Table 4. (continued)

Risk for Low, Mid, and High Proportion of Teeth With CAL ≥5 mm According to Alcohol Consumption and Other Variables

		Model 3	
Teeth With CAL ≥5 mm		Univariate	Multivariate
None	High	OR (95% CI)	OR (95% CI)
394	47	1	1
106	10	0.8 (0.4-1.6)	0.6 (0.3-1.4)
57	15	2.2 (1.2-4.2)†	1.4 (0.6-3.3)
67	24	3.0 (1.7-5.2)‡	1.2 (0.5-2.6)
458	38	1	1
53	8	1.8 (0.8-4.1)	1.9 (0.7-5.1)
70	18	3.1 (1.7-5.7)‡	2.8 (1.2-6.8)*
43	32	9.0 (5.1-15.8)‡	4.9 (1.9-12.2)†
461	60	1	1
114	14	0.9 (0.5-1.7)	0.6 (0.3-1.1)
49	22	3.5 (2.0-6.1)‡	2.0 (1.0-3.9)*
206	61	1	1
418	35	0.3 (0.2-0.4)‡	1.0 (0.4-2.3)
		1.1 (1.0-1.1)‡	1.0 (1.0-1.1)
		0.8 (0.8-0.9)‡	0.9 (0.8-1.0)†
		4.3 (3.0-6.2)‡	2.9 (1.9-4.4)‡

DNA RESEARCH 12, 39–51 (2005)

Characterization of the Structure and Expression of Mouse *Itpa* Gene and its Related Sequences in the Mouse Genome

Mehrdad BEHMANESH,[1,2] Kunihiko SAKUMI,[1] Daisuke TSUCHIMOTO,[1] Kumiko TORISU,[1] Yoko OHNISHI-HONDA,[1] Derrick E. RANCOURT,[3] and Yusaku NAKABEPPU[1,*]

Division of Neurofunctional Genomics, Department of Immunobiology and Neuroscience, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan,[1] Department of Genetics, School of Sciences, Tarbiat Modarres University, Tehran, Iran,[2] and Southern Alberta Cancer Research Center, Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, Alberta, Canada, T2N 4N1[3]

(Received 20 October 2004; revised 3 December 2004)

Abstract

In the mouse genome, we found one processed *Itpa* gene-like sequence and two processed *Itpa* pseudogenes as well as the *Itpa* gene itself with introns, located on chromosome 2F3, which was isolated by a retro-recombination method. We also identified three types (A, B, C) of *Itpa* transcripts in mouse tissues. The processed *Itpa* gene-like sequence located on chromosome 2E1 has a complete open reading frame for exactly the same polypeptide as ITPA encoded by the type A transcript, with a polyadenylation signal. However, no transcribed sequence derived from the *Itpa* gene-like sequence was detectable in any of the mouse tissues examined, thus naming the sequence as *Itpa* processed pseudogene α. The type A *Itpa* mRNA, which was expressed in all mouse tissues examined, only encodes mouse ITPA polypeptide consisting of 198 amino acid residues with a capacity to hydrolyze dITP into dIMP. *Itpa* mRNA was detected in all tissues examined, and its expression is especially high in the testis, brain, and thymus. ITPA protein was mostly detected in the cytoplasm, to a lesser extent in the nuclei of neurons in the brain, and also those of hepatocytes, epithelial cells lining the bile duct, and endothelial cells lining the portal vein in the liver.

Key words: ITPase; ITP; XTP; pseudogene; retro-recombination; oxidative deamination

1. Introduction

The accumulation of modified or damaged bases in genomic DNA is a major cause of altered genetic information that results in mutagenesis or even programmed cell death.[1] It has been established that such damaged bases in genomic DNA arise from two independent pathways: one is a consequence of the direct modification of the normal bases in the DNA and the other is incorporation of modified nucleotides generated in resident nucleotide pools.[2,3]

The incorporation of dUTP into DNA which arises from the deamination of dCTP or as a byproduct of UTP biosynthesis is a major cause for shortening the Okazaki fragment during nascent strand synthesis that occurs during base excision repair by uracil DNA glycosylase.[4] dUTP is thus detrimental to organisms and it is eliminated by a specific nucleotidase, dUTPase, from the resident nucleotide pools.[5–7] 8-Oxo-2′-deoxyguanosine triphosphate (8-oxo-dGTP) is one of the major causes of spontaneous mutagenesis, because 8-oxo-dGTP is formed by the spontaneous oxidation of dGTP in the nucleotide pool and it is incorporated into the nascent strand opposite both adenine and cytosine in the template strand during DNA replication.[8] We have demonstrated that 8-oxo-dGTP is specifically hydrolyzed by MutT family proteins from prokaryotes to humans, and, as a result, organisms maintained a low spontaneous mutation rate.[9–12]

Genome projects revealed the existence of many MutT-like proteins with different substrate specificities,[13–17] and a structure-based approach recently identified another novel enzyme, namely inosine triphosphate pyrophosphatase (ITPase), which hydrolyzes deaminated purine nucleoside triphosphates such as ITP, dITP, and XTP.[18–20] In *Escherichia coli*, a mutant of the *rdgB* gene coding ITPase protein is viable but it shows synthetic lethality with *recA* or *recBC* mutation.[21] Recently, the lethality of *rdgB recA* or *rdgB recBC* double mutants has been shown to be suppressed by the inactivation of endonuclease V (EndoV) which initiates the excision of

Communicated by Hideo Shinagawa

* To whom correspondence should be addressed. Tel. +81 92-642-6800, Fax. +81 92-642-6791, E-mail: yusaku@bioreg.kyushu-u.ac.jp

deoxyinosine (dI) or deoxyxanthosine (dX) incorporated into DNA.[22] It is likely that an ITPase deficiency results in the accumulation of its substrate nucleotides, dITP or dXTP in the nucleotide pools, thus causing an increased accumulation of dI or dX into DNA, and further excision repair initiated by EndoV leads to chromosomal fragmentation in *recA* or *recBC* mutants. Following this study,[22] the missense mutants of the *dut* gene encoding dUTPase, were reported to exhibit synthetic lethality with *recA* or *recBC* mutations, and synthetic lethality was suppressed by inactivation of the *ung* gene encoding uracil DNA glycosylase, thus confirming that the accumulation of abnormal nucleotides in the nucleotide pools increased chromosomal fragmentation as a consequence of the excision repair of such abnormal bases incorporated into DNA.[7]

The first reported deficiency in the ITPase in a human was characterized by an elevated ITP level in erythrocytes as an inherited abnormality.[23] However, such individuals do not exhibit any abnormal phenotype even with accumulation of ITP in erythrocytes. A cDNA for human ITPase was isolated and its gene *ITPA* was identified,[19] and the structure of the *hITPA* gene and nucleotide alterations responsible for the ITPase deficiency were reported.[24,25]

In contrast to *E. coli*, the biological importance of ITPase protein and the pathological consequences of its deficiency in humans or mammals have not yet been elucidated. To explore the biological significance of ITPase in mammals, disruption of the mouse *Itpa* gene and characterization of such mutant mice is considered to be one of the best experimental approaches. During isolation of the *Itpa* gene from the mouse genome, we found one processed *Itpa* gene-like sequence and two processed *Itpa* pseudogenes as well as the *Itpa* gene itself with introns, which was isolated by a retro-recombination method in the present study.

2. Materials and Methods

2.1. RT-PCR

Total RNAs from CCE mouse embryonic stem (ES) cells derived from 129SvEv mouse,[26] serum-starved and serum-stimulated BALB/c3T3 cells,[27,28] and C57BL/6J adult mouse tissue specimens were prepared using ISOGEN (Nippon Gene) according to the manufacturer's instructions. Total RNAs from BALB/c mouse tissues were purchased from Clontech. Total RNA (10 μg) was treated with 20 units of RNase-free deoxyribonuclease I (DNase I) (Amersham Biosciences) at 37°C for 15 min in 100 μl of reaction buffer containing 40 mM Tris-HCl (pH 7.5) and 6 mM $MgCl_2$. The treated RNA was purified by extraction with Phenol/$CHCl_3$ (1:1) followed by ethanol precipitation. The purified RNA (2 μg) was used for synthesis of the first strand cDNA using a First-strand cDNA synthesis Kit (Amersham Biosciences) according to the manufacturer's instructions. PCR was performed in 20-μl reaction mixture containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.3 μl of the first-strand cDNA, 0.4 U of recombinant Taq DNA polymerase (Takara), 4 μM of each primer (Table 1), and 200 μM of each deoxynucleoside triphosphate. The initial denaturation was performed at 95°C for 1 min and the amplification was performed by 27, 32, 40 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 20 sec, and extension at 72°C for 60 sec, followed by a final extension at 72°C for 5 min. PCR products were subjected to agarose gel electrophoresis, and the band intensity on the gel stained with ethidium bromide was measured using the LAS1000-plus Luminescent Image Analysis System (FUJI FILM).

Table 1. The synthetic oligonucleotides used in this study. These were obtained from Greiner Japan and Hokkaido System Science.

Oligomer	Seuence
mItpa_EX1f	GACGGGAAACGCCAAGAAGCTG
mItpa_EX2r	CAAAGTGCATGGAAAATTATCTCCG
mItpa_EX2f	TCAATGCACTTTGGAGGCTCAG
mItpa_EX3r	TGGAAATCTCATCCGGTTCTCC
mItpa_EX3f	CCATACAGAAGTGTCGGGAGGC
mItpa_EX4r	AGGTATCTTCCACCAGGACAGG
mItpa_EX4f	CTTTAACGCACTTGGGGGACTG
mItpa_EX5r	AGGCTTCAGCTTCTGTAGGAAC
mItpa_EX5f	AAATGGTTCCTACAGAAGCTGA
mItpa_EX6r	GTGCATAGGCCGATTTGTCTTC
mItpa_EX6f	CAGCCAGTGCTCCTGTTCAGAG
mItpa_EX7r	CGGCTGCCTCGTGGCATCAC
mItpa_EX7f	TCCCTGCTTTCAGCCTGATG
mItpa_EX8r	TCTTCTCAGACTTTGGCATCTC
3-1mIT	GATATCGCTAGTGGTCACCAGCCCCAG
3-2mIT	GAGCCAACTGCTCTTCCAAGAACTG
3-3mIT	AAGTCTTCTAGAATTTACATTTGC
3-4mIT	CAGACACACCATTTATTTTGGAAA
5-1mIT	*AAGCTTGCC* ATGGCTGCGTCTTTGGTCG[a]
5-3mIT	AAAGGGTTACTGGGCATTCTGG
Pα-45	CCCAATAAGAAACTGGCATCTTTC
Pα-46	AGTTATAGTATAAATAATATCCAAGA
3-1hIT	GAAGTCAAGCTGCCAAACTGC
3-2hIT	CAGACAGGCCGGTGAGGCTACTTG
5-1hIT	GTAACCGGGGATCACCATGGC
5-3hIT	TGGACGCCAAGGAGTTTTCGGTG

[a]Italic letters indicate a sequence added to create restriction sites for *Hin*dIII and *Nco* I.

2.2. Genomic PCR

Genomic DNA (25 μg) from CCE cells was treated with RNase A (20 units, Sigma), which was boiled for 15 min to inactivate contaminated DNase, in 150 μl of a reaction mixture containing 5 mM Tris-HCl (pH 7.6) and

0.5 mM EDTA at 37°C for 30 min to eliminate the residual RNA completely. DNA was purified by extraction with Phenol/$CHCl_3$ followed by ethanol precipitation. A PCR reaction was performed with 100 ng of the purified genomic DNA with appropriate primers (Table 1), as described in the RT-PCR section. The PCR products were subcloned into pT7 Blue T vector (Novagen) plasmid and their nucleotide sequences were confirmed by sequencing the insert from the both strands using ABI Sequencing Kits and a model 3100 automated DNA sequencer (ABI), according to the manufacturer's instructions.

2.3. *Isolation of genomic clones for* Itpa *gene*

A PCR fragment carrying intron 2 or 3 of *Itpa* amplified from CCE genomic DNA was subcloned into the multi-cloning site of the πANγ plasmid, and each plasmid was introduced into *E. coli* MC1061[P3] cells.[29] Each cell was infected with λ phage genomic library constructed in λTK phage with 129 SvJ mouse genomic DNA.[29] Four out of 6 clones isolated with a probe plasmid carrying intron 2, and 8 out of 14 clones isolated with a probe plasmid carrying intron 3, using the retro-recombination methods,[29] contained the appropriate probe sequence and identical adjacent sequences for each other, respectively.

2.4. *Plasmid construction*

Mouse *Itpa* cDNA was amplified from the mouse testis library (ML 4007AB, Clontech) and from total RNA prepared from CCE cells using 5-3 mIT and 3-3 mIT primers (Table 1). Human *ITPA* cDNA was amplified from a human leukocyte library (HL 4021AB, Clontech) using 5-3 hIT and 3-2 hIT primers (Table 1). PCR products were subcloned into pT7 Blue T Vector. The entire coding region of type A *Itpa* cDNA (615 bp) was amplified with 5-1 mIT and 3-1 mIT primers, in which a *Nco* I site at the initiation site and an *Eco*RV site after the termination codon was introduced, respectively. The PCR product digested with *Nco* I and *Eco*RV was subcloned into a *Nco* I/*Bam*HI site of the pET8c and pET32a plasmids (Novagen), in which a *Bam*HI site was converted to a blunt end by filling in with Klenow fragment. Obtained plasmids were designated pET8c:mITPA and pET32a:mITPA, respectively. A *Nco* I-*Bam*HI fragment of hITPA cDNA was also subcloned into a *Nco* I/*Bam*HI site of pET32a or pET8c plasmid yielding pET32a:hITPA and pET8c:hITPA, respectively.

2.5. *Immunodetection of* ITPA

Rabbit antiserum against the fusion protein Trx-hITPA were prepared as described previously.[30] The antiserum was able to detect both Trx-hITPA and mITPA proteins with almost the same efficiency, and thus was designated as anti-ITPA. Western blotting analyses were performed as previously described,[31] using anti-ITPA serum (1/500 dilution). Immunohistochemistry with anti-ITPA were performed as follows. The mice were deeply anesthetized with 5% pentobarbital, and perfused transcardially with 10 ml of heparinized saline (0.9%) followed by 30 ml of phosphate buffer (0.1 M) containing 4% paraformaldehyde. Tissue specimens were fixed in 4% paraformaldehyde at 4°C for 12 to 24 hr and embedded in paraffin. The sections (4 μm) were deparaffinized, pretreated in 3% hydrogen peroxide in methanol, and subjected to immunohistochemistry with anti-ITPA serum (1/500 dilution). The sections were processed by Vectastain ABC KITs (Vector Laboratories) with an anti-rabbit biotinylated secondary antibody, and the peroxidase reaction product was detected using 3,3′-diaminobenzidine-tetrahydrochloride (Sigma). The slides were subjected to counterstaining by hematoxylin. Digital images were acquired using Axioskop2 plus equipped with an AxioCam (Carl Zeiss).

3. Results and Discussion

3.1. *Isolation of Mouse cDNA and Genomic Sequences Homologous to Human* ITPA *cDNA sequence*

We searched the DNA sequence databases to retrieve the mouse cDNA or genome sequences homologous to human *ITPA* cDNA sequence (accession no. AF219116), using the BLASTN program (http://www.ncbi.nlm.nih.gov/BLAST/), and found many mouse expressed sequence tags (ESTs) and several mouse genomic sequences. Among the many EST sequences retrieved, the longest cDNA sequence (accession no. AK008279) which encodes a polypeptide highly homologous to the amino acid sequence of human ITPA protein (hITPA) was selected, and it was used for the BLASTN search with the Ensembl Genome Data Resources for mouse (http://www.ensembl.org/Mus_musculus/). The alignment revealed that there are several different mouse chromosomes carrying highly homologous but not identical sequences to that of AK008279 itself, thus suggesting that there are multiple genes for mouse ITPA protein (mITPA) or related proteins.

To isolate cDNA or genomic sequences which encode mITPA, we amplified the sequences from genomic DNA and cDNA prepared from a mouse ES cell line CCE cells which was established from a 129 SvEv mouse, using two different primer sets for the AK008279 EST sequence (Fig. 1A, Table 1). From the RNase A-treated genomic DNA, a major single fragment approximately 1100 bp in length was amplified with primer set I (5-1 mIT and 3-3 mIT), but not with primer set II (5-3 mIT and 3-3 mIT; Fig. 1B, lanes 2, 7). In primer set I, the 5′-primer hybridizes a region including the initiation codon while the 5′ primer in the latter set hybridizes its 5′-untranslated region (UTR). In contrast, two major fragments of approximately 1100 and 830 bp in length from primer set I,

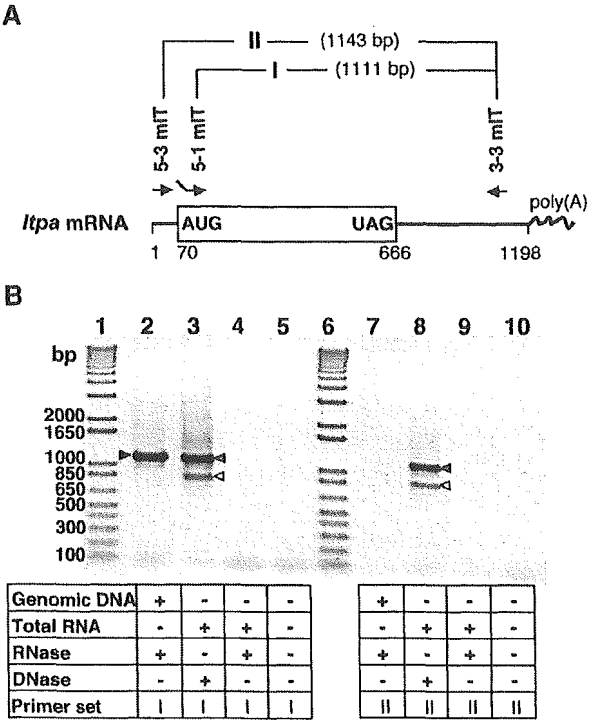


Figure 1. Amplification of the mouse *Itpa* cDNA-related sequences from genomic DNA and total RNA prepared from CCE ES cells. *A*. Primer sets for mouse *Itpa* cDNA. Two primer sets (I and II) were designed based on a mouse *Itpa* cDNA (accession no. AK008279). Primer set I with primers 5-1 mIT and 3-3 mIT, in which the 5′-primer hybridizes into the initiation codon with 9-base extra sequence (see Table 1), was expected to amplify the 1111-bp fragment, while primer set II with primers 5-3 mIT and 3-3 mIT, in which the 5′-primer hybridizes into 5′-UTR sequence, was expected to amplify the 1143-bp fragment from *Itpa* cDNA. The coding regions for ITPA and poly(A) sequences are shown with an open box and a wavy line, respectively. *B*. Agarose gel electrophoresis of the PCR products. To avoid cross-contamination of genomic DNA or RNA into total RNA or genomic DNA prepared for PCR templates, genomic DNA was treated with DNase-free RNase A, while total RNA preparation was pre-treated with RNase-free DNase I before cDNA synthesis. PCR was performed with a template and primer set shown under each lane. Lanes 1, 6, size markers. The arrowheads indicate the major PCR products in each reaction: closed, 1108-bp genomic PCR product; gray, 1111- or 1143-bp RT-PCR product; open, 832- or 864-bp RT-PCR product.

or 1140 and 860 bp in length from primer set II were amplified from the cDNA prepared from the DNase I-treated RNA, respectively (Fig. 1B, lanes 3, 8).

A sequence analysis of subcloned PCR products (102 independent clones) obtained by primer set II, revealed three different types of cDNA sequences (types A, B, and C) to be amplified from the RNA prepared from the ES cells (Fig. 2A). In the RT-PCR products, the ratio of each cDNA identified was type A : B : C = 1 : 0.3 : 0.05. The longest type A cDNA isolated was 1143 bp length with an open reading frame (ORF) of 597 nucleotides, whose sequence was identical to AK008279 and was predicted to encode a polypeptide of 198 amino acid residues with a molecular weight of 21,883 Da. The amino acid sequence was 89.9% identical to that of human ITPA (194 aa). Type B and C cDNAs were 864 and 720 bp length, respectively, and had the same 5′- and 3′-UTR sequences as did the type A cDNA. Both of the cDNAs had the same reading frame as that of the type A, but they were shorter than that of type A, and were 318 and 168 nucleotides in length, respectively. Both type B and C cDNA shared the first ATG codon and the stop codon (TAG) with type A cDNA, and were predicted to encode polypeptides with 105 or 55 amino acid residues missing central parts (94–186 aa or 44–186 aa) of the full-length mITPA protein, respectively (Fig. 2A). These results were confirmed by a sequence analysis of RT-PCR products (100 clones) obtained by primer set I, whose length was 1111, 832, or 688 bp, respectively (data not shown).

A sequence analysis of 50 clones of DNA fragments amplified from genomic DNA with primer set I revealed a single 1108-bp fragment whose sequence was 99% identical to that of the type A *Itpa* cDNA (Fig. 2B). The more 5′- and 3′-regions of the genomic sequence were amplified using a new primer set based on the mouse DNA sequence from clone RP23-16O15 on chromosome 2 (accession no. AL672251) which contained exactly the same sequence shown in Fig. 2B (Table 1). The genomic sequence (1135 bp) retained the same ORF of 597 nucleotides which encodes the same polypeptide as does the type A cDNA, with a canonical polyadenylation signal (AATAAA) in the same position as in the type A cDNA. There are three silent nucleotide substitutions in the ORF of the genomic sequence in comparison to that in type A cDNA (Fig. 2B). In addition, one of the five repeats of the CAA triplet found in the 3′-UTR of the type A *Itpa* cDNA sequence was missing in the genomic sequence. These results strongly suggest that there are indeed multiple genomic sequences and transcripts which are capable of encoding mITPA protein or its variants.

3.2. *Transcripts for mouse* ITPA *expressed in adult mouse tissue*

Since the genomic sequence amplified with the primer set I has no intron-like sequence with a complete ORF for mITPA protein, it is likely that this genomic sequence is a processed type gene for the mITPA protein. Although we did not obtain such a sequence in the RT-PCR products from the CCE ES cells with primer set I, it is possible that the genomic sequence is transcribed in other types of cells or tissues. To clarify this point, we amplified cDNA prepared from various adult mouse tissues using primer set I, and the RT-PCR products were digested with two different restriction enzymes, *Apa*LI and *Hae* II. The former digests the processed gene-like sequence but not the three types of cDNA, while the latter digests only the

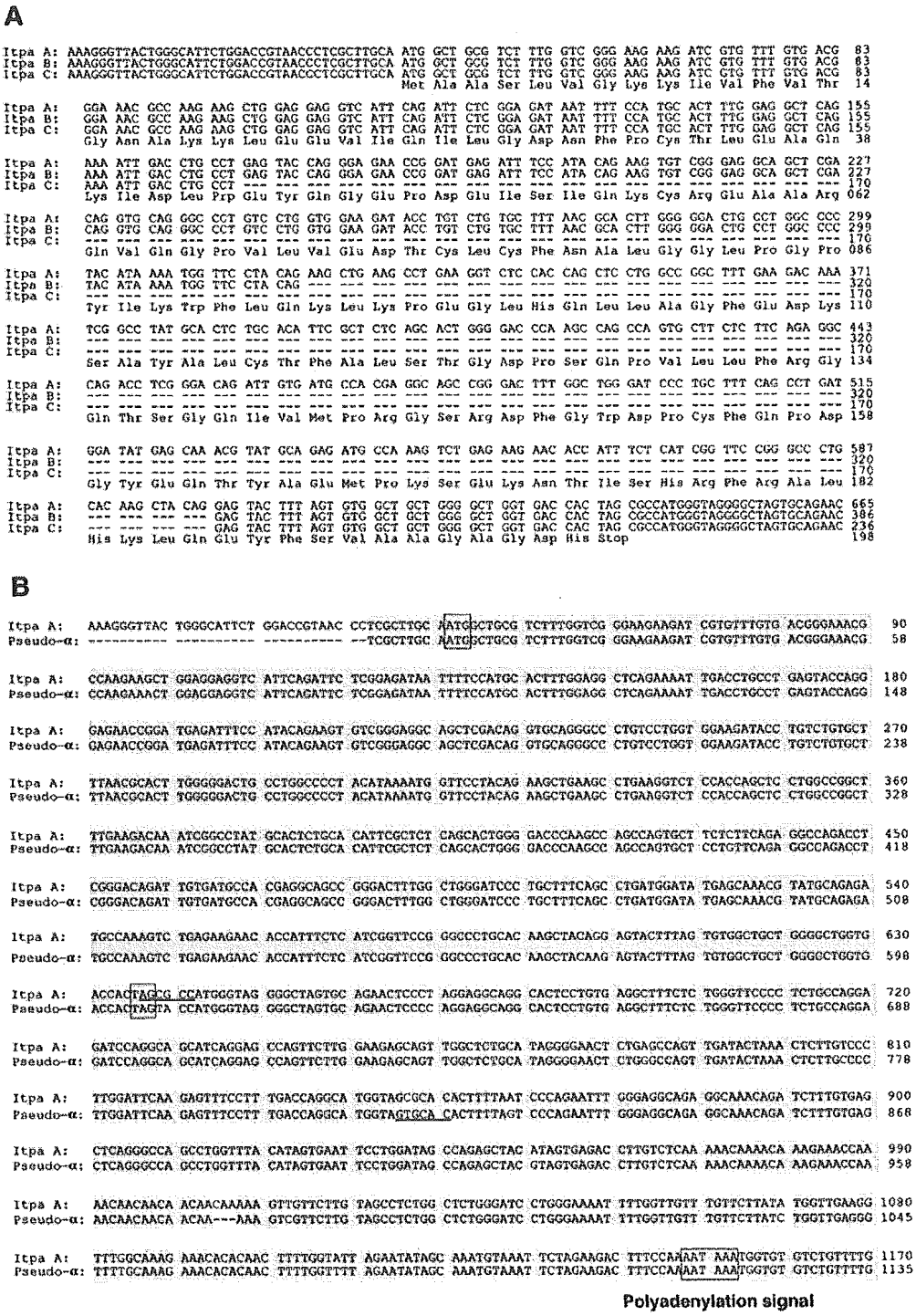


Figure 2. Mouse *Itpa* cDNA-related sequences. *A*. Alignment of the coding sequences of three types of mouse *Itpa* cDNA and amino acid sequences of their translation products (accession no. AB100501). The number of nucleotides and amino acid residues are shown on the right. *B*. Alignment of DNA sequences for type A *Itpa* cDNA and the genomic PCR product containing *Itpa* cDNA-like sequence. Type A *Itpa* cDNA was amplified from CCE RNA with a primer set of 5-3 mIT and 3-4 mIT, and the genomic PCR product was obtained from CCE genomic DNA with the primer set of Pα-45 and Pα-46 (Table 1) designed based on a mouse genome sequence (accession no. AL672251) which contained a sequence identical to that of the genomic PCR product shown in Fig. 1B (lane 2). Identical nucleotides between the two sequences are shaded and the initiation and stop codons and the polyadenylation signal are boxed. The recognition sequences for *Hae* II in the *Itpa* cDNA and *Apa*LI in the genomic PCR product are underlined.

three types of cDNA for mITPA (Fig. 2B, 3A). A single 1108-bp fragment was amplified from CCE genomic DNA pre-treated with DNase-free RNase A, while two or three bands corresponding to 1111-, 832-, or 688-bp fragments were detected in the RT-PCR products from all the tissues examined as well as from the CCE cells (Fig. 3B).

*Apa*LI digestion of the 1108-bp genomic PCR product yielded 814- and 294-bp fragments as predicted from its sequence (Fig. 3C, lane 10), while no such digested band was detected in the RT-PCR products from all adult mouse tissues examined as well as from the CCE ES cells (Fig. 3C, lanes 2–9). On the other hand, *Hae* II digestion produced 609- and 502-bp fragments from all RT-PCR products (Fig. 3D, lanes 2–9), and the genomic PCR product remained undigested (Fig. 3D, lane 10). From the cDNA prepared from the lung, cerebrum and cerebellum as well as from CCE cells, three fragments — 1111-bp, 832-bp, and, to a much lesser extent, 688-bp fragments — were amplified, and the latter two fragments were also digested by *Hae* II but not by *Apa*LI. These results were confirmed by simultaneous digestion with the two enzymes (Fig. 3E).

As a result, we concluded that the 1108-bp genomic sequence highly homologous to the mouse *Itpa* cDNA sequence is not transcribed in any of the adult mouse tissues examined so far, and thus we designated the genomic sequence as *Itpa* processed pseudogene α (accession no. AB100502). Furthermore, we confirmed that the three types of *Itpa* transcripts identified in CCE ES cells were also expressed in adult mouse tissue.

3.3. *Type A* Itpa *mRNA encodes a functional deoxyinosine triphosphate pyrophosphatase*

In order to confirm that a polypeptide encoded by the type A *Itpa* transcript possesses the capability of hydrolyzing nucleotides such as ITP or dITP, the type A cDNA was placed under the control of T7 promoter in pET8c or pET32a. In *E. coli* cells harboring pET8c:mITPA or pET32a:mITPA plasmids, a single 22-kDa or 34-kDa polypeptide was expressed after IPTG induction, respectively (Fig. 4A, lanes 5, 11). The former corresponds to the native mITPA protein with the expected molecular weight of 21,883 Da, and the latter is a fusion protein with thioredoxin (Trx-mITPA) whose molecular weight is predicted to be 33,558 Da. Crude extracts prepared from *E. coli* cells harboring vector itself or pET8c:mITPA, pET32a:mITPA were incubated with dITP or dATP, and the reaction products were separated on a DEAE column attached to an HPLC system.[12] Extracts from *E. coli* cells harboring vector itself had no detectable activity to hydrolyze dITP (Fig. 4B, top panel; Fig. 4C, open circles and squares), while the extracts from *E. coli* cells expressing mITPA efficiently converted dITP to dIMP (Fig. 4B, bottom panel; Fig. 4C,

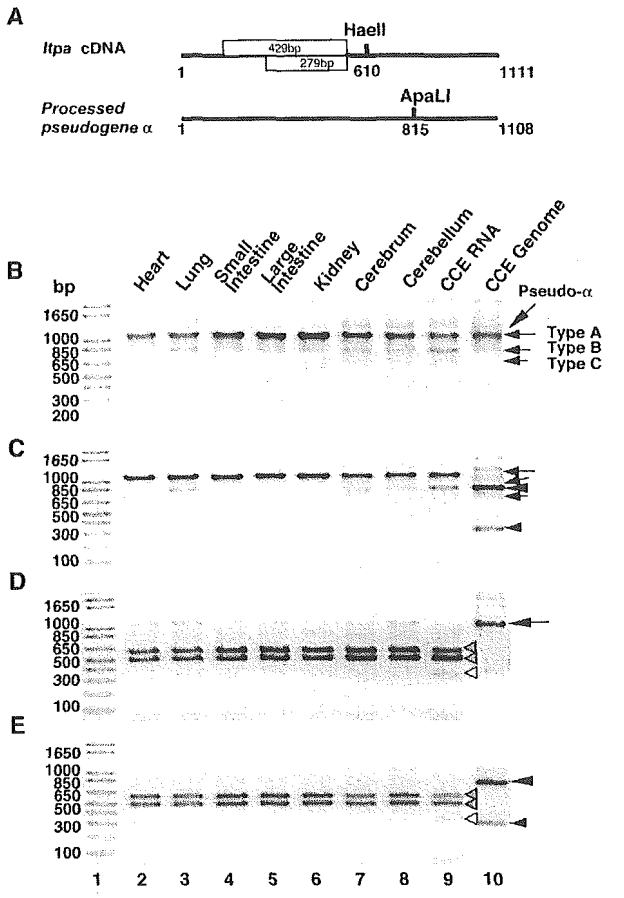
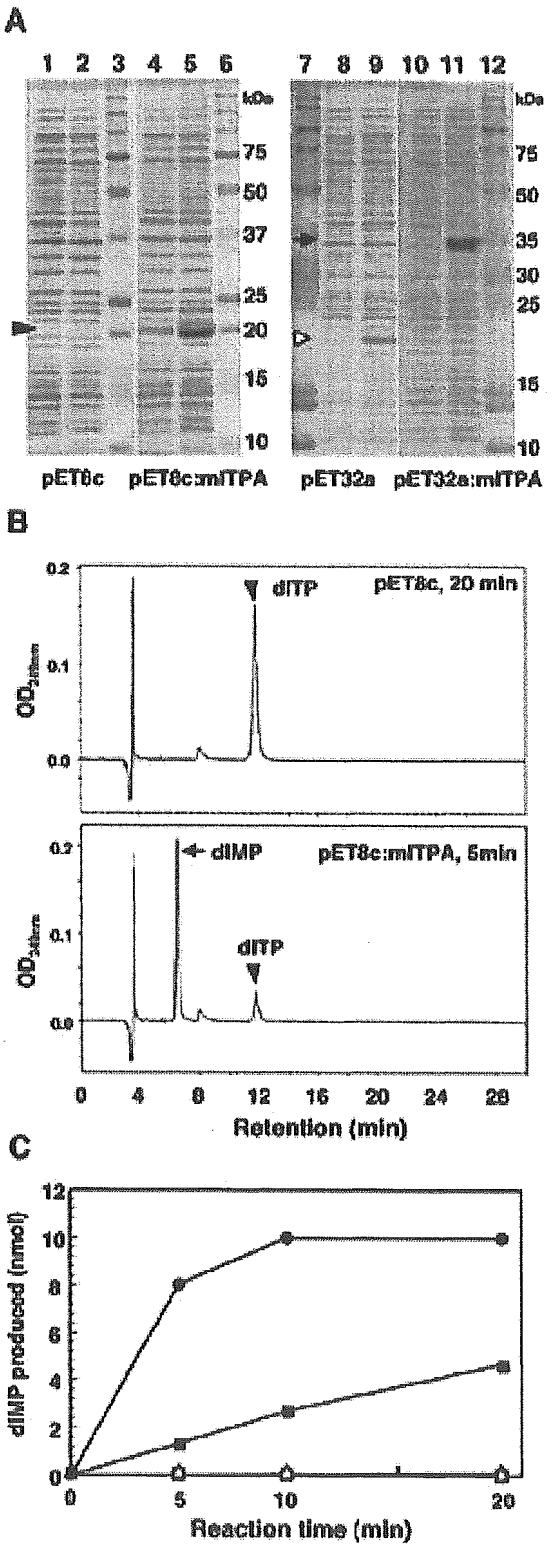


Figure 3. Transcripts for mouse ITPA expressed in adult mouse tissue. *A.* Schematic representation of the DNA sequences for type A *Itpa* cDNA (top) and the genomic PCR product containing *Itpa* cDNA-like sequence (bottom). The recognition site for *Hae* II or *Apa*LI in each sequence is shown, and sequences missing in type B and C *Itpa* cDNA are shown with open boxes. *B.* Agarose gel electrophoresis of the RT-PCR products from adult mouse tissue. cDNAs reverse-transcribed from total RNA prepared from C57BL/6J adult mouse tissue specimens (lanes 2–8) and CCE cells (lane 9), were used as templates for RT-PCR with primer set I shown in Fig. 1A. RNAs from lung, heart and kidney as well as CCE cells were pre-treated with RNase-free DNase I before cDNA synthesis in order to avoid any contamination of genomic DNA. Genomic DNA from CCE cells was treated with DNase-free RNase A (lane 10). *C.* Digestion of the PCR products with *Apa*LI. *D.* Digestion of the PCR products with *Hae* II. *E.* Digestion of the PCR products with *Apa*LI and *Hae* II. The longest arrow indicates the genomic PCR fragment of 1108 bp. RT-PCR products for type A (1111 bp), B (832 bp), and C (688 bp) are also shown with three different sizes of arrows. Two large and small closed arrowheads indicate the digested bands (814, 294 bp) derived from the genomic PCR fragment. Gray arrowheads indicate *Hae* II-digested RT-PCR fragments (609, 502 bp) derived from type A *Itpa* cDNA, open arrowheads indicate the *Hae* II-digested RT-PCR fragment (330 bp) derived from type B *Itpa* cDNA.



closed circles). Extracts containing Trx-mITPA also hydrolyzed dITP to dIMP (Fig. 4C, closed squares). dATP was barely hydrolyzed by extracts prepared from *E. coli* cells with or without mITPA (Fig. 4C, triangles), thus demonstrating that mITPA encoded by the type A *Itpa* mRNA specifically hydrolyzes dITP into dIMP.

3.4. Isolation and characterization of mouse Itpa *gene*

A detailed analysis of mouse genomic sequences revealed that there are several pseudogene-like sequences highly homologous to mouse *Itpa* cDNA in the mouse genome (Table 2). In order to isolate the functional mouse *Itpa* gene, we initially intended to isolate unique intronic sequences for mouse *Itpa* gene. Based on the genomic structure of the human *ITPA* gene consisting of 8 exons and 7 introns,[24,25] each exon of the mouse *Itpa* gene was predicted based on the type A mouse *Itpa* cDNA sequence, and each intron sequence, except for introns 1 and 6, was amplified from genomic DNA prepared from CCE ES cells by two primers hybridized to the adjacent exons (Table 1).

With the genomic fragment containing the intron 2 or 3 as a probe, we applied the retro-recombination method,[29] in order to isolate the genomic sequences encompassing *Itpa* gene from a λTK phage genomic library derived from

Figure 4. Type A *Itpa* mRNA encodes functional deoxyinosine triphosphate pyrophosphatase. *A.* The expression of recombinant mITPA proteins in *E. coli* cells. Isopropyl β-D-thiogalactoside (IPTG, 1 mM) was added to the exponentially growing cultures of *E. coli* BL 21 cells carrying pET8c (lanes 1, 2), pET8c:mITPA (lanes 4, 5) or pET32a (lanes 8, 9, pET32a:mITPA (lanes 10, 11), and each culture was further incubated at 37°C for 3 hr. An aliquot was harvested before (lanes 1, 4, 8, 10) and after the addition of IPTG (lanes 2, 5, 9, 11). Whole cell extracts prepared from each cell were subjected to 15% (Lanes 1–6) or 12.5% SDS-PAGE (lanes 7–12), and the gels were stained with Coomassie brilliant blue. Lanes 3, 6, 7, 12, molecular weight markers. The closed arrowhead indicates a band corresponding to the 22-kDa mITPA protein, an arrow indicates the 34-kDa Trx-mITPA fusion protein, and the open arrowhead indicates the Trx itself (20 kDa). *B.* Hydrolysis of dITP by the recombinant mITPA. dITP (1 mM) was incubated in reaction mixture (10 μl) contained 50 mM Tris-HCl, (pH 8.5), 50 mM $MgCl_2$, 1 mM DTT, with 50 ng of whole cell extracts prepared from IPTG-induced *E. coli* cells carrying pET8c for 20 min (top panel) or pET8c:mITPA for 5 min (bottom panel), at 30°C. The reaction was terminated by adding 5 mM EDTA, then the products were separated on a TSK-Gel DEAE-2SW column, as previously described.[12] *C.* mITPA specifically hydrolyzes dITP but not dATP. Whole cell extracts (50 ng of protein) were incubated with 1 mM dITP or dATP for the time noted, and the products were separated on a TSK-Gel DEAE-2SW column. The amount of nucleotides was determined by UV absorbance at 249 nm. Closed circles, dITP reacted with extracts prepared from cell carrying pET8c:mITPA; open circles, dITP reacted with extracts prepared from cells carrying pET8c; closed squares, dITP reacted with extracts prepared from cells carrying pET32a:mITPA; open squares, dITP reacted with extracts prepared from cells carrying pET32a; closed triangles, dATP reacted with extracts prepared from cells carrying pET8c:mITPA; open triangles, dATP reacted with extracts prepared from cell carrying pET8c.

Table 2. Restriction fragments containing a homologous sequence with mouse type A *Itpa* cDNA predicted from the mouse genome database.

Nomenclature	Chromosome	DDBJ/EMBL/GenBank accession no.	Restriction fragment size (kbp) *Eco*RI	*Hin*dIII	*Xba*I
Itpa gene	2F3	AL772162 AB101662	3.17, 7.4, 9.95	0.84, 0.91, 2.35, 21	4.7, 10.4
Itpa pseudogene α	2E1	AL672251 AB100502	5.9, 12.7	4.8	4.7, 7.4
Itpa pseudogene β	9E4	AC091531	11.8, >2.6	9.8	12.5
Itpa pseudogene γ	XA3.2	AL672147	8.8	4	4.4

129 SvJ mouse. Fourteen clones were isolated and classified into three groups as shown in Fig. 5A. The sequences of PCR products and λTK phage clones together with the mouse DNA sequence from clone RP23-175J8 on chromosome 2 (accession no. AL772162) revealed that mouse *Itpa* gene consists of 8 exons and 7 introns as expected from the genomic structure of the human *ITPA* gene, and it spans about 13.8 kb of the region on the F3 band of chromosome 2 (Table 2, Fig. 5A,B). The alignments of type B and C *Itpa* cDNA sequences with the genomic sequence revealed that the two transcripts are generated by an unusual joining between exon 5 and exon 8 or between exon 3 and exon 8, respectively (Fig. 5A).

Type A mRNA, a major transcript from *Itpa* gene, is generated by splicing at canonical splice sites with GT-AG dinucleotides at the splice junctions.[32,33] However, type B and C transcripts are most likely generated by splicing at non-canonical splices sites with GA-AG or AA-AG dinucleotides at each unusual splice junction, respectively (Fig. 5B).[33] In both cases, the same 3′-acceptor site in the exon 8 was used. The 5′-donor sequence (*GAGTAC*) in exon 3 was also found at the 5′-end of alternatively spliced exon 8, while the sequence (*CTACAG*) at the 3′-end of alternatively spliced exon 5 was also present at the 3′-acceptor site in exon 8, thus suggesting that these repeated sequences may be involved in such unusual splicing.

To confirm the genomic structure for the mouse *Itpa* gene, genomic DNA prepared from ES cells was digested with restriction enzymes *Eco*RI, *Hin*dIII or *Xba*I, and subjected to Southern blot hybridization with a DNA fragment containing the entire coding region from type A cDNA (731 bp) as a probe (Fig. 6). In addition to the bands expected from the genomic sequence for *Itpa* (Table 2; Fig. 6, arrows), three extra bands for each digested sample were identified which likely represent the processed pseudogenes α, β, and γ, respectively (Fig. 6, arrowheads). Among them, the *Itpa* pseudogene α located on chromosome 2E1 has a complete ORF for exactly the same polypeptide as ITPA encoded by the type A transcript, with a polyadenylation signal (Fig. 2B). However, we could not detect any transcribed sequence derived from the *Itpa* pseudogene α in CCE ES cells or from any mouse tissues examined, as shown in Fig. 3, thus concluding that the mouse has only one functional *Itpa* gene in its genome.

The existence of such processed pseudogenes in the mice were confirmed by Southern blot analysis of 129, C57BL/6J and BALB/c mouse strains, and essentially the same results were obtained from all strains examined as well as for CCE ES cells (data not shown). Although there was one more pseudogene-like sequence on chromosome 6 in the mouse genome database, we could not identify any band corresponding to the sequence in our Southern blots (Fig. 6). This may be because the sequence has much less homology with the *Itpa* cDNA probe used.

3.5. Expression of mouse Itpa *gene in mouse tissues and its association with cell proliferation*

Expression of *Itpa* gene in adult mouse tissues was examined by a Northern blot analysis (Fig. 7A, top panel). A band corresponding to 1.35 kb in length was detected in all examined tissues, however, its levels significantly varied from tissue to tissue. Measurements of the radioactivities of bands on the blot revealed the highest expression to be seen in the testis and brain, followed by the thymus (Fig. 7A, lanes 2, 13, 14), and most other tissues except for smooth muscle, the salivary gland and stomach, expressed about a 50% of the level of *Itpa* mRNA detected in the brain or testis (Fig. 7A, bottom panel).

We then monitored the expression of *Itpa* in serum-starved BALB/c 3T3 cells with or without serum stimulation (Fig. 7B). A quantitative analysis of the Northern blots revealed the level of *Itpa* mRNA in quiescent cells to be low, however, this level increased twofold within 15 hr after serum stimulation and then gradually decreased after entering the S phase, 18 hr after serum stimulation, and then returned to the basal level after 24 hr or later (Fig. 7C).

The genomic sequence of the *Itpa* gene revealed a

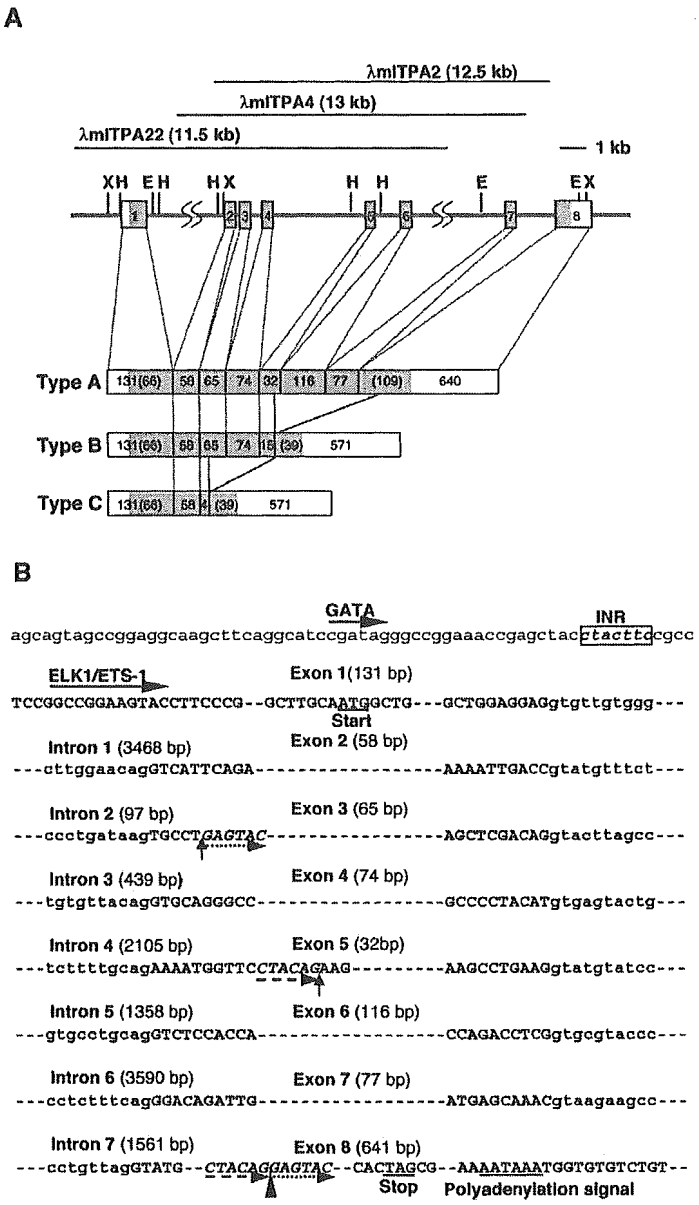


Figure 5. Genomic organization and alternative splicing of the mouse *Itpa* gene. *A*. A schematic diagram of the structure of the *Itpa* gene and its transcripts. The alignment of the DNA fragments derived from mouse genomic libraries is shown in the upper part. The approximate insertion size of each clone is also shown in parentheses. In the middle part, the structure of the gene, together with the appropriate restriction enzyme sites (E: *Eco*RI, H: *Hin*dIII, X: *Xba*I), is shown. The boxes represent the exons for the *Itpa* gene and the shaded regions represent the protein-coding region. In the bottom part, three types of alternatively spliced *Itpa* mRNAs are shown. The sequence derived from each exon is shown as a box with a number of bases consisting of each exon, and the dashed lines indicate the normal splicing sites, while solid lines indicate the unusual splicing sites found in type B and type C transcripts. A number of bases for the protein-coding region derived from exon 1 or exon 8 are shown in parenthesis. *B*. The nucleotide sequences of intron/exon boundaries of *Itpa* gene. The nucleotide sequences of the exons and parts of the introns determined by a comparison of the sequence of type A *Itpa* cDNA with the mouse genomic sequence (Accession nos. AL772162, AB101662), are shown in bold uppercase and lowercase, respectively, and the flanking sequences in plain lowercase. The genomic sequences shown were confirmed by sequencing of phage clones from a λTK phage genomic DNA library derived from 129SvJ mouse or genomic PCR products. The start of exon 1 was based on the most 5′-extended EST clone for *Itpa* mRNA (Accession no. BB654208). The initiation codon ATG, the termination codon TGA and a putative polyadenylation signal are underlined. The arrows indicate the unusual 5′-splicing sites in the exon 3 and 5, and the arrowhead indicates the unusual 3′-splicing site in the exon 8, found in type B and type C transcripts. The dotted lines indicate the repeated sequences at the unusual splicing junctions.

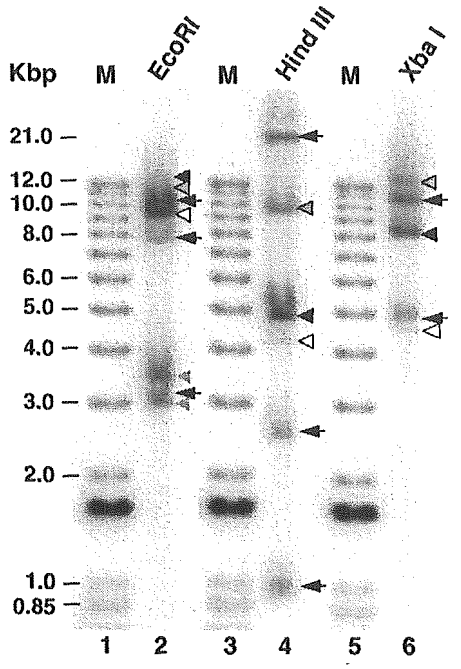


Figure 6. Southern blot analyses of genomic DNA for mouse *Itpa*-related sequences. Genomic DNA (10 μg) prepared from CCE ES cells was digested with *Eco*RI (lane 2), *Hin*dIII (lane 4) or *Xba*I (lane 6), and was subjected to Southern blot analyses with ^{32}P-labeled fragment containing the entire coding region from the type A *Itpa* cDNA (731 bp). Southern hybridization was done as previously described,[27] and the membrane was washed twice in 2× SSC/0.1% SDS, and then once in 0.2×SSC/0.1% SDS. The arrows indicate bands corresponding to the expected fragments from *Itpa* gene, closed arrowheads indicate bands corresponding to the expected fragments from pseudogene α, the gray arrowheads with solid line indicates the bands corresponding to the expected fragments from pseudogene β, while the gray arrowheads without solid line indicate band(s) whose size could not be predicted precisely because of incompleteness of the database. The open arrowheads indicate bands corresponding to the expected fragments from pseudogene γ (Table 2).

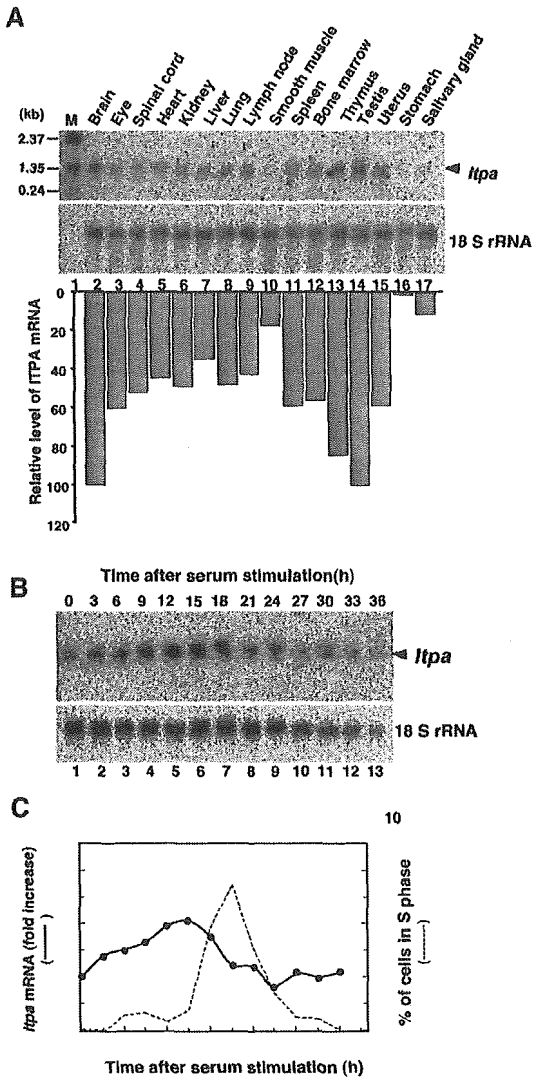


Figure 7. Expression of mouse *Itpa* gene in mouse tissue and its association with cell proliferation. *A*. The expression of *Itpa* mRNA in various types of adult mouse tissue. Total RNAs (20 μg each) extracted from various types of mouse tissue were electrophoresed, transferred onto a Hybond^{TM}-N^{+} nylon membrane, and subsequently probed with ^{32}P-labeled fragment containing the entire coding region from the type A *Itpa* cDNA (731 bp) (top panel) and the 18S rRNA probe[49] (middle panel), as previously described.[50] The arrowhead indicates *Itpa* mRNA. In the bottom panel, the relative amounts of *Itpa* mRNA to *18S rRNA* were calculated based on the radioactivity. The ratio of the relative amount of each transcript to that in the testis is shown. *B*. The expression of *Itpa* mRNA in quiescent and serum-stimulated BALB/c3T3 cells. The total RNA isolated was subjected to Northern blot analyses to determine the expression amount of *Itpa* mRNA and 18S rRNA. *C*. The relative amounts of *Itpa* mRNA to 18S rRNA were calculated based on radioactivity. The ratio of the relative amount of each transcript to that in the quiescent cells is shown (closed circle). The percentages of the cells in the S phase of the cell cycle were determined by flow cytometry as previously described[28] and then were plotted with a dotted line.

TATA-less promoter with a sequence (CTACTTC) which exactly matched the consensus sequence for the initiator (PyPyANT/APyPy), just upstream of the 5′ end of the longest *Itpa* cDNA (Fig. 5B),[34] indicating that the *Itpa* gene is likely to be one of the housekeeping genes which are ubiquitously expressed. A consensus-like sequence for ELK1 or ETS-1 proteins (ACMGGAAGTNC, ACMGGAWRTT), which are known to be activated during serum stimulation,[35,36] and one for GATA transcription factors (WGATAR)[37] were found in the *Itpa* promoter region (Fig. 5B). Those factors may be responsible for the inducible expression of *Itpa* in some tissues.

3.6. Expression and distribution of mouse ITPA protein in mouse tissue

In extract prepared from mouse brain, only a single band corresponding to a polypeptide with a molecular

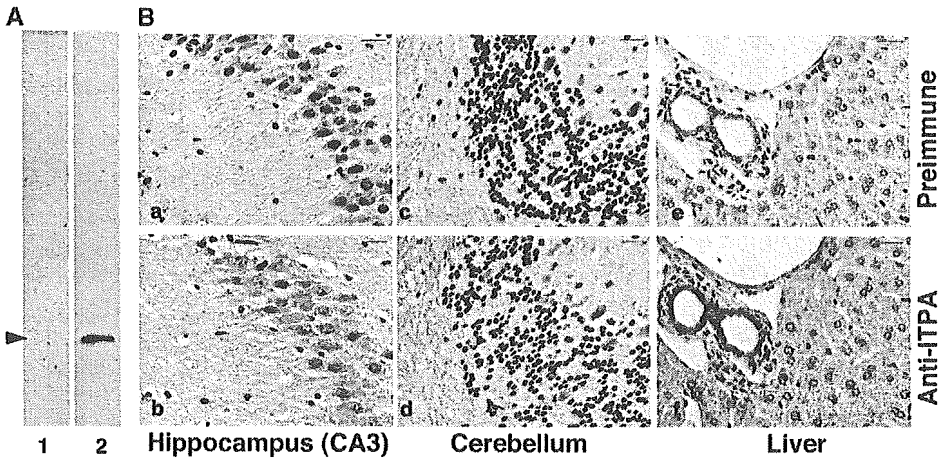


Figure 8. The expression of mouse ITPA protein in mouse tissues. *A.* Western blotting analysis. Brain extracts (20 μg protein) were separated on 12.5% SDS-PAGE and subjected to Western blotting as previously described,[31] with preimmune serum (lane 1) or anti-ITPA serum (lane 2). The arrowhead indicates the 22-kDa polypeptide. *B.* Immunohistochemistry. Brain (panels a–d) and liver (panels e, f) sections embedded in paraffin blocks were subjected to immunohistochemical analyses with the anti-ITPA serum (panels b, d, f) and preimmune serum (panels a, c, e) as controls. The nuclei were counterstained with hematoxylin (blue).

weight of 22 kDa was reacted with the anti-ITPA serum but not with preimmune serum (Fig. 8A), thus indicating that the anti-ITPA serum specifically reacts with mouse ITPA in the extract. We could not detect smaller polypeptides in either cellular or tissue extracts reacted with the anti-ITPA serum; therefore, it remains to be established as to whether type B or C *Itpa* transcripts produce any polypeptide *in vivo*.

Next, brain and liver sections embedded in paraffin blocks were subjected to immunohistochemistry using the anti-ITPA and preimmune sera as controls (Fig. 8B). In the brain, the anti-ITPA serum exhibited apparent immunoreactivity throughout the brain section in comparison to the preimmune serum (data not shown). In the brain section, most of the hippocampal neurons in CA1 to CA3 and DG exhibited relatively strong ITPA immunoreactivity, especially in the soma of CA3 pyramidal cells and mossy fibers (Fig. 8Ba, b). In the cerebellum, the cell bodies of Purkinje cells and mossy fibers in the cerebellar white matter exhibited a strong ITPA immunoreactivity, while granule cell bodies exhibited a relatively weak immunoreactivity (Fig. 8c, d). In the liver, hepatocytes exhibited an evenly distributed ITPA immunoreactivity mainly in the cytoplasm, and a significantly strong immunoreactivity was seen in the cytoplasm and nuclei of the epithelial cells lined bile ducts, and to a lesser extent those of the endothelial cells lined the portal vein (Fig. 8Be, f).

The expression profile of *Itpa* suggests that the ITPA function may be required for proliferative tissues as well as postmitotic neurons. In neurons, ITPA protein was mostly detected in the cytoplasm, to a lesser extent in the nucleus and also in nerve fibers. A particularly high level of expression was apparent in hippocampal CA3 pyramidal cells and Purkinje cells in the cerebellum. Those neurons were postmitotic, thus suggesting that ITPA may function to sanitize the nucleotide pools for RNA synthesis or DNA synthesis in the mitochondria. However, the PSORT II program (http://psort.nibb.ac.jp/) to predict the protein localization sites in cells,[38] predicted that human and mouse ITPA proteins are likely to mostly localize in the cytoplasm and some in the nucleus but not in the mitochondria, and thus ITPA may not function in mitochondria.

On the other hand, the localization of ITPA in nerve fibers may indicate that ITP/dITP or XTP/dXTP, substrate nucleotides for ITPA, may be toxic for the neural function when such abnormal nucleotides accumulated in the nerve fibers. It is well known that nitric oxide (NO), a neurotransmitter,[39] promotes the deamination of various molecules such as nucleic acids and proteins,[40–42] thus increasing the intracellular concentration of ITP/dITP or XTP/dXTP which can be generated by the deamination of ATP/dATP or GTP/dGTP.[20] It has recently been shown that ITP or XTP disturbs the small G protein function through competition with GTP,[43–45] thus suggesting that the hydrolysis of ITP/XTP by ITPA is critical for maintaining such signal transduction through small G proteins, especially in neurons. As a result, ITPA may protect the neurons from damage caused by these abnormal nucleotides.

In the liver, a much higher level of ITPA was detected in the epithelial cell-lined bile duct and also in the endothelial cell-lined portal vein than in hepatocytes. In the endothelial cells, again NO plays an important role in regulating vasoconstriction, thus a high level of ITPA may be required.[46] In the bile duct, a high level of ITPA was detected in the nuclei of epithelial cells, suggesting that those cells have greater exposure to deamination by chemicals excreted through the bile duct as well as

NO.[47,48]

In conclusion, among various genomic sequences highly homologous to human ITPA coding sequences in the mouse, the mouse ortholog (*mItpa*) for hITPA was successfully isolated by the retro-recombination method, while a processed-*Itpa* gene like sequence and two pseudogenes were also identified in the mouse genome. Since no transcribed sequence derived from the processed *Itpa*-like sequence was found in any of the examined mouse tissue specimens, we thus conclude that the mouse has only one functional *Itpa* gene, which is highly expressed in the testis, brain and thymus.

Acknowledgements: We thank Dr. Daniel Nathans for BALB/c 3T3 cells, Dr. Motoya Katsuki for CCE ES cells, Drs. Masato Furuichi and Yoshimichi Nakatsu for their helpful discussions, Setsuko Kitamura and Keiko Aiura for their technical assistance, Dr. B. Quinn for comments on the manuscript. This work was supported by grants from CREST, Japan Science and Technology Agency, the Ministry of Education, Culture, Sports, Science, and Technology of Japan (grant number: 16012248), and the Japan Society for the Promotion of Science (grant numbers: 15590347, 16390119).

References

1. Nakabeppu, Y., Tsuchimoto, D., Furuichi, M., and Sakumi, K. 2004, The Defense Mechanisms in Mammalian Cells Against Oxidative Damage in Nucleic Acids and Their Involvement in the Suppression of Mutagenesis and Cell Death, *Free Radic. Res.*, **38**, 423–429.
2. Kamiya, H. and Kasai, H. 1995, Formation of 2-hydroxydeoxyadenosine triphosphate, an oxidatively damaged nucleotide, and its incorporation by DNA polymerases. Steady-state kinetics of the incorporation, *J. Biol. Chem.*, **270**, 19446–19450.
3. Sekiguchi, M. 1996, MutT-related error avoidance mechanism for DNA synthesis, *Genes Cells*, **1**, 139–145.
4. Tye, B. K., Nyman, P. O., Lehman, I. R., Hochhauser, S., and Weiss, B. 1977, Transient accumulation of Okazaki fragments as a result of uracil incorporation into nascent DNA, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 154–157.
5. Hochhauser, S. J. and Weiss, B. 1978, *Escherichia coli* mutants deficient in deoxyuridine triphosphatase, *J. Bacteriol.*, **134**, 157–166.
6. Gadsden, M. H., McIntosh, E. M., Game, J. C., Wilson, P. J., and Haynes, R. H. 1993, dUTP pyrophosphatase is an essential enzyme in *Saccharomyces cerevisiae*, *Embo J.*, **12**, 4425–4431.
7. Kouzminova, E. A. and Kuzminov, A. 2004, Chromosomal fragmentation in dUTPase-deficient mutants of *Escherichia coli* and its recombinational repair, *Mol. Microbiol.*, **51**, 1279–1295.
8. Maki, H. and Sekiguchi, M. 1992, MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis, *Nature*, **355**, 273–275.
9. Fujii, Y., Shimokawa, H., Sekiguchi, M., and Nakabeppu, Y. 1999, Functional significance of the conserved residues for the 23-residue module among MTH1 and MutT family proteins, *J. Biol. Chem.*, **274**, 38251–38259.
10. Tsuzuki, T., Egashira, A., Igarashi, H. et al. 2001, Spontaneous tumorigenesis in mice defective in the *MTH1* gene encoding 8-oxo-dGTPase, *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 11456–11461.
11. Nakabeppu, Y. 2001, Molecular genetics and structural biology of human MutT homolog, MTH1, *Mutat. Res.*, **477**, 59–70.
12. Sakai, Y., Furuichi, M., Takahashi, M. et al. 2002, A molecular basis for the selective recognition of 2-hydroxy-dATP and 8-Oxo-dGTP by human MTH1, *J. Biol. Chem.*, **277**, 8579–8587.
13. Bessman, M. J., Frick, D. N., and O'Handley, S. F. 1996, The MutT proteins or "Nudix" hydrolases, a family of versatile, widely distributed, "housecleaning" enzymes, *J. Biol. Chem.*, **271**, 25059–25062.
14. Ishibashi, T., Hayakawa, H., and Sekiguchi, M. 2003, A novel mechanism for preventing mutations caused by oxidation of guanine nucleotides, *EMBO Rep.*, **4**, 479–483.
15. Cai, J. P., Ishibashi, T., Takagi, Y., Hayakawa, H., and Sekiguchi, M. 2003, Mouse MTH2 protein which prevents mutations caused by 8-oxoguanine nucleotides. Biochem, *Biophys. Res. Commun.*, **305**, 1073–1077.
16. Kamiya, H., Iida, E., and Harashima, H. 2004, Important amino acids in the phosphohydrolase module of *Escherichia coli* Orf135, *Biochem. Biophys. Res. Commun.*, **323**, 1063–1068.
17. Nunoshiba, T., Ishida, R., Sasaki, S., Iwai, S., Nakabeppu, Y., and Yamamoto, K. 2004, A novel Nudix hydrolase for oxidized purine nucleoside triphosphates encoded by *ORFYLR151c* (*PCD1* gene) in *Saccharomyces cerevisiae*, *Nucleic Acids Res.*, **32**, 5339–5348.
18. Hwang, K. Y., Chung, J. H., Kim, S. H., Han, Y. S., and Cho, Y. 1999, Structure-based identification of a novel NTPase from *Methanococcus jannaschii*, *Nat. Struct. Biol.*, **6**, 691–696.
19. Lin, S., McLennan, A. G., Ying, K. et al. 2001, Cloning, expression, and characterization of a human inosine triphosphate pyrophosphatase encoded by the *ITPA* gene, *J. Biol. Chem.*, **276**, 18695–18701.
20. Chung, J. H., Park, H. Y., Lee, J. H., and Jang, Y. 2002, Identification of the dITP- and XTP-hydrolyzing protein from *Escherichia coli*, *J. Biochem. Mol. Biol.*, **35**, 403–408.
21. Clyman, J. and Cunningham, R. P. 1987, *Escherichia coli* K-12 mutants in which viability is dependent on recA function, *J. Bacteriol.*, **169**, 4203–4210.
22. Bradshaw, J. S. and Kuzminov, A. 2003, RdgB acts to avoid chromosome fragmentation in *Escherichia coli*, *Mol. Microbiol.*, **48**, 1711–1725.
23. Vanderheiden, B. S. 1964, Inosine triphosphate in human erythrocytes: a genetic treat. *Proc. Xth. Congress Int. Soc. Blood. Transf., Stockholm*, pp. 540–548.
24. Cao, H. and Hegele, R. A. 2002, DNA polymorphisms in *ITPA* including basis of inosine triphosphatase deficiency, *J. Hum. Genet.*, **47**, 620–622.
25. Sumi, S., Marinaki, A. M., Arenas, M. et al. 2002, Genetic basis of inosine triphosphate pyrophosphohydrolase deficiency, *Hum. Genet.*, **111**, 360–367.
26. Hirano, S., Tominaga, Y., Ichinoe, A. et al. 2003, Muta-

tor Phenotype of MUTYH-null Mouse Embryonic Stem Cells, *J. Biol. Chem.*, **278**, 38121–38124.
27. Ide, Y., Tsuchimoto, D., Tominaga, Y., Iwamoto, Y., and Nakabeppu, Y. 2003, Characterization of the genomic structure and expression of the mouse *Apex2* gene, *Genomics*, **81**, 47–57.
28. Nakabeppu, Y., Oda, S., and Sekiguchi, M. 1993, Proliferative activation of quiescent Rat-1A cells by ΔFosB, *Mol. Cell. Biol.*, **13**, 4157–4166.
29. Woltjen, K., Bain, G., and Rancourt, D. E. 2000, Retrorecombination screening of a mouse embryonic stem cell genomic library, *Nucleic Acids Res.*, **28**, E41.
30. Nakabeppu, Y. and Nathans, D. 1991, A naturally occurring truncated form of FosB that inhibits Fos/Jun transcriptional activity, *Cell*, **64**, 751–759.
31. Tsuchimoto, D., Sakai, Y., Sakumi, K. et al. 2001, Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen, *Nucleic Acids Res.*, **29**, 2349–2360.
32. Burset, M., Seledtsov, I. A., and Solovyev, V. V. 2000, Analysis of canonical and non-canonical splice sites in mammalian genomes, *Nucleic Acids Res.*, **28**, 4364–4375.
33. Burset, M., Seledtsov, I. A., and Solovyev, V. V. 2001, SpliceDB: database of canonical and non-canonical mammalian splice sites, *Nucleic Acids Res.*, **29**, 255–259.
34. Butler, J. E. and Kadonaga, J. T. 2002, The RNA polymerase II core promoter: a key component in the regulation of gene expression, *Genes Dev.*, **16**, 2583–2592.
35. Rao, V. N. and Reddy, E. S. 1992, A divergent ets-related protein, elk-1, recognizes similar c-ets-1 proto-oncogene target sequences and acts as a transcriptional activator, *Oncogene*, **7**, 65–70.
36. Woods, D. B., Ghysdael, J., and Owen, M. J. 1992, Identification of nucleotide preferences in DNA sequences recognised specifically by c-Ets-1 protein, *Nucleic Acids Res.*, **20**, 699–704.
37. Merika, M. and Orkin, S. H. 1993, DNA-binding specificity of GATA family transcription factors, *Mol. Cell Biol.*, **13**, 3999–4010.
38. Nakai, K. and Horton, P. 1999, PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization, *Trends Biochem. Sci.*, **24**, 34–36.
39. Boehning, D. and Snyder, S. H. 2003, Novel neural modulators, *Annu. Rev. Neurosci.*, **26**, 105–131.
40. Burney, S., Tamir, S., Gal, A., and Tannenbaum, S. R. 1997, A mechanistic analysis of nitric oxide-induced cellular toxicity, *Nitric Oxide*, **1**, 130–144.
41. Kow, Y. W. 2002, Repair of deaminated bases in DNA, *Free Radic. Biol. & Med.*, **33**, 886–893.
42. Dedon, P. C. and Tannenbaum, S. R. 2004, Reactive nitrogen species in the chemical biology of inflammation, *Arch. Biochem. Biophys.*, **423**, 12–22.
43. Klinker, J. F. and Seifert, R. 1997, Functionally nonequivalent interactions of guanosine 5′-triphosphate, inosine 5′-triphosphate, and xanthosine 5′-triphosphate with the retinal G-protein, transducin, and with G_i-proteins in HL-60 leukemia cell membranes, *Biochem. Pharmacol.*, **54**, 551–562.
44. Seifert, R., Gether, U., Wenzel-Seifert, K., and Kobilka, B. K. 1999, Effects of guanine, inosine, and xanthine nucleotides on β_2-adrenergic receptor/G_s interactions: evidence for multiple receptor conformations, *Mol. Pharmacol.*, **56**, 348–358.
45. Liu, H. Y. and Seifert, R. 2002, Distinct interactions of $G_{s\alpha\text{-long}'}$, $G_{s\alpha\text{-short}'}$, and $G_{\alpha olf}$ with GTP, ITP, and XTP, *Biochem. Pharmacol.*, **64**, 583–593.
46. Cahill, P. A., Redmond, E. M., and Sitzmann, J. V. 2001, Endothelial dysfunction in cirrhosis and portal hypertension, *Pharmacol. Ther.*, **89**, 273–293.
47. Spirli, C., Fabris, L., Duner, E. et al. 2003, Cytokine-stimulated nitric oxide production inhibits adenylyl cyclase and cAMP-dependent secretion in cholangiocytes, *Gastroenterol.*, **124**, 737–753.
48. Trauner, M. 2003, When bile ducts say NO: the good, the bad, and the ugly, *Gastroenterol.*, **124**, 847–851.
49. Financsek, I., Mizumoto, K., and Muramatsu, M. 1982, Nucleotide sequence of the transcription initiation region of a rat ribosomal RNA gene, *Gene*, **18**, 115–122.
50. Oda, S., Nishida, J., Nakabeppu, Y., and Sekiguchi, M. 1995, Stabilization of cyclin E and cdk2 mRNAs at G1/S transition in Rat-1A cells emerging from the G0 state, *Oncogene*, **10**, 1343–1351.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
DNA Repair 4 (2005) 826–835
DNA REPAIR
www.elsevier.com/locate/dnarepair

XRCC1 interactions with multiple DNA glycosylases: A model for its recruitment to base excision repair

Anna Campalans[a,1], Stéphanie Marsin[a,1], Yusaku Nakabeppu[b], Timothy R. O'Connor[c], Serge Boiteux[a], J. Pablo Radicella[a,*]

[a] *Département de Radiobiologie et Radiopathologie, CEA, UMR217 CNRS/CEA, 18 Route du Panorama, F-92265 Fontenay aux Roses, France*
[b] *Department of Immunobiology and Neuroscience, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-8582, Japan*
[c] *Division of Biology, Beckman Research Institute of the City of Hope, Duarte, CA 91010, USA*

Received 18 February 2005; received in revised form 26 April 2005; accepted 27 April 2005
Available online 31 May 2005

Abstract

Repair of chemically modified bases in DNA is accomplished through base excision repair (BER). This pathway is initiated by a specific DNA glycosylase that recognizes and excises the altered base to yield an abasic (AP) site. After cleavage of the AP site by APE1, repair proceeds through re-synthesis and ligation steps. In mammalian cells, the XRCC1 protein, essential for the maintenance of genomic stability, is involved in both base excision and single-strand break repair. XRCC1 participates in the first step of BER by interacting with the human DNA glycosylases hOGG1 and NEIL1. To analyze the possibility of a general mechanism involving the interaction of XRCC1 with DNA glycosylases we used XRCC1 to pull-down DNA glycosylases activities from human cell extracts. XRCC1 co-purifies with DNA glycosylase activities capable of excising hypoxanthine and dihydrothymine, in addition to 8-oxoguanine, but not uracil. Biochemical analyses with the purified proteins confirmed the interactions between XRCC1 and MPG, hNTH1 or hNEIL2. Furthermore, XRCC1 stimulates the activities of these enzymes. In vivo localization studies show that after genotoxic treatments these DNA glycosylases can be found associated with XRCC1 foci. Our results support a BER model in which XRCC1 is recruited to the repair of alkylated or oxidized bases by the enzyme recognizing the lesion. XRCC1 would then coordinate the subsequent enzymatic steps and modulate the activities of all the proteins involved.

Keywords: XRCC1; Base excision repair; DNA glycosylases

1. Introduction

Base damage, abasic (AP) sites and strand breaks, are common lesions threatening the integrity of cellular DNA. These various types of lesions can be spontaneously generated, or the consequence of the reaction of DNA with endogenous or environmental compounds [1]. Base excision repair (BER) is the main enzymatic pathway for the repair of modified bases and abasic sites [2]. Single-strand breaks, which can also be generated as BER intermediates, are repaired by the single-strand break repair (SSBR) pathway [3]. In human cells' BER, the damaged base is recognised by a specific DNA glycosylase that removes the base to generate an AP site, which in turn is the substrate for APE1, the major AP endonuclease. This latter enzyme is responsible for the recognition and cleavage of AP sites. Repair of the majority of the incised AP sites is completed by addition of a nucleotide to the 3′-OH and removal of the 5′-dRP by DNA polymerase β (POLβ). The resulting nick is sealed by the action of the DNA ligase 3 (LIG3)–XRCC1 complex. The SSBR pathway shares several enzymes with BER. Indeed, after recognition of the break by PARP1, XRCC1 is recruited to the break [4,5] and depending on their nature, enzymes required for the cleaning of the DNA extremities, APE1 or polynucleotide kinase (PNK), prepare the substrate for the synthesis and/or ligation steps performed by POLβ and XRCC1–LIG3, respectively.

* Corresponding author. Tel.: +33 1 46 54 88 57; fax: +33 1 46 54 88 57.
E-mail address: jpradicella@cea.fr (J.P. Radicella).
[1] These authors contributed equally to this work.

1568-7864/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.dnarep.2005.04.014

In spite of its lack of known enzymatic activity, the involvement of XRCC1 in both, BER and SSBR pathways is supported by substantial biochemical evidence. Its participation in the later steps of both pathways was shown through its interactions with all the enzymes common to both pathways: PARP1 and 2 [6–8], POLβ [6,9,10] and LIG3 [11]. More recent results have implicated XRCC1 in enzymatic steps specific for each of the pathways. For SSBR, the interaction with and stimulation of PNK [12], as well as its association with tyrosyl DNA phosphodiesterase (TDP1) [13], show that XRCC1 interacts with all the known SSBR proteins. Likewise, its physical and functional interaction with APE1 [14] suggested a similar situation for BER of AP sites. More recently, two reports indicate the involvement of XRCC1 in all the steps of the repair of oxidised bases by its interactions with the human 8-oxoguanine (8-oxoG) DNA glycosylase hOGG1 [15] and hNEIL1 [16]. Because of its multiple interactions, it was proposed that XRCC1 acts as a scaffold protein coordinating BER and SSBR enzymatic steps to avoid the exposure of potentially toxic DNA intermediates to the cellular milieu [17,18]. As shown by the recruitment of XRCC1 by PARP1 in the case of SSBR [4,5], previous results on BER of AP sites [14] and 8-oxoG [15] suggested a model by which XRCC1 is recruited by the enzyme initiating repair and stays associated to the substrate throughout the repair process. Here, we extend this model to the human BER of other modified bases, namely oxidised pyrimidines and alkylated bases.

2. Materials and methods

2.1. Plasmid constructions

A plasmid clone HRC08117 carrying full-length cDNA for human NEIL1 (DDBJ/EMBL/GenBank accession no. AK026055) in pME18SFL3 (DDBJ/EMBL/GenBank accession no. AB009864) was obtained from Dr. Sumio Sugano (Tokyo University), and a plasmid clone NT2RI2003751 carrying full-length cDNA for human NEIL2 (DDBJ/EMBL/GenBank accession no. AK056206) in pME18SFL3 was obtained from NITE Biological Resource Center (NBRC), Japan. hNEIL2 cDNA was amplified from plasmid pME18SFL3-hNEIL2 with addition of the coding sequence for 6 histidines at its 3′-end and inserted into the EcoR1/Sma1 site of pTRc99A vector (Amersham) to create plasmid pPR363. To generate the yellow fusion protein (YFP)-tagged XRCC1 expression construct, human XRCC1 cDNA was amplified by PCR with 5′ and 3′ primers containing *Eco*RI and *Apa*I sites, respectively. The amplified DNA fragment was subcloned into the *Eco*RI and *Apa*I sites of pEYFP-N1 (CLONTECH). To generate the myc-tagged proteins hNTH1, hNEIL1, hNEIL2 and MPG, the cDNAs were amplified with 5′ and 3′ primers containing *Eco*RI and *Xho*I sites, respectively. The PCR products were subcloned into pCDNA3.1/myc-His A(+) (Invitrogen). The OGG1-myc construct was previously described [15].

2.2. Oligonucleotide substrates

The 34mer oligodeoxynucleotides containing an 8-oxoguanine, a dihydrothymine (DHT) (both gifts from J. Cadet, CEA, Grenoble), an inosine (Sigma–Genosys) or a uracil (Sigma–Genosys) were labelled at the 5′ end using [γ-^{32}P]ATP (3000 Ci/mmol; Amersham) and T4 polynucleotide kinase (New England Biolabs). The ^{32}P-labeled strand was hybridised with a complementary sequence containing a cytosine (C) opposite to 8-oxoG and to DHT a thymine opposite to inosine or a guanine opposite to uracil.

2.3. Enzymes

The recombinant hOGG1, XRCC1 and MPG proteins were purified as described [15,19,20]. hOGG1 was further purified by fixation to an heparin column. The hNTH1-pET14b expression vector was kindly given by Ian Hickson (Oxford University). *Escherichia coli* BL21 CodonPlus (Stratagene) harbouring the plasmid was grown at 37 °C until $OD_{600} = 0.3$. After the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to 0.5 mM, the cell culture was continued at 20 °C for 16 h. Cells from 3 l of culture ($OD_{600} = 3$) were harvested by centrifugation, resuspended and sonicated in 300 ml ice-cold lysis buffer (50 mM Tris–HCl, 500 mM NaCl, 10% glycerol, 0.01% TritonX100, 1 mM DTT, 0.1 mM EDTA, pH 8) with antipaïne 2.5 μg/ml, aprotrinine 2.5 μg/ml, leupeptine 2.5 μg/ml and lysozyme 1 mg/ml. The lysate was centrifuged at 100,000 × *g* for 30 min at 4 °C, and the supernatant containing his-tagged hNTH1 was incubated with 2 ml of nickel resin (Amersham). The column was washed first with lysis buffer containing 40 mM imidazol, and then equilibrated with A100 buffer (50 mM $NaPO_4$, 100 mM NaCl, 5% glycerol, 0.01% TritonX100, 1 mM DTT, 0.1 mM EDTA, pH 7.3) containing 40 mM imidazol. hNTH1 was eluted in A100 buffer containing 300 mM imidazol. The protein was loaded on a 1 ml Hitrap heparin column (Amersham) equilibrated in A100 buffer and eluted with a linear NaCl gradient in buffer A and collected at 400 mM NaCl. The protein was desalted on PD10 column (Amersham) equilibrated in A100 buffer, loaded on a 1 ml Hitrap SP column (Amersham), eluted with a linear NaCl gradient in buffer A and collected at 400 mM NaCl.

To purify hNEIL2, *E. coli* BL21 CodonPlus cells harbouring pPR363 were grown at 37 °C until $OD_{600} = 0.5$. After addition of 0.2 mM IPTG, the cell culture was continued at 20 °C for 16 h. Bacteria from 3 l of culture ($OD_{600} = 3$) were harvested by centrifugation, resuspended and sonicated in 300 ml ice-cold phosphate lysis buffer (50 mM NaPO4, 500 mM NaCl, 10% glycerol, 0.01% TritonX100, 1 mM DTT, 0.1 mM EDTA, pH 7.3) with antipaïne 2.5 μg/ml, aprotrinine 2.5 μg/ml, leupeptine 2.5 μg/ml and lysozyme 1 mg/ml. The lysate was centrifuged at 100,000 × *g* for 30 min at 4 °C, and the supernatant containing his-tagged hNEIL2 was incubated with 1 ml of nickel resin (Amersham). The column was washed with lysis buffer contain-

ing 2 mM imidazol, and then equilibrated with A50 buffer (50 mM $NaPO_4$, 50 mM NaCl, 5% glycerol, 0.01% TritonX100, 1 mM DTT, 0.1 mM EDTA, pH 7.3) containing 2 mM Imidazol. hNEIL2 was eluted in A50 buffer containing 300 mM Imidazol. The protein was loaded on a 1 ml Hitrap SP column (Amersham) equilibrated in A50 buffer and eluted with a linear NaCl gradient in buffer A and collected at 500 mM NaCl. The protein was desalted on PD10 column (Amersham) equilibrated in A50 buffer, loaded on a 1 ml Hitrap heparin column (Amersham), eluted with a linear NaCl gradient in buffer A and collected at 600 mM NaCl.

2.4. MBP fusion proteins expression

Plasmids pMAL-c2x harbouring the different XRCC1 domains [15] were introduced in BL21-Rosetta cells (Invitrogen). Cultures and inductions were performed according to the manufacturer's instructions (New England Biolabs). Pellet of induced cells were resuspended and sonicated at OD = 30 in MBP lysis buffer (20 mM $NaPO_4$, 1 mM EDTA, 1 mM DTT, 0.1% nonidet NP40, pH 7.3) supplemented with 500 mM NaCl, antipaïne 2.5 μg/ml, aprotrinine 2.5 μg/ml, leupeptine 2.5 μg/ml and lysozyme 1 mg/ml. Lysates were centrifuged at 4 °C during 30 min at 10,000 × *g* and supernatants used in MBP pull-down assays.

For each interaction, 30 μl of Amylose Resin (New England Biolabs), equilibrated in MBP lysis buffer was incubated, for 20 min at 4 °C in a shaker, with an excess (1 ml) of protein extract prepared as described above. The amylose resin was recovered by centrifugation and washed twice with 1 ml of MBP lysis buffer containing 500 mM NaCl. The amylose resine was then alternatively used for interaction with HeLa crude extracts or purified glycosylase.

2.5. MBP pull-down assays with HeLa crude extracts

Exponential growing HeLa cells were washed twice in PBS and harvested. The pellet was resuspended and sonicated in 4 volumes of HeLa lysis buffer (20 mM Tris–Hcl, 250 mM NaCl, 1 mM EDTA, pH 8). Lysates were centrifuged at 4 °C for 45 min at 100,000 × *g* and protein concentrations were estimated by Bradford analysis. A 30 μl of MBP or MBP–XRCC1 interacting resin was then equilibrated in MBP lysis buffer containing 200 mM NaCl and incubated in the presence or absence of 200 μg of HeLa crude extract for 2 h at 4 °C. The resin was washed twice with 1 ml of MBP lysis buffer containing 200 mM NaCl and then equilibrated in the reaction buffer corresponding to the glycosylase activity tested. The reactions were performed directly in the resin by addition of the labelled oligonucleotide substrate.

2.6. MBP pull-down assays with purified proteins

A 1.5 μg of purified His-tagged hNTH1 or hNEIL2 in 60 μl of MBP lysis buffer was added to 30 μl of MBP-fusion protein binding resin and incubation was pursued for 20 min at 4 °C. The resin was washed twice with 1 ml of MBP lysis buffer containing 50 mM NaCl. Bound proteins were separated by SDS-10% PAGE, transferred to nitrocellulose membranes and analysed by Western blot analysis with peroxydase coupled antibody (Roche). In parallel, we checked by Coomassie staining that the same amount of each MBP-fused proteins was bound to Amylose.

2.7. DNA glycosylase assays

In a standard reaction (10 μl final volume), 50 fmol of the indicated labelled duplex was incubated with various amounts of glycosylase in glycosylase specific reaction buffer: OGG1 reaction buffer (50 mM Hepes, 2 mM EDTA, 1 mM DTT, 50 mM NaCl, 5% glycerol, 0.1 mg/ml bovine serum albumin (BSA), pH 7.6); MPG reaction buffer (50 mM Hepes–KOH, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 5% glycerol, 0.1 mg/ml BSA, pH 7.8); NTH reaction buffer (40 mM Hepes–KOH, 100 mM Kcl, 0.5 mM EDTA, 1 mM DTT, 0.2 mg/ml BSA, pH 7.8); UNG reaction buffer (50 mM Hepes–KOH, 20 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.8). Reactions were carried out at 37 °C for the indicated time, and stopped on ice with NaOH (0.1N final concentration). After 5 min incubation at 37 °C, reactions were stopped with formamide dye and heated for 5 min at 95 °C. Reaction products were resolved by denaturing 20% PAGE (19:1 acrylamide:bisacrylamide). Gels were scanned and band intensities quantified using a Storm Phosphorimager (Molecular Dynamics).

The borohydride trapping assays were carried out at 37 °C for 15 min with 10 fmol of DHT:C labelled duplex and various amounts of hNTH1 in the presence of the indicated amounts of XRCC1. $NaBH_4$ at a final concentration of 50 mM was then added and incubation continued for 10 min at 37 °C. The reaction was terminated by addition of Laemli buffer (1X final) and heated at 95 °C for 5 min. Trapped complexes were separated by SDS-PAGE. The gels were dried and scanned using a Storm Phosphorimager (Molecular Dynamics).

2.8. Human cells transfection and treatment

HeLa cells were grown on glass coverslips in DMEM (Cambrex) containing 10% foetal calf serum, glutamine and penicilline/streptomycine and transfected using Effectene Transfection Reagent (QIAGEN) according to the manufacturer's protocol. Twenty-four hours after transfection, cells were rinsed in PBS and treated with either 5 mM H_2O_2 in PBS at room temperature for 30 min or 10 μg/ml MMS in DMEM for 20 min at 37 °C. Rinsed coverslips were then incubated for 1 h in drug-free medium at 37 °C.

2.9. Immunofluorescence and confocal microscopy

Cells were fixed with 4% paraformaldehyde for 15 min at 37 °C, rinsed with PBS and permeabilized at room temperature in PBS–0.1% Triton for 10 min. Cells were incubated in

blocking solution (PBS, 0.1% Triton, 3% BSA, 1% normal goat serum) at 37 °C for 1 h. For labelling the myc-tagged proteins, cells were incubated for 1 h at 37 °C with a mouse monoclonal anti-myc (0.1 μg/μl; clon 9E10, oncogene) in blocking solution. Cells were then washed three times for 10 min in PBS-0.1% Triton and incubated with Alexa Fluor 594 goat anti-mouse IgG Antibodies (Molecular Probes) at 1:1000 dilution in blocking solution for 1 h at 37 °C. Nuclear DNA was counterstained with 4′,6′-diamidino-2-phenylindole (DAPI, 1 μg/ml). Coverslips were mounted in DAKO fluorescent mounting medium. Image acquisition was performed with a Leica DM RxA2 confocal microscope (Wetzlar, Germany).

Images were processed with the MetaMorph image processing software (Universal Imaging, West Chester, PA). In order to measure the co-localisation we applied a median filter (3), followed by an FFT (fast Fourier transformation; high pass, sensitivity 6) and the images were tresholded (autotreshold for light objects) in order to remove the background. To allow a better visualisation of the common elements between the green and red images, the thresholded images were binarized and processed with the arithmetic–logical AND tool to obtain the image labelled AND in the figures. The "measure of co-localization" tool from MetaMorph was used to calculate the percentage of co-localisation between the two fluorochromes.

3. Results

3.1. DNA glycosylase activities associated with XRCC1 in human whole cell extracts

To explore whether in addition to hOGG1 [15] other DNA glycosylases interact with XRCC1, a MBP–XRCC1 fusion protein was used in pull-down experiments on HeLa whole cell extracts (WCE). The protein fractions associated with the MBP–XRCC1 fusion were analyzed for their DNA glycosylase activities on duplex oligonucleotides carrying modified bases. To validate the experimental protocol, we first tested the known XRCC1–hOGG1 interaction by using an 8-oxoG substrate on the MBP–XRCC1 pulled-down fraction and confirmed the association of the 8-oxoG DNA glycosylase activity to XRCC1 (Fig. 1A). Using as substrate an oligonucleotide duplex with an inosine residue, an hypoxanthine DNA glycosylase activity was specifically recovered from the HeLa WCE when the pull-down experiment was performed with the MBP–XRCC1 fusion but not the MBP alone (Fig. 1B). When a DNA carrying dihydrothymidine, an oxidized pyrimidine, was used, a cleavage activity was also specifically detected as co-purifying with the MBP–XRCC1 fusion (Fig. 1C). Although a uracil DNA glycosylase activity was detected under the same experimental conditions, similar levels of activity were recovered associated to MBP without XRCC1. Because of the high activity detected in human cell extracts, this is likely to be due to the non-specific binding of a human uracil DNA glycosylase to the resin. Indeed, increasing the stringency of the washes resulted in the loss of uracil DNA glycosylase activity pulled-down by both MBP and MBP–XRCC1 columns (data not shown). The experiments described above show that in cellular extracts XRCC1 can associate, either directly or indirectly, with proteins displaying DNA glycosylase activities on alkylated or oxidized bases. This suggests that DNA glycosylases other than hOGG1 could exist in protein complexes with XRCC1 at the initiation steps of BER. The candidate proteins for carrying the DNA glycosylase activities are MPG in the case of inosine substrates [20] and either hNTH1, hNEIL1 or hNEIL2 for the oxidized pyrimidines [21].

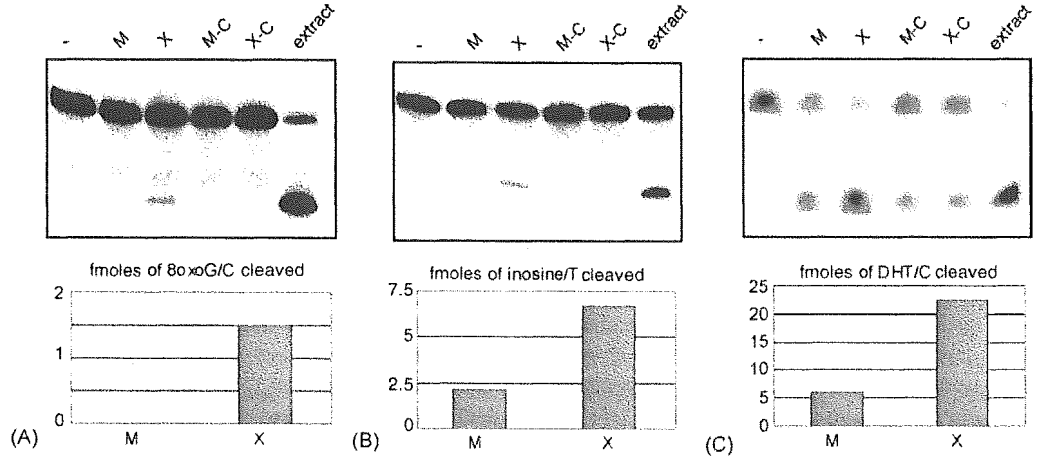


Fig. 1. DNA Glycosylase activities associated with XRCC1. Amylose resins loaded with MBP (M) or MBP–XRCC1 (X) were incubated with HeLa crude extract. After several washes, the resin was incubated for 30 min at 37 °C in the presence of the appropriate labelled substrate. The reactions were stopped and the products resolved by PAGE. (A) 8-oxoG glycosylase activity; (B) inosine glycosylase activity; (C) DHT glycosylase activity. M-C and X-C correspond to parallel experiments where the HeLa cell extract was omitted to control for bacterial-derived contaminating activities. Lane labelled "extract" corresponds to the whole cell extract (3 μg of protein) activity.

3.2. XRCC1 interactions with hNTH1, hNEIL2 and MPG

Purified hNTH1, hNEIL2 and MPG were used to test for a direct interaction between XRCC1 with the DNA glycosylases. The fusion MBP–XRCC1 protein was bound to an amylose resin and incubated in the presence of either purified histidine-tagged hNTH1 or hNEIL2, or MPG. Pull-down assays were then performed and the presence of the glycosylases was analyzed by immunoblotting with an antibody against the histidine tag or by their activity in the case of MPG. hNTH1, hNEIL2 (Fig. 2A) and MPG (Fig. 2B, left panel) are recovered by the XRCC1-bound resin but not the MBP-bound resin. This result indicates a direct physical interaction between XRCC1 and the two oxidized pyrimidine-specific DNA glycosylases hNTH1 and hNEIL2 as well as with MPG. In order to define the domains of XRCC1 necessary for the interactions, similar experiments were carried out using a series of XRCC1 protein fragments fused to MBP (Fig. 2A and C). Both hNTH1 and hNEIL2 DNA glycosylases, were retained on the matrix when the BRCT1, either alone (BRCT1) or in combination with other domains (N-B1 or H-B1) were fused to MBP, indicating that the BRCT1 domain is sufficient for the interaction with these DNA glycosylases. However, hNTH1 and hNEIL2 were also recovered by the MBP fused to NTD plus the hinge domains (N–H). No retention was observed for the MBP fused to either the NTD or the BRCT2 domains alone. These results suggest therefore two potentially independent interacting sites in XRCC1 for these glycosylases. Interestingly, the same pattern of interacting XRCC1 regions was found for APE1 and hOGG1 [15]. Pull-down experiments from HeLa cell extracts showed that the hypoxanthine DNA glycosylase activity was recovered using the NB1 region of XRCC1 spanning the N-terminal, first linker and BRCT1 domains (Fig. 2B, right panel). Therefore, hOGG1, hNTH1, hNEIL2 and possibly MPG are able to interact with the same region of XRCC1 (Fig. 2C), suggesting a common mechanism for the initiation steps of base repair.

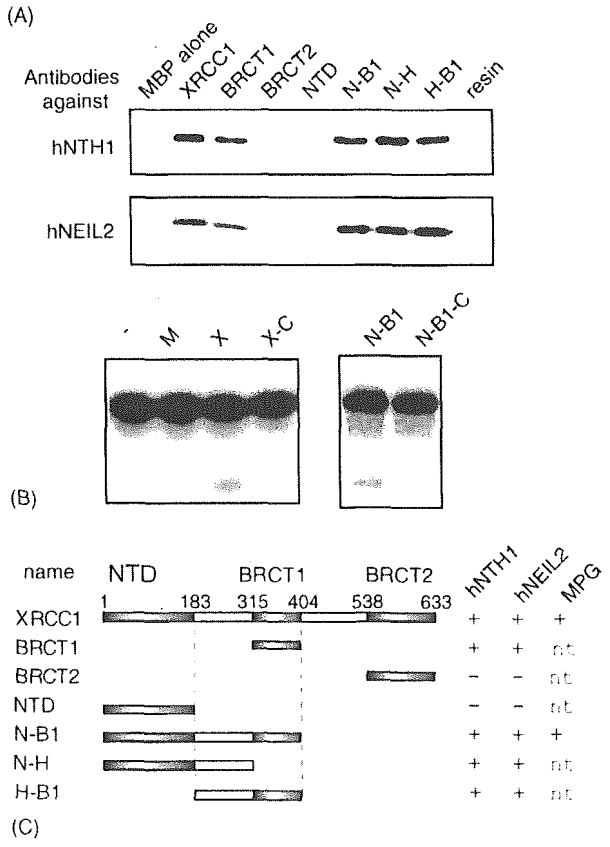


Fig. 2. XRCC1 interacting domains with several DNA glycosylases. MBP either alone or fused to the indicated XRCC1 fragments was expressed in *E. coli* and fixed to an amylose resin. (A) hNTH1 or hNEIL2 His-tagged fusions purified from *E. coli* were put in contact with the protein bound to the resin. After washes (see Section 2), the proteins associated to the resin were separated by SDS-PAGE, transferred to a membrane and immunoblotted with antibodies against His_6 tag; (B) purified MPG (left panel) or HeLa cell extract (right panel) were put in contact with resin associated with MBP fusion proteins, and after washes, hypoxanthine DNA glycosylase activity assays were performed (see Section 2); (C) regions of XRCC1 expressed fused to MBP. Interacting domains (+) or not (−) are indicated; nt: not tested.

3.3. Stimulation of the DNA glycosylase activities by XRCC1

XRCC1 is capable of modulating the enzymatic activities of several of its protein partners involved in the recognition of the lesion [7,12,14,15]. To investigate whether this could be extended to its interactions with the glycosylases shown above, we performed enzymatic assays with purified hNTH1 and MPG using as substrates oligonucleotides harbouring a single DHT or hypoxanthine residue, respectively. For MPG, the incubation in the presence of increasing amounts of XRCC1 stimulated the removal of hypoxanthine, yielding an AP site that is revealed by its cleavage with 0.1N NaOH. An excess of XRCC1 gave up to a five-fold stimulation of the MPG DNA glycosylase activity (Fig. 3A). XRCC1 by itself had no effect on the substrate showing that the stimulation observed is not due to a DNA glycosylase contaminating the XRCC1 preparation. In the case of hNTH1, the activity was monitored by the capacity of the protein to form a covalent complex with an oligonucleotide carrying a DHT residue in the presence of a reducing agent. Purified hNTH1, in the presence or absence of the indicated amounts of XRCC1, was incubated with the radiolabelled duplex DNA harbouring the lesion and $NaBH_4$. The formation of the Schiff base intermediate was followed by denaturing PAGE. Fig. 3B shows the stimulation of the formation of the covalent complex by increasing amounts of XRCC1. Therefore, as previously shown for APE1 and hOGG1, XRCC1 is able to physically and functionally interact with enzymes initiating the BER process.

3.4. DNA glycosylases sub-cellular localisation

The different DNA glycosylases analysed were tagged with a *myc* epitope and expressed in HeLa cells to follow their

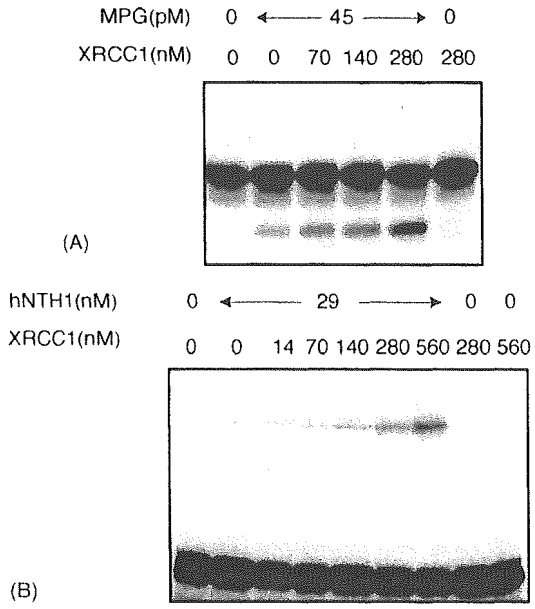


Fig. 3. XRCC1 stimulates MPG and hNTH1 glycosylases activities. (A) Effect of XRCC1 on MPG hypoxanthine DNA glycosylase activity. 5′-End-labelled 34mer duplex containing an inosine residue was incubated with (+) or without (−) MPG and XRCC1 at the indicated concentrations. The products of the reactions were separated by denaturing 20% PAGE; (B) effect of XRCC1 on borohydride trapping of hNTH1. 5′-End-labelled 34mer duplex containing a DHT residue was incubated with (+) or without (−) hNTH1 and XRCC1 at the indicated concentrations. After incubation at 37 °C, 50 mM of sodium borohydride was added. The reactions were pursued for 5 min at 37 °C, stopped in 1× SDS sample buffer and heated for 5 min at 100 °C. The trapped complexes were separated from free substrate by SDS-PAGE 10%.

sub-cellular localisation by immunohistochemistry. Transfections with the empty vector gave no detectable signal (data not shown). Our previous experience with *myc*-tagged hOGG1 nuclear and mitochondrial forms had shown that the myc tag does not interfere with the subcellular localisation of the DNA glycosylase [22]. Consistently with the described nuclear localisation sequences (NLS), α-hOGG1-*myc* [23], hNTH1-*myc* [24] and hNEIL1-*myc* [25] were mainly found in the nucleus (Fig. 4). That was also the case for MPG-*myc* which carries a potential NLS in its N-terminal region (PSORT programme, Version 6.4, http://psort.nibb.ac.jp) and for NEIL2-*myc* for which no putative NLS is detected.

For hNTH1-*myc*, a fraction of the protein is localised to the mitochondria in about 10% of the cells (Fig. 4). This was confirmed by using an antibody against hNTH1 (data not shown). For hNEIL1-*myc*, the signal was nuclear with some cells displaying a cytoplasmic fusion protein. In the case of hNEIL2-*myc*, the fusion protein was exclusively localised to the nuclei. MPG-*myc* also showed some cytoplasmic signal in addition to an abundant nuclear labelling. However, as shown by confocal microscopy, MPG-*myc* is not randomly distributed in the nucleoplasm. Indeed, it is mainly located in the periphery of the nucleus and around nucleoli (Fig. 4).

3.5. Co-localisation of DNA glycosylases and XRCC1 in nuclear foci

XRCC1 is known to re-localise and form distinct foci upon H_2O_2 treatment of the cells [5]. These foci are associated with the repair of single-strand breaks. We wished to investigate if some XRCC1 foci are associated with DNA glycosylases and therefore with the removal of damaged bases. By using the "measure co-localization" tool of the MetaMorph software on confocal images (AND images on Figs. 5 and 6) the degree of co-localisation between YFP-XRCC1 and the various glycosylases can be analysed. One hour after H_2O_2 treatment about 10% of the XRCC1 foci were found to contain also the DNA glycosylase studied in cells expressing XRCC1-YFP and either hOGG1-*myc*, hNTH1-*myc*, hNEIL1-*myc* or hNEIL2-*myc* fusions (Fig. 5, compare the AND column between panels A and B). The limited degree of co-localisation of each glycosylase with XRCC1 is likely to be due to the variety of DNA damages induced by the treatment and the transient nature of the interactions.

To perform similar tests with MPG, cells expressing XRCC1-YFP and MPG-*myc* were treated with MMS to specifically generate alkylated bases in DNA and the localisation of the fusion proteins was analysed with a confocal microscope one hour after treatment. In this case, although the distribution of MPG does not change significantly after the treatment, XRCC1 is redistributed and forms foci that co-localise with MPG in the periphery of the nucleus and around nucleoli (Fig. 6, compare the AND images between control and MMS). Around 50% of the XRCC1 foci induced by the treatment do contain MPG. These observations, together with the in vitro interactions described above, suggest that the DNA glycosylase recruits XRCC1 to the repair process.

4. Discussion

Repair of DNA lesions even as simple as modified bases involves multiple biochemical steps in most cases carried out by several different proteins. As for other DNA transactions, multiple dynamic protein interactions coordinate the ordered progression of DNA repair events to prevent incomplete repair or exposure of the substrate to new lesions. Several lines of evidence favour the dynamic assembly and disassembly of the DNA processing machinery at the time it is needed instead of the existence of preformed complexes [26]. In the case of BER, as in that of SSBR, XRCC1 is proposed to act as scaffold for the various enzymes required to accomplish repair [27]. This is supported by the demonstrated interactions of XRCC1 with all BER enzymes involved in the repair of AP sites, including APE1, PARP, POLβ and LIG3 (for review [3]). More recently, the finding of XRCC1 interactions with two DNA glycosylases, hOGG1 and NEIL1 [15,16] extended these observations to the repair of some modified bases. Based on these interactions and the XRCC1