Cloning of cDNA Encoding PAS-4 Glycoprotein, an Integral Glycoprotein of Bovine Mammary Epithelial Cell Membrane

Sik Hwangbo, Soo-Won Lee* and Chouemon Kanno[1]

Faculty of Life Science and Technology, Sungkyunkwan University, Suwon, Kyunggi 440-746, Korea

ABSTRACT : Bovine PAS-4 is an integral membrane glycoprotein expressed in mammary epithelial cells. Complementary DNA (cDNA) cloning of PAS-4 was performed by reverse-transcriptase polymerase chain reaction (RT-PCR) with oligonucleotide probes based on it's amino terminal and internal tryptic-peptides. The cloned PAS-4 cDNA was 1,852 nucleotides (nt) long and its open reading frame (ORF) was encoded 1,413 base long. The deduced amino acid sequence indicated that PAS-4 consisted of 471 amino acid residues with molecular weight of 52,796, bearing 8 potential *N*-glycosylation sites and 9 cysteine residues. Partial bovine CD36 cDNA from liver also was sequenced and the homology of both nucleotide sequence was 94%. Most of the identical amino acid residues were in the luminal/extracellular domains. Contrary to PAS-4, bovine liver CD36 displays 6 potential *N*-glycosylation sites, which were located, except for those at positions 101 and 171, at same positions as PAS-4 cDNA. Cysteine residues of PAS-4 and CD36 were same at position and in numbers. Northern blot analysis showed that PAS-4 was widely expressed, although its mRNA steady-state levels vary considerably among the analyzed cell types. PAS-4 possessed hydrophobic amino acid segments near the amino- and carboxyl-termini. Two short cytoplasmic tails of the amino- and carboxyl-terminal ends constituted of a 5-7 and 8-11 amino acid residues, respectively. ***(Asian-Aust. J. Anim. Sci. 2001. Vol 15, No. 4 : 576-584)***

Key Words : cDNA, PAS-4, CD36, Glycoprotein, MFGM, RT-PCR

INTRODUCTION

PAS-4 is a 78kDa integral membrane glycoprotein of the milk fat globule membrane (MFGM) that coats discrete milk fat droplets and is derived from the apical surface of mammary epithelial cells during the milk secretion process (Freudenstein et al., 1979; Greenwalt et al., 1985; Greenwalt and Mather, 1985). A component corresponding to PAS-4 has been found in the membrane fraction of lactating mammary gland (Kanno et al., 1982). PAS-4 glycoprotein is one of the major components constituting of MFGM and resistant to proteolysis when present in the intact membrane (Greenwalt et al., 1985; Greenwalt and Mather, 1985). The function of this protein remains unknown. Partial sequence of PAS-4 has been printed by Greenwalt et al. (1990) to be related to CD36 found in the surface of endothelial cells and platelet (Clemetson and McGregor, 1987).

CD36 has recently been shown to mediate cellular interactions with the extracelluar matrix components, collagen, thrombospondin, oxidized low density lipoprotein, and long chain fatty acids (Leung et al., 1992; Endelmann et al., 1993). Human CD36 cDNA and genomic DNA have been cloned and characterized (Oquendo et al., 1989; Wyler et al., 1993). Immunoprecipitation of CD36 from COS cells showed the presence of an 83 kDa molecule, which was consistent with the glycosylated form found on the surface of C32 melanoma cells. Another CD36 cDNA have been isolated from HEL cells using degenerated oligonucleotides and the placental cDNA as hybridization probes (Taylor et al., 1993). All the cDNA from HEL cells and placenta has the same predicted amino acid sequence for CD36. However, the detection of a smaller partial cDNA from butt coat RNA pools of HEL cells using the polymerase chain reaction (PCR) suggests the existence of a CD36 variant in the cellular fraction, probably derived from monocytes (Tang et al., 1994).

This study describes cDNA cloning of PAS-4 from bovine mammary gland and CD36 from bovine liver. Moreover, the expression of PAS-4 in bovine internal organs, and the relationship between CD36 of bovine liver and PAS-4 of mammary epithelial cell are also described.

MATERIALS AND METHODS

Materials

Fresh normal milk was obtained from Holstein cows of the university herd and immediately cooled to 4°C. Bovine tissues were obtained from Holstein cows at the Meat Center of Nasu-branch (Japan). Trypsin (from bovine pancreas, EC 3.4.21.4), ampicillin, isopropyl-β-D(-)-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-β-D-galactoside, xylene cyanol, Ficoll 400, bovine serum albumin (BSA), polyvinylpyrolidine, phenol, Diethylpyrocarbonate (DEPC), RNase A and denatured salmon DNA were purchased from Sigma (U.S.A.). Agar

* Corresponding Author: Soo-Won Lee. Tel: +82-31-290-7805, Fax: +82-31-290-7815, E-mail : leesw1949@skku.ac.kr

[1] Department of applied Biochemistry, Faculty of Agriculture, Utsunomiya University, Utsunomiya, Tochigi-Ken 321, Japan.

Received July 5, 2001; Accepted October 10, 2001

noble, yeast extract and bacto trypton were purchased from Difco Lab. (U.S.A.) while cDNA cloning kit was purchased from Clontech Lab. Inc. (U.S.A.). Restriction enzymes were purchased from Bio Labs. (New England). Agarose, 1 kbp DNA ladder, Taq polymerase, TRlzol™ reagent and DH5α competent cells were purchased from Gibco BRIL (U.S.A.). Geneclean II was purchased from Bio 101 Inc. (U.S.A.), and *p*-formaldehyde was from Merck (U.S.A.). Oligo (*dT*) 30, T7 sequencing kit, DNA ligation kit Ver. II and acylamino acid-releasing enzyme (EC 3.4.19.1.) were purchased from Takara Shuzo (Japan). TA cloning™ kit was purchased from Invitrogen Co. (U.S.A. PVP-K90, ethidium bromide, chloroform, isopropyl alcohol, isoamyl alcohol, ethanol and Triton X-100 were purchased from Wako Chemical (Japan). MOPS was purchased from Dojindo Chemical Co. (Japan). HEL cells were obtained from the Riken Cell Bank (Japan). Hybond™-N^+ nylon membrane was purchased from Amersham Life Sci. (U.S.A). MicroSpin™ S-200 HR column (0.5×2.0 cm) was purchased from Pharmacia Biotech (Sweden). Thermal cycler was purchased from Nippon Genetics Co. (QTP-I, Japan). Mupid-2 electrophoresis apparatus was purchased from Cosmo-Bio and Capcell Pak C8 was from Shiseido (Japan). All other chemicals were of analytical grade.

Preparation of bovine PAS-4

Bovine MFGM was prepared from bovine milk by method 2 described by Kanno and Kim (Kanno and Kim, 1990). PAS-4 was purified by the method of Kanno et al. (1995).

Deblocking and sequencing of N-terminal amino acid

A glycoproteins of the N-acetylated protein were separated by SDS-PAGE and electroblotted onto the PVDF membrane. Deblocking with acylamino-acid releasing enzyme was performed as previously described (Hirano et al., 1992).

Internal peptide sequence of PAS-4

PAS-4 was digested for 6 h at 37°C with trypsin (1/100 of the substrate) in a 100 mM Tris-acetate buffer (pH 7.6) containing 5 M urea. The digested PAS-4 was dissolved in 0.1% trifluoroacetic acid (TFA) and submitted to an SP-8700 HPLC system (Spectra Physic, Sweden) using a reversed-phase column of Capcell Pak C8 (Shiseido, 5 µm, 4.6×250 mm). The peptide was eluted with a linear gradient of acetonitrile containing 0.1% TFA from 18 to 40% (1%/min) at a flow rate of 1 ml/min and 40°C and monitored at 230 nm. The eluted peptide was collected and subjected to protein sequencing analysis (Applied Biosystems 470A).

Preparation of RNA

Total RNA was extracted and purified by the guanidinium-phenol-chloroform method described by Chomczynski and Sacchi (1987). mRNA was prepared by oligo dT column chromatography for Northern blot analysis.

Amplification of 5'-frame PAS-4 cDNA

An 18 mer oligonucleotides (table 1, primer P4-5, sense) designed based on the N-terminal amino acid sequence of bovine PAS-4 and a 21-mer oligonucleotide (table 1, primer P4-3, antisense) based on the internal peptide sequence were synthesized and total RNA was extracted as described above. First-strand cDNA was synthesized by using a cDNA cloning. PCR was performed using 1 µl of cDNA pool, 100 pmol of sense and anti-sense primers, 0.5 mM dNTP, and 5 U of Taq DNA polymerase. Amplification was performed for 30 cycles, under cycle conditions of 94°C for 1 min, 52°C for 2 min, and 72°C for

Table 1. Sequences of oligonucleotide probe and their location used for RT-PCR and RACE-PCR

Primers	Sequences	Location[1]
P4-5 (sense)	5'-TGTAATCGGAACTGTGGG-3'	Nt 3-21
P4-3 (antisense)	5'-GAATGGATCGGTATAGCCCCA-3'	Nt 555-535
MA-1 (antisense)	5'-CAATAGTTCTTTCAAAGTTC-3'	Nt 534-515
MA-2 (sense)	5'-ACTGAGGATGACAAGTTCACC-3'	Nt 385-405
L-1 (sense)	5'-GGGTGCAATCGAAACTGTG-3'	Nt 1-19
L-2 (antisense)	5'-CCTTGCATAAATGTATTTG-3'	Nt 464-446
L-3 (antisense)	5'-GTAGTAATAGGATATGGAAC-3'	Nt 584-565
L-4 (sense)	5'-CAAATCTTCTATGTTTC-3'	Nt 492-508
L-5 (antisense)	5'-GAACATTTCCGCCTTCTC-3'	Nt 1289-1272
L-6 (antisense)	5'-ATTTACTCTCTTTGATCTG-3'	Nt 1423-1395
AP-1	5'-CCATCCTAATACGACTCACTATAGGGC-3'	
AP-2	5'-ACTCACTATAGGGCTCGAGCGGC-3'	

[1] Primer P4-5 and P4-3 are synthesized based on N-terminal and internal amino acid sequence of PAS-4, respectively, for RT-PCR of PAS-4. Oligonucleotides MA-1, MA-2 and L-1 to L-6 were synthesized according to the sequence showing figure 3. Nucleotide positions are from PAS-4 cDNA sequence in figure 3. Primers AP-1 and AP-2 were marathon adaptor primers.

3 min with a thermal cycler. After amplification, 3.5 μl of the PCR products were applied to agarose gel electrophoresis (1% gel). DNA was stained with ethidium bromide at 0.3 μg/ml of 0.5×TBE buffer.

5'-RACE and 3'-RACE PCR

Marathon cDNA adapter was ligated to the ds cDNA obtained after second-strand cDNA synthesis for 5'-RACE PCR. Diluted solution (1/50) of the above reaction mixture was added to a PCR mixture containing each oligonucleotide primer (MA-1) and the AP-1 adapter primer (table 1). PCR reaction was performed at 94°C for 30 s and at 68°C for 4 min for 30 cycles. Amplified products were separated by agarose gel electrophoresis (1% agarose), cut from the gel, purified with glass-milk, and ligated to the pCR^{TM} II vector, and the mixture was used to transform. Several colonies were randomly selected and sequenced. 3'-RACE PCR was performed with sense primer (MA-2) and AP-2 adapter primer (table 1) and the treatment of amplified products was performed as described for 5'-RACE PCR.

Sequencing of PCR products

The amplified PAS-4 cDNA fragments were ligated into the pCR^{TM} II vector and transformed into DH5α competent cells. After subcloning, sequencing reaction were performed with the dideoxy termination method (Sanger, 1977) using the T7 DNA polymerase according to manufacturer's protocol. The sequencing primers, P4-5, P4-3, MA-1, and MA-2, are shown in table 1.

Northern blot analysis

Twenty micrograms of total RNA extracted from mammary gland, liver, heart, spleen, lung, and HEL cells were electrophoresized on agarose gel (1%) containing formaldehyde (16.7%) and transferred to Hybond^{TM}-N^{+} nylon membranes. PAS-4 cDNA was labeled by [^{32}P]dCTP by the random primer method (Thomas, 1980). After hybridization, the membrane was washed and autoradiographed for 24 h at room temperature.

RT-PCR cloning of bovine CD36 from bovine liver tissues

Bovine CD36 cDNA was amplified with sense and antisense primer shown in table 1 and figure 3 (L-1 to L-6). Amplification was performed for 30 cycles with cycle condition of 94°C for 30 s, 52°C for 30 s, and 72°C for 1 min. Purification, ligation subcloning, and sequencing of amplified products were performed as above described for 5'-frame PAS-4 cDNA cloning.

Agarose gel electrophoresis

Agarose gel electrophoresis was performed by the method of Johnson and Grossman (1977) using Mupid-2 apparatus. The concentration of the agarose was 1% dependent on DNA size while electrophoresis was performed at 100 V with 0.5×Tris-borate EDTA buffer (45 mM Tris-borate, pH 8.0, 1 mM EDTA). The DNA was visualized with ethidium bromide (0.3 mg/ml) and the size was evaluated by using a 1 kilobase pair (kbp) DNA ladder.

RESULTS

N-terminal amino acid residue of PAS-4 was blocked and therefore, it was deblocked with acylamino acid releasing enzyme before sequencing. N-terminal amino acid sequence of PAS-4 was determined as AcX-Cys-Asn-Arg-Asn-Cys-Gly-, which was to be nearly identical with the published N-terminal sequence of human placental CD36 cDNA (Oquendo et al., 1989). Internal peptide sequence of PAS-4 was determined by digesting the PAS-4 with trypsin. Sequence of internal tryptic peptide of PAS-4 at 43.2 min eluted by reverse phase HPLC was determined as -Leu-Leu-Trp-Gly-Tyr-Thr-Asp-Pro-Phe-Leu- by sequence analysis (figure 1). Oligonucleotide primers for PCR of PAS-4 cDNA were synthesized based on the amino acid sequences of the N-terminal and internal tryptic peptides of PAS-4 (table 1, primer P4-5 and P3-3).

The PCR product of primer P4-5/P4-3 gave a single band of 555 base pair (bp) on 1% agarose gel electrophoresis (figure 2, lane 1). This PCR product consisted of 555 base pair was later found to be corresponding to PAS-4 cDNA in which its nt ranges from 181st to 736th as shown in figure 3.

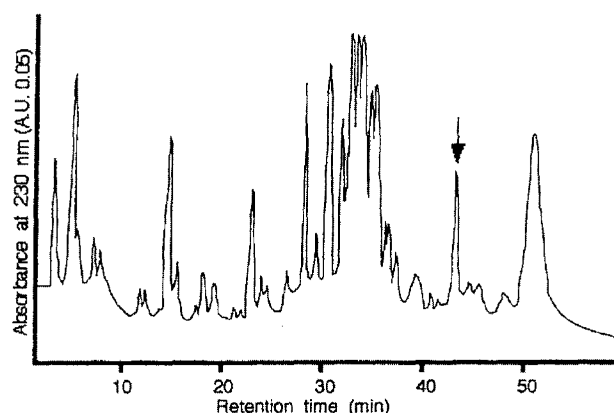

Figure 1. Analysis of tryptic peptides of the PAS-4 glycoprotein on reverse phase HPLC
PAS-4 was hydrolyzed with trypsin (1/100 of the substrate) for 6 h at 37°C and analyzed by HPLC on C8-reversed phase column. The peptide, which eluted at approximately 43.2 min (marked by the arrow) was applied to a protein sequencer.

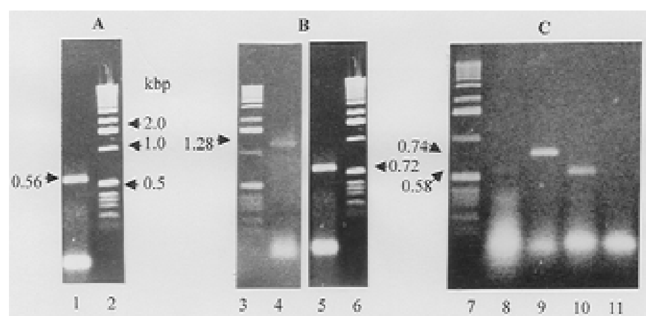

Figure 2. RT-PCR products for PAS-4 cDNA (A) and its RACE PCR products of bovine mammary gland (B), an RT-PCR products from CD36 cDNA of bovine liver (C) Lanes 2, 3, 6 and 7, ladder; 1, PCR product from primer P4-5/P4-3; 4, MA-2/AP-1; 5, MA-1/AP-2; 8, L-4/L-6; 9, L-4/L-5; 10, L-1/L-3; 11, L-1/L-2.
Products in lane 1 (PCR product), 4 (3´-RACE PCR product), 5 (5´-RACE PCR product), 8 to 11 (3´-frame and 5´frame PCR product of bovine liver CD36 cDNA) were sequenced. The concentration of agarose was 1.0%.

To obtain the full length of cDNA of PAS-4, 5’-and 3’-RACE PCR was done using MA-1 (antisense) and MA-2 (sense) PAS-4 specific primers, respectively. By 5’-RACE PCR a fragment of 720 bp was amplified (figure 2, lane 5) and 1.28 kbp fragment (figure 2, lane 4) by 3’-RACE PCR cDNA. Both fragments were subcloned and sequenced. The results are shown in figure 3. The full length ORF of PAS-4 cDNA was 1413 nt long from nt 178 to 1591. The overlapping sequences read forward from primer MA-2 and backward from primer MA-1 were completely identical. The deduced amino acid sequence consisted of 471 amino acid residues and its molecular weight was calculated to be the 52,796 by Genetyx-Mac (Ver. 7.3#1). The internal tryptic peptide sequence was found in the deduced amino acid sequence from nt 177th to 188th, and N-terminal amino acid was Gly (figure 3).

Partial bovine liver CD36 cDNA was prepared by RT-PCR cloning using combination of PAS-4 specific primers. Figure 2-C shows the amplified products of RT-PCR from total RNA of bovine liver. PCR products obtained by primer combinations L-1/L-3 and L-4/L-5 were calculated to be 0.58 kbp and 0.74 kbp (figure 2-C, lanes 10 and 9), respectively. However, primer combinations L-1/L-2, and L-4/L-6 failed to amplify (figure 2-C, lanes 8 and 11). The results shown by primers L-2 and L-6 were not acceptable for amplifying CD36 cDNA of liver, these primers being different sequence between PAS-4 and CD36. These RT-PCR products of about 0.58 and 0.74 kbp (figure 2, lanes 10 and 9) were purified and sequenced. Figure 4 shows the partial nucleotide sequence of the CD36 cDNA of bovine liver constructed from these two fragments which overlapped 92nt. All the three clones used for sequencing had the same nucleotide sequence, therefore, the possibility of PCR error was negative. The deduced amino acid sequence also is shown in figure 4.

Alignment of the amino acid sequences of PAS-4 of bovine mammary gland and CD36 of bovine liver is shown figure 5 in comparison with those of human placental CD36 (Taylor et al., 1993). PAS-4 and CD36 share 90.4% of the amino acids and most of the identical amino acid residues were localized in the luminal/extracellular domains. PAS-4 does 8 potential N-glycosylation sites, while liver CD36 displays 6 potential N-glycosylation sites, which were well conserved between them. Cysteine residues were strictly conserved between PAS-4 and CD36 (figure 5). All the nine cysteine residues of PAS-4 were at the same position as those of human placenta CD36 cDNA, and the one at 465th of human placenta CD36 cDNA exhibits a point mutation.

Northern blot analysis revealed the presence of a 2.9 kbp transcript in all internal organs examined (figure 6). Substantial variations of the mRNA expression were low amounts, whereas higher level expression was found in mammary gland, kidney liver, and HEL cells.

Hydropathy scale was inferred from the amino acid sequence of PAS-4 and compared with that of CD36 of bovine liver (Kyte and Doolittle, 1982). Hydropathy analysis suggested that the location of the strong hydrophobic region ranges from residues 7 to 28 located close to the amino-terminus and from residues 448 to 460 located near the carboxy-terminus (figure 7). The predicted extracellular domain is predominant in hydrophobic residues as well as the transmembrane domain.

DISCUSSION

In this report, we clarified the full length of cDNA sequence encoding of PAS-4 glycoprotein of bovine mammary gland. The human placental CD36 (Oquendo et al., 1989) and PAS-4 cDNA had 86% identity in nucleotide sequence and 82% identity in amino acid sequence. Cysteine residues of PAS-4 were located at the identical positions as human placental CD36, except for 465st amino acid residue (figure 5). The carboxyl-terminal half of the extracellular segment is rich in proline and contains 6 out of 9 cysteine residues. All of the cysteine residues in the extracellular positions are confined to a segment 242nd to 331st of amino acid residues. PAS-4 was contained 8 potential N-linked glycosylation sites at positions 78, 101, 171, 204, 234, 246, 320 and 416. However, human placenta CD36 cDNA displays 10 potential N-linked glycosylation sites, which are well conserved in PAS-4 and CD36 except for those positions 132, 165, 171 and 219 (Oquendo et al., 1989). PAS-4 of bovine mammary gland and CD36 of human platelet have different molecular masses of 78 and

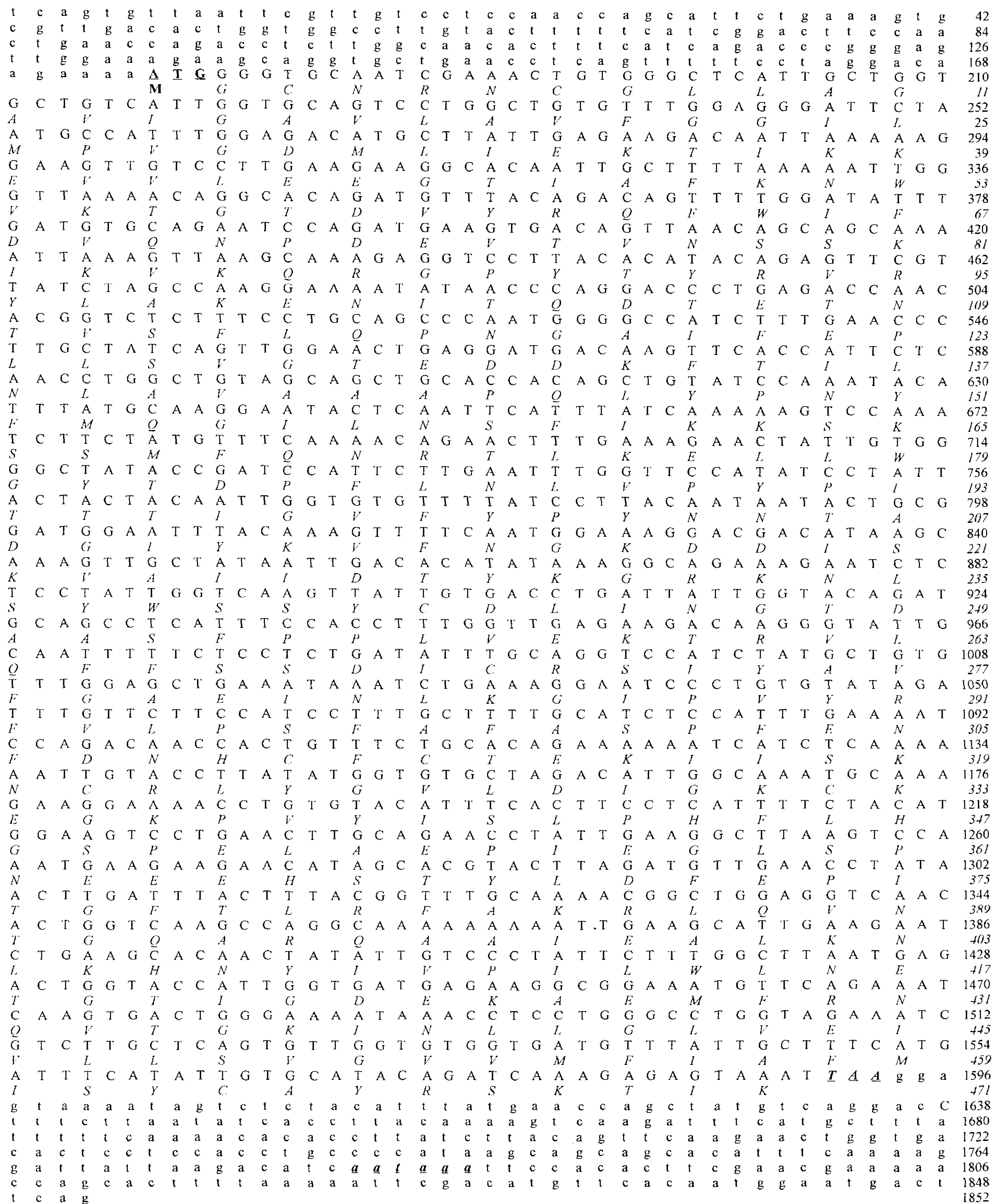

Figure 3. Nucleotide sequence of PAS-4 cDNA of bovine mammary gland and its deduced amino acid sequence
Nucleitides are numbered at the right of each line, starting with the first base following the start cordon (bold type underlines). The deduced amino acid sequence is shown below the DNA sequence with the amino acid numbered at the right of each lines. Nucleotide sequence, -aataaa-, was potential polyadenylation site (italic type underlines).

```
G G C T G T A A T C G G A A C T G T T G G T T C G T C A C T G G T T C T G T C A T T    42
G     C     N     R     N     C     G     L     V     T     G     A     V     I        14
G G T G C A G T T G T G G C T C T G T T C G G A G G G A T T C T A A T G C C A G T T    84
G     A     V     V     A     L     F     G     G     I     L     M     P     V        28
G G A G A C A T G C T T A T T G T G A A G A C A A T T A A A A A G G A A T C G G T C   126
G     D     M     L     I     V     K     T     I     K     K     E     S     V        42
C T T G A A G A A G G T A C A A T T G C T T T T A A A A A T T G G G T T A A A G C A   168
L     E     E     G     T     I     A     F     K     T     W     V     K     A        56
G G T G C T A A C G T T T A C A G A C A A T T T T G G A T A T T T G A T T T G C A G   210
G     A     N     V     Y     R     Q     F     W     I     F     D     V     Q        70
A A T C C A G A G G A A G T G T C A G T C A A T A T C A G C A A A A T T A A A G T T   252
N     P     E     E     V     A     V     N     S     S     K     I     K     V        84
A A G C A A A G A G G T C C T T A C A C G T A C A G A G T C C G T T A T C T A G C T   294
K     Q     R     G     P     Y     T     Y     R     V     R     Y     L     A        98
A A G G A T A T A A T A A T C C A A G A C C T C A A G A C C C A C A C A G T C T C T   336
K     D     I     I     S     Q     D     S     K     T     H     T     V     S       112
T T C C A A C G G C C C A A A G G C G C C A T C T T T G A A T C C T C G C T A T C A   378
F     Q     R     P     K     G     A     I     F     E     P     S     L     S       126
G T T G G A A C C A G A G A T G A C A C A T T C A C T G T T C T C A A C C T G G C T   420
V     G     T     R     D     D     T     F     T     V     L     N     L     A       140
G T A G A C G C T A C A C C A C C T C T C T A T C C A A G C A C A T T T T T T C A G   462
V     A     A     T     P     P     L     Y     P     S     T     F     V     Q       154
A T A A T A C T C A A T G T A T T T A T C A A A A A A C T C A A C T C T T T T A T G   504
I     I     L     N     V     F     I     K     K     L     N     S     F     M       168
T T T C A A A A A A G A A C T C T G A A A G A A C T A T T G T G G G G C T A T A C G   546
F     Q     N     R     T     L     K     E     L     L     W     G     Y     T       182
G A T C C A T T C T T G A A T T T G G T T C C A T A T C C T A T T A C T A C T A C A   588
D     P     F     L     N     L     V     P     Y     P     I     T     T     T       196
A T T G G T G T G T T T T A T C C T T A C A A T A A T A C T G C G G A T G G A A T T   630
I     G     V     F     Y     P     Y     N     N     T     A     D     G     I       210
T A C A A A G T T T T C A A T G G A A A G G A C G A C A T A A G C A A A G T T G C T   672
Y     K     V     F     N     G     K     D     D     I     S     K     V     A       224
A T A A T T G A C A C A T A G A A A G G C A G A A A G A A T C T C T C C T A T T G G   714
I     I     D     T     Y     K     G     R     K     N     L     S     Y     W       238
T C A A G T T A T T G T G A C C T G A T T A A T G G T A C A G A T G C A G C C T C A   756
S     S     Y     C     D     L     I     N     G     T     D     A     A     S       252
T T T C C A C C T T T G G T T G A G A A G A C A A G G T T A T T G C A A T T T T T C   798
F     P     P     L     V     E     K     T     R     V     L     Q     F     F       266
T T C T C C T C T G A T A T T T G C A G G T C C A T C T A T G C T G T G T T T G G A   840
S     S     D     I     C     R     S     I     Y     A     V     F     G     A       280
G C T G A A A T A A A T C T G A A A G G A A T C C C T G T G T A T T T T G T T C T T   882
E     I     N     L     K     G     I     P     V     Y     R     F     V     L       294
C C A T C C T T T G C T T T T G C A T C T C C A T T T G A A A A T C C A G A C A A C   924
P     S     F     A     F     A     S     P     F     E     N     P     D     N       308
C A C T G T T T C T G C A C A G A A A A A A T C A T C T C A A A A A A T T G T A C C   966
H     C     F     C     T     E     K     I     I     S     K     N     C     T       322
T T A T A T G G T G T G C T A G A C A T T G G C A A A T G C A A A G A A G G A A A A  1008
L     Y     G     V     L     D     I     G     K     C     K     E     G     K       336
C C T G T G T A C A T T T C A C T T C C T C A T T T T C T A C A T G G A A G T C C T  1050
P     V     Y     I     S     L     P     H     F     L     H     G     S     P       350
G A A C T T G C A G A A C C T A T T G A A G G C T T A A G T C C A A A T G A A G A A  1092
E     L     A     E     P     I     E     G     L     S     P     N     E     E       364
G A A C A T A A C A C G T A C T T A G A T G T T G A A C C T A T A A C T G G A T T T  1134
E     H     N     T     Y     L     D     V     E     P     I     T     G     F       378
A C T T T A C G G T T T G C A A A A C G G C T G C A G G T T A A C A C T G G T C A A  1176
T     L     R     F     A     K     R     L     Q     V     N     T     G     Q       392
G C C A G G C A A A A A A A A A T T G A A G C A T T G A A G A A T C T G A A G C A C  1218
A     R     Q     K     K     I     E     A     L     K     N     L     K     H       406
A A C T A T A T T T T C C C T A T T C T T T G G C T T A A T G A G A C T G G T A C C  1260
N     Y     I     V     P     I     L     W     L     N     E     T     G     T       420
A T T G G T G A T G A G A A G G C A A A C A T G T T C                                1287
I     G     D     E     K     A     N     M     F                                     429
```

Figure. 4. Nucleotide sequences CD36 cDNA of bovine liver and its deduced amino acid
Nucleotides are numbered at the right of each lines. The deduced amino acid sequence is shown below the DNA sequence with the amino acid numbered at the right of each lines.

85 kDa, respectively, both are reduced to 57 kDa after deglycosylation with endoglycosidase F (Greenwalt et al., 1990). This suggested that cell type specific processing gives rise to different glycosylated forms of these glycoproteins.

Partial CD36 cDNA sequence of bovine liver was determined by RT-PCR. The homology search result of nucleotide sequences revealed 94% identity (90% for deduced amino acid sequences) to that of the bovine PAS-4 cDNA. The identity between the two proteins were confirmed throughout the whole molecule, however, primer combinations L-1/L-2, and L-4/L-6, on the basis of the bovine PAS-4 cDNA, failed to be amplified (figure 2-C, lanes 8 and 11). The results showed the different sequence between PAS-4 and CD36 cDNA. PAS-4 of bovine mammary gland is localized at the apical surface of luminal epithelial cells, while the basement membrane is completely negative (Pechoux et al., 1994). Therefore, monoclonal

```
PAS-4  G C N R N C G L L A G A V I G A V L A V F G G I L M P V G D M L I E K T I K K E V V   42
LCD36  G C N R N C G L V T G A V I G A V V A L F G G I L M P V G D M L I V K T I K K E S V
HCD36  G C D R N C G L I A G A V I G A V L A V F G G I L M P V G D L L I Q K T I K K Q V V

       L E E G T I A F K N W V K T G T D V Y R Q F W I F D V Q N P D E V T V N S S K I K V   84
       L E E G T I A F K N W V K A G A N V Y R Q F W I F D V Q N P E E V A V N S S K I K V
       L E E G T I A F K N W V K T G T E V Y R Q F W I F D V Q N P Q E V M N N S S N I Q V

       K Q R G P Y T Y R V R Y L A K E N I T Q D T E T N T V S F L Q P N G A I F E P L L S   126
       K Q R G P Y T Y R V R Y L A K D I I S Q D S K T H T V S F Q R P K G A I F E P S L S
       K Q R G P Y T Y R V R F L A K E M V T Q D A E D N T V S F L Q P N G A I F E P S L S

       V G T E D D K F T I L N L A V A A A P Q L Y P N Y F M Q G I L N S F I K K S K S S M   168
       V G T R D D T F T V L N L A V A A T P P L Y P S T F V Q I I L N V F I K K L N S F M
       V G T E A D N E I V L N L A V A A A S H I Y Q N Q F V Q M I L N S L I N K S K S S M

       F Q N R T L K E L L W G Y T D P F L N L V P Y P I T T T I G V F Y P Y N N I A D G I   210
       F Q K R T L K E L L W G Y T D P F L N L V P Y P I T T T I G V F Y P Y N N T A D G I
       F Q V R T L R E L L W G Y R D P F L S L V P Y P V T T T V G L F Y P Y N N T A D G V

       Y K V F N G K D D I S K V A I I D T Y K G R K N L S Y W S S Y C D L I N G T D A A S   252
       Y K V F N G K D D I S K V A I I D T Y K G R K N L S Y W S S Y C D L I N G T D A A S
       Y K V F N G K D N I S K V A I I D T Y K G K R N L S Y W E S H C D M I N G T A A A S

       F P P L V E K T R V L Q F F S S D I C R S I Y A V F G A E I N L K G I P V Y R F V L   294
       F P P L V E K T R V L Q F F S S D I C R S I Y A V F G A E I N L K G I P V Y R F V L
       F P P F V E K S Q V L Q F F S S D I C R S I Y A V F E S D V N L K G I P V Y R F V L

       P S F A F A S P F E N F D N H C F C T E K I I S K N C T L Y G V L D I G K C K E G K   336
       P S F A F A S P F E N P D N H C F C T E K I I S K N C T L Y G V L D I G K C K E G K
       P S K A F A S P V E N P D N Y C F C T E K I I S K N C T S Y G V L D I S K C K E G R

       P V Y I S L P H F L H G S P E L A E P I E S L S P N E E E H S T Y L D F E P I T G F   378
       P V Y I S L P H F L H G S P E L A E P I E G L S P N E E E H N T Y L D V E P I T G F
       P V Y I S L P H F L Y A S P D V S E P I D G L N P N E E E H R T Y L D I E P I T G F

       T L R F A K R L Q V N T G Q A R Q K K I E A L K N L K H N Y I V P I L W L N E T G T   420
       T L R F A K R L Q V N T G Q A R Q K K I E A L K N L K H N Y I V P I L W L N E T G T
       T L Q F A K R L Q V N L L V K P S E K I Q V L K N L K R N Y I V P I L W L N E T G T

       I G D E K A E M F R N Q V T G K I N L L G L V E I V L L S V G V V M F I A F M I S Y   462
       I G D E K A N M F                                                                   429
       I G D E K A N M F R S Q V T G K I N L L G L I E M I L L S V G V V M F V A F M I S Y

       C A Y R S K T I K                                                                   471
       C A C R S K T I K
```

Figure 5. Homology of amino acid sequences among PAS-4 from bovine mammary gland (PAS-4) and CD36 from bovine liver (L CD36) and human placenta (H CD36)
Amino acids are numbered at the right of each line. The strongly mutation region between bovine PAS-4 and human CD36 shown by the open box.
Underlines indicate the potential sites for asparagine-linked glycosylation.
Data for human placenta CD36 cDNA are cites from Taylor et al.

antibody (MABs) OKM5, OKM8, and NLO7 only recognized human CD36-specific epitopes and did not bind bovine PAS-4 (Okenhouse et al., 1993; Alessio et al., 1991). The adhesion of parasitized erythrocytes binding to bovine PAS-4 appeared to be independent of the epitopes recognized by MAbs OKM5 and OKM8 since these MAbs

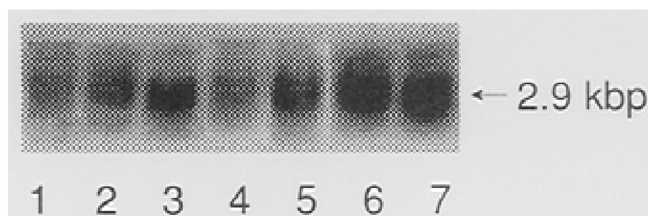

Figure 6. Northern blot analysis of various bovine tissues and HEL cell 20 μg of total RNA separated from tissues and cell were hybridized with the PAS-4 cDNA. Lanes 1, spleen; 2, lung; 3, kidney; 4, heart; 5, liver; 6, HEL ell; 7, mammary gland.

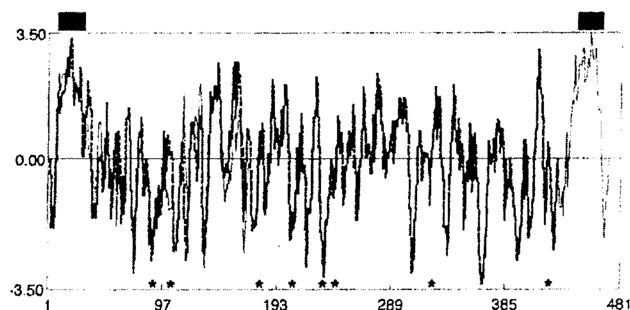

Figure 7. Hydropathy plot of bovine PAS-4 and CD-36
The hydropathy of 5 consecutive amino acid was calculated using the algorithm of Kyte and Doolittle (1982). Black boxes indicate the position of potential transmembrane segments. A solid line indicates the PAS-4 from bovine mammary gland and a dotted line indicates CD 36 from bovine liver.
* potential asparagine-linked glycosylation sites of PAS-4.

do not recognized bovine PAS-4 by enzyme-linked immunosorbent or radioimmunoassay (Okenhouse et al., 1993).

This difference between mammary gland and liver may be due to the organ specific functions. It is remarkable that changes in the expression and processing of surface proteins related to mammary gland and other organs of bovine. In other words, it's described that these proteins are homologous components, but organ specific mutation or different components, respectively.

Hydropathy scale followed by Kyte and Dolittle index (Kyte and Doolittle, 1982) was inferred from the amino acid sequence of PAS-4 from mammary gland and compared with that of CD36 of bovine liver. Hydropathy analysis suggested that the location of the strong hydrophobic region ranges from residues 7 to 28 located close to the amino-terminus and residues from 448 to 460 located near the carboxy-terminus (figure 7). The predicted extracellular domain is predominant in hydrophobic residues as well as the transmembrane domain. The carboxyl-terminal half of the extracellular segment of PAS-4 contains 6 of the 9 cysteines. The other three cysteines are located within the proposed intracelluar segments found at the amino (at 2nd and 6th residues) and carboxyl terminal (at 463rd residues). The hydrophobicity index of PAS-4 was estimated to be 1,226 cal/residues, which comparable favorably with other well-characterized integral membrane glycoproteins, 1,120 and 1,185 for bovine rhodopsin and bacteriorhodopsin, respectively (Capaldi, 1982). This index, which is obtained from the molar free energies for the transfer of individual amino acids from aqueous solution to organic solvent, is correlated with the overall hydrophobic character of proteins (Vega et al., 1991). The extramembranous segment of PAS-4 exhibiting multiple stretches of hydrophobic amino acids, may be embedded in it or form small hydrophobic pockets as it seems to be too short to span the membrane. Hydropathy analysis revealed other hydrophobic regions 111-127, 134-147, 185-200, 287-300, and 448-460 (figure 7). In the case of human CD36, the region 183-203 is thought to be membrane-associated (Greenwalt et al., 1992).

The distinct hydrophobic region, residues 448 to 460 that follows the luminal domain, is a possible candidate to act as a stop transfer signal and to anchor the protein to the membrane. The broad tissue distribution of PAS-4 and homologous protein CD36 suggests that these proteins might be a housekeeping glycoprotein (Vega et al., 1991; Okenhouse et al., 1989). At present it is difficult to speculate the function of PAS-4, but one could guess on the basis of its structural homology to CD36 that PAS-4 could act as an extracellular receptor. Furthermore, the research for the promoter region of PAS-4 and CD36 from bovine and the functions of these proteins function are under investigation.

ACKNOWLEDGEMENTS

This work in part was supported by Korea Research Foundation Grant (KRF-99-005-G00004). The technical assistance of Dr. N. Azuma and M. Ametani is gratefully acknowleged.

REFERENCES

Alessio, M., D. Ghigo, G. Garbaringo, M. Geuna and F. Malavasi. 1991. Analysis of the human CD36 leucocyte differentiation antigen by means of the monoclonal antibody. Cellular Immunol. 137:487.

Capaldi, R. A. 1982. Structure of intrinsic membrane proteins. Trends Biochem. Sci. 7:292.

Chomczynski, P. and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidium thiocynate-PhOH-chloroform extraction. Anal. Biochem. 162:156.

Clemetson, K. J. and J. L. McGregor. 1987. In: Platelets in

Biology and Pathology (Ed. E. E. Maclntyre and J. L. Gordon). Vol. III, Academic Press, New York, pp. 160-212.

Endelmann, G., L. W. Stanton, K. S. Madden, C. M. Bryant, R. T. White and A. A. Protter. 1993. CD36 is a receptor for oxiized low density lipoprotein. J. Biol. Chem. 268:1811.

Freudenstein, C., T. W. Kennan, W. N. Eigel, M. Sasaki, J. Stadler and W. W. Franke. 1979. Preparation and characterization of the inner coat material associated with fat globule membranes from bovine and human milk. Exp. Cell Res. 118:277.

Greenwalt, D. E., V. G. Johnson and I. H. Mather. 1985. Specific antibodies to PAS-IV, a glycoprotein of bovine milk-fat-globule membrane, bin to a similar protein in cardiac endothelial cells and epithelial cells of lung bronchioles. Biochem. J. 228:233.

Greenwalt, D. E. and I. H. Mather. 1985. Characterization of an apically derived epithelial membrane glycoprotein from bovine milk, which is expressed in capillary endothelial in diverse tisues. J. Cell Biol. 100:397.

Greenwalt, D. E., R. H. Lipsky, C. F. Ockenhouse, H. Ikeda, N. N. Tandon and G. A. Jamieson. 1992. Membrane glycoprotein CD36: a review of its roles in adherence, signal transduction, and transfusion medicine. Blood 80:1105.

Greenwalt, D. E., K. W. K. Watt, O. T. So and N. Jiwani. 1990. PAS-4, an integral membrane protein of mammary epithelial cells, in related to platelet and endothelial cell CD36 (GP IV). Biochemistry 29:7054.

Hirano, H., S. Komatsu, H. Takakura, F. Sakiyama and S. Tsunasawa. 1992. Deblocking and subsequent microsequence analysis of N^a-blocked proteins electroblotted onto PVDF membrane. J. Biochem. 111:754.

Joheson, P. H. and L. I. Grossman. 1977. Electrophoresis of DNA in agarose gels. Optimizing separations of conformational isomers of double- and single-stranded DNAs. Biochmistry 16:4217.

Kanno, C., H. Hattori and K. Yamauchi. 1982. Isolation and characterization of plasma membrane from lactating bovine mammary gland. Biochem. Biophys. Acta 689:121.

Kanno, C. and D. H. Kim. 1990. A simple procedure for the preparation of bovine milk fat globule membrane and a comparison of its composition, enzymatic activities, and electrophoresis with those prepared by other methods. Agric. Biol. Chem. 54:2845.

Kanno, C., S. Hwangbo and N. Azuma. 1995. Rapid and simple procedure for purifying PAS-4 glycoprotein from bovine milk fat globule membrane. Biosci. Biotech. Biochem. 59:848.

Kyte, J. and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105.

Leung, L. L., W. Li, J. L. McGregor, G. Albrecht and R. J. Howard. 1992, CD36 peptides enhance or inhibit CD36-thrombospondin binding. J. Biol. Chem. 267:18244.

Okenhouse, C. F., C. Magowan and J. D. Chulay. 1989. Activation of monocytes and platelets by monoclonal antibodies or malaris-infected erythrocytes binding to the CD36 surface receptor *in vitro*. J. Clin. Invest. 84:468.

Okenhouse, C. F., N. N. Tandon, G. A. Jamieson and D. E, Greenwalt. 1993. Antigenic and functional differences in adhesionof plasmodium falciparum-infected erythrocytes to human and bovine CD36. Infec. Immun. 61:2229.

Oquendo, P., E. Hundt, J. Lawler and B. Seed. 1989. CD36 directly mediates cytoadherence of plasmodium falciparum parasitized rythrocytes. Cell 58:95.

Pechoux, C., P. Clezardin, R. Dante, C. M. Serre, M. Clerget, N. Bertin, J. Lawler, P. D. Delmas, J. L. Vauzelle and L. Frappart. 1994. Localization of thrombospondin, CD36 and CD51 during prenatan development of the human mammary gland. Differentiation 57:133.

Sanger, F., S. Nicken and A. R. Coulson. 1977. DNA sequencing with chain terminating. Pros. Natl. Acad. Sci. USA. 74:5463.

Tang, Y., K. T. Tayler, D. A. Sorbieski, E. S. Medved and R. H. Lipsky. 1994. Identification of a human CD36 isoform produced by exon skipping conservation of exon organization and pre-mRNA splicing patterns with a CD36 gene family member, CLA-1. J. Biol. Chem. 269:6011.

Taylor, K. T., Y. Tang, D. A. Sobieski and R. H. Lipsky. 1993. Characterization of two alternatively spliced 5'-untranslated exons of the human CD36 gene in different cell types. Gene (Amst.). 133:205.

Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA. 77, 5201-5205.

Vega, M. A., B. Segui-Real, J. A. Garcia, C. Cales, F. Rodriguez, J. Vaderkerckhove and I. V. Sandoval. 1991. Cloning, sequencing, and expression of a cDNA encoding rat LIMP II, a novel 74-kDa lysosomal membran protein related to the surface adhesion protein CD36. J. Biol. Chem. 166:16818.

Wyler, B., L. Daviet, H. Bortkiewicz, J. C. Bordet and J. L. McGregor. 1993. Cloning of the cDNA encoding human platelet CD36: comparison to PCR amplified fragmens of monocyte, endotherial and HEL cells. Thromb. Haemastasis. 70:500.