Biochemical and Biophysical Research Communications **284,** 118–125 (2001)
doi:10.1006/bbrc.2001.4908, available online at http://www.idealibrary.com on IDEAL®

A Novel Early Estrogen-Regulated Gene *gec1* Encodes a Protein Related to GABARAP

Sandrine Vernier-Magnin, Stéphanie Muller, Myriam Sallot, Jean Radom, Jean-François Musard, Pascale Adami, Philippe Dulieu, Jean-Paul Rémy-Martin, Michèle Jouvenot,[1] and Annick Fraichard
Institut d'Etude et de Transfert de Gènes, Bâtiment INSERM, Besançon, France; and
Laboratoire de Biochimie Biologie Moléculaire, Faculté des Sciences, Besançon, France

Received March 26, 2001

We have isolated, in guinea-pig endometrial cells, an estrogen-induced 1.8 kb RNA called *gec1*. Screening of a guinea-pig genomic library led to identification of *gec1* gene consisting of 4 exons and 3 introns. Exon 1 contains the 5′UTR and the ATG initiation codon. A guinea-pig *gec1* cDNA was obtained by 5′-RACE. The 351 bp coding sequence shares 76.8% identity with that of the human *GABARAP* 924 bp cDNA while UTRs of the two cDNAs differ. A *gec1* probe from the 3′UTR revealed a 1.9 kb mRNA in human tissues and a human *GEC1* cDNA was isolated from placenta. Its coding sequence shares 93 and 79% identity with that of guinea-pig *gec1* and human *GABARAP*, respectively. The human and guinea-pig GEC1 proteins have 100% identity. GEC1 and GABARAP proteins have 87% identity and N terminus featuring a tubulin binding motif. Thus, estrogen-regulated *gec1* is a new gene which could encode a microtubule-associated protein. © 2001 Academic Press

Key Words: **estrogen; primary response gene; *gec1*; GABARAP (GABA$_A$-receptor-associated protein), microtubule-associated protein; molecular cloning; genomic organization; LD-PCR; RACE.**

Estrogens, potent mitogens for breast and uterine epithelia, are responsible for tumor-promoting action on breast and endometrial cancers (1, 2). Estrogen action on cell proliferation is mediated by receptors ERα and ERβ (3, 4) which act as transcription factors and regulate the expression of target genes (5). Therefore, identification of estrogen-regulated genes in target tissues is an essential step in understanding the molecular mechanisms of estrogen action on normal cell growth and consequently on tumor cell growth.

Based upon the kinetics of the response and the binding of receptors to the target genes, a classification of steroid-regulated genes into three groups has been proposed: primary response genes, secondary response genes and delayed primary response genes (6, 7). The primary response genes are directly activated by steroid–receptor complexes, without the requirement for *de novo* protein synthesis and their expression arises at the beginning of the cascade of transcriptional events. Among the genes directly activated by estrogen in the uterus, there are proto-oncogenes encoding transcription factors, such as c-*myc*, c-*fos* and c-*jun* (8–10) and some genes with an incompletely defined function (11, 12).

Our laboratory has developed an *in vitro* model of guinea-pig endometrial glandular epithelial cells (GEC). It has been demonstrated that the cultured cells are estrogen responsive (13) and that 17β-estradiol (E_2) induces *c-fos* gene expression within 2 h when it acts in association with either epidermal growth factor plus insulin or a protein synthesis inhibitor such as cycloheximide (14, 15). In an effort to identify genes that may be regulated by E_2 in the same conditions as *c-fos*, a cDNA library has been constructed from poly(A)$^+$ RNAs extracted from GEC stimulated with E_2 (10^{-8} M) plus cycloheximide (Chx, 10 μg/ml) for 2 h using the λgt10 cloning system. By differential screening, one estrogen-regulated sequence, called *gec1*, has been identified. The initial *gec1* cDNA (833 bp) has been partially sequenced and corresponded to the 3′ end of the 1.8 kb mRNA. Furthermore, as previously reported (16) and shown in Fig. 1, the level of *gec1* mRNA in GEC treated with E_2 plus Chx for 2 h was significantly higher than that observed in control Chx treated cells. The fold induction by E_2 was 2.3. The early E_2 action on *gec1* gene in GEC is similar to the E_2 action on *c-fos* gene (15) and

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AF 012920 (guinea-pig *gec1* cDNA), AF 312680 (guinea-pig *gec1* genomic DNA), and AF287012 (human *GEC1* cDNA).

[1] To whom correspondence should be addressed at Laboratoire de Biochimie Biologie Moléculaire, Faculté des Sciences, 16, route de Gray, 25030 Besançon cedex, France. Fax: 33(0)3.81.66.62.67. E-mail: michele.jouvenot@univ-fcomte.fr.

0006-291X/01 $35.00

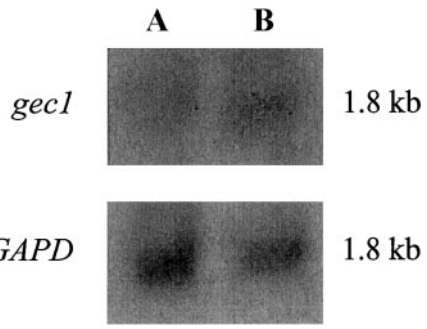

FIG. 1. Induction of *gec1* mRNA by E_2. 15 μg of total RNAs from Chx-treated cells (lane A) or E_2 plus Chx-treated cells (lane B) for 2 h were electrophoresed on agarose gel and transferred onto nylon filters. The filters were then hybridized with the *gec1* DNA probe. They were rehybridized with the *GAPD* probe. The *gec1* mRNA levels were normalized to the level of *GAPD* mRNAs.

the *gec1* gene appears to be an early estrogen-regulated gene.

The aim of this study was to identify the guinea-pig *gec1* genomic organization, its full length cDNA and its human counterpart.

MATERIALS AND METHODS

Isolation and characterization of genomic clones. The 833 bp *gec1* cDNA was totally sequenced. Then, a 755 bp fragment was amplified by PCR using P1 and P2 primers, subcloned into pGEM-T vector (Promega) and used as *gec1* probe 1. The guinea-pig genomic library in the λFIX bacteriophage vector (Stratagene) was screened (1.5×10^6 phage plaques) with this probe 1. Hybridization was performed overnight at 42°C in DIG Easy Hybridization Buffer (Roche Diagnostics). Filters were washed twice in 2× SSC, 0.1% SDS at room temperature for 5 min and twice in 0.1× SSC, 0.1% SDS at 68°C for 15 min. Positive clones were plaque purified, and phage DNA was isolated (Reddy *et al.*, 1988). Two clones were isolated and one clone called *gec1a*, containing a 19 kb insert, was chosen for further analysis. The genomic clone was digested or partially digested with selected restriction enzymes (*Sac*I and *Bgl*II). Then, the restriction fragments were subcloned into pBluescript II KS+ (pBS, Stratagene) or PQE60 (Qiagen).

Southern blot analysis. Fifty nanograms of purified phage DNA digested or not with *Sac*I and *Bgl*II were electrophoresed on 0.8% agarose gel and blotted onto nylon filter (Dig User Manual, Roche Diagnostics). The filter was hybridized and washed as described above.

Long distance-PCR. Long range PCRs using 300 ng of purified phage DNA as templates were performed according to the manufacturer's specifications (Expand 20 kb PCR system, Roche Diagnostics). Primers were chosen on restriction fragments generated by *Sac*I digestion of the 19 kb insert, subcloned into pBS and partially sequenced.

Screening of human cDNA library. A human placenta cDNA library (Clontech) was screened (400,000 phage plaques) using probe 1 as described in isolation and characterization of genomic clones. Thirty-two clones were isolated and excised as a p-triplex vector and one was chosen for further analysis (1.2 kb in size) and sequenced.

Probes and primers. The glyceraldehyde-3-phosphate dehydrogenase (*GAPD*) cDNA probe and the *β actin* cDNA probe were from Clontech. For hybridization of Northern blot, these probes as well as the *gec1* probe 1 were labeled with [^{32}P]-dCTP by nick translation. For all other hybridization experiments, the probes were labeled by incorporation of digoxigenin 11-dUTP during PCR (PCR DIG probe Synthesis kit, Roche diagnostics) with purified DNA fragments as template.

Probe 1: a 755 bp *gec1* specific probe was amplified using primers P1 (5′ AACTGTAGTAGCGGACCATACCTGG 3′) and P2 (5′ GAACCTGAAGCTGACAGGACACAC 3′) chosen on the 833 bp fragment from the λgt10 *gec1* clone. Probe 2: a 1042 bp *gec1* specific probe was amplified using primers P3 (5′ GAGGTATTAAACAAGTGTCTATTGTGCAC 3′) and P4 (5′ CAGTGAGGTCAGAGGGCA 3′) chosen on the *Sac*I 2314 bp fragment from the *gec1* genomic clone. Probe 3: a 892 bp *gec1* specific probe was amplified using primers P5 (5′ TGTGCTGTAGCAGAAATTGTAGG 3′) and P6 (5′ GGTGTGGACAGAGGTGAGTCTAA 3′) chosen on the *Bgl*II 4129 bp fragment from the genomic clone *gec1*.

RT-PCR and 5′-RACE. Five micrograms of total RNA extracted from guinea-pig's liver or from human placenta were treated with 5 units of DNase I (Promega) for 30 min at 37°C and then, reverse transcribed in a 20 μl reaction containing: 500 ng of oligo(dT) primer, 0.8 mM dNTPs, 10 mM DTT, 2 μl of 5× first strand buffer and 10 units of SuperScript II reverse transcriptase (Life Technologies). The sample was incubated for 1 h at 42°C. The PCR (50 μl) contained: 2 μl of RT-PCR products, 200 μM of each dNTP, 5 μl of 10× Mg-free *Taq* polymerase buffer, 2.5 mM $MgCl_2$, 0.1 μM of sense and reverse primers and 2 units of *Taq* polymerase (Promega). The PCR protocol was: 2 min at 94°C, then 35 cycles (1 min at 94°C, 1 min at 55°C, 1 min at 72°C) and finally 7 min at 72°C.

A new 5′-RACE technique described by Ranasinghe and Hobbs (18) was used with some modifications. The first-strand cDNA was synthesized in a classic reverse transcription reaction as described above. Then, the reaction volume was made up to 100 μl with TE (10 mM Tris–HCl, 1 mM EDTA pH 8) and 1 unit of RNase I (Epicentre) was added. The sample was incubated for 30 min at 37°C to make the cDNA–mRNA hybrid blunt-ended. A phenol extraction was performed to inactivate the enzyme and the cDNA–mRNA hybrid was ethanol precipitated after adding 0.1 vol of 3 M sodium acetate pH 5.2 and 2 μl of pellet-paint coprecipitant (Novagen). After centrifugation, the pellet was washed, dried and dissolved in 6.5 μl of TE. The cDNA–mRNA hybrid was ligated to the pBS digested by *Eco*RV. After ligation, the PCR amplification was carried out between a sequence specific *gec1* primer (5′ CAGGTGGATCCTCTTCCGGAT 3′) and T3 or T7 primers of the plasmid. The PCR conditions with 3 μl of ligated cDNA–mRNA hybrids were identical to those described in the RT-PCR protocol. The PCR products were analyzed on 1% agarose gel. Fragments of interest were subcloned into the pGemT-Easy vector (Promega).

Guinea-pig tissue collection. Ovaries, uterus, kidney, lung, liver and brain were rapidly removed from a mature female guinea pig (of the Hartley albinos variety) killed by decapitation on the day of the vaginal opening. The different organs were dissected free of adipose tissue (if necessary) and washed in phosphate-buffered saline (at 4°C). These organs, except the uterus, were immediately frozen in liquid nitrogen and stored at −70°C until use. Before freezing, the uterus was dissociated and the endometrium was separated from the myometrium and stored at −70°C.

RNA extraction and Northern blotting. To study the expression of mRNAs in guinea-pig tissues, a pool of each tissue was prepared from four guinea pigs and RNAs were extracted using the cesium chloride gradient method (19). Poly(A)$^+$ mRNAs were selected by two passages over oligo(dT) cellulose columns and quantified at 260 nm. Samples of poly(A)$^+$ mRNAs (5 μg) were denatured (20), electrophoresed in 1.1% agarose gels, and blotted onto nylon filters (Zeta-probe^{GT}, Bio-Rad) according to the vacuGene method (Pharmacia). The filters were baked (80°C, 1 h) and kept desiccated until use. The filters were prehybridized, hybridized with ^{32}P-labeled *gec1* cDNA probe 1 and washed as previously described (14, 21). They were then

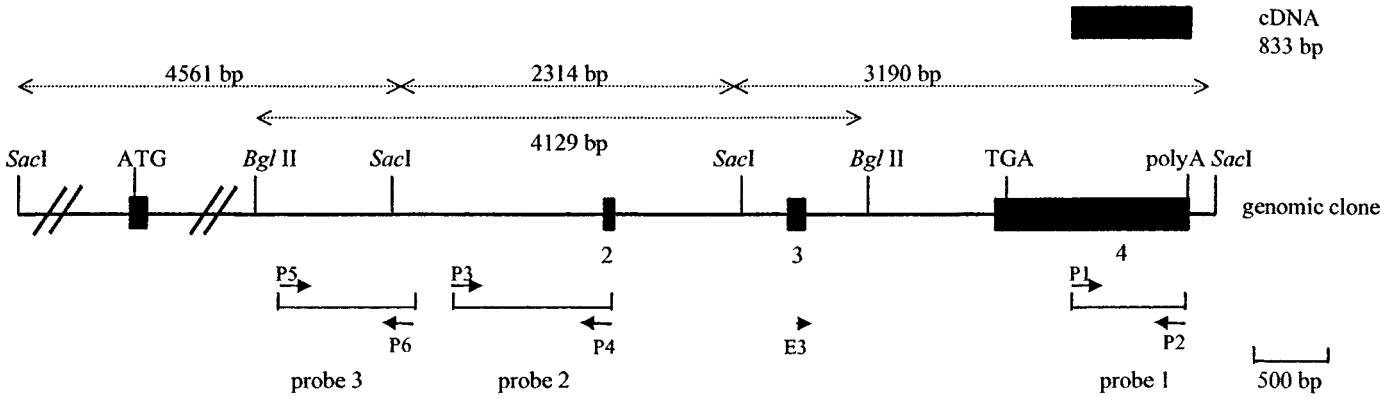

FIG. 2. Organization of the guinea-pig *gec1* gene. A partial restriction endonuclease map of the genomic *gec1* clone is shown. The lines represent introns, and the boxes represent exons. Above, the initial λgt10 *gec1* clone is represented. The position of the translation initiation codon (ATG), the translation termination codon (TGA), polyadenylation site, restriction sites, probes, and primers used are indicated.

exposed to X-ray films with an intensifying screen at −80°C. The filters were dehybridized and rehybridized with the GAPD probe used as control. Two human multiple tissue Northern blots were purchased from Clontech. The blots were prehybridized and hybridized in "Expresshybrid solution" (Clontech) according to the manufacturer's recommendations. In a first hybridization, the ^{32}P-labeled *gec1* cDNA probe 1 was used and the washing conditions were the following: twice in 2× SSC, 0.1% SDS at 50°C for 15 min, twice in 2× SSC, 0.1% SDS at room temperature for 15 min, four times in 0.5× SSC, 0.1% SDS at 50°C and four times in 0.1× SSC, 0.1% SDS at 50°C for 15 min. After autoradiography, the filters were rehybridized with a *β actin* cDNA probe used as control.

DNA sequence determination and analysis. Genomics and cDNA clones were sequenced using the dye terminator cycle sequencing kit (Amersham) on a ABI 373 automated sequencer (Perkin–Elmer) by walking along the sequence using synthetic oligonucleotides as primers. The nucleotide sequences homologies search through the EMBL/GenBank were performed by using the Advanced Blast software. Deduced amino acid sequences were analyzed by using the BCM Search Launcher software.

RESULTS

Characterization of guinea-pig gec1 genomic structure. The λgt10 cDNA clone was incomplete in relation to the estimated transcript size (1.8 kb). Moreover, no open reading frame was found in the available sequence. Initial 5′-RACE experiments could not identify further the *gec1* sequence. Thus, a genomic guinea-pig library was screened to progress into the sequence and identify putative introns. Using the *gec1* probe 1, labeled with 11-dUTP digoxigenin, one positive clone was obtained first. The 19 kb insert from this clone was cut into 6 fragments (6600, 4561, 3190, 2314, 1100, and 810 bp) by *Sac*I digestion. One of them (3190 bp) hybridized with the probe 1 and was totally sequenced. The 833 bp *gec1* cDNA determined from the λgt10 clone was located between the nucleotide 1969 and 2801 of the 3190 bp fragment. Additional nucleotides were obtained and two predicted exons, called exon 3 (119 bp) and exon 4 (1322 bp), were positioned by software BCM Gene Finder on the 3190 bp sequence (Fig. 2). The putative position of these two exons was confirmed by RT-PCR analysis with primers E3 and P2 (Fig. 2) chosen on the guinea-pig 3190 bp sequence and with RNA extracted from guinea-pig liver as template. Indeed, the amplification produced a fragment which size confirmed the junctions exon–intron (data not shown).

At this stage, the known cDNA sequence was 1441 bp in size and incomplete in 5′.

To localize the *Sac*I fragment upstream the 3190 bp, partial digestions were performed. One 2314 bp fragment was estimated near the 3190 bp (data not shown). It was subcloned into the pBS vector and partially sequenced. Long Distance-PCRs were carried out with primers chosen on the 2314 bp and 3190 bp fragments: P7, a sense primer of the 2314 bp fragment (5′-GAG GTA TTA AAC AAG TGT CTA TTG TGC AC-3′); P8, a sense primer of the 3190 bp fragment (5′-AGT AAG AAT GAT ACC AAT GTG TAT TGT GTT-3′); P9, a reverse primer of the 3190 bp fragment (5′-ATC AGC TGA TGC AGG AAG G-3′). Amplification was obtained only with the sense primer P7 on the 2314 bp fragment and the reverse primer P9 on the 3190 bp fragment. According to this result, the 2314 bp fragment appeared to be localized upstream of the 3190 bp fragment. Total sequencing and software analysis of the 2314 bp fragment gave a determination of an additional exon (79 nucleotides in size), called exon 2. RT-PCR, using a forward primer localized on the 5′ end of exon 2 and the reverse primer P2 localized on the end of exon 4 confirmed the position of the 3 exons (data not shown).

The 19 kb genomic insert was digested by *Bgl*II and the restriction fragments were then hybridized with probe 2 specific for the 2314 bp fragment. Only one 4129 bp fragment hybridized with this probe (data not shown). This result as well as the presence of a *Bgl*II site in the 3190 bp fragment suggested that the 4129

TABLE 1

DNA Sequences of Exon–Intron Junctions in the *gec1* Gene

Exon	Size (bp)	5′ splice donor	3′ splice acceptor	Intron size (bp)
1	>90	TCCCG**gt**gagcctccctggg	ctcctgtcttccc**ag**GTCAT	4177
2	79	CACTG**gt**aacgttgttttcc	ctttcctacattc**ag**TTGGC	1047
3	119	ATGAG**gt**aatggtcctgtca	tatctatttttct**ag**GACAA	1304
4	1322	ACTCTaactgagttcagtact		

Note. Exon sequences are shown in uppercase letters and intron sequences in lower case letters.

bp fragment overlapped both a part of the 3190 bp fragment as well as and the 2314 bp fragment (Fig. 2). Sequencing of this *Bgl*II fragment gave a determination of further 919 bp upstream of the 2314 bp *Sac*I fragment. In the same manner, using probe 3 specific for the *Bgl*II fragment, a 4561 bp *Sac*I fragment was located upstream of the 2314 bp fragment. This 4561 bp fragment allowed determination of 3642 bp upstream the *Bgl*II 4129 bp fragment. It contained exon 1 (>90 bp) with an ATG initiation codon (Fig. 2).

These overall results led us to propose an organization of the *gec1* gene (Fig. 2) with 4 exons. Intron/exon junctions are presented in Table 1 and have the conserved GT and AG dinucleotides present at their donor and acceptor sites.

Identification of the 5′ end of guinea-pig gec1 cDNA. To identify the 5′ *gec1* cDNA end, total RNAs were extracted from guinea-pig liver and used as template to perform another 5′-RACE experiment, as outlined under Materials and Methods. A specific *gec1* primer derived from exon 3 and T3 or T7 primer from pBS vector were used for PCR amplification of the ligated cDNA-mRNA hybrids. One RACE-PCR product of 521 bp, obtained with T3 and *gec1* specific primer, hybridized with probe 2. Sequencing of this 521 bp fragment gave the 5′ part of exon 3, the exon 2 and an upstream 401 bp sequence. This sequence confirmed the 3′ end of exon 1 and contained the ATG initiation codon at position +312. As shown in Fig. 3, the 1921 bp *gec1* cDNA contains a 5′ untranslated region of 311 bp, an open reading frame (ORF) of 351bp, and a 3′ untranslated region of 1259 bp. The open reading frame encodes a protein of 117 amino acid residues with a calculated molecular weight of 14162 Da and an isoelectric point (p*I*) of 8.67. The results of a search in GenBank databases are presented in Fig. 4. The *gec1* cDNA (Accession No. AF012920) has high identity with the whole sequence (UTRs and ORF) of a recently reported 1835 bp cDNA (Hashimoto *et al.*, 2000, unpublished sequence in GenBank, Accession No. AB041648) isolated from mouse brain cDNA library using the oligo-capping method. The *gec1* ORF (351 bp) shows 76.8% identity with the coding sequence of the 924bp human *GABARAP* cDNA (Accession No. NM_007278) but the UTRs of both cDNAs (guinea-pig *gec1* cDNA and human *GABARAP* cDNA) have no identity. As shown in Fig. 5, the predicted guinea-pig GEC1 protein shares 100% identity with the unnamed mouse protein and 87% identity with human GABARAP, an ubiquitous protein interacting with the γ2 subunit of $GABA_A$-receptor and with tubulin and microtubules (22, 23). GEC1 and GABARAP proteins have respectively 60 and 64% similarity with light chain-3 (LC3) of microtubule-associated protein (MAPs) 1A and 1B and 75% similarity with GATE-16, a membrane transport modulator (24). The GEC1 N terminus, like that of GABARAP, is highly positively charged and features a putative tubulin binding motif (25).

gec1 mRNA expression in guinea pig and human tissues. A ^{32}P-labeled guinea-pig *gec1* probe 1 was used to determine by Northern blot analysis the *gec1* expression in guinea-pig tissues. The results of a representative experiment with guinea-pig poly(A)$^+$ mRNAs are reported in Fig. 6A. A single 1.8 kb transcript was detected in all tissues examined. However, endometrium, liver and lung expressed the highest levels of *gec1* mRNA. Probe 1 was also used to investigate the expression of *gec1* counterpart in human tissues (Fig. 6B). Only one transcript of 1.9 kb was detected in the different tissues examined except in the lung where the control *β actin* mRNA level was very low.

Isolation of a partial human GEC1 cDNA. As a 1.9 kb transcript was detected in all human tissues examined (Fig. 6B), a human placenta cDNA library was subsequently screened with the *gec1* probe 1 to obtain the *gec1* human counterpart. Thirty two clones were isolated and the longest one, containing 1081 nucleotides, was sequenced. It contained a polyadenylation signal but no ORF. It shared 71.3% identity with the 3′ untranslated region of guinea-pig *gec1* and 82% identity with that of the unnamed mouse cDNA (Accession No. AB041648), but no homology with the human *GABARAP* cDNA. At this stage, the new human cDNA appeared to be likely the 3′ UTR of human *GEC1* mRNA. Moreover, this 1081 bp cDNA fitted with a sequence localized to human chromosome 12 (Accession No. AC006514) while *GABARAP* gene has been localized to chromosome 17 (23). By analysis of the unordered sequences located to chromosome 12, a pu-

```
1                                                                   CA GTG CAA TCT    11
12   AGC GCG CTG GGG GCG TGG CGT GGG CGG GTC AAC GTG ACG TCA CAA CAC CGC CGC ACC GCC   71
72   CCC GTG TTT TTG TGC TCC TGC TCG CGG GGA AGC CTC TGG GAA TTT TTT CAC CTG GCT GTT  131
132  TCC CCG CTC TGG TCC CGA GGC TTT CCG AAC CCT TGG TCA AGG CTG TGG CAT CAG GAG CTG  191
192  CCG TTT CTA GAA CGA TCC TGG GAG CCC TGA AAT CAG CCT GAG GCA GCG AGA CGC GCG AGT  251
252  CTT CAT CAA GGG AGG CAG GCC TCG CGC GGG GAT CTC TGG AAA GCA AGC CTG CGC CCC ATC  311
312  ATG AAG TTC CAG TAC AAG GAG GAC CAT CCC TTC GAG TAT CGG AAA AAA GAA GGA GAG AAG  371
     M   K   F   Q   Y   K   E   D   H   P   F   E   Y   R   K   K   E   G   E   K
372  ATC CGA AAG AAG TAC CCA GAC AGG GTC CCG GTC ATT GTG GAG AAG GCT CCT AAG GCC AGG  431
     I   R   K   K   Y   P   D   R   V   P   V   I   V   E   K   A   P   K   A   R
432  GTG CCT GAT CTG GAC AAG AGG AAG TAC CTT GTG CCC TCT GAC CTC ACT GTT GGC CAG TTC  491
     V   P   D   L   D   K   R   K   Y   L   V   P   S   D   L   T   V   G   Q   F
492  TAT TTC TTA ATC CGG AAG AGG ATC CAC CTG AGA CCT GAG GAC GCC TTA TTC TTC TTT GTC  551
     Y   F   L   I   R   K   R   I   H   L   R   P   E   D   A   L   F   F   F   V
552  AAC AAC ACT ATC CCT CCC ACT AGC GCA ACC ATG GGC CAA CTG TAT GAG GAC AAC CAT GAA  611
     N   N   T   I   P   P   T   S   A   T   M   G   Q   L   Y   E   D   N   H   E
612  GAA GAC TAT TTT CTG TAC GTG GCC TAC AGT GAT GAG AGT GTC TAT GGG AAG TGA GTG GTG  671
     E   D   Y   F   L   Y   V   A   Y   S   D   E   S   V   Y   G   K   *
672  GAA ACC CAG CAG ATG GGA GCA CCT GGA CTC GGG GGT GGG GCG GGC TGT GAG TGG GAC TTG  731
732  AGG GAA GAG AAG GAG GGC TCC CAC CAT GGA GAA GAC AGA AGG TGA AGA CAT TGG AAA CAT  791
792  TAC ACT GCA CAC ATT GGC GTC ATA TTT TCA CAT GCT CAG TTG ATG TTT TTT TGC TGC TTC  851
852  TTT GGC CCA GGG AGA AAG CAT GTC AGG ACA GAG CTG CTG GAC TGG CTT TGA TAA GAG AGN  911
912  ATA GAG ATA GCC TAA GTT CAT GGC ATT CCC AAG AAN TAA GTT TTT TTG TGA TTT NAT AGT  971
972  AGG TGA GAA GGT GGG CAG GTT AAA GCC AGA GGT GAA TGT AAT CCA GTA GCA ATT CCC ATT 1031
1032 TTT TTG GGC AGA GTT TCA ATT TTT GAC ATT TGC ACA AAA CGA AAT AGG AAA GGG GAC AAG 1091
1092 AAT TCC TTA GGA ACT GTA GTA GCG GAC CAT ACC TGG GGA CCA AAA CAT ACC TAC AGT AAT 1151
1152 TAA AGC ACC ATA GCC TTT GAT TGC TTC TGT TTC TCC CCT TCC TGC ATC AGC TGA TAG TCT 1211
1212 CTT CTC GTA TAT GAA TGT CAC ACC CCA GAG GTG GCA GTA GAC ACC CAC ACA TTC AGA AGG 1271
1272 AAC TGT CTT AAC TTT GGA AAA TGG CAG TCC TAC AGA AGG TGA TGC CTG TTA GTT TTG ATA 1331
1332 GCA TTA AAC TGG AAA TGA CTT GAG TGC TGA GCT TTT GTC TTA AAC TGC TCT CTC TCT CGT 1391
1392 GCC CCT TAT CCC ACA CCT ATT TTG GAA TTT CCT GAA CCC TGG GCA CTT TCA TTT GCA GGC 1451
1452 TAA AGC CAG TTG TTT CTT GTT ACC TTT AGC AGT TTC CAT GTA CAG AGA CAG AGG TGT CTT 1511
1512 TCA TGT GGA AGG GGC AGG TAT GAG GTA GGT AGA TTA TCT ACA CTT TGC ACA CGT CCA GAA 1571
1572 TTT TTG CTC CCT CGC TGT TGT CCC TCA GTC TCA CGC ACA GAC AGG AAA AAG CTA AGC AAT 1631
1632 GGC CAG CTG CTT CCC TCC TTT GGT TTT CAT CCA CTG CAG CTG CTA CTT AGA AAT GTT TGG 1691
1692 AGG GAT GAC GTT AGT AAT TCA TGG GAG TTT AAT TCG TTT ATT TTC ATT TAA GGG TTG AGG 1751
1752 GTG GGG GTG GGG AGC AGG ACT TGC GCT CAG ATG TGA CAT TTC ATC TCT GCT AAT GAA AAG 1811
1812 GAT CCT TCC ATT GGG GGA AAT CCG TGT GTC CTG TCA GCT TCA GGT TCT TGT ATA ATG AGG 1871
1872 TCA ATC TTG TCA GGC CAA GGA AAT AAA ATA ATT GCT TAT TGT AAA ACT CT              1921
```

FIG. 3. cDNA sequence and protein translation of guinea-pig *gec1*. The primer used for 5′-RACE is double underlined. The polyadenylation signal is underlined and the translation stop codon is marked with an asterisk. Nucleotides 1 to 401, 402 to 480, 481 to 599, and 600 to 1921 correspond to exons 1, 2, 3, and 4, respectively. The sequence is available under GenBank Accession No. AF012920.

tative ORF for the human *GEC1* gene was identified. Then, using two human primers (a sense primer upstream this putative ORF, a reverse primer at the 3′ end of the 1081 bp sequence), and human placenta mRNAs as template, a cDNA was amplified (Accession No. AF287012). As shown in Fig. 4, the human cloned 1638 bp cDNA has a 3′ UTR of 1277 bp, a coding sequence of 351 and 10 bp upstream the ATG initiation codon. Like the initial 1081 bp cDNA, the complete 1277 bp 3′ UTR has no identity with the 3′ UTR *GABARAP* cDNA. The 351 bp coding sequence has more identity with the guinea-pig *gec1* cDNA (93%) than with the human *GABARAP* cDNA (79%). As shown in Fig. 5, the protein encoded by the human 1638 bp cDNA shares 100% identity with the guinea-pig GEC1 protein and 87% identity with the human GABARAP protein. In conclusion, it can be said that the partial human cloned cDNA corresponds to the human *GEC1* gene encoding a protein related to GABARAP.

DISCUSSION

A previous screening of a guinea-pig cDNA library constructed from mRNAs of cultured GEC led to the isolation of the *gec1* cDNA corresponding to an estrogen-regulated 1.8 kb mRNA (16). As this cDNA lacked the 5′ end, several 5′-RACE reactions were per-

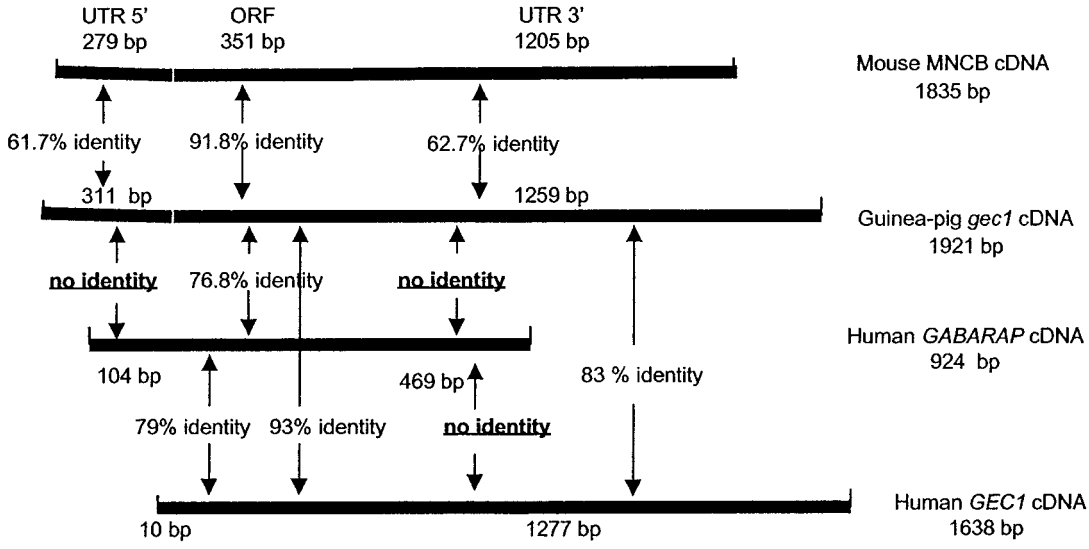

FIG. 4. Comparison of cDNA sequences of guinea-pig *gec1*, mouse MNCB, human *GABARAP*, and human *GEC1*. Two compared sequences are related by an arrow. ORF, open reading frame; UTR, untranslated region.

formed but did not succeed, suggesting that secondary structures could hinder the reverse transcription step. Thus, a guinea-pig genomic DNA library was screened with probe 1 corresponding to 755 bp of the initial cDNA. This screening allowed isolation of one clone with a 19 kb insert. By sequencing of three *Sac*I fragments (10065 bp), we identified: the *gec1* 351 bp coding region included in 4 exons, 3 introns, 1538 bp upstream the ATG initiation codon and 389 bp downstream the *gec1* gene. The initial cDNA was 833 bp in size and a 5′-RACE procedure led to identification of 1921 bp cDNA with 311 bp upstream the ATG initiation codon, suggesting a 401 bp size for exon 1. Comparison of this cDNA with the genomic sequence confirmed all the exon-intron junctions. Furthermore, the guinea-pig *gec1* cDNA appears to be full length since its size matches the mRNA size (1.8 kb) estimated from Northern blot analysis.

The computer analysis of the genomic 1538 bp sequence upstream the ATG revealed a typical TATA box and a transcriptional initiation site included in an initiator element (Inr), $PyPyA_{+1}N(TA)PyPy$. The TATA box and the transcriptional initiation site are located respectively at 363 and 310 bp upstream the ATG. As we have previously demonstrated that the *gec1* mRNA level increases after stimulation of GEC by E_2 plus Chx, *gec1* gene induction may represent a primary response to estrogen. According to a classical model, the estrogen effects on primary response genes depend on the association of the estrogen–receptor complex with specific DNA sequences called ERE (estrogen response element) (26). ERE elements are frequently localized in position 5′ of the promoter, but can also exist in the 3′ position of the coding sequence (27) or in the coding sequence (28). The computer analysis of the *gec1* gene and of the flanking regions revealed a puta-

FIG. 5. Alignment of guinea-pig or human GEC1, mouse MNCB (GenBank Accession No. AB041648), human GABARAP (GenBank Accession No. NM007278), bovine or human GATE-16 (GenBank Accession Nos. AF020262 and AJ010569, respectively), and human LC3 (GenBank Accession No. AF303888) proteins. Identities are indicated by black background and similar residues are shadowed.

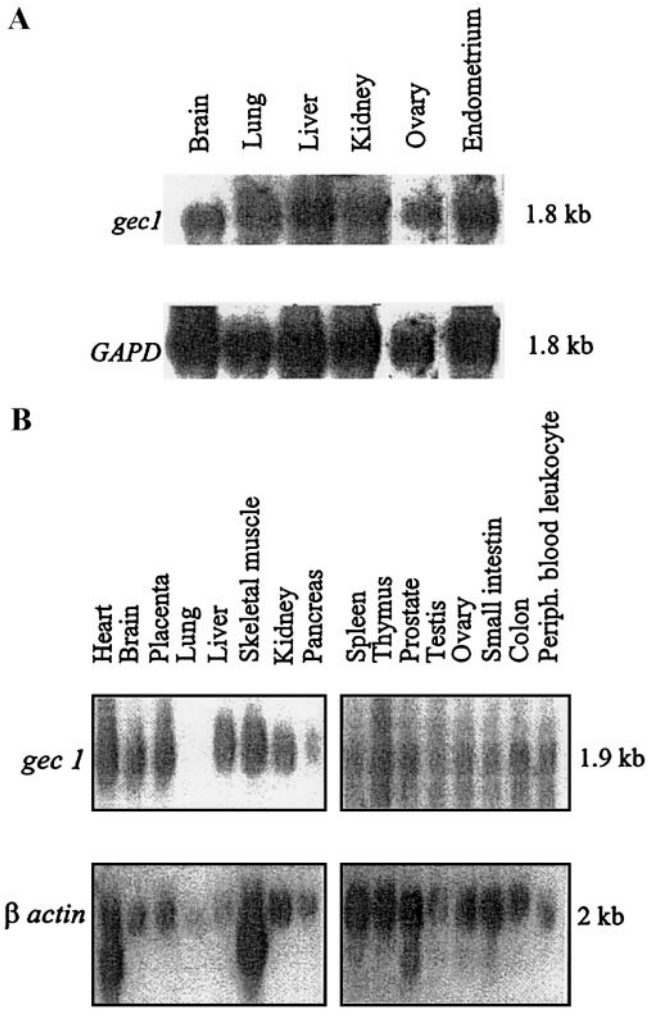

FIG. 6. (A) Northern blot of 5 μg poly(A)$^+$ mRNAs isolated from guinea-pig tissues was hybridized sequentially with the ^{32}P-labeled *gec1* probe 1 and human *GAPD* probe. (B) Expression of *gec1* counterpart in human tissues. Two human multiple tissue Northern blots (Clontech) with 2 μg of poly(A)$^+$ mRNAs were hybridized sequentially with the ^{32}P-labeled *gec1* probe 1 and human *β actin* probe.

tive ERE (GGTCAACGTGACG) which differs only by one nucleotide from the consensus (GGTCANNNTGACC). This putative ERE is located at 275 bp upstream the ATG, i.e., in the 5′ UTR. Experiments are in progress to identify the promoter region and *cis*-regulatory elements involved in *gec1* transcriptional control.

The coding sequences of guinea-pig *gec1* cDNA and human *GABARAP* cDNA have 76.8% identity while 5′ UTRs (311 bp for *gec1* and 104 bp for *GABARAP*) and 3′ UTRs (1259 bp for *gec1* and 469 bp for *GABARAP*) differ strongly. The results of human Northern blot hybridization are in agreement with this sequence analysis. Indeed, probe 1 corresponding to the 3′ UTR of guinea-pig *gec1* detected one 1.9 kb transcript while a *GABARAP* cDNA probe has revealed one 0.9 kb transcript (22). Furthermore, probe 1 allowed isolation of a 1081 nt cDNA from human placenta which had identity with the 3′ UTR of guinea-pig *gec1*, but no identity with the 3′ UTR of human *GABARAP* cDNA and which appeared to be the 3′ UTR of the human *GEC1* gene. Using the genomic sequences in the databases and RT-PCR experiments, we further identified the human *GEC1* cDNA and its ORF. The human and guinea-pig GEC1 proteins have 100% identity and share 87% identity with the human GABARAP. According to this sequence analysis and the distinct chromosomal localizations, *GEC1* and *GABARAP* genes appear to be different. In view of these overall results, it can be concluded that *GEC1* is a new gene coding a protein related to GABARAP.

Both proteins GEC1 and GABARAP have a putative tubulin-binding N-terminal motif and similarity with MAP light chain-3. In response to different extracellular signals, MAPs can be phosphorylated by various protein kinases, including cAMP-dependent kinase, CaM kinase, PKC, tyrosine kinase (29) and casein kinase II (30). GEC1 and GABARAP proteins could also be substrates of protein kinases. Indeed, the two proteins contain a potential tyrosine kinase phosphorylation site at residue 106 which is included in a 100–106 motif (HEEDYFLY for GEC1 and HEEDFFLY for GABARAP). These motifs differ only by one residue from the consensus pattern [R or K –X(2)-D or E-X(3)-Y]. GABARAP, differently from GEC1, has a serine residue which could be phosphorylated by casein kinase II (16–19 SEGE motif) or by cAMP-dependent protein kinase (13–16 KRRS motif).

The sequence between amino acids 36 and 68 of GABARAP is important for its $GABA_A$-receptor-binding activity. As GEC1 contains an identical sequence, its interaction with the $GABA_A$ receptor needs to be investigated. However, both proteins are widely expressed, suggesting their involvement in biological events other than interaction with $GABA_A$ receptors.

The *gec1* gene, expressed in all tissues investigated and induced by E_2 in GEC, could encode a new MAP or a new component of MAP complexes, acting as linker protein between membrane receptors and microtubules. There is little information concerning the estrogen action on MAP expression. However, it has been demonstrated that estrogen significantly increases MAP-2 and Tau protein levels in female rat pituitary cells and suggested that these quantitative changes could be related to cell proliferation (31). Further studies on the regulation and function of *gec1* gene will help to clarify the mechanism of estrogen action in normal and cancer cell growth.

ACKNOWLEDGMENTS

This research work was supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche (MesR) and by grants from MesR (EA 2272), the Ligue Nationale Contre le Cancer

(Comité du Doubs and Comité du Jura) and INSERM (CRI N° 4U001B). We thank F. Poncet and A. Laroche for technical assistance and Dr Lundström-Baudais for reviewing the English.

REFERENCES

1. Hong, W. K., and Sporn, M. B. (1997) Recent advances in chemoprevention of cancer. *Science* **278,** 1073–1077.
2. Hulka, B. S. (1997) Epidemiologic analysis of breast and gynecologic cancers. *Prog. Clin. Biol. Res.* **396,** 17–29.
3. Evans, R. M. (1988) The steroid and thyroid hormone receptor superfamily. *Science* **240,** 889–895.
4. Kuiper, G. G., Enmark, E., Pelto-Huikkon, M., Nilsson, S., and Gustafsson, J. A. (1996) Cloning of a novel estrogen receptor expressed in rat prostate and ovary. *Proc. Natl. Acad. Sci. USA* **93,** 5925–5930.
5. Beato, M. (1988) Gene regulation by steroid hormone. *Cell* **56,** 335–344.
6. Rories, C., and Spelsberg, T. C. (1989) Ovarian steroid action on gene expression: Mechanisms and models. *Annu. Rev. Physiol.* **51,** 653–681.
7. Dean, D. M., and Sanders, M. M. (1996) Ten years after: Reclassification of steroid responsive genes. *Mol. Endocrinol.* **10,** 1489–1495.
8. Dubik, D., and Shiu, R. P. (1992) Mechanism of estrogen activation of c-*myc* oncogene expression. *Oncogene* **7,** 1587–1594.
9. Weisz, A., and Bresciani, F. (1993) Estrogen regulation of protooncogenes coding for nuclear proteins. *Crit. Rev. Oncogenesis* **4,** 361–388.
10. Bigsby, R. M., and Li, A. (1994) Differentially regulated immediate early genes in the rat uterus. *Endocrinology* **134,** 1820–1826.
11. Orimo, A., Inoue, S., Ikeda, K., Noji, N., and Muramatsu, M. (1995) Molecular cloning, structure and expression of mouse estrogen-responsive finger protein Efp. Co-localization with estrogen receptor mRNA in target organs. *J. Biol. Chem.* **270,** 24406–24413.
12. Everett, L. M., Li, A., Deveraju, G., Caperell-Grant, A., and Bigsby, R. M. (1997) A novel estrogen-enhanced transcript identified in rat uterus by differential display. *Endocrinology* **138,** 2836–2841.
13. Alkhalaf, M., Chaminadas, G., Propper, A., and Adessi, G. L. (1987) Compared effects of estrogens and oestrone sulphate on the epithelial cell surface of guinea-pig endometrium (*in vivo* and *in vitro* studies). *Exp. Clin. Endocrinol.* **89,** 201–210.
14. Jouvenot, M., Pellerin, I., Alkhalaf, M., Marechal, G., Royez, M., and Adessi, G. L. (1990) Effects of 17β-estradiol and growth factors on c-*fos* gene expression in endometrial epithelial cells in primary culture. *Mol. Cell. Endocrinol.* **72,** 149–157.
15. Pellerin, I., Vuillermoz, C., Jouvenot, M., Royez, M., Ordener, C., Marechal, G., and Adessi, G. L. (1992) Superinduction of c-*fos* gene expression in cultured guinea-pig endometrial cells requires priming by a cycloheximide-dependent mechanism. *Endocrinology* **131,** 1094–1100.
16. Pellerin, I., Vuillermoz, C., Jouvenot, Ordener, C., Royez, M., and Adessi, G. L. (1993) Identification and characterization of an early estrogen-induced RNA in cultured guinea-pig endometrial cells. *Mol. Cell. Endocrinol.* **90,** 17–21.
17. Reddy, K. J., Kuwabara, T., and Sherman, L. A. (1988) A simple and efficient procedure for the isolation of high-quality phage λDNA using a DEAE-cellulose column. *Anal. Biochem.* **168,** 324–331.
18. Ranasinghe, C., and Hobbs, A. A. (1998) A simple method to obtain the 5′ ends of mRNA sequences by direct ligation of cDNA-RNA hybrids to a plasmid vector. Elsevier Trends Journals Technical Tips Online: T01519.
19. Chirgwin, J. M., Przbyla, A. E., McDonald, R. J., and Rutter, W. J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18,** 5294.
20. Thomas, P. S. (1980) Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
21. Sallot, M., Ordener, C., Lascombe, I., Propper, A., Adessi, G. L., and Jouvenot, M. (1996) Differential EGF action on nuclear protooncogenes in human endometrial carcinoma RL95–2 cells. *Anticancer Res.* **16,** 401–406.
22. Wang, H., Bedford, F. K., Brandon, N. J., Moss, S. J., and Olsen, R. W. (1999) $GABA_A$-receptor-associated protein links $GABA_A$ receptors and the cytoskeleton. *Nature* **397,** 69–72.
23. Wang, H., and Olsen, R. W. (2000) Binding of the GABA(A) Receptor-associated protein (GABARAP) to microtubules and microfilaments suggests involvement of the cytoskeleton in GABARAPGABA(A) receptor interaction. *J. Neurochem.* **75,** 644–655.
24. Sagiv, Y., Legesse-Miller, A., Porat, A., and Elazar, Z. (2000) GATE-16, a membrane transport modulator, interacts with NSF and the Golgi v-SNARE GOS-28. *EMBO J.* **19,** 1494–1504.
25. Mann, S. S., and Hammarback, J. A. (1994) Molecular characterization of light chain 3. *J. Biol. Chem.* **269,** 11492–11497.
26. Parker, M. G. (1995) Structure and function of estrogen receptors. *Vit. Horm.* **51,** 267–287.
27. Hyder, S. M., Stancel, G. M., Nawaz, Z., McDonnels, D. P., and Loose-Mitchell, D. S. (1992) Identification of an estrogen response element in the 3′ flanking region of the murine c-*fos* protooncogene. *J. Biol. Chem.* **270,** 18047–18054.
28. Hyder, S. M., Nawaz, Z., Chiapetta, C., Yokoyama, K., and Stancel, G. M. (1995) The protooncogene c-*jun* contains an unusual estrogen-inducible enhancer within the coding sequence. *J. Biol. Chem.* **270,** 8506–8513.
29. Jefferson, A. B, and Schulman, H. (1991) Phosphorylation of microtubule-associated protein-2 in GH3 cells. *J. Biol. Chem.* **266,** 346–354.
30. Sapir, T., Cahana, A., Seger, R., Nekhai, S., and Reiner, O. (1999) LIS1 is a microtubule-associated phosphoprotein. *Eur. J. Biochem.* **265,** 181–188.
31. Matsuno, A., Takekoshi, S., Sanno, N., Utsunomiya, H., Ohsugi, Y., Saito, N., Kanemitsu, H., Tamura, A., Nagashima, T., Osamura, R. Y., and Watanabe, K. (1997) Modulation of protein kinases and microtubule-associated proteins and changes in ultrastructure in female rat pituitary cells: Effects of estrogen and bromocriptine. *J. Histochem. Cytochem.* **45,** 805–813.