JA (2007). The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair. *DNA Repair* 6: 410-428.
- Kraemer KH, Patronas NJ, Schiffmann R, Brooks BP, et al. (2007). *Xeroderma pigmentosum*, trichothiodystrophy and Cockayne syndrome: a complex genotype-phenotype relationship. *Neuroscience* 145: 1388-1396.
- Leibeling D, Laspe P and Emmert S (2006). Nucleotide excision repair and cancer. *J. Mol. Histol.* 37: 225-238.
- Machado CR, de Oliveira RL, Boiteux S, Praekelt UM, et al. (1996). Thi1, a thiamine biosynthetic gene in *Arabidopsis thaliana*, complements bacterial defects in DNA repair. *Plant Mol. Biol.* 31: 585-593.
- Makarova E, Goes TS, Marcatto AL, Leite MF, et al. (2003). Serological differentiation of acute and chronic schistosomiasis using *Schistosoma mansoni* recombinant protein RP26. *Parasitol. Int.* 52: 269-279.
- Mandavilli BS, Santos JH and Van Houten (2002). Mitochondrial DNA repair and aging. *Mutat. Res.* 509: 127-151.
- Marti TM and Fleck O (2004). DNA repair nucleases. *Cell Mol. Life Sci.* 61: 336-354.
- Pang Q, Hays JB and Rajagopal I (1992). A plant cDNA that partially complements *Escherichia coli* recA mutations predicts a polypeptide not strongly homologous to RecA proteins. *Proc. Natl. Acad. Sci. U. S. A.* 89: 8073-8077.
- Perkins EL, Sterling JF, Hashem VI and Resnick MA (1999). Yeast and human genes that affect the *Escherichia coli* SOS response. *Proc. Natl. Acad. Sci. U. S. A.* 96: 2204-2209.
- Sancar A (1996). DNA excision repair. *Annu. Rev. Biochem.* 65: 43-81.
- Towbin H, Staehelin T and Gordon J (1992). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. *Biotechnology* 24: 145-149.
- Tudek B, Van Zeeland AA, Kusmierek JT and Laval J (1998). Activity of *Escherichia coli* DNA-glycosylases on DNA damaged by methylating and ethylating agents and influence of 3-substituted adenine derivatives. *Mutat. Res.* 407: 169-176.
- van Hoffen A, Balajee AS, Van Zeeland AA and Mullenders LH (2003). Nucleotide excision repair and its interplay with transcription. *Toxicology* 193: 79-90.
- Verhoeven EE, van Kesteren M, Moolenaar GF, Visse R, et al. (2000). Catalytic sites for 3' and 5' incision of *Escherichia coli* nucleotide excision repair are both located in UvrC. *J. Biol. Chem.* 275: 5120-5123.
- Vinson RK and Hales BF (2002). DNA repair during organogenesis. *Mutat. Res.* 509: 79-91.
- Wood RD, Robins P and Lindahl T (1988). Complementation of the *Xeroderma pigmentosum* DNA repair defect in cell-free extracts. *Cell* 53: 97-106.