AD_____

GRANT NUMBER DAMD17-94-J-4416

TITLE: Loss of DNA Base Excision Repair Capacity in Initiation and Progression of Breast Cancer

PRINCIPAL INVESTIGATOR: Robert Boorstein, M.D., Ph.D.

CONTRACTING ORGANIZATION: New York University Medical Center
New York, New York 10010-2598

REPORT DATE: March 1997

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
*OMB No. 0704-0188*3

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1997	Final (1 Nov 94 - 31 Jan 97)

4. TITLE AND SUBTITLE
Loss of DNA Base Excision Repair Capacity in Initiation and Progression of Breast Cancer

5. FUNDING NUMBERS
DAMD17-94-J-4416

6. AUTHOR(S)
Robert Boorstein, M.D., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
New York University Medical Center
New York, New York 10010-2598

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

Heritable alterations in Several genes including p53, ATM, BRCA1 and BRCA2 are known to play a role in cellular responses to DNA damage. Inherited or acquired alterations in these genes are thought to contribute to the development of human breast cancer. This study has focused on the purification and cloning of the mammalian homologue of the *E.Coli* repair enzyme endonuclease III, with the goal of determining whether this enzyme specically also plays a role in human breast cancer. Endonuclease III and its mammalian homologues are responsible for repairing several major lesions produced by oxidative damage to DNA.

During the grant period, we have purified a functional analog of endonuclease III 5,000 fold from calf thymus, and cloned and expressed the human full length cDNA. Ongoing studies will allow us to determine whether this enzyme plays a specific role in genetic or environmentally produced breast cancer by characterizing its expression at the molecular level.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer, DNA Repair, DNA glycosylases	31
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

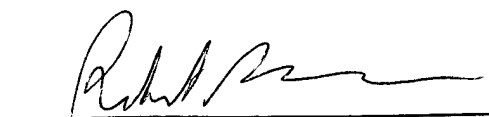
____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

RB For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.



PI - Signature

Jul 97

Date

4) TABLE OF CONTENTS

(5) INTRODUCTION

Base Excision repair. Excision repair is an enzymatically mediated system which effects removal of damaged portions of genomic DNA followed by resynthesis of the DNA to its original state. The system has evolved to repair damage caused by physico-chemical stresses ubiquitous to living organisms and is present throughout phylogenetic kingdoms. Two broad categories of DNA repair systems have been described. In nucleotide excision repair, "bulky" lesions such as those produced by UV induced dimerization of pyrimidines are recognized by a multienzyme complex and the segment of DNA including the DNA modification is removed by the action of an endonuclease complex. In base excision repair (the subject of this proposal) modified bases are recognized by specific DNA glycosylases which produce cleavage of the N-glycosyl bond. The enzymatically-mediated release of the modified base from the DNA backbone leaves an abasic (AP) site. These AP sites are then cleaved by endonucleases and the resulting single strand breaks are subsequently repaired by the sequential action of nucleases, DNA polymerases and DNA ligase, similarly to nucleotide excision repair (Friedberg, 1984).

Oxidative damage to DNA produces a variety of base modifications which are repaired by glycosylase type activities. Oxidative damage to the 5-6 double bond of thymine produces thymine glycol while damage to cytosine may produce cytosine glycol or its dehydrated counterpart, 5-hydroxycytosine (Dizdaroglu, et al., 1986). These modifications are repaired by the mammalian analogue of endonuclease III. If unrepaired these lesions may either block replication (which might lead to aberrant recombination, slippage during synthesis, etc.) or result in a point mutation. 5-hydroxymethyluracil is formed by oxidation of the methyl group of thymine and or by oxidation and deamination of methylcytosine (via a 5-hydroxymethylcytosine intermediate) (Frenkel et al, 1985; Cannon-Carlson, et al., 1989). HmUra is mutagenic when formed via the latter route. The incomplete repair of any of these lesions recognized by each of these DNA glycosylases might result in point mutations or more complex mutagenic alterations. Such mutations might activate oncogenes or produce dominant or loss of function mutations in tumor suppressor genes.

Cancer as a result of DNA repair deficiency. The importance of DNA repair to human well-being through the prevention of cancer is emphasized by the demonstration of a DNA repair defect in the cancer prone hereditary disease *Xeroderma pigmentosum*. Other well characterized inherited cancer prone syndromes, including Bloom's syndrome, Fanconi's anemia and ataxia telangiectasia (AT), have been linked to defects in the ability to respond normally to specific forms of DNA damage. Although these syndromes are rare autosomal recessive disorders, it would be of significant consequence to the general population if heterozygotes also have a propensity to develop tumors. Indeed, it has been proposed that up to 5% of patients who develop breast cancer in early years are AT heterozygotes (Swift, et al, 1991). Such a result suggests that even a partial deficiency of a DNA repair function might accelerate the initiation or progression of cancer.

The role of oxidative and spontaneous DNA damage in tumorigenesis. The development of breast cancer results from the combined effects of heritable, hormonal, environmental and nutritional factors. Heritable alterations in several genes including p53, AT, the BRCA1 locus, and other, as yet undetermined, loci are thought to contribute to the development of human breast cancer (Borresen, 1992; Easton, 1993). Acquired mutations in tumor suppressor genes (including p53) and proto-oncogenes are also thought to play a role in the development of breast cancer. The source of such acquired mutations might be either "spontaneous" damage to DNA, such as deamination of adenine or cytosine, aberrant methylation, or oxidative modifications of pyrimidines including thymine glycol, 5-

hydroxycytosine, 5-hydroxymethyluracil or of guanine to 8-oxoguanine (Breimer and Lindahl, 1984; Teebor et al., 1988). Each of these lesions is repaired via the base excision repair pathway which is initiated by removal of the modified base through the action of a DNA-glycosylase specific for each lesion . While the factors determining the qualitative and quantitative yields of such base damage are not well characterized, epidemiological evidence has linked prooxidant dietary and serum factors to an increased risk of cancer , and antioxidant micronutrients in diet and serum to a reduced risk of cancer (Ames, 1993; Stephens, et al., 1988).

The paradigm: Cancer as a result of acquired mutations in DNA repair genes. We propose that DNA glycosylases function as tumor suppressor genes protecting against the development of breast cancer. Partial loss of repair capacity of a specific DNA glycosylase might result from an acquired or inherited mutation in one allele for the DNA glycosylase gene. Such a deficiency might promote further mutations in other tumor suppressor genes, proto-oncogenes, in the initially mutated repair gene, or in other repair genes. If during this process, both copies of a DNA glycosylase gene are inactivated, the incipient tumor cell might have the equivalent of a mutator phenotype in which the progression towards malignancy would be further accelerated. If this paradigm proves true, then the determination of DNA repair capacity might ultimately have prognostic or diagnostic significance in the evaluation of breast cancer. This paradigm is based on the supposition that although DNA glycosylases are necessary for the evolutionary survival of a species, loss of repair function is not incompatible with cell viability. Cells in culture can be viable missing at least some of these activities. Mutants of nearly every DNA glycosylase have been isolated in *E.coli*, which are viable but have a mutator phenotype (Tye, et al., 1977). We have isolated a mammalian cell line completely lacking hmUra-DNA glycosylase which has normal growth characteristics (Boorstein, et al., 1992).

Working Hypothesis. The overall goal of this project is to study the role of enzymes which effect the repair of oxidative damage and spontaneous deamination of DNA bases in the initiation and progression of breast cancer. The hypothesis we are testing is whether diminution or loss of DNA repair capacity leads to increased mutagenesis in genes responsible for the initiation and progression of breast cancer. Acquired or inherited loss of one copy of a DNA-glycosylase gene, might, via a gene dosage effect, lead to diminished repair capacity and contribute to the production of mutations in tumor suppressor genes, proto-oncogenes or in other DNA repair genes. The determination of DNA repair capacity may have prognostic or diagnostic significance in the evaluation of breast cancer.

The obvious limitation of our original proposal is that studies of activity in breast cancer cells may not be sufficiently reliable to confirm or rebut this important hypothesis. Cells in culture may lose repair capacity, although the original tumor may contain normal levels. Alternatively, loss of some activity, especially heterozygosity, may not be detectable. Thus, it is ultimately necessary to isolate the genes for these potentially critical genes and study them at the molecular level.

Overall Approach. During the first year of this two year proposal, we have pursued our initial aim of measuring DNA glycosylase activities in breast cancer cells. We then modified and modified the aims of our research efforts according to the Statement of Work (REVISED). Recognizing the limitations of the approach described in our original proposal, we focused our efforts on have developed improved assays for the measurement of DNA glycosylase activities, which give markedly improved sensitivity. To date, these efforts have not resulted in development of new assays with sufficient sensitivity or specificity to use on clinical material. However, we have made substantial progress by cloning of one of the DNA glycosylase activities involved in oxidative damage. The cloning of the DNA repair

enzymes will ultimately prove necessary for the successful determination of the role of DNA repair enzymes in breast cancer. In the future, we will continue these efforts and utilize the newly isolated genes and relevant molecular approaches to complete the initial studies.

(6) BODY

Our original goal was to determine whether one or more of six DNA-glycosylase activities is significantly reduced or absent in breast tumor cells. Uracil-DNA glycosylase and hypoxanthine-DNA glycosylase are being measured as examples of enzymes responsible for repairing spontaneous deamination damage. 5-Hydroxymethyluracil-DNA glycosylase, 5-hydroxymethylcytosine-DNA glycosylase, mammalian "endonuclease III", and 8-oxoguanine-DNA glycosylase are measured as examples of enzymes responsible for repairing oxidative damage. These studies utilize breast tumor cells in culture. Compared with human breast tumor material from surgical specimens, tumor cell lines are more easily acquired, and are not contaminated by normal or necrotic tissue. Assays can readily be done in batches and in replicate.

Our goal WASs to determine whether such cell lines have significant reduction or complete absence of the DNA glycosylase activities which repair oxidative or deamination damage by measuring the specific activity of each of the six enzymes. Ideally one would like to be able to determine whether there are mutations in the genes coding for these enzymes, loss of homozygosity for these genes, or alterations in gene expression. However, at the time of the initial submission, no human or any other mammalian DNA glycosylase genes for these enzymes had been reliably cloned (with the possible exception of uracil -DNA glycosylase) (Meyer-Siegler, et al., 1991; Muller, and Caradonna, 1991; Olsen, et al., 1989) and antibodies to the enzymes were not readily available. Therefore, it was not yet possible to assess gene expression by any other technique than direct measurement of enzyme activity.

Interpretation of results and future directions The absence of any one of the DNA glycosylase activities in any tumor cell line will be a significant positive result. Such a loss could result from inactivation of both alleles via a combination of either germline (spontaneous or inherited) or somatic mutations. This inactivation scheme is analogous to the tumor suppressor paradigm described for the Rb gene and the p53 gene. It may also be possible to detect an intermediate level of enzyme activity resulting from a reduction to heterozygosity. This might be the result of germline heterozygosity or an acquired inactivation during the process of tumorigenesis in breast cells.

From an epidemiological standpoint, either result would warrant investigation of breast cancer patients and their families with the goal of determining whether there are subpopulations of individuals heterozygous for one or more of the DNA glycosylase genes. It may be ultimately be possible to detect subpopulations of individuals bearing germline mutations putting them at increased risk for cancer, and early preventive screening could be performed. If the enzymes which repair oxidative damage are seen to be lost, we will have developed a mechanistic rationale to provide women at risk with currently available or as yet undeveloped antioxidant supplementation. Even if only one or two percent of breast cancers result from each reduced enzyme activity measured here, the total contribution such loss of DNA repair capacity makes to the development of breast cancer could be substantial.

In the long run, the detection of germline or somatic mutations in DNA repair genes might also be useful in predicting the clinical course of breast cancer and the response to therapy. The loss of a specific repair activity might prove to be correlated with a poor clinical outcome, or lack of response to radiation or chemotherapy.

During the previous granting period we therefore focused on continuing our efforts to clone two of the genes for the DNA-glycosylase activities under study in this project. Our

purification and cloning of the mammalian "endonuclease III" enzyme has been published in two parts (see Appendix). We are also attempting to isolate the gene for hydroxymethyluracil-DNA glycosylase (hmUDG), and in the past grant period, did a series of experiments designed to try to develop a new complementation strategy.

1. Tumor cell acquisition and growth.

During the initial grant interval we brought the following breast lines into our laboratory and characterized their growth characteristics. Our initial studies were on the following lines. MCF-7 (ER+), ZR-75-1, BT-20, and MDA-MB-231: We found their growth characteristics (growth rate and cycling fraction) to be variable, and realized that obtaining sufficient numbers of cells to perform enzymatic assays on crude extracts, using our existing methodology at the time, was likely to be difficult if not impossible. Therefore, we devoted significant effort to increasing the sensitivity of our assays and to the cloning of the novel DNA repair genes.

2. Development of assays for DNA glycosylase activities: Assays of lyases. Historically, we had measured endonuclease III activity by measuring release of 3H-containing modified bases from polynucleotides. With Dr. Cunningham, Dr. Teebor and I have developed a methodology to measure endonuclease activity by reducing enzyme bound to a P32-labelled oligonucleotide substrate to make an irreversible covalent linkage that could be visualized on a gel. This assay is highly specific and sensitive, and was the basis for our successful purification of the mammalian enzyme. These results have been reported in the manuscript in the attached Appendix.

Dr. Teebor has developed the hypothesis that several enzymes including endonuclease III, 8-oxoG DNA glycosylase, and hypoxanthine DNA glycosylase, are in fact "lyases" as opposed to true "glycosylases" such as uracil-DNA glycosylase and hmUra-DNA glycosylase. Drs. Teebor, Cunningham and myself are working on a manuscript describing the mechanism of action of endonuclease III given this hypothesis. Given this hypothesis, the methodology we have used to purify and measure mammalian "endonuclease 111" is being utilized to measure methodology is not being extended to 8-oxoG DNA glycosylase, and hypoxanthine DNA glycosylase. Preliminary experiments indicate that the methodology can be reliably used to identify 8-oxoG DNA glycosylase in a highly sensitive and specific manner. However, we have not yet developed either assay to enable us to quantitatively measure glycosylase/lyase activities in a quantitative manner.

3. Cloning of DNA repair genes.

a. Purification and characterization of the endonuclease III like activity of mammalian cells. We initially purified a functional analog of endonuclease III 5,000 fold from calf thymus. Like endonuclease III, the mammalian enzyme demonstrates DNA-glycosylase activity against both pyrimidine hydrates (which are produced by ultraviolet radiation) and thymine glycol (which is produced by ionizing radiation or oxidative stress). It also contains an AP lyase activity (which catalyzes DNA strand cleavage at AP sites via β-elimination). The functional similarity of the bacterial and mammalian enzymes suggested a strategy for definitive identification of the mammalian protein based on the nature of its putative enzyme-substrate (ES) intermediate. Prokaryotic DNA glycosylase/AP lyases had been shown to function through *N*-acylimine (Schiff's base) ES intermediates. Chemical reduction of such

intermediates to stable secondary amines results in the irreversible cross linking of the repair enzymes to the oligodeoxynucleotides containing the lesions which they repair. We incubated endonuclease III with a ^{32}P-labeled thymine glycol-containing oligodeoxynucleotide in the presence of $NaCNBH_3$. This resulted in an increase in the apparent molecular weight of the enzyme by SDS-PAGE, and phosphorimaging confirmed irreversible cross linking between endonuclease III and DNA. Identical treatment of the most purified mammalian enzyme fraction resulted in reductive cross linking of the oligodeoxynucleotide to a predominant 31 kD species. Amino acid analysis of this mammalian 31 kD enzyme revealed homology to *E. coli* endonuclease III, to theoretical proteins of *S. cerevisiae* and *C. elegans.* and to the translated sequences of rat and human 3'ESTs.

The human 3'EST was then used to isolate the cDNA clone encoding the human enzyme from a human spleen cDNA library, by use of the GENETRAPPER selection. The 5'terminus of the cDNA was confirmed by 5'RACE, and the full length cDNA was shown to be expressed in human splenic cells and human T cells via Northern analysis. The human cDNA was then expressed as a GST-fusion protein, which demonstrated thymine glycol-DNA glycosylase activity and, after incubation with $NaCNBH_3$, became irreversibly cross-linked to a thymine glycol-containing oligodeoxynucleotide. The gene for the human enzyme was localized to chromosome 16p13.2-.3.

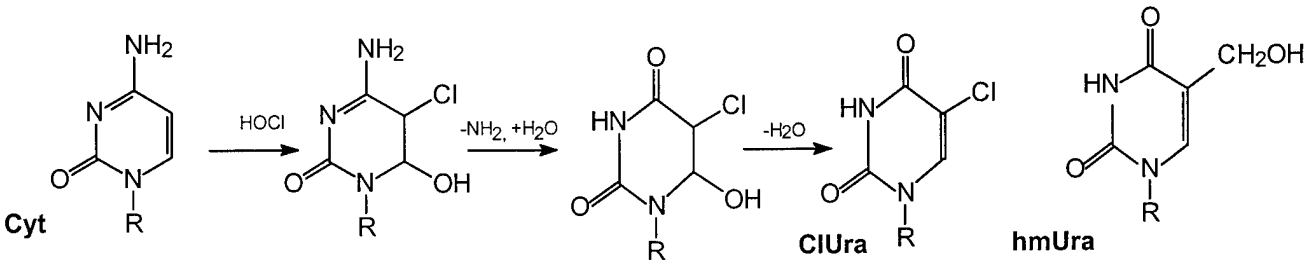
Amino acids within the active site, DNA binding domains and [4Fe-4S] cluster of the *E. Coli* endonuclease III were shown to be conserved in the human enzyme. Additionally, sequences encoding putative endonuclease III homologues are present in the genomes of bacteria, archaea and eukarya. The ubiquitous distribution of these proteins suggests that the 5,6 double bond of pyrimidines is subject to oxidation, reduction and/or hydration in the DNA of organisms of all biologic domains, and that the resulting modified pyrimidines are deleterious to the organism.

Ongoing studies will allow us to determine whether this enzyme plays a specific role in genetic or environmentally produced breast cancer by characterizing its expression at the molecular level. We are determining whether breast tumor cell lines and tissues contain the normal gene, and whether it is expressed in normal amounts. In addition, we have utilized the human cDNA to isolate the murine cDNA and are in the process of developing a knockout mouse. Studies in such animals will allow us to determine whether the absence of this DNA repair activity leads to increased spontaneous breast cancers, or to increased cancers as a result of radiation or other oxidative stresses. We will be able to determine whether the gene is present by Southern analysis or PCR, and whether it is expressed, by Northern or RT PCR. Studies of potential mutations can be undertaken. In addition, we have developed an antibody against the human enzyme to perform immunohistochemistry and Western blot studies.

b. A novel strategy for the cloning of the hydroxymethyluracil-DNA glycosylase gene of mammalian cells. It was suggested to us by B. Ames that hmUDG might repair biologically important lesions other than hmUra. He suggested that Chlorouracil might be a substrate for hmUDG. Chlorouracil, he has proposed, is formed in DNA as a result of reaction with hypochlorite, a product of inflammatory cells (see Figure 1). He has demonstrated its formation in DNA in solution (personal communication). Chlorouracil has a Cl moiety at the identical position where hmUra has a hydroxymethyl moiety. One would predict that

formation of Chlorouracil, as a deamination product of cytosine, would be toxic or mutagenic. If hmUDG has a role to protect cells against Chlorouracil, then one would predict that the repair deficient cells would be hypersensitive to hypochlorite. Our preliminary results show that this is the case, with our hmUDG deficient cells being 10 fold more sensitive to hypochlorite than wild type cells in terms of toxicity, and 3 fold more sensitive in terms of mutagenicity. We are currently determining with Dr. Ames whether repair deficient cells are unable to repair this new lesion.

Figure 1. Proposed formation of Chlorouracil



These experiments are designed to demonstrate a previously unknown biological role for hmUDG, directly related to repair of damage formed during the inflammatory process.

These findings were designed to be used for a new strategy to clone the hmUDG gene, in parallel with the direct efforts of our co-investigators to purify the gene directly. Briefly, the finding that the V79*mut*1 cells are sensitive to hypochlorite, along with a demonstration of a lack of chlorouracil repair activity, would make it possible to attempt to clone the gene by a straightforward complementation approach. Briefly, V79*mut*1 cells could be transfected with a human cDNA library with the human cDNA on a vector containing another selectable marker such as neo^r. Neomycin resistant cells that are able to survive in hypochlorite would be screened for recovery of glycosylase activity. The gene should then be able to be readily isolated from these transfected cells by the proximity of the gene to the selectable marker. A genomic library will could be constructed from the transfected cells, and sequences adjacent to the selectable marker will be mapped. These in turn could be used to isolate cDNA clones from the original library that will presumptively contain the repair gene. They could be proved to contain the repair gene by repeat transfection of V79*mut*1 cells or directly by expression in an expression system.

Unfortunately, since the time these studies were proposed, we have demonstrated that the CldUrd is not repaired by hmUDG or by any other enzyme. We have performed additional studies on the mechanism of toxicity of CldUrd, about which we are currently preparing an additional manuscript.

(7) CONCLUSIONS

Damage to DNA, the genetic material of cells, has been implicated in the development of human cancers. Such damage might occur as a result of environmental exposures, diet, or the normal processes of aging. Highly specific enzymatic pathways have evolved to repair such damage. Unrepaired damage presumably leads to changes in important cellular genes, which, when they malfunction, ultimately result in the formation of

cancers. The hypothesis we are testing is whether diminution or loss of the capacity to repair specific classes of such damage is responsible for the initiation and progression of breast cancer.

In this study, we initially isolated one such DNA repair protein. Then, we cloned the human gene for this important DNA repair activity. Having now isolated this human repair gene, we will be able to determine whether it plays a specific role in the development of breast cancer. Ongoing studies will allow us to determine whether this enzyme plays a specific role in genetic or environmentally produced breast cancer by characterizing its expression at the molecular level. In the long run, the determination of DNA repair capacity in individual patients may have prognostic or diagnostic significance in the evaluation and treatment of breast cancer.

From an epidemiological standpoint, our goal is to determine whether there are subpopulations of individuals heterozygous for one or more of the DNA glycosylase genes. It may be ultimately be possible to detect subpopulations of individuals bearing germline mutations putting them at increased risk for cancer, and early preventive screening could be performed. If the enzymes which repair oxidative damage are seen to be lost, we will have developed a mechanistic rationale to provide women at risk with currently available or as yet undeveloped antioxidant supplementation. Even if only one or two percent of breast cancers result from each reduced enzyme activity, the total contribution such loss of DNA repair capacity makes to the development of breast cancer could be substantial. In the long run, the detection of germline or somatic mutations in DNA repair genes might also be useful in predicting the clinical course of breast cancer and the response to therapy. The loss of a specific repair activity might prove to be correlated with a poor clinical outcome, or lack of response to radiation or chemotherapy.

(8) REFERENCES

Ames, B., Proc. Natl.. Acad. Sci. U.S.A., 88:8460-8464, 1991.

Band, V. and Sager, R., Proc. Natl.. Acad. Sci. U.S.A., 86:1249-1253, 1989.

Bessho, T., et al., J. Biol. Chem., 268:19416-19421, 1993.

Boorstein, R.J., *et al.,* Biochemistry, 29:10455-10460, 1990.

Boorstein, R.J., Chiu, L.-N., and Teebor, G.W., Nucl. Acids Res., 17:7653-7661, 1989.

Boorstein, R.J., Chiu, L.-N., and Teebor, G.W., Molec. Cell. Biol., 12: 5536-5540, 1992.

Boorstein, R.J., Levy, D.D. and Teebor, G.W., Mutat. Res., 183: 257-263, 1987.

Boorstein, R. J., Levy, D.D. and Teebor, G.W. Cancer Res., 47: 4372-4377, 1987.

Borresen, A-L., Acta Oncol. 31:151-1555, 1992. .

Breimer, L. H., and Lindahl, T., J. Biol. Chem. 259: 5543-5548, 1984.

Cannon, S., Cummmings, A., and Teebor, G.W., Biochem. Biophys. Res. Comm., 151: 1173-1179, 1988.

Cannon-Carlson, S.V., Gokhale, N.V., and Teebor, G.W., J. Biol. Chem., 264:13306-13312, 1989.

Dianov, G. and Lindahl, T., Nucl. Acids Res., 19:3829-3833, 1991.

Dizdaroglu, M., et al., Biochem. J., 235:531-536, 1986.

Easton, D.F., et al., Am. J. Hum. Genet., 52:678-701, 1993

Frenkel, K., et al., Biochemistry **24**:4527-4533, 1985.

Friedberg, Errol C., *DNA Repair.* New York:Freeman, 1984.

Hackett, A.J., et al., J. Natl. Canc. Inst., 58:1795-1806, 1977.

Meyer-Siegler, K., et al., Proc. Natl.. Acad. Sci. U.S.A., 88:8460-8464, 1991.
Muller, S.J., and Caradonna, S., Biochim. Biophys. Acta., 1088:197-207, 1991
Olsen, L.C., et al., EMBO Jour., 8:3121-3125, 1989.
Stevens, R.G., et al., New Eng. J. Med., 319:1047-1052, 1988.
Swift, M., et al., New Eng. J. Med., 325:1831-1836, 1991.
SO N-Engl-J-Med. 1991 Dec 26. 325(26). P 1831-6.
Tchou, J., and Grollman, A.P., Mutat. Res., 299:277-287, 1993.
Teebor, G.W., Boorstein, R.J. and Cadet, J., Int. J. Radiat. Biol., 54:131-150, 1988.
Tye, B.-K., et al., Proc. Natl.. Acad. Sci. U.S.A., 74:154-157, 1977.

(9) APPENDIX:

Manuscripts (3).

1. Hilbert TP, Boorstein RJ, Xing D, Kung HC, Bolton PH, Cunningham RP, and Teebor GW. Purification of a Pyrimidine Hydrate DNA-glycosylase/AP lyase from Calf Thymus, a Mammalian Analogue of *Escherichia Coli* Endonuclease III, Biochemistry 1996;35: 2505-2511.

2. Hilbert TP, ChaungW, Boorstein RJ, CunninghamRP, Teebor GW. Cloning and expression of the cDNA encoding the human homologue of the DNA repair enzyme, *Escherichiae coli* endonuclease III. Journal of Biological Chemistry 1997;272:6733-6740.

3. Chaung W, Mi L-J, Boorstein RJ. The p53 status of Chinese hamster V79 cells frequently used for studies on DNA damage and DNA repair. Nucleic Acids Research 1997;25: 992-994.

Personnel supported by the project:

Robert Boorstein
George Teebor
Wenren Chaung
Li-Jun Mi

Reprinted from Biochemistry, 1996, 35.

Purification of a Mammalian Homologue of *Escherichia coli* Endonuclease III: Identification of a Bovine Pyrimidine Hydrate-Thymine Glycol DNA-Glycosylase/AP Lyase by Irreversible Cross Linking to a Thymine Glycol-Containing Oligodeoxynucleotide[†]

Timothy P. Hilbert,[‡] Robert J. Boorstein,[‡] Hsiang Chuan Kung,[§] Philip H. Bolton,[§] Dongxia Xing,[∥] Richard P. Cunningham,[⊥] and George W. Teebor*[,‡]

Department of Pathology and Kaplan Comprehensive Cancer Center, New York University Medical Center, 550 First Avenue, New York City, New York 10016, Department of Chemistry, Wesleyan University, Middletown, Connecticut 06459, and Departments of Chemistry and Biological Sciences, SUNY at Albany, Albany, New York 12222

Received October 20, 1995; Revised Manuscript Received December 20, 1995[⊗]

ABSTRACT: We purified a homologue of the *Escherichia coli* DNA repair enzyme endonuclease III 5000-fold from calf thymus which, like endonuclease III, demonstrates DNA-glycosylase activity against pyrimidine hydrates and thymine glycol and AP lyase activity (DNA strand cleavage at AP sites via β-elimination). The functional similarity between the enzymes suggested a strategy for definitive identification of the bovine protein based on the nature of its enzyme−substrate (ES) intermediate. Prokaryotic DNA glycosylase/AP lyases function through *N*-acylimine (Schiff's base) ES intermediates which, upon chemical reduction to stable secondary amines, irreversibly cross link the enzyme to oligodeoxynucleotides containing substrate modified bases. We incubated endonuclease III with a ^{32}P-labeled thymine glycol-containing oligodeoxynucleotide in the presence of $NaCNBH_3$. This resulted in an increase in the apparent molecular weight of the enzyme by SDS−PAGE. Phosphorimaging confirmed irreversible cross linking between enzyme and DNA. Identical treatment of the most purified bovine enzyme fraction resulted in irreversible cross linking of the oligodeoxynucleotide to a predominant 31 kDa species. Amino acid analysis of the 31 kDa species revealed homology to the predicted amino acid sequence of a *Caenorhabditis elegans* 27.8 kDa protein which, in turn, has homology to endonuclease III. The translated amino acid sequences of two partial 3′ cDNAs, from *Homo sapiens* and *Rattus sp.*, also demonstrate homology to the *C. elegans* and bovine sequences suggesting a homologous family of endonuclease III-like DNA repair enzymes is present throughout phylogeny.

The DNA repair enzyme *Escherichia coli* endonuclease III was initially identified by its nicking activity directed against UV-irradiated DNA (Radman, 1976). Subsequently, it was shown that nicking of UV-irradiated DNA resulted from two enzymatic activities, a DNA-glycosylase which released pyrimidine hydrates from the DNA backbone, yielding an apyrimidinic (AP) site (Boorstein et al., 1989), and an activity which effected strand cleavage via β-elimination of the 3′ phosphate group of the apyrimidinic sugar residue (Bailly & Verly, 1987; Kim & Linn, 1988; Mazumder et al., 1991). The latter activity has been termed an AP lyase to distinguish it from AP endonucleases, such as exonuclease III or endonuclease IV, which catalyze strand cleavage via hydrolysis of phosphodiester bonds (Bailly & Verly, 1989). Endonuclease III is one of a group of enzymes,

[†] Supported by NIH-CA 16669 and CA 49869 (G.W.T.), CA 16087 (Kaplan Cancer Center), GM 07308 (MSTP for T.P.H.), GM46312 (R.P.C.), USAMRDC DAMD 17-94-J-4416 (R.J.B.), and, in part, DOE-DE-FG02-93ER61598 (P.H.B.).

* Corresponding author.
[‡] New York University Medical Center.
[§] Wesleyan University.
[∥] Department of Chemistry, SUNY at Albany.
[⊥] Department of Biological Sciences, SUNY at Albany.

[⊗] Abstract published in *Advance ACS Abstracts,* February 15, 1996.

0006-2960/96/0435-2505$12.00/0

including T4 endonuclease V and the *E. coli* Fpg protein (MutM), which demonstrate both DNA-glycosylase and AP lyase activities (Demple & Harrison, 1994; Dodson et al., 1994).

In addition to excising pyrimidine hydrates, the DNA-glycosylase activity of endonuclease III also excises pyrimidine glycols, ring-contracted pyrimidine derivatives, such as 5-hydroxymethylhydantoin, and urea residues composed of the N1−C2−N3 atoms of the pyrimidine skeleton (Strniste & Wallace, 1975; Demple & Linn, 1980; Breimer & Lindahl, 1984; Cunningham & Weiss, 1985). Enzyme activities functionally analogous to endonuclease III have been identified in bacteria other than *E. coli*, in yeast, and in mammalian cells and tissues through the use of UV-irradiated, chemically oxidized, and γ-irradiated DNA as substrates (Brent, 1973; Bacchetti & Benne, 1975; Duker & Teebor, 1975; Ness & Nissen-Meyer, 1978; Doetsch et al., 1986). We showed that extracts of Hela cells contained a thymine glycol DNA-glycosylase (Higgins et al., 1987). We also demonstrated that both endonuclease III and Hela cell extracts released cytosine hydrate (as well as its deamination product, uracil hydrate) from UV-irradiated DNA (Boorstein et al., 1989). Kim and Linn (1989) described two, or possibly three, UV-endonuclease activities in Hela cells by monitoring the nicking of UV-irradiated circular DNA. Huq et al. (1992) reported a 25-fold purification of an endonuclease III-like activity from calf thymus and stated that the *N*-terminal sequence of this protein was not homologous to other known proteins.

We undertook the purification of a mammalian endonuclease III-like enzyme from calf thymus by monitoring its DNA glycosylase activity against UV-induced pyrimidine hydrates. The assay measures release of ^{3}H-labeled pyrimidine hydrates and is reproducible and linear with respect to time and protein concentration. The substrate is easily prepared, and, most importantly, the chemical identity of the enzymatically released photoproducts can be corroborated by HPLC analysis. Calf thymus was chosen as the source of enzyme because it contains endonuclease III-like activity and because large amounts of very fresh tissue are available.

A novel approach to the definitive identification of the mammalian enzyme was the application of a chemical reaction which results in the irreversible cross linking of the enzyme to its DNA substrate. *N*-Acylimine (Schiff's base) enzyme−substrate (ES) intermediates are characteristic of the prokaryotic DNA glycosylase/AP lyases described to date. Such intermediates can be irreversibly stabilized through chemical reduction to secondary amines. In such a way T4 endonuclease V (Dodson et al., 1993) and the *E. coli* Fpg protein (Tchou & Grollman, 1995) were irreversibly cross-linked to substrate oligodeoxynucleotides containing a cyclobutane dimer and an 8-oxoguanine residue, respectively. The reductive cross linking of enzyme to an oligodeoxynucleotide permits identification of the mammalian protein by two experimental parameters. The first is an increase in the apparent molecular weight of the enzyme as determined by SDS−PAGE. Second, if the oligodeoxynucleotide is 5′-end-labeled with ^{32}P, the irreversibly cross-linked protein−DNA complex can be detected by autoradiography or phosphorimaging after SDS−PAGE.

On the basis of the results obtained with endonuclease V and the Fpg protein, we anticipated successful irreversible cross linking of *E. coli* endonuclease III to an oligodeoxynucleotide containing one of the enzyme's known substrates, thymine glycol. Assuming that the mammalian enzyme also functions through a *N*-acylimine ES intermediate, we could then apply the reductive cross linking reaction to the purified mammalian enzyme fractions using the same oligodeoxynucleotide. This would permit isolation of the correct protein species from a SDS−polyacrylamide gel in sufficient amount for primary amino acid sequencing.

EXPERIMENTAL PROCEDURES

Buffers. Homogenization buffer: 25 mM HEPES, pH 7.5, 15 mM NaCl, 1 mM DTT, 2 mM EDTA, 0.5 mg/mL leupeptin, 0.7 mg/mL pepstatin, 0.2 mM phenylmethanesulfonyl flouride. HDE: 25 mM HEPES, pH 7.5, 1 mM DTT, 2 mM EDTA.

Enzyme. E. coli endonuclease III was purified from *E. coli* strain UC6444 carrying the plasmid pHIT1 as previously described (Asahara et al., 1989).

Radionucleotides. [5,5′-^{3}H]Deoxycytidine 5′-triphosphate (15−30 Ci/mmol) and [*methyl*-^{3}H]thymidine 5′-triphosphate (70−90 Ci/mmol) were purchased from Du Pont/NEN).

Oligodeoxynucleotides. Alternating poly(dG-dC) and poly(dA-dT) were purchased from Pharmacia.

Purification of a Pyrimidine Hydrate DNA-Glycosylase from Calf Thymus. All purification procedures were carried out at 4 °C, unless otherwise indicated. Freshly obtained calf thymus (1.2 kg) was homogenized in a Waring Blendor in 4.8 L of homogenization buffer and further fragmented by sonication in 300 mL aliquots for 3 min at 70% power using a Heat Systems model W-375 sonicator equipped with a model 305 high gain horn. NaCl (4 M) was added to a final concentration of 320 mM and the gelatinous precipitate removed manually by spooling using a 10 mL glass pipette as a stirring rod. The remaining solution was cleared by centrifugation at 10000*g*, filtered through cheesecloth, and diluted with 1.7 volumes of HDE to produce fraction I (4000 mL).

Fraction I was batch extracted with 450 mL (packed volume) of cation exchange resin (SP Fast-flow, Pharmacia) preequilibrated with HDE containing 150 mM NaCl. After the beads settled the supernatant was discarded, and the beads were poured into an XK 26/60 column (Pharmacia). They were washed with 500 mL of HDE containing 150 mM NaCl, followed by a 2 L gradient from 150 to 700 mM NaCl at 4 mL/min. Twenty milliliter column fractions were collected and assayed. Fractions 45−75 were pooled to yield fraction II (620 mL).

Solid ammonium sulfate was added to fraction II, which contained approximately 350 mM NaCl, to a final saturation of 21% (120 g/L solution). The sample was centrifuged at 12000*g* for 20 min to remove precipitate and the supernatant applied to a C 26/40 column (Pharmacia), containing 150 mL (bed volume) of Octylsepharose 4 Fast-flow media (Pharmacia) preequilibrated with HDE, 21% ammonium sulfate, and 300 mM NaCl. The column was washed with 150 mL of HDE, 21% ammonium sulfate, and 300 mM NaCl followed by a 1.5 L gradient beginning with HDE, 21% ammonium sulfate, and 300 mM NaCl and finishing with HDE containing neither ammonium sulfate nor NaCl at 3 mL/min, collected in 20 mL fractions. One milliliter aliquots of the column fractions were dialyzed into HDE and 125 mM NaCl and assayed for enzymatic activity. Active

fractions (31−44) were pooled and dialyzed into HDE and 125 mM NaCl (fraction III, 280 mL).

Fraction III was concentrated by loading onto an HR 10/10 Mono S column (Pharmacia) and eluting via a step increase in NaCl concentration to HDE and 0.5 M NaCl. One milliliter fractions were collected and assayed, and 12 active fractions were pooled. The 12 mL sample was divided into 3 × 4 mL aliquots each of which were fractionated via gel filtration chromatography through a Hiload 26/60, Superdex 75 pg column (Pharmacia), run in HDE and 350 mM NaCl (2.5 mL/min), and collected in 2.5 mL fractions. The gel filtration column was precalibrated with the Gel Filtration Low Molecular Weight Calibration Kit from Pharmacia. Active fractions (70−75, approximately 29 kDa) from each of three column runs were pooled to 45 mL which was diluted from 350 to 125 mM NaCl with 1.8 volumes of HDE. The sample was then loaded onto a HR 5/5 MonoS column (Pharmacia) and concentrated via step elution with HDE and 0.5 M NaCl. Enzymatic activity eluted in six 0.5 mL fractions which were pooled to yield fraction IV (3 mL).

Fraction IV was diluted to 100 mM NaCl with four volumes of HDE, loaded onto a 1 mL single-stranded DNA-cellulose (ssDNA-cellulose, Sigma) HR 5/5 column (Pharmacia) and eluted with a 12.5 mL gradient (100−600 mM NaCl) (0.2 mL/min). Fractions 15−17 were pooled to yield fraction V (1.5 mL).

Preparation of Substrates for DNA-Glycosylase Assays. Poly(dG-[^{3}H]dC) was produced as described previously (Boorstein et al., 1989), by nick translation of poly(dG-dC) (Pharmacia) with [5,5′-^{3}H]dCTP (Du Pont/NEN), and purified using Nick-Spin columns (Pharmacia). Poly(dG-[^{3}H]-dC) produced in this manner had a specific activity of 1.2 × 10^6 cpm/μg. This DNA was then exposed to 400 kJ/m^2 of UV radiation at 254 nm (two 15 W germicidal bulbs) to induce the formation of cytosine hydrate. UV flux was quantitated using a UVX 54 radiometer (UVP Inc., San Gabriel, CA).

Poly(dA-[^{3}H]dT) was produced by the nick-translation of poly(dA-dT) with [*methyl*-^{3}H]dTTP, followed by oxidation of the alternating copolymer with osmium tetroxide to form thymine glycol residues (Higgins et al., 1987). The radiolabeled, oxidized DNA was purified by passing it twice through Nick-Spin columns (Pharmacia). Thymine glycol-containing poly(dA-[^{3}H]dT) produced in this manner had a specific activity of approximately 7 × 10^6 cpm/μg.

DNA-Glycosylase Assays. Pyrimidine hydrate and thymine glycol DNA-glycosylase assays were carried out against UV-irradiated and oxidized DNA substrates, respectively, as follows: enzyme aliquots were incubated with 0.1 μg of substrate DNA in a reaction mixture containing 15 mM HEPES, pH 7.5, 75 mM NaCl, 10 mM EDTA, and 1 mM DTT in a volume of 60 μL for specified periods of time up to 3 h at 37 °C. Reactions were terminated by the addition of 25 μL of 25 mg/mL BSA and 2 mL of acetone, which precipitated both the protein and DNA, leaving in solution only the free modified bases which had been enzymatically cleaved from the DNA backbone. After centrifugation at 8000*g* for 15 min the supernatant was dried, resuspended in water, and analyzed by liquid scintillation counting.

At each step the chemical identity of the released radioactive product was proven to be cytosine by HPLC. The free cytosine hydrate released by the enzyme is unstable, rapidly eliminating water, and is recovered as free cytosine (Boorstein et al., 1989). One unit of enzyme released 1 pmol of cytosine hydrate from 0.1 μg of UV-irradiated poly(dG-[^{3}H]-dC) in 1 min. Enzyme assays lasted from 15 min to 3 h, depending upon the specific activity of the enzyme during the different phases of the purification.

[^{3}H]Thymine glycol released from the oxidized poly(dA-[^{3}H]dT) was identified by HPLC as previously described (Higgins et al., 1987).

AP Nicking Assay. AP-site containing DNA was prepared and nicking activity assayed as described previously (Cunningham & Weiss, 1985). The assay is done in 10 mM EDTA to preclude any Mg^{2+}-dependent AP endonuclease from acting on the substrate.

Preparation of Thymine Glycol-Containing Oligodeoxynucleotide for Cross Linking Studies. Thymine glycol-containing single-stranded oligodeoxynucleotide was prepared as described previously (Kao et al., 1993). The oxidation was carried out on 50 OD_{260} of d(CGCGAT-ACGCC). The complementary 11-mer was synthesized by conventional means.

Cross Linking of Enzyme to Oligodeoxynucleotide. Twenty picomoles of the appropriate oligodeoxynucleotide, either thymine glycol-containing or complementary, was 5′-end-labeled using T4 kinase (Gibco BRL) and [γ-^{32}P]ATP, according to the manufacturer's recommendations, and purified using a Nuc Trap Push Column (Stratagene) preequilibrated in 20 mM HEPES, pH 7.5, 50 mM NaCl, and 5 mM EDTA. The radiolabeled oligodeoxynucleotide was then combined with 200 pmol of nonradioactive oligodeoxynucleotide, and the complementary strand was added at a 1:1 ratio and placed on ice for 30 min. Enzyme was reacted with the substrate double-stranded oligodeoxynucleotide in a total volume of 300 μL under the following reaction conditions: 37.3 mM $NaCNBH_3$, 20 mM HEPES, pH 7.5, 46.5 mM KCl, 5 mM EDTA, 1.5 uM oligodeoxynucleotide, and 15 ng/μL protein. In the case of *E. coli* endonuclease III, this represented a 4-fold molar excess of substrate oligodeoxynucleotide to enzyme. After incubation at 37 °C for 2 h, samples were quick frozen on dry ice, lyophilized, resuspended, and boiled in 35 μL of 1× SDS−PAGE loading buffer and separated by electrophoresis on a 15% Tricine-SDS gel. Following electrophoresis, the gel was stained with Coomassie Blue, wrapped in plastic, and analyzed via phosphorimaging.

Gel Electrophoresis. All samples were lyophilized to dryness and resuspended in standard SDS loading buffer prior to electrophoresis. Fifteen percent Tricine gels were prepared (Shagger et al., 1987) and run using the Mini-Protein II electrophoresis system (Bio-Rad). Gels were run at 90 V for approximately 5 h, completion being determined by the progress of prestained low molecular weight electrophoresis standards (Bio-Rad). Gels were then stained with Coomassie Blue.

Amino Acid Sequence Analysis. Fractions from the ssDNA cellulose column (fraction V) were run on a 15% Tricine-SDS gel and stained with Coomassie Blue. The predominant band, identical to the band which shifted after reductive coupling to the thymine glycol-containing oligodeoxynucleotide, was excised from the gel and sent to the W. M. Keck Foundation microsequencing facility at Yale University, New Haven, CT.

At Yale, the protein was subjected to proteolytic digestion followed by purification on HPLC using a reverse-phase

Table 1: Summary of Purification of a Pyrimidine Hydrate DNA-Glycosylase from Calf Thymus[a]

fraction	total protein (mg)	volume (mL)	total activity (pmol/min)	specific activity [pmol/(min·mg)]	purification (fold)	yield (%)
I	52000	4000	2920	0.056		
II	2420	720	1020	0.421	7.5	34.9
III	264	350	370	1.40	25	12.7
IV	1.3	3.0	36	27.7	495	1.2
V	0.1	1.5	29	290	5180	1.0

[a] Purification steps and fractions are described in the text.

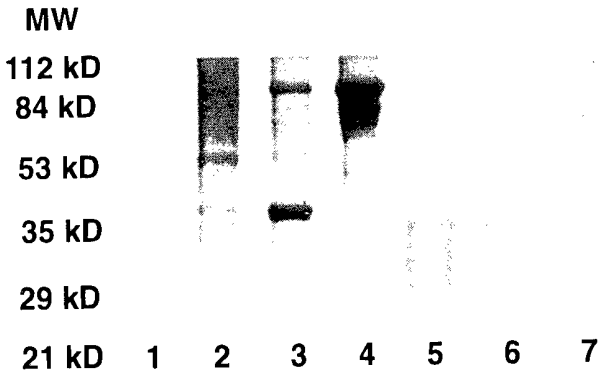


FIGURE 1: SDS−PAGE analysis of the purification fractions. Lanes 1 and 7 contain molecular weight markers. Lanes 2−5 contain fractions I−IV, respectively (described in the text and Table 1). Lane 6 contains material from ssDNA cellulose column fraction 17, which was then pooled with fractions 15 and 16 to yield fraction V.

microbore C18 column. Individual peaks were assayed for purity by laser desorption mass spectroscopy. After a 16 h hydrolysis, amino acid analysis was carried out on a Beckman Model 6300 ion-exchange instrument (Rosenfeld et al., 1992; Elliott et al., 1993; Williams & Stone, 1995; Williams et al., 1995).

The sequence homologies were obtained via the BLAST (Altschul et al., 1990) Network Service of the National Center for Biotechnology Information which accesses the Brookhaven, Swiss, PIR, and GenBank data bases.

RESULTS

Purification of the Mammalian Enzyme. A mammalian homologue of *E. coli* endonuclease III was purified from fresh calf thymus on the basis of its pyrimidine hydrate DNA-glycosylase activity. After the final purification step, ss-DNA-cellulose chromatography, the enzyme was purified approximately 5000-fold as estimated by the specific activity of the pyrimidine hydrate DNA-glycosylase and the yield was approximately 1% (Table 1 and Figure 1).

Coelution of DNA-Glycosylase Activities. Successive fractions from the ssDNA-cellulose column were assayed simultaneously for pyrimidine hydrate and thymine glycol DNA-glycosylase activities, both of which have been demonstrated for endonuclease III (Higgins et al., 1987; Boorstein et al., 1989). Figure 2A documents the coelution of the two activities.

Coelution of DNA-Glycosylase and Mg^{2+}-Independent AP Site Nicking Activity. Comparable ssDNA-cellulose purified material from another calf thymus preparation was assayed simultaneously for Mg^{2+}-independent AP-nicking activity and pyrimidine hydrate DNA-glycosylase activity, both of which are also previously documented activities of *E. coli* endonuclease III (Cunningham & Weiss, 1985). The co-elution of these two activities is shown in Figure 2B.

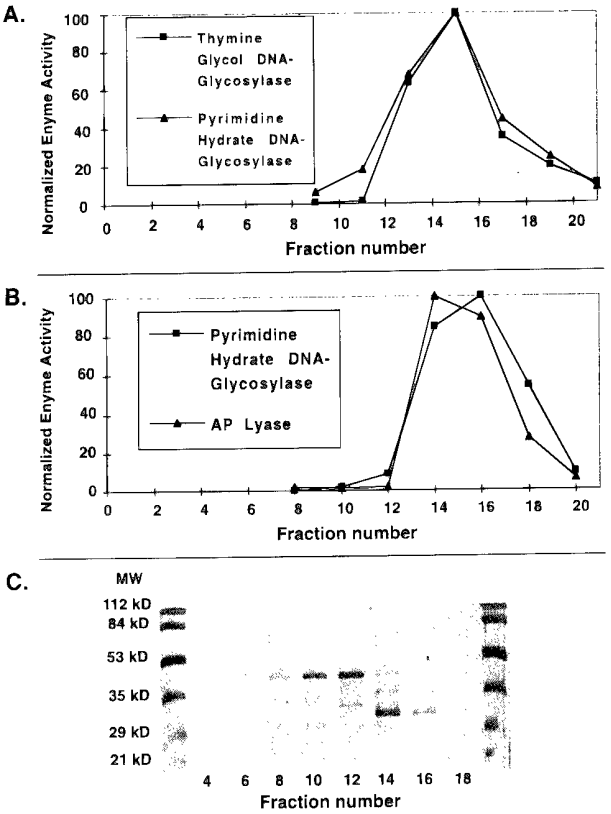


FIGURE 2: (A) Coelution of pyrimidine hydrate and thymine glycol DNA-glycosylase activities from the ssDNA cellulose column. Fractions from the ssDNA cellulose column were assayed simultaneously for both enzyme activities. Activities were normalized by dividing the activity of each fraction by the activity of the fraction with maximum activity. (B) Coelution of pyrimidine hydrate DNA glycosylase activity and AP lyase activity from the ssDNA cellulose column. A second calf thymus preparation was purified through ssDNA cellulose, and elution fractions were analyzed for both enzyme activities, normalized as in Figure 2A. (C) SDS−PAGE analysis of ssDNA cellulose elution fractions (fraction V). A 25 μL aliquot from each of the indicated fractions shown in panel A was analyzed by SDS−PAGE. Fractions 14−18 contain the predominant 31 kDa species. The extreme left and right lanes contain molecular weight markers.

Estimation of the Molecular Weight of the Mammalian Enzyme. The molecular radius of the mammalian DNA-glycosylase, as determined by gel filtration, was approximately 29 kDa. Although ssDNA-cellulose fractions with peak enzymatic activity contained more than one protein species, a predominant band of apparent molecular mass of 31 kDa was present on SDS−PAGE analysis. Moreover, when 25 μL aliquots of successive ssDNA-cellulose column fractions were subjected to electrophoresis and stained with Coomassie Blue, the elution profile of this predominant 31 kDa species, as judged by the intensity of staining, corresponded to that of the two DNA-glycosylase activities (Figure 2C).

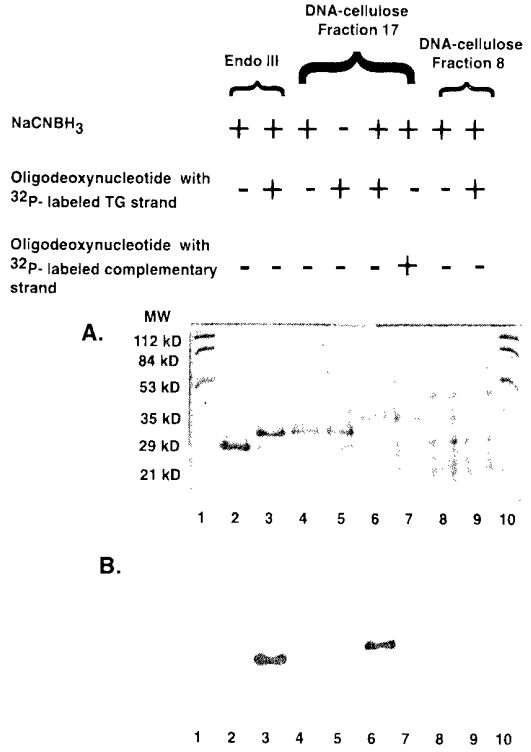


FIGURE 3: (A) SDS−PAGE analysis of *E. coli* endonuclease III and the bovine enzyme after incubation with the thymine glycol-containing oligodeoxynucleotide and $NaCNBH_3$. Lanes 1 and 10 contain molecular weight markers. Lane 2 contains the product of the reaction of *E. coli* endonuclease III and $NaCNBH_3$. Lane 3 contains the product of the same reaction mixture as lane 2 with addition of duplex 5′-^{32}P-labeled oligodeoxynucleotide containing a single thymine glycol (TG) residue. Lane 4 contains the product of fraction 17 eluted from the ssDNA cellulose column and incubated with $NaCNBH_3$ but no oligodeoxynucleotide. Lane 5 contains the product of elution fraction 17 incubated with the 5′-^{32}P-oligodeoxynucleotide but no $NaCNBH_3$. Lane 6 contains the product of elution fraction 17 incubated with both the oligodeoxynucleotide and $NaCNBH_3$. Lane 7 is the same mixture as 6 except that the complementary (non-thymine glycol-containing) oligodeoxynucleotide was 5′-labeled with ^{32}P. Lanes 8 and 9 contain the products of the incubation of ssDNA-cellulose fraction 8, which did not exhibit enzymatic activity, alone or with oligodeoxynucleotide and $NaCNBH_3$, respectively. (B) Phosphorimage of the SDS−PAGE gel of panel A. The lanes are identical to those of panel A.

Reductive Cross Linking of the Enzymes to a Thymine Glycol-Containing DNA Oligodeoxynucleotide. Incubation of purified *E. coli* endonuclease III with duplex DNA (the thymine glycol-containing oligodeoxynucleotide annealed to its complementary strand) in the presence of $NaCNBH_3$ resulted in an increase in the apparent molecular mass of the enzyme as determined by SDS−PAGE (Figure 3A). Lane 2 demonstrates endonuclease III incubated with substrate DNA in the absence of $NaCNBH_3$ and lane 3 in the presence of $NaCNBH_3$. The increase in the apparent molecular mass of the endonuclease III is the result of irreversible cross linking of the enzyme to the oligodeoxynucleotide.

The reductive cross linking reaction was also performed on the most purified preparation of the calf thymus pyrimidine hydrate DNA-glycosylase. A 75 μL aliquot of fraction 17 eluted from the ssDNA cellulose column (Figure 2A) containing purified enzyme of maximal specific activity was incubated with the thymine glycol-containing oligodeoxynucleotide in the presence of $NaCNBH_3$ along with appropriate controls. Lanes 4 and 5 represent ssDNA fraction 17 incubated with $NaCNBH_3$ and no substrate DNA and substrate DNA in the absence of $NaCNBH_3$ respectively. The apparent molecular mass of 31 kDa, as first shown in Figure 2C, did not change under either of these incubations. However, when the reaction mixture contained both substrate DNA and $NaCNBH_3$, the predominant 31 kDa Coomassie Blue-stained band shifted to an apparent molecular mass of 35 kDa, as shown in lanes 6 and 7. As an additional control, fraction 8 eluting from the ssDNA cellulose column (Figure 2A), which contained no enzymatic activity, was also exposed to the conditions of the coupling reaction. Lanes 8 and 9 of Figure 3A contain protein from this fraction incubated with the oligodeoxynucleotide in the presence or absence of $NaCNBH_3$. As can readily be seen, no shift of the visible protein bands occurred.

Additional proof of irreversible cross linking is demonstrated in Figure 3B. The thymine glycol-containing oligodeoxynucleotide had been 5′-end-labeled with ^{32}P prior to the coupling reactions. A phosphorimage of the gel in Figure 3A demonstrated only two bands. Figure 3B shows a single band in lane 3 which corresponds to the position of the shifted cross linked endonuclease III. The single band in lane 6 corresponds to the position of the predominant Coomassie-Blue stained species from calf thymus, which had also shifted after cross linking. Under the denaturing conditions (boiling) used to prepare samples for the SDS gel, the complementary oligodeoxynucleotide strand does not remain associated with the protein. When the complementary, rather then the thymine glycol-containing oligodeoxynucleotide was ^{32}P-labeled, the protein shifted (Figure 3A, lane 7) but did not appear on the phosphorimage (Figure 3B, lane 7). Thus, lanes 6 and 7 of Figure 3B prove that the bovine enzyme cross linked only to the oligodeoxynucleotide containing the thymine glycol residue, thereby confirming that the irreversible cross linking resulting from chemical reduction is exclusively dependent upon formation of a specific ES intermediate. There was no evidence of any binding of oligodeoxynucleotide to the proteins which did not contain enzyme activity, further corroborating that the reductive cross linking reaction was absolutely specific.

Amino Acid Sequence Data. Four peptides derived from a proteolytic digest of the purified bovine protein were sequenced yielding sequences of 14, 15, 22, and 23 amino acids. None of these sequences demonstrated direct similarity to *E. coli* endonuclease III by initial BLAST analysis. However, the 22 amino acid peptide sequence demonstrated considerable similarity to a portion of two predicted full length protein sequences from *Caenorhabditis elegans* (accession no. Z05874) (Wilson et al., 1994) with $P(N) = 0.00053$ and *Saccharomyces cerevesiae* (accession no. L05146) with $P(N) = 0.0063$). Both the *C. elegans* and *S. cerevesiae* proteins, in turn, bear similarity to *E. coli* endonuclease III (accession no. J02857). When compared with the sequence of endonuclease III via BLAST, the *C. elegans* and the *S. cerevesiae* sequence yielded $P(N)$ values of 9.1×10^{-25} and 1.9×10^{-7}, respectively. This same bovine polypeptide demonstrates an even greater degree of similarity to two recently submitted partial 3′ cDNA sequences, from *Homo sapiens* (accession no. F04657) with $P(N) = 6.8 \times 10^{-9}$ and *Rattus sp.* (accession no. H33255) with $P(N) = 1.8 \times 10^{-7}$.

Figure 4 demonstrates the alignment of the *E. coli* endonuclease III amino acid sequence, with the primary

```
A:1 MNKAKRLEILTRLRENNPHPTTELNFSSPFELLIAVLLSAQATDVSVNKATAKLYPVANTPAAMLELGVE 70
B:1  MRKDMIAPVDTMGCHKLADPLAAPPVHRFQVLVALMLSSQTRDEVNAAAMKRLKDHGLSIGKILEFKVP 69
C:          PVDQLGAEHCFDPSA                                          LTVDSILQTDDS

A:71 GVKTYIKTIGLYNSKAENIIKTCRILLEQHNGEVPEDRAALEALPGVGRKTANVVLNTAFGWPT.IAV 137
B:70 DLETILCPVGFYKRKAVYLQKTAKILKDDFSGDIPDSLDGLCALPGVGPKMANLVMQIAWGECVGIAV 137
C:   TLGALIVPVGF                                                  QGTVNGIAV

A:138 DTHIFRVCNRTQFAPGKN.VEQVEEKLLKVVPAEFKVDCHHWLILHGRYTCIARKPRCGSCIIEDLC 203
B:138 DTHVHRIS.NRLGWIKTSTPEKTQKALEILLPKSEWQPINHLLVGFGQMQCQPVRPKCGTCLCRFTC 203
C:    XTHVP                              *LWSEINGLLVGFGQQTCLPIRP
D:                                     WLPR?LWHEINGLLVGFGQQTCLPVHPRCHACLNQALC
E:       HRIANRLKWTKKMTKSPEETRRNLE?WLPRVLWSEINGLLVGFGQ?ICLPVHPRCQACL?KALC

A:204 EYKEKVDI 211
B:204 PSSTAKNVKSETEETSTSIEVKQEVEDEFEDEKPAKKIKKTRKTRTKIEVKTESET 259
C:
D:    PAAQGL
E:    PAAQGL
```

FIGURE 4: Amino acid sequence alignment of *E. coli* endonuclease III (line A), *C. elegans* translated protein (line B), bovine primary amino acid sequences for peptides of 15, 23, 14, and 22 amino acids, respectively (line C), and *H. sapiens* and *Rattus sp.* sequences obtained by translation of partial cDNA sequence (lines D and E, respectively). X in sequence C represents an indeterminate amino acid residue. ? in sequences D and E represents indeterminate nucleotide sequences. The six amino acid region presented in boldface and italic constitutes a portion of the active site of endonuclease III. A 22 amino acid region of near identity between the predicted *C. elegans* sequence, the primary bovine sequence, and the *H. sapiens* and *Rattus sp.* translated partial cDNA sequences is presented in boldface. The four cysteine residues presented in double underlined type represent the ligands of the iron−sulfur cluster of *E. coli* endonuclease III.

amino acid sequences of the bovine polypeptides, and the predicted amino acid sequences of the *C. elegans*, *H. sapiens*, and *Rattus sp.* proteins derived by translation of their respective nucleotide sequences. The boldface sequence marked with the asterisk represents the 22 amino acid bovine polypeptide found to be most similar to the *C. elegans* and the *H. sapiens* and *Rattus sp.* sequences, as determined by the BLAST program using the default Blosum 62 as the algorithmic matrix and the default expected cutoff value of 10. The other peptides were aligned by the BLAST program after we raised the expected cutoff value from 10 to 100.

DISCUSSION

We purified a mammalian pyrimidine hydrate DNA-glycosylase 5000-fold from calf thymus (summarized in Table 1). The most purified fractions, after elution from ssDNA cellulose, also demonstrated thymine glycol DNA-glycosylase and AP lyase activities. The AP nicking assay, performed in the presence of 10 mM EDTA, has previously been shown to specifically correspond to β-elimination (Mazumder et al., 1991). The elution profiles of the two glycosylase activities and the AP lyase activities were superimposable (Figure 2A,B). The fact that these activities coeluted in the most purified calf thymus fractions strongly suggests that they are contained within the same protein.

When identical volumes of successive column fractions were analyzed by SDS−PAGE, there was strong correspondence between the intensity of staining of a predominant 31 kDa band in the active fractions and the enzyme activities, but several other protein species were also present which could have represented the bovine pyrimidine hydrate-thymine glycol DNA-glycosylase/AP lyase.

Therefore, to definitively identify the bovine enzyme, we took advantage of a reductive cross linking reaction which had already been applied to T4 endonuclease V and the *E. coli* Fpg protein. We first demonstrated that, in the presence of $NaCNBH_3$, purified *E. coli* endonuclease III would form a stable cross link to an oligodeoxynucleotide containing one of its substrates, thymine glycol. The apparent increase in molecular weight of the purified enzyme (Figure 3A, lane 3) together with the phosphorimaging data (Figure 3B, lane 3) demonstrated unequivocally that the bacterial enzyme was irreversibly cross linked to the substrate oligodeoxynucleotide. Thus, *E. coli* endonuclease III was cross linked to a substrate DNA oligodeoxynucleotide in a manner analogous to T4 endonuclease V and the *E. coli* Fpg protein, confirming that it also functions via an *N*-acylimine ES intermediate.

We applied the same reaction to the most purified bovine enzyme fraction and showed that only the predominant 31 kDa protein species was irreversibly cross linked to the same thymine glycol-containing oligodeoxynucleotide. The specificity of the reaction was confirmed by separately 5′-end-labeling either the thymine glycol-containing or complementary strand of the substrate double-stranded oligodeoxynucleotide. The increase in the apparent molecular mass of the 31 kDa protein occurred independently of which DNA strand was labeled; however, only when the thymine glycol-containing oligodeoxynucleotide was labeled did a band corresponding to the shifted protein appear on phosphorimage of the gel (Figure 3B). That this strategy enabled us to successfully identify the bovine analog of *E. coli* endonuclease III in a relatively complex mixture of mammalian proteins was contingent on the fact, unproven until now, that the mammalian enzyme and the bacterial enzyme both function through a *N*-acylimine ES intermediate.

The primary amino acid sequence data confirms that the purified 31 kDa protein species we identified by reductive cross linking is a mammalian homologue of endonuclease III. The aligned sequences of Figure 4 demonstrate the homology between the bovine and *C. elegans* proteins extending into the region which constitutes the iron−sulfur cluster of *E. coli* endonuclease III (Thayer et al., 1995). This iron−sulfur cluster motif contains four cysteine residues at endonuclease III positions 187, 194, 197, and 203 and has been shown to be a DNA binding domain. The *H. sapiens* and *Rattus sp.* partial 3′ cDNA sequences also contain four cysteine residues which align with those of *E. coli* endonuclease III and the *C. elegans* sequence. Thus, it seems probable that the *E. coli*, *C. elegans*, and mammalian enzymes all share a common mode of DNA binding. A second bovine peptide and the *C. elegans* predicted protein both align with a region containing a known active site amino

acid of endonuclease III, aspartic acid 138 (Figure 4, bold, italics). Another critical active site residue in *E. coli* endonuclease III is lysine 120 which probably contributes the ϵ-amino group necessary for the formation of the *N*-acylimine ES intermediate (Thayer et al., 1995). Since we have demonstrated such an ES intermediate for the bovine enzyme, it is probable that mammalian DNA glycosylase/AP lyases will prove to have a lysine residue as part of their active sites. In conclusion, given the similarities in amino acid sequence, including active sites and DNA binding domains among the *E. coli* endonuclease III, the purified bovine enzyme and the predicted sequences of the *C. elegans*, *H. sapiens*, and *Rattus sp.* proteins, we suggest that there is a homologous family of endonuclease III-like DNA repair enzymes present throughout phylogeny.

ACKNOWLEDGMENT

We thank Mr. Archie Cummings for invaluable technical assistance during the purification of the bovine enzyme. We also thank Maria Thayer for helpful discussions on protein similarities. We thank Ken Williams and Kathy Stone of the W. M. Keck Foundation Biotechnology Resource Laboratory at Yale University for carrying out the protein sequencing and for helpful discussion.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Meyers, E. W., & Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
Asahara, H., Wistort, P. M., Bank, J. F., Bakerian, R. H., & Cunningham, R. P. (1989) *Biochemistry 28*, 4444−4449.
Bacchetti, S., & Benne, R. (1975) *Biochim. Biophys. Acta 390*, 285−297.
Bailly, V., & Verly, W. G. (1987) *Biochemical J. 242*, 565−572.
Bailly, V., & Verly, W. G. (1989) *Nucleic Acids Res. 17*, 3617−3618.
Boorstein, R. J., Hilbert, T. P., Cadet, J., Cunningham, R. P., & Teebor, G. W. (1989) *Biochemistry 28*, 6164−6170.
Breimer, L. H., & Lindahl, T. (1984) *J. Biol. Chem. 259*, 5543−5548.
Brent, T. P. (1973) *Biophys. J. 13*, 399−401.
Cunningham, R. P., & Weiss, B. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 474−478.
Demple, B., & Linn, S. (1980) *Nature 287*, 203−208.
Demple, B., & Harrison, L. (1994) *Annu. Rev. Biochem. 63*, 915−948.
Dizdaroglu, M., Laval, J., & Boiteux, S. (1993) *Biochemistry 32*, 12105−12111.
Dodson, M. L., Schrock, R. D., III, & Lloyd, R. S. (1993) *Biochemistry 32*, 8284−8290.
Dodson, M. L., Michaels, M. L., & Lloyd, R. S. (1994) *J. Biol. Chem. 269*, 32709−32712.
Doetsch P. W., Helland, D. E., & Haseltine, W. A. (1986) *Biochemistry 25*, 2212−2220.
Duker, N. J., & Teebor, G. W. (1975) *Nature 255*, 82−84.
Elliott, J. I., Stone, K. L., & Williams, K. R. (1993) *Anal. Biochem. 211*, 94−101.
Higgins, S. A., Frenkel, K., Cummings, A., & Teebor, G. W. (1987) *Biochemistry 26*, 1683−1688.
Huq, I., Haukanes, B. I., & Helland, D. E. (1992) *Eur. J. Biochem. 206*, 833−839.
Kao, J. Y., Goljer, I., Phan, T. A., & Bolton, P. H. (1993) *J. Biol. Chem. 268*, 17787−17793.
Kim, J., & Linn, S. (1989) *J. Biol. Chem. 264*, 2739−2745.
Mazumder, A., Gerlt, J. A., Absolon, J. M., Stubbe, J., Cunningham, R. P., Withka, J., & Bolton, P. H. (1991) *Biochemistry 30*, 1119−1126.
Ness, I. F., & Nissen-Meyer, J. (1978) *Biochim. Biophys. Acta 520*, 111−121.
Radman, M. (1976) *J. Biol. Chem. 251*, 1438−1445.
Rosenfeld, J., Capdeville, J., Guillemot, J., & Ferrara, P. (1992) *Anal. Biochem. 203*, 173−179.
Shagger, H., & von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
Strniste, G. F., & Wallace, S. S. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 1997−2001.
Tchou, J., & Grollman, A. P. (1995) *J. Biol. Chem. 270*, 11671−11677.
Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P., & Tainer, J. A. (1995) *EMBO J. 14*, 4108−4120.
Williams, K. R., & Stone, K. L. (1995) in *Techniques in Protein Chemistry VI* (Crabb, J., Ed.) pp 143−152, Academic Press, New York.
Williams, K. R., Samander, S., Stone, K. L., Saylor, M., & Rush, J. (1995) in *Protein Protocol Handbook* (Walker, J., Ed.) Humana Press.
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., et al. (1994) *Nature 368*, 32−38.

BI952516E

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 10, Issue of March 7, pp. 6733–6740, 1997
Printed in U.S.A.

Cloning and Expression of the cDNA Encoding the Human Homologue of the DNA Repair Enzyme, *Escherichia coli* Endonuclease III*

(Received for publication, September 27, 1996, and in revised form, December 17, 1996)

Timothy P. Hilbert‡, Wenren Chaung‡, Robert J. Boorstein‡, Richard P. Cunningham§, and George W. Teebor‡¶

From the ‡Department of Pathology and Kaplan Comprehensive Cancer Center, New York University Medical Center, New York, New York 10016 and the §Department of Biological Sciences, SUNY, Albany, New York 12222

We previously purified a bovine pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase. The amino acid sequence of tryptic bovine peptides was homologous to *Escherichia coli* endonuclease III, theoretical proteins of *Saccharomyces cerevisiae* and *Caenorhabditis elegans*, and the translated sequences of rat and human 3′-expressed sequence tags (3′-ESTs) (Hilbert, T. P., Boorstein, R. J., Kung, H. C., Bolton, P. H., Xing, D., Cunningham, R. P., Teebor, G. W. (1996) *Biochemistry* 35, 2505–2511). Now the human 3′-EST was used to isolate the cDNA clone encoding the human enzyme, which, when expressed as a GST-fusion protein, demonstrated thymine glycol-DNA glycosylase activity and, after incubation with $NaCNBH_3$, became irreversibly cross-linked to a thymine glycol-containing oligodeoxynucleotide, a reaction characteristic of DNA glycosylase/AP lyases. Amino acids within the active site, DNA binding domains, and [4Fe-4S] cluster of endonuclease III are conserved in the human enzyme. The gene for the human enzyme was localized to chromosome 16p13.2-.3. Genomic sequences encoding putative endonuclease III homologues are present in bacteria, archeons, and eukaryotes. The ubiquitous distribution of endonuclease III-like proteins suggests that the 5,6-double bond of pyrimidines is subject to oxidation, reduction, and/or hydration in the DNA of organisms of all biologic domains and that the resulting modified pyrimidines are deleterious to the organism.

When a pyrimidine residue in cellular DNA becomes modified by oxidation, reduction, or hydration of its 5,6-double bond, repair is initiated by a DNA-glycosylase activity that cleaves the *N*-glycosyl bond of the damaged residue, releasing the modified base and creating an abasic (AP) site in the DNA backbone. Such DNA glycosylase activities have been identified in bacteria, yeast, and mammalian species (1–8) The first such enzyme described was *Escherichia coli* endonuclease III, which was identified not on the basis of its DNA glycosylase activity, but rather because it nicked UV-irradiated DNA (9). For this reason it was termed an endonuclease, because it was thought that nicking resulted from enzyme-catalyzed hydrolysis of internucleotide phosphodiester bonds at sites of DNA damage. It has since been determined that the enzyme nicks DNA not via hydrolysis, but by catalyzing β-elimination of the 3′-phosphate group at the AP site formed as a result of the enzyme's DNA glycosylase activity (10–12). The modified base that was enzymatically released from UV-irradiated DNA proved to be cytosine and/or uracil hydrate (8). Enzymes that effect base release together with strand cleavage via β-elimination are now termed DNA glycosylase/AP lyases and, in addition to endonuclease III, include the Fpg protein of *E. coli* (13), the OGG1 protein of *Saccharomyces cerevisiae* (14, 15), and T4 endonuclease V (16).

DNA glycosylase/AP lyases function through *N*-acylimine (Schiff's base) enzyme-substrate intermediates (17). Such enzyme-substrate intermediates can be chemically reduced to stable secondary amines, resulting in irreversible cross-linking of the enzymes to their particular substrates (13, 16–18). We previously used this cross-linking reaction to definitively identify a pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase purified from calf thymus. Incubation, done under reducing conditions, of a ^{32}P-labeled oligodeoxynucleotide containing a single thymine glycol (5,6-dihydroxy-5,6-dihydrothymine) residue with a 5000-fold purified enzyme preparation resulted in cross-linking of a predominant 31-kDa protein to the oligodeoxynucleotide as determined by SDS-PAGE[1] analysis and phosphor imaging. Tryptic digestion of this protein, followed by microsequencing of several of the resulting peptides demonstrated that the bovine enzyme was homologous to theoretical proteins translated from the genomic DNA of *S. cerevisiae* and *Caenorhabditis elegans*. Both of these theoretical proteins, in turn, were homologues of *E. coli* endonuclease III. The bovine peptide amino acid sequences were also homologous to the translated sequences of 3′-ESTs from *H. sapiens* brain tissue (accession number F04657) and *Rattus sp.* PC 12 cells (accession number H33255) (18). In the current study, we used probes based upon the homologous human 3′-EST, to isolate clones that encode the human homologue of *E. coli* endonuclease III from a splenic cDNA library. Once determined, the cDNA sequence was used to express the enzyme as a functional

* This work was supported by National Institutes of Health Grants CA 16669 and 49689 (to G. W. T.), CA 16087 (to Kaplan Cancer Center), GM 07308 (Medical Scientist Training Program for T. P. H.), U. S. Army Medical Research and Development Command DAMD 17-94-J-4416 (to R. J. B.), and GM 46312 (to R. P. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U81285.

¶ To whom correspondence should be addressed: Dept. of Pathology, New York University Medical Center, 550 First Ave., New York, NY 10016. Tel.: 212-263-5473; Fax: 212-263-8211; E-mail: george.teebor@mcccm.med.nyu.edu.

[1] The abbreviations used are: PAGE, polyacrylamide gel electrophoresis; EST, expressed sequence tag; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; IPTG, isopropyl-1-thio-β-D-galactopyranoside; GST, glutathione *S*-transferase; ORF, open reading frame; HhH, helix-hairpin-helix; FISH, fluorescence *in situ* hybridization.

recombinant protein and to determine the chromosomal localization of the human gene.

EXPERIMENTAL PROCEDURES

Radionucleotides—[α-^{32}P]dCTP (3000 Ci/mmol), [γ-^{32}P]dATP (3000 Ci/mmol), and [*methyl*-^{3}H]TTP (70–90 Ci/mmol) were purchased from DuPont NEN.

Cloning of the cDNA—Oligodeoxynucleotides based upon the human 3′-EST sequence (accession number F04657) were used to isolate homologous clones from a Superscript human spleen cDNA library in the pCMV-SPORT plasmid vector (Life Technologies, Inc.) using the GENETRAPPER cDNA positive selection system (Life Technologies), according to the manufacturer's protocol. Briefly, the amplified double-stranded cDNA library was made single-stranded by treatment with the Gene II product (phage F1) endonuclease and *E. coli* exonuclease III and then hybridized to a biotinylated sense strand-specific oligodeoxynucleotide, P1 (5′-GTGGCACGAGATCAATGGACTCTTG). The cDNA-oligodeoxynucleotide hybrids were captured using streptavidin paramagnetic beads. Nonspecifically bound cDNAs were washed away at high stringency, and specifically bound cDNAs were eluted from the paramagnetic beads by denaturing the cDNA-oligodeoxynucleotide hybrids. Selected cDNA clones were then made double-stranded via repair, which was primed by a second sequence-specific oligodeoxynucleotide, P2 (5′-ATCATTGGACTCTGGGTGGGC). The selected repaired plasmids were electroporated into the *E. coli* strain DH5α and plated onto Lennox L agar plates containing 50 μg/ml ampicillin (LB/amp agar).

After 20 h of incubation at 37 °C, colonies were analyzed for the presence of the desired cDNA insert via colony PCR, according to the manufacturer's protocol, using a second set of 3′-EST-specific primers (P3, 5′-CAACAGGCGTGGCTTCCTGAAGCG; P4, 5′-GGTGGGCTTCGGCCAGCAGACCTGT) to maximize specificity of the selection procedure. PCR was conducted as follows: 1 cycle of 95 °C for 2 min and 37 cycles of 94 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min, followed by a final cycle of 10 min at 72 °C. PCR products were then analyzed by electrophoresis in a 1.2% agarose gel. Colonies that proved positive through the first PCR, by virtue of the production of a 180-base pair product, were subjected to a second round of colony PCR in order to determine the size of the inserts using T7 and SP6-specific primers (5′-TAATACGACTCACTACTATAGGAGA and 5′-AGCTATTTAGGTGACACTATAG, respectively). Of the 23 colonies obtained, 10 proved, through colony PCR and sequencing analysis, to contain the sequence of interest.

Isolation of Longer cDNA Clones via a Second GENETRAPPER Selection—In order to isolate additional cDNA clones that contained long inserts and thus had a higher probability of containing the full-length cDNA sequence, the GENETRAPPER cDNA selection system was used a second time, substituting a second set of oligodeoxynucleotides for capture (P5, 5′-ACAGAGACTGCGTGTGGCCTATGAG) and repair (P6, 5′-AAGAGAGCCTGCAGCAGAAGC) of the selected clones. These primers were based not upon the human 3′-EST sequence but were specific for the 3′ portion of previously sequenced cDNA inserts and therefore were specific for the 5′ portion of the mRNA. Colonies were again screened, and insert size was determined by PCR as described above. However, rather then using the T7 primer, an additional sequence-specific primer, P7 (5′-CACCTTGCTCCAGAAACC), was used as a primer in PCR with the SP6 primer to determine the size of the plasmid inserts. PCR-positive colonies that contained the largest inserts were sequenced.

5′-RACE Analysis—Additionally, to confirm the sequence of the 5′-terminus of the mRNA, the 5′-RACE System (Life Technologies) was used to amplify the 5′-terminus of the message for sequencing. The manufacturer's protocol for GC-rich cDNAs was followed. Briefly, 2.5 pmol of a gene-specific primer P8 (5′-CATCAGTGACAGCAGCACCT) was hybridized to 100 ng of human spleen poly(A)$^+$ RNA (Clontech) and cDNA was synthesized using Superscript II Reverse Transcriptase (Life Technologies). The RNA was then degraded with RNase, and the cDNA was isolated. A poly(dC) tail was then added to the 3′-terminus of the purified cDNA using dCTP and TdT, and the cDNA region corresponding to the 5′-end of the mRNA was amplified by two successive rounds of PCR using additional gene-specific primers P9 (5′-CATAGGCCACACGCAGTCTC) and P10 (5′-CTTCTGCTGCAGCCTCTCTTC), together with the anchor primers supplied by the manufacturer.

The second round of PCR yielded a single amplified product that, when analyzed by electrophoresis on a 1.2% agarose gel, corresponded in size to what was expected on the basis of the longest GENETRAPPER-isolated cDNA sequences. The PCR product was gel-purified and cloned into the pCR II cloning vector (Invitrogen) using the TA cloning kit (Invitrogen), electroporated into the *E. coli* strain DH5α, and plated onto LB/amp agar plates. Colonies were used to inoculate Lennox L broth cultures containing 50 μg/ml ampicillin (LB/amp broth), and the inserts of 10 isolated plasmids were sequenced.

DNA Sequencing—Plasmid DNA was purified for sequencing using the QIAprep Spin Plasmid Miniprep kit (QIAGEN) from 5 ml of LB/amp broth cultures, containing 50 μg/ml ampicillin incubated for 16 h at 37 °C. DNA sequencing was carried out by the New York University Kaplan Cancer Center sequencing facility, using a model 373 automated DNA sequencer (ABI), and model 800 Lab Station (ABI).

Construction of a GST Fusion Protein in pGEX-2T—The DNA sequence encoding amino acids (8–304) of the open reading frame (Fig. 1) were amplified via PCR from 50 ng of the purified cDNA containing plasmid via PCR using the following primers: P11 (5′-CTTGGATCCATGCTGACCCGGAGCCGGAGC) and P12 (5′-CTCGAATTCGAGCCATGCGGCCCTCCGAGA). These primers were designed to incorporate *Bam*HI and *Eco*RI restriction sites into the 5′- and 3′-ends of the sense strand, respectively. PCR was conducted as follows: 1 cycle of 95 °C for 2 min and 35 cycles of 94 °C for 1 min, 65 °C for 1 min, and 72 °C for 2 min, followed by a final cycle of 10 min at 72 °C. The resulting PCR product was digested with *Bam*HI and *Eco*RI, gel-purified, ligated into gel-purified pGEX-2T vector (Pharmacia Biotech Inc.) that had previously been digested with *Bam*HI and *Eco*RI, and electroporated into the *E. coli* strain NB42. Colonies were selected via growth on LB agar/amp plates, and the presence of the appropriate insert was verified via colony PCR as described above, using primers P3 and P4. Expression of the full-length fusion protein was confirmed via the induction of log phase (A_{590} = 0.6) 5-ml LB/amp broth cultures with 0.1 mM IPTG for 4 h at 37 °C. To prepare total cell SDS lysates, 1-ml aliquots of induced and uninduced cultures were centrifuged at 5000 × *g* for 2 min, the supernatant was discarded, and the pelleted bacteria were resuspended in 100 μl of SDS-PAGE loading buffer and heated at 95 °C for 5 min. Thirty μl of each sample was then analyzed on a 15% Tricine gel. After the gels were stained with Coomassie Blue, induced and uninduced samples were compared to demonstrate the expression of the full-length (65-kDa) fusion protein. Bacterial lysates produced in an identical manner were also run on the SDS-PAGE gel in Fig. 3 in order to demonstrate induction of the GST fusion protein.

Protein Expression and Purification—600 ml of LB/amp broth were inoculated with 10 ml of overnight cultures. Bacteria were grown at 37 °C until the A_{590} reached 0.6. Expression of the fusion protein was induced by incubation with 0.1 mM IPTG for 5 h at 30 °C (the lower temperature was used to increase the solubility of the fusion protein). Bacteria were then placed on ice for 1 h and pelleted by centrifugation at 3200 × *g* in 250-ml centrifuge tubes (Corning) for 10 min. The supernatant was discarded, and the pellet was resuspended in 20 ml of sonication buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 5 mM EDTA, 0.5% Triton X-100, 0.25 mM phenylmethylsulfonyl fluoride, 0.1 mg/ml aprotinin). The bacteria were transferred to a 30-ml Corex centrifuge tube and sonicated for 2 min at 70% power using a Heat Systems model W-375 sonicator equipped with a model 419 standard tapered microtip. The sonicate was then centrifuged for 15 min at 10,000 × *g*, and the supernatant was transferred to a 50-ml plastic centrifuge tube containing 1.2 ml of glutathione-agarose 4B affinity medium (volume of medium was measured as a slurry in 20% ethanol, as supplied by the manufacturer) prewashed with 2 × 40 ml of wash buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 5 mM EDTA, 0.5% Triton X-100). The sample was incubated on ice with agitation for 30 min to allow adsorption of the fusion protein. The affinity medium was then pelleted by centrifugation for 2 min at 950 × *g*. The supernatant was removed by pipetting, and the affinity medium was washed once with 20 ml of sonication buffer and 4 times with 40 ml of wash buffer by thorough resuspension of the beads in the appropriate buffer followed by centrifugation at 950 × *g* for 1 min. After the final wash, the affinity medium was resuspended in 1 ml of wash buffer, transferred to a 2-ml plastic tube, and centrifuged again at 950 × *g* for 1 min to pellet the beads. The supernatant was removed, and the beads were resuspended in 1 ml of glutathione-agarose elution buffer (100 mM Tris, pH 8.0, 500 mM NaCl, 2.5 mM EDTA, 0.1% Triton X-100, 20 mM glutathione (Sigma)) and incubated for 12 h on ice with agitation. Beads were then quickly pelleted by centrifugation at 950 × *g*, and the supernatant that contained the eluted fusion protein was transferred to a fresh tube. All purification procedures from sonication through elution of the fusion protein were carried out at 4 °C. The purification yielded 9.9 mg of fusion protein.

As a control, the 26-kDa glutathione *S*-transferase (GST) of *Schistosoma japonicum* was expressed from the pGEX-2T vector (without a fusion insert) in the bacterial strain NB42 according to the same pro-

cedure described for the fusion protein. Twelve mg of purified GST was purified from 600 ml of induced bacterial culture.

Purification of E. coli Endonuclease III—Endonuclease III was purified from *E. coli* strain UC6444 carrying the plasmid pHIT1 as described previously (19).

Spectrophotometry—Spectrophotometric measurements of proteins were made in elution buffer (100 mM Tris, pH 8.0, 500 mM NaCl, 2.5 mM EDTA, 0.1% Triton X-100, 20 mM glutathione) in a quartz cuvette. The optical absorption spectra of the GST fusion protein and the unfused GST protein were recorded between 200 and 700 nm using a Spectronic Genesystems 5 spectrophotometer (Milton Roy). In order to allow comparison of the absorption spectra of the purified GST fusion protein and purified GST (see Fig. 6), the purified proteins were diluted prior to analysis with glutathione-agarose elution buffer to the same absolute protein concentration (5.5 mg/ml).

FISH Analysis—FISH Analysis was performed by SeeDNA Biotech Inc. (Dept. of Biology, York University, Ontario, Canada). Lymphocytes isolated from human blood were cultured in α-minimal essential medium supplemented with 10% fetal calf serum and phytohemagglutinin at 37 °C for 68–72 h. The lymphocyte cultures were treated with BrdUrd (0.18 mg/ml; Sigma) to synchronize the cell population. The synchronized cells were washed 3 times with serum-free medium to release the block and recultured at 37 °C for 6 h in α-minimal essential medium with thymidine (2.5 mg/ml; Sigma). Cells were harvested and slides were made by using standard procedures, including hypotonic treatment, fixing, and air drying.

To produce a probe for FISH analysis, a 1.1-kilobase pair fragment containing the entire cDNA sequence was excised from an isolated cDNA clone using *Eco*RI and *Hin*dIII, purified, and labeled with biotin-14-dATP using the BioNick labeling kit (Life Technologies) (20). The procedure for FISH analysis was performed according to the previously reported procedures of Heng *et al.* (21, 22). Briefly, slides were baked at 55 °C for 1 h. After RNase treatment, the slides were denatured in 70% formamide, 2 × SSC for 2 min at 70 °C followed by dehydration with ethanol. Probes were denatured at 75 °C for 5 min in a hybridization solution containing 50% formamide, 10% dextran sulfate, and human *Cot*I-restricted DNA. Probes were loaded on the denatured chromosomal slides. After overnight hybridization, slides were washed and analyzed. FISH signals and the DAP1 banding pattern were recorded separately by taking photographs. Chromosomal localization was achieved by superimposing FISH signals with DAP1-banded chromosomes (22).

Northern Blot Analysis—Two μg of mRNA, isolated from 293T cells using the FastTrack 2.0 mRNA isolation system (Invitrogen), 1 μg of human spleen Poly(A)$^+$ RNA (Clontech), and 5 μg of 0.24–9.5-kilobase pair RNA ladder (Life Technologies) were electrophoresed on an 11 × 14-cm 1.0% agarose-formaldehyde gel. The gel was rinsed with deionized water, and RNA was transferred to a Nytran membrane (Schleicher & Schuell) using the Turboblotter rapid downward transfer system (Schleicher & Schuell), according to the manufacturer's specifications. Following transfer, the membrane was gently washed in 2 × SSC for 5 min, dried on a fresh sheet of filter paper, and baked at 80 °C for 1 h. The portion of the membrane that contained the molecular weight markers was cut away and stained by treatment with 5% acetic acid for 15 min and 0.5 M sodium acetate, pH 5.2, with 0.04% methylene blue for 10 min, followed by destaining with water. The baked filter was incubated in prehybridization solution (in 50% formamide, 3 × SSC, 0.1 M Tris, pH 7.4, 5 × Denhardt's solution) for 4 h at 42 °C, followed by hybridization overnight at 42 °C with 2 × 10^6 cpm of radiolabeled probe/ml of hybridization solution (50% formamide, 3 × SSC, 0.1 M Tris, pH 7.4, 5 × Denhardt's solution, 10% dextran sulfate). Following hybridization, the membrane was washed three times for 30 min at 50 °C, successively with 1 × SSC, 0.1% SDS; 0.5 × SSC, 0.1% SDS; and 0.1 × SSC, 0.1% SDS. The membrane was exposed to x-ray film for 24 h at −70 °C. The autoradiogram was matched to the prestained markers to determine the size of the native mRNA. Before hybridization with the cDNA-specific probe, the Northern blot membrane was analyzed by hybridization to a β-actin-specific probe to confirm the integrity of the mRNA. After hybridization to the β-actin probe detected an mRNA species of the predicted size (approximately 2.1 kilobase pairs), the membrane was stripped by boiling for 30 min in 0.1 × SSC, 0.5% SDS and probed according to an identical procedure with the probe specific for the human homologue of endonuclease III (Fig. 1).

Preparation of Probes for Northern Blot Analysis—The β-actin probe was produced by PCR with sequence-specific primers (Clontech) against cDNA made from the RNA of cells taken from a sample of a human bone marrow aspirate. PCR was conducted as follows: 1 cycle of 95 °C for 2 min and 35 cycles of 94 °C for 1 min, 60 °C for 1 min, 72 °C for 1 min, followed by a final cycle of 10 min at 72 °C. The probe was then radiolabeled using the Random Primed DNA Labeling kit (Boehringer Mannheim) and [α-^{32}P]dCTP, and it was purified using Nick-Spin columns (Pharmacia). The specific probe for the human homologue of endonuclease III was prepared by excising the full-length cDNA sequence shown in Fig. 1 from the 2 μg of purified plasmid DNA via restriction with *Eco*RI and *Bam*HI followed by gel purification of the restricted fragment. The probe was radiolabeled and hybridized to the Northern blot membrane as described.

DNA Glycosylase Assay—Poly(dA-[^{3}H]dT) was produced by nick translation of the alternating copolymer poly(dA-dT) (Pharmacia) with [5′,5-^{3}H]TTP followed by oxidation with osmium tetroxide to form thymine glycol residues (23). Thymine glycol-containing poly(dA-[^{3}H]dT) produced in this manner had a specific activity of approximately 1.4 × 10^7 dpm/μg. Thymine glycol DNA-glycosylase assays were carried out against oxidized DNA and the released radioactive product proven to be thymine glycol by high pressure liquid chromatography analysis as described previously (23).

Sodium Cyanoborohydride-mediated Cross-linking of Fusion Protein to a Thymine Glycol-Containing Oligodeoxynucleotide—A double-stranded oligodeoxynucleotide containing a single thymine glycol-residue was prepared as described previously (18, 24). The thymine glycol-containing strand was 5′-end-labeled with [γ-^{32}P]dATP, using T4 kinase (Life Technologies) according to the manufacturer's recommendations, and purified using a ChromaSpin-10 column (Clontech).

The purified GST fusion protein, the nonfusion GST protein, and *E. coli* endonuclease III were reacted with the substrate double-stranded oligodeoxynucleotide in a total volume of 50 μl under the following reaction conditions: 37.3 mM $NaCNBH_3$, 20 mM HEPES, pH 7.5, 46.5 mM KCl, 5 mM EDTA, a 4.0 μM concentration of each oligodeoxynucleotide, and 40 ng/μl protein. In the case of *E. coli* endonuclease III, this represented approximately a 4-fold molar excess of substrate deoxyoligonucleotide to enzyme. After incubation at room temperature for 2 h, a 25-μl volume of 3 × SDS-PAGE loading buffer was added to each sample. Samples were then heated to 90 °C for 5 min and separated by electrophoresis on a 15% Tricine-SDS gel. Following electrophoresis, the gel was stained with Coomassie Blue, wrapped in plastic, and analyzed via autoradiography.

Gel Electrophoresis—Prior to electrophoresis all samples were incubated at 95 °C for 5 min in standard SDS-PAGE loading buffer. Fifteen percent Tricine gels (25) were prepared and run using the Mini-Protein II electrophoresis system (Bio-Rad). Gels were run at 90 V for approximately 5 h, completion being determined by the progress of prestained low molecular weight electrophoresis standards (Bio-Rad). Gels were then stained with Coomassie Blue.

RESULTS

Fig. 1 presents the nucleotide sequence of a cDNA of 1045 base pairs, which contains a putative open reading frame (ORF) of 912 base pairs. This ORF encodes a protein of 304 amino acids with a calculated molecular mass of 33,569 and a calculated pI of 9.85, which is the human homologue of *E. coli* endonuclease III. The nucleotide sequence data presented in Fig. 1 were obtained from two sources. The sequence of nucleotides 6–1045 was obtained by analysis of clones isolated from a cDNA library, using probes based upon the sequence of the previously described human 3′-EST. The sequence of nucleotides 1–5 was obtained by sequencing the products of 5′-RACE, performed using gene-specific primers based upon the sequence of the longest cDNA clones.

Previously we reported the sequence of four peptides obtained by proteolysis of a purified bovine pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase (18). The sequences of those four peptides as well as that of one additional peptide (GEGGEGAEHLQAP) derived from the same purified protein are also included in Fig. 1, aligned with the homologous sequences encoded within the ORF of the human cDNA.

The 1045-base pair sequence of Fig. 1 probably represents most, if not all, of the entire full-length cDNA. The Northern blot analysis (Fig. 2) of human splenic and 293T cell (human) mRNA each demonstrate a predominant mRNA species of approximately 1.1–1.2 kilobase pairs, which hybridized to a ^{32}P-labeled probe containing the entire sequence of the ORF. The

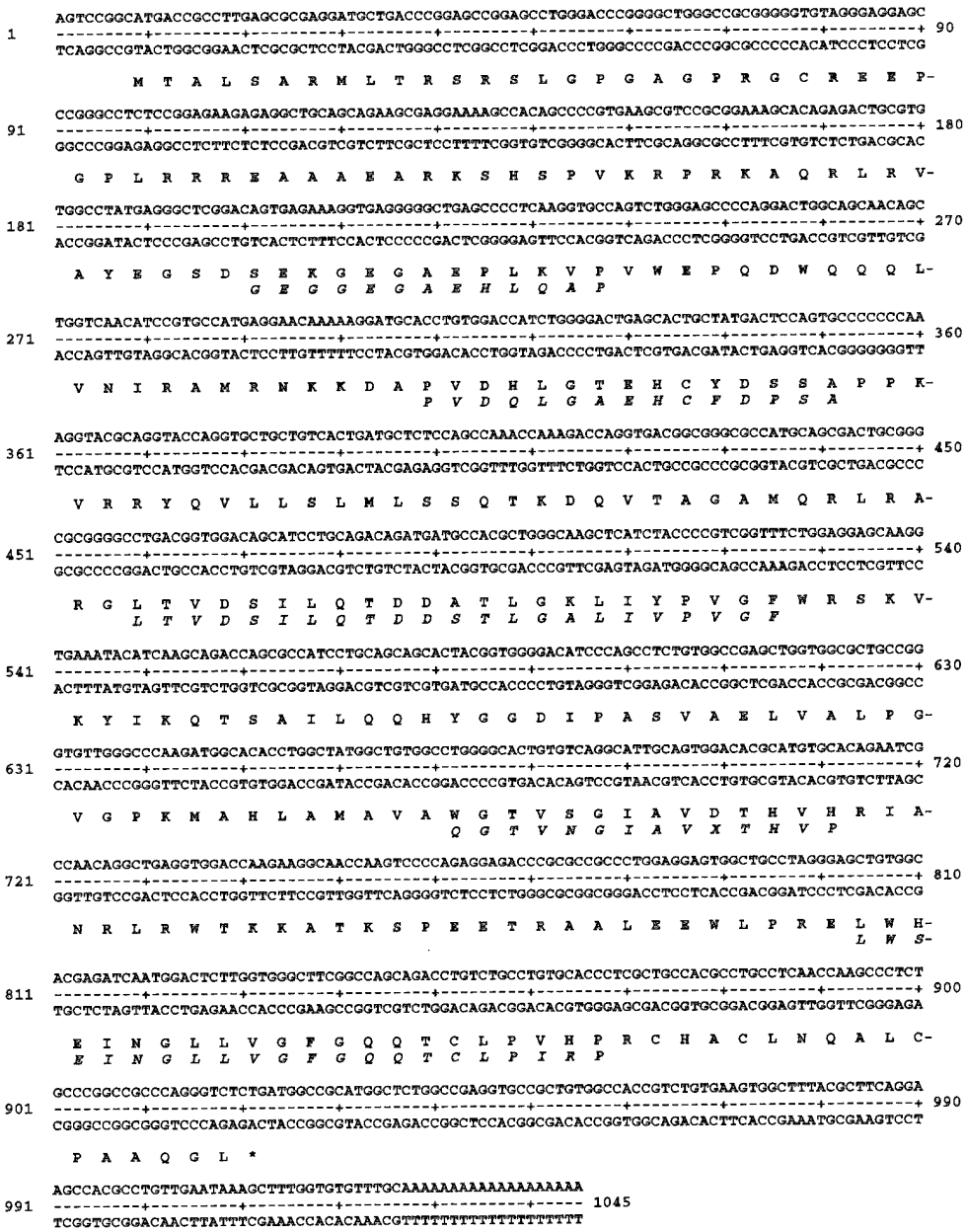


FIG. 1. **Nucleotide and deduced amino acid sequence of the human pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase.** The sequences of peptides obtained by proteolytic digestion of purified bovine pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase are in *italic type* and are aligned with the homologous human amino acid sequence.

difference of approximately 50–150 nucleotides in length between the cDNA sequence presented in Fig. 1 and the native mRNA can be explained by the expected presence of a poly(A) tail of approximately the same length on the native species and perhaps a few more nucleotides 5′ to the first AUG codon.

Fig. 2, *lane 3*, which contains mRNA extracted from 293T cells, shows a second faint band of higher M_r. Although we think this band is nonspecific, we cannot fully exclude the possibility that it represents mRNA encoding a protein similar to human endonuclease III. Such a situation is present in *S. cerevisiae*, which contains two homologues of *E. coli* endonuclease III, one of which is thought be nuclear, the second mitochondrial (see "Discussion" and Fig. 8).

To demonstrate that the cDNA sequence of Fig. 1 encoded a functional homologue of endonuclease III, a GST fusion protein was constructed consisting of amino acid residues 8–304 of the ORF fused to the C terminus of the 30-kDa GST protein. SDS-PAGE analysis of the IPTG-induced, affinity-purified fusion protein (Fig. 3) revealed a predominant 65-kDa full-length protein. Two additional lower molecular weight protein species were present in the purified preparation. We believe these to be fragments of the 65-kDa protein that arose through abortive synthesis of the full-length protein or proteolysis occurring before, during, and possibly after cell lysis and affinity purification, due to the action of contaminating cellular proteases.

As demonstrated previously, *E. coli* endonuclease III can be specifically, irreversibly cross-linked to a thymine glycol-containing oligodeoxynucleotide via the reductive stabilization of its characteristic enzyme-substrate intermediate (18). To further confirm that the ORF presented in Fig. 1 encoded a fully functional homologue of *E. coli* endonuclease III, the cross-linking reaction, as described under "Experimental Procedures," was applied to the purified GST-fusion protein. The results of this reaction are illustrated in Fig. 4. When aliquots of the purified GST fusion protein that had been incubated with a ^{32}P-labeled thymine glycol-containing oligodeoxynucleotide

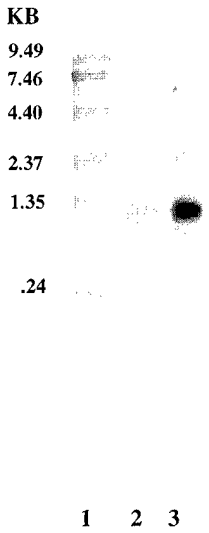


FIG. 2. **Northern blot analysis.** Northern blot analysis was performed against 1 μg of mRNA from human spleen (*lane 2*) and 2 μg of mRNA from human 293T cells (*lane 3*), using the full-length ^{32}P-labeled cDNA for the human pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase as a probe. Methylene blue-stained RNA markers are shown in *lane 1*.

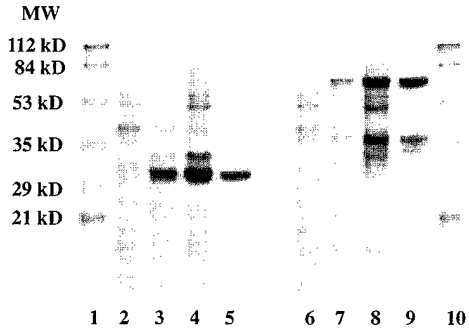


FIG. 3. **Expression and purification of the recombinant human pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase.** SDS-PAGE analysis of the GST fusion protein is shown. *Lane 2*, total SDS lysate from uninduced *E. coli* containing the pGEX-2T vector; *lane 3*, total SDS lysate of the same *E. coli* after induction by IPTG for 5 h; *lane 4*, soluble fraction obtained by centrifugation of induced *E. coli* disrupted by sonication; *lane 5*, purified GST protein after elution from glutathione-agarose affinity media; *lane 6*, total SDS lysate from uninduced *E. coli* containing the pGEX-2T vector into which the sequence encoding the human enzyme had been cloned; *lane 7*, total SDS lysate from induced *E. coli* containing the recombinant pGEX-2T vector; *lane 8*, the soluble fraction of induced disrupted *E. coli* containing the recombinant pGEX-2T vector; *lane 9*, purified GST fusion protein after elution from affinity medium. *Lanes 1* and *10* are M_r markers.

in the absence (*lane 6*) or presence (*lane 7*) of sodium cyanoborohydride ($NaCNBH_3$) were compared by SDS-PAGE analysis, it became evident that a portion of the protein had been irreversibly cross-linked to the oligodeoxynucleotide. This is manifested by an increase in the apparent molecular weight of the enzyme resulting in formation of the doublet shown in *lane 7*. This shift is analogous to that observed when endonuclease III was subjected to the same reductive cross-linking reaction (*lane 3*) and compared with native endonuclease III (*lane 2*). No shift of the major protein species was observed when the non-fusion GST protein (*lane 3*) was incubated under reducing conditions with the thymine glycol-containing oligodeoxynucleotide (*lane 4*).

An autoradiogram of the gel in Fig. 4*A* is presented in Fig. 4*B*. As described, the thymine glycol-containing oligodeoxynucleotide had been 5′-end-labeled with ^{32}P prior to incubation with the proteins. Thus, cross-linking was confirmed by

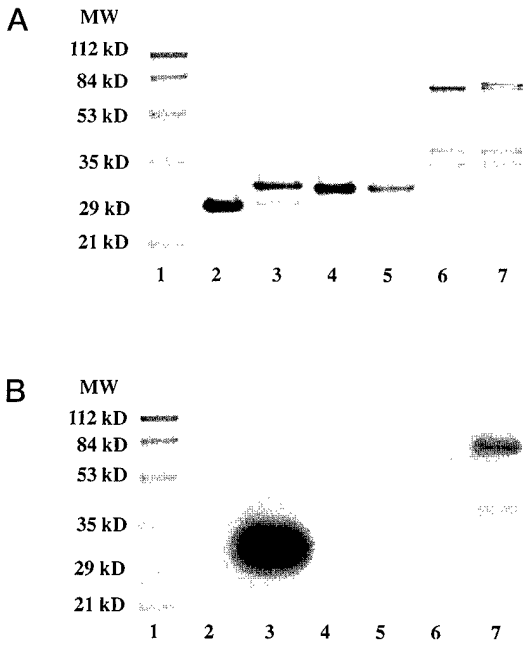


FIG. 4. **SDS-PAGE analysis of *E. coli* endonuclease III and the human GST fusion protein after incubation with the thymine glycol-containing oligodeoxynucleotide and $NaCNBH_3$.** *A*, *lane 1* contains M_r markers. *Lane 2* contains the product of the incubation of *E. coli* endonuclease III with $NaCNBH_3$. *Lane 3* contains the product of the same incubation mixture as *lane 2* with the addition of duplex ^{32}P-5′-end-labeled oligodeoxynucleotide containing a single thymine glycol residue. *Lane 4* contains the product of the incubation the purified non-fusion GST protein (Fig. 3, *lane 5*) with $NaCNBH_3$ but no oligodeoxynucleotide. *Lane 5* contains the product of the incubation of the same purified non-fusion GST protein with $NaCNBH_3$ and the ^{32}P-5′-end-labeled oligodeoxynucleotide. *Lanes 6* and *7* contain the products of the incubation of the purified GST fusion protein (Fig. 3, *lane 9*) with $NaCNBH_3$ alone or with $NaCNBH_3$ and oligodeoxynucleotide, respectively. *B*, phosphor image of the SDS-PAGE gel of Fig. 4*A*. The *lanes* are identical to those described in *A*. The M_r in *lane 1* are not radiolabeled but are the same Coomassie-stained markers shown in Fig. 4*A*.

this autoradiogram in which predominant radioactive species are present only in *lanes 2* (*E. coli* endonuclease III plus $NaCNBH_3$) and *7* (GST fusion plus $NaCNBH_3$), which correspond in apparent M_r to the shifted species seen on the Coomassie Blue-stained gel. Also evident on the autoradiogram in *lane 7* are two visible, but less intense, lower molecular weight bands that correspond in position to presumed degradation products of the fusion protein present even after affinity purification (Fig. 3). Presumably these represent cross-linked, partially degraded fusion protein.

After purification, the fusion protein was also analyzed for thymine glycol-DNA glycosylase activity. Fig. 5 presents the *V versus* $[E_t]$ plot in which thymine glycol release is expressed as a function of increasing content of fusion protein. The release of thymine glycol is linear with respect to fusion protein concentration over the amount of protein used. Based on the results of this plot, the specific enzymatic activity of the fusion protein was calculated to be about 1–2% that of genetically engineered *E. coli* endonuclease III using the same assay (latter assay data not shown). This reduced level of activity is apparently quite common among GST fusion proteins.[2] GST protein that contained no C-terminal fusion was induced and purified in a manner identical to the fusion protein and assayed for enzymatic activity. This non-fusion GST protein did not demonstrate detectable thymine glycol-DNA glycosylase activity at a protein concentration 3 orders of magnitude higher than that

[2] R. Schneider, personal communication.

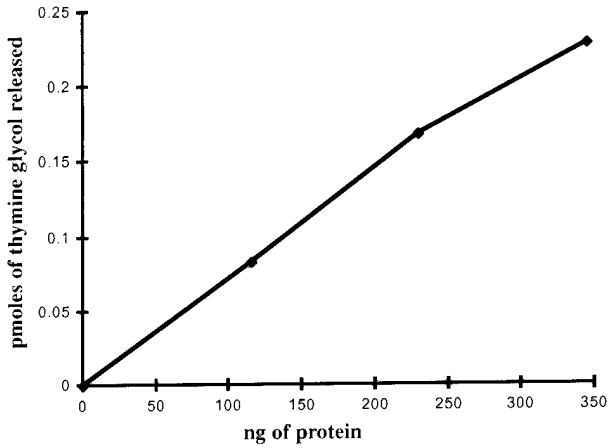


FIG. 5. ***V versus* [E_t] plot.** The amount of thymine glycol released after incubation of oxidized alternating poly(dA-dT) for 20 min with recombinant protein is shown. The *points* represent the average of two determinations. There was less than 5% variability among duplicate samples.

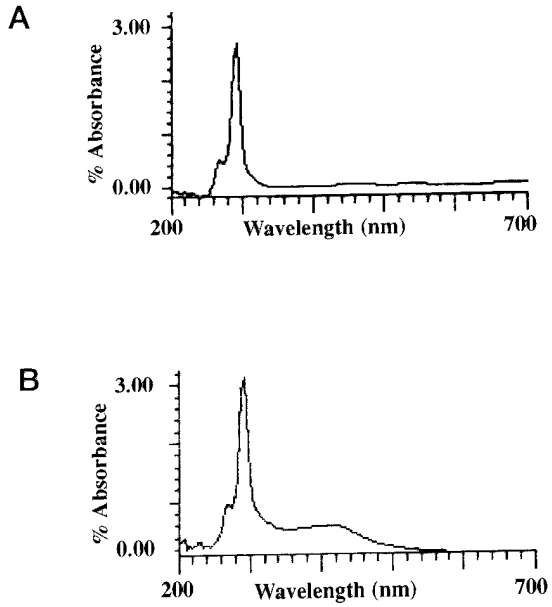


FIG. 6. **Spectroscopic analysis.** *A*, optical absorption spectrum of the purified human pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase-GST fusion protein. *B*, optical absorption spectrum of the purified non-fusion GST protein of *S. japonicum*.

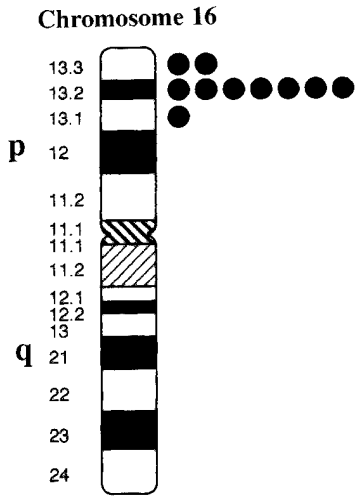


FIG. 7. **Histogram of FISH analysis results.** Ten mitotic figures to which the FISH probe had bound were analyzed to determine the precise position of the gene on chromosome 16. Each *dot* represents the position of the human gene as determined through one such analysis.

at which the fusion protein was assayed.

As documented previously, *E. coli* endonuclease III contains an iron-sulfur cluster in which a cubane [4Fe-4S] moiety is liganded by four cysteine residues. This domain produces a distinctive absorbance at 410 nm (26). Conservation of this [4Fe-4S] cluster in the human enzyme was inferred on the basis of the cDNA sequence of Fig. 1, since the putative ORF contains the appropriate four cysteine residues at amino acid positions 282, 289, 292, and 300, and confirmed by taking an absorption spectrum of the purified GST-fusion protein, which revealed that it too absorbed strongly at 410 nm (Fig. 6).

Although purified *E. coli* endonuclease III has a characteristic absorption peak at 410 nm and might be expected to appear blue in solution, the color of solutions containing approximately 0.5 mg/ml or greater of purified endonuclease III are typically yellow-brown (19). Similarly, a solution of the purified GST fusion protein at similar concentrations of protein was also yellow, while a solution of the simultaneously purified non-fusion GST protein was colorless.

In order to determine the chromosomal localization of the gene encoding the mammalian enzyme, FISH analysis was performed as described under "Experimental Procedures." Under the conditions used, hybridization efficiency for our probe was approximately 70% (*i.e.* among 100 mitotic spreads analyzed, 70 demonstrated binding of the probe to one pair of chromosomes). DAP1 banding was used to identify the chromosome pair to which the probe had bound (chromosome 16). The precise localization of the gene (16p13.2) was determined by the summary analysis of 10 pairs of photographs in which the probe signal was matched with the results of DAP1 banding (Fig. 7). There was no additional locus detected by FISH analysis. These results taken together with the presence of a single mRNA species on Northern analysis indicates that the gene for human endonuclease III is a single copy gene.

DISCUSSION

The human sequence of Fig. 1 shows a remarkable similarity to that of several other putative homologues of the *E. coli* endonuclease III (National Center for Biotechnology Information (NCBI) sequence ID 119329) found in representative species of all three biologic domains. In bacteria they have been found in both Gram-negative (*Hemophilus influenza*, NCBI sequence ID 1169526) and Gram-positive (*Bacillus subtilis*, NCBI sequence ID 729418) organisms; among archeons, in *Methanococcus jannaschii* (NCBI sequence ID 1510694); and among eukaryotes, in *Schizosaccharomyces pombe* (NCBI sequence ID 1065894), *S. cerevisiae* (NCBI sequence ID 1419843 and 401436), *C. elegans* (NCBI sequence ID 974795), *Rattus sp.* (accession number H33255), and *Homo sapiens* (accession number F04657). The *S. cerevisiae* genome encodes two distinct theoretical homologues of *E. coli* endonuclease III. The alignment of the nine putative homologous sequences using the program Clustal W (version 1.5) (Fig. 8) reveals that a core sequence of amino acids is remarkably well conserved. In bacteria, the core sequence comprises virtually the entire protein. In contrast, the proteins of archeons and eukaryotes have unique extensions at their N and/or C termini. For the sake of clarity, these extensions have been omitted from Fig. 8.

Based upon similarities among several bacterial DNA glycosylases, site-directed mutagenesis studies, and molecular modeling, Thayer *et al.* (26) identified several regions and residues within the core sequence of amino acids of *E. coli* endonuclease III that could be involved in DNA binding and catalysis. The

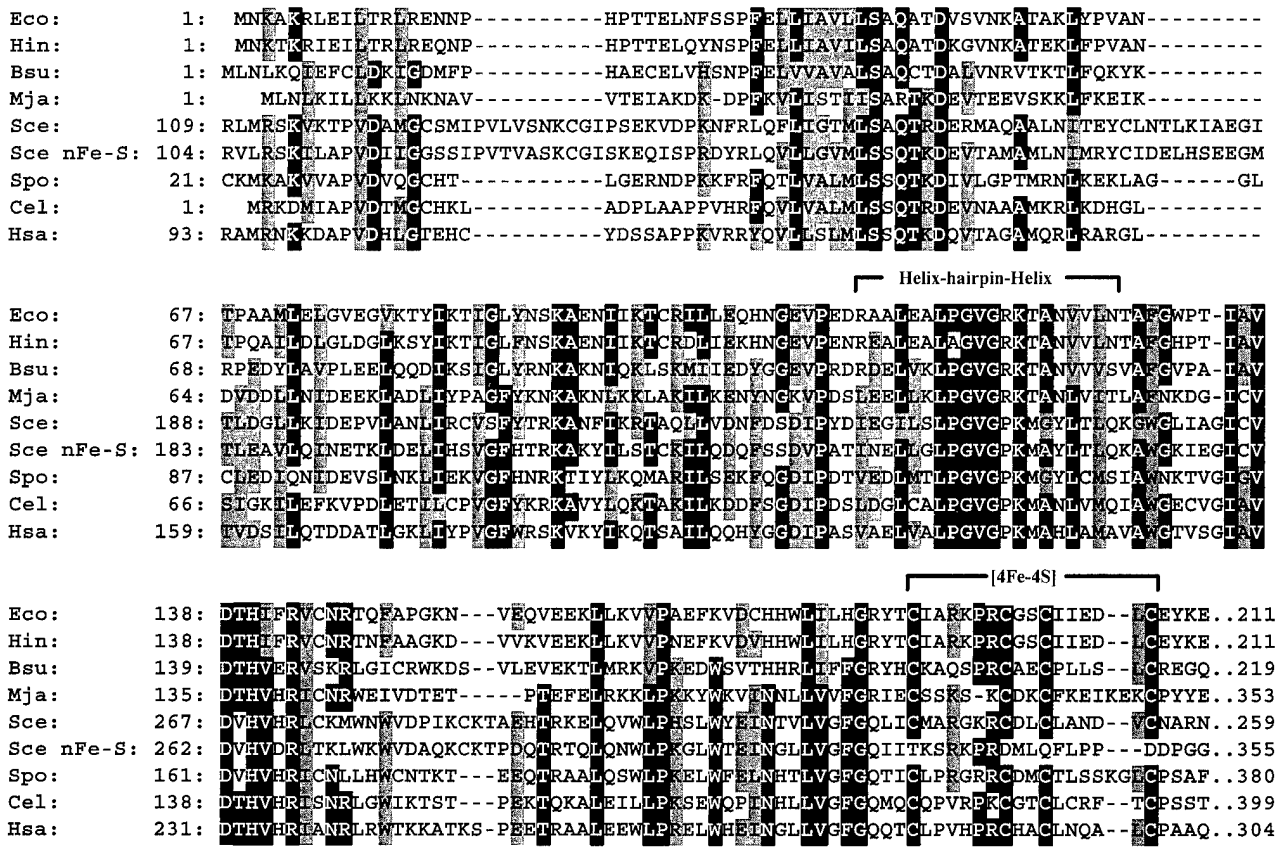


FIG. 8. **Alignment of the amino acid sequence of *E. coli* endonuclease III with those of putatively homologous proteins from three evolutionary domains.** The amino acid sequence of *E. coli* endonuclease III (*Eco*) is aligned with homologous sequences from *H. influenza* (*Hin*), *B. subtilis* (*Bsu*), *M. jannaschii* (*Mja*), *S. pombe* (*Spo*), *C. elegans* (*Cel*), *H. sapiens* (*Hsa*), and two unique homologous sequences from *S. cerevisiae* (*Sce* and *Sce nFe-S*). Residues in *black boxes* indicate identical sequences. Residues in *gray boxes* indicate conservative substitution. *Dashes* denote gaps in sequence introduced to maximize alignment. *Numbers* in the *left column* refer to the first amino acid residue in each line of the respective protein sequences. *Numbers* in the *lower right* indicate the total number of amino acid residues in each protein sequence. In archeons and eukaryotes, the proteins that are homologous to *E. coli* endonuclease III have unique extensions at their N and/or C termini. For the sake of clarity, these extensions have been omitted from the figure. Alignment of residues 83–304 of the human enzyme with residues 2–209 of the *E. coli* enzyme demonstrates that there is 29.3% identity and 51.9% similarity between the two proteins.

region surrounding glutamine 41 (residue numbers refer to the *E. coli* endonuclease III amino acid sequence unless otherwise indicated) may form a portion of the substrate binding pocket, in which the damaged pyrimidine fits when in the "flipped out" conformation that the enzyme recognizes. The Helix-hairpin-helix (HhH) motif encoded by the residues surrounding the central LPGVG sequence (residues 114–118) is thought to function in nonspecific DNA recognition. Recently, Doherty *et al.* (27) have extended this analysis and shown that similar HhH motifs occur in 14 homologous families of DNA-binding proteins, including DNA glycosylases, DNA polymerases, and "flap" endonucleases. Lysine 120 appears to be the nucleophile in the active site of endonuclease III that contributes the ϵ-amino group necessary for the formation of the *N*-acylimine enzyme-substrate intermediate, characteristic of DNA glycosylase/AP lyases. Aspartic acid 138 has also been implicated as a functional active site residue. All of these residues appear to be well conserved in all of the nine sequences shown. The structure of the *E. coli* endonuclease III was recently solved (26), and, in light of the high degree of conservation of critical residues, it is likely that the common core sequence of all members of the endonuclease III family will have a similar three-dimensional structure.

In addition to the previously mentioned residues, four highly conserved cysteine residues (187, 194, 197, and 203) have been identified within this common core sequence that contribute to the [4Fe-4S] cluster of *E. coli* endonuclease III. Examination of the aligned sequences in Fig. 8 reveals that in *E. coli* endonuclease III and five of its eight putative homologues, including the human enzyme, these four cysteines are arranged according to the consensus sequence Cys-X_6-Cys-X_2-Cys-X_5-Cys. A similar but slightly modified sequence appears in *S. pombe* (Cys-X_6-Cys-X_2-Cys-X_7-Cys) and *M. jannaschii* (Cys-X_5-Cys-X_2-Cys-X_7-Cys). Thayer *et al.* (26) suggested that basic amino acid residues between the first two cysteines of the [4Fe-4S] cluster may form a loop that functions in the nonspecific binding of DNA. While Fig. 8 does not indicate absolute conservation of these residues, some conservation is apparent, especially with respect to arginine 193.

As mentioned previously, the genome of *S. cerevisiae* encodes two putative homologues of *E. coli* endonuclease III, one of which (designated Sce non-Fe-S in Fig. 8 (NCBI sequence ID 1419843)) lacks the four-cysteine [4Fe-4S] motif completely and presents an obvious exception to this consensus sequence. However, this sequence also encodes a putative mitochondrial leader sequence (28). Whether pairs of endonuclease III-like proteins, with and without [4Fe-4S] clusters, are present in other eukaryotic organisms and whether the non-Fe-S proteins are mitochondrial remains to be determined.

This interesting question notwithstanding, the presence of endonuclease III-like enzymes in representative species of all three evolutionary domains suggests that the genomic DNA of organisms throughout phylogeny is subject to endogenous stresses that attack the 5,6-double bonds of pyrimidine resi-

dues. Previously well characterized substrates of endonuclease III include oxidized pyrimidines such as thymine glycol and 5-hydroxycytosine and hydrates of cytosine and uracil. The oxidation of DNA bases has been primarily attributed to reactive oxygen species formed as byproducts of oxidative metabolism and inflammation. The formation of pyrimidine hydrates has been primarily attributed to the action of UV radiation (reviewed in Ref. 29). The archeon *M. jannaschii* lives beneath the sea and therefore is not exposed to direct sunlight. Furthermore, it is characterized by a reducing rather than an oxidizing metabolism (30). The identification of a homologue of endonuclease III in the genome of this organism suggests that pyrimidines with reduced 5,6-double bonds such as 5,6-dihydrothymine may be formed spontaneously in archeon genomic DNA. Perhaps within this evolutionary domain, it is primarily the formation of such reduced rather than oxidized or photohydrated pyrimidine residues that has promoted the conservation of an endonuclease III-like enzyme.

At this time, the specific contribution that the human pyrimidine hydrate-thymine glycol DNA glycosylase/AP lyase activity makes to the maintenance of the genome is uncertain. The human gene encoding this enzyme was localized to the locus 16p13.2-.3 by FISH analysis (Fig. 7). The accuracy of this localization was corroborated through the identification of genomic data base nucleotide sequence (accession number L48777) obtained by exon trapping from this same region of chromosome 16 (31), which is 94.1% identical to nucleotides 699–799 of the sequence of Fig. 1. The chromosomal locus of the human endonuclease III homologue is in very close proximity to that of another DNA base excision repair enzyme, 3-methylpurine DNA glycosylase as well as the DNA nucleotide excision gene, ERCC-4. There is no apparent homology among these three proteins, so it seems unlikely that their localization to the same chromosomal region is the result of gene duplication and divergence. Loss of heterozygosity in this region has been reported to occur in 22% of human hepatocellular carcinomas (32). Whether any or all of these DNA repair proteins act as tumor suppressors for human hepatocarcinogenesis remains to be determined.

Acknowledgments—We thank Philip H. Bolton for supplying us with thymine glycol-containing oligodeoxynucleotide, Ajay Chheda for assistance with computer-assisted sequence retrieval, and Archie Cummings, Jr., for technical assistance. We thank Dr. Bernard Goldschmidt of the New York University Skirball Institute Sequencing Facility for DNA sequence.

REFERENCES

1. Brent, T. P. (1973) *Biophys. J.* **13,** 399–401
2. Bacchetti, S., and Benne, R. (1975) *Biochim. Biophys. Acta* **390,** 285–297
3. Duker, N. J., and Teebor, G. W. (1975) *Nature* **255,** 82–84
4. Ness, I. F., and Nissen-Meyer, J. (1978) *Biochim. Biophys. Acta* **520,** 111–121
5. Demple, B., and Linn, S. (1980) *Nature* **287,** 203–208
6. Cunningham, R. P., and Weiss, B. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 474–478
7. Doetsch P. W., Helland, D. E., and Haseltine, W. A. (1986) *Biochemistry* **25,** 2212–2220
8. Boorstein, R. J., Hilbert, T. P., Cadet, J., Cunningham, R. P., and Teebor, G. W. (1989) *Biochemistry* **28,** 6164–6170
9. Radman, M. (1976) *J. Biol. Chem.* **251,** 1438–1445
10. Bailly, V., and Verly, W. G. (1987) *Biochem. J.* **242,** 565–572
11. Kim, J., and Linn, S. (1989) *J. Biol. Chem.* **264,** 2739–2745
12. Mazumder, A., Gerlt, J. A., Absolon, J. M., Stubbe, J., Cunningham, R. P., Withka, J., and Bolton, P. H. (1991) *Biochemistry* **30,** 1119–1126
13. Tchou, J., and Grollman, A. P. (1995) *J. Biol. Chem.* **270,** 11671–11677
14. van der Kemp, P. A., Thomas, D., Barbey, R., de Oliveira, R., and Boiteux, S. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 5197–5202
15. Nash, H. M., Bruner, S. D., Scharer, O. D., Kawate, T., Addona, T. A., Spooner, E., Lane, W. S., and Verdine, G. L. (1996) *Curr. Biol.* **6,** 968–980
16. Dodson, M. L., Schrock, R. D., III, and Lloyd, R. S. (1993) *Biochemistry* **32,** 8284–8290
17. Dodson, M. L., Michaels, M. L., and Lloyd, R. S. (1994) *J. Biol. Chem.* **269,** 32709–32712
18. Hilbert, T. P., Boorstein, R. J., Kung, H. C., Bolton, P. H., Xing, D., Cunningham, R. P., and Teebor, G. W. (1996) *Biochemistry* **35,** 2505–2511
19. Asahara, H., Wistort, P. M., Bank, J. F., Bakerian, R. H., and Cunningham, R. P. (1989) *Biochemistry* **28,** 4444–4449
20. Heng, H. H. Q., Squire, J., Tsui, L.-C. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 9509–9513
21. Heng, H. H. Q., and Tsui, L.-C. (1993) *Chromosoma* **102,** 325–332
22. Heng, H. H. Q., and Tsui, L.-C. (1994) in *Methods in Molecular Biology: In Situ Hybridization Protocols* (Choo, K. H. A., ed) pp. 35–49, Humana Press, Clifton, NJ
23. Higgins, S. A., Frenkel, K., Cummings, A., and Teebor, G. W. (1987) *Biochemistry* **26,** 1683–1688
24. Kao, J. Y., Goljer, I., Phan, T. A., and Bolton, P. H. (1993) *J. Biol. Chem.* **268,** 17787–17793
25. Shagger, H., and von Jagow, G. (1987) *Anal. Biochem.* **166,** 368–379
26. Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P., and Tainer, J. A. (1995) *EMBO J.* **14,** 4108–4120
27. Doherty, A. J., Serpell, L. C., and Ponting, C. P. (1996) *Nucleic Acids Res.* **24,** 2488–2498
28. Ouellette, B. F, Clark, M. W., Keng T., Storms, R. K., Zhong, W., Zeng, B., Fortin N., Delaney S., Barton, A., Kaback, D. B., and Bussey, H. (1993) *Genome* **36,** 32–42
29. Teebor, G. W. (1995) in *DNA Repair Mechanisms and Cancer* (Vos, J.-M. H., ed.) pp 99–123, R. G. Landes Co., Austin, TX
30. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, G. G., Sutton, J. A. B., FitzGerald, L. M., and Clayton, R. A. (1996) *Science* **273,** 1058–1073
31. Burn, T. C., Connors, T. D., Klinger, K. W., and Landes, G. M. (1995) *Gene (Amst.)* **161,** 183–187
32. Sakai, K., Nagahara, H., Abe, K., Obata, H., (1992) *J. Gastroenterol. Hepatol.* **7,** 288–292

992–994 Nucleic Acids Research, 1997, Vol. 25, No. 5

The p53 status of Chinese hamster V79 cells frequently used for studies on DNA damage and DNA repair

Wenren Chaung, Li-Jun Mi and Robert J. Boorstein*

Department of Pathology, New York University School of Medicine, Sackler Institute of Graduate Biomedical Sciences, and The Rita and Stanley Kaplan Cancer Center, New York, NY 10016, USA

Received November 25, 1996; Revised and Accepted January 21, 1997

DDBJ/EMBL/GenBank accession nos Y08900, Y08901

ABSTRACT

Chinese hamster lung fibroblast V79 cells have been widely used in studies of DNA damage and DNA repair. Since the p53 gene is involved in normal responses to DNA damage, we have analyzed the molecular genetics and functional status of p53 in V79 cells and primary Chinese hamster embryonic fibroblast (CHEF) cells. The coding product of the p53 gene in CHEF cells was 76 and 75% homologous to human and mouse p53 respectively, and was 95% homologous to the Syrian hamster cells. The V79 p53 sequence contained two point mutations located within a presumed DNA binding domain, as compared with the CHEF cells. Additional immunocytochemical and molecular studies confirmed that the p53 protein in V79 cells was mutated and nonfunctional. Our results indicate that caution should be used in interpreting studies of DNA damage, DNA repair and apoptosis in V79 cells.

INTRODUCTION

The V79 cell has been widely used in studies on X-ray, UV radiation and oxidizing agent induced DNA damage and DNA repair. Over 300 primary studies using V79 cells have been published in the past 20 years (1–7).

The protein product of p53 gene has been suggested to act as 'the guardian of the genome' (8). Evidence suggests that p53 temporarily halts the cell cycle in response to DNA damage to allow time for DNA to be repaired (9). For example, p53 is activated in response to DNA damage, and overexpression of wild type p53 induces a pronounced accumulation of the *mdm2* gene product at mRNA and protein levels (10,11). It has also been suggested that p53 works through Gadd45 and perhaps can directly stimulate the repair machinery as well (12). Another function of the p53 product is to mediate the apoptosis (13,14).

To clarify the role of the p53 tumor suppresser gene product in the response of V79 cells to DNA damage, p53 cDNA of V79 cells was cloned and sequenced. p53 cDNA was also collected from Chinese hamster embryonic fibroblast (CHEF) cells (at early passage P7, known to contain wild type p53) and used as the control. In this report we compare the p53 cDNA sequence of CHEF and V79 cells, check the homology between different species, and define the mutations in V79 cells. Immunohistochemical and RNA dot blot analyses were also used to determine the biological function of p53 in V79 cells. These studies have implications for the interpretation and generalizability of studies of the mutagenicity of DNA damaging agents to V79 cells.

MATERIALS AND METHODS

Cells and cell lines

CHEF/P7 cell was a gift from Dr John Lehman (Albany Medical College). V79 cell at low passage was maintained as previously described (6). The $p53^{-/-}$ human promyelocytic HL-60 cells (15) and mouse embryo fibroblasts (MEF) known to have wild type p53 (16) were generously provided by Dr Robert Carroll (NYUMC).

Immunohistochemical analysis and RNA dot blotting assay of p53 protein

The status of p53 protein in V79 and CHEF cell was analyzed by immunocytochemistry (17) using the p53 monoclonal antibody DO-1 (Santa Cruz Biotechnology) and pAb240 (Oncogene Science). The detection of the *mdm2* mRNA level was performed by the modified RNA dot blotting method described by Kline *et al.* (18). Two oligonucleotides homologous to the hamster *mdm2* cDNA were synthesized (5′-CCAGCTTCGGAACAAGAGAC-3′ and 5′-GGTGGAAGGGGAGGATTCATT-3′) and used as the primers. The PCR product from the two *mdm2* primer sets then was used as the probe. The β-actin primer sets were purchased from CLONTECH and their PCR product was used as an internal control probe. The densitometric analysis was done using the 'NIH Image 1.6 on the Macintosh' program.

cDNA library screening, sequencing and analysis

The constructed cDNA library from each cell line was selected with a PCR amplified probe using the primer set suggested by Legros (19). Sequencing was carried out directly on the positive cDNA clones using method described elsewhere (20). To preclude the possibility that *Taq* polymerase errors might be interpreted as mutations, the p53 coding region from the mutiple cDNA clones

*To whom correspondence should be addressed. Tel: +1 212 263 8530; Fax: +1 212 263 8211; Email: robert.boorstein@ccmail.med.nyu.edu

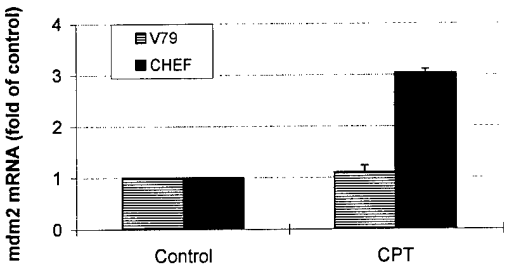


Figure 1. Densitometric analysis of *mdm2* mRNA induction by camptothecin in CHEF cells and V79 cells from RNA dot blotting experiment. The relative amount of *mdm2* mRNA is normalized relative to β-actin mRNA for each condition. The data plotted is the average of three determinations.

was sequenced twice to confirm the mutations in the V79 and CHEF cells. The analysis programs used were 'ALIGN' from EERIE-Nimes, France and 'BCM Multiple sequence alignments program-CLUSTAL-W' at Baylor College of Medicine (21,22).

RESULTS AND DISCUSSION

The V79 cell line has been widely used to study the toxicity, mutagenicity and repair of a wide variety of DNA damaging agents. We have utilized these cells to study the toxicity of the thymidine analogue 5-hydroxymethyl-2′-deoxyuridine (hmdUrd), and to study the repair of hydroxymethyluracil (hmUra) from DNA (5,6). We have found that hmdUrd is able to induce apoptosis in the V79 cells but not in V79*mut1* (hmUra-DNA glycosylase deficient line) cells (23). Since the p53 protein is elevated in response to other types of DNA damage (24), is involved in binding to and reannealing strand breaks (25), and is integrally involved in apoptosis (13), it became necessary to ask whether the p53 gene and its expression were altered in these cells.

First, we evaluated the status of p53 in the V79 and CHEF/P7 cells by immunocytochemistry. V79 cells stained strongly positive with the anti p53 monoclonal antibody PAb-240. PAb-240 recognizes the epitope between AA213 and 217 which is exposed only when p53 is denatured or mutated. PAb-240 did not stain either CHEF cells, HL-60 cells ($p53^{-/-}$) or MEF cells (which have only wild type p53). Although antibody p53 DO-1 recognizes both wild type and mutant type of p53, the strongly positive staining inside the nucleus of V79 cells suggested the p53 protein was mutated in V79 cells. In contrast, CHEF cells showed very weak staining mainly in the cytoplasm, indicating wild type p53. This result was confirmed by Dr Lehman (personal communication) and is consistent with other studies (26).

Next, we cloned and sequenced the complete cDNA from the CHEF and V79 cells. The CHEF sequence contained 2041 bp (GenBank No. Y08900), while the V79 cDNA contained 2073 bp (GenBank No. Y08901). These two cell lines showed 98% homology to each other. Two mutations were found in the coding region of the p53 cDNA (#136 leucine [CTA]→glutamine [CAA]; #138 cysteine [TGC]→tryptophan [TGG]). These mutations, within the p53 DNA binding domain, presumably affect the ability of p53 to bind to DNA and thus regulate gene expression (27). The mutated sites in V79 cells are located within the evolutionary conserved box #2 (27,28) and is within one of the 'hotspots' found in human p53 described by Vogelstein and Kinzler (29). When compared with the Syrian hamster cell (GenBank No. M75144), the Chinese hamster cells showed 91% identity of the cDNA and 95% homology (only 19 amino acid sequences are different) in the coding region. None of these amino acid differences were found in five major conserved domains. When compared with other species, CHEF cells showed 76 and 75% homologous to human (30) and mouse (31,32) respectively.

Since p53 protein is able to activate transcription of the *mdm2* gene and elevates its mRNA level (11), we demonstrated that V79 cells do not have a functional p53 product by showing that these cells failed to induce *mdm2* gene product with camptothecin (CPT) treatment (Fig. 1). The control, CPT treated CHEF cells, generated an obvious induction of *mdm2* mRNA and is consistent with previous studies with normal human fibroblasts lines (11). This finding suggested that two point mutations found in the DNA binding domain of the p53 protein in V79 cells could possibly cause the loss of the ability to bind specific p53-binding sequences and thereby limit the ability of p53 to regulate the expression of other genes (27,29).

Our results therefore indicate that V79 cells, a widely used cell line for studies of DNA repair and DNA damage, do not express normal p53 protein and fail to show induction of *mdm2* gene product response to DNA damage. V79 cells are known to be immortal, have a shortened cell cycle, and are readily mutagenized to make stable mutant lines deficient in DNA repair enzymes and related DNA damage response functions. While these properties have made these cells extremely useful, they also raise questions about the generalizability of results obtained. As a consequence, studies of mutagenesis and related studies of DNA damage and DNA repair in these cells thus must, therefore, be interpreted with caution as a result of the disruption of normal DNA damage response pathways.

ACKNOWLEDGEMENTS

We would like to thank Dr Herman Yee for the immunohistochemistry staining and Dr John Lehman for helpful communications and discussions. This work was supported by NIH grants CA51060, AM07421 and USAMRDC 17-94-J-4416, by the Rita and Stanley Kaplan Cancer Center, the Irma T. Hirschl Trust and a gift from Charles and Helen Lazarus.

REFERENCES

1 Bradley, M. O., Bhuyan, B., Francis, M. C., Langenbach, R., Peterson, A. and Huberman, E. (1981) *Mutation Res.*, **87**, 81–142.
2 Langenbach, R., Hix, C., Oglesby, L. and Allen, J. (1983) *Ann. N.Y. Acad. Sci.*, **407**, 258–266.
3 Peterson, A. R., Danenberg, P. V., Ibric, L. L. and Peterson, H. (1985) *Basic Life Sci.*, **31**, 313–334.
4 Jacobson-Kram, D. (1986) *Environ. Mutagenesis*, **8**, 161–169.
5 Boorstein, R. J., Levy, D. and Teebor, G. W. (1987) *Cancer Res.*, **47**, 4372–4377.
6 Boorstein, R. J., Chiu, L. N. and Teebor, G. W. (1992) *Mol. Cell. Biol.*, **12**, 5536–5540.
7 Whitacre, C. M., Hashimoto, H., Tsai, M. L., Chatterjee, S., Berger, S. J. and Berger, N. A. (1995) *Cancer Res.*, **55**, 3697–3701.
8 Lane, D. P. (1992) *Nature (London)*, **358**, 15–16.
9 Oren, M. (1992) *FASEB J.*, **6**, 3169–3176.
10 Barak, Y. and Oren, M. (1992) *EMBO J.*, **11**, 2115–2121.
11 Price, B. D. and Park, S. J. (1994) *Cancer Res.*, **54**, 896–899.
12 Smith, M. L., Chen, I. T., Zhan, Q., Bae, I., Chen, C. Y., Gilmer, T. M., Kastan, M. B., O'Connor, P. M. and Fornace Jr, A. J. (1994) *Science*, **266**, 1376–1380.

13 Lowe, S. W., Schmitt, E. M., Smith, S. W., Osborne, B. A. and Jacks, T. (1993) *Nature*, **362**, 847–849.
14 Merritt, A. J., Potten, C. S., Kemp, C. J., Hickman, J. A., Balmain, A., Lane, D. P. and Hall, P. A. (1994) *Cancer Res.*, **54**, 614–617.
15 Wolf, D. and Rotter, V. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 790–794.
16 Appleman, L. J., Uyeki, J. and Frey, A. B. (1995) *Int. J. Cancer*, **61**, 887–894.
17 Vojtesek, B., Bartek, J., Midgley, C. A. and Lane, D. P. (1992) *J. Immunol. Methods*, **151**, 237–244.
18 Kline, E., Chiang, S. J., Lattora, D. and Chaung, W. (1992) *J. Biochem.* **111**, 168–174.
19 Legros, Y., McIntyre, P. and Soussi, T. (1992) *Gene*, **112**, 247–250.
20 Chaung, W. and Boorstein, R. J. (1997) *Mutation Res.*, **373**, 125–137.
21 Higgins, D. G., Bleasby, A. J. and Fuchs, R. (1992) *Comp. Appl. Biosci.*, **8**, 189–191.
22 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) *Nucleic Acids Res.*, **22**, 4673–4680.
23 Mi, L.-J., Chaung, W., Horowitz, R., Teebor, G. W. and Boorstein, R. J. manuscript in preparation.
24 Kastan, M. B., Onyekwere, O., Sidransky, D., Bogelstein, B. and Craig, R. W. (1991) *Cancer Res.*, **51,** 6304–6311.
25 Lee, S., Elenbaas, B., Levine, A. and Griffith, J. (1995) *Cell*, **81**, 1013–1020.
26 Moro, F., Ottaggio, L., Bonatti, S., Simili, M., Miele, M., Bozzo, S. and Abbondandolo, A. (1995) *Carcinogenesis*, **16,** 2435–2440.
27 Crook, T., Marston, N. J., Sara, E. A. and Vousden, K. H. (1994) *Cell*, **79**, 817–827.
28 Soussi, T., Caron de Fromentel, C. and May, P. (1990) *Oncogene*, **5**, 945–952.
29 Vogelstein, B. and Kinzler, K. W. (1992) *Cell*, **70**, 523–526.
30 Harlow, E., Williamson, N. M., Ralston, R., Helfman, D. M. and Adams, T. E. (1985) *Mol. Cell. Biol.*, **5**, 1601–1610.
31 Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S. and Givol, D. (1983) *Nature*, **306**, 594–597.
32 Pennica, D., Goeddel, D. V., Hayflick, J. S., Reich, N. C., Anderson, C. W. and Levine, A. J. (1984) *Virology*, **134**, 477–482.