cDNA cloning and expression of a hamster α-thrombin receptor coupled to Ca²⁺ mobilization

Ulla B. Rasmussen[1], Valérie Vouret-Craviari[2], Sophie Jallat[1], Yasmin Schlesinger[1], Gilles Pagès[2], Andrea Pavirani[1], Jean-Pierre Lecocq[1], Jacques Pouysségur[2] and Ellen Van Obberghen-Schilling[2]

[1]Transgene, S.A., 11 rue de Molsheim, 67082 Strasbourg, France and [2]Centre de Biochimie, C.N.R.S., Parc Valrose, 06034 Nice, France

The serine protease α-thrombin (thrombin) potently stimulates G-protein-coupled signaling pathways and DNA synthesis in CCL39 hamster lung fibroblasts. To clone a thrombin receptor cDNA, selective amplification of mRNA sequences displaying homology to the transmembrane domains of G-protein-coupled receptor genes was performed by polymerase chain reaction. Using reverse transcribed poly(A)⁺ RNA from CCL39 cells and degenerate primers corresponding to conserved regions of several phospholipase C-coupled receptors, three novel putative receptor sequences were identified. One corresponds to an mRNA transcript of 3.4 kb in CCL39 cells and a relatively abundant cDNA. Microinjection of RNA transcribed in vitro from this cDNA in Xenopus oocytes leads to the expression of a functional thrombin receptor. The hamster thrombin receptor consists of 427 amino acid residues with 8 hydrophobic domains, including one at the extreme N-terminus that is likely to represent a signal peptide. A thrombin consensus cleavage site is present in the N-terminal extracellular region of the receptor sequence followed by a negatively charged cluster of residues present in a number of proteins that interact with the anion-binding exosite of thrombin.

α-Thrombin receptor; G-protein; Phospholipase C; Polymerase chain reaction; Oocyte expression; Hamster fibroblast

1. INTRODUCTION

Procoagulant α-thrombin (thrombin) plays a central role in hemostasis by converting fibrinogen to fibrin monomers that polymerize and form the matrix of blood clots. Thrombin is also involved in the process of wound healing and in pathological disorders involving the vascular or perivascular wall such as atherosclerosis by virtue of its effects on various cell types including platelets, endothelial cells, vascular smooth muscle, and fibroblasts (for review see [1]). Extensive studies by several laboratories have shown that thrombin elicits cellular responses via guanine nucleotide-binding proteins (G-proteins). In a line of Chinese hamster lung fibroblasts, CCL39 cells, G-proteins mediate thrombin activation of phosphoinositide-specific phospholipase C (PLC) [2,3] and inhibition of adenylate cyclase [4]. In addition to these rapid effects, thrombin potently stimulates DNA synthesis in CCL39 cells in a pertussis toxin-sensitive manner [5], indicating that one or more G-proteins participate in the stimulation of cell cycle re-entry.

Isolation and characterization of surface receptors for thrombin that interact with intracellular effectors has been a difficult task. Whereas thrombin appears to behave similarly to other agonists that activate G-protein-coupled receptors, in terms of the dose dependency and ligand-induced desensitization of its effects [6–8], there is considerable circumstantial evidence that proteolysis plays a role in the activation process. An active catalytic site on the enzyme is required for stimulation of biological responses, in CCL39 cells [9], as well as in platelets [10]. Furthermore, enzymes with similar specificity for arginyl residues such as trypsin, are able to mimic thrombin's action in platelets [6], CCL39 cells (unpublished observations) and human erythroleukemic cells (HEL) [8]. As a first step in isolating a cDNA clone encoding the receptor, we have previously demonstrated functional expression of thrombin receptors in Xenopus oocytes following microinjection of CCL39 cell mRNA [11]. In the present study, we have employed the technique of PCR to selectively amplify members of the G-protein-coupled receptor family expressed by CCL39 cells [12]. A cDNA clone has been isolated with structural features and functional properties of a thrombin receptor coupled to PLC when expressed in Xenopus oocytes.

2. MATERIALS AND METHODS

2.1. Materials
Highly purified thrombin (3209 NIH U/mg) was a generous gift of Dr J.W. Fenton II (New York State Department of Health, Albany, NY, USA). Serotonin was from Sigma (St. Louis, MO, USA).

Correspondence address: E. Van Obberghen-Schilling, Centre de Biochimie, C.N.R.S., Parc Valrose, 06043 Nice Cedex, France. Fax: (33) (93) 52 99 17.

2.2. RNA isolation

RNA was isolated by rapid homogenization of cultured cells or tissues in a solution containing 3 M LiCl/6 M urea, 10 mM sodium acetate (pH 5.2) and 0.05% SDS, followed by precipitation overnight at 4°C. Poly(A)$^+$ RNA was selected by oligo d(T)-cellulose affinity chromatography. CCL39 cell poly(A)$^+$ RNA ≥ 3 kb, enriched in thrombin receptor mRNA, was obtained by sucrose density gradient centrifugation (10–30%) for 16 h at 4°C as described [11].

2.3. Northern analysis

For Northern blot analyses, the indicated amounts of RNA were eletrophoresed on 1% agarose gels containing 0.66 M formaldehyde and transferred to Hybond-N hybridization membrane (Amersham, France). Blots were hybridized as described by Church [13] with probes labeled by random priming with ^{32}P.

2.4. PCR and subcloning

Enriched CCL39 cell poly(A)$^+$ RNA (1 μg) was reverse transcribed by random hexanucleotide priming. A 2 μl aliquot of single strand cDNA was directly amplified in 50 μl containing 10 mM Tris (pH 8.3), 10 mM KCl, 1.6 mM MgCl$_2$, 1 mM DTT, 200 mM dNTP's, 300 ng each of the 3' and 5' primers and 2.5 U of Taq polymerase (AmpliTaq DNA polymerase, Perkin-Elmer-Cetus, Norwalk, CT, USA). PCR amplifications were performed as follows: denaturation at 93°C/3 min, followed by 5 cycles of 45°C/2 min, 72°C/4 min, 93°C/1.5 min, then 25 cycles of 55°C/2 min, 72°C/4 min, 93°C/1.5 min ending with an extension step at 72°C for 10 min. One μl of this reaction product was re-amplified using the same program. Primer pairs used for PCR were designed by alignment of nucleotide sequences in transmembrane domains II, III, VI and VII of the following PLC-coupled receptors: hamster α-1 adrenergic [14], rat muscarinic m1 and m3 [15], rat 5-HT1c [16] hamster 5-HT2 [17], bovine substance K [18] receptors, and the MAS oncogene [19].

Primers that successfully amplified the thrombin receptor sequence are located in transmembrane domains II and III (see legend to Fig. 2 for exact positions); their sequences are:

(5' primer) ATGTCGACTTC(C/A)T(G/C)(A/G)T(G/C)AICCTG-GCCA(T/G)(C/T)GCTGA including a SalI restriction site at the 5' end, and

(3' primer) GAAAGCTTCGTACC(T/G)GTC(C/A)AI(A/G)(G/C)-(A/T)(A/G)ATG including a HindIII restriction site at the 5' end.

After phenol extraction and ethanol precipitation, the amplified DNA was cut with SalI and HindIII and electrophoresed on a 1% low melting point agarose gel. A broad band of about 200 bp was purified, ligated into the M13 vector pTG130 and used to transform E. coli NM522 using standard protocols. Of 7 randomly picked white plaques, 5 were sequenced and all contained the same insert sequence which was 186 bp long and showed homology to the G-protein-coupled superfamily between transmembrane domain II and III. This insert, E2-03, was labeled with ^{32}P by random priming and used for probing Northern blots and a cDNA library.

2.5. cDNA isolation and characterization

cDNA was synthesized using enriched CCL39 cell poly(A)$^+$ RNA primed with oligo(dT), and inserted with EcoRI linkers in the bacteriophage expression vector λZAP (Stratagene, La Jolla, CA, USA). The ^{32}P-labeled insert, E2-03, was used as a probe to screen 10^6 recombinants under stringent hybridization conditions. Of 96 positive clones obtained, 5 were chosen for phagemid excision, and sequenced using the Sequenase system (United States Biochemical Corporation, Cleveland, OH, USA). One clone with an insert of 2.8 kb containing a 1.28 kb open reading frame (pTG4403) was used for further analysis.

2.6. In vitro RNA synthesis and expression in Xenopus oocytes

Phagemids pTG4403 and pTG4402 corresponding to the thrombin receptor, and as control pCHSR coding for the hamster 5-HT2 receptor [17], were linearized and used for in vitro transcription with RNA polymerase (T7 or T3) and the Stratagene RNA Transcription Buffer

Kit. Transcription was carried out as described by the manufacturer for non-radioactive transcripts except that capping of the transcript was performed in the presence of 200 μM m7G(5')ppp(5')G (New England Biolabs, Beverly, MA, USA) (15 min, 37°C) prior to the elongation reaction. The transcription product was resuspended in H$_2$O at the indicated concentration for microinjection in collagenase-treated Dumont stage V and VI Xenopus oocytes. Isolation and preparation of oocytes for microinjection has previously been detailed [11]. Thrombin or 5HT-2 receptor expression in microinjected oocytes was determined by measuring agonist-stimulated ^{45}Ca$^+$ efflux in oocytes pre-loaded for 4–6 h with the isotope (Calcium-45; 50 μCi/ml, Amersham), as described [11].

3. RESULTS AND DISCUSSION

3.1. Amplification and mRNA expression of G-protein-coupled receptor sequences

Initial attempts to isolate a cDNA clone for the thrombin receptor by expression cloning in the Xenopus oocyte system were unsuccessful. Consequently, we chose to selectively amplify cDNA's encoding G-protein-coupled receptors by PCR, using fractionated CCL39 cell mRNA enriched in thrombin receptor message [11] as starting material. Since thrombin has been shown to potently activate PLC and Ca^{2+} mobilization in CCL39 fibroblasts, as well as in a number of other cell types via G-protein-mediated pathways, several degenerate nucleotide primer pairs were designed with homology to sequences within transmembrane domains II, III, VI and VII of G-protein-coupled receptors that activate PLC. Following low stringency PCR, amplified DNA fragments with the expected size were subcloned and their translated sequences were examined for the presence of G-protein-coupled receptor features (highly conserved residues and characteristic hydropathic profile). Among 90 sequences analyzed (several of which were identical), three corresponded to novel sequences for putative receptors of this family; all three were obtained with different primer pairs. In addition, sequences for the previously published hamster α-1-adrenergic [14] and 5-HT2 receptors [17], both coupled to the activation of PLC, were identified.

One of the three novel PCR-generated fragments, designated E2-03 (see section 2), hybridized to an abundant transcript of 3.4 kb on Northern blots of CCL39 cell RNA (Fig. 1a). Probes corresponding to the other two novel sequences, failed to hybridize significantly to CCL39 cell mRNA. Poly(A)$^+$ RNA isolated from human erythroleukemic (HEL) cells, a cell line that displays several markers indicative of megakaryocytic differentiation [20] and responsiveness to thrombin also hybridized to the putative thrombin receptor probe. In the human cells, however, the major transcript was slightly larger (≈ 3.5 kb) than that seen in hamster cells. In addition, less abundant smaller species were also present. A unique transcript of 3.4 kb was detected in a number of rat tissues (Fig. 1b), among which the expression level was highest in the lung.
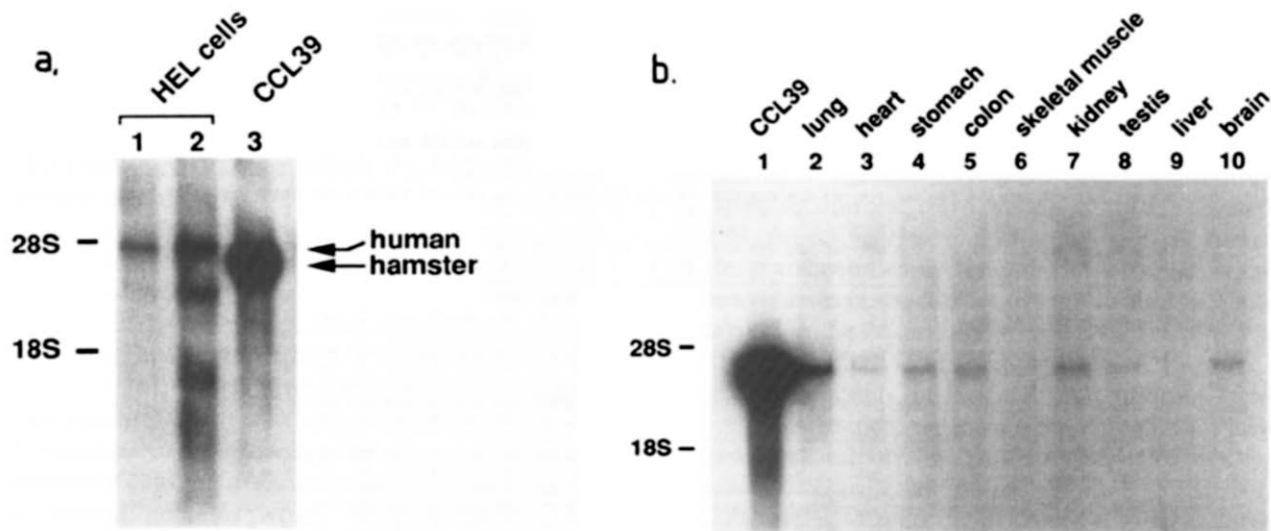
Fig. 1. (a) A Northern blot of poly(A)+ RNA from human erythroleukemia cells (HEL) (4 μg in lane 1; 10 μg in lane 2) and CCL39 cells (10 μg in lane 3) was hybridized with the ^{32}P-labeled PCR fragment, E2-03. (b) RNA blot analysis of the putative thrombin receptor expression was performed on poly(A)+ RNA from the indicated adult rat tissues, and CCL39 cells. The amount of RNA deposited was 12 μg for all rat tissues, except 5 μg for skeletal muscle, and 10 μg for CCL39 cell RNA. A ≈ 400 bp PstI fragment of pTG4403 located in the 5' end of the coding region was used as probe.

3.2. Functional expression of the thrombin receptor in Xenopus oocytes

A CCL-39 cell library constructed with the same thrombin receptor-enriched mRNA that was used for PCR, was screened with the three novel PCR-generated probes. No positive clones were detected with the two probes that failed to hybridize to CCL39 mRNA, whereas the E2-03 probe hybridized to several clones (0.01% in the 'enriched' library).

To determine whether cDNA clones isolated with E2-03 encode a thrombin receptor the corresponding RNA was synthesized in vitro and tested for functional expression in Xenopus oocytes. Expression of thrombin receptors coupled to Ca^{2+} mobilization in oocytes microinjected with CCL39 cell mRNA has previously been demonstrated using a Ca^{2+} efflux assay [11]. This assay is based on the observation that activation of PLC by G-protein-coupled receptors leads to an increase in intracellular Ca^{2+} concentrations which results in an efflux of Ca^{2+} measured by $^{45}Ca^{2+}$ release in oocytes preloaded with the isotope. As shown in Table I, oocytes microinjected with RNA transcribed from a cDNA clone carrying a 2.8 kbp insert (pTG4403) responded to thrombin. A similar increase in Ca^{2+} efflux was elicited by 5-HT in oocytes injected with RNA coding for the cloned 5-HT2 receptor. Under the same conditions, non-injected oocytes, or oocytes injected with RNA transcribed from a truncated thrombin receptor cDNA clone, pTG4402 (lacking an initiation codon) were refractory to stimulation by thrombin.

3.3. Characterization of the thrombin receptor cDNA clone

Several thrombin receptor cDNA clones were

characterized. The nucleotide and deduced amino acid sequence of one of these cDNAs that gave a positive response in the oocyte assay (pTG4403) is shown in Fig. 2a. pTG4403 contains a 5' untranslated region of 36 bp, an open reading frame of 1281 bp, and a long 3' untranslated region of 1539 bp ending in a poly(A)+ tail. The sequence AATACA is present 21 bases upstream from the polyadenylation site in pTG4403, and in at least 4 other independent cDNA's that were sequenced. This hexamer has been observed, albeit rarely (2%), in the 3' termini of vertebrate mRNA's and may serve as a functional polyadenylation signal [21]. The open reading frame initiates at a 'Kozak' consensus sequence [22] and encodes a 427 amino acid protein (calculated M_r = 47 475), in the proper reading frame with respect to the primers designed for PCR. Eight hydrophobic segments are predicted from the hydropathy plot of the deduced amino acid sequence (Fig. 2b). Based on

Table I

Expression of thrombin receptors in Xenopus oocytes

RNA	Thrombin	5-HT
pTG4403	2.12 ± 0.30 (n = 19)	
Non-injected	0.72 ± 0.08 (n = 16)	
pTG4402	0.95 ± 0.13 (n = 7)	
pCHSR		3.34 ± 0.59 (n = 8)

$^{45}Ca^{2+}$ efflux was determined on single collagenase-treated oocytes 48 h after microinjection with 50 nl of RNA (0.5 mg/ml) corresponding to: the putative thrombin receptor, clones pTG4403 (see Fig. 2) and pTG4402 (lacking 94 bp of 5' sequence as compared to pTG4403), or to the cloned 5-HT2 receptor. Oocytes were stimulated for 2 min with thrombin (10^{-7} M) or 5-HT (10^{-6} M), as indicated. Results are presented as the fold-stimulation of $^{45}Ca^{2+}$ released during a 2 min incubation in the presence of agonist. The data presented were obtained in 3 independent experiments using different animal donors.

a.

```
CCGCGGGCGCCCGGCCGCGCGGGCAGCCTCGGGACA ATG GGG CCG CAG CGC TTG CTG CTC GTC GCC GCT GGC CTC AGC'CTG TGC GGT  87
                                      M   G   P   Q   R   L   L   L   V   A   A   G   L   S   L   C   G   17

CCC TTG CTG TCT TCC CGT GTC CCC GTG CGC CAG CCA GAA TCT GAG ATG ACA GAT GCC ACT GTG AAC CCC CGC TCA TTT TTT CTT  171
 P   L   L   S   S   R   V   P   V   R   Q   P   E   S   E   H   T   D   A   T   V   N   P   R   S   F   F   L   45

AGG AAT CCC GGT GAA AAT ACA TTT GAA CTG ATT CCC CTG GGG GAT GAG GAG GAG AAA AAT GAA AGC ACC TTG CCA GAG GGC AGG  255
 R   N   P   G   E   N   T   F   E   L   I   P   L   G   D   E   E   E   K   N   E   S   T   L   P   E   G   R   73

GCA ATC TAC TTA AAT AAA AGC CAC TCT CCT GCA CCA CTT GCT CCC TTC ATC TCT GAA GAT GCC TCA GGG TAT CTG ACC AGC CCC  339
 A   I   Y   L   N   K   S   H   S   P   A   P   L   A   P   F   I   S   E   D   A   S   G   Y   L   T   S   P   101

TGG CTG AGG CTC TTC ATC CCC TCG GTT TAC ACC TTT GTG TTT GTC GTC AGC CTT CCT CTG AAC ATC CTG GCC ATC GCA GTG TTT  423
 W   L   R   L   F   I   P   S   V   Y   T   F   V   F   V   V   S   L   P   L   N   I   L   A   I   A   V   F   129
                                                                   I

GTC TTG AAG ATG AAG GTC AAA AAG CCG GCT GTG GTA TAC ATG CTG CAC CTG GCC ATG GCT GAC GTG CTG TTC GTG TCC GTG CTT  507
 V   L   K   H   K   V   K   K   P   A   V   V   Y   H   L   H   L   A   H   A   D   V   L   F   V   S   V   L   157
                                                                              II

CCC CTG AAA ATC AGC TAC TAC TTT TCC GGC AGT GAT TGG CAG TTT GGG TCT GGG ATG TGT CGC TTC GCC ACT GCA GCG TTT TAT  591
 P   L   K   I   S   Y   Y   F   S   G   S   D   W   Q   F   G   S   G   H   C   R   F   A   T   A   A   F   Y   185

TGT AAC ATG TAC GCC TCC ATC ATG CTC ATG ACT GTC ATA AGC ATT GAC CGG TTC CTG GCA GTC GTG TAC CCC ATC CAG TCC CTG  675
 C   N   H   Y   A   S   I   M   L   M   T   V   I   S   I   D   R   F   L   A   V   V   Y   P   I   Q   S   L   213
         III

TCC TGG CGC ACT CTG GGC AGA GCT AAC TTC ACT TGC CTG GTC ATT TGG GTG ATG GCC ATC ATG GGG GTG GTG CCT CTT CTC CTC  759
 S   W   R   T   L   G   R   A   N   F   T   C   L   V   I   W   V   M   A   I   M   G   V   V   P   L   L   L   241
                                                  IV

AAG GAG CAG ACC ACC CGG GTT CCA GGG CTC AAC ATC ACC ACC TGC CAT GAT GTC CTC AAT GAG ACC CTG CTG CAG GGC TTT TAC  843
 K   E   Q   T   T   R   V   P   G   L   N   I   T   T   C   H   D   V   L   N   E   T   L   L   Q   G   F   Y   269

TCC TAC TAC TTC TCG GCC TTC TCC GCT GTC TTC TTT CTT GTG CCA TTG ATC ATT TCC ACG ATA TGC TAC ATG TCC ATC ATT CGG  927
 S   Y   Y   F   S   A   F   S   A   V   F   F   L   V   P   L   I   I   S   T   I   C   Y   H   S   I   I   R   297
                                                  V

TGT CTC AGC TCC TCC TCA GTT GCC AAC CGC AGC AAG AAG TCC CGG GCT TTG TTC CTG TCC GCT GCT GTG TTC TGC GTC TTC ATC  1011
 C . L   S   S   S   S   V   A   N   R   S   K   K   S   R   A   L   F   L   S   A   A   V   F   C   V   F   I   325
                                                      VI

GTC TGC TTT GGA CCC ACC AAC GTC CTC CTG ATC ATG CAC TAC CTG CTC CTG TCT GAC AGC CCT GCC ACA GAG AAG GCC TAC TTC  1095
 V   C   F   G   P   T   N   V   L   L   I   H   H   Y   L   L   L   S   D   S   P   A   T   E   K   A   Y   F   353

GCT TAC CTC CTC TGC GTC TGT GTG AGC AGC GTG AGC TGC TGC ATC GAT CCC TTG ATT TAC TAC TAC GCC TCC TCC GAG TGC CAG  1179
 A   Y   L   L   C   V   C   V   S   S   V   S   C   C   I   D   P   L   I   Y   Y   Y   A   S   S   E   C   Q   381
                             VII

AGG ACC CTC TAC GGC ATC TTG TGC TGC AAA GAA AGA AGC TCC GAC CCG AAC AGT TAC AAC AGC ACC GGC CAG CTG ATG CCG AGT AAA  1263
 R   T   L   Y   G   I   L   C   C   K   E   S   S   D   P   N   S   Y   N   S   T   G   Q   L   M   P   S   K   409

ATG GAT ACC TGC TCT AGT CAC CTG AAT AAC AGC ATA TAC AAA AAG CTA TTA GCT TAG GGGAGGATTGCTAGAAGGTCACATAGAAAAGGTTG  1355
 M   D   T   C   S   S   H   L   N   N   S   I   Y   K   K   L   L   A   *                                        427
GAAAAGTGGACAGCCCGGGATCCTGGCAAGAACTATATTTACTTCACAAGCCATGACTGACTTGCATGTCTACTTCTTATTAGTGCTATCAAGCATAATTATCAGAAAGGT  1466
AATGGGAACCAGAGAGAGATTTCCAGTGCTGCGGTAGCAACCGTCACTTCTTGTATATCTAGGTGACTTAATTATATAGGTGGTGTGCGCACACATATTTGCAATGCAGTA  1577
TGAAATCAATGCTTTGACACTTTCTTGTTTATTCCCTAGCAATTACTATGGAAATAATTGGATTCTCTGAGTTATAATAAACAAAGTCTGTTAACACTGAACAGCTGAAGG  1688
TATCTAATGATAGGGAAGGAGTCCATAGTTTAGACTTCACACAGCTTTTGCCCATATGTATTTTTTTCAGTTATTTGATGGTAATGTTTAAAGTATAGAAGGGTGAAACAG  1799
TATTATCTGTATAGGGGAAGATCTAGTGCTTTTCATCTTGAATTACCAAAGCCGTTGGAAAGCCTATTTTGATATGGGGTACCATTTTTGCAATTTACATACTGAATACAT  1910
CGACCAAGACTGAGCATAAAACTCACCAGGATCATAAGAAACCTTACAAAGCAGCCAAGGGGGCCTCTGACAGCAGATCCATACTTAACCAGCTGTGCCTCCCAAGAGAGT  2021
GGTGATGGGGACCACCAGGCCCACTCCTTCCTCCAAGGCATCCCATTAGCTATGAGTTGACTGTATCTGGATTGAAGCTGGCAAGATAGGATATGACATCCAGGGAGGGAA  2132
TAGCCACAAAAGTCGTCTGCACATCTGAGGAGCAGAGAAGGTGTGTTTATATCCAGTAGCTGTCCTGCAAGGCTGGCTCTTGCACAGACACACCCATGTGCCCTGGGTCAC  2243
TCTGGTGGATACTGGGCCTGCAGACTGACCACTGCAGGAGATTAACTGAGTCTTGCCTTTACTCTAGCAAAGCAGATAGCTGACATGGCCAATGTGATGTATACACTGCAA  2354
ATACACCATTGTACATAGAAGTGTGTCAGCCACACATCGCCAAGCATATTTCACATAAGCAAGGCCTAACAGCTAAACAGCTTTGGAGATCTGAGTTTCTGCATCAGTAGC  2465
TGTAGGTTAGGATAAAAACACAGTTTAAGATGTATATTTTTAAGAAACAAACCCTTAAGTCTACAATAATTAGACACTATTTATTTACAAATGTTTTATTAAAAATTGCCC  2576
AAATCAGACAGGCATTGAGGCACATGCCTTTAAGCCCAGCACTTGGGAGGCAGAGGTAGGCAGATCTCTGTGCATTCAGGTCAGCCTGGTCTACATGATGACTTCTAGGTC  2687
ATCTAGGGCTATATAGTGAGACCTTGGCTTAAAACAATAATGACATCAAAACAAAACAACAACAACAACAACAAAAAAAAAACCCACATTTTTAGAAAGCAAGAAAACACA  2798
AAAAGTTACCTGAATCCAGTTTTAAGAAATACATAGGTCAGTCTGGTTAAAAAAAAA  2856
```
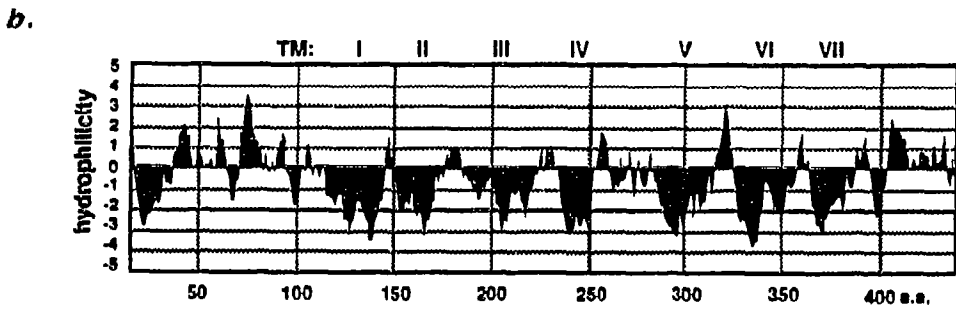
b.

Fig. 2. (a) Nucleotide and derived amino acid sequence of the hamster thrombin receptor cDNA pTG4403. A putative signal peptide (italics), 7 putative transmembrane domains (TM I-VII) and four potential N-glycosylation sites (*) are indicated under the amino acid sequence. The potential thrombin cleavage site (R41) is denoted by an arrow; details of the sequence surrounding the putative thrombin cleavage site are shown in Fig. 3. The primer pair used for PCR (see Section 2) amplifies the region between nucleotide position 460 and 646 (including the primer sequences).
(b) Hydophilicity of the translated 427 amino acid sequence from pTG4403 was determined using the Kyte–Doolittle scale [35].

analogy with the topographic model for other members of this receptor family, the first hydrophobic domain is likely to represent a signal peptide, with a probable signal peptidase cleavage site after Serine-22 or Serine-21 [23]. Four potential N-glycosylation sites are present, including two in the N-terminal extracellular domain, and two in the second extracellular loop which may be involved in receptor localization [24].

Since the proteolytic activity of thrombin is required for its biological effect on CCL39 cells, the extracellular

domains were searched for consensus thrombin cleavage sites and sequence similarly with other proteins that interact specifically with the enzyme. Results of these analysis are presented in Fig. 3. A thrombin cleavage site typically contains the sequence P4-P3-P2-(Arg/Lys)-P1'-P2' where P2, P3 and P4 are hydrophobic residues and P1' and P2' are non-acidic, particularly when P2 is a Proline [25]. At Arginine-41 in the translated sequence of pTG4403, a characteristic thrombin cleavage site is present. A putative thrombin cleavage site at this position was also found in the human thrombin receptor which was reported during the course of this work [26]. A negatively charged cluster of residues is located approximately 20 amino acids C-terminal to the cleavage site.

A negatively charged domain or acidic stretch is present in several proteins which bind thrombin, and found near the cleavage site in proteins cleaved by thrombin. This domain presumably interacts with the anion-binding exosite of thrombin and contributes to the enzyme's specificity. In fibrinogen Aα a negatively charged domain is located about 20 amino acids C-terminal to the thrombin cleavage site, similar to the thrombin receptor (Fig. 3). For Factor VIII, an acidic domain about 20 residues C-terminal to the thrombin cleavage site is aligned with the thrombin receptor in Fig. 3, but other acidic domains are present in the molecule approximately 20 residues N-terminal to thrombin cleavage sites (positions 372/373 and 1689/1690) [27]. Heparin cofactor II and hirudin, which both complex thrombin, also possess a negatively charged cluster of amino acids. In fact it has been shown that the acidic C-terminus of hirudin interacts with the anion-binding exosite of thrombin [28], and that synthetic peptides derived from the C-terminal 20 residues of hirudin possess anticoagulant activity [29]. Interesting sequence similarity exists between residues 56-60 in hirudin (FEEIP) and residues 53-57 (FELIP) in the hamster thrombin receptor, residues that are absent in the human thrombin receptor sequence (Fig. 3).

Overall amino acid sequence identity between the hamster and the recently reported human thrombin receptor is 79%. This is low compared to 87% overall sequence identity between the hamster [30] and human [31]) β2 adrenergic receptor, and >90% between the

rat and human forms of the muscarinic m1, dopamine D2 receptors and the *mas* oncogene. The cytoplasmic loops, which presumably play a role in G-protein coupling, are more highly conserved indicating that the hamster and human thrombin receptors utilize the same intracellular signalling systems. Sequence identities in the first, second and third intracellular loops are 100, 95 and 93 respectively, whereas they range from 68 to 91% in the transmembrane domains. The lowest homology between human and hamster thrombin receptor is observed in the N-terminal extracellular domain which is 55% (excluding the signal peptide). Altogether these results suggest that the hamster sequence may represent a different receptor isoform coupled to PLC, however further studies including genomic Southern analyses are required to clarify this point.

Concerning amino acid identity of the thrombin receptor with the transmembrane domains of the 7 receptors aligned for primer design (see Section 2), values ranged from: 17-30% in TMI, 23-41% in TMII, 23-36% in TMIII, 11-32% in TMIV, 23-37% in TMV, 22-30% in TMVI and 14-32% in TMVII. Apart from its human homolog, the hamster thrombin receptor displayed highest overall identity (27%) to the human cytomegalovirus sequence HCMVU28 [32], with 55, 41 and 32% in TMII, III, and VII, three domains that have been implicated in ligand binding in the case of adrenergic receptors [24]. It is interesting to note that the HCMVU28 sequence has a negatively charged cluster of residues in the N-terminal domain (DYDED), but a putative thrombin cleavage site in the short N-terminal domain is absent.

At positions 177 and 256 in the first and second extracellular domains, two Cysteins are present that are highly conserved in G-protein-coupled receptors and may be involved in an essential disulfide bridge important for maintaining receptor structure [33]. Two Lysines in a basic surrounding are present in the third intracellular loop (position 309 and 310); a similar motif is present at this position in many receptors of this family [32] and may be involved in G-protein coupling. It is noteworthy that this third intracellular loop, which varies in length among the different receptors of this gene family, is only 14 residues long in the thrombin receptor. Potential sites for modification of the
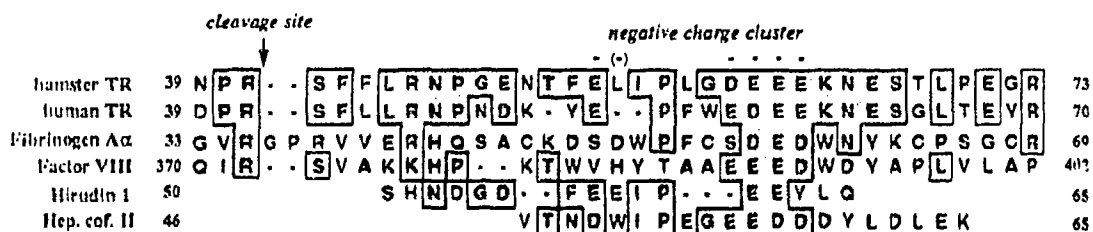

Fig. 3. Alignment of the extracellular N-terminal domain of the hamster thrombin receptor with other thrombin proteins that interact with. Proteins aligned include: the hamster thrombin receptor, the human thrombin receptor [26], human Aα fibrinogen [36], human factor VIII [27], hirudin variant I [37], and heparin cofactor II [38]. A thrombin cleavage site and a common negative charge cluster are shown above the sequences. Homologies to the hamster thrombin receptor are boxed; Asp (D) and Glu (E) are boxed together since they are both negatively charged, as well as Lys (K) and Arg (R) which are both positively charged.

thrombin receptor activity and/or desensitization by phosphorylation are present including 3 Threonine and 10 Serine residues in the cytoplasmic C-terminal domain, six of which are in pairs next to a charged amino acid, and 6 Serines in the third intracellular loop. A potential tyrosine kinase phosphorylation site [34] is located at Tyrosine 399.

In conclusion, the cDNA we have isolated encodes a functional thrombin receptor with structural features of a G-protein-coupled receptor. A thrombin consensus cleavage site and negative charge cluster are present in the N-terminal extracellular domain of the receptor, consistent with recent evidence that thrombin receptor activation involves a novel proteolytic mechanism by which thrombin cleaves its receptor to create a new amino terminal that functions as a ligand to activate the receptor [26]. In this study by Vu et al. a synthetic peptide of 14 residues corresponding to the new receptor N-terminal was shown to activate the cloned thrombin receptor in oocytes an induce platelet activation. Subsequently, we have found that a peptide corresponding to only 7 residues proximal to the thrombin cleavage site in the hamster sequence (Serine-42–Proline-48) both activates PLC and inhibits adenylate cyclase in CCL39 cells (manuscript in preparation). These results suggest that either the cloned receptor is coupled to at least two distinct G proteins, or that more than one receptor subtype is present in these cells and can be activated by the peptide. We are currently investigating these possibilities. Interestingly, the receptor peptide does not stimulate DNA synthesis in CCL39 cells but rather it potentiates the action of the tyrosine kinase-activating fibroblast growth factor (in preparation). Therefore we believe that mitogenic stimulation of cells by thrombin requires yet an additional pathway involving a tyrosine kinase.

REFERENCES

[1] Fenton II, J.W. (1988) Seminars in thrombosis and hemostasis 14, 234–240.

[2] Paris, S. and Pouysségur, J. (1986) EMBO J. 5, 55–60.

[3] Paris, S. and Pouysségur, J. (1987) J. Biol. Chem. 262, 1970–1976.

[4] Magnaldo, I., Pouysségur, J. and Paris, S. (1988) Biochem. J. 253, 711–719.

[5] Chambard, J.C., Paris, S., L'Allemain, G. and Pouysségur, J. (1987) Nature 326, 800–803.

[6] Martin, B.M., Feinman, R.D. and Detwiler, T.C. (1975) Biochemistry 14, 1308–1314.

[7] Paris, S., Magnaldo, I. and Pouysségur, J. (1988) J. Biol. Chem. 263, 11250–11256.

[8] Brass, L.F., Manning, D.R., Williams, A.G., Woolkalis, M.J. and Poncz, M. (1991) J. Biol. Chem. 266, 958–965.

[9] Van Obberghen-Schilling, E., Chambard, J.C., Paris, S., L'Allemain, G. and Pouysségur, J. (1985) EMBO J. 4, 2927–2932.

[10] Davey, M.G. and Luscher, E.F. (1967) Nature 216, 857–858.

[11] Van Obberghen-Schilling, E., Chambard, J.-C., Lory, P., Nargeot, J. and Pouysségur, J. (1990) FEBS Lett. 262, 330–334.

[12] Parmentier, M., Libert, F., Maenhaut, C., Lefort, A., Gérard, C., Perret, J., Van Sande, J., Dumont, J.E. and Vassart, G. (1989) Science 246, 1620–1622.

[13] Church, G.M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991–1995.

[14] Cotecchia, S., Schwinn, D.A., Randall, R.R., Lefkowitz, R.J., Caron, M.G. and Kobilka, B.K. (1988) Proc. Natl. Acad. Sci. USA 85, 7159–7163.

[15] Bonner, T.I., Buckley, N.J., Young, A.C. and Brann, M.R. (1987) Science 237, 527–532.

[16] Julius, D., MacDermott, A.B., Axel, R. and Jessel, T.M. (1988) Science 241, 558–564.

[17] Chambard, J.C., Van Obberghen-Schilling, E., Haslam, R.J., Vouret, V. and Pouysségur, J. (1990) Nucleic Acids Res. 18, 5282.

[18] Masu, Y., Nakayama, K., Tamaki, H., Harada, Y., Kuno, M. and Nakanishi, S. (1987) Nature 329, 836–838.

[19] Young, D., Waitches, G., Birchmeier, C., Fassano, O. and Wigler, M. (1986) Cell 45, 711–719.

[20] Tabilio, A., Rosa, J.-P., Testa, U., Kieffer, N., Nurden, A.T., Del Canizo, M.C., Breton-Gorius, J. and Vainchenker, W. (1984) EMBO J. 3, 453–459.

[21] Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Cell 41, 349–359.

[22] Kozak, M. (1987) Nucleic Acids Res. 15, 8125–8148.

[23] Von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.

[24] Lefkowitz, R.J. and Caron, M.G. (1988) J. Biol. Chem. 263, 4993–4996.

[25] Chang, J.Y. (1985) Eur. J. Biochem. 151, 217–224.

[26] Vu, T.K.H., Hung, D., Wheaton, V.I. and Coughlin, S.R. (1991) Cell 64, 1057–1068.

[27] Vehar, G.A., Keyt, B., Eaton, D., Rodriguez, H., O'Brien, D.P., Rotblat, F., Oppermann, H., Keck, R., Wood, W.I., Harkins, R.N., Tuddenham, E.G.D., Lawn, R.M. and Capon, D.J. (1984) Nature 312, 337–342.

[28] Chang, J.-Y., Ngai, P.K., Rink, H., Dennis, S. and Schlaeppi, J.-M. (1990) FEBS Lett. 261, 287–290.

[29] Maraganore, J.M., Chao, B., Joseph, M.L., Jablonski, J. and Ramachandran, K.L. (1989) J. Biol. Chem. 264, 8692–8698.

[30] Dixon, R.F., Koblika, B.K., Strader, D.J., Benovic, J.L., Dohlman, H.G., Frielle, T., Bolanowski, M.A., Bennett, C.D., Rands, E., Diehl, R.E., Mumford, R.A., Slater, E.E., Sigal, I.S., Caron, M.G., Lefkowitz, R.J. and Strader, C.D. (1986) Nature 321, 75–79.

[31] Kobilka, B.K., Dixon, R.A.F., Frielle, T., Dohlman, H.G., Bolanowski, M.A., Sigal, I.S., Yang-Feng, T.L., Francke, U., Caron, M.G. and Lefkowitz, R.J. (1987) Proc. Natl. Acad. Sci. USA 84, 46–50.

[32] Chee, M.S., Satchwell, S.C., Preddie, E., Weston, K.M. and Barrell, B.G. (1990) Nature 344, 774–777.

[33] Dixon, R.A.F., Sigal, I.S. and Strader, C.D. (1988) Cold Spring Harbor Symposia on Quantative Biology 53, 487–497.

[34] Cooper, J.A., Esch, F.S., Taylor, S.S. and Hunter, T. (1984) J. Biol. Chem. 259, 7835–7841.

[35] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.

[36] Kant, J.A., Lord, S.T. and Crabtree, G.R. (1983) Proc. Natl. Acad. Sci. USA 80, 3953–3957.

[37] Dodt, J., Müller, H.-P., Seemüller, U. and Chang, J.-Y. (1984) FEBS Lett. 165, 180–184.

[38] Blinder, M.A., Marasa, J.C., Reynolds, C.H., Deaven, L.L. and Tollefsen, D.M. (1988) Biochemistry 27, 752–754.