Physical and biochemical properties of mammalian DNase X proteins: non-AUG translation initiation of porcine and bovine mRNAs for DNase X

Daisuke SHIOKAWA*, Yukari SHIKA*, Kazuki SAITO*, Kosuke YAMAZAKI* and Sei-ichi TANUMA*†[1]

*Department of Biochemistry, Faculty of Pharmaceutical Sciences, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan, and †Genome and Drug Research Center, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan

DNase X is the first human DNase protein identified as being homologous with DNase I. In the present study we describe the isolation of several mammalian DNase X cDNAs and the molecular characterization of their coding proteins. A sequence comparison reveals some conserved characteristics: all the mammalian DNase X proteins have an N-terminal signal peptide, a potential *N*-linked glycosylation site and a C-terminal hydrophobic domain. Human DNase X, ectopically expressed in HeLa S3 cells, is located in the ER (endoplasmic reticulum) and is modified by an *N*-linked glycosylation at Asn-243. Gene expression analyses show that the high expression level in muscular tissues, a known feature of human *DNASE X*, is also observed in mouse *DNase X*. Interestingly, the translation of porcine and bovine DNase X proteins occurs in the absence of an in-frame AUG initiation codon. We show that their mRNAs utilize a conserved CUG triplet for translation initiation.

Key words: cDNA cloning, DNase X, glycosylation, muscle, translation initiation.

INTRODUCTION

DNase I is a well-characterized divalent cation-dependent endonuclease originally isolated from bovine pancreas and thereafter, DNase activities identical or similar to those of DNase I have been detected in a wide variety of mammalian tissues, cells and body fluids [1–3]. Until recently, most such DNases have been regarded as being DNase I. However, recent experimental progress has revealed the existence of several genes for DNase I-like DNases in mammalian genomes. At present, the DNase I family has expanded to include four similar but distinct enzymes: DNase I, DNase X/Xib, DNase γ/DNAS1L3, and DNAS1L2 [4–8].

In a previous study [9], it was revealed that the physical and enzymatic features of the four DNases are similar: (i) all DNase I family proteins contain hydrophobic precursor peptides in their N-termini; (ii) two essential histidine residues of DNase I are conserved in the other members; (iii) all are activated by Ca^{2+} and Mg^{2+} in a synergistic manner and are strongly inhibited by Zn^{2+}; and (iv) members hydrolyse DNA endonucleolytically to produce 3′ OH/5′ P ends. By contrast, the expression profiles of their coding genes are found to be quite different in human tissues [9] therefore DNases of the DNase I family have been considered to play unique physiological roles *in vivo*.

DNASE X, the first gene shown to encode a DNase I-like protein, was identified at the q28 region of the human X chromosome [4–6]. Interestingly, *DNASE X* is expressed at high levels in cardiac and skeletal muscles, in which the rest of the family members are scarcely expressed [4–6,9,10]. Furthermore, DNase X has an extra hydrophobic stretch in its C-terminus [4–6,9], and this has been regarded as the most outstanding structural feature of human DNase X [9].

These observations suggest that DNase X plays a role in some muscular functions at its C-terminal HD (hydrophobic domain). However, our knowledge of DNase X is still too sparse to predict its physiological function and it is not known whether the muscle-specific expression and the existence of the C-terminal HD are common features of DNase X in mammals. In the present study, we determined the primary structures of the DNase X proteins of monkey, porcine, bovine, mouse, rat and hamster origins by cloning their full-length cDNAs and reveal some conserved physical and biochemical characteristics of mammalian DNase X proteins. Furthermore, we have shown that the translation of porcine and bovine DNase X proteins occurs in the absence of in-frame AUG initiation codons.

EXPERIMENTAL

cDNA cloning

The EST (expressed sequence tag) subdivision of the National Center for Biotechnology Information GenBank® database was searched with the deduced amino acid sequence of human DNase X (GenBank® accession number X90392) using the Tblastn program. As a result, we identified several draft sequences potentially coding for porcine, mouse and rat DNase X proteins (results not shown). A sequence alignment revealed several conserved regions at the nt sequence level. On the basis of this result, we designed two types of primer set, types A and B, which were used in the following RACE (rapid amplification of cDNA ends) reactions. First-PCR primers for 5′ RACE were 5′-TCAGCTCCACCTCCACGGGGTAGTGGTC-3′ type A and 5′-TCAGTTCCACTTCCACAGGATAATGGTC-3′ type B. Nested-PCR primers for 5′ RACE were 5′-GCCAGGATCCGAACTAAGGTGTCCATCAC-3′ type A and 5′-GCTAGGATCTGAACTAAGGTATCCATCAC-3′ type B. First-PCR primers for 3′ RACE were 5′-GCCTTTCGCATCTGCGCCTTCAATGCCC-3′ type A and 5′-GCCTTTCGTATCTGTGCCTTCAATGCCC-3′ type B. Nested-PCR primers for 3′ RACE were 5′-CTTCGGACTCAGGCTGGCTTCCACTGGG-3′ type A and 5′-CTCCGGACTAAGGCAGGCTTCCACTGGG-3′ type B. The type A set was used to determine the full-length cDNA sequences of monkey,

Abbreviations used: DPBS, Dulbecco's modified phosphate-buffered saline; ER, endoplasmic reticulum; FGF-2, fibroblast growth factor-2; HD, hydrophobic domain; ORF, open reading frame; PDI, protein disulphide isomerase; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcription-PCR; SR, sarcoplasmic reticulum.

[1] To whom correspondence should be addressed (email tanuma@rs.noda.tus.ac.jp).

© 2005 Biochemical Society

porcine and bovine DNase X proteins, and the type B set to determine those for the mouse, rat and hamster DNase X proteins. Firstly, we amplified internal cDNA fragments by PCR using the first-PCR primers for 5′ and 3′ RACE, and then the 5′ and 3′ cDNA ends were isolated by the RACE reactions using a SMART™ RACE cDNA Amplification kit (Clontech) according to the manufacturer's protocol. Total RNAs used in the SMART™ RACE were prepared from COS-7 cells (monkey), liver (porcine, bovine, mouse and rat) and CHO-K1 (Chinese hamster ovary-K1) cells as described in the gene expression analyses section.

The PCR products were subcloned into pBluescript KS^+ (Stratagene) and the nt sequences were determined on both strands by cycle sequencing using a 7-deaza Thermo Sequenase kit (Amersham Biosciences) and a DSQ2000 DNA sequencer (Shimadzu, Kyoto, Japan).

Construction of expression vectors

phDNase-X-Myc-His_6, an expression vector for C-terminal Myc- and His_6-tagged form of human DNase X, was constructed previously [9]. phDNase-X-N243A-Myc-His_6, an expression plasmid for DNase X with a point mutation converting Asn-243 into Ala-243, was generated from phDNase-X-Myc-His_6 using a long and accurate-PCR *in vitro* Mutagenesis kit (Takara) according to the manufacturer's protocol. Expression vectors for bovine and porcine DNase X proteins were generated by cloning PCR fragments covering nts 170–1135 and 151–1134 of bovine and porcine cDNAs (see Figure 4A) respectively into the EcoRV site of pcDNA3.1-myc-his B (Invitrogen) in-frame with the following Myc and His_6 tags. The cDNA fragments, listed in Figure 5(D), were amplified by PCR and the vectors used to define the translation initiation sites were constructed as described above. Mutant vectors were generated using PCR primers containing the indicated base mutations.

Cell culture, transfection and Western blot analysis

HeLa S3 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10 % (v/v) foetal calf serum, 100 units/ml penicillin and 0.1 mg/ml streptomycin. The cells, cultured in 60 mm dishes at approx. 80 % confluence, were transfected with 1 μg of expression vector using FuGENE 6 transfection reagent (Roche). At 24 h post-transfection, the cells were harvested and the Myc-His_6-tagged recombinant proteins were detected by Western blotting using an anti-Myc monoclonal antibody (Invitrogen) as described previously [9].

Indirect immunofluorescence

The Myc and His_6-tagged forms of human, porcine and bovine DNase X proteins, expressed in HeLa S3 cells, were detected *in situ* using an anti-Myc-tag monoclonal antibody (1 μg/ml, Invitrogen) as described previously, with some modifications [11]. Briefly, cells grown on sterile coverslips were fixed using cold methanol and after blocking with DPBS (Dulbecco's modified phosphate-buffered saline) containing 5 % (v/v) foetal calf serum, the resulting cells were incubated with the primary antibody for 60 min at room temperature in a sealed humid chamber. The cells were then washed with DPBS and incubated with highly cross-adsorbed Alexa Fluor® 488 goat anti-(mouse IgG) (Molecular Probes) for 30 min at room temperature (1:2000 dilution). After washing with DPBS, the cells were observed using a fluorescence microscope (BX60, Olympus). Double staining of DNase X and ER (endoplasmic reticulum) marker proteins were performed using the following combinations of primary antibodies: anti-Myc-tag mouse monoclonal (Invitrogen) 1 μg/ml and anti-calreticulin rabbit polyclonal (Sigma) 1:1000; anti-Myc-tag rabbit polyclonal (Cell Signaling Technology) 1:200 and anti-PDI (protein disulphide isomerase) (Medical & Biological Laboratories) mouse monoclonal 5 μg/ml. A highly cross-adsorbed Alexa Fluor® 594 goat anti-rabbit IgG (Molecular Probes) at 1:2000 was used to detect the rabbit antibodies. The resulting cells were counter-stained with 1 μM Hoechst 33342 (Molecular Probes) and observed under a confocal laser scanning microscope (TCS SP2, Leica). Digital images were processed using Photoshop 7.0 software (Adobe).

Table 1 Summary of the data for cDNAs and the deduced proteins

	cDNA length*	5′ noncoding	Amino acid number†	Molecular mass†	pI†‡
Human	2641 bp	794 bp	302 (284)	33 892 (31 966)	5.38 (5.27)
Monkey	1603 bp	230 bp	302 (284)	33 945 (31 991)	5.26 (5.16)
Porcine	1224 bp	186 bp	315 (286)	34 986 (32 079)	5.69 (5.62)
Bovine	2258 bp	184 bp	316 (288)	35 168 (32 313)	5.47 (5.13)
Mouse	1519 bp	222 bp	314 (277)	35 604 (31 646)	6.80 (6.74)
Rat	1512 bp	226 bp	312 (275)	35 181 (31 318)	7.63 (7.15)
Hamster	1957 bp	264 bp	304 (280)	34 358 (31 855)	5.49 (5.52)

* nt numbers do not cover the regions for poly(A) tail.
† Values calculated for mature forms are shown in parentheses.
‡ pI, isoelectric point.

Gene expression analyses

Total RNAs were prepared from freshly isolated tissues using an RNeasy kit (Qiagen) in combination with an RNase-free DNase set (Qiagen) according to the manufacturer's instructions. The cDNAs were generated from the DNA-free RNA samples and RT-PCR (reverse transcription-PCR) was performed as described previously [9]. The primers used were as follows: *DNase X*, 5′-GATATCATGGTGCTTCAGGAG-3′ sense and 5′-TGCATTGAAGTCTCCAAGCAG-3′ antisense; and β-actin, 5′-CACCTTCTACAATGAGCTGCG-3′ sense and 5′-CTTGCT-GATCCACATCTGCTG-3′ antisense. For the quantification of the *DNase X* transcripts, real-time PCR was performed using a QuantiTect SYBR Green PCR kit (Qiagen) and an ABI Prism 7700 sequence Detector (Applied Bio Systems). All assays were performed in triplicate and normalized for β-actin mRNA.

Cell-free transcription/translation system

DNase X proteins were generated *in vitro* using a TNT Quick Coupled Transcription/Translation System (Promega) according to the manufacturer's instrucutions. Briefly, 1 μg of each expression vector was added to an aliquot of the TNT Quick Master Mix and incubated in a 50 μl of reaction volume for 60 min at 30 °C. Aliquots of equal volume of the resulting mixtures were separated by SDS/PAGE (12 % gel) and the DNase X proteins were detected by Western blotting as described above.

RESULTS AND DISCUSSION

Determination of primary structures of mammalian DNase X proteins

The C-terminal HD has been considered to be the most outstanding feature of human DNase X, however, it is not clear whether this is widely conserved among mammals or is observed only in humans. To better understand the nature of DNase X, we determined the primary structures of DNase X proteins of several mammals. The full-length cDNAs were isolated by the RACE method and their sequence information is summarized in Table 1. The detailed experimental procedure was described in the Experimental section.

© 2005 Biochemical Society

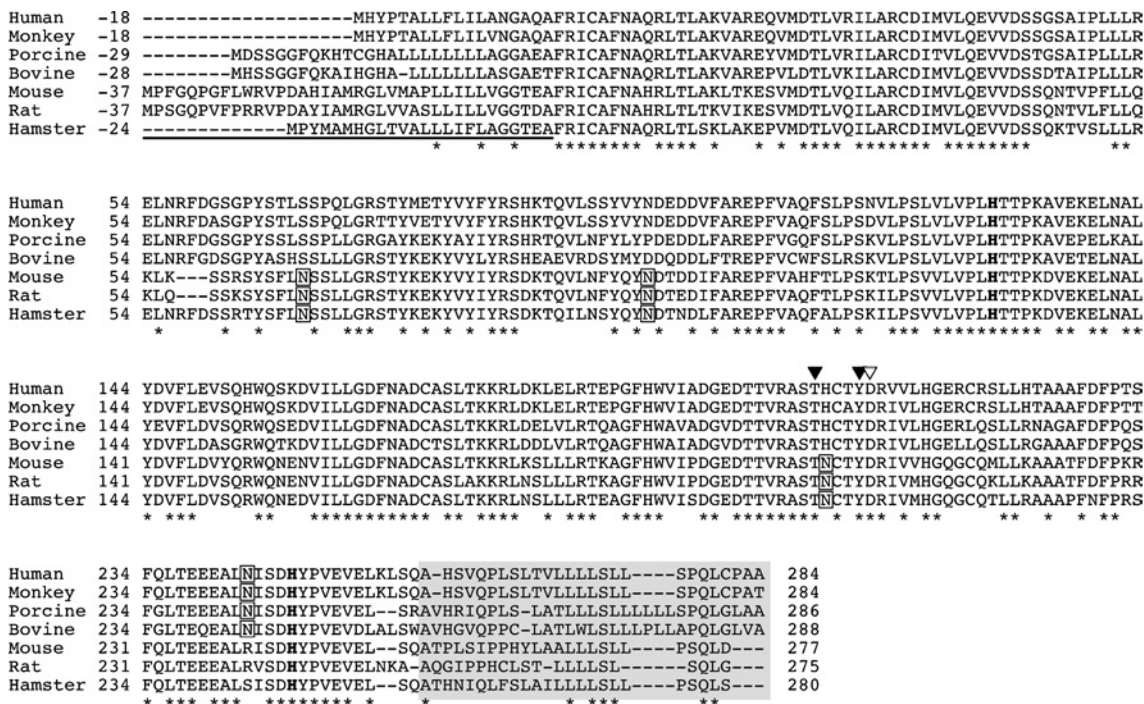

Figure 1 Comparison of the deduced amino acid sequences of mammalian DNase X proteins

Sequence alignment of the indicated DNase X proteins was performed using the GENETYX-MAC version 11 program. Identical residues are marked by asterisks. Dashes represent gaps introduced for better alignment. Putative precursor peptides, predicted by the PSORT II program, are underlined. Two conserved His residues (possible active sites) are shown by bold letters. Asp-209, which is considered to be hydrogen-bonded to His-247, is marked by an open triangle. Thr-204 and Thr-208, which may interact with the phosphate immediately 3′ to the cutting phosphate, are marked by closed triangles. Potential Asn-linked glycosylation sites are outlined. C-terminal hydrophobic domains are shaded. The amino acid numbering begins at the N-terminal postulated for the mature enzymes with negative numbers for the precursor peptides. The GenBank® accession numbers for DNase X cDNAs are X90392 (human), DQ116781 (monkey), DQ116782 (porcine), DQ116783 (bovine), DQ116784 (mouse), DQ116785 (rat) and DQ116786 (hamster).

The human DNASE X mRNA contains a long 5′ leader sequence (794 bp, GenBank® accession number X90392) coded by noncoding exons [6]. However, the 5′ untranslated regions of the DNase X mRNAs of nonhumans are not very long being in the range of 184–264 bp (Table 1). These observations may be simply accounted for by species differences. Alternatively, this may be due to the difference between the tissue sources used for mRNA preparation, because the basal level expression of the DNase X mRNA is observed widely in murine tissues (see Figure 3) and the mRNAs used in RACE reactions were isolated from non-muscular tissues or cultured cell lines (as described in the Experimental section). It is possible that the expression of the genes coding for DNase X are controlled by multiple promoters, tissue-specific and nonspecific, with distinct transcriptional origins.

The ORFs (open reading frames) of the clones were found to encode 302–316 amino acid polypeptides, whose calculated molecular masses range from 33 945–35 604 (Table 1). Amino acid sequences deduced from the cDNAs are shown in Figure 1. Amino acid identities between any two of the indicated species are found to be in the range of 63.3 % (bovine versus rat) to 97.5 % (human versus monkey) and the scores are summarized in Table 2. As described below, we found that translation initiation of porcine and bovine mRNAs occurs at a conserved CUG codon that usually encodes the amino acid Leu (Figure 5). In the present study, Met instead of Leu has tentatively been assigned as the first amino acid of porcine and bovine DNase X proteins, because such non-AUG translation initiation is usually caused by the pairing of the anticodon of the methionyl initiator tRNA with certain codons that differ by one base from the classical AUG [12–14]. However, recent studies showed that translation initiation at CUG with Leu is also possible [15,16]. Therefore, more studies are required to identify the N-terminal amino acid.

Table 2 Amino acid identities of mammalian DNase X proteins

Percentage identities are calculated for mature forms.

	Percentage identity					
	Human	Monkey	Porcine	Bovine	Mouse	Rat
Monkey	97.5					
Porcine	81.0	79.9				
Bovine	75.8	74.7	80.9			
Mouse	70.9	70.5	69.3	66.1		
Rat	69.5	69.1	66.8	63.3	87.5	
Hamster	75.8	75.1	72.8	68.9	83.6	82.3

A sequence comparison clarified some important structural features of the DNase X proteins (Figure 1). All the DNase X proteins contain N-terminal signal sequences (underlined) and potential *N*-linked glycosylation sites (outlined). Importantly, HD has been shown to be located at the C-terminus of all the mammalian DNase X proteins (shaded area), suggesting its essential importance in the specific function of DNase X.

Localization of human DNase X in ER and occurrence of *N*-linked glycosylation at Asn-243

The occurrences of N-terminal signal peptides in DNase X proteins suggest that their nascent polypeptides are translocated

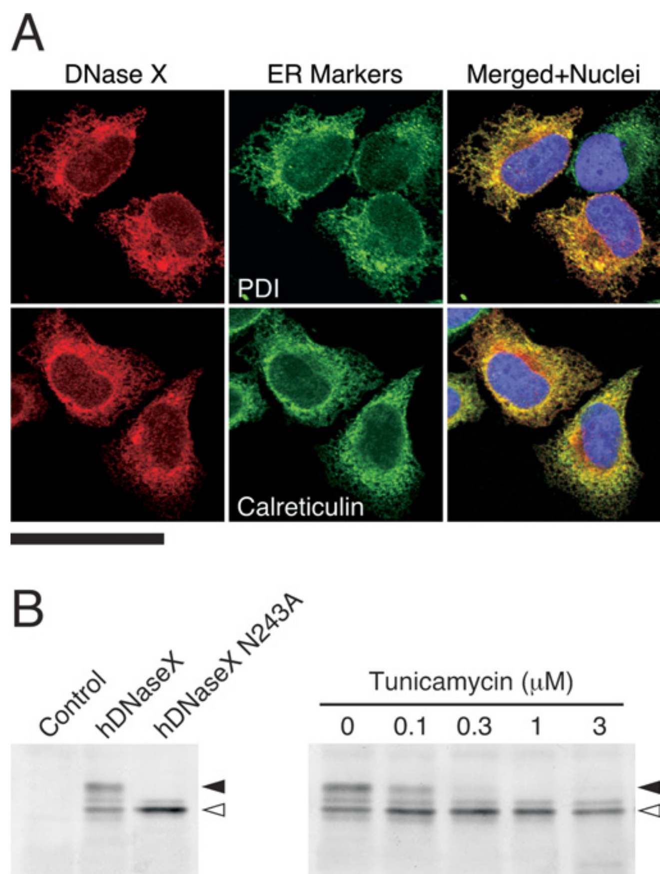

Figure 2 *N*-linked glycosylation at Asn-243 of human DNase X protein

(**A**) Colocalization of DNase X with ER marker proteins. HeLa S3 cells, transfected with phDNase-X-Myc-His$_6$, were fixed, permeabilized and double-stained with antibodies for Myc-tag (red) and indicated ER marker proteins (green) as described in the Experimental section. The fluorescence images were observed by confocal laser scanning microscopy. The yellow areas in the merged images show colocalization of the two proteins. Scale bar represents 50 μm. (**B**) Effects of N243A point mutation (left panel) and tunicamycin (right panel) on the attachment of an *N*-linked sugar chain. HeLa S3 cells were transfected with expression vectors for the indicated DNases (left panel). HeLa S3 cells, transfected with an expression vector for wild-type DNase X, were cultured for 24 h in the presence of increasing doses of tunicamycin (right panel). The resulting cells were analysed by Western blotting as described in the Experimental section. Closed and open arrow heads represent glycosylated and nonglycosylated forms of DNase X proteins respectively.

into the ER and directed to the secretory pathway. Thus we determined the subcellular localization of human DNase X exogenously expressed in HeLa S3 cells. The results obtained by confocal microscopy clearly showed the ER-specific localization of DNase X. The immunofluorescence images of DNase X are well matched with those of two independent ER markers, PDI and calreticulin (Figure 2A). Muscle cells are known to contain SR (sarcoplasmic reticulum), a muscle specific ER-related organelle, in which Ca^{2+}, an important cofactor for the DNase X activity, is stored [17,18]. Therefore, it might be possible that DNase X is targeted to the SR and plays a role in some SR-related events in muscle cells.

Human DNase X contains a potential *N*-linked glycosylation site at Asn-243, and the consensus sequence for this modification Asn-Xaa-Ser/Thr, is shown to be conserved in the corresponding regions of monkey, porcine and bovine DNase X proteins (Figure 1). Furthermore, the ER is the cellular compartment in which the addition of oligosaccharides takes place. Therefore we next investigated whether modification by glycosylation occurs at the Asn-243 of human DNase X. DNase X expressed in HeLa S3 cells was detected as two bands, whereas that carrying an N243A mutation appeared as a single band at the same position as the lower band of the wild-type protein (Figure 2B, left panel). The apparent molecular mass of the lower band (35 kDa) is in agreement with that calculated for its mature form (without the signal sequence) including the Myc and His$_6$ tags, suggesting that the upper band represents DNase X with an *N*-linked sugar chain at Asn-243. This was confirmed by the observation that the appearance of the upper band is strongly suppressed by the addition of tunicamycin, an inhibitor of *N*-linked glycosylation, in a dose-dependent manner (Figure 2B, right panel). Taken together, these results indicate the occurrence of *N*-linked glycosylation at Asn-243 of human DNase X. Based on the known three-dimensional structure and the catalytic mechanism of bovine pancreatic DNase I [19–21], it is suggested that the Asp-209 and His-247 pair of human DNase X proteins act as a general base in its catalysis. Furthermore, Thr-204 and Tyr-208 are considered to make contacts with the phosphate immediately 3′ to the scissile phosphate of substrate DNA. This implies that both Asp-243 and Asp-205 of human and rodent DNase X respectively, are in proximity to the active site (Figure 1). Namely, although the possible *N*-glycosylation sites of the mammalian DNase X proteins are different, they may be conserved at the ternary structure level.

We further analysed the localization of DNase X in other cellular compartments including the Golgi, lysosomes and mitochondria. However, we could not observe the colocalization of DNase X with any of the protein markers for such organelles (results not shown). Furthermore, we obtained no results to suggest that DNase X is anchored to the plasma membrane by the C-terminal HD. There are at least two possible explanations for the results: (i) DNase X is located exclusively in the ER and (ii) the C-terminal Myc-tag, the epitope used for immunological detection, is inaccessible due to modifications or removal by proteolysis. To evaluate these possibilities, we are now preparing monoclonal antibodies specific for DNase X.

Expression profile of *DNase X* in mice

In previous studies, the expression level of *DNASE X* was shown to be high in heart and skeletal muscles in human tissues [4–6,9,10]. To investigate whether the muscle-specific expression is an important feature of the genes for DNase X, we determined the expression profile of *DNase X* in several murine tissues. RT-PCR showed that *DNase X* is expressed at high levels in heart and skeletal muscles, low levels in the brain and thymus, and at intermediate levels in the other tissues (Figure 3A). We further performed quantitative real-time PCR, and revealed that the expression levels of *DNase X* in heart and skeletal muscles are about 4-fold and 21-fold higher respectively than that in the liver (Figure 3B). These results strongly suggest that the high level of expression in muscular tissues is an important feature of the genes for DNase X. Note that a basal level of expression of *DNase X* is widely seen in the tissues tested and the same observation was reported for human *DNASE X* previously [4–6,9,10]. Therefore although DNase X is highly expressed in muscle, it seems to be involved in general cellular actions, rather than in those that are tightly restricted in muscle cells.

Translation of bovine and porcine mRNAs in the absence of in-frame AUG codon

Figure 4(A) shows the nt sequence in the vicinity of the 5′ and 3′ ends of the putative coding regions of bovine and porcine cDNAs. Their deduced amino acid sequences show acceptable homologies with those deduced for other species (Table 2), more importantly, they include two conserved active His residues (Figure 1). These results imply that these two proteins are fully

© 2005 Biochemical Society

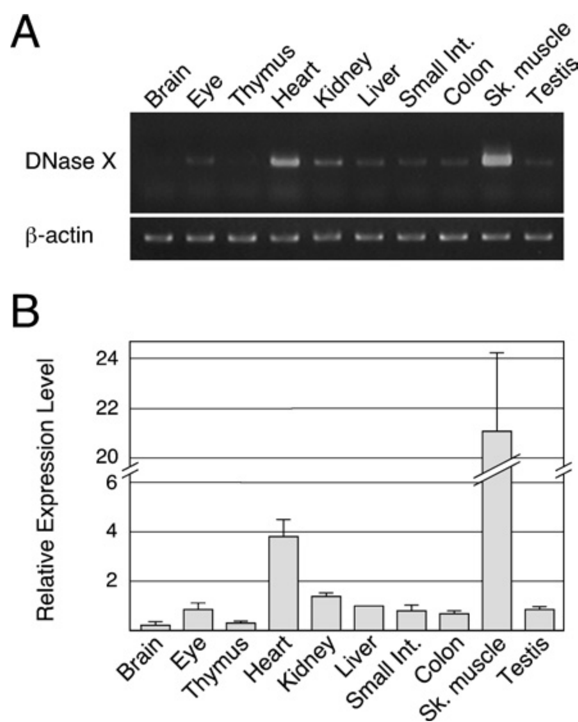

Figure 3 Tissue distribution of *DNase X* mRNA expression

(**A**) RT-PCR analysis of DNase X mRNA expression in adult mouse tissues. The identities of the RNAs are indicated at the top of each lane. β-Actin signals shown in the lower panel serve as an internal control. (**B**) Quantitative real-time PCR of DNase X transcript. The cDNA pools used in the RT-PCR were also analysed by real-time PCR as described in the ExperimentaL section. The levels of DNase X mRNA were normalized for β-actin and are indicated relative to the basal DNase X expression level in the liver. Values are shown as means $\pm$ S.D. ($n = 3$). Abbreviations used in this Figure: Int, intestine; Sk, skeletal.

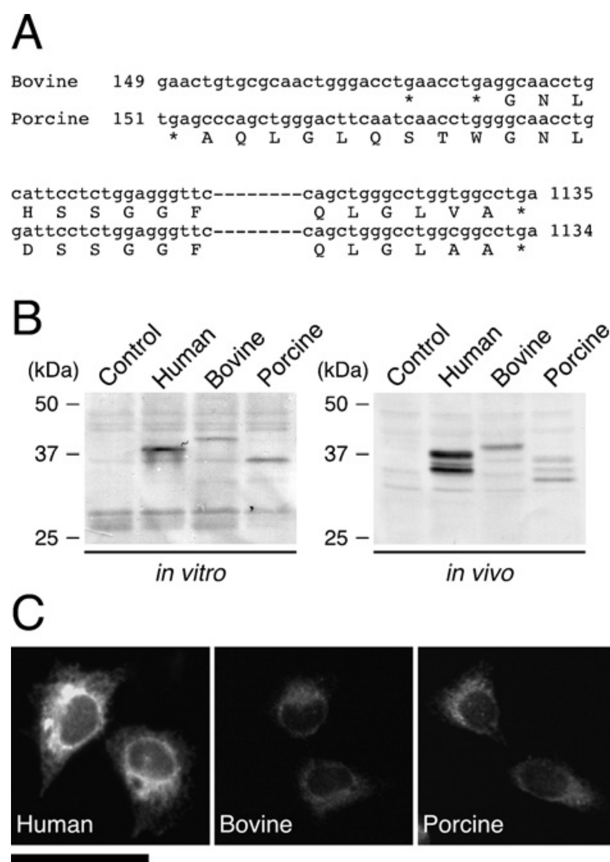

Figure 4 Translation of bovine and porcine mRNAs for DNase X *in vitro* and *in vivo*

(**A**) nt and predicted amino acid sequences of bovine and porcine cDNAs in the vicinity of the 5′ and 3′ in-frame stop codons. nt numbers are shown at both ends. Asterisks indicate the in-frame stop codons. (**B**) Western blotting of recombinant DNase X proteins translated *in vitro* (left panel) and *in vivo* (right panel). The DNase X proteins of the indicated species were generated using a cell-free transcription/translation system (*in vitro*) and in HeLa S3 cells (*in vivo*) using an expression vector for each DNase. The resulting proteins were detected by Western blotting using the anti-Myc-tag antibody, as described in the Experimental section. The control represents the results obtained when using an empty vector. (**C**) Immunofluorescence analysis of DNase X proteins. HeLa S3 cells were transfected with expression vectors for the indicated DNase X proteins. The cells were fixed, immunostained with the anti-Myc-tag antibody and observed by fluorescence microscopy as described in the Experimental section. Scale bar represents 50 μm.

functional, however, no in-frame AUG codons are found in their mRNAs in the vicinity of the expected translation initiation sites (Figure 4A). To determine the capacities of these mRNAs to encode DNase X, we generated expression vectors containing their putative coding sequences and subjected them to an *in vitro* transcription/translation system. To exclude the possibility of artificial translation in the *in vitro* system, the expression of their coding products was also examined in HeLa S3 cells *in vivo*. Figure 4(B) shows the expression of bovine and porcine DNase X proteins *in vitro* and *in vivo*. The expression levels of these proteins were relatively low as compared with that of human DNase X, however, the apparent molecular masses of the three DNase X proteins were found to be similar.

Indirect immunofluorescence showed the cytoplasmic localization of bovine and porcine DNase X proteins (Figure 4C). By comparing their images with that obtained for human DNase X, the subcellular localization of bovine and porcine proteins was suggested to be in the ER. These results clearly show the translation initiation of bovine and porcine DNase X proteins in the absence of in-frame AUG codons. Furthermore, taking their apparent molecular masses and ER localization into account the translation initiation sites for these DNase X proteins are likely to be located closely upstream of the regions coding for signal peptides.

Translation initiation from conserved CUG in porcine and bovine mRNAs coding for DNase X

The capacity of mammalian systems to initiate translation at non-AUG triplets was assessed in previous studies and the results showed that CUG and ACG are the most efficient codons for translation initiation [14,22,23]. We found that two CUG codons are located in porcine mRNA closely downstream of the 5′ in-frame stop codon (Figures 4A and 5D); we evaluated their potential for translation initiation. We generated several expression constructs, inserted with the cDNA fragments (1–9) listed in Figure 5(D), for this purpose. To detect a small difference in apparent molecular mass by Western blotting, almost three out of four of the ORFs of the cDNA fragments were deleted from the 3′ ends. As shown in Figure 5(A), proteins coded by ORFs, in which either the upstream (lane 2) or downstream (lane 3) CTG triplet was replaced by an ATG codon showed slow and identical electromobilities respectively, as compared with that produced by the wild-type construct (lane 1). The replacement of the downstream CTG with TTT, which has no potential for translation initiation, resulted in no protein synthesis (lanes 5 and 6), however, the inactivation of the upstream CTG by this mutation apparently had no effect (lane 4). These results suggest that translation of porcine DNase X is initiated from the

© 2005 Biochemical Society

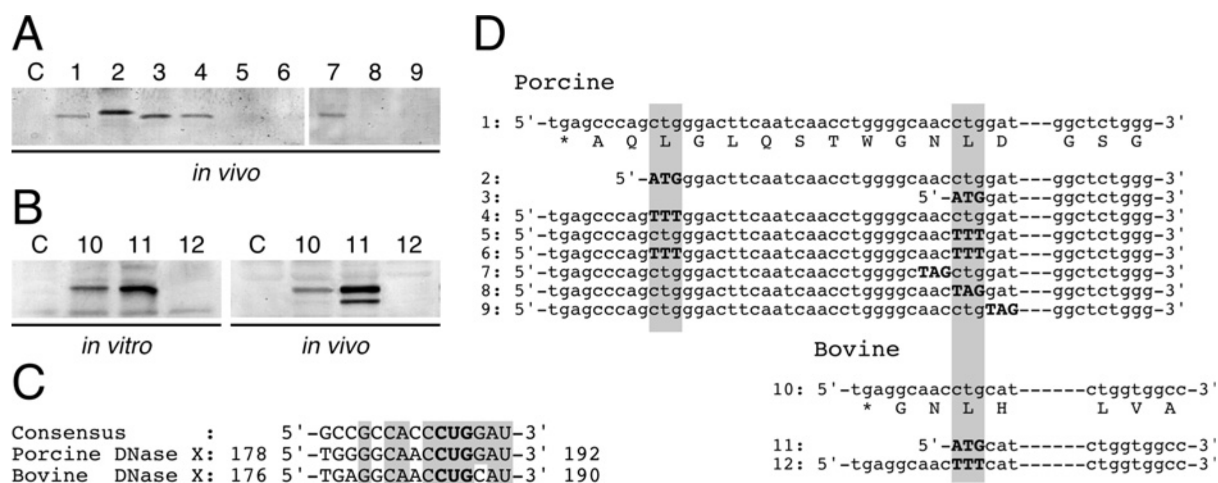

Figure 5 Effects of base mutations on the translation of porcine and bovine mRNAs for DNase X

(**A**) Identification of initiation CUG codon in porcine DNase X mRNA. HeLa S3 cells were transfected with expression vectors carrying partial cDNA fragments for porcine DNase X with the base mutations listed in (**D**). The DNase X proteins expressed in HeLa S3 cells (*in vivo*) were detected by Western blotting using the anti-Myc-tag antibody, as described in the Experimental section. (**B**) Translation initiation of bovine mRNA from conserved CUG codon. Bovine DNase X proteins were generated in HeLa S3 cells (*in vivo*) or a cell-free translation system (*in vitro*) using expression vectors carrying indicated base mutations (**D**). The bovine DNase X proteins were detected by Western blotting, as described in the Experimental section. The identities of expression vectors are indicated at the top. (**C**) Represents the control experiments performed using an empty vector. (**C**) nt sequences surrounding initiator CUG codon in porcine and bovine mRNAs. The porcine and bovine DNase X sequences are compared with that shown to be suitable for CUG initiation (consensus). Conserved bases are shaded. The initiation CUG codons are bold. nt numbers are shown at both ends. (**D**) List of base mutations tested in this study. The indicated cDNA fragments, generated by PCR, were cloned into a pcDNA vector in-frame with the following Myc and His_6 tags. The C-terminal amino acids coded by the porcine and bovine cDNAs were Glu-63 and Ala-288 respectively. Mutated bases are shown in bold capital letters. In-frame CTG codons are shaded. The identification numbers of the vectors are given on the left.

downstream CUG, corresponding to CTG at nts 187–189 of the cDNA sequence (Figure 4A and 5D). This conclusion is supported by the observation that the replacement of the initiation CTG or the 3′ adjacent GAT with a stop codon, TAG completely abrogated protein translation, whereas the same mutation at the 5′ neighbouring AAC had no effect (Figure 5A).

We found that the bovine DNase X cDNA sequence contains an in-frame CTG at nts 185–187, the position corresponding to the initiation CTG for porcine DNase X (Figure 4A and 5D). Bovine DNase X proteins, coded for by vectors containing the wild-type (Figure 5D, 10) or an ATG-initiated (Figure 5D, 11) ORF, appeared to have the same molecular mass when generated using both *in vitro* and *in vivo* expression systems (Figure 5B). Furthermore, no protein synthesis was observed for a vector carrying a TTT mutation at the conserved CTG triplet (Figure 5B, lane 12). On the basis of these results, the translation initiation of porcine and bovine DNase X proteins was shown to occur at the conserved CUG codon.

Most of the eukaryotic genes identified thus far have been shown to utilize an initiation AUG triplet for their protein synthesis. In some rare cases, translation initiation also occurs at some non-AUG codons such as CUG, ACG and less frequently at GUG [14,22,23]. Although CUG is the most effective and frequently used codon for non-AUG translation initiation, by no means does every CUG function as an initiation codon in a mRNA. The efficiency of initiation codon recognition is greatly influenced by neighbouring codons and the most favourable sequence for non-AUG translation initiation was determined previously [24,25]. A sequence comparison revealed that the nt sequences surrounding the initiation CUG for bovine and porcine DNase X proteins are highly similar to the prototype sequence for non-AUG translation initiation (Figure 5C).

Translation initiation at a non-AUG initiator is strongly stimulated when the immediate downstream codon is GAU, in which A at +5 bp is strongly beneficial and at +6 bp only U has a positive effect (the first nt of the initiation codon is defined as +1). At +4 bp, the porcine DNase X sequence matches the rule, but the bovine sequence contains C instead of G. Although only G exerts a positive influence at +4 bp, it has been suggested that the combination of the following A and U at +5 bp and +6 bp respectively, can compensate for the absence of G at +4 bp [26]. In fact, the occurrence of a CAU codon immediately downstream of the CUG initiator was reported for human T-cell leukaemia virus type-1 Tax protein [26].

The functional relevance of the non-AUG-initiated protein synthesis has yet to be fully understood. However, some mammalian genes such as *hck*, *int-2*, *FGF-2* (fibroblast growth factor-2), and *Bag-1* have non-AUG initiators in addition to their AUG initiation codons and produce alternative products with altered properties for example, a different subcellular localization [27–30]. By contrast, eukaryotic genes, in which translation is initiated only at non-AUG codons, are quite rare. No DNase X protein synthesis was observed when the conserved CUG was inactivated (Figure 5A and 5B) suggesting that there is no alternative initiator codon in porcine and bovine DNase X mRNAs. To the best of our knowledge, this is the first study to identify mammalian genes utilizing a CUG initiator exclusively for protein synthesis. Non-AUG translation initiators are found most frequently in mRNAs for certain regulatory proteins, such as oncoproteins and growth factors and their translation levels are often altered in response to particular cellular conditions or affected by upstream factors. Translation initiation from the CUG codon of c-Myc mRNA is enhanced by lowering the methionine level in the cells [31] and the relative abundances of CUG- and AUG-initiated *FGF-2* products are controlled by eukaryotic initiation factor 4F [32]. Therefore the expression levels of bovine and porcine DNase X proteins may be controlled at the translation level and this may provide some specific advantage for these animals.

In summary, we have isolated and characterized several mammalian cDNAs for DNase X and demonstrated that the high expression level in muscular tissues and the occurrence of a HD in the C-terminus are conserved features of DNase X. We have also

© 2005 Biochemical Society

revealed the ER localization and *N*-linked glycosylation at Asn-243 of human DNase X. Furthermore, we have provided evidence that porcine and bovine mRNAs for DNase X utilize a conserved CUG for translation initiation. At present, little is known about the regulatory mechanisms of non-AUG-initiated protein synthesis. Therefore our results may also be useful for understanding the physiological significance of non-AUG translation initiation.

This work was funded in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

1 Moore, S. (1981) Pancreatic DNase. In The enzymes (Boyer, P. D., ed.), pp. 281–296, Academic Press, New York
2 Price, P. A. (1975) The essential role of Ca^{2+} in the activity of bovine pancreatic deoxyribonuclease. J. Biol. Chem. **250**, 1981–1986
3 Campbell, V. W. and Jackson, D. A. (1980) The effect of divalent cations on the mode of action of DNase I. The initial reaction products produced from covalently closed circular DNA. J. Biol. Chem. **255**, 3726–3735
4 Pergolizzi, R., Appierto, V., Bosetti, A., DeBellis, G. L., Rovida, E. and Biunno, I. (1996) Cloning of a gene encoding a DNase I-like endonuclease in the human Xq28 region. Gene **168**, 267–270
5 Parrish, J. E., Ciccodicola, A., Wehhert, M., Cox, G. F., Chen, E. and Nelson, D. L. (1995) A muscle-specific DNase I-like gene in human Xq28. Hum. Mol. Genet. **4**, 1557–1564
6 Coy, J. F., Velhagen, I., Himmele, R., Delius, H., Poustka, A. and Zentgraf, H. (1996) Isolation, differential splicing and protein expression of a DNase on the human X chromosome. Cell Death Differ. **3**, 199–206
7 Shiokawa, D. and Tanuma, S. (1998) Molecular cloning and expression of a cDNA encoding an apoptotic endonuclease DNase γ. Biochem. J. **332**, 713–720
8 Rodriguez, A. M., Rodin, D., Nomura, H., Morton, C. C., Weremowicz, S. and Schneider, M. C. (1997) Identification, localization, and expression of two novel human genes similar to deoxyribonuclease I. Genomics **42**, 507–513
9 Shiokawa, D. and Tanuma, S. (2001) Characterization of human DNase I family endonucleases and activation of DNase γ during apoptosis. Biochemistry **40**, 143–152
10 Los, M., Neubuser, D., Coy, J. F., Mozoluk, M., Poustka, A. and Schulze-Osthoff, K. (2000) Functional characterization of DNase X, a novel endonuclease expressed in muscle cells. Biochemistry **39**, 7365–7373
11 Shiokawa, D., Kobayashi, T. and Tanuma, S. (2002) Involvement of DNase γ in apoptosis associated with myogenic differentiation of C2C12 cells. J. Biol. Chem. **277**, 31031–31037
12 Kozak, M. (1992) A consideration of alternative models for the initiation of translation in eukaryotes. Crit. Rev. Biochem. Mol. Biol. **27**, 385–402
13 Kozak, M. (1989) The scanning model for translation: an update. J. Cell Biol. **108**, 229–241
14 Peabody, D. S. (1989) Translation initiation at non-AUG triplets in mammalian cells. J. Biol. Chem. **264**, 5031–5035
15 Schwab, S. R., Li, K. C., Kang, C. and Shastri, N. (2003) Constitutive display of cryptic translation products by MHC class I molecules. Science (Washington, DC) **301**, 1367–1371
16 Schwab, S. R., Shugart, J. A., Horng, T., Malarkannan, S. and Shastri, N. (2004) Unanticipated antigens: translation initiation at CUG with leucine. PLoS Biol. **2**, 1774–1784
17 Martonosi, A. N. and Pikula, S. (2003) The network of calcium regulation in muscle. Acta Biochim. Pol. **50**, 1–30
18 Toyoshima, C. and Inesi, G. (2004) Structural basis of ion pumping by Ca^{2+}-ATPase of the sarcoplasmic reticulum. Annu. Rev. Biochem. **73**, 269–292
19 Lahm, A. and Suck, D. (1991) DNase I-induced DNA conformation. 2 Å structure of a DNase I-octamer complex. J. Mol. Biol. **222**, 645–667
20 Jones, S. J., Worrall, A. F. and Connolly, B. A. (1996) Site-directed mutagenesis of the catalytic residues of bovine pancreatic deoxyribonuclease I. J. Mol. Biol. **264**, 1154–1163
21 Evans, S. J., Shipstone, E. J., Maughan, W. N. and Connolly, B. A. (1999) Site-directed mutagenesis of phosphate-contacting amino acids of bovine pancreatic deoxyribonuclease I. Biochemistry **38**, 3902–3909
22 Hann, S. R., King, M. W., Bentley, D. L., Anderson, C. W. and Eisenman, R. N. (1988) A non-AUG translational initiation in c-myc exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. Cell **52**, 185–195
23 Peabody, D. S. (1987) Translation initiation at an ACG triplet in mammalian cells. J. Biol. Chem. **262**, 11847–11851
24 Boeck, R. and Kolakofsky, D. (1994) Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis. EMBO J. **13**, 3608–3617
25 Grunert, S. and Jackson, R. J. (1994) The immediate downstream codon strongly influences the efficiency of utilization of eukaryotic translation initiation codons. EMBO J. **13**, 3618–3630
26 Corcelette, S., Masse, T. and Madjar, J. J. (2000) Initiation of translation by non-AUG codons in human T-cell lymphotropic virus type I mRNA encoding both Rex and Tax regulatory proteins. Nucleic Acids Res. **28**, 1625–1634
27 Lock, P., Ralph, S., Stanley, E., Boulet, I., Ramsay, R. and Dunn, A. R. (1991) Two isoforms of murine hck, generated by utilization of alternative translational initiation codons, exhibit different patterns of subcellular localization. Mol. Cell. Biol. **11**, 4363–4370
28 Acland, P., Dixon, M., Peters, G. and Dickson, C. (1990) Subcellular fate of the int-2 oncoprotein is determined by choice of initiation codon. Nature (London) **343**, 662–665
29 Bugler, B., Amalric, F. and Prats, H. (1991) Alternative initiation of translation determines cytoplasmic or nuclear localization of basic fibroblast growth factor. Mol. Cell. Biol. **11**, 573–577
30 Packham, G., Brimmell, M. and Cleveland, J. L. (1997) Mammalian cells express two differently localized Bag-1 isoforms generated by alternative translation initiation. Biochem. J. **328**, 807–813
31 Hann, S. R., Sloan-Brown, K. and Spotts, G. D. (1992) Translational activation of the non-AUG-initiated c-myc 1 protein at high cell densities due to methionine deprivation. Genes Dev. **6**, 1229–1240
32 Kevil, C., Carter, P., Hu, B. and DeBenedetti, A. (1995) Translational enhancement of FGF-2 by eIF-4 factors, and alternate utilization of CUG and AUG codons for translation initiation. Oncogene **11**, 2339–2348

Received 12 July 2005/15 August 2005; accepted 17 August 2005
Published as BJ Immediate Publication 15 August 2005, doi:10.1042/BJ20051114

© 2005 Biochemical Society