J. Biochem. 121, 868-875 (1997)

Isolation and Molecular Cloning of Epidermal- and Hair Follicle-Specific Peptidylarginine Deiminase (Type III) from Rat[1]

Takayuki Nishijyo,* Akira Kawada,† Takuya Kanno,* Masakazu Shiraiwa,* and Hidenari Takahara*,[2]

*Laboratory of Biochemistry, School of Agriculture, Ibaraki University, Ami-machi, Inashiki-gun, Ibaraki 300-03; and †Department of Dermatology, National Defense Medical College, Tokorozawa, Saitama 359

Received for publication, November 25, 1996

Peptidylarginine deiminase (PAD) is a post-translational modification enzyme that catalyzes deimination of arginine residues of proteins in the presence of calcium ions. There are three types of PAD in rodent tissues: PAD types I, II, and III [Terakawa *et al.* (1991) *J. Biochem.* 110, 661-666]. Type III enzyme was detected only in the epidermis and in hair follicles. In this study, we have purified PAD type III from 2-day-old rat epidermis by a four-step procedure that included soybean trypsin inhibitor-affinity chromatography. The enzyme was purified about 600-fold from the crude extract and the recovery was 23%. The final preparation of the enzyme gave only a single protein band on SDS-PAGE and showed an apparent molecular weight of 76,000. Subsequently, we cloned and sequenced the full-length cDNA encoding rat PAD type III by reverse transcription-polymerase chain reaction (RT-PCR) using degenerate oligonucleotide primers designed from the internal amino acid sequences and by the rapid amplification of the cDNA ends method. The composite cDNA sequence contained a 5′ untranslated region of 42 bp, an open reading frame of 1,995 bases that encoded 664 amino acids (M_r=75,036), a 3′ untranslated region of 1,063 bp, and part of a poly(A)$^+$ tail. The entire reading frame sequence of rat PAD type III showed 51% homology with that of rat PAD type II, and the C-terminal region is highly conserved between the two types. The cloned gene was expressed in *Escherichia coli* cells to produce PAD type III, which had not only enzymatic activity, but also immunoreactivity against specific antibodies toward PAD type II. Furthermore, the specific expression of the enzyme in the epidermis and hair follicles was confirmed by RT-PCR assays of mRNAs from several tissues.

Key words: cDNA cloning, peptidylarginine deiminase type III, purification, rat epidermis, tissue-specific expression.

About thirty years ago, Rogers (*1*) first found citrulline as a constituent of proteins from the inner root sheath of hair follicles and medulla of hair and quills. Citrulline residues were also found in several proteins from epidermis (*2-4*). Following the finding of citrulline residues in these proteins, it became clear that these tissues contain an enzyme, named peptidylarginine deiminase [EC 3.5.3.15: PAD], that converts the guanido side chain of certain arginine residues to the ureido side chains of citrulline in the presence of calcium ions (*5*). Recently, we have reported that there are three types of PAD in mouse tissues and we proposed designating them PAD types I, II, and III according to their elution order in anion-exchange column chromatography (*6*). Among these variants, PAD type II has been extensively studied. It is widely distributed in various kinds of tissues, such as brain (*7*), pituitary (*8*), salivary gland (*9*), pancreas (*9*), skeletal muscle (*10*), uterus (*9*), and yolk-sac erythroid cells (*11*), and the PAD type II enzymes purified from skeletal muscle (*12, 13*) and brain (*7, 14*) have been well characterized. The amino acid sequences of PAD type II of rat (*15*) and mouse (*16*) have been deduced from the nucleotide sequences. On the other hand, types I and III are present exclusively in the epidermis and hair follicles, but little is known about the enzymatic properties and cellular localization, because of the extremely limited amounts of these types available for study. Here we describe the isolation of PAD type III from newborn rat epidermis and its expression from a full-length cDNA. We also report the epidermal- and hair follicle-specific expression of the enzyme mRNA.

MATERIALS AND METHODS

Materials—Bz-L-Arg-*O*-Et, Bz-L-Arg, and soybean trypsin inhibitor (Kunitz) (STI) were purchased from

[1] This study was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan. The nucleotide sequence data reported in this paper have been submitted to the DDBJ, EMBL, and GenBank nucleotide sequence databases with the accession number D88034.

[2] To whom correspondence should be addressed. Tel/Fax: +81-298-88-8681, E-mail: takahara@ami.ipc.ibaraki.ac.jp

Abbreviations: Bz-L-Arg, benzoyl-L-arginine; Bz-L-Arg-*O*-Et, benzoyl-L-arginine-*O*-ethylester; G3PDH, glyceraldehyde-3-phosphate dehydrogenase; PAD, peptidylarginine deiminase; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcription-polymerase chain reaction; STI, soybean trypsin inhibitor (Kunitz).

Sigma Chemical. Q-Sepharose Fast Flow, aminohexyl-Sepharose 4B, CNBr-activated Sepharose 4B, random hexamer primer, *Not*I-d(T)$_{18}$, and *pKK223-3* vector were obtained from Pharmacia Biotechnology. STI coupled to CNBr-activated Sepharose 4B was prepared as described previously (*17*). *Staphylococcus aureus* V8 protease and Immobilon PSQ PVDF membrane were obtained from Millipore. Restriction enzymes and other enzymes for molecular biology were purchased from Takara Shuzo, Nippon Gene, and New England Biolabs, and were used as recommended by the manufacturers. The plasmid pBluescript SK and the primers (GA)$_{10}$*Xho*I(T)$_{18}$ and (GA)$_4$ *Xho*I(T)$_5$ were purchased from Stratagene Cloning Systems. Synthetic oligonucleotides were prepared with an Applied Biosystems Model 391 PCR-MATE.

Tissue Preparations—Skin obtained from 2-day-old rat (Sprague Dawley) was attached to surgical tape and soaked in cold 0.25 M NH_4Cl, pH 9.0, for 10 min. The dermis was pealed from the epidermis, and epidermal cells were obtained from the tape by gentle scraping parallel to the surface, frozen in liquid nitrogen, and kept at −80°C until use. All other tissues were collected from 10-week-old rats.

Purification of PAD Type III—Purification of PAD type III from the neonatal rat epidermis was carried out by the method described previously (*9*) with some modifications. The epidermis was homogenized in 5 volumes of 20 mM Tris-HCl, pH 7.6, containing 10 mM 2-mercaptoethanol, 1 mM EDTA, and 0.43 mM phenylmethylsulfonyl fluoride (buffer A) with an Ultra Turrax at 4°C. The homogenate was fractionated by ammonium sulfate precipitation (30-50%). The precipitate formed was collected by centrifugation, dissolved in buffer A, dialyzed against the buffer, and then applied to a column of Q-Sepharose Fast Flow. The column was washed with buffer A until the absorbance at 280 nm of the effluent was less than 0.1, and then the enzyme was eluted with a linear concentration gradient of NaCl from 0 to 1.0 M in buffer A. Enzymatic activity emerged as two peaks at about 0.3 and 0.5 M NaCl as shown in Fig. 1. Active fractions which corresponded to type III were concentrated by ammonium sulfate precipitation, dialyzed against buffer A containing 10 mM $CaCl_2$, 10 mM dithiothreitol, and 10% (v/v) glycerol, and applied to a column of STI-Sepharose previously equilibrated with the same buffer. The column was washed with the buffer, and the adsorbed enzyme was eluted with buffer A containing 10 mM EGTA and 10% glycerol. The active fractions from the STI-Sepharose affinity chromatography were applied to a column of aminohexyl-Sepharose previously equilibrated with buffer A containing 10% glycerol. The column was washed with the buffer until the absorbance at 280 nm of the effluent was less than 0.05 and the enzyme was eluted with a linear concentration gradient of NaCl from 0 to 1.0 M in buffer A containing 10% glycerol. The active fractions that emerged at about 0.8 M NaCl were pooled at −80°C until use. Rat PAD type II for comparative studies was purified from the skeletal muscle as described previously (*9*).

Assays of PAD Activity and Protein—PAD activity was assayed by measuring the formation of citrullyl residue from Bz-L-Arg-*O*-Et or Bz-L-Arg as described previously (*12*). One unit of the enzyme was defined as the amount of enzyme that catalyzed the formation of 1 μmol of citrullyl residues from arginyl residues in 1 h at 55°C. The concentration of protein was determined by the method of Read and Northcote (*18*) using bovine serum albumin as a standard.

SDS-PAGE—SDS-PAGE was performed on a slab gel by the method of Laemmli (*19*).

Peptide Mapping by In Situ V8 Protease Digestion in SDS-PAGE and Internal Amino Acid Sequencing—Principally, peptide mapping was carried out by the *in situ* V8 protease digestion method of Cleveland *et al.* (*20*). After the first SDS-PAGE (9% acrylamide, 0.24% bisacrylamide), the gel was briefly stained in a solution of 0.1% Coomassie Brilliant Blue, 40% (v/v) methanol, and 10% (v/v) acetic acid, and destained in a solution of 5% methanol and 10% acetic acid. Then, the band of the PAD was cut out, soaked in 125 mM Tris-HCl, pH 6.8, containing 0.1% SDS, and subjected to the second SDS-PAGE. The second SDS-PAGE was performed on a slab gel (15% acrylamide, 0.4% bisacrylamide). The gel slice was put into the bottom of the sample well and overlaid with 125 mM Tris-HCl, pH 6.8, containing 0.1% SDS, 10% glycerol, and 1 μg of *S. aureus* V8 protease. Hydrolysis of the PAD by the protease occurred in the stacking gel, and then the peptides were separated by the subsequent electrophoresis. After electrotransfer onto the PVDF membrane, the blots were stained with 0.1% Coomassie Brilliant Blue in 50% methanol/10% acetic acid and destained with 50% methanol/10% acetic acid. Finally, the peptides of interest were subjected to amino acid sequence analysis with an automated peptide sequencer (Applied Biosystems, model 477A). By this procedure, the sequences of two peptides were determined. They were: Peptide 1, VYGTPGVDIYV; Peptide 2, LGESDIIDIPQLFKSEKRKAVAFFPDLVNMXVLGKXLGIPKPFGP.

Preparation of Poly(A)$^+$ RNA from Rat Epidermis—Total RNA from rat epidermis was isolated according to Chomczynski and Sacchi (*21*). Poly(A)$^+$ RNA was isolated from the total RNA using a Poly(A)$^+$ Quick mRNA Purification Kit (Stratagene).

cDNA Cloning of PAD Type III by RT-PCR—Single-stranded cDNA was synthesized from 1 μg of poly(A)$^+$ RNA using random hexamers with SuperScript™ RNase H$^-$ Reverse Transcriptase (GIBCO BRL). The single-stranded cDNA product was subjected to PCR using sense and antisense primers designed from the internal amino acid sequences (Table II). The conditions for PCR were modified from those of the standard method (*22*). The PCR reaction mixture (50 μl) contained 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 0.5 mM of each nucleotide (dATP, dGTP, dTTP, and dCTP), 2 units of Taq DNA polymerase (Takara Shuzo), 50 pmol of each primer, and 1-5 μl of the above single-stranded cDNA product. The reaction mixture was overlaid with 50 μl of mineral oil and then subjected to 40 cycles of PCR. Cycling conditions were 94°C (1 min) for denaturation, 50°C (1 min) for annealing, and 72°C (3 min) for elongation. The sample was extracted once with chloroform to remove mineral oil and analyzed by 1% agarose gel electrophoresis. The DNA product was eluted from the gel and cloned into dT-tailed vectors prepared from pBluescript SK by the method of Marchuk *et al.* (*23*). Plasmid DNAs were isolated from 5 different clones and sequenced by the dideoxy chain termination method (*24*) in an automated DNA sequencer (Pharmacia, model ALF II).

Determination of the Full-Length cDNA Sequence—A partial rat PAD type III cDNA sequence with a total size of approximately 1,700 bp was determined by the above experiments. The sequences from the internal region to the 3′-end and to the 5′-end were obtained after generating PCR fragments by the method of rapid amplification of the cDNA ends (RACE) (*25*) with the primer sets listed in Table II. For obtaining the 3′-terminal region of the cDNA, rat epidermis poly(A)$^+$ RNA was reverse-transcribed with the $(GA)_{10}$*Xho*I$(T)_{18}$ primer. The single-stranded cDNA product was subjected to PCR using PRI-3 as a sense primer and $(GA)_4$*Xho*I$(T)_8$ as an antisense primer. The PCR was run for 40 cycles with the following schedule: 94°C for 1 min, 55°C for 1 min, and 72°C for 7 min. To obtain the 5′-end cDNA fragment, the poly(A)$^+$ RNA was reverse transcribed with PRI-4C and then a poly (dA) tail was added to the cDNA with terminal deoxynucleotidyl-transferase (*25*). The single-stranded cDNA product was subjected to PCR using *Not*I-d$(T)_{18}$ as a sense primer and PRI-5C as an antisense primer. The PCR was run for 40 cycles with the following schedule: 94°C for 1 min, 45°C for 2 min, and 72°C for 3 min. Amplified cDNAs were cloned into dT-tailed vectors and sequenced as described above.

Construction of Expression Plasmid for Rat PAD Type III—We constructed an expression plasmid for rat PAD type III according to the reported method for high-level expression of a recombinant PAD in *Escherichia coli* (*26*). First, two segments of the enzyme cDNA, which overlapped and covered the full-coding region, were amplified by RT-PCR as described above. Primers for amplification of the 5′-terminal region (nucleotides 43-1451 in Fig. 4) were designed as follows: sense primer, 5′-ACTGAATTCAGGATATTACTATGGAGGATTGATCT *ATGTCTCTGCAGAGGACTGTACG*-3′ which created an *Eco*RI site (indicated by a single underline) and included a pair of Shine-Dalgarno sequences and a short preceding cistron (indicated by a double underline) inserted into the adjacent deduced amino terminal sequence (indicated in italics); antisense primer, 5′-GCTCAGAAATTCATCCACGT-3′. The 3′-terminal coding region of the cDNA (nucleotides 1075-2076 in Fig. 4) was amplified with the following primer pair; sense primer, 5′-GACCGATGGATCCAGGATGA-3′; antisense primer, 5′-GGAAAGCTT*CTATCCTGGGGACGGATGG*-3′, which created a *Hin*dIII site (indicated by an underline) and included the complementary nucleotide sequence to that encoding the carboxyl terminus (indicated in italics). These PCRs were run for 40 cycles with the following schedule; 94°C for 1 min, 56°C for 1 min, and 72°C for 3 min. The *Eco*RI/*Xho*I-digested product of the 5′-terminal region and the *Hin*dIII/*Xho*I-digested product of the 3′-terminal region were ligated to each other, and inserted into the *Eco*RI/*Hin*dIII-digested *pKK223-3* vector. The resulting plasmid, named *pKKrPADIII*, was transformed into the host *E. coli* JM105 cells and its nucleotide sequence was confirmed by DNA sequencing as described above.

Detection of Rat PAD Type III Produced in E. coli by Immunoblotting—Preparation of cell-free extracts of *E. coli* cells harboring pKKrPADIII and immunochemical detection of the enzyme in the extracts were carried out according to the method reported previously (*26*). Rabbit anti-mouse skeletal muscle PAD serum (*9*) was preadsorbed with boiled lysates of the *E. coli* cells and then used for the immunoblotting.

Analysis of Expression of PAD Type III in Various Tissues—For the analysis of the expression of PAD type III mRNA in various tissues, we carried out semi-quantitative RT-PCR analysis. The amount of cDNA was confirmed by RT-PCR of glyceraldehyde-3-phosphate dehydrogenase (G3PDH) mRNA of each tissue. The cDNAs were synthesized from mRNAs of each tissue using random hexamer as described above. Primer pairs for rat PAD type III (sense primer, 5′-ATGTCTCTGCAGAGGACTGT-3′ at nucleotides 43-62 in Fig. 4; antisense primer, 5′-GCAAATGCTCGTGACTGGAG-3′ at nucleotides 333-362 in Fig. 4) and for rat G3PDH [sense primer, 5′-GTGAAGGTCGGAGTCAACGGAT-3′ at nucleotides 34-55 (*27*)]; antisense primer, 5′-CCAAATTCATTGTCATACCAGGA-3′ at nucleotides 958-980 (*27*) were used for the amplification of products of 310 bp (PAD type III) and 947 bp (G3PDH). Amplification was carried out for 1 min at 95°C, for 1 min at 50°C and for 3 min at 72°C. The authenticity of the PCR products was confirmed by subcloning and DNA sequencing. To confirm that the amplification proceeded exponentially, aliquots were removed after 25, 30, and 35 cycles. The reaction products were electrophoresed on a 2% agarose gel, stained with ethidium bromide, and photographed.

RESULTS AND DISCUSSION

Purification of PAD Type III from Newborn Rat Epidermis—As shown in Fig. 1, the PAD activity of the extract from rat epidermis was resolved into two peaks on a Q-Fast Flow anion-exchange column. The elution profile was similar to that of the murine epidermal cells reported previously (*6*). Furthermore, the substrate specificities of the peaks were similar to those of murine PAD types I and III (*6*). The first peak (type I) which emerged at about 0.3 M NaCl showed activity towards Bz-L-Arg as well as Bz-L-Arg-*O*-Et, whereas the second peak (type III) eluted at around 0.5 M NaCl preferentially deiminated Bz-L-Arg-*O*-Et. In this study, we collected the PAD fraction that corresponded to type III and further purified it by affinity column chromatographies on STI-Sepharose and aminohexyl-Sepharose. Table I summarizes the purification of the PAD type III. Starting from 65 g of epidermis, about 0.2 mg of purified enzyme with 23.4% yield of the activity was obtained. At the final stage, the enzyme had a specific activity of 47.6 units per mg of protein with Bz-L-Arg-*O*-Et as a substrate. SDS-PAGE of the sample at the final stage is shown in Fig. 2A. The final product showed a single band of 76,000 apparent molecular weight. We compared the molecular weight of types II and III on SDS-PAGE. As shown in Fig. 2A, the molecular weight of type III was smaller than that of type II, which was estimated to be 83,000 from SDS-PAGE. Watanabe *et al.* (*13*) also reported that the molecular weight of rat skeletal muscle PAD (type II) was 83,000, based on SDS-PAGE. Furthermore, they demonstrated that the calculation of the molecular weight (M_r=75,122) based on the deduced amino acid sequence of rat skeletal muscle PAD gave a value smaller than that estimated by SDS-PAGE (*15*). The same discrepancy was also observed with mouse PAD type II (*9, 16*). On the other hand, as described below, the calculated molecular weight of rat PAD type III based on its putative amino

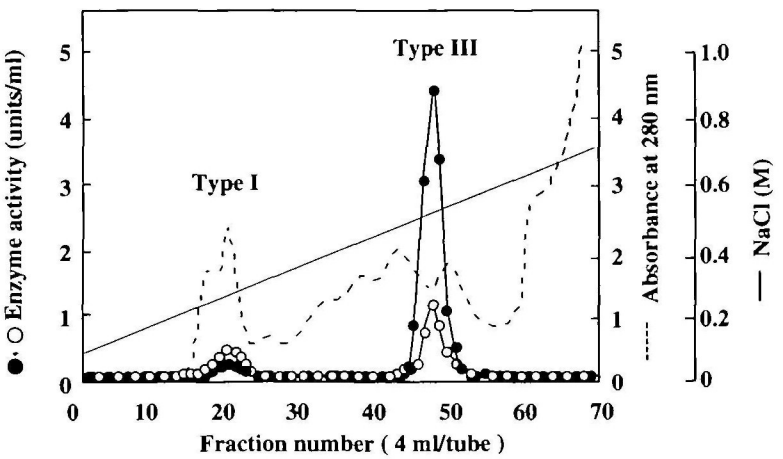

Fig. 1. **Q-Sepharose Fast Flow column chromatography of the extract from newborn rat epidermis.** The enzyme activity of each fraction was measured by using two different substrates, Bz-L-Arg-*O*-Et (○) and Bz-L-Arg (●) as described in "MATERIALS AND METHODS." Two isoforms of PAD (types I and III) were separately eluted on the column, and the active fractions of type III were collected for further purification.

TABLE I. **Purification of PAD type III from rat epidermis.**

Step	Total protein (mg)	Total activity (units)[a]	Specific activity (units/mg of protein)	Purity (fold)	Yield (%)
Crude extract	465	37.2	0.083	1.00	100
$(NH_4)_2SO_4$ (0-50%)	338	33.5	0.099	1.24	90.1
Q-Sepharose Fast Flow	65.9	30.0	0.455	5.69	80.6
STI-Sepharose	2.57	11.3	4.40	55.0	30.4
Aminohexyl-Sepharose	0.183	8.71	47.6	595.0	23.4

[a] One unit was defined as the amount of enzyme that catalyzed the formation of 1 μmol of Bz-L-Arg-*O*-Et for 1 h at 55°C.

acid sequence is very close to that derived from SDS-PAGE (75,036 *versus* 76,000).

Amino-Terminal and Internal Amino Acid Sequences—The N-terminal sequence of rat PAD type III could not be determined because of blocking. Therefore, we determined the amino acid sequences of internal peptides obtained by Cleveland's hydrolysis method (*20*). Figure 2B showed the peptides derived from types II (lane 1) and III (lane 2) on SDS-PAGE. Some differences were observed when the peptide profiles from types II and III were compared. Thus, we selected similar and distinct peptides for amino acid sequence analyses. The peptides were transferred onto PVDF membrane, and two (Peptides 1 and 2 in Fig. 2B) from type III were subjected to amino acid sequencing. The N-terminal amino acid sequence of Peptide 1 (36 kDa) was VYGTPGVDIYV, whereas that of Peptide 2 (12 kDa) was LGESDIIDIPQLFKSEKRKAVAFFPDLVNMXVLGKXL-GIPKPFGP. The sequence of Peptide 1 is different from any known protein (including PAD type II) or gene product sequence in the SwissProt and GenBank databases. On the other hand, the sequence of Peptide 2 shares 65% identity with that of the C-terminal region (amino acid residues 559-606) of rat PAD type II (*15*). From these findings, we presumed that Peptide 1 is located in the upstream region of the enzyme relative to Peptide 2. Using these data, we designed primers for RT-PCR cloning of rat PAD type III cDNA (Table II).

Cloning and Analysis of Full-Length cDNA—Figure 3 shows the strategy for cloning rat PAD type III cDNA. RT-PCR with poly(A)$^+$ RNA from the newborn rat epidermis was carried out according to the method described in "MATERIALS AND METHODS." The RT-PCR using a primer pair PRI-1/PRI-2C produced an approximately 1.7 kbp cDNA fragment (*pPAD3-1*). This fragment was subcloned into the dT-tailed vector and sequenced. The sequence

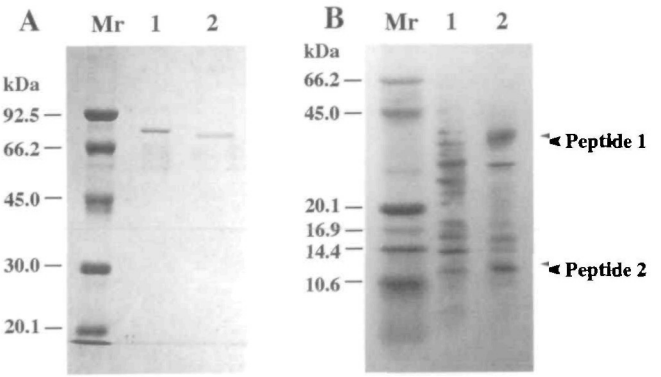

Fig. 2. **SDS-PAGE (A) and peptide map (B) of the purified PAD type III.** A: Samples were run on a 9% gel and stained with Coomassie Brilliant Blue. Lane 1, the final preparation of the enzyme chromatographed on an aminohexyl-Sepharose column (1 μg); lane 2, PAD type II purified from the rat skeletal muscle (2 μg); lane Mr, molecular weight markers. B: The hydrolysates of the PADs by *S. aureus* V8 protease were subjected to SDS-PAGE by Cleveland's method (*21*). Lane 1, the hydrolysates from rat PAD type II; lane 2, the hydrolysates from rat PAD type III; lane Mr, molecular weight markers. Peptide 1 (36 kDa) and Peptide 2 (12 kDa) showed by arrows were analyzed by the protein sequencer.

analyses identified a 1,683 bp clone encoding a peptide of 561 amino acids, including sequences identical to those of Peptides 1 and 2 (Fig. 4). Therefore, the PCR product was judged to correspond to a PAD type III. The rest of the 3′-end region of the PAD cDNA was extended by the 3′-RACE method as described in "MATERIALS AND METHODS." As shown in Fig. 3, approximately 1.4 kbp of the amplified fragment (*pPAD3-2*) covered the rest of the 3′-terminal region of the cDNA and also overlapped with the 3′-terminal region of *pPAD3-1*. Finally, we determined

the rest of the sequence of the 5′-terminus by the 5′-RACE method as described in "MATERIALS AND METHODS." As shown in Fig. 3, we obtained a cDNA fragment (*pPAD3-3*) corresponding to the 5′-terminal region of the cDNA. *pPAD3-3* overlapped *pPAD3-1* and was composed of a coding region (310 bp) and a 5′-untranslated region (42 bp). The nucleotide sequence of the cDNA thus determined is summarized in Fig. 4 along with the deduced amino acid sequence. The rat PAD type III cDNA consists of a coding region of 1,995 bp, a 3′-noncoding region of 1,063 bp and a 5′-noncoding region of 42 bp. The initiation methionine was assigned based on the following findings: (1) the sequence of GCCGGCATG at the region of the translational initiation is very similar to the consensus sequence GCC(A/G)CCATG for the eukaryotic translation initiation site (*28*); (2) a nonsense codon (TAG) was found in-frame upstream of the initiation codon (nucleotides 34 to 36). The 3′-untranslated region includes a poly(A)$^+$ tail and a consensus sequence (AATAAA) for polyadenylation (*29*). Thus, it is likely that the nucleotide sequence determined by this study covers the full-length cDNA encoding rat PAD type III.

Analysis of the Deduced Amino Acid Sequence—Rat PAD type III cDNA encodes a putative protein of 664 amino acid residues with a calculated molecular weight of 75,036 Da. This is very close to the result of SDS-PAGE (Fig. 2A). The N-terminal sequence of rat PAD type III was not determined in this study because of blocking. N^a-acetylation of the N-terminal amino acid residue is one of the predominant blocking modifications. A review of the published N-terminal sequences of N^a-acetylated proteins showed that small amino acids such as Gly, Ala, and Ser, adjacent to the first Met are mostly N^a-acetylated (*30*). Thus, we speculated that the enzyme undergoes amino-terminal processing (removal of the first methionine and subsequent *N*-acetylation of the serine residue adjacent to the methionine). Figure 5 shows the alignment of the amino acid sequence of the type III with that of the type II. Alignment of the amino acid sequence with that of rat PAD type II yielded an identity of 50.8%. However, the homology between the two PADs is variable within the molecule. The N-terminal region of the molecule shows a low degree of homology, whereas the C-terminal region is highly conserved and shows only minor differences within the last 323 amino acid. This conservation may be related to the function of the region of the protein which is responsible for the catalytic activity.

TABLE II. **Primers used for cloning rat PAD type III.**

Name	Nucleotide sequence
PRI-1[a]	5′-ACCCCAGGCGTGGACATCTACGTG-3′
PRI-2C[b]	5′-GGCCCGAAGGG(C/T)TTNGG(G/A/T)AT-3′
PRI-3[c]	5′-GACATCATTGACATCCCACAGCTCTT-3′
PRI-4C[d]	5′-ATCCCTGGGAAGCCTGCATC-3′
PRI-5C[e]	5′-GCAAATGCTCGTGACTGGAG-3′
$(GA)_{10}$*Xho*I$(T)_{18}$[f]	5′-$(GA)_{10}$ACTAGTCTCGAG$(T)_{18}$-3′
$(GA)_4$*Xho*I$(T)_5$[g]	5′-$(GA)_4$ACTAGTCTCGAG$(T)_5$-3′
*Not*I-d$(T)_{18}$[h]	5′-AACTGGAAGAATTCGCGGCCGCAGGAA-$(T)_{18}$-3′

[a]PRI-1 was a sense primer designed from TPGVDIYV of Peptide 1. [b]PRI-2C was an antisense primer designed from IPKPFGP of Peptide 2. [c]PRI-3 (at nucleotides 1735-1760 in Fig. 4) was a part of the sequence of *pPAD3-1* and used as a sense primer for 3′-RACE. [d]PRI-4C (at nucleotides 814-833 in Fig. 4) was a primer used for single-stranded cDNA synthesis for 5′-RACE. [e]PRI-5C (at nucleotides 333-352 in Fig. 4) was an antisense primer that was complementary to a part of *pPAD3-1*. [f]$(GA)_{10}$*Xho*I$(T)_{18}$ was a primer used for single-stranded cDNA synthesis for 3′-RACE. [g]$(GA)_4$*Xho*I-$(T)_5$ was an antisense primer for 3′-RACE. [h]*Not*I-d$(T)_{18}$ was a sense primer for 5′-RACE.

Expression of PAD Type III in E. coli Cells—In the previous paper, we reported the construction of a two-cistron expression plasmid (*pKKPAD4*) for high-level gene expression of the mouse skeletal muscle PAD (type II) gene in *E. coli* (*26*). *pKKrPAD III* constructed in this study had the same promoter sequence as *pKKPAD4*. The values of specific activity of *E. coli* cells harboring *pKKrPAD III* for Bz-L-Arg-*O*-Et and Bz-L-Arg were 0.86±0.06 and 0.15±0.02 unit/mg, respectively. Figure 6 shows the results of the SDS-PAGE and immunoblotting of an extract of *E. coli* cells harboring *pKKrPAD III* with rabbit anti-mouse PAD type II antibody. An extra protein band of about 76 kDa could be seen in the extracts from *E. coli* cells harboring the expression plasmid (Fig. 6A). Furthermore, the extra protein was immunoreactive with the PAD type II antibody (Fig. 6B). These results show that the cloned gene was expressed in *E. coli* cells to produce PAD type III, which shares common antigenicity with the PAD type II.

Specific Expression of PAD Type III mRNA in Epidermal and Hair Follicle Cells—Previously, we found the enzyme activity of PAD type III only in the epidermal and hair follicle cells of all tissues tested (*6*). To evaluate whether the localization of the enzyme in the above tissues is a result of the specific expression of the PAD type III mRNA, we compared the enzyme mRNA levels in several tissues by RT-PCR. Poly(A)$^+$ RNAs from the tissues were reverse-transcribed to cDNA and amplified using the specific primers for rat PAD type III. As shown in Fig. 7, extremely high expression levels were found in epidermis and hair follicle. In liver, pancreas, spleen, and ovary, the mRNA of the enzyme was only slightly expressed. Brain, heart, lung, stomach, spleen, kidney, uterus, and skeletal

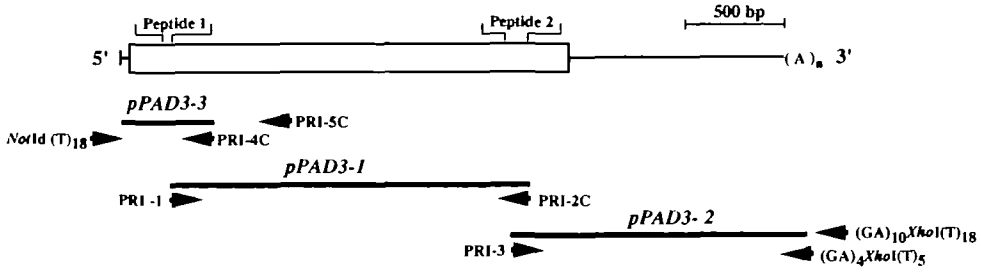

Fig. 3. **Cloning strategy of rat PAD type III by RT-PCR, 3′-RACE, and 5′-RACE.** The clone *pPAD3-1* shown by thick bars was obtained by RT-PCR. *pPAD3-2* and *pPAD3-3* were cloned by 3′- and 5′-RACE, respectively. These sequences were combined to construct the composite cDNA in which the open box and lines indicate the coding and noncoding sequences, respectively. The locations of Peptides 1 and 2 are shown on the open box. Horizontal arrows denote primers used for RT-PCR, 3′-RACE, and 5′-RACE amplifications.

```
GAGTGCTGGGTTTGGTGGCCAAAGCCCCCATCCTAGGCCGGCATGTCTCTGCAGAGGACTGTACGTGTATCCCTGGAACACCCAACCAGC   90
                                    M  S  L  Q  R  T  V  R  V  S  L  E  H  P  T  S              16

GCTGTGTGTGTGGCAGGGGTTGAGACCATCGTGGACATTTATGGGTCGGTTCCTGAGGGCACAGACATGTTTGAGGTCTATGGGACACCG  180
A  V  C  V  A  G  V  E  T  I  V  D  I  Y  G  S  V  P  E  G  T  D  M  F  E  V  Y  G  T  P     46

GGAGTGGACATCTACGTGTCTCCCAGCATGGAGAGGAACCGAGAGCGAGCTGACACCAGGCGGTGGTGCTTCAACAAGGGGCTGGAGATC  270
G  V  D  I  Y  V  S  P  S  M  E  R  N  R  E  R  A  D  T  R  R  W  C  F  N  K  G  L  E  I     76

ATCGTGGTCATGAACTCTCCTAGCAATGACCTCAACGACAGTCACGTTCAAATAGCTTACCACTCCAGTCACGAGCATTTGCCCCTGGCC  360
I  V  V  M  N  S  P  S  N  D  L  N  D  S  H  V  Q  I  A  Y  H  S  S  H  E  H  L  P  L  A     106

TATGCAGTACTCTACCTCACCTGTGTGGACATCACCCTGGACTGTGACATGAACTGTGCTGACCGGCAGGACAGGAGCTTCGTGGATAAG  450
Y  A  V  L  Y  L  T  C  V  D  I  T  L  D  C  D  M  N  C  A  D  R  Q  D  R  S  F  V  D  K     136

AGGCAGTGGATGTGGGGACCTGATGGGTACGGAGCCATCTTGCTGGTGAACTGTGATCGGGACGAGGTCAGCTCTGATGCCCAAGATAAC  540
R  Q  W  M  W  G  P  D  G  Y  G  A  I  L  L  V  N  C  D  R  D  E  V  S  S  D  A  Q  D  N     166

TGTGACCAATGTGTGCGCTGCCTGCAAGATCTGGAAGACATGTCCGTCATGGTCCTACGGACCCAGGGCCCCGAATCTCTCTTTGACGAC  630
C  D  Q  C  V  R  C  L  Q  D  L  E  D  M  S  V  M  V  L  R  T  Q  G  P  E  S  L  F  D  D     196

CACAGGCTTATCCTCCACACCTCCAGCTGTGATGCGGAGCGGGCACGGGTCTTCCACGTCTGTGGCCCCGAGGATTCCTGCGAAGCCTAT  720
H  R  L  I  L  H  T  S  S  C  D  A  E  R  A  R  V  F  H  V  C  G  P  E  D  S  C  E  A  Y     226

AGGTGTGTTCTGGGCCCGGACCGGATGTCCTATGAGGTACCCCGCCTGAAGGGCTATGAGGAGCGCTTCTACGTGGAGGGCCTCTCCTTC  810
R  C  V  L  G  P  D  R  M  S  Y  E  V  P  R  L  K  G  Y  E  E  R  F  Y  V  E  G  L  S  F     256

CCTGATGCAGGCTTCCCAGGGATCCTCTCCTTCCATATCACCCTGCTGGATGACTCCAATGAGGACTACTCCGAGACCCCGATCTTCACG  900
P  D  A  G  F  P  G  I  L  S  F  H  I  T  L  L  D  D  S  N  E  D  Y  S  E  T  P  I  F  T     286

GACACAGTGGTGTTCCGAGTGGCACCCTGGATCATGACGCCCAGCACCCTGCCACCCCTGGAGGTGTACGTGTGCCAGGTGAGGAATAAC  990
D  T  V  V  F  R  V  A  P  W  I  M  T  P  S  T  L  P  P  L  E  V  Y  V  C  Q  V  R  N  N     316

ACCTGCTTTGTGGAAGCTGTGGAGGAGCTGGCCAGGAAGGCAGGCTGCAAACTCACCATCTGTCCTCAGGCAGAGAACCGCAATGACCGA 1080
T  C  F  V  E  A  V  E  E  L  A  R  K  A  G  C  K  L  T  I  C  P  Q  A  E  N  R  N  D  R     346

TGGATCCAGGATGAGATGGAGCTGGGGTACACGCAGGCGCCACACAAGACTTTGCCTGTGGTCTTTGATTCTCCGAGGAATGGGGAGCTG 1170
W  I  Q  D  E  M  E  L  G  Y  T  Q  A  P  H  K  T  L  P  V  V  F  D  S  P  R  N  G  E  L     376

CAGGGTTTCCCTTACAAAAGGATCCTGGGTCTAGATTTTGGCTACGTGACTCGAGAACCTCCAGACAGTTCTGTGAGCGGTCTGGACTCC 1260
Q  G  F  P  Y  K  R  I  L  G  L  D  F  G  Y  V  T  R  E  P  P  D  S  S  V  S  G  L  D  S     406

TTTGGCAACCTGGAGGTCAGCCCGCCCGTGGTGGCCAATGGGAAGGAGTATCCACTGGGGAGGATTCTCATCGGGGGCAACCTGCCTGGG 1350
F  G  N  L  E  V  S  P  P  V  V  A  N  G  K  E  Y  P  L  G  R  I  L  I  G  G  N  L  P  G     436

TCGAGAGGCCGCAGAGTTACCCAGGTGGTACGAGACTTCCTCCATGCCCAGAAGGTGCAACCTCTGGTGGAGTTGTTTGTGGACTGGCTG 1440
S  R  G  R  R  V  T  Q  V  V  R  D  F  L  H  A  Q  K  V  Q  P  L  V  E  L  F  V  D  W  L     466

GCTGTGGGCCACGTGGATGAATTTCTGAGCTTCGTCCCTGCACCAGATGGAAAGGGCTTCCGGCTCCTCCTGGCCAGCCCTGGTGCCTGC 1530
A  V  G  H  V  D  E  F  L  S  F  V  P  A  P  D  G  K  G  F  R  L  L  L  A  S  P  G  A  C     496

TTCAGGCTATTCCAGGAAAAGCAGAAGTGGGGCCATGGCCGGTCCCTCCTTTTCGAGGGGGTCATTGGTGACAGGCGGGTCCAGACCATC 1620
F  R  L  F  Q  E  K  Q  K  W  G  H  G  R  S  L  L  F  E  G  V  I  G  D  R  R  V  Q  T  I     526

TCCATCAATCAGGTTCTTTCCAACCAGAGCCTTATCAACTTTAATAAGTTTGCACAGAGCTGTATTGACTGGAACCGTGAGGTGCTGAAG 1710
S  I  N  Q  V  L  S  N  Q  S  L  I  N  F  N  K  F  A  Q  S  C  I  D  W  N  R  E  V  L  K     556

CGGGAGCTAGGCCTGGGGGAGAGCGACATCATTGACATCCCACAGCTCTTCAAGTCGGAGAAGAGGAAGGCGGTGGCGTTCTTCCCGGAC 1800
R  E  L  G  L  G  E  S  D  I  I  D  I  P  Q  L  F  K  S  E  K  R  K  A  V  A  F  F  P  D     586

CTGGTGAACATGCTGGTACTGGGCAAGCACCTGGGCATCCCTAAGCCCTTCGGGCCCATCATCAATGGCCGCTGCTGCCTGGAGGAGAAG 1890
L  V  N  M  L  V  L  G  K  H  L  G  I  P  K  P  F  G  P  I  I  N  G  R  C  C  L  E  E  K     616

GTGCGCTCCCTGCTGGAGCCGCTGGGTCTGCACTGCACCTTCATTGATGACTTCACACCGTACCACATGCTGCACGGGGAGGTTCACTGT 1980
V  R  S  L  L  E  P  L  G  L  H  C  T  F  I  D  D  F  T  P  Y  H  M  L  H  G  E  V  H  C     646

GGCACCAACGTGCGCCGGGAGCCCTTTGCCTTCAAGTGGTGGCACATGGTGCCCTGAGCCGCCCGCCACCCATCCCCCATCCGTCCCCCA 2070
G  T  N  V  R  R  E  P  F  A  F  K  W  W  H  M  V  P  end                                     664

GGATAGGGCATCCGCCCGGCAGGTGGAGACTGCGACAGACCCCAGAGCACCCAGAGACTCCCAGTTTCTAGACTGAATGACCCAAGCTTC 2160
TCAGAAGCCTTTTCTCTTGAAATGTCCGCATCCACCTGCTGCCCATTCAGTTCCCAGACATGAATCCTCTTGGTGCCCCCAAAGAAGGGT 2250
CCCATGTCAACCAAACTCTCCTGCGGGTCGAGATGATCCATGGGGATAACTTTAAATACCATTCACCCCCCCCCCACCCCCGCAAACTCC 2340
AAAACTCCTCTTTGTGGGGAAAAATCCGCTTGGGTATCTAGAAGCATTGACCTTGGCTTGTGCTTCCTATGGTGGAAGGCCTGGGGCGGT 2430
GGGGGTTGGTGGGGTTGTTTTGGCCATTGCCTCCTGCCCTGCAGGGGATGGTTGTAAAGTCAGACTATTTCACACCCTCACTACTCGACT 2520
AGAGGGAGTCTACCCTCCCAGGAAGCACCCCGGGTCCATTCAGGTCAAAGTGTGTCCATCATGTCCACTGCAAGCTCCCGGCAGGACCGG 2610
CCTTTACTGCAAGCCAGCCTGGGCTATCACACGGTGGGGACCTGTCACACAATGGGGACCTCTAGATGCTTTTGCAATGACAGATGTCCC 2700
TTAGCTTTCCAGGCTTCCCATCATCCGATGTCTGGGAATCTGGTTGTCCAAATCCTTTTGGAGTTTCTCTGAAGTGAATGATTGGGGGGC 2790
TGGATTCCCAGGGGAGAACGCCTTTCCTAAGACAGGATTGACCCTCGGGGTCCCACCATCTGAAGGGGTTCCTCCATCAGCTTGCCTGGG 2880
GGCTCTGTGCAGAGGAGACCCCAGGCTGACTTAGCCCAGAGCCTTCTGCCTATACAGCCATCCCGTAGTAGAGGCCCGGGAAGTGCCACC 2970
CTGGGTGCCATCAAGGCAGGACTTAGAGGGACAGTTAAGCCTCTCCACATACTGGGAGAAGGGAGGGATGTGGCTGTCTCAGTATGGGTC 3060
TTCTGTGAGTTCAGCTGGCCAAAAAATAAAACAGAGGCTCAAAAAAAAAAAAAAAA                                    3115
```

Fig. 4. **Nucleotide sequence and deduced amino acid sequence of rat PAD type III.** Nucleotides and amino acids are numbered at the right. An in-frame stop codon in the 5′-untranslated sequence is indicated below the nucleotide sequence. A putative polyadenylation signal in the 3′-untranslated sequence is boxed. The determined amino acid sequences of Peptides 1 and 2 are double-underlined.

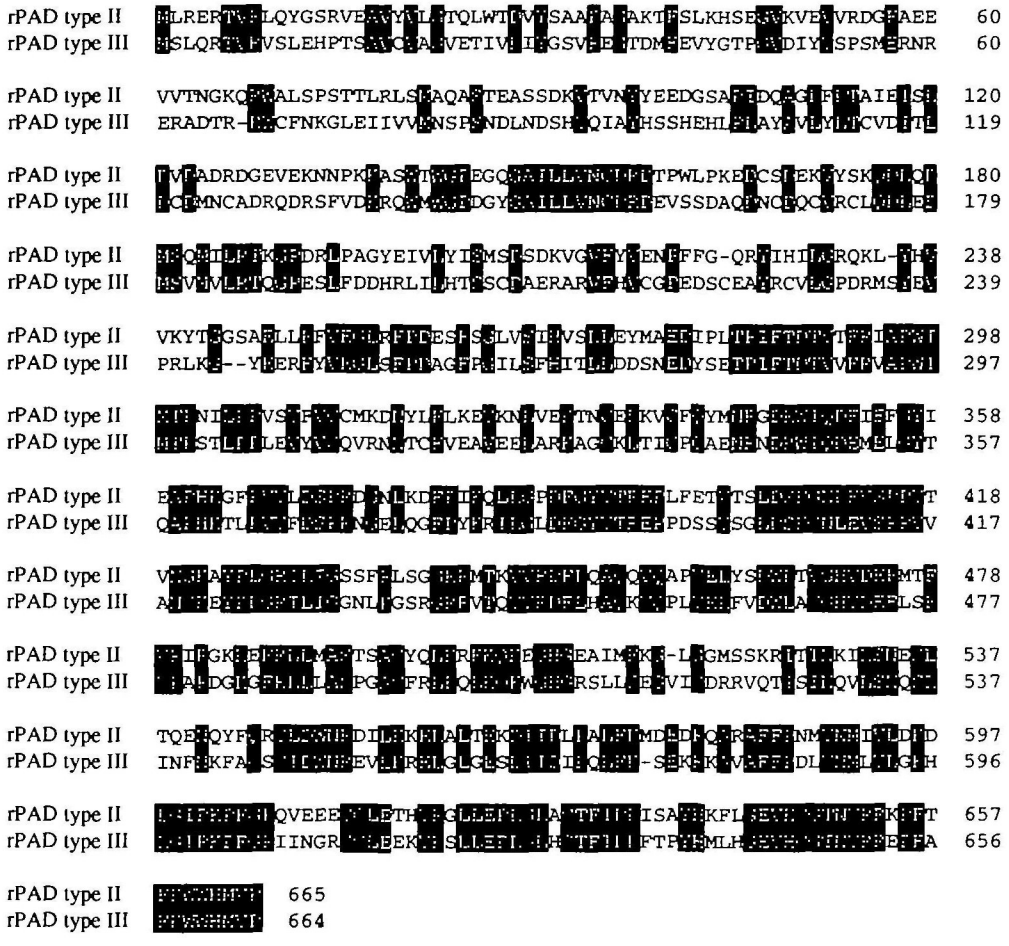

Fig. 5. **Sequence comparison between rat PAD type II and III.** Amino acids encoded by rat PAD type II (*15*) and III are aligned, and identical residues are shown by black boxes. Dashes indicate gaps inserted to maximize alignment.

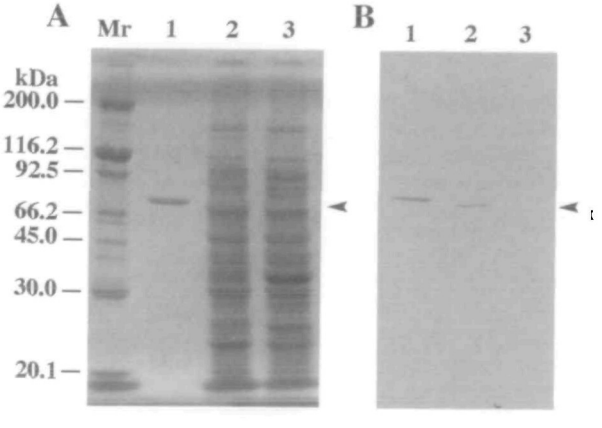

Fig. 6. **Detection of PAD type III in *E. coli* extracts by SDS-PAGE (A) and immunoblotting (B) analyses.** Lane Mr, molecular weight markers; lanes 1, purified PAD type II (1 μg for SDS-PAGE and 0.1 μg for immunoblotting analyses); lanes 2, lysate (25 μg protein for SDS-PAGE and 5 μg protein for immunoblotting analyses) of JM 105 cells harboring the plasmid *pKKrPAD III*; lanes 3, lysate (25 μg protein for SDS-PAGE and 5 μg protein for immunoblotting analyses) of JM 105 cells harboring the control plasmid *pKK223-3*. An extra band of 76 kDa in the extract of *pKKrPAD III* on SDS-PAGE and an immunoreactive band corresponding to PAD type III are indicated by arrows.

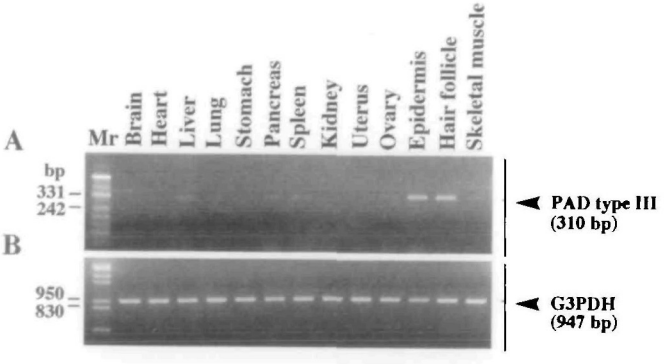

Fig. 7. **Analysis of PAD type III expression in several rat tissues by semi-quantitative RT-PCR.** The PAD type III mRNA of each tissue was analyzed by RT-PCR as described in "MATERIALS AND METHODS." The PAD type III mRNA-amplified fragment is represented by 310 bp bands, whereas G3PDH controls are represented by 947 bp bands. To confirm that the amplification proceeded exponentially, aliquots were removed after 25, 30, and 35 cycles and assayed. This figure shows the reaction products after 30 cycles of the PCR.

muscle displayed trace or undetectable levels of the PAD mRNA. In parallel analyses, we employed G3PDH primers for RT-PCR, and these data indicate that different amounts of mRNA applied did not account for the observed differences in the PAD expression by RT-PCR, as similar levels of G3PDH could be detected in all tissues. These data agreed with the previous findings on the tissue distribution of the PAD type III activity (*6*). The cloning of PAD type III

should extend our understanding of the function of this enzyme in epidermis and hair follicle, and the availability of recombinant PAD will help us to study its enzymatic and physicochemical properties in detail.

We are grateful to Dr. Thanh Vu, Stanford University, for his helpful suggestions.

REFERENCES

1. Rogers, G.E. (1962) Occurrence of citrulline in proteins. *Nature* **194**, 1149-1151
2. Sugawara, K. (1979) Presence of citrulline in the membranous proteins of stratum corneum of newborn rat and cow snout. *Agric. Biol. Chem.* **43**, 2215-2217
3. Harding, C.R. and Scott, I.R. (1983) Histidine-rich proteins (filaggrins): structural and functional heterogeneity during epidermal differentiation. *J. Mol. Biol.* **170**, 651-673
4. Senshu, T., Akiyama, K., Kan, S., Asaga, H., Ishigami, A., and Manabe, M. (1995) Detection of deiminated proteins in rat skin. *J. Invest. Dermatol.* **105**, 163-169
5. Fujisaki, M. and Sugawara, K. (1981) Properties of peptidylarginine deiminase from the epidermis of newborn rats. *J. Biochem.* **89**, 257-263
6. Terakawa, H., Takahara, H., and Sugawara, K. (1991) Three types of mouse peptidylarginine deiminase: characterization and tissue distribution. *J. Biochem.* **110**, 661-666
7. Kubilus, J. and Baden, H.P. (1983) Purification and properties of a brain enzyme which deiminates proteins. *Biochim. Biophys. Acta* **745**, 285-291
8. Senshu, T., Akiyama, K., Nagata, S., Watanabe, K., and Hikichi, K. (1989) Peptidylarginine deiminase in rat pituitary: sex difference, estrous cycle-related changes, and estrogen dependence. *Endocrinology* **124**, 2666-2670
9. Takahara, H., Tsuchida, M., Kusubata, M., Akutsu, K., Tagami, S., and Sugawara, K. (1989) Peptidylarginine deiminase of the mouse. Distribution, properties, and immunocytochemical localization. *J. Biol. Chem.* **264**, 13361-13368
10. Sugawara, K., Oikawa, Y., and Ouchi, T. (1982) Identification and properties of peptidylarginine deiminase from rabbit skeletal muscle. *J. Biochem.* **91**, 1065-1071
11. Koike, H., Shiraiwa, M., Sugawara, K., Kohsaka, T., and Takahara, H. (1995) Existence and differential changes of peptidylarginine deiminase type II in mouse yolk-sac erythroid cells. *Biosci. Biotech. Biochem.* **59**, 552-554
12. Takahara, H., Oikawa, Y., and Sugawara, K. (1983) Purification and characterization of peptidylarginine deiminase from rabbit skeletal muscle. *J. Biochem.* **94**, 1945-1953
13. Watanabe, K., Akiyama, K., Hikichi, K., Ohtsuka, R., Okuyama, A., and Senshu, T. (1988) Combined biochemical and immunochemical comparison of peptidylarginine deiminases present in various tissues. *Biochim. Biophys. Acta* **966**, 375-383
14. Lamensa, J.W. and Moscarello, M.A. (1993) Deimination of human myelin basic protein by a peptidylarginine deiminase from bovine brain. *J. Neurochem.* **61**, 987-996
15. Watanabe, K. and Senshu, T. (1989) Isolation and characterization of cDNA clones encoding rat skeletal muscle peptidylarginine deiminase. *J. Biol. Chem.* **264**, 15255-15260
16. Tsuchida, M., Takahara, H., Minami, N., Arai, T., Kobayashi, Y., Tsujimoto, H., Fukazawa, C., and Sugawara, K. (1993) cDNA nucleotide sequence and primary structure of mouse uterine peptidylarginine deiminase. Detection of a 3′-untranslated nucleotide sequence common to the mRNA of transiently expressed genes and rapid turnover of this enzyme's mRNA in the estrous cycle. *Eur. J. Biochem.* **215**, 677-685
17. Takahara, H., Okamoto, H., and Sugawara, K. (1986) Affinity chromatography of peptidylarginine deiminase from rabbit skeletal muscle on a column of soybean trypsin inhibitor (Kunitz)-Sepharose. *J. Biochem.* **99**, 1417-1424
18. Read, S.M. and Northcote, D.H. (1981) Minimization of variation in the response to different proteins of the Coomassie blue G dye-binding assay for protein. *Anal. Biochem.* **116**, 53-64
19. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
20. Cleveland, D.W., Fischer, S.G., Kirschner, M.W., and Laemmli, U.K. (1977) Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *J. Biol. Chem.* **252**, 1102-1106
21. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156-159
22. McPherson, M.J., Hames, B.D., and Taylor, G.R. (1995) *PCR 2. A Practical Approach*, Oxford University Press, Oxford, NY
23. Marchuk, D., Drumm, M., Saulino, A., and Collins, F.S. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19**, 1154
24. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
25. Frohman, M.A., Dush, M.K., and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002
26. Ohsugi, I., Takahara, H., Shiraiwa, M., and Sugawara, K. (1995) Expression of mouse uterine peptidylarginine deiminase in *Escherichia coli*: construction of expression plasmid and properties of the recombinant enzyme. *Arch. Biochem. Biophys.* **317**, 62-68
27. Tso, J.Y., Sun, X.H., Kao, T.H., Reece, K.S., and Wu, R. (1985) Isolation and characterization of rat and human glyceraldehyde-3-phosphate dehydrogenase cDNAs: genomic complexity and molecular evolution of the gene. *Nucleic Acids Res.* **13**, 2485-2502
28. Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12**, 857-872
29. Proudfoot, N.J. and Brownlee, G.G. (1976) 3′ non-coding region sequences in eukaryotic messenger RNA. *Nature* **263**, 211-214
30. Tsunasawa, S. and Sakiyama, F. (1984) Amino-terminal acetylation of proteins: an overview. *Methods Enzymol.* **106**, 165-170