CHUGOKUGAKUEN J. 2006
Vol. 5, pp. 23-27
Copyright© 2006 by Chugokugakuen

CHUGOKUGAKUEN
Journal

http://www.cjc.ac.jp/

Original Article

cDNA Cloning and Analysis of the Chicken Homolog of E.coli Endonuclease Ⅲ

**Shuji Seki[1], Altaf H. Sarker[2], Takashi Nakamura[3]
Yuichi Seki[4] and Shogo Ikeda[5]**

[1] *Department of Human Nutrition, Faculty of Contemporary Life Science, Chugokugakuen University, Okayama 700-0197, Japan,*
[2] *Department of Cellular and Molecular Biology, Life Science Division, Lawrence Berkeley National Laboratory, University of California, California, USA,*
[3] *Department of Molecular Biology, Okayama University Graduate School of Medicine and Dentistry, Okayama 700-8558, Japan,*
[4] *Tsuyama Medical Examination Center of Chugoku Occupational Health Association, Tsuyama 708-0001, Japan,*
[5] *Department of Biological Chemistry, Okayama University of Science, Okayama 700-0005, Japan*

We have cloned and analyzed the cDNA for the chicken homolog of *E. coli* endonuclease Ⅲ (chNthl1; chicken homolog of *E. coli* endonuclease Ⅲ-like 1). The cDNA was isolated from a chicken embryonic cDNA library using cDNA of mouse endonuclease Ⅲ (mNthl1) as a probe. The obtained cDNA was 1211 nucleotide long encoding a protein consisting of 281 amino acids with a predicted molecular mass of 32 kDa. The amino acid sequence exhibits 66% and 67% homologies with those of the human and mouse homologs, respectively, and the homologies are mostly concentrated within the C-terminal, 2/3 region of the sequence. The two motifs consisting of a helix-hairpin-helix and an iron-sulphur [4Fe-4S] cluster were also preserved in the chicken homolog suggesting similar enzymatic function in chicken cells as that of endonuclease Ⅲ in *E. coli.*

Key Words: chicken Nthl1, cDNA cloning, endonuclease Ⅲ homolog, chicken

Introduction

Reactive oxygen species generated by ionizing radiation and also as a consequence of normal metabolism induce DNA lesions that are removed by ubiquitous DNA repair systems [1]. Among the DNA base-lesions thymine residue is especially susceptible to oxidative DNA damage and the major detectable product is thymine-glycol [2]. *E. coli* endonuclease Ⅲ (EndoⅢ), a bifunctional repair enzyme, initiates repair of this lesion by its DNA glycosylase activity which is followed by its AP lyase activity. EndoⅢ was also reported to be involved in the repair of other damaged pyrimidine derivatives such as 5, 6-dihydrothymine, 5-hydroxy dihydrothymine, 5-hydroxycytosine, 5-hydroxyuracil, uracil glycol, beta-ureidobutyric acid, 5-hydroxy-5-methyl hydantoin and urea [1, 3, 4] which are cytotoxic or mutagenic, otherwise they are repaired properly [1, 3, 4, 5].

Enzymes functionally and structurally homologous to endoⅢ have been reported from both prokaryotic and eukaryotic cells [1, 6, 7, 8, 9]. Previously we have cloned and characterized the mouse (mNthl1) and human (hNTH1) homologs of endoⅢ and determined their genomic organization [10, 11, 12].

[1] Corresponding author.
Shuji Seki
Department of Human Nutrition, Faculty of Contemporary Life Science, Chugokugakuen University, Okayama 700-0197, Japan

In the present paper we report the cloning of the chicken homolog (chNthl1) using mNthl1 cDNA as a probe and the sequence analysis. Studies on the repair enzyme using chicken cells may be valuable not only for understanding the structure and function of the enzyme in chicken cells, but also for studying its general biological function by using homologous recombination-proficient chicken B cell lines, and to examine the phenotype caused by the disruption of chNthl1 gene.

Materials and Methods

Materials

The reagents used in the present experiments were obtained as previously described [11]. The chicken Embryo cDNA Library (10-day White Leghorn Embryo fibroblast) was obtained from Clontech Laboratories Japan Ltd., Tokyo, Japan, and Chick Embryo Lambda cDNA Library (#937405; 5-day-old chick embryos) was obtained from Stratagene, La Jolla, CA 92037 USA. Mouse Nthl1 cDNA (accession no. AB006812) was prepared as previously described [11].

```
chNthl1 M-CAAAPRGGGRA-ARRLGAATAG---SR--VPSA-APR--YS--R-RTR-RV--P-------IAY-E----AEP----K 48
hNTH1   M-TALSARMLTRSRSLGPGAGPRG---CR-EEPGPLRRREAAAEAR-KSHSPVKRPRKAQRLRVAY-E-GSDSEKGE--G 70
mNthl1  M-NS-GVRMVTRSRSRATRIASEG---CR-EE---LAPREAAAEGR-KSHRPVRHPRRTQKTHVAY-E-AANGEEGE--D 66
S. cer  MREESRSRKRKHIPVDIEEVEVRSKYFKKNERTVEL-VKENKINKDLQNYGGVNIDWIKALKPIEYFEWIESRTCDDPRT 79

chNthl1 PE-SL--GP-KWEPENW-QQQ-L---ERIREMRRHRDAPVDEMG--VDKCY--DTSA-PP-QV-MR-Y--QVLLSLMLSS 109
hNTH1   AE-PL-KVP-VWEPQDW-QQQ-L---VNIRAMRNKKDAPVDHLG--TEHCY--DSSA-PP-KV-RR-Y--QVLLSLMLSS 132
mNthl1  AE-PL-KVP-VWEPQNW-QQQ-L---ANIRIMRSKKDAPVDQLG--AEHCY--DASA-PP-KV-RR-Y--QVLLSLMLSS 128
S. cer  WGRPITKEEMINDSGAKVPESFLPIYNRVRLMRSKVKTPVDAMGCSMIPVLVSNKCGIPSEKVDPKNFRLQFLIGTMLSA 159

chNthl1 QTKD----Q----VTSAAMLRLR-QRGLTVDSILQMDDATLGQIIYPVGFWRNKVKYIKQTTAILKQKYGGDIPGTVEEL 180
hNTH1   QTKD----Q----VTAGAMQRLR-ARGLTVDSILQTDDATLGKLIYPVGFWRSKVKYIKQTSAILQQHYGGDIPASVAEL 203
mNthl1  QTKD----Q----VTAGAMQRLR-ARGLTVESILQTDDDTLGRLIYPVGFWRNKVKYIKQTTAILQQRYEGDIPASVAEL 199
S. cer  QTRDERMAQAALNITEYCLNTLKIAEGITLDGLLKIDEPVLANLIRCVSFYTRKANFIKRTAQLLVDNFDSDIPYDIEGI 239

        Helix-hairpin-Helix
chNthl1 VKLPGVGPKMAHLAMNIAWNSVSGIAVDTHVHRITNRLKWVKKET-RYPEETRVALEDWLPRDLWREINWLLVGFGQQTC 259
hNTH1   VALPGVGPKMAHLAMAVAWGTVSGIAVDTHVHRIANRLRWTKKAT-KSPEETRAALEEWLPRELWHEINGLLVGFGQQTC 282
mNthl1  VALPGVGPKMAHLAMAVAWGTISGIAVDTHVHRIANRLRWTKKMT-KTPEETRKNLEEWLPRVLWSEVNGLLVGFGQQIC 278
S. cer  LSLPGVGPKMGYLTLQKGWGLIAGICVDVHVHRLCKMWNWVDPIKCKTAEHTRKELQVWLPHSLWYEINTVLVGFGQLIC 319
                 o                                                                      *

          4Fe-4S
chNthl1 LPVNPRCKECLNQDIC--------------------------P----------AA---KRF            281
hNTH1   LPVHPRCHACLNQALC--------------------------P----------AA---QGL            304
mNthl1  LPVHPRCQACLNKALC--------------------------P----------AA---QDL            300
S. cer  MARGKRCDLCLANDVCNARNEKLIESSKFHQLEDKEDIEKVYSHWLDTVTNGITTERHKKK            380
              *   *    *
```

Fig. 1. Alignment of the amino acid sequence of chNth1 with those of the eukaryotic homologs. The amino acid sequence of chicken endonuclease III homolog (chNthl1) is aligned with homologous sequences (abbreviated name and accession number in brackets) from human (hNTH1; AB001575), mouse (mNthl1; AB009371) and S. cerevisiae (Sec; Z74785). Amino acids that are identical in all the four of the sequences are shaded in black. Numbers in the right column refer to the last amino acid residue in each line of the respective protein sequence. Solid lines indicate the helix-hairpin-helix motif and the four-cysteine [4Fe-4S] motif. The proposed active site lysine residue is indicated by an open circle and the four cysteine residues responsible for [4Fe-4S] motif are indicated by star.

HhH motif [4Fe-4S] binding

chicken

Human

E.coli

S. cerevisiae (Ntg1)

S. cerevisiae (Ntg2)

S. pombe

C. ele.

100 aa

HhH motif

```
chicken                 177:VEELVKL PGVGP KMAHLAMNIA WNSVSGIAVDTHV:211
Human                   200:VAELVAL PGVGP KMAHLAMAVA WGTVSGIAVDTHV:234
E. coli                 108:RAALEAL PGVGR KTANVVLNTA -FGWPTIAVDTHI:141
S. cerevisiae(Ntg1)     231:INELLGL PGVGP KMAYLTLQKA WGKIEGICVDVHV:265
S. cerevisiae(Ntg2)     236:IEGILSL GGVGP KMGYLTLQKG WGLIAGICVDVHV:270
S. pombe                130:VEDLMTL PGVGP KMGYLCMSIA WNKTVGIGVDVHV:164
C. elegans              107:LDGLCAL PGVGP KMANLVMQIA WGECVGIAVDTHV:141
                            helix   hairpin  helix
```

[4Fe-4S] binding

```
chicken                 259:CLPVNPRCKEC-L-NQDIC:275
Human                   282:CLPVHPRCHAC-L-NQALC:298
E. coli                 187:CIARKPRCGSC-I-IEDLC:203
S. cerevisiae(Ntg2)     317:CMARGKRCDLC-L-ANDVC:333
S. pombe                210:CLPRGRRCDMCTLSSKGLC:228
C. elegans              187:CQPVRPKCGTC-L-CRFTC:203
                            *      *  *     *
```

Cys

Fe S S Fe Fe Fe S

Cy s

Cy s

Cy s

Fig. 2. Schematic comparison of amino acid sequences among E. coli endonuclease Ⅲ and its eukaryotic homologs, and sequence comparisons of HhH motif and four cysteine [4Fe-4S] motif among them.

Cloning of the chNthl1 cDNA

Chicken Embryo cDNA Library (Clontech) was screened using ^{32}P-labeled mouse Nthl1 cDNA as a probe with low stringent hybridization condition. mNthl1 cDNA was labelled with [α-^{32}P] dCTP (3000 Ci/mmol) using the random priming method. Hybridization was carried out at 42℃ overnight in 6x SSC (1x SSC: 150mM NaCl, and 15mM sodium citrate, pH 7.0), 5x Denhardt's solution, 0.5% SDS, denatured salmon testes DNA at 100 μg/ml and the labeled cDNA probe at 4.5 x 10^5cpm/ml. After washing at room temperature for 10min in 2x SSC and 0.1 % SDS and then twice at 45℃ for 1h in 1x SSC and 0.1% SDS, the membrane was air-dried, and exposed on Kodak X-Ray film for autoradiography. One positive clone was obtained from 5x 10^5 independent plaques. The cDNA obtained was a defective one lacking a 3'-region. The full-length cDNA was obtained by recloning of another embryonic cDNA library (Stratagene) using the defective cDNA as a probe. The inserts in the Uni-ZAP XR vector in the isolated bacteriophage clone were excised with helper phage infection and recircularized to generate a subclone in the pBluescript SK (−) phagemid vector. The nucleotide sequence of the insert of the subclone was determined by the dideoxy chain termination method [13] using either M13 universal primers or oligonucleotide primers synthesized according to the sequences determined using AmpliTaq FS Ready Mix (Perkin-Elmer) by an Applied Biosystem Model 373A DNA sequencer. The isolated full-length cDNA (designated as chNthl1 cDNA) was 1211 bp long and contained an open reading frame of 843 bp. The nucleotide sequence data reported in this paper have been submitted to the DDBJ/GenBank/EMBL Data Bank under the accession number AB284187.

Results and Discussion

Cloning and characterization of the chNthl1 cDNA

Eighteen positive colonies were picked up by screening the chicken embryonic cDNA library (Stratagene) with the defective clone as described in Materials and Methods. Among them, at least four independent clones were present. Two of the four were defective clones, one was a splice variant (lack of exon 5) and the remainder was the clone containing an entire coding region. Sequencing of the latter clone showed that the cDNA inserted in the clone was 1211 nucleotide long with an open reading frame of 843bp encoding 281 amino acids. The cDNA terminated with a poly (A) tail, 20 nucleotides downstream from an authentic polyadenylation signal, AATAAA [14]. The chNthl1 protein encoded by the cDNA has a calculated molecular weight of 31,900 and a calculated pI of 10.07.

Analysis of the deduced amino acid sequence of chNthl1

The deduced amino acid sequence of the chNthl1 was compared with those of human, mouse and Saccharomyces cereviciae homologs of *E. coli* endonucleaseⅢ (Fig. 1). The homologies of the amino acid sequence deduced from the chNthl1 with that of *E. coli* endonucleaseⅢ (accession no. J02857) and those of the human, mouse and Saccharomyces cereviciae homologs (accession nos. AB001575, AB006812 and Z74785, respectively) are 29% identity/178 amino acids, 66% identity/285 amino acids, 67% identity/269 amino acids and 37% identity/240 amino acids, respectively. The homologies are mostly concentrated within the C-terminal, 2/3 part of the sequence with the exception of the Saccharomyces cereviciae homolog.

Genomic cloning of chNthl1 gene was performed (unpublished results). Comparing the genomic sequence of chNthl1 and cDNA sequences of the other cDNA clones with the present chNthl1 cDNA, several sequence variances were observed. In order to tentatively assign a base position within the cDNA, the adenine residue in the translation initiation codon of the chNthl1 cDNA was designated as position +1 and the 5' neighboring residue cytosine as position −1. With this as a starting point, positive numbers are given to positions downstream from the starting point and negative numbers to those upstream. The observed sequence variances are as follows: position 117, T to C (Valine, synonymous substitution); position 155, C to T (Proline to Leucine); position 864, T to C; position 1000, T to G; position 1064, A to G; position 1080, G to A. Poly (A) attached to position 1202 (T) or 1207 (A). The sequence variability in the chicken Nthl1 cDNAs seems to be high comparing with that in human NTHL1 or mouse Nthl1 cDNAs, although the reason is not clear.

The helix-hairpin-helix (HhH) motif, which is common in enzymes having DNA glycosylase activity and is involved in interaction with DNA, is preserved in chNthl1 (Figs. 1 and 2) [11, 15]. The HhH DNA binding motif is found in the amino acid sequence positions from 168 to 197. Furthermore, the motif of the HhH-GDP superfamily base excision DNA repair protein, which is present in a diverse range of structurally related DNA repair proteins, is found in positions from 103 to 241 [16]. The superfamily is called the HhH-GPD family after its hallmark Helix-hairpin-helix and Gly/Pro rich loop followed by a conserved aspartate [17]. The four cysteine residue motif (Cys-X_6-Cys-X_2-Cys-X_5-Cys) that constitutes the iron/sulphur(4Fe-4S) cluster is well preserved in the chicken homolog at the C-terminal region (from positions 259 to 275) (Figs. 1 and 2). The key residues responsible for the formation of enzyme-substrate intermediate (Lys212 for hNTHL1 and Lys120 for endoⅢ) and involved in catalysis (Asp231 for hNTHL1 and Asp138 for endoⅢ) are well preserved in the chicken homolog in positions 189 and 208, respectively, suggesting that the chicken homolog has an endonuclease Ⅲ-like function [10].

Studies using chicken cells on the repair enzyme may be valuable not only for understanding the structure and function of the enzyme in chicken cells, but also for studying its general biological function by using homologous recombination-proficient chicken B cell lines, and to examine the phenotype caused by the disruption of chNthl1 gene.

Acknowledgements and notes. This work had been conducted largely in Department of Molecular Biology, Institute of Cellular and Molecular Biology, Okayama University Medical School, Okayama 700-8558, Japan. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture and in part by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports and Culture of Japan (Repair, Recombination and Mutagenesis, 08280101).

The nucleotide sequence data reported in this paper have been deposited with the GSDB, DDBJ, EMBL, and NCBI nucleotide databases under the Accession Number AB284187.

References

1. Friedberg EC, Walker GC and Siede W: DNA Repair and Mutagenesis. ASM Press, Washington DC (1995).
2. Cathcart R, Schwiers E, Saul RL and Ames BN: Thymine glycol and thymidine glycol in human and rat urine: a possible assay for oxidative DNA damage. Proc Natl Acad Sci USA (1984) **81**, 5633-5637.
3. Krokan HE, Standal R and Slupphaug G: DNA glycosylases in the base excision repair of DNA. Biochem J (1997) **325**, 1-16.
4. Roldan-Arjona T, Anselmino C and Lindahl T: Molecular cloning and functional analysis of a schizosaccharomyces pombe homologue of Escherichia coli endonuclease Ⅲ. Nucleic Acids Res (1996) **24**, 3307-3312.
5. Zuo S, Boorstein RJ and Teebor GW: Oxidative damage to 5-methylcytosine in DNA. Nucleic Acids Res (1995) **23**, 3239-3243.
6. Demple B and Harrison L: Repair of oxidative damage to DNA: enzymology and biology. Annu Rev Biochem (1994) **63**, 915-948.
7. Bacchetti S and Benne R: Purification and characterization of an endonuclease from calf thymus acting on irradiated DNA. Biochim Biophys Acta (1975) **390**, 285-297.
8. Kim J and Linn S: Purification and characterization of UV endonuclease I and II from murine plasmacytoma cells. J Biol Chem (1989) **264**, 2739-2745.
9. Doetsch PW, Helland DE and Haseltine WA: Mechanism of action of a mammalian DNA repair endonuclease. Biochemistry (1986) **25**, 2212-2220.
10. Ikeda S, Biswas T, Roy F, Izumi T, Boldogh I, Kurosky A, Sarker AH, Seki S, and Mitra S: Purification and characterization of a human homolog (hNTH1) of Escherichia coli endonuclease Ⅲ: Direct identification of lys-212 as the nucleophilic residue. J Biol Chem (1998) **273**, 21585-21593.
11. Sarker AH, Ikeda S, Nakano H, Terato H, Ide H, Akiyama K, Tsutsui K, Bo Z, Kubo K, Yamamoto K, Yasui A, Yoshida MC, and Seki S: Cloning and characterization of a mouse homologue (mNthl1) of Escherichia coli endonuclease Ⅲ. J Mol Biol (1998) **282**, 761-774.
12. Imai K, Sarker AH, Akiyama K, Ikeda S, Yao M, Tsutsui K, Shohmori T and Seki S: Genomic structure and sequence of a human homologue (NTHL1/NTH1) of Escherichia coli endonuclease Ⅲ with those of the adjacent parts of TSC2 and SLC9A3R2 genes. Gene (1998) **222**, 287-295.
13. Sanger F, Nicklen S and Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Nat Acad Sci USA (1977) **74**, 5463-5467.
14. Proudfoot NJ and Brownlee GG: 3'Non-coding region sequences in eukaryotic messenger RNA. Nature (1976) **263**, 211-214.
15. CDD: PF00633. 12 (CDD; Conserved Protein Domain Database): pfam entry Helix Hairpin Helix (HhH).
16. CDD: PF00730: pfam entry HhH-GPD superfamily base excision DNA repair protein.
17. Bruner SD, Norman DP and Verdine GL: Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine. Nature (2000) **403**, 859-866.

Accepted March 31, 2006.