

cDNA Cloning and Analysis of the Chicken Homolog (chApex1) of APEX Nuclease, a Multifunctional DNA Repair Enzyme

Shuji Seki¹⁾, Takashi Nakamura²⁾, Altaf H. Sarker³⁾, Yuichi Seki⁴⁾ and Shogo Ikeda⁵⁾

¹⁾Department of Human Nutrition, Faculty of Contemporary Life Science, Chugokugakuen University, Okayama 701-0197, Japan,

²⁾Department of Molecular Biology, Okayama University Graduate School of Medicine and Dentistry, Okayama 700-8558, Japan,

³⁾Department of Cellular and Molecular Biology, Life Science Division, Lawrence Berkeley National Laboratory, University of California, California, USA,

⁴⁾Tsuyama Medical Examination Center of Chugoku Occupational Health Association, Tsuyama 708-0016, Japan,

⁵⁾Department of Biological Chemistry, Okayama University of Science, Okayama 700-0005, Japan

We have cloned and analyzed the cDNA (chApex1) for the chicken homolog of APEX nuclease, a multifunctional DNA repair enzyme. Chicken Apex cDNA fragments were first amplified from Chick Embryo Lambda cDNA Library by nested PCR (polymerase chain reaction) using primers constructed on the basis of information of highly conserved regions of the amino acid sequences among mouse, rat and human APEX nuclease. Almost full-size chicken Apex cDNA was cloned from the Chick Embryo Lambda cDNA Library using the amplified chicken Apex cDNA fragments as a probe. The cloned cDNA was 1751 nucleotide long encoding a protein consisting of 300 amino acids with a predicted molecular mass of 33 kDa. The amino acid sequence exhibits a high homology with the sequence of the mammalian APEX nuclease (64%, 62% and 63% identities/300 amino acids with those of the human, mouse and rat APEX nuclease, respectively). The homologies of the amino acid sequences between chicken and mammalian APEX nucleases occurred along nearly the entire length of the sequences.

Key Words: chApex1 cDNA cloning, chicken APEX nuclease, multifunctional DNA repair enzyme, exonuclease III homolog, chicken

Introduction

AP sites resulting from loss of bases are the most frequent lesions occurring in vivo in cellular DNA [1-4]. Single-strand breaks with 3' termini blocked by nucleotide fragments are also produced by free radical pathways caused by ionizing radiation or other sources of oxygen radicals [1-4]. These lesions

must be corrected to restore genetic integrity. APEX nuclease is a multifunctional DNA repair enzyme having 5' apurinic/aprimidinic (AP) endonuclease, DNA 3' repair diesterase, 3'→5' exonuclease, and DNA 3' phosphatase activities, and is the major AP endonuclease possibly involved in repair of AP sites and single-strand DNA breaks with 3'-blocked termini [5-10]. Curran's group [11] reported a redox factor (Ref-1) which stimulates the DNA binding activity of Fos-Jun heterodimers, Jun-Jun homodimers and HeLa cell AP-1 protein as well as that of several other transcription factors including NF- κ -B, Myb and the member of the ATF/

¹⁾ Corresponding author.

Shuji Seki

Department of Human Nutrition, Faculty of Contemporary Life Science, Chugokugakuen University, 83, Niwase, Okayama 701-0197, Japan

CREB family. They cloned Ref1 cDNA, and showed that Ref1 and the major AP endonuclease (APEX nuclease, HAP1, Ape) are the identical protein [11].

Known homologs of APEX1 (APEX nuclease or APEX1, Ape protein, HAP1, Ref1) are listed in UniGene (NCBI, 12), information hyperlinked over proteins [iHOP, 13] and others.

In the present paper, we report on the cloning of the chicken homolog (chApex1 cDNA) of APEX cDNA 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 conditional disruption of chApex1 gene.

Materials and Methods

Materials

The reagents used in the present experiments were obtained as previously described [14]. Chick Embryo Lambda cDNA Library (#937405; 5-day-old chick embryos) was obtained from Stratagene, La Jolla, CA, USA.

Cloning of chicken Apex1 cDNA

Chicken Apex cDNA fragments were first amplified from Chick Embryo Lambda cDNA Library by nested PCR (polymerase chain reaction) using primers constructed on the basis of information of highly conserved regions of the amino acid sequences among mouse, rat and human APEX nuclease. Three sense-strand primers F1, F2 and F3, and 3 antisense-strand primers R1, R2 and R3 were synthesized. In the following primer sequences, W indicates an equal mixture of A and T, R a mixture of A and G, M a mixture of A and C, Y a mixture of C and T, S a mixture of C and G, H a mixture of A, T and C, N a mixture of A, C, G and T, respectively.

F1: 5'-AARATHHTGYWSNTGGAAAYGT-3'

F2: 5'-TGYTNCARGARACNAARTG-3'

F3: 5'-GARCAYGAYCARGARGGNMG-3'

R1: 5'-WNNWNWNSSNSWNWSNWS-3'

R2: 5'-WNNWNWSNWNSSNSWSS-3'

R3: 5'-SWWNWNWSWSSWNWNWSNWW-3'

Nested PCR was conducted using these primers.

Finally, amplified fragments F3/R3 were sequenced,

and the amino acid sequence deduced from the sequence of the PCR-amplified fragment showed a high homology with the mammalian amino acid sequences between F3 and R3. Almost full-length chicken Apex cDNA was cloned from the Chick Embryo Lambda cDNA Library using the amplified chicken Apex cDNA fragments as a probe.

The nucleotide sequences were determined by the dideoxy chain termination method [15] using either

```

GCCGACTCCGGTCTGTGTGTGCTCACTTCTGTCTTCTCCTGCCCGCCGCCATGTT    60
GTAGTTTTCAGCTCCACTTCCGTTTTCCAGGCCTCGCCCGCCTTCGCCACATCCGGT    120
CTTCACTCCGCTCACTTCCGTTCTCGACAGCACTTCCGGCGGTGGGACGCACGAGGGA    180
GGCGCGCGGATGCCGAAACGGAGCAAAAAGGGCGAAGATGGCGAGGCAGAGGTGGTGT    240
      M P K R S K K G E D G E A E V V S    17
GGCGAAGTCCCCCGCGAGGCGCGCCCCATACGTGACCCCGCTGGCGGAAGAGAC    300
A K S P R E A A A P Y V D P P V R E E T    37
GGCCGACGGCGCCCTACAACCTTTAAGGTCACTCATGGAACGTGGATGGATGGATGAGG    360
A D G R P Y N F K V T S W N V D G L R A    57
ATGGTCCGAAAGGGGGCTGCAGTGGCTGCAGTCGGAGCCCGGACGTTGTTTGCCT    420
W V R K G G L Q W L Q S E A P D V V C L    77
TCAGGAGACCAATGTGGGGCGGAATCGATCCCATCTGAGCTGCCCACTGCCCATCT    480
Q E T K C G A E S I P S E L S Q L P H L    97
GCCCATAAAGTTTGGGGTCCGCCGTGGGGCGTTCGGGTACACGGGGTGGGACTGCT    540
P H K F W G S A V G R S G Y S G V G L L    117
CAGCGCACCCGCCCATCCGCGTCAACCACGGCATCGGCATAGAGGAGCAGCAGCCGA    600
S R T A P I R V T H G I G I E E H D A E    137
AGGCCGGGTCTGACGGCCGAATCCCTCCGCTACGTGGTCAAGCCTACGTCGCCAA    660
G R V L T A E F P S V Y V V S A Y V P N    157
TTCCGGTCCGGTCTCAACCGCTCCAATACCGCAACGTTGGGACGCTGCCTTAAATC    720
S G R G L N R L Q Y R Q R W D G A F K S    177
CTTCTGCAACGTTTGGACGCCAAAAACCGTCTGCTCCTGCGGGGACCTCAACGTGGC    780
F L Q R L D A Q K P V V L C G D L N V A    197
CCACCGTGAATCGACCTCAGGAACCCCAAGCAACGGAGTCCCGCGATTACCCCA    840
H R E I D L R N P K S N R R S P G F T Q    217
AGAAGAACGGGACGCTTTCGGGGCCCTTTGGATGGGGGTTTTGGATTCCCTCCGCT    900
E E R D A F G A L L D G G L F L S F R L    237
CCTTTACCCGACGTCGCCAACGCTTACACCTTTTGGACCTATATGGGGGCGCCCGGA    960
L Y P D V P N A Y T F W T Y M G G A R E    257
GGCAACGTGGGCTGGAGGTTGATTATTTCTCCTCTCCACGCTACTCCGGGAGCGCT    1020
R N V F R L D Y F L L S T R L E A L    277
GTGCACTCAAGATCCGCTCGCGGCCATGGGAAGCGACCACTGCCCATCACGCTCTA    1080
C D S K I R S A A M G S D H C P I T L Y    297
CCTGGCCATAGGGCCCTAGAGAGGGCGATCCCGCCAAAATCCCATAGCGGGACCCG    1140
L A L *    300
AAATCCATTACGTGGCTCGAAATCCCATATCTAGTCCCAAAATCCCATATCTAGACC    1200
CAAAATCCCATACGTGGCTCAAAATCCCATTAATGGGACCCAAAATCCCATACGTGGC    1260
TCGAAATCCCATACGTGGCTCGAAATCCCATTAACGTACCCAAAACCCCACTAACAAGA    1320
CCCAAAATCCCATATGTGGCTCGAAATCCCATACATAGACCCAAAATCCCATCCGCG    1380
GCCATGGGAAGCGACCACTGCCCATCACGCTCTACCTGGCCCTATAGGGCCCTATAGGG    1440
GGCGATCCCGCCAAAATCCCATAGCGGGACCCGAAATCCCATTAATGTCGCTGAAAT    1500
CCCATACGTGGCTCGAAATCCCATACGTAGACCCAAAATCCCATACGTGGCTCGAAA    1560
TCCCATTAACGTACCCAAAATCCCATACGTAGACCTGAAATCCCATTAATGTACCCCAA    1620
ATCCCATGCTAGCCAAAATCCCATGACCACAACCCAAAATCCCGTTACGTAGATC    1680
CAAAATCCCATACATGGCTGGAAATCCCACTACGTAGACCCAAAATCCCAATAAGTGA    1740
CCCAAGACCC (A) 34    1751

```

Fig. 1 Nucleotide sequence of the cDNA for the chicken homolog (chApex1) of APEX nuclease. The deduced amino acid sequence in the standard one-letter code is shown beneath the nucleotide sequence. The translation termination codon and polyadenylation signal are indicated with an asterisk and single underlining, respectively.

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 chApex1 cDNA) was 1751 nucleotide long and contained an open reading frame of 900 nucleotides. The nucleotide sequence data reported in this paper have been submitted to the DDBJ/GenBank/EMBL Data Bank under the accession number AB450280.

Results and Discussion

Cloning and characterization of the chApex1 cDNA

Fourteen positive colonies were picked up by screening the chicken embryonic cDNA library (Stratagene) with the PCR-amplified fragments, and all of them were sequenced. Among them eight clones were independent clones, and the clone 5 was

selected for extensive sequencing. Sequencing of the clone 5 showed that the cDNA inserted was 1751 nucleotide long with an open reading frame of 900 nucleotides with 190 nucleotides of 5'-flanking (untranslated) and 661 nucleotides of 3'-flanking regions (Fig. 1). The cDNA was terminated with a poly(A) tail, 16 nucleotides downstream of an authentic polyadenylation signal, AATAAA [16].

Analysis of the deduced amino acid sequence of chApex1

The cDNA, designated as chApex1 cDNA, contained an open reading frame encoding 300 amino acids with a predicted molecular mass of 33 kDa. The deduced amino acid sequence of chicken APEX nuclease was compared with those of *E. coli* exonuclease III [18], ExoA protein of *S. pneumoniae* [19] and some of mammalian APEX nuclease. Amino acid sequence homology between chicken APEX nuclease and exonuclease III occurs from the 45th amino acid of the former and the first amino acid of the latter

chApex1	MPKRSKKGED	GAEVVSAKS	PRAAAPYVD	PPVREETADG	RPYNFKVTSW	NVDGLRAWVR	60'
						* . * . * . * . * . * .	
ExoIII					MKFVSF	NINGLRARPH	16"
chApex1	KGGLQWLQSE	APDVVCLQET	KCGAESIPSE	LSQLPHLPHK	FWGSAVGRSG	YSGVGLLSRT	120'
 * * * .	* * * * * .	* *	*	* * . * * . * * . * * . . .	
ExoIII	QLEAI-VEKH	QPDVIGLQET	KVHDDMFP--	LEEVAKLGYN	VF--YHGQKG	HYGVALLTKE	71"
chApex1	APIRVTHGIG	IEEHDAEGRV	LTAEFPS---	-VYVVSAYVP	NSGRGLNRLQ	Y--RQRWDGA	174'
	. * * . * * .	* . * * . * * .	* * . . . *	
ExoIII	TPIAVRRGFP	GDDEEAQRR	IMAEIPSLLG	NVTVINGYFP	QGESRDHP	IK FPAKAQFYQN	131"
chApex1	FKSFLQ-RLD	AQKPVVLCGD	LNVAHRE--I	DLRNPKSNR-	---RSPGFTQ	EERDAFGALL	227'
 * .	* . . * * .	* * . * . . .	* * .	. . * .	* * * . . . * .	
ExoIII	LQNYLETELK	RDNPVLMGD	MNISPTDLDI	GIGEENRKRW	LRTGKCSFLP	EEREWMDRLM	191"
chApex1	DGGFLDSFRL	LYPDVPNAYT	FWTYMGGARE	RNVGWRLDYF	LLSTRL----	REALCDSKIR	283'
	. * . * . * * .	* *	* * * * .	* * * . * . * .	* . * . * * .	
ExoIII	SWGLVDTRFR	ANPQTADRF	SWFDYRSKGF	DNRGLRIDLL	LASQPLAEC	VETGIDYEIR	251"
chApex1	SAAMGSDHCP	ITLYLAL					300'
	* . . * * * .	* .					
ExoIII	SMEKPSDHAP	VWATFRR					268'

Fig. 2 Comparison of chicken Apex nuclease and *E. coli* exonuclease III sequences. The amino acid sequences of chicken APEX nuclease (upper; chApex1) and exonuclease III (lower; ExoIII) are aligned to give maximum correspondence. Dashes indicate gaps produced by this alignment in one or the other sequence. Exact matches and similar amino acids are indicated with asterisks and dots, respectively.

(27.8% identity/266 amino acids) (Fig. 2). The amino acid sequence of chicken APEX nuclease exhibits a high homology with the sequence of mammalian APEX nuclease (64%, 62% and 63% identities/300 amino acids with those of the human, mouse and rat APEX nuclease, respectively) (Fig. 3). The amino acid sequence homology occurred along nearly the entire length of the sequences, although the homology is somewhat low in N-terminal 44 amino acids. The region between the 45th amino acid (phenylalanine) and the carboxyl terminal (leucine) of chicken APEX nuclease corresponds to the XthA region of mammalian APEX nuclease (refer to the accession number NP_033817) [8]. The probable

metal binding sites of APEX nuclease [17] are preserved in the chicken APEX nuclease at positions 51 (asparagine), 79 (glutamic acid), 193 (aspartic acid), 195 (asparagine), 290 (aspartic acid) and 291 (histidine). The possible proton acceptor site and possibly important site for substrate are also preserved at positions 291 (histidine) and 195 (asparagine) respectively.

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

chApex	1	MPKRSKKG---EDGE---AEVVSAKS-----PREAAA--P--YMDPPVREETADGRPY	43
hAPEX	1	MPKRGGKGAVAEDGDELRTPEPAKSKKTAAKKNDKEAAGEGPALYEDPPDHKTSPSGKPA	60
mApex	1	MPKRGGKAAA-DDGEEPKESEPETKSKSGAAKTEKEAAGEGPVLYEDPPDQKTSPSGKSA	59
rApex	1	MPKRGRAAA-EDGEEPKSEPETKSKSGAAKTEKEAAGEGPVLYEDPPDQKTSASGKSA	59
chApex	44	NEKVTISWNVDGLRAWVRKGGLOWLQSEAPDVVCLQETKCGAESIPSEISQLPHLPHKFWG	103
hAPEX	61	TLKICSWNVDGLRAWIKKKGLDWVKEAPDILCLQETKCSENKLPALQELPGLSHQYWS	120
mApex	60	TLKICSWNVDGLRAWIKKKGLDWVKEAPDILCLQETKCSENKLPALQELPGLTHQYWS	119
rApex	60	TLKICSWNVDGLRAWIKKKGLDWVKEAPDILCLQETKCSENKLPALQELPGLTHQYWS	119
chApex	104	SAVGRSGYSGVGLLSRTAPIRVTHGIGIEEHDAEGRVLTAEFPSVYVVSAYVPNSGRGLN	163
hAPEX	121	APSDKEGYSGVGLLSRQCPLKVSYGIGDEEHDOEGRVIVAEFDSFVLMTAYVPNAGRGLV	180
mApex	120	APSDKEGYSGVGLLSRQCPLKVSYGIGEEHDOEGRVIVAEFESFVLMTAYVPNAGRGLV	179
rApex	120	APSDKEGYSGVGLLSRQCPLKVSYGIGEEHDOEGRVIVAEFESFILMTAYVPNAGRGLV	179
chApex	164	RLEYRQRWDGAFKSFQRDAQKPVVLCGDLNVAHREIDLNPKSNRRSPGFQTQERDAF	223
hAPEX	181	RLEYRQRWDEAFKFLKGLASRKPLVLCGDLNVAHEEIDLNPKGNKKNAGFTTQERQGF	240
mApex	180	RLEYRQRWDEAFKFLKDLASRKPLVLCGDLNVAHEEIDLNPKGNKKNAGFTTQERQGF	239
rApex	180	RLEYRQRWDEAFKFLKDLASRKPLVLCGDLNVAHEEIDLNPKGNKKNAGFTTQERQGF	239
chApex	224	GALLDGGFLLDSFRLLYPDVPNAYTFWTYMGGARERNVWRLDYFLLSSTRIREALCDSKI	282
hAPEX	241	GELLQAVPLADSFRLYPNTPYAYTFWTYMNRASKNVWRLDYFLLSHSLLPALCDSKI	300
mApex	240	GELLQAVPLADSFRLYPNTAYAYTFWTYMNRASKNVWRLDYFLLSHSLLPALCDSKI	299
rApex	240	GEMLQAVPLADSFRLYPNTAYAYTFWTYMNRASKNVWRLDYFLLSHSLLPALCDSKI	299
chApex	283	RSAAAGSDHCPITLYLAI	300
hAPEX	301	RSKALGSDHCPITLYLAI	318
mApex	300	RSKALGSDHCPITLYLAI	317
rApex	300	RSKALGSDHCPITLYLAI	317

Fig. 3 Alignment of the amino acid sequence of chicken APEX nuclease with those of the eukaryotic homologs. The amino acid sequence of chicken APEX nuclease (chApex; chApex1) is aligned with homologous sequences (abbreviated name and accession number in brackets) from human (hAPEX; hAPEX1; D90373), mouse (mApex; mApex1; D90374) and rat (rApex; rApex1; D44495). Amino acids that are identical in all the four of the sequences are surrounded. Numbers in the right column refer to the last amino acid residue in each line of the respective protein sequence.

the conditional disruption of chApex1 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, Japan. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan 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 AB450280.

References

1. Friedberg EC, Walker GC and Siede W: DNA Repair and Mutagenesis. ASM Press, Washington DC (1995).
2. Sancar A and Sancar GB: Annu Rev. Biochem. (1988) **57**, 29–67.
3. Lindahl T (1982) Annu. Rev Biochem **51**, 61–87.
4. Kunkel TA, Shearman CW and Loeb LA: Nature (1981) **291**, 349–351.
5. Seki S and Oda T: An exonuclease possibly involved in the initiation of repair of bleomycin-damaged DNA in mouse ascites sarcoma cells. Carcinogenesis (1988) **9**, 2239–2244.
6. Seki S, Ikeda S, Watanabe S, Hatsushika M, Tsutsui K, Akiyama K and Zhang B: A mouse DNA repair enzyme (APEX nuclease) having exonuclease and apurinic/aprimidinic endonuclease activities: purification and characterization. Biochim Biophys Acta (1991) **1079**, 57–64.
7. Seki S, Akiyama K, Watanabe S, Hatsushika M, Ikeda S and Tsutsui K: cDNA and deduced amino acid sequence of a mouse DNA repair enzyme (APEX nuclease) with significant homology to *Escherichia coli* exonuclease III. J Biol Chem (1991) **266**, 20797–20802.
8. Seki S, Hatsushika M, Watanabe S, Akiyama K, Nagao K and Tsutsui K: cDNA cloning, sequencing, expression and possible domain structure of human APEX nuclease homologous to *Escherichia coli* exonuclease III. Biochim Biophys Acta (1992) **1131**, 287–299.
9. OMIN (Online Mendelian Inheritance in Man). APEX nuclease. Accession No. 107748, 2007.
10. MGI (Mouse genome informatics) ID. MGI: 88042. Apex1. 2007.
11. Xanthoudakis S, Miao G, Wang F, Pan EYC and Curran T: Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme, EMBO J (1992) **11**, 3323–3335.
12. UniGene (NCBI) Hs. 73722. APEX nuclease (multifunctional DNA repair enzyme) 1 (APEX1), 2008.
13. iHOP-Information Hyperlinked over Proteins. (iHOP provides the network of genes and proteins as a natural way of accessing the millions of abstracts in PubMed.), 2008.
14. 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 (mNth1) of *Escherichia coli* endonuclease III. J Mol Biol (1998) **282**, 761–774.
15. Sanger F, Nicklen S and Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Nat Acad Sci USA (1977) **74**, 5463–5467.
16. Proudfoot NJ and Brownlee GG: Nature (1976) **263**, 211–214.
17. UniProtKB/Swiss-Prot entry P27695.
18. Saporito SM, Smith-White BJ and Cunningham RP: Nucleotide sequence of the xth gene *Escherichia coli* K-12. J Bacteriol (1988) **170**, 4542–4547.
19. Puyet A, Greenberg B and Lacks SA: The *exoA* gene of *Streptococcus pneumoniae* and its product, a DNA exonuclease with apurinic endonuclease activity. J Bacteriol (1989) **171**, 2278–2286.

Accepted March 30, 2008.