PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/22010

Please be advised that this information was generated on 2018-11-16 and may be subject to change.
Expression Cloning and Chromosomai Mapping of the Leukocyte Activation Antigen CD97, a New Seven-Span Transmembrane Molecule of the Secretin Receptor Superfamily with an Unusual Extracellular Domain

Jörg Hamann,2*+ Wolfram Eichler, Dörte Hamann, Harold M. J. Kerstens, Pino J. Poddighe, Jan M. N. Hoovers, Enno Hartmann, Michael Strauss, and René A. W. van Lier*

*Central Laboratory of The Netherlands Red Cross Blood Transfusion Service and Laboratory for Experimental and Clinical Immunology, University of Amsterdam, The Netherlands; †Max Planck Society, Berlin-Buch, Germany; ‡Institute of Zoology, University of Leipzig, Germany; §Institute of Pathology, Academic Hospital Nijmegen, The Netherlands; ‡Institute of Anthropogenetics, Academic Medical Center, University of Amsterdam, The Netherlands; and ^Max Delbrück Center for Molecular Medicine, Berlin-Buch, Germany

CD97 is a monomeric glycoprotein of 75 to 85 kDa that is induced rapidly on the surface of most leukocytes upon activation. We herein report the isolation of a cDNA encoding human CD97 by expression cloning in COS cells. The 3-kb cDNA clone encodes a mature polypeptide chain of 722 amino acids with a predicted molecular mass of 79 kDa. Within the C-terminal part of the protein, a region with seven hydrophobic segments was identified, suggesting that CD97 is a seven-span transmembrane molecule. Sequence comparison indicates that CD97 is the first leukocyte Ag in a recently described superfamily that includes the receptors for secretin, calcitonin, and other mammalian and insect peptide hormones. Different from these receptors, CD97 has an extended extracellular region of 433 amino acids that possesses three N-terminal epidermal growth factor-like domains, two of them with a calcium-binding site, and a single Arg-Gly-Asp (RGD) motif. The existence of structural elements characteristic for extracellular matrix proteins in a seven-span transmembrane molecule makes CD97 a receptor potentially involved in both adhesion and signaling processes early after leukocyte activation. The gene encoding CD97 is localized on chromosome 19 (19p13.12–13.2). The Journal of Immunology, 1995, 155: 1942–1950.

Leukocyte activation evokes coordinated changes in the expression of specific membrane molecules that allow cells to respond to environmental stimuli in an appropriate manner. Membrane receptors are involved in distinct processes such as adhesion, migration, proliferation, differentiation, and effector functions (1, 2). Expression levels of these surface molecules are regulated via several mechanisms. First, lymphocyte activation is accompanied by the transcriptional activation of several genes (3). One of the earliest induced genes encoding membrane molecules is the CD69 gene, which encodes a type II integral membrane protein with a C-type lectin domain (4, 5). CD69 mRNA can be detected already within 1 h after lymphocyte stimulation, whereas CD69 protein can be demonstrated at the cell surface within 2 to 4 h. Second, cellular activation initiates post-translational protein modifications that affect distribution and/or conformation of surface receptors. In this respect, mechanisms have been documented, such as fusion of intracellular granules with the cell membrane and the induction of neoepitopes on receptors that are already expressed on resting cells (6–8).

Recently, we have described a new activation Ag that was clustered during the Fifth International Workshop on Leukocyte Differentiation Antigens as CD97 (9, 10). Four mAbs, BL-Ac/F2, VIM3, VIM3B, and VIM3C, defined...
CD97 as a single-chain glycoprotein of 75 to 85 kDa. Whereas granulocytes and monocytes constitutively express CD97, resting lymphocytes bear the Ag only at low density. Cellular activation up-regulates surface expression of CD97 on B and T cells within 1 to 4 h. Markedly, these rapid kinetics even exceed those of the early activation Ag CD69. To begin to understand the function of CD97, we have isolated a cDNA encoding human CD97 by expression cloning. We describe herein the molecular structure of the Ag and discuss its possible function. Furthermore, the chromosomal localization of the CD97 gene is shown.

Materials and Methods

cDNA cloning and sequencing of CD97

The cDNA coding for human CD97 was isolated by expression cloning (11, 12), as previously described (4). In brief, a cDNA library from human PBMC, stimulated for 6 h with PMA, was transfected into COS cells by the DEAE-dextran method. Cells expressing CD97 were separated by immunomagnetic selection with the mAb BL-Ac/F2 (9). Plasmid DNA was extracted from the sorted cells and used to transfected fresh COS cells by spheroplast fusion, followed by immunomagnetic selection and plasmid recovery. After one additional round of screening, plasmid DNA prepared from single bacterial colonies was transfected individually into COS cells and analyzed for expression of CD97 by immunoperoxidase staining.

An isolated cDNA clone missing a large part of the 3' sequence was used to screen the enriched cDNA library (15). Therefore, plasmid DNA from COS cells that had been processed through one round of screening was transformed into Escherichia coli. BAR5 microcultures filters (Schleicher & Schuell, Inc., Dassel, Germany) were lifted from 2000 colonies and hybridized with the 32P-labeled incomplete clone for 20 h at 37°C in a buffer consisting of 6X SSC, 5X Denhardt's solution, 0.1% SDS, 10 mM EDTA, and 0.1 mg/ml herring sperm DNA. The filters were washed with 2X SSC/0.1% SDS at 50°C and exposed to autoradiography.

Sequencing was performed on both strands with the fmol DNA-sequencing system (Promega Corp., Madison, WI) and a series of primers differing in 20 nucleotides. Sequencing was performed on both strands with the fmol DNA-sequencing system (Promega Corp., Madison, WI) and a series of primers differing in 20 nucleotides.

RNA blot analysis

Total RNA was prepared from human PBMC, either untreated or stimulated with 1 µg/ml PHA at 1, 3, 8, and 24 h before isolation, with RNaseOut (Cinna/Biotex, Friendswood, TX). For each sample, 10 µg RNA was separated electrophoretically on a 1.2% agarose/formaldehyde gel, transferred to a GeneScreen Plus nylon membrane (DuPont, Boston, MA), UV cross-linked, and prehybridized with 1.5X SSPE, 1% SDS, 0.5% nonfat dry milk, 0.5 mg/ml herring sperm DNA, and 10% dextran sulfate. Hybridization was performed in the same buffer with 32P-labeled CD97 cDNA insert at 60°C for 20 h. Before autoradiography, the blot was washed with 0.5X SSC/0.1% SDS at 50°C. Densitometry was performed after exposure of the blot on a PhosphorImager BAS 2000 (Fuji, Tokyo, Japan). In parallel, surface expression of CD97 was analyzed at the indicated time points with biotinylated mAb BL-Ac/F2 (9) and streptavidin-phycoerythrin on a FACScan (Becton Dickinson, Mountain View, CA).

Chromosomal mapping of the CD97 gene

The CD97 gene was localized by fluorescence in situ hybridization (FISH). Metaphase spreads from human PBL were denatured in 70% formamide at 70°C for 3 min, permeabilized with 100 µg/ml proteinase K for 15 min at 37°C, and subsequently dehydrated in an ascending alcohol series. Chromosomes were hybridized with CD97 cDNA labeled by nick translation with biotin 14-dATP (Life Technologies, Inc., Gaithersburg, MD) at a concentration of 2 ng/µl in hybridization solution containing 60% formamide, 10% dextran sulfate, 2X SSC, 50 ng/µl herring sperm DNA, and a 50-fold excess Cot-1 DNA. The probe was de­natured at 95°C for 5 min and pre-annealed with Cot-1 DNA at 37°C for 30 min. Hybridization was performed in 10 µl hybridization solution at 37°C overnight, followed by posthybridization washes with 60% form­amide at 42°C.

Demonstration of the probe was conducted by using a recently described amplified method for in situ hybridization signals, on the basis of the deposition of biotinylated tyramine (14). All incubations were in PBS/0.05% Tween 20 containing 1% nonfat dry milk at 37°C for 20 min. After pre-incubation, biotinylated CD97 cDNA was detected by using subsequently mouse anti-biotin Ig (1/1000; Dakopatts, Glostrup, Denmark), biotinylated horseradish peroxidase antibody (1/100; Vector Laboratories, Burlingame, CA), avidin-biotin-horseradish peroxidase complex (1/100; Vector Laboratories), biotinylated tyramine/H2O2 (0.7 µmol/0.1% in PBS, 5 min), and fluorescein-labeled avidin (1/1000 in 4X SSC; Vector Laboratories). The slides were counterstained with 4,6-diamidino-2-phenylindole (DAPI) and analyzed on a Leitz DMRB/F fluorescence microscope with filterblocks A (DAPI) and I (fluorescein). Kodak EPL 400 films were used for photography. The map position of CD97 was determined by measuring the relative position of the probe on the chromosome expressed as the fractional length of the whole chromosome relative to pter (FL-pter).

Results

Isolation and expression of a cDNA encoding human CD97

Expression cloning in COS cells was used to isolate a cDNA encoding the human CD97 Ag. A cDNA library from PMA-activated PBMC (4) was expressed transiently in COS cells, and after 3 days, cells expressing CD97 were enriched by immunomagnetic selection with the mAb BL-Ac/F2 (9). Plasmid DNA was extracted from the sorted cells, and the cloned and expression and immunomagnetic selection was repeated twice. The following analysis of

Abbreviations used in this paper: FISH, fluorescence in situ hybridization; ECF, epidermal growth factor; FL-pter, fractional length-pter; RGD, Arg-Gly-Asp.
CD97 is a seven-span transmembrane glycoprotein related to the secretin receptor superfamily

The nucleotide sequence of the CD97 cDNA, with a length of 2921 bp, is shown in Figure 2. The first seven nucleotides have been supplemented from the 5′ end of the incomplete 573-bp clone, whose sequence was found to be identical with the corresponding region of the 3-kb clone. The CD97 cDNA contains a single open reading frame of 2226 bp encoding a polypeptide chain of 742 amino acids. The open reading frame is preceded by a 5′ untranslated region of 70 bp and followed by a 3′ untranslated region of 625 bp, with a polyadenylation signal (ATTAAA) 13 bp upstream of the poly(A) tail. The N-terminal methionine is followed by a hydrophobic signal sequence. A possible cleavage site behind Thr^{20} (15) gives evidence that the mature polypeptide probably consists of 722 amino acids, with a predicted molecular mass of 79 kDa. However, in vitro translation of the cDNA revealed a product of approximately 63 kDa, corresponding in size with the nonglycosylated CD97 that was immunoprecipitated from transfected COS cells (Fig. 1, B and C).

The discrepancy between the predicted molecular mass and the biochemical findings raises the question on the structure of the CD97 Ag. A hydropathy plot (16) identified seven hydrophobic segments within the C-terminal part of the molecule (Fig. 3A). This suggests that CD97 is a multispanning membrane protein, with an extended extracellular part of 433 amino acids, a 243-amino-acid region that passes the membrane seven times, and a cytoplasmic part of 46 amino acids (Fig. 3B). The tendency to migrate faster than predicted on the basis of m.w. is a well-known feature of many multispanning molecules (17).

The seven-span transmembrane region shows no homology with the corresponding region from classical protein G-coupled receptors (18). However, a significant homology was found with the membrane-spanning region of members from the secretin receptor group (Fig. 4). This recently described superfamily includes peptide receptors with specificity for secretin (19), calcitonin (20), parathyroid hormone (21), vasoactive intestinal peptide (22), glucagon-like peptide 1 (23), growth hormone-releasing hormone (24), glucagon (25), pituitary adenylate cyclase-activating polypeptide (26), corticotropin-releasing factor (27), and insect diuretic hormone (28). The sequence identity of 23 to 25% between CD97 and the receptors aligned in Figure 4 is restricted largely to the seven-span transmembrane region and involves numerous hydrophilic amino acid residues. Four highly conserved cysteines preceding the first transmembrane domain of the related proteins do not exist within the deduced amino acid sequence of CD97. Only the extracellular part of the diuretic hormone receptor from Acheta domesticus (GenBank accession number U15959) is weakly homologous with the corresponding region of CD97.
mucosa, an expressed sequence tag has recently been found (GenBank accession number D25916) that is identical with nucleotide 2412–2742 from the 3' noncoding region. This cDNA could possibly originate from mucosa- localized lymphocytes.
The extracellular region of CD97 contains three epidermal growth factor-like domains and a single Arg-Gly-Asp motif

The extracellular region of CD97 is approximately 300 amino acids longer than those from other known members of the secretin receptor superfamily. By analyzing the deduced amino acid sequence, three N-terminal EGF-like domains were identified (Figs. 2 and 3B). Remarkably, the sequence DXD/NEXnD/Nhydroxy, is formed by amino acids, including the β-hydroxylated Asp/Asn that bind the calcium ion.

A search for established consensus sites within the extracellular region of CD97 revealed the existence of eight potential N-glycosylation sites, Asn-Xaa-Ser/Thr. This is in agreement with the reduction of the molecular mass of natural CD97 by approximately 15 to 20 kDa by N-glycanase treatment (Fig. 3B). Markedly, a single Arg-Gly-Asp (RGD) motif was found at amino acids 274–276. RGD motifs are known as potential recognition sequences for integrins (33). The C-terminal cytoplasmic part of CD97 is rather small (46 amino acids) and contains no potential phosphorylation sites.

Transcription of CD97 is up-regulated transiently during lymphocyte activation

To investigate transcription of the CD97 gene during lymphocyte activation, RNA was analyzed from PBMC that had been stimulated with PHA for different periods of activation.

**FIGURE 4.** Amino acid sequence alignment of CD97 with the receptors for secretin (rn SecR, rat) (19), calcitonin (ss CaR, pig) (20), parathyroid hormone (dm PTHR, opossum) (21), and diuretic hormone (ad DHR, cricket). Amino acids conserved between CD97 and other receptors are shaded. The asterisks indicate positions where amino acids are identical between all sequences.
time. Hybridization with a CD97 cDNA probe revealed an mRNA of approximately 3 kb, consistent in size with the CD97 cDNA clone (Fig. 6A). The transcription of the CD97 gene at the various time points was quantified by densitometry and compared with the expression of the protein on the cell surface (Fig. 6B). In line with the weak expression on resting cells, low levels of CD97 mRNA could be detected in nonactivated cells. After stimulation, CD97 mRNA showed a transient increase between 1 and 8 h, with a maximum at 3 h. Thereafter, transcription declined to basic levels within a few hours.

**CD97 is encoded by a gene on chromosome 19**

Finally, the chromosomal localization of the CD97 gene was studied by FISH on human metaphase chromosomes by using biotinylated CD97 cDNA. The probe was hybridized to chromosomal DNA and then visualized by a recently described amplification method for in situ hybridization signals, on the basis of the deposition of biotinylated tyramine (17). As illustrated in Figure 7, clear hybridization signals were detected on the short arm of chromosome 19. In most metaphases, both chromatids of two chromosomes were labeled. The map position of CD97 was determined by length measurement of 10 chromosomes, with a 95% confidence interval. The mean FL-pter value of CD97 is 0.22 to 0.26, and corresponds to bands 19pl3.12–13.2 (Fig. 8). The gene encoding the human glucagon receptor, another member of the secretin receptor superfamily, has recently been assigned to chromosome 17 (34).

**Discussion**

In this study, we describe the molecular structure of the human CD97 Ag. By expression cloning in COS cells (11, 12), a cDNA was identified that encodes a mature protein of 722 amino acids. Near to the C-terminus, an extended hydrophobic region formed by seven distinct transmembrane segments was found that characterizes receptors coupled to G proteins (35). Structural analysis of new seven-span transmembrane molecules has recently expanded the main group of these receptors with a superfamily designated after the first described secretin receptor (19). This superfamily involves proteins that bind biologic active peptide hormones as secretin, calcitonin, parathyroid hormone, vasoactive intestinal peptide, glucagon-like peptide 1, growth hormone-releasing hormone, glucagon, pituitary adenylate cyclase-activating polypeptide, corticotropin-releasing factor, and insect diuretic hormone (19–28). Homology analysis indicates that CD97 is the first leukocyte Ag characterized that belongs to this group. However, sequence comparison with members from the secretin receptor superfamily also revealed remarkable differences. The amino acid homology is restricted largely to the seven-transmembrane region, and is smaller than between the other proteins within the superfamily. Four highly conserved cysteines preceding the first transmembrane domain, and probably important for the tertiary structure of
FIGURE 7. A, Chromosomal localization of the CD97 gene. Human metaphases were hybridized with biotinylated CD97 cDNA. Chromosomes 19 are indicated by arrowheads. B–D: Partial metaphases.

FIGURE 8. Histogram of chromosome 19 with FL-pter of CD97 after FISH.

the extracellular ligand binding region, do not exist within the CD97 polypeptide chain. In contrast to all other members of the secretin receptor superfamily, which have a rather small extracellular region of approximately 150 amino acids, that of CD97 extends 433 amino acids.

Most surprising was the finding of structural similarities between the extracellular domain of CD97 and that from extracellular matrix proteins (36). Among those, the existence of EGF-like domains, with a calcium binding site in parallel with an RGD motif, has been described for nidogen (37), fibrillin-1 and -2 (38, 39), and fibulin-2 (40). EGF domains from many proteins are known to be involved in interactions with other proteins (32). In a subclass of EGF-like domains, these interactions are stabilized by Ca\(^{2+}\) bound to a conserved sequence (30–32). RGD motifs can serve as specific recognition sites for members of the integrin family of adhesion molecules (33). The interaction of integrins with various extracellular matrix proteins, plasma proteins, and, as recently shown, also pathogenic ligands is based on RGD recognition (41). The structural findings therefore suggest an adhesive capacity for the extracellular region of CD97.

Interestingly, within the CD97 molecule, the assumed adhesive capacity is combined with a potential signal-
helpful in explaining the evolution of the CD97 Ag, and in evaluating functional similarities with other, structurally related molecules.

Acknowledgments
We thank Dr. Arthur J. Verhoeven, Heleen M. Kuiper, and Dr. Johan Stenfå for comments on the manuscript, and Tino Podstufka for experimental support.

References


